

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE**  
**SCIENTIFIQUE**

**UNIVERSITE M'HAMED BOUGARA-BOUMERDES**



**Faculté de Technologie**

**Thèse de Doctorat**

Présentée par :

**BOUKHOBZA Fatima Zohra**

En vue de l'obtention du diplôme de **DOCTORAT - LMD** en :

**Filière : Télécommunication**

**Spécialité : Réseaux et Télécommunications**

**Sélection des Paramètres Pertinents pour la  
Reconnaissance des Expressions Faciales en Mode  
Indépendant des Personnes**

**Devant le jury composé de :**

|                           |            |                         |               |
|---------------------------|------------|-------------------------|---------------|
| Mme. ACHELI Dalila        | Professeur | Université de Boumerdes | Président     |
| M. ROUABAH Khaled         | Professeur | Université de M'sila    | Directeur     |
| M. HACINE GHARBI Abdenour | MCA        | Université de BBA       | Co- Directeur |
| M. MERAIHI Yassine        | Professeur | Université de Boumerdes | Examineur     |
| M. MEBARKIA Kamel         | MCA        | Université de Sétif 1   | Examineur     |
| Mme. BELKACEM Samia       | MCA        | Université de Boumerdes | Examinatrice  |
| M. Kadri Abid             | Radiologue | ABID Unique Center      | Invité        |

Année Universitaire 2024/2025

*À mon frère Rahim,*

*Que Dieu t'accorde Sa miséricorde et t'accueille  
en Son Paradis.*

اللهم اغفر له وارحمه وأسكنه فسيح جناتك

## **REMERCIEMENTS**

Je rends tout d'abord grâce à Allah, le Tout-Puissant, pour m'avoir accordé la force, la patience et la persévérance nécessaires à l'accomplissement de ce travail.

J'exprime toute ma gratitude à Monsieur ROUABAH Khaled, mon Directeur de thèse, ainsi qu'à Monsieur HACINE GHARBI Abdenour, mon Co-Directeur, pour leur encadrement, leurs précieux conseils et leur disponibilité tout au long de ce parcours.

J'adresse également mes remerciements à Monsieur Philippe Ravier, enseignant à l'Université d'Orléans, France, pour son expertise et ses précieux échanges qui ont contribué à enrichir cette recherche.

Je remercie sincèrement les membres du jury d'avoir accepté d'évaluer ce travail et pour l'intérêt qu'ils y portent.

Mes remerciements vont aussi à mes enseignants, qui ont su me guider et m'apporter les connaissances nécessaires tout au long de mon parcours académique.

Ma profonde reconnaissance va à mes parents, pour leur soutien sans faille et leurs innombrables sacrifices. À ma sœur Selma et à toute ma famille, merci pour votre présence et vos encouragements constants.

Enfin, je remercie chaleureusement toutes les personnes qui, de près ou de loin, m'ont accompagnée et soutenue dans cette aventure, par leurs conseils, leur aide ou leurs encouragements.

## ملخص

تُعَدُّ عملية التعرف التلقائي على تعابير الوجه مهمةً أساسيةً في مجالات عدة مثل علم النفس، والتفاعل بين الإنسان والحاسوب، والمراقبة. يتكون نظام التعرف على تعابير الوجه بشكل رئيسي من مرحلتين: مرحلة التعلم ومرحلة التعرف. في المرحلة الأولى، يهدف التعلم إلى نمذجة التعابير الوجهية المختلفة من خلال استخراج الخصائص المميزة. أما في المرحلة الثانية، فتُستخدم هذه الخصائص لتصنيف صورة وجه جديدة بناءً على تعابيرها العاطفية باستخدام تقنيات مختلفة.

ومع ذلك، تُظهر الأنظمة التي تعمل في وضع مستقل عن الشخص، أي دون تخصيص مسبق للأفراد، معدلات تعرف أقل عادةً. يعود هذا الانخفاض في الأداء غالبًا إلى التباين في ملامح الوجه بين الأشخاص (العمر، الجنس، الشكل)، مما يسبب تشويشًا في عملية التعرف على التعابير. ومن أجل التغلب على هذه الصعوبة، يصبح اختيار الخصائص الملائمة أمرًا بالغ الأهمية.

تهدف هذه الأطروحة إلى تحسين أداء هذه الأنظمة من خلال اختيار مجموعة من الخصائص المناسبة للتعرف على التعابير، بحيث تُميز بفعالية بين التعابير مع تقليل تأثير التغيرات بين الأفراد. بالاعتماد على تقنيات اختيار الخصائص مثل أساليب الفلترة واللف والدمج، يسعى هذا العمل إلى تصميم نظام أكثر قوة وكفاءة في البيئات التي تتطلب فيها التعرف على تعابير الوجه بشكل موثوق ومستقل عن هوية الأفراد.

**الكلمات المفتاحية:** التعرف على تعابير الوجه، أنظمة العواطف، وضع مستقل عن الشخص، اختيار الخصائص، استخراج الخصائص، التصنيف

## ABSTRACT

Automatic Facial Expression Recognition (FER) is an essential task in various fields, such as psychology, human-computer interaction, and surveillance. An FER system generally operates in two main phases: the learning phase and the recognition phase. In the learning phase, the system aims to model different facial expressions by extracting discriminative features. In the recognition phase, the system classifies a new facial image according to its emotional expression using techniques such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), or Artificial Neural Networks (ANN).

However, FER systems operating in a person-independent mode, meaning without prior customization for specific individuals, typically result in lower recognition rates. This drop in performance is often due to the variability in human facial features (such as age, gender and morphology), which creates interference when recognizing expressions. To overcome this challenge, selecting relevant features becomes a crucial factor.

This thesis aims to enhance the performance of FER systems by selecting a set of relevant features for expression recognition that effectively distinguish between expressions while minimizing the influence of individual differences. By leveraging feature selection techniques such as Filter, Wrapper, or Embedded methods, this work seeks to design a more robust and efficient system in contexts where facial expressions must be reliably recognized independently of subjects' identities.

**Keywords :** Facial expression recognition, Emotion systems, Person-independent mode, Feature selection, Feature extraction, Classification.

## RESUME

La reconnaissance automatique des expressions faciales est une tâche essentielle dans de nombreux domaines tels que la psychologie, l'interaction homme-machine, et la surveillance. Un système de REF se divise principalement en deux phases : la phase d'apprentissage et la phase de reconnaissance. Dans un premier temps, l'apprentissage vise à modéliser les différentes expressions faciales en extrayant des caractéristiques discriminantes. Dans un second temps, la phase de reconnaissance permet de classifier une nouvelle image faciale en fonction de son expression émotionnelle à l'aide de techniques telles que les classifieurs K-Nearest Neighbors (KNN), Support Vector Machines (SVM), ou encore les réseaux de neurones artificiels (ANN).

Toutefois, les systèmes REF opérant en mode indépendant de la personne, c'est-à-dire sans personnalisation préalable aux individus, présentent généralement des taux de reconnaissance plus faibles. Cette baisse de performance est souvent due à la variabilité des visages selon les personnes (âge, sexe, morphologie), créant des interférences lors de la reconnaissance des expressions. Afin de surmonter cette difficulté, la sélection de paramètres pertinents devient un enjeu majeur.

Cette thèse vise à améliorer les performances de ces systèmes en sélectionnant un ensemble de paramètres pertinents pour la reconnaissance des expressions, qui permettent de distinguer efficacement les expressions tout en minimisant l'influence des variations entre individus. En s'appuyant sur des techniques de sélection de caractéristiques telles que les méthodes de type filter, wrapper ou embedded, ce travail vise à concevoir un système plus robuste et performant dans des contextes où les expressions faciales doivent être reconnues de manière fiable et indépendante de l'identité des sujets.

**Mots-clés :** Reconnaissance des expressions faciales, Systèmes d'émotions, Mode indépendant, Sélection de paramètres, Extraction de caractéristiques, Classification.

# TABLE DES MATIERES

|                                                                                                       |      |
|-------------------------------------------------------------------------------------------------------|------|
| REMERCIEMENTS                                                                                         | i    |
| ملخص                                                                                                  | ii   |
| ABSTRACT                                                                                              | iii  |
| RESUME                                                                                                | iv   |
| LISTE DES ACRONYMES                                                                                   | viii |
| LISTE DES FIGURES                                                                                     | xiv  |
| LISTE DES TABLEAUX                                                                                    | xv   |
| INTRODUCTION GENERALE                                                                                 | 1    |
| CHAPITRE I : RECONNAISSANCE AUTOMATIQUE DES EXPRESSIONS FACIALES                                      | 3    |
| I.1. Introduction                                                                                     | 4    |
| I.2. Emotions                                                                                         | 4    |
| I.2.1. Définition des émotions                                                                        | 4    |
| I.2.2. Théories des émotions                                                                          | 4    |
| I.2.3. Le processus émotionnel                                                                        | 6    |
| I.2.4. Communication des émotions                                                                     | 7    |
| I.3. Les expressions faciales des émotions                                                            | 9    |
| I.3.1. Définition                                                                                     | 9    |
| I.3.2. Physiologie du visage                                                                          | 12   |
| I.3.3. Aspect universel des émotions ‘Universalité des expressions faciales’                          | 13   |
| I.4. Système de reconnaissance automatique des expressions faciales                                   | 15   |
| I.4.1. Processus d’un système de reconnaissance des expressions faciales                              | 15   |
| I.4.1.1. Acquisition d’images                                                                         | 15   |
| I.4.1.2. Le prétraitement                                                                             | 15   |
| I.4.1.3. Extraction des caractéristiques faciales                                                     | 18   |
| I.4.1.3.1. Méthodes d’extraction des caractéristiques faciales pour les images statiques              | 18   |
| I.4.1.3.2. Méthodes d’extraction des caractéristiques faciales pour les séquences d’images dynamiques | 27   |
| I.4.1.4. Classification des expressions faciales                                                      | 31   |
| I.4.2. Bases de données REF                                                                           | 32   |
| I.4.3. REF en mode indépendant de la personne                                                         | 37   |
| I.4.3.1. Défis de la REF en mode indépendant de la personne                                           | 38   |
| I.5. Conclusion                                                                                       | 40   |

|                                                                                                                                                               |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| CHAPITRE II : SELECTION DES PARAMETRES PERTINENTS ET LEUR IMPACT SUR LA RECONNAISSANCE DES EXPRESSIONS FACIALES                                               | 41 |
| II.1. Introduction                                                                                                                                            | 42 |
| II.2. Concepts et définitions de la sélection de caractéristiques                                                                                             | 42 |
| II.2.1. Définition de la sélection de caractéristiques                                                                                                        | 42 |
| II.2.2. Processus de sélection des caractéristiques                                                                                                           | 42 |
| II.2.2.1. Génération de sous-ensembles                                                                                                                        | 43 |
| II.2.2.2. Évaluation des sous-ensembles                                                                                                                       | 44 |
| II.2.2.3. Critères d'Arrêt                                                                                                                                    | 45 |
| II.2.2.4. Validation des Résultats                                                                                                                            | 45 |
| II.2.3. L'influence des variables redondantes et corrélées en sélection de caractéristiques                                                                   | 46 |
| II.2.3.1. Effets de la corrélation sur la redondance                                                                                                          | 46 |
| II.2.3.1.1. Corrélation positive                                                                                                                              | 46 |
| II.2.3.1.2. Corrélation négative                                                                                                                              | 46 |
| II.2.3.2. Utilité des variables individuelles en contexte de combinaison                                                                                      | 46 |
| II.2.3.2.1. Variables apparemment inutiles                                                                                                                    | 46 |
| II.2.3.2.2. Approches de sélection                                                                                                                            | 46 |
| II.2.4. Typologie des méthodes de sélection de caractéristiques                                                                                               | 47 |
| II.2.4.1. Méthodes de filtrage                                                                                                                                | 47 |
| II.2.4.2. Méthodes enveloppantes (wrapper)                                                                                                                    | 49 |
| II.2.4.3. Méthodes Intégrées (Embedded Methods)                                                                                                               | 51 |
| II.2.5. Méthodes d'optimisation en sélection de caractéristiques                                                                                              | 53 |
| II.2.5.1. Méthodes traditionnelles d'optimisation                                                                                                             | 53 |
| II.2.5.1.1. Recherche exhaustive                                                                                                                              | 53 |
| II.2.5.1.2. Méthodes séquentielles                                                                                                                            | 53 |
| II.2.5.2. Méthodes inspirées de la nature                                                                                                                     | 53 |
| II.2.5.2.1. Algorithmes génétiques (Genetic Algorithms, GA)                                                                                                   | 54 |
| II.2.5.2.2. Optimisation par essaim particulaire (Particle Swarm Optimization, PSO)                                                                           | 54 |
| II.2.5.2.3. Optimisation par colonies de fourmis (Ant Colony Optimization, ACO)                                                                               | 54 |
| II.3. Pourquoi la sélection de caractéristiques est cruciale pour la REF                                                                                      | 54 |
| II.4. Application de la sélection de caractéristiques à la reconnaissance des expressions faciales                                                            | 55 |
| II.4.1. Sélection de caractéristiques basées sur des méthodes filtres                                                                                         | 55 |
| II.4.2. Sélection de caractéristiques basées sur des méthodes wrapper                                                                                         | 57 |
| II.4.3. Sélection de caractéristiques basées sur des méthodes embedded                                                                                        | 59 |
| II.4.4. Sélection de caractéristiques basées sur des méthodes hybride                                                                                         | 61 |
| II.5. Sélection de caractéristiques pour la reconnaissance des expressions faciales en mode indépendant de la personne : analyse et comparaison des approches | 61 |
| II.6. Conclusion                                                                                                                                              | 67 |
| CHAPITRE III : IMPLEMENTATION ET RESULTATS                                                                                                                    | 71 |
| III.1. Introduction                                                                                                                                           | 72 |
| III.2. Sélection de caractéristiques locales à l'aide de l'approche wrapper pour la reconnaissance des expressions faciales                                   | 72 |
| III.2.1. Principes du système REF proposé                                                                                                                     | 75 |



|                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------|-----|
| III.2.2. Résultats expérimentaux et discussion                                                                                           | 79  |
| III.2.2.1. L'impact de l'ajout d'une approche de sélection de caractéristiques sur la performance du système REF                         | 79  |
| III.2.2.2. Analyse du phénomène de la malédiction de la dimensionnalité                                                                  | 82  |
| III.2.2.3. Pertinence des régions faciales locales sur la performance du système REF                                                     | 84  |
| III.2.3. Comparaison avec les méthodes d'état de l'art                                                                                   | 85  |
| III.3. Un nouvel algorithme de reconnaissance des expressions faciales basé sur l'extraction et la sélection de caractéristiques par DWT | 86  |
| III.3.1. Le modèle proposé                                                                                                               | 87  |
| III.3.1.1. Prétraitement des images                                                                                                      | 88  |
| III.3.1.2. Extraction de caractéristiques                                                                                                | 88  |
| III.3.1.3. Sélection de caractéristiques                                                                                                 | 93  |
| III.3.1.4. Reconnaissance émotionnelle                                                                                                   | 95  |
| III.3.2. Résultats expérimentaux et discussion                                                                                           | 96  |
| III.3.2.1. Préparation des données                                                                                                       | 96  |
| III.3.2.2. Expériences                                                                                                                   | 97  |
| III.3.2.3. Résultats et discussion                                                                                                       | 98  |
| III.3.2.4. Évaluation des performances                                                                                                   | 99  |
| III.3.2.5. Comparaison des performances                                                                                                  | 100 |
| III.4. Conclusion                                                                                                                        | 102 |
| CONCLUSION GENERALE                                                                                                                      | 103 |
| REFERENCES BIBLIOGRAPHIQUES                                                                                                              | 105 |

## LISTE DES ACRONYMES

|                 |                                            |                                                          |
|-----------------|--------------------------------------------|----------------------------------------------------------|
| <b>2D :</b>     | Two-Dimensional                            | - Deux-Dimensionnel                                      |
| <b>2D-DCT :</b> | Two-Dimensional Discrete Cosine Transform  | - Transformée en Cosinus Discrète Bidimensionnelle       |
| <b>2D-DWT :</b> | Two-Dimensional Discrete Wavelet Transform | - Transformation en Ondelettes Discrète Bidimensionnelle |
| <b>4-OF :</b>   | 4-Optimal Filters                          | - 4 Filtres Optimaux                                     |

### A

|                                |                                                 |                                                                              |
|--------------------------------|-------------------------------------------------|------------------------------------------------------------------------------|
| <b>AAM :</b>                   | Active Appearance Model                         | - Modèle d'Apparence Active                                                  |
| <b>ACO :</b>                   | Ant Colony Optimization                         | - Optimisation par Colonies de Fourmis                                       |
| <b>AFEW :</b>                  | The Acted Facial Expressions in the Wild        | - Expressions Faciales Jouées en Conditions Réelles                          |
| <b>ANN :</b>                   | Artificial Neural Networks                      | - Réseaux de Neurones Artificiels                                            |
| <b>ASM :</b>                   | Active Shape Model                              | - Modèle de Formes Actives                                                   |
| <b>ASSM :</b>                  | Active and Statistical Shape Models             | - Modèles de Formes Actives et Statistiques                                  |
| <b>AU :</b>                    | Action Unit                                     | - Unité d'Action                                                             |
| <b>AUS:</b>                    | Action Unit Selectivity based Feature Selection | - Sélection de Caractéristiques Basée sur la Sélectivité des Unités d'Action |
| <b>A<math>\beta</math>CH :</b> | Adaptive $\beta$ -Hill Climbing                 | - Montée de Colline Adaptive $\beta$                                         |

### B

|                  |                                            |                                                                             |
|------------------|--------------------------------------------|-----------------------------------------------------------------------------|
| <b>BGWO :</b>    | Binary Gray Wolf Optimization              | - Optimisation par Meute de Loups Gris en Binaire                           |
| <b>BioID :</b>   | BioID Face Database                        | - Base de Données BioID pour la Reconnaissance Faciale                      |
| <b>BN :</b>      | Bayesian Network                           | - Réseau Bayésien                                                           |
| <b>BU-3DFE :</b> | Binghamton University 3D Facial Expression | - Base de Données des Expressions Faciales 3D de l'Université de Binghamton |

### C

|                   |                                                                                      |                                                                                                 |
|-------------------|--------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| <b>C3D :</b>      | Convolutional 3D Network                                                             | - Réseau Convolutif 3D                                                                          |
| <b>CAS-PEAL :</b> | Chinese Academy of Sciences - Pose, Expression, Accessories, and Lighting (Database) | - Académie Chinoise des Sciences - Pose, Expression, Accessoires et Éclairage (Base de données) |
| <b>CASME :</b>    | Chinese Academy of Sciences Micro-expression Database                                | - Base de Données des Micro-Expressions de l'Académie Chinoise des Sciences                     |
| <b>CASME2 :</b>   | Chinese Academy of Sciences Micro-expression Database (Version 2)                    | - Base de Données des Micro-Expressions de l'Académie Chinoise des Sciences (Version 2)         |
| <b>CCA :</b>      | Canonical Correlation Analysis                                                       | - Analyse Canonique des Corrélations                                                            |
| <b>CK+ :</b>      | Cohn-Kanade Plus                                                                     | - Base de données Cohn-Kanade Plus                                                              |
| <b>CNN :</b>      | Convolutional Neural Networks                                                        | - Réseaux de Neurones Convolutifs                                                               |
| <b>ConvLSTM :</b> | Convolutional Long Short-Term Memory                                                 | - Mémoire à Long Court Terme Convolutionnelle                                                   |
| <b>CFS :</b>      | Correlation-based Feature Selection                                                  | - Sélection de Caractéristiques Basée sur la Corrélations                                       |

## D

|                 |                                                          |                                                                              |
|-----------------|----------------------------------------------------------|------------------------------------------------------------------------------|
| <b>D-SIFT :</b> | Discriminative Scale Invariant Feature Transform         | - Transformation des Caractéristiques Invariantes à l'Échelle Discriminative |
| <b>DA :</b>     | Augmentation des Données                                 | - Data Augmentation                                                          |
| <b>DBN :</b>    | Dynamic Bayesian Network                                 | - Réseau Bayésien Dynamique                                                  |
| <b>DCPN :</b>   | Deep Coupled Pendulum Network                            | - Réseau Profond de Pendule Couplé                                           |
| <b>DCT :</b>    | Discrete Cosine Transform                                | - Transformée en Cosinus Discrète                                            |
| <b>DCNN :</b>   | Deep Convolutional Neural Network                        | - Réseau de Neurones Convolutifs Profond                                     |
| <b>DEMO :</b>   | Multi-Objective Differential Evolution                   | - Évolution Différentielle Multi-Objectif                                    |
| <b>DISFA :</b>  | Denver intensity of spontaneous facial action (Database) | - /                                                                          |
| <b>DNN :</b>    | Deep Neural Network                                      | - Réseau de Neurones Profond                                                 |
| <b>DT :</b>     | Decision Tree                                            | - Arbre de Décision                                                          |
| <b>DWE :</b>    | Discret Wavelet Energy                                   | - Énergie des Ondelettes Discrètes                                           |
| <b>DWT :</b>    | Discrete Wavelet Transform                               | - Transformation en Ondelettes Discrètes                                     |

## E

|               |                                 |                                                                  |
|---------------|---------------------------------|------------------------------------------------------------------|
| <b>EBGM :</b> | Elastic Bunch Graph Matching    | - Appariement Graphique à Faisceau Élastique                     |
| <b>EHMM :</b> | Embedded Hidden Markov Model    | - Modèle de Markov Caché Intégré                                 |
| <b>ELM :</b>  | Extreme Learning Machine        | - Machine d'Apprentissage Extrême                                |
| <b>ExpW :</b> | Expression in-the-Wild Database | - Base de Données des Expressions Faciales en Conditions Réelles |

## F

|                  |                                             |                                                                   |
|------------------|---------------------------------------------|-------------------------------------------------------------------|
| <b>FACS :</b>    | Facial Action Coding System                 | - Système de Codage des Actions Faciales                          |
| <b>FER2013 :</b> | Facial Expression Recognition 2013 database | - Base de Données de Reconnaissance des Expressions Faciales 2013 |
| <b>FDM :</b>     | Feature Disentangling Machine               | - Machine de Désentrelacement des Caractéristiques                |
| <b>FHOFO :</b>   | Fuzzy Histogram of Optical Flow Orientation | - Histogramme Flou de l'Orientation du Flux Optique               |
| <b>FNN :</b>     | Feedforward Neural Networks                 | - Réseaux de Neurones à Propagation Avant                         |
| <b>FOS:</b>      | Face-Occupancy-based Feature Selection      | - Sélection de Caractéristiques Basée sur l'Occupation du Visage  |
| <b>FS:</b>       | Feature Selection                           | - Sélection de Caractéristiques                                   |
| <b>FSV :</b>     | Feature selection via concave minimization  | - Sélection de Caractéristiques par Minimisation Concave          |
| <b>FSVM :</b>    | Fuzzy Support Vector Machine                | - Machine à Vecteurs de Support Floue                             |

## G

|                  |                                                     |                                                                                               |
|------------------|-----------------------------------------------------|-----------------------------------------------------------------------------------------------|
| <b>GA :</b>      | Genetic Algorithms                                  | - Algorithmes Génétiques                                                                      |
| <b>GA-SIFT :</b> | Geometric Algebra-Scale Invariant Feature Transform | - Transformation des Caractéristiques Invariantes à l'Échelle Basée sur l'Algèbre Géométrique |
| <b>GAN :</b>     | Generative Adversarial Network                      | - Réseau Antagoniste Génératif                                                                |

|                     |                                                                                             |                                                                                                        |
|---------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| <b>GEMEP-FERA :</b> | Geneva Multimodal Emotion Portrayal - Facial Expression Recognition and Analysis (Database) | - Base de Données GEMEP pour la Reconnaissance et l'Analyse des Expressions Faciales (Base de données) |
| <b>GPA :</b>        | Generalized Procrustes Analysis                                                             | - Analyse de Procruste Généralisée                                                                     |
| <b>GP :</b>         | Genetic Programming                                                                         | - Programmation Génétique                                                                              |
| <b>GR :</b>         | Gain Ratio                                                                                  | - Ratio de Gain                                                                                        |
| <b>GRU :</b>        | Gated Recurrent Unit                                                                        | - Unité Récurrente à Portes                                                                            |
| <b>GPU :</b>        | Graphics Processing Unit                                                                    | - Unité de Traitement Graphique                                                                        |
| <b>H</b>            |                                                                                             |                                                                                                        |
| <b>HAC :</b>        | Hierarchical Agglomerative Clustering                                                       | - Clustering Hiérarchique Agglomératif                                                                 |
| <b>HLACLF :</b>     | Higher-Order Local Auto-Correlation like features                                           | - Caractéristiques de Type Auto-Corrélation Locale d'Ordre Supérieur                                   |
| <b>HMM :</b>        | Hidden Markov Model                                                                         | - Modèle de Markov Caché                                                                               |
| <b>HOG :</b>        | Histogram of Oriented Gradients                                                             | - Histogramme des Gradients Orientés                                                                   |
| <b>HPED :</b>       | Histograms of Prominent Edge Directions                                                     | - Histogrammes des Directions de Bords Prominents                                                      |
| <b>hvnLBP :</b>     | horizontal-vertical neighborhood local binary pattern                                       | - Motif Binaire Local du Voisinage Horizontal-Vertical                                                 |
| <b>I</b>            |                                                                                             |                                                                                                        |
| <b>ISED :</b>       | Indian Spontaneous Expression Database                                                      | - Base de Données des Expressions Spontanées Indiennes                                                 |
| <b>IG :</b>         | Information Gain                                                                            | - Gain d'Information                                                                                   |
| <b>IUTWSVM :</b>    | Iterative Universum Twin Support Vector Machine                                             | - Machine à Vecteurs de Support Jumeaux Universum Itérative                                            |
| <b>J</b>            |                                                                                             |                                                                                                        |
| <b>J48 :</b>        | J48 Decision Tree Algorithm                                                                 | - Algorithme d'Arbre de Décision J48                                                                   |
| <b>JAFFE :</b>      | Japanese Female Facial Expression Database                                                  | - Base de Données des Expressions Faciales Féminines Japonaises                                        |
| <b>K</b>            |                                                                                             |                                                                                                        |
| <b>KDEF :</b>       | Karolinska Directed Emotional Faces database                                                | - Base de Données des Visages Émotionnels Dirigés de Karolinska                                        |
| <b>KLT :</b>        | Kanade-Lucas-Tomaci (Feature Tracker)                                                       | - Kanade-Lucas-Tomasi (Suivi de Caractéristiques)                                                      |
| <b>KNN :</b>        | K-Nearest Neighbors                                                                         | - K plus proches voisins                                                                               |
| <b>L</b>            |                                                                                             |                                                                                                        |
| <b>LAHCRR :</b>     | Late Acceptance Hill Climbing through Redundancy and Relevancy                              | - Montée de Colline à Acceptation Tardive Basée sur la Redondance et la Pertinence                     |
| <b>LBP :</b>        | Local Binary Patterns                                                                       | - Motifs binaires locaux                                                                               |
| <b>LBPH :</b>       | Local Binary Patterns Histogram                                                             | - Histogramme des Motifs Binaires Locaux                                                               |
| <b>LDA :</b>        | Linear Discriminant Analysis                                                                | - Analyse Discriminante Linéaire                                                                       |
| <b>LHCMA :</b>      | Late Hill Climbing based Memetic Algorithm                                                  | - Algorithme Mémétique Basé sur la Montée de Colline Tardive                                           |

|                    |                                                                            |                                                                                            |
|--------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| <b>LNEP :</b>      | Local Neighborhood Encoded Pattern                                         | - Motif Encodé du Voisinage Local                                                          |
| <b>LPDP :</b>      | Local Prominent Directional Pattern                                        | - Motif Directionnel Local Prominent                                                       |
| <b>LSTM :</b>      | Long Short-Term Memory                                                     | - Mémoire à Long Court Terme                                                               |
| <b>LWE :</b>       | Log Wavelet Energy                                                         | - Énergie des Ondelettes Logarithmiques                                                    |
| <b>M</b>           |                                                                            |                                                                                            |
| <b>MA :</b>        | Memetic Algorithms                                                         | - Algorithmes Mémétiques                                                                   |
| <b>MIFS :</b>      | Mutual Information Feature Selection                                       | - Méthode d'Information Mutuelle pour la Sélection des caractéristiques                    |
| <b>ML :</b>        | Machine Learning                                                           | - Apprentissage Automatique                                                                |
| <b>MLP :</b>       | Multi-Layer Perceptron                                                     | - Perceptron Multi-Couche                                                                  |
| <b>MMI :</b>       | MMI Facial Expression Database                                             | - Base de Données d'Expressions Faciales MMI                                               |
| <b>MOCBGWO :</b>   | Multiobjective and Opposition-Based Competitive Binary Gray Wolf Optimizer | - Optimiseur de Loup Gris Binaire Compétitif Basé sur l'Opposition et Multi-Objectifs      |
| <b>MPEG-4 :</b>    | Moving Picture Experts Group 4                                             | - Groupe d'Experts en Images Animées 4                                                     |
| <b>mGA :</b>       | Micro Genetic Algorithm                                                    | - Algorithme Génétique Micro                                                               |
| <b>mRMR :</b>      | Minimum Redundancy Maximum Relevance                                       | - Redondance Minimale et Pertinence Maximale                                               |
| <b>MSE :</b>       | Mean Squared Error                                                         | - Erreur Quadratique Moyenne                                                               |
| <b>MTCNN :</b>     | Multi-task Cascade Convolution Neural Network                              | - Réseau de Neurones Convolutifs en Cascade Multi-Tâches                                   |
| <b>Multi-PIE :</b> | Multiple Pose, Illumination, and Expression Database (Database)            | - Base de Données Multi-PIE (Multiples Poses, Éclairages et Expressions) (Base de données) |
| <b>N</b>           |                                                                            |                                                                                            |
| <b>NB :</b>        | Naïve Bayes                                                                | - Classificateur de Bayes Naïf                                                             |
| <b>NSGA II :</b>   | Non-dominated Sorted Genetic Algorithm II                                  | - Algorithme Génétique à Tri Non-Dominé II                                                 |
| <b>P</b>           |                                                                            |                                                                                            |
| <b>PCA :</b>       | Principal Component Analysis                                               | - Analyse en Composantes Principales                                                       |
| <b>PCC :</b>       | Pearson Correlation Coefficient                                            | - Coefficient de Corrélation de Pearson                                                    |
| <b>PDM :</b>       | Point Distribution Model                                                   | - Modèle de Distribution de Points                                                         |
| <b>PHOG :</b>      | Pyramid Histogram of Oriented Gradients                                    | - Histogramme Pyramidale des Gradients Orientés                                            |
| <b>PNN :</b>       | Probabilistic Neural Network                                               | - Réseau de Neurones Probabiliste                                                          |
| <b>PPDN :</b>      | Peak-Piloted Deep Network                                                  | - Réseau Profond Guidé par les Pics                                                        |
| <b>PSO :</b>       | Particle Swarm Optimization                                                | - Optimisation par Essaim Particulaire                                                     |
| <b>PWFP :</b>      | Pair-wise Feature Proximity                                                | - Proximité des Caractéristiques par Paires                                                |
| <b>Q</b>           |                                                                            |                                                                                            |
| <b>QMI :</b>       | Quadratic Mutual Information                                               | - Information Mutuelle Quadratique                                                         |

## R

|                 |                                      |                                                               |
|-----------------|--------------------------------------|---------------------------------------------------------------|
| <b>RAF-DB :</b> | Real-world Affective Face Database   | - Base de Données des Visages Affectifs en Conditions Réelles |
| <b>RaFD :</b>   | Radboud Faces Database               | - Base de Données des Visages de Radboud                      |
| <b>ReLU :</b>   | Rectified Linear Unit                | - Unité Linéaire Rectifiée                                    |
| <b>REF :</b>    | Facial Expression Recognition        | - Reconnaissance des Expressions Faciales                     |
| <b>ResNet :</b> | Residual Neural Networks             | - Réseau Neuronal Résiduel                                    |
| <b>RBF :</b>    | Radial Basis Function                | - Fonction de Base Radiale                                    |
| <b>RBFNN :</b>  | Radial Basis Function Neural Network | - Réseau de Neurones à Fonction de Base Radiale               |
| <b>RGB :</b>    | Red, Green, Blue (color model)       | - Rouge, Vert, Bleu (Modèle de couleur)                       |
| <b>RF :</b>     | Random Forest                        | - Forêt Aléatoire                                             |
| <b>RNA :</b>    | Artificial Neural Networks           | - Réseaux de Neurones Artificiels                             |
| <b>RNN :</b>    | Recurrent Neural Networks            | - Réseaux de Neurones Récurents                               |
| <b>ROI :</b>    | Region of Interest                   | - Région d'Intérêt                                            |

## S

|                     |                                                                        |                                                                                                               |
|---------------------|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| <b>SA :</b>         | Simulated Annealing                                                    | - Recuit Simulé                                                                                               |
| <b>SBoF-SSIFT :</b> | Spatial Bag of Features with Spatial Scale-Invariant Feature Transform | - Sac de Caractéristiques Spatiales avec Transformation des Caractéristiques Invariantes à l'Échelle Spatiale |
| <b>SBoF-SSURF :</b> | Spatial Bag of Features with Spatial Speeded-Up Robust Transform       | - Sac de Caractéristiques Spatiales avec Transformation Robuste Accélérée Spatiale                            |
| <b>SBoFs :</b>      | Spatial Bag of Features                                                | - Sac de Caractéristiques Spatiales                                                                           |
| <b>SBS :</b>        | Sequential Backward Selection                                          | - Sélection Séquentielle Arrière                                                                              |
| <b>SCA :</b>        | Sine Cosine Algorithm                                                  | - Algorithme Sinus-Cosinus                                                                                    |
| <b>SFFS :</b>       | Sequential Floating Forward Selection                                  | - Sélection Séquentielle Flottante Avant                                                                      |
| <b>SFEW :</b>       | Static Facial Expressions in the Wild                                  | - Expressions Faciales Statiques en Conditions Réelles                                                        |
| <b>SFHSA :</b>      | Supervised Filter Harmony Search Algorithm                             | - Algorithme de Recherche Harmonique avec Filtrage Supervisé                                                  |
| <b>SFS:</b>         | Sequential Forward Selection                                           | - Sélection Séquentielle Avant                                                                                |
| <b>SIFT :</b>       | Scale-Invariant Feature Transform                                      | - Transformation des Caractéristiques Invariantes à l'Échelle                                                 |
| <b>SHM :</b>        | Switching Hypothesis Measurement                                       | - Changement d'Hypothèses de Mesure                                                                           |
| <b>SMIC :</b>       | Spontaneous Micro-expression Database                                  | - Base de Données des Micro-Expressions Spontanées                                                            |
| <b>SNA :</b>        | <b>Autonomic Nervous System</b>                                        | - <b>Système Nerveux Autonome</b>                                                                             |
| <b>SRC :</b>        | Sparse Representation-based Classification                             | - Classification Basée sur la Représentation Parcimonieuse                                                    |
| <b>SURF :</b>       | Speeded-Up Robust Features                                             | - Caractéristiques Robustes Accélérées                                                                        |
| <b>SVR:</b>         | Support Vector Regression                                              | - Régression à Vecteurs de Support                                                                            |
| <b>SVM :</b>        | Support Vector Machines                                                | - Machines à Vecteurs de Support                                                                              |
| <b>SVD :</b>        | Singular Value Decomposition                                           | - Décomposition en Valeurs Singulières                                                                        |

## T

|              |                              |                                     |
|--------------|------------------------------|-------------------------------------|
| <b>TLD :</b> | Tracking Learning Detection  | - Suivi-Apprentissage-Détection     |
| <b>TIM :</b> | Temporal Interpolation Model | - Modèle d'Interpolation Temporelle |

## U

|                    |                                                                                                           |                                                                                                                                             |
|--------------------|-----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| <b>uLBP :</b>      | Uniform Local Binary Patterns                                                                             | - Motifs Binaires Locaux Uniformes                                                                                                          |
| <b>USTC-NVIE :</b> | University of Science and<br>Technology of China - Natural<br>Visible and Infrared Expression<br>Database | - Base de Données d'Expressions Faciales<br>Naturelles en Visible et Infrarouge de<br>l'Université des Sciences et<br>Technologies de Chine |

## V

|              |                       |                                |
|--------------|-----------------------|--------------------------------|
| <b>VGG :</b> | Visual Geometry Group | - Groupe de Géométrie Visuelle |
|--------------|-----------------------|--------------------------------|

## W

|                  |                                                             |                                                                                                     |
|------------------|-------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| <b>WCC :</b>     | Wavelet Cepstral Coefficient                                | - Coefficient Cepstral par Ondelettes                                                               |
| <b>WFACOFs :</b> | Wrapper-Filter Ant Colony<br>Optimization Feature Selection | - Sélection de Caractéristiques par<br>Optimisation de Colonie de Fourmis en<br>Mode Wrapper-Filter |
| <b>WLD :</b>     | Weber Local Descriptor                                      | - Descripteur Local de Weber                                                                        |

## LISTE DES FIGURES

|                                                                                                                                                                                                                |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure I.1: Différents canaux de communication des émotions. ....                                                                                                                                              | 7   |
| Figure I.2: Les muscles peauciers du visage [16]. ....                                                                                                                                                         | 10  |
| Figure I.3: Les trois phases de l'activation musculaire du visage. ....                                                                                                                                        | 12  |
| Figure I.4: Schéma du système REF conventionnel [27] ....                                                                                                                                                      | 15  |
| Figure I.5: Schéma de repérage des caractéristiques faciales utilisé dans la mise en œuvre classique de l'ASM [36]. ....                                                                                       | 19  |
| Figure I.6: Modèle d'Apparence Active (une image d'entraînement étiquetée fournit un patch sans forme et un ensemble de points) [41]. ....                                                                     | 21  |
| Figure I.7: Scale Invariant Feature Transform (les positions initiales de la caractéristique faciale, la caractéristique SIFT et les positions du SIFT amélioré) [52]. ....                                    | 22  |
| Figure I.8: Méthode d'extraction de caractéristiques de modèle binaire local (LBP) [60]. ....                                                                                                                  | 24  |
| Figure I.9: Histogramme de l'extraction du gradient orienté à partir du visage [66]. ....                                                                                                                      | 26  |
| Figure I.10: Le modèle temporel de l'expression « sourire » [70]. ....                                                                                                                                         | 27  |
| Figure I.11: Flux optique d'une séquence issue de la base de données CK+ [27]. ....                                                                                                                            | 28  |
| Figure I.12: Exemples de bases de données de la reconnaissance des expressions faciales [27]. ....                                                                                                             | 33  |
| Figure II.1: Les quatre étapes clés de la sélection des caractéristiques [138]. ....                                                                                                                           | 43  |
| Figure II.2: Approche basée sur les filtres [142]. ....                                                                                                                                                        | 47  |
| Figure II.3: Approche basée sur les wrappers [142]. ....                                                                                                                                                       | 49  |
| Figure III.1: Un aperçu schématique du système REF proposé dans [200]. ....                                                                                                                                    | 73  |
| Figure III.2: Normalisation des repères. ....                                                                                                                                                                  | 74  |
| Figure III.3: Aperçu schématique de notre système REF proposé. ....                                                                                                                                            | 75  |
| Figure III.4: Les trois premières itérations du processus de sélection des caractéristiques intégrées dans notre système REF proposé. ....                                                                     | 78  |
| Figure III.5: (a) Comparaison de la précision entre le système REF de Silva et al. [200] et le système proposé. ....                                                                                           | 82  |
| Figure III.6: Variation de la certitude en fonction du nombre de caractéristiques ....                                                                                                                         | 83  |
| Figure III.7: Caractéristiques les plus pertinentes en fonction de leur position sur le visage en utilisant (classificateur LDA). ....                                                                         | 84  |
| Figure III.8: Organigramme du système proposé. ....                                                                                                                                                            | 88  |
| Figure III.9: Schéma fonctionnel de la méthode d'extraction de caractéristiques proposée. ....                                                                                                                 | 89  |
| Figure III.10: Famille des ondelettes de Daubechies. ....                                                                                                                                                      | 89  |
| Figure III.11: Famille des ondelettes de Coiflet. ....                                                                                                                                                         | 90  |
| Figure III.12: Famille des ondelettes de Symlet. ....                                                                                                                                                          | 90  |
| Figure III.13: Décomposition 2D-DWT à un niveau sur une image d'entrée. ....                                                                                                                                   | 91  |
| Figure III.14: Concept général de l'algorithme de sélection de caractéristiques adopté dans notre système proposé. ....                                                                                        | 94  |
| Figure III.15: Représentation du concept du noyau RBF dans le SVM. ....                                                                                                                                        | 95  |
| Figure III.16: Variation du taux de reconnaissance (émotion de surprise) en fonction du nombre de caractéristiques. ....                                                                                       | 100 |
| Figure III.17: Comparaison entre les taux de reconnaissance des expressions faciales (en %) obtenus par notre modèle proposé et ceux obtenus par le modèle basé sur le LBP dans la base de données JAFFE. .... | 101 |
| Figure III.18: Comparaison entre les taux de reconnaissance des expressions faciales (en %) obtenus par notre modèle proposé et ceux obtenus par le modèle basé sur le LBP dans la base de données CK+. ....   | 101 |



## LISTE DES TABLEAUX

|                                                                                                                                                                                                                           |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Tableau I.1: Les principales théories historiques et courante de l'émotion.....                                                                                                                                           | 5   |
| Tableau I.2: Catégorisation de communication des émotions via le canal auditif.....                                                                                                                                       | 8   |
| Tableau I.3: Catégorisation de communication des émotions via le canal visuel. ....                                                                                                                                       | 8   |
| Tableau I.4: Catégorisation de communication des émotions via le canal corporel. ....                                                                                                                                     | 8   |
| Tableau I.5: Classification des muscles du visage responsable de l'apparition des expressions faciales. ....                                                                                                              | 10  |
| Tableau I.6: Liste des émotions de base proposées par différents auteurs. ....                                                                                                                                            | 13  |
| Tableau I.7: Les caractéristique des expressions faciales des émotions de base (Ekman & Wezowski) .....                                                                                                                   | 14  |
| Tableau I.8: Les techniques couramment utilisées pour le prétraitement des données.....                                                                                                                                   | 16  |
| Tableau I.9: Les modèles de machine learning les plus connus dans la littérature.....                                                                                                                                     | 31  |
| Tableau I.10: Aperçu de bases de données de la reconnaissance des expressions faciales. ....                                                                                                                              | 33  |
| Tableau I.11: Les bases de données les plus connues dédiées à l'étude des micro-expressions faciales. ....                                                                                                                | 37  |
| Tableau II.1: Un algorithme de filtrage généralisé [138].....                                                                                                                                                             | 48  |
| Tableau II.2: Un algorithme wrapper généralisé [138].....                                                                                                                                                                 | 50  |
| Tableau II.3: Un algorithme généralisé de type intégré (embedded). ....                                                                                                                                                   | 51  |
| Tableau II.4: Avantages et inconvénients des approches de sélection de caractéristiques [138]. ...                                                                                                                        | 52  |
| Tableau II.5: Tableau récapitulatif des travaux de recherche adoptant des méthodes de sélection de caractéristiques pour les systèmes de reconnaissance des expressions faciales en mode indépendant de la personne. .... | 68  |
| Tableau III.1: Nombre de points de repère extraits pour chaque région du visage.....                                                                                                                                      | 73  |
| Tableau III.2: Spécifications techniques et environnement.....                                                                                                                                                            | 79  |
| Tableau III.3: Résultats de classification en utilisant le classificateur KNN. ....                                                                                                                                       | 80  |
| Tableau III.4: Résultats de classification utilisant le classificateur DT.....                                                                                                                                            | 80  |
| Tableau III.5: Résultats de classification utilisant le classificateur NB.....                                                                                                                                            | 80  |
| Tableau III.6: Résultats de classification utilisant le classificateur LDA.....                                                                                                                                           | 80  |
| Tableau III.7: Comparaison des performances entre le système de REF de Silva et al. [200] et le système proposé. ....                                                                                                     | 81  |
| Tableau III.8: Comparaison des temps de réponse entre le système REF de Silva et al. [200] et le système proposé.....                                                                                                     | 81  |
| Tableau III.9: Les dix premières caractéristiques sélectionnées “ <i>f<sub>si</sub></i> ” et leurs valeurs de certitude obtenues en utilisant le classificateur LDA pour le cas optimal. ....                             | 84  |
| Tableau III.10: Comparaison entre les performances du système REF proposé et celles des systèmes les plus pertinents rapportés dans la littérature. ....                                                                  | 85  |
| Tableau III.11: Nombre d'images (par émotion) utilisées dans nos expériences à partir des bases de données CK+ et JAFFE. ....                                                                                             | 97  |
| Tableau III.12: Résultats de performance obtenus pour chaque émotion.....                                                                                                                                                 | 98  |
| Tableau III.13: Taux de reconnaissance et nombre de caractéristiques obtenues pour chaque émotion de la base de données JAFFE avant et après l'adoption de l'approche de sélection.....                                   | 99  |
| Tableau III.14: Taux de reconnaissance et nombre de caractéristiques obtenues pour chaque émotion dans la base de données CK+ avant et après l'adoption de l'approche de sélection.....                                   | 99  |
| Tableau III.15: Résultats de la simulation du critère de temps de calcul pour les bases de données JAFFE et CK+. ....                                                                                                     | 102 |

# INTRODUCTION GENERALE

---

## Motivation

Avec les avancées technologiques en intelligence artificielle et en vision par ordinateur, la reconnaissance des expressions faciales (REF) a suscité un intérêt croissant dans divers domaines. Les systèmes REF sont déployés dans des applications variées telles que la psychologie, la sécurité, les interactions homme-machine et la réalité virtuelle. Cependant, leur performance est souvent limitée par la variabilité interindividuelle, ce qui entrave leur généralisation à des populations diverses. Une REF robuste et indépendante de la personne permettrait d'améliorer significativement les performances dans des applications critiques telles que la surveillance sécuritaire, le diagnostic psychologique et les interactions homme-machine adaptatives. Cette motivation nous pousse à explorer des solutions permettant de rendre les systèmes REF plus performants et généralisables.

## Problématique

L'un des principaux obstacles à la reconnaissance efficace des expressions faciales en mode indépendant est la grande variabilité interindividuelle, notamment en termes de morphologie faciale, de couleur de peau, d'âge et de sexe. Cette variabilité rend difficile l'extraction de caractéristiques universellement discriminantes pour chaque expression.

Par ailleurs, les systèmes de reconnaissance doivent concilier précision et efficacité computationnelle. Réduire la dimensionnalité des données sans compromettre la performance de classification est un enjeu clé pour rendre ces systèmes plus rapides et adaptés aux applications en temps réel.

Dans ce contexte, la question essentielle à laquelle nous tentons de répondre est la suivante : comment sélectionner les caractéristiques les plus pertinentes afin d'optimiser la reconnaissance des expressions faciales tout en assurant une robustesse en mode indépendant des personnes ?

## Contribution

Cette thèse se concentre sur la sélection de paramètres pertinents afin d'améliorer la performance des systèmes de reconnaissance des expressions faciales en mode indépendant des personnes. L'objectif principal est de réduire la dimensionnalité des données tout en conservant les caractéristiques les plus discriminantes pour les expressions émotionnelles, minimisant ainsi l'impact des variations individuelles.

Les contributions de cette thèse sont les suivantes :

- **Une approche de sélection locale des caractéristiques basée sur une méthode wrapper pour la REF :** Nous évaluons l'impact de la sélection des caractéristiques locales, notamment celles issues des régions du visage comme les sourcils, les yeux et la bouche. Nos expériences sur la base de données MUG montrent une amélioration significative des performances, avec une réduction de 50% du nombre de caractéristiques et une précision allant jusqu'à 100% avec LDA.
- **Un nouvel algorithme de reconnaissance des expressions faciales basé sur l'extraction et la sélection des caractéristiques issues de la transformation en ondelettes discrètes (DWT) :** Nous proposons une méthode combinant l'extraction des caractéristiques par DWT avec une approche wrapper pour la sélection des traits les plus pertinents. Les tests réalisés sur les bases JAFFE et CK+ en mode indépendant des personnes démontrent une bonne amélioration du taux de reconnaissance et du temps de réponse, permettant ainsi le développement d'un système REF en temps réel.

## Plan du document

Le présent travail est structuré en trois chapitres principaux :

### 1. Reconnaissance automatique des expressions faciales

Ce chapitre introduit les concepts fondamentaux de la reconnaissance des expressions faciales, les bases théoriques des émotions et les différentes techniques d'analyse faciale.

### 2. Sélection des Paramètres Pertinents et leur Impact sur la Reconnaissance des Expressions Faciales

Ce chapitre présente une vue d'ensemble des approches de sélection des caractéristiques et explore leur application à la reconnaissance des expressions faciales. Il analyse les méthodes issues de la littérature et évalue leur impact sur la réduction de la dimensionnalité et l'optimisation des performances de reconnaissance.

### 3. Implémentation et résultats

Ce dernier chapitre décrit les expérimentations réalisées, l'analyse des performances des approches proposées et une discussion des résultats obtenus.

En conclusion, cette thèse apporte une contribution significative à l'amélioration des systèmes de reconnaissance des expressions faciales, en mettant l'accent sur la sélection des paramètres les plus discriminants pour assurer une robustesse accrue en mode indépendant des personnes. Les travaux réalisés ouvrent également des perspectives pour l'optimisation de systèmes REF adaptatifs et plus efficaces.

# Chapitre I

## Reconnaissance Automatique Des Expressions Faciales

---

|                                                                                                       |    |
|-------------------------------------------------------------------------------------------------------|----|
| I.1. Introduction                                                                                     | 4  |
| I.2. Emotions                                                                                         | 4  |
| I.2.1. Définition des émotions                                                                        | 4  |
| I.2.2. Théories des émotions                                                                          | 4  |
| I.2.3. Le processus émotionnel                                                                        | 6  |
| I.2.4. Communication des émotions                                                                     | 7  |
| I.3. Les expressions faciales des émotions                                                            | 9  |
| I.3.1. Définition                                                                                     | 9  |
| I.3.2. Physiologie du visage                                                                          | 12 |
| I.3.3. Aspect universel des émotions ‘Universalité des expressions faciales’                          | 13 |
| I.4. Système de reconnaissance automatique des expressions faciales                                   | 15 |
| I.4.1. Processus d’un système de reconnaissance des expressions faciales                              | 15 |
| I.4.1.1. Acquisition d’images                                                                         | 15 |
| I.4.1.2. Le prétraitement                                                                             | 15 |
| I.4.1.3. Extraction des caractéristiques faciales                                                     | 18 |
| I.4.1.3.1. Méthodes d’extraction des caractéristiques faciales pour les images statiques              | 18 |
| I.4.1.3.2. Méthodes d’extraction des caractéristiques faciales pour les séquences d’images dynamiques | 27 |
| I.4.1.4. Classification des expressions faciales                                                      | 31 |
| I.4.2. Bases de données REF                                                                           | 32 |
| I.4.3. REF en mode indépendant de la personne                                                         | 37 |
| I.4.3.1. Défis de la REF en mode indépendant de la personne                                           | 38 |
| I.5. Conclusion                                                                                       | 40 |

## I.1. Introduction

L'intérêt pour l'intégration des émotions en intelligence artificielle remonte à 1981, lorsque Sloman et Croucher ont souligné pour la première fois la nécessité de prendre en compte la dimension émotionnelle dans les systèmes informatiques. Depuis, la reconnaissance automatique des expressions faciales a connu des avancées significatives, en particulier grâce aux progrès de l'apprentissage automatique.

Dans ce chapitre, nous explorons les fondements théoriques des émotions, leur expression à travers les traits du visage et les approches utilisées pour leur reconnaissance automatique. Nous commencerons par explorer les concepts fondamentaux des émotions, leurs théories et les mécanismes sous-jacents à leur expression. Ensuite, nous examinerons les approches classiques utilisées pour l'analyse et la reconnaissance des expressions faciales. Enfin, nous discuterons des principaux défis liés à la REF en mode indépendant des personnes, un enjeu majeur pour améliorer la robustesse et la généralisation des systèmes.

## I.2. Emotions

### I.2.1. Définition des émotions

Dans la vie courante, le terme émotion désigne le centre de toute relation, exprimée sur un mode verbal ou analogique [1]. Bien que tout le monde sache ce qu'est une émotion, mais sa définition scientifique reste difficile. « On peut affirmer sans exagération que, scientifiquement, nous ne comprenons absolument rien aux émotions, que nous n'avons pas l'ombre d'une théorie sur les émotions en général ou de telle émotion en particulier » [2]. Le terme 'émotion' au grand siècle désignait essentiellement des manifestations collectives, comme mouvement, troubles et sédition, avec un sens proche de celui qu'on donne encore aujourd'hui au terme 'émeute'. La première apparition du mot 'émotion' était dans le dictionnaire Littré (1883) avec la signification de « mouvement qui se passe dans la population », alors que la notion de « mouvement moral qui trouble et agite » n'apparaît que comme un sens dérivé[3]. Le terme 'émotion' est donc d'origine récente. Le petit Robert propose : « L'émotion est un état affectif intense, caractérisé par une brusque perturbation physique et mentale où sont abolies, en présence de certaines excitations ou représentations très vives, les réactions appropriées d'adaptation à l'événement. (Au sens affaibli) : État affectif, plaisir ou douleur, nettement prononcé. »[1], alors que Larousse nous en offre la définition suivante : « Trouble subit, agitation passagère causés par un sentiment vif de peur, de surprise, de joie, etc. » ou encore « Réaction affective transitoire d'assez grande intensité, habituellement provoquée par une stimulation venue de l'environnement ».

Une grande variété de définitions du mot 'émotion' a été proposée par les chercheurs spécialistes. En 1981, Kleinginna et Kleinginna ont recensé 92 définitions, qu'ils ont regroupées en 11 catégories : affectives, cognitives, liées aux stimuli externes, physiologiques, comportementales, perturbatrices, adaptatives, multi-aspects, restrictives, motivationnelles et sceptiques. Ces définitions mettent en lumière les différentes approches adoptées pour caractériser l'émotion, qu'elle soit perçue comme une réponse physiologique, un processus cognitif ou encore un mécanisme d'adaptation [4].

De manière synthétique, l'émotion peut être définie comme une réaction affective et corporelle déclenchée par un stimulus externe ou une pensée.

### I.2.2. Théories des émotions

L'étude de la littérature en psychologie des émotions est une phase importante pour une meilleure compréhension de la perception humaine [5]. En retour, cela permet d'améliorer la qualité de l'automatisation du traitement émotionnel humain. Dans ce contexte, nous avons collecté les principales théories historiques et courantes de l'émotion, tout en faisant les regroupées en catégories ([Tableau I.1](#))

- **Théories Physiologiques** : soutiennent que l'activation physiologique entraînent le déclenchement et le déroulement du processus émotionnel.
- **Théories Adaptatives** : suggèrent que les émotions permettent l'adaptation de l'organisme à son environnement.
- **Théories Neuropsychologiques** : affirment que les processus nerveux centraux sont responsables des émotions.
- **Théories Expressives** : focalisent sur la détermination des émotions de base en étudiant les expressions faciales émotionnelles.
- **Théories Cognitivo-physiologique** : considèrent le processus émotionnel comme résultant d'une activation physiologique et une activité cognitive.
- **Théories de l'évaluation** : envisagent la cognition de la séquence émotionnelle comme la perception et l'évaluation de la signification d'un événement pour le bien-être d'une personne [6].
- **Théories Schématiques** : suggèrent l'existence d'un ensemble d'informations émotionnelles stockées en mémoire responsables de la formation des émotions.

**Tableau I.1: Les principales théories historiques et courante de l'émotion.**

|                     |                                            |                                      |
|---------------------|--------------------------------------------|--------------------------------------|
| Théories classiques | Théories Physiologiques                    | Théorie périphéraliste               |
|                     |                                            | Théorie centraliste                  |
|                     | Théorie biologique et adaptative           | Théorie darwinienne (évolutionniste) |
|                     | Théories Neuropsychologiques               | Le circuit de Papez                  |
|                     |                                            | Théorie de Panksepp                  |
|                     |                                            | Théorie de LeDoux                    |
|                     |                                            | Théorie des marqueurs somatiques     |
| Théories courantes  | Théories Expressives (Rétroaction Faciale) | Théorie des émotions de base         |
|                     |                                            | Théorie de l'amplification           |
|                     |                                            | Théorie différentielle               |
|                     | Théories cognitivo-physiologique           | Théorie bifactorielle                |
|                     | Théories de l'évaluation [6]               | Théorie de l'évaluation cognitive    |
|                     |                                            | Théories dimensionnelles             |
|                     |                                            | Théorie de l'émotion en composantes  |

|  |                           |                                           |
|--|---------------------------|-------------------------------------------|
|  | Théories schématiques [6] | Théorie des émotions perceptuelle motrice |
|  |                           | Modèle de la mémoire en réseau            |
|  |                           | Théorie bio-informationnelle des émotions |

### I.2.3. Le processus émotionnel

Comme montre la section (I.2.2), les auteurs ne s'accordent pas sur un processus émotionnel exact. De ce fait, il nous semble pertinent de faire appel à un modèle suivant le dérouls temporel prototypique le plus généralement impliqué dans les processus émotionnels. Le modèle intégratif proposé par Planalp (1999) vise à réguler les interactions entre l'individu et son environnement, il présente le processus émotionnel en cinq phases essentielles : (1) l'apparition d'un stimulus, (2) l'évaluation émotionnelle, (3) changements au niveau du système nerveux autonome, (4) une tendance à l'action, une réponse émotionnelle et un sentiment subjectif, (5) la régulation émotionnelle. [7]

- *Apparition d'un stimulus :*

La première phase de ce processus est le déclenchement de l'émotion suite à un stimulus interne ou externe. Le stimulus peut être présenté sous forme d'une image, un son, un goût, une odeur ou encore un souvenir ou une bonne ou mauvaise nouvelle.

- *L'évaluation émotionnelle :*

La deuxième phase est l'évaluation émotionnelle. Un même stimulus peut avoir différentes significations pour le sujet, selon le contexte et l'état interne du sujet [8], d'où une évaluation est nécessaire pour confère à l'événement une signification particulière.

- *Les changements au niveau du système nerveux autonome :*

La troisième étape est caractérisée par le fonctionnement du système nerveux autonome (SNA). Le SNA se décompose en deux sous-systèmes : le système nerveux sympathique (l'excitateur) et le système nerveux parasympathique (l'inhibiteur). Ces deux systèmes interagissent de manière antagoniste pour générer des niveaux variables d'activation physiologique[9].

Remarque : Cette étape peut ne pas apparaître (elle ne semble pas obligatoire) et peut apparaître différemment d'un individu à l'autre ou d'une situation à l'autre.

- *La tendance à l'action, la réponse émotionnelle et le sentiment subjectif :*

La quatrième phase est subdivisée en trois étapes distinctes : la tendance à l'action, la réponse émotionnelle et le sentiment subjectif.

La tendance à l'action, qui représente des informations portant sur l'action que l'individu s'apprête à effectuer rationnellement avec son environnement. Selon Frijda, une tendance à l'action peut mener à plusieurs comportements différents ce qui conduit à une variabilité des réponses émotionnelles[10].

Les réponses émotionnelles peuvent bien être des réponses physiologiques (changement de la fréquence cardiaque, changement du rythme respiratoire, changement de la température du corps, la conductance de la peau ...), comportementales (comportements expressifs (facial, vocal, gestuel, postural) ou des

actions effectuées pour atteindre les buts poursuivis), motrices (sauter de joie, taper des mains, fuir, éviter...) ou cognitives (changements attentionnels, biais de mémoire).

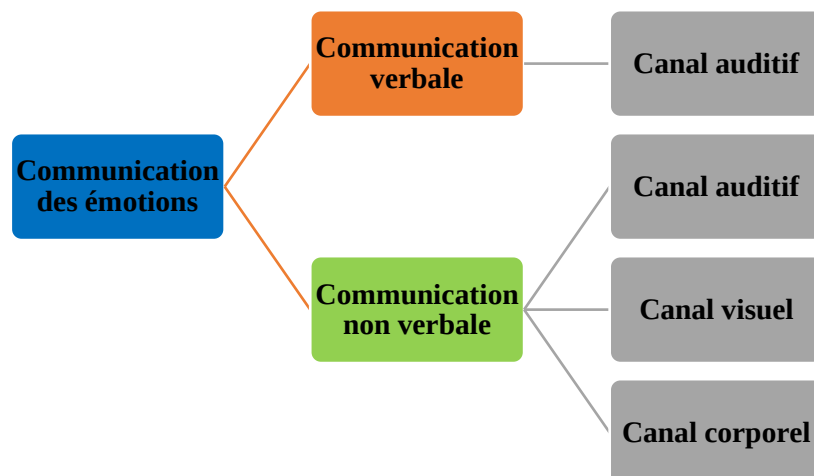
Finalement, après « avoir » une émotion, l'individu « ressent » l'émotion, en autre terme il ressent qu'il est dans un état différent (se sentir heureux, se sentir peur, se sentir en colère...), c'est ce qu'on appelle le sentiment subjectif. Ce dernier peut être envisagé comme l'infime partie du processus émotionnel qui atteint le seuil de la conscience et de la verbalisation.

- *La régulation émotionnelle :*

La régulation des émotions a été définie comme des activités dans lesquelles les individus s'engagent pour influencer la nature de leurs émotions, le moment et la façon dont ils en font l'expérience et l'expriment[11] in [12]. Elle se concentre sur la diminution des expériences négatives et l'amplification des expériences positives. Ce processus puisse être conscient ou inconscient et automatique. De nombreuses techniques ont été étudiées, Ryan et Deci en 2017, ont proposé la théorie de l'autodétermination « self-determination theory's »[13] in [12]. Cette dernière décrit trois types de régulation des émotions : la régulation intégrative « integrative emotion regulation », qui considère les émotions comme informations qui sont portées de la conscience ; la régulation contrôlée « controlled emotion regulation », qui se concentre sur la diminution des émotions par l'évitement, la suppression ou la réévaluation ; et la régulation des émotions amotivées « amotivated emotion regulation », dans laquelle les émotions sont incontrôlées.

#### I.2.4. Communication des émotions

L'expérience émotionnelle se caractérise par un changement dans l'environnement qui se manifeste par des modifications internes (physiologique : tension, détente, ...) ou externes où l'être humain communique ses émotions via de multiples canaux. Ces derniers constituent un support pour véhiculer l'état émotionnel entre l'émetteur et le récepteur. Le schéma ci-dessous (Figure I.11) présente les différents canaux verbaux et non-verbaux assurant cette tâche (l'encodage ou le décodage du message émotionnel) [14], [15].



**Figure I.1: Différents canaux de communication des émotions.**

Par opposition à la communication verbale, la communication non verbale permet la communication entre personnes sans l'usage de mots. Elle représente l'ensemble des signes qui transitent par des



expressions du visage, des gestes, des postures, du ton de la voix, des mouvements du corps... Les (Tableaux I.2, I.3 et I.4), représentent une catégorisation de communication des émotions par les différents canaux (visuel, auditif et corporel).

**Tableau I.2: Catégorisation de communication des émotions via le canal auditif.**

| Le canal auditif              |                                                                |
|-------------------------------|----------------------------------------------------------------|
| Message verbal en face à face |                                                                |
| Signaux conversationnels      | Accentuation de mots, ponctuation, marqueurs d'une question... |
| Emissions vocales             | Le bâillement, le rire, le cri, la toux...                     |

Remarque : la différence entre le canal auditif de la communication verbale et celui de la communication non verbale est que le premier transporte le contenu du discours alors que le deuxième transporte la façon de parler.

**Tableau I.3: Catégorisation de communication des émotions via le canal visuel.**

| Le canal visuel                   |                      |                                                       |
|-----------------------------------|----------------------|-------------------------------------------------------|
| Signaux émis au niveau de la face | Expressions faciales |                                                       |
|                                   | Regard               | Les mouvements des yeux (direction, fréquence, durée) |
|                                   |                      | La dilatation de la pupille                           |

**Tableau I.4: Catégorisation de communication des émotions via le canal corporel.**

| Le canal corporel        |                                          |
|--------------------------|------------------------------------------|
| Actions comportementales | Posture, gestes ...                      |
| Actions motrices         | Sauter de joie, taper des mains, fuir... |

Les travaux de recherche en psychologie ont démontré que les moyens non verbaux et en particulier les modifications au niveau de la face sont le principal moyen d'expression des émotions, cette notion ait été combattue dès les premières études à la suite des travaux de Darwin (1872). Selon l'école de Palo Alto, les moyens verbaux ne porteraient que sur 30 à 35% de la signification du message, tandis que les moyens non verbaux véhiculeraient 65 à 70% de l'information. Les travaux du professeur de psychologie à l'université de Californie Albert Mehrabian (1967), ont montré aussi que la partie non verbale aurait une importance fondamentale, 93 % de communication non verbale et 7 % de communication verbale. Ce rapport se décline comme suit : 55 % du message émotionnel est communiqué par l'expression faciale alors que 38% par le paralangage et 7 % seulement par le canal

linguistique. Par conséquent, les expressions faciales sont le moyen le plus couramment utilisé par les humains pour transmettre leur état émotionnel.

« La parole peut dissimuler la réalité, alors que l'expression la révèle » **Jean Cosnier**

Étant le visage la partie la plus expressive et communicative de l'être humain, ce travail de recherche s'intéresse particulièrement à la reconnaissance des émotions à travers les expressions faciales. Dans ce qui suit, nous présenterons un état de l'art sur les expressions faciales des émotions.

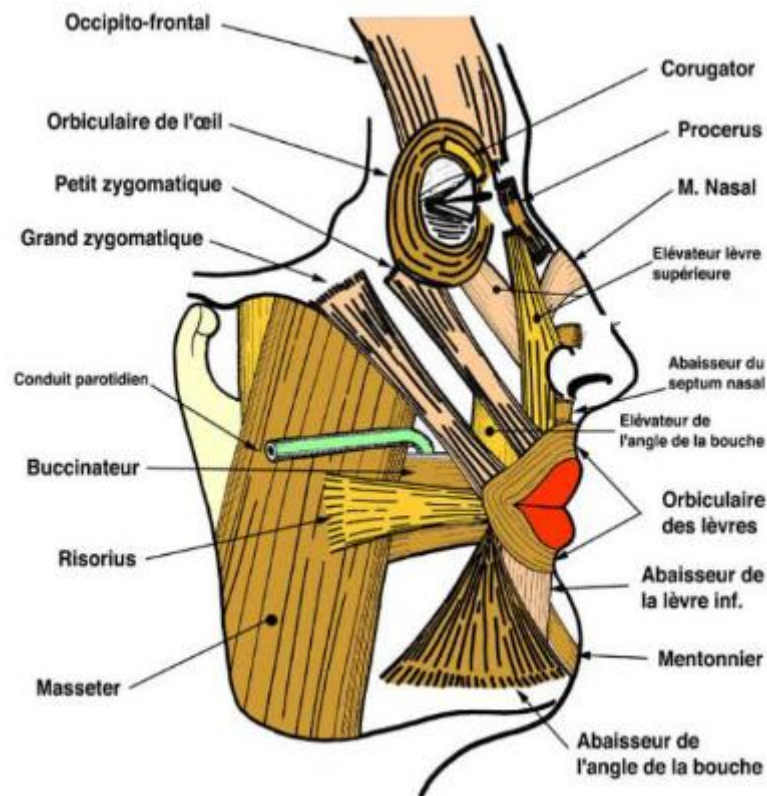
### **I.3. Les expressions faciales des émotions**

Le visage est l'un des canaux les plus puissants de communication non verbale. Les sourcils, le nez, la bouche et les yeux sont des régions de visage qui portent autant d'informations qui permettent de faire plusieurs déductions et de récupérer plusieurs informations sur les différents états émotionnels.

#### **I.3.1. Définition**

Une expression faciale est une manifestation visible de l'état émotionnel, c'est un changement dans le visage, résulte de l'activation et le mouvement d'un ou plusieurs muscles composant le visage. Ce dernier comprend un ensemble d'environ 20 muscles peauciers (Figure I.2) [16] répartis en cinq groupes[17] :

- Muscle épicroïque
- Les muscles orbitaires
- Les muscles du nez
- Les muscles de la bouche et du menton
- Muscle du cou



**Figure I.2: Les muscles peauciers du visage [16]**

Les muscles peauciers du visage sont des muscles dont au moins une des extrémités s'attache à la partie intérieure de la peau, formant une nappe presque ininterrompue qui donnent des indices observables de l'état interne. La classification de ces muscles[16] [17] est détaillée dans le tableau ci-dessous (Tableau I.5).

**Tableau I.5: Classification des muscles du visage responsable de l'apparition des expressions faciales.**

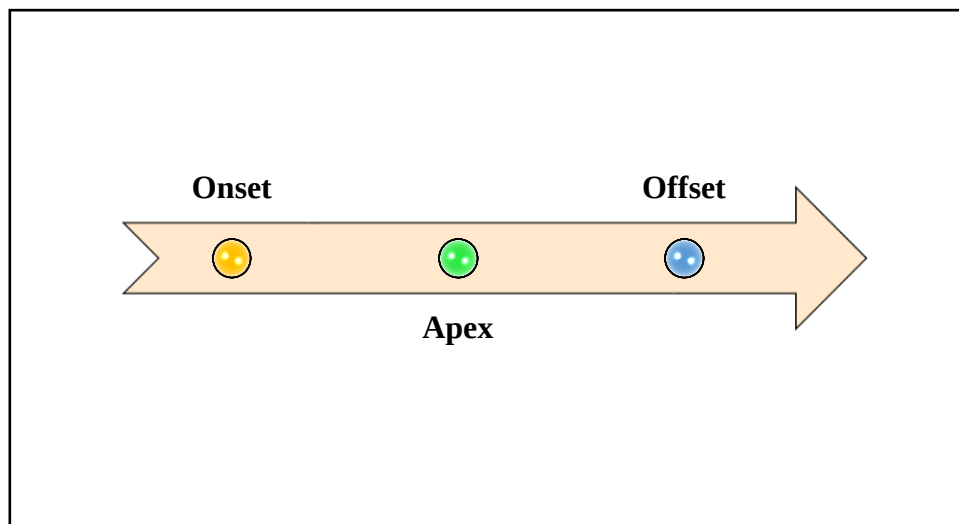
|          |                        | <b>Muscles</b>                 | <b>Description</b>                                                                        | <b>Action</b>                                                                   | <b>Mimique</b>                                                 |
|----------|------------------------|--------------------------------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------|
| <b>1</b> | Muscle épicroânien     | Le frontalis                   | Un muscle qui recouvre le front                                                           | - élève les sourcils<br>- plisse le front                                       | - L'attention<br>- La surprise                                 |
| <b>2</b> | Les muscles orbitaires | Le muscle orbiculaire de l'œil | Un muscle plat, disposé en ellipse à grand axe transversal autour de la fente palpébrale. | - ferme l'œil                                                                   | - Le souci<br>- La préoccupation                               |
|          |                        | Le sourcilier                  | Un petit muscle pair profond situé sur le bord interne du sourcil                         | - rapproche les sourcils<br>- détermine les rides verticales inter sourcilières | - La douleur<br>- La colère<br>- L'attention<br>- L'impatience |

|   |                                       |                                  |                                                                                                                                                                                    |                                                                                |                                         |
|---|---------------------------------------|----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|-----------------------------------------|
| 3 | Les muscles du nez                    | Le muscle transverse du nez      | Un petit muscle aplati et de forme triangulaire, situé au niveau des ailes du nez.                                                                                                 | - dilate les narines<br>- détermine les rides horizontales                     | - mauvaise humeur                       |
|   |                                       | Le muscle élévateur naso-labiale | Un muscle mince et triangulaire, situé dans le sillon nasogénien. Il s'étend du rebord médial de l'orbite à la peau du bord postérieur de l'aile du nez et de la lèvre supérieure. | -élève la lèvre supérieure et l'aile du nez                                    | -Le sourire gingivale<br>- La tristesse |
|   |                                       | Le pyramidal                     | Un petit muscle situé sur la partie supérieure du dos du nez entre les yeux.                                                                                                       | -abaisse l'espace inter-sourcilier                                             | - L'agressivité                         |
| 4 | Les muscles de la bouche et du menton | Orbiculaire de la bouche         | Un sphincter non homogène situé dans l'épaisseur des lèvres, il comprend deux parties (interne et externe) constitués de nombreuses fibres musculaires.                            | -ferme la bouche<br>-projette en avant les lèvres (occlusion des lèvres)       | - La parole et la mastication           |
|   |                                       | Releveur de la lèvre supérieure  | Un muscle pair, mince, aplati et quadrilatère, situé au-dessus de la lèvre supérieure.                                                                                             | - relève la lèvre supérieure                                                   | - La tristesse                          |
|   |                                       | Le petit zygomatique             | Un muscle pair et inconstant de la lèvre supérieure.                                                                                                                               | - tire la lèvre supérieure vers l'arrière, le haut et l'extérieur.             | -Le mépris                              |
|   |                                       | Le grand zygomatique             | Un muscle fin situé dans la partie supérieure de la joue.                                                                                                                          | - dilate la fente orale.                                                       | -La joie                                |
|   |                                       | Buccinateur                      | Un muscle profond de forme rectangulaire qui se situe au niveau de la joue, entre les deux mâchoires.                                                                              | -allonge la bouche transversalement.                                           | - Le souffler                           |
|   |                                       | Risorius                         | Un muscle spécial à l'homme, plat, triangulaire à sommet antérieur, situé au niveau de la joue et s'attache aux commissures des lèvres.                                            | -attire la commissure des lèvres en arrière (Le coin de la bouche sur le côté) | -La joie                                |
|   |                                       | Muscle abaisseur de              | Un muscle pair, triangulaire aplati,                                                                                                                                               | - abaisse l'angle de la bouche                                                 | - La tristesse                          |

|   |               |                      |                                                                                                                         |                                                                           |                                                        |
|---|---------------|----------------------|-------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|--------------------------------------------------------|
|   |               | l'angle de la bouche | situé au-dessous de la commissure des lèvres.                                                                           |                                                                           |                                                        |
|   |               | Muscle mentonnier    | Un muscle pair et triangulaire, situé au centre du menton.                                                              | -élève les parties molles du menton et indirectement la lèvre inférieure. | -Le doute<br>-L'indécision                             |
| 5 | Muscle du cou | Le platysma          | Un muscle plat, quadrangulaire en forme de feuille, assez mince, situé superficiellement sur la face antérieure du cou. | - dépression de la mâchoire                                               | -La douleur<br>-La torture<br>-L'effroi<br>-La terreur |

### I.3.2. Physiologie du visage

Comme mentionné précédemment, le visage est composé de plusieurs muscles qui se déforment par activation volontaire ou non volontaire, pour former un changement perceptible visuellement sur le visage qui est l'expression faciale émotionnelle. L'activation musculaire du visage passe par un processus temporel linéaire divisé en trois phases (Figure I.3) [18].



**Figure I.3: Les trois phases de l'activation musculaire du visage.**

*Phase 01 : s'appelle la phase d'attaque « Onset »*

C'est la période pendant laquelle l'activité du muscle passe de la valeur nulle à sa valeur maximale.

*Phase 02 : s'appelle la phase de soutien « Apex »*

C'est la période pendant laquelle l'activité stagne à son maximum.

*Phase 03 : s'appelle la phase de relâchement « Offset »*

C'est la période pendant laquelle l'activité du muscle baisse jusqu'au niveau initial.

### I.3.3. Aspect universel des émotions ‘Universalité des expressions faciales’

Depuis les années 60, plusieurs chercheurs ont démontré l’universalité de certaines expressions émotionnelles perçues comme fondamentales pour l’individu [19]. Charles Darwin, est le premier à traiter l’universalité des expressions du visage, cet auteur rédige en 1872 un livre intitulé "The Expression of the Emotions in Man and Animals" dans lequel il émet la théorie de l’universalité des émotions chez l’homme et les animaux. Plusieurs chercheurs ont par la suite essayé de prouver cette théorie, dont les plus représentatifs sont Izard, Tomkins, Plutchik, et Ekman. Les émotions de base, sont souvent représentées comme des émotions innées et indépendantes des cultures mais leur nombre reste toujours un débat (varie selon les différentes études). Le (Tableau I.6) présente l’ensemble d’émotions de base selon différents auteurs.

**Tableau I.6: Liste des émotions de base proposées par différents auteurs.**

| Auteur   | Référence | Emotions de base                                                                                            |
|----------|-----------|-------------------------------------------------------------------------------------------------------------|
| Izard    | [20]      | <u>Dix émotions</u> : surprise, joie, colère, dégoût, peur, honte, tristesse, mépris, intérêt, culpabilité. |
| Tomkins  | [21]      | <u>Neuf émotions</u> : surprise, joie, colère, dégoût, peur, mépris, honte, intérêt, anxiété.               |
| Plutchik | [22]      | <u>Huit émotions</u> : surprise, joie, colère, dégoût, peur, tristesse, acceptation, espérance.             |
| Ekman    | [23]      | <u>Six émotions</u> : surprise, joie, colère, dégoût, peur, tristesse.                                      |

Plus récemment, le psychologue américain Paul Ekman, l'un des spécialistes les plus influents dans l'étude des émotions, affirme qu'il y a six émotions de base : la joie, la tristesse, le dégoût, la peur, la colère et la surprise [23]. Chacune de ces émotions est caractérisée par des expressions faciales résultant des divers changements qui apparaissent sur le visage. Nous verrons dans le tableau ci-dessous (Tableau I.7) le lien qui existe entre ces six émotions de base et la déformation des différentes régions du visage[24] .

Tableau I.7: Les caractéristique des expressions faciales des émotions de base (Ekman &amp; Wezowski)

| Zone<br>Emotion  | Yeux                               | Nez                    | Bouche                                                                                              | Front                           | Sourcils                                      | Paupières                                                          | Joues                |
|------------------|------------------------------------|------------------------|-----------------------------------------------------------------------------------------------------|---------------------------------|-----------------------------------------------|--------------------------------------------------------------------|----------------------|
| <b>Surprise</b>  | Yeux élargis                       | /                      | Bouche ouverte                                                                                      | Rides horizontales sur le front | Sourcils levés (parties internes et externes) | /                                                                  | /                    |
| <b>Joie</b>      | Rides montantes aux coins des yeux | /                      | Les coins des lèvres tirés vers le haut (avec possibilité d'ouverture de la bouche)                 | /                               | /                                             | Les rides sous la paupière inférieure                              | Haussement des joues |
| <b>Colère</b>    | /                                  | /                      | Rétrécissement des lèvres                                                                           | /                               | Les sourcils se baissent                      | /                                                                  | /                    |
| <b>Dégoût</b>    | /                                  | Rides au sommet du nez | Lèvre supérieure levée                                                                              | /                               | /                                             | /                                                                  | /                    |
| <b>Peur</b>      | Yeux ouverts et pupilles larges    | /                      | -Lèvres légèrement étirées horizontalement jusqu'aux oreilles<br>-Lèvre inférieure souvent abaissée | /                               | Sourcils levés et rapprochés                  | -Paupières supérieures soulevées<br>-Paupières inférieures tendues | /                    |
| <b>Tristesse</b> | Yeux brillants                     | /                      | Coins de la bouche descendent                                                                       | Rides sur le front              | Abaissement et rapprochement des sourcils     | Paupières supérieures tombantes                                    | /                    |

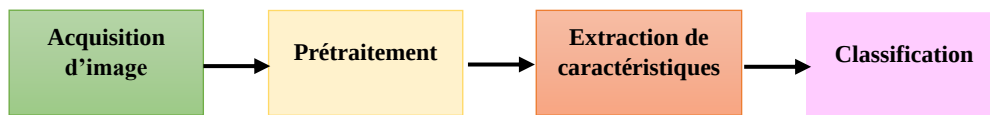
## I.4. Système de reconnaissance automatique des expressions faciales

Les expressions faciales constituent l'un des moyens les plus puissants, naturels et universels pour les êtres humains de communiquer leurs émotions et intentions. En raison de leur importance pratique, de nombreuses recherches ont été effectuées sur l'analyse automatique des expressions faciales. Ces recherches sont cruciales pour des domaines tels que les robots sociaux, les traitements médicaux, la surveillance de la fatigue des conducteurs, et d'autres systèmes d'interaction homme-machine. Dans les domaines de la vision par ordinateur et de l'apprentissage automatique, divers systèmes de reconnaissance des expressions faciales ont été développés pour extraire les informations émotionnelles à partir des représentations faciales[25].

Un système de reconnaissance automatique des expressions faciales analyse et rapporte le comportement facial d'un sujet en classant ses expressions faciales selon l'une des sept expressions universelles des émotions (peur, colère, dégoût, tristesse, joie, surprise et neutre)[26].

### I.4.1. Processus d'un système de reconnaissance des expressions faciales

Les systèmes REF reposent sur des techniques de traitement d'images et d'analyse de caractéristiques. En général, ces systèmes comprennent les étapes suivantes : acquisition d'images, prétraitement, extraction de caractéristiques, classification, comme illustré dans la (Figure I.4)[27].



**Figure I.4: Schéma du système REF conventionnel** [27]

#### I.4.1.1. Acquisition d'images

Les images utilisées pour la reconnaissance des expressions faciales peuvent être des images statiques ou des séquences d'images. Les séquences d'images, qui capturent une série de photos prises à différents moments, offrent l'avantage de saisir les aspects dynamiques des expressions faciales, ce qui enrichit les informations disponibles par rapport à une image fixe. Cela permet de mieux comprendre les variations et les transitions des expressions au fil du temps.

En ce qui concerne les dimensions spatiale, chromatique et temporelle des images, les séquences en niveaux de gris (2D) sont actuellement les plus couramment utilisées pour les systèmes de reconnaissance automatique des expressions faciales. Ces images sont populaires en raison de leur simplicité et de leur efficacité pour traiter des données essentielles à l'analyse des expressions.

Cependant, avec l'évolution technologique et la réduction des coûts des équipements d'acquisition d'images couleur, les images en couleur pourraient devenir de plus en plus courantes dans ce domaine. Les images couleur offrent des avantages supplémentaires, notamment la capacité de capturer des indices émotionnels plus subtils, tels que les changements de teint comme les rougeurs, qui peuvent fournir des informations précieuses sur l'état émotionnel d'une personne. Cela ouvre la voie à des systèmes de reconnaissance des expressions faciales plus précis et plus complets à l'avenir [28].

#### I.4.1.2. Le prétraitement

Le prétraitement est une étape fondamentale, tant dans les systèmes REF que dans tout système reposant sur l'apprentissage automatique. Traiter des données brutes sans une préparation adéquate peut conduire à des résultats imprécis ou non pertinents. Il est donc crucial de s'assurer de la qualité des données avant d'en extraire les caractéristiques nécessaires. Les techniques suivantes sont les techniques de



prétraitement les plus couramment utilisées et les plus efficaces, telles que documentées dans les études examinées [27] [28].

- **Détection de visage :** En général, la détection de visage est la première étape d'un système de reconnaissance des expressions faciales (REF). Elle constitue un aspect essentiel du prétraitement en permettant d'identifier et de localiser les visages dans une image ou une séquence d'images. Cette procédure délimite la région d'intérêt (ROI) en se concentrant sur les zones pertinentes et en excluant les parties non importantes. En isolant les visages, les algorithmes de reconnaissance peuvent ainsi se focaliser sur les caractéristiques faciales spécifiques, ce qui améliore la précision des analyses ultérieures[27].
- **Transformations géométriques :** Les transformations géométriques jouent un rôle clé dans le prétraitement des images pour assurer une uniformité des données. Ces ajustements, comprenant la rotation, le redressement, le redimensionnement et le recadrage, permettent de corriger les variations d'orientation et de taille des visages détectés. Bien que les visages puissent être identifiés dans une image d'entrée, ils ne sont pas toujours dans des conditions idéales pour l'analyse. Des problèmes tels que des déviations de rotation, des erreurs d'échelle ou du bruit peuvent affecter les régions d'intérêt (ROI) détectées. Il est donc essentiel d'aligner et de normaliser les images de manière à ce que les visages à classer soient aussi géométriquement similaires que possible à ceux utilisés pour entraîner le classificateur. Cette normalisation améliore la précision des systèmes de reconnaissance des expressions faciales en assurant une cohérence dans les données traitées[27].
- **Traitement d'image :** Le traitement d'image est une étape clé du prétraitement, visant à optimiser la qualité des images avant leur analyse. Cette phase comprend des opérations telles que la réduction du bruit, la correction de l'éclairage, l'ajustement du contraste, et l'accentuation des détails essentiels. Même si une région d'intérêt (ROI) est correctement ajustée géométriquement, cela peut ne pas suffire pour préparer les données de manière optimale. Le traitement d'image permet de corriger les imperfections et de mettre en évidence les caractéristiques pertinentes, rendant ainsi les images plus adaptées aux algorithmes de reconnaissance. Ces améliorations assurent des images nettes et cohérentes, facilitant une analyse plus précise des expressions faciales et améliorant la performance des systèmes de reconnaissance[27].

Le (Tableau I.8) présente les techniques couramment utilisées pour chaque étape de prétraitement [27].

**Tableau I.8: Les techniques couramment utilisées pour le prétraitement des données.**

| Prétraitement       | Techniques couramment utilisées     | Description                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Détection de visage | Détecteur de visage Viola–Jones[29] | Le détecteur de visage Viola–Jones est une méthode d'apprentissage automatique qui utilise une fonction en cascade entraînée à partir de nombreuses images avec visages (positives) et sans visages (négatives). Cette approche repose sur des caractéristiques de type Haar appliquées à l'ensemble des images d'entraînement pour déterminer le seuil optimal permettant de classer les visages. |
|                     | Détecteur de visage de Dlib [30]    | Le détecteur de visage de Dlib emploie un ensemble d'arbres de régression pour localiser 68 points de repère faciaux à partir d'un échantillon réduit de valeurs d'intensité issues d'une image                                                                                                                                                                                                    |

|                              |                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              |                                                                                                                                                                                                                                                                                                                                                                                                | d'entrée, ce qui permet de déterminer la présence des visages.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|                              | Multi-task Cascade Convolution Neural Network (MTCNN) [31]                                                                                                                                                                                                                                                                                                                                     | Le MTCNN fonctionne en trois étapes distinctes pour assurer une détection précise des visages. À chaque étape, le visage est analysé à l'aide d'un réseau de neurones convolutifs (CNN). La première étape détecte les fenêtres candidates et calcule les vecteurs de régression des boîtes englobantes, tout en fusionnant les candidats qui se chevauchent fortement. La deuxième étape utilise un CNN supplémentaire pour éliminer la majorité des faux positifs. Enfin, la troisième étape, qui est similaire à la deuxième, fournit également les positions de cinq points de repère faciaux.                 |
|                              | Bien que le détecteur de visage Viola–Jones soit le plus couramment utilisé dans les études examinées, le MTCNN offre de meilleures performances dans des conditions difficiles, tout en conservant une vitesse en temps réel. Viola–Jones est privilégié dans les environnements contrôlés où il fonctionne efficacement, tandis que MTCNN est préféré pour les environnements non contrôlés. |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Transformations géométriques | Correction de la rotation                                                                                                                                                                                                                                                                                                                                                                      | Pour corriger la rotation, la méthode la plus courante consiste à utiliser les points de repère faciaux générés par le détecteur de visage. En général, les études prennent en compte deux points de repère faciaux qui, lorsqu'un visage est aligné, forment un angle nul avec l'axe horizontal. Pour aligner un visage tourné, une transformation de rotation est appliquée afin que ces deux points de repère soient alignés avec l'axe horizontal, jusqu'à ce que l'angle entre eux soit nul.                                                                                                                  |
|                              | Suppression de l'arrière-plan                                                                                                                                                                                                                                                                                                                                                                  | Le bruit dans les régions d'intérêt (ROIs) détectées est principalement dû à l'arrière-plan. Il est essentiel d'éliminer cet arrière-plan des ROIs originales, car il peut diminuer la précision du classificateur en ajoutant une variable supplémentaire : la distinction entre avant-plan et arrière-plan. Bien que ce type de bruit soit souvent négligé, certaines études ont tenté de recadrer les ROIs pour mieux filtrer l'arrière-plan. La méthode la plus courante consiste à utiliser les points de repère faciaux et la boîte englobante fournis par le détecteur de visage. En mesurant les distances |

|                    |                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--------------------|---------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                    |                           | entre les points de repère faciaux pertinents, on peut réduire la taille de la ROI fournie par le détecteur et ainsi éliminer le bruit de l'arrière-plan.                                                                                                                                                                                                                                                                                                |
| Traitement d'image | Lissage des images        | Le lissage des images est crucial pour réduire le bruit et améliorer la fiabilité des données. Les filtres bilatéraux, qui conservent les contours nets, et les filtres gaussiens, qui atténuent le bruit gaussien, sont couramment utilisés. Pour la reconnaissance des expressions faciales, le filtre bilatéral est préféré afin de préserver les traits distinctifs du visage.                                                                       |
|                    | Égalisation d'histogramme | L'égalisation d'histogramme ajuste la répartition des intensités dans une image pour uniformiser l'éclairage, améliorant ainsi le contraste et mettant en valeur les traits faciaux. Bien que cette technique réduise l'impact des variations d'éclairage, elle peut aussi accentuer le bruit de fond.                                                                                                                                                   |
|                    | Augmentation des Données  | Les petites bases de données pour la reconnaissance des émotions peuvent entraîner un surapprentissage, où le modèle fonctionne bien sur les données d'entraînement mais échoue sur de nouvelles données. L'augmentation des données (DA) résout ce problème en augmentant la quantité de données via des modifications telles que le recadrage, la rotation et le changement de luminosité, tout en veillant à ne pas altérer le sens des échantillons. |

#### I.4.1.3. Extraction des caractéristiques faciales

L'extraction des caractéristiques faciales consiste à extraire des traits distinctifs des images de visages afin de représenter efficacement les expressions faciales. Selon le type d'images en entrée, les méthodes d'extraction peuvent être classées en deux catégories : celles appliquées aux images statiques et celles utilisées pour les séquences d'images dynamiques[32].

##### I.4.1.3.1. Méthodes d'extraction des caractéristiques faciales pour les images statiques

Pour les images statiques, il existe deux types de méthodes d'extraction des caractéristiques faciales : les méthodes basées sur les caractéristiques géométriques et les méthodes basées sur l'apparence [32].

###### I.4.1.3.1.1. Méthodes basées sur les caractéristiques géométriques

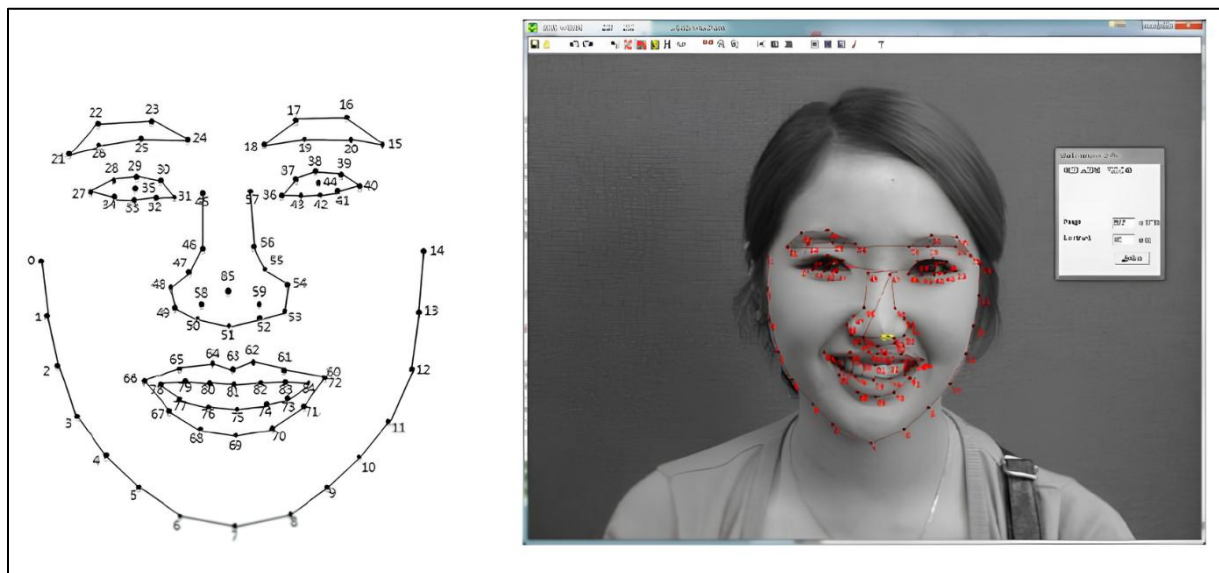
L'extraction des caractéristiques faciales cherche à obtenir la meilleure représentation possible de l'image du visage pour la reconnaissance. Dans les systèmes basés sur les caractéristiques géométriques, les points faciaux d'une image sont utilisés de différentes manières pour créer un vecteur de

caractéristiques permettant de reconnaître les expressions faciales. Par exemple, les distances entre les points caractéristiques et les tailles relatives des principales composantes du visage sont calculées pour former ce vecteur. Ces points peuvent aussi être utilisés pour créer une représentation géométrique du visage. L'approche géométrique présente des avantages et des inconvénients. Une des principales difficultés est l'initialisation et le suivi précis et en temps réel des points caractéristiques du visage. Toute erreur dans ce processus s'accumule lors de l'extraction des caractéristiques géométriques. De plus, la résolution de l'image, l'orientation de la tête, les lunettes, la présence de barbe, etc., peuvent affecter ce processus. Cependant, une fois que les points sont correctement initialisés et suivis, les caractéristiques géométriques obtenues sont robustes face aux variations d'échelle, de taille, d'orientation de la tête et de texture du visage dues à l'âge, entre autres[33].

Il existe trois principales méthodes d'extraction basées sur les caractéristiques géométriques : les modèles de formes actifs (ASM), les modèles d'apparence actifs (AAM) et la transformation des caractéristiques invariantes à l'échelle (SIFT)[32].

#### - Modèle de Formes Actives (ASM)

Le modèle de formes actives (ASM), introduit par Cootes et ses collègues [34], est une technique d'appariement des caractéristiques basée sur un modèle statistique. Une ASM se compose d'un modèle de distribution de points (PDM) qui apprend les variations de formes possibles, ainsi que de plusieurs modèles flexibles qui capturent les niveaux de gris autour de divers points de repère caractéristiques. La méthode ASM comporte deux étapes principales : d'abord, les modèles de forme sont construits à partir d'échantillons d'entraînement avec des points de repère annotés ; ensuite, des modèles de texture locale sont créés pour chaque point de repère. Une procédure de recherche itérative adapte ensuite le modèle aux nouvelles images en déformant l'exemple initial. Un schéma de repérage des caractéristiques faciales, utilisé dans la mise en œuvre classique de l'ASM, est présenté dans la (Figure I.5) [76]. Anderson et al. [35], ont proposé une version améliorée, appelée modèles de formes actives et statistiques (ASSM), pour la reconnaissance faciale, avec des applications potentielles pour la REF.



**Figure I.5: Schéma de repérage des caractéristiques faciales utilisé dans la mise en œuvre classique de l'ASM [36]**

Divers travaux de recherche ont exploré l'utilisation des modèles de formes actives (ASM) pour la reconnaissance des expressions faciales.

Shbib et Zhou [37], ont développé un système de reconnaissance automatique des expressions faciales en utilisant le modèle de formes actives (ASM) couplé à un classificateur SVM. Pour détecter et suivre

les visages, ils ont appliqué le classificateur Adaboost et les caractéristiques Haar-Like, permettant une initiation automatique de l'ASM dans les visages détectés. En utilisant la technique de mise en forme ASM, ils ont extrait des points de repère faciaux précis et discriminants. L'analyse géométrique des expressions faciales s'est basée sur le déplacement de ces points de repère par rapport à la forme moyenne de l'ASM. Avec le classificateur SVM, ils ont obtenu un taux de reconnaissance de 93 %. Des méthodes avancées telles que le descripteur Viola-Jones et l'Adaboost ont également été utilisées pour optimiser les performances du système en temps réel.

Kim [38], a proposé un système de reconnaissance des expressions faciales utilisant les points de repère du modèle de formes actives (ASM) et un algorithme de classification basé sur l'apparence, le modèle de Markov caché intégré (EHMM). Il a utilisé les informations ASM pour la normalisation des images faciales et pour les facteurs de pondération des probabilités, calculés via la divergence de Kullback-Leibler. Ensuite, l'algorithme EHMM, avec le vecteur de caractéristiques 2D-DCT, a été employé pour classifier les états émotionnels. Les tests sur les bases de données CK et JAFFE ont montré des améliorations de performance de 6,5 % et 2,5 %, respectivement, par rapport aux méthodes précédentes basées sur l'alignement facial ASM.

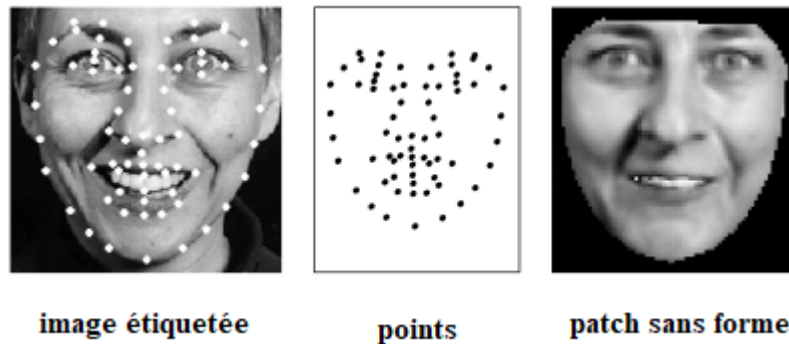
Si et al. [39], ont proposé un nouvel algorithme de reconnaissance des expressions faciales basé sur une machine à automates finis pour résoudre le problème de l'impact des erreurs de détection des caractéristiques. Les étapes de mise en œuvre incluent le prétraitement des images, la détection des visages basée sur Adaboost, la localisation des caractéristiques faciales avec ASM, la normalisation et la classification des expressions faciales via une machine à automates finis. Les résultats expérimentaux montrent que ce nouveau système de reconnaissance faciale possède une grande précision, une stabilité élevée et une vitesse d'exécution rapide. Cela représente une nouvelle méthode d'amélioration de la reconnaissance des expressions faciales, favorisant le développement de l'intelligence artificielle et de l'interaction homme-machine.

Suchitra et al. [40], ont proposé une méthode de reconnaissance des émotions en temps réel à partir d'images faciales. Cette méthode se décompose en trois étapes : la détection du visage à l'aide de la cascade de Haar, l'extraction des caractéristiques en utilisant le modèle de formes actives (ASM) pour extraire 26 points faciaux, et la classification des émotions (colère, dégoût, bonheur, neutre et surprise) via le classificateur Adaboost. La nouveauté de cette méthode réside dans son implémentation en temps réel sur un Raspberry Pi II, atteignant une précision moyenne de 94 %. Monté sur un robot mobile, le Raspberry Pi II peut reconnaître les émotions dynamiquement en temps réel dans des environnements sociaux ou de service où la reconnaissance des émotions joue un rôle crucial. Les résultats montrent une précision globale de 94 % avec un temps de traitement moyen de 120 ms sur une plateforme Linux utilisant le Raspberry Pi II (ARM1176JZF, 900 MHz). Ce système est utile dans diverses applications sociales telles que les hôpitaux ou les maisons de retraite.

#### - **Modèle d'Apparence Active (AAM)**

Le modèle d'apparence active (AAM) est une technique de vision par ordinateur qui combine les informations de forme et de texture d'une image faciale pour une reconnaissance précise des traits (Figure I.6). Développé par Cootes et al. en 2001 [41], l'AAM améliore le modèle de formes actives (ASM) en intégrant ces deux types d'informations. L'AAM construit un modèle statistique à partir de données d'entraînement, puis utilise ce modèle pour ajuster les données de test. Contrairement à l'ASM, l'AAM analyse également les textures locales pour comprendre les relations entre forme et texture.

Des versions avancées de l'AAM, telles que l'AAM basé sur l'histogramme des gradients orientés (HOG) [42], l'AAM dense [43], et l'AAM basé sur la régression [44], ont été développées récemment, offrant des performances améliorées pour la reconnaissance des expressions faciales.



**Figure I.6: Modèle d'Apparence Active (une image d'entraînement étiquetée fournit un patch sans forme et un ensemble de points) [41].**

À l'instar de l'ASM, de nombreuses recherches ont étudié l'application des modèles AAM pour la reconnaissance des expressions faciales.

Zheng et Liu [45], ont proposé une méthode de reconnaissance des expressions faciales utilisant le AAM pour extraire les paramètres de forme et de texture des points caractéristiques du visage. Cette technique permet de capturer les informations spécifiques à chaque expression et de les classer à l'aide d'un algorithme d'apprentissage automatique. Ils ont constaté que retenir trop d'informations introduit du bruit, ce qui affecte la précision de la classification. En ajustant la dimension du vecteur de caractéristiques, ils parviennent à maintenir un taux d'identification élevé tout en réduisant la dimensionnalité. Les résultats expérimentaux démontrent que cette approche améliore la précision et la robustesse par rapport aux méthodes précédentes, tout en conservant un bon taux d'identification malgré la réduction de la dimension des caractéristiques.

Wang et al. [46], ont proposé une méthode en utilisant l'AAM pour extraire les régions faciales selon le Facial Action Coding System (FACS). Cette méthode se compose de trois étapes : l'extraction des régions faciales avec AAM, l'extraction des caractéristiques par transformation par ondelettes de Gabor, et la reconnaissance des expressions via des Machines à Vecteurs de Support (SVM). L'approche se distingue par sa capacité à minimiser les effets des variations de taille des régions faciales, des poses de la tête et des conditions d'éclairage, ce qui améliore la précision de la reconnaissance. Les expérimentations sur les bases de données JAFFE et Yale ont montré des taux de reconnaissance atteignant 100%, en particulier pour les expressions de bonheur et de surprise, démontrant l'efficacité et la robustesse de cette méthode pour l'interaction homme-machine et d'autres applications en intelligence artificielle.

Ko et Sim [47], ont proposé une méthode en combinant l'AAM et le Système FACS avec un Réseau Bayésien Dynamique (DBN) et un Filtre de Kalman. Cette approche permet de modéliser les phases temporelles des expressions faciales et d'améliorer la précision de reconnaissance, atteignant plus de 90 % sur le jeu de données BioID. En utilisant AAM, les auteurs facilitent l'extraction des caractéristiques émotionnelles variées, représentées par des mesures géométriques, d'apparence et de surface. Le DBN, quant à lui, assure la modélisation des incertitudes et des dynamiques du comportement facial, augmentant ainsi la précision et la robustesse de la méthode.

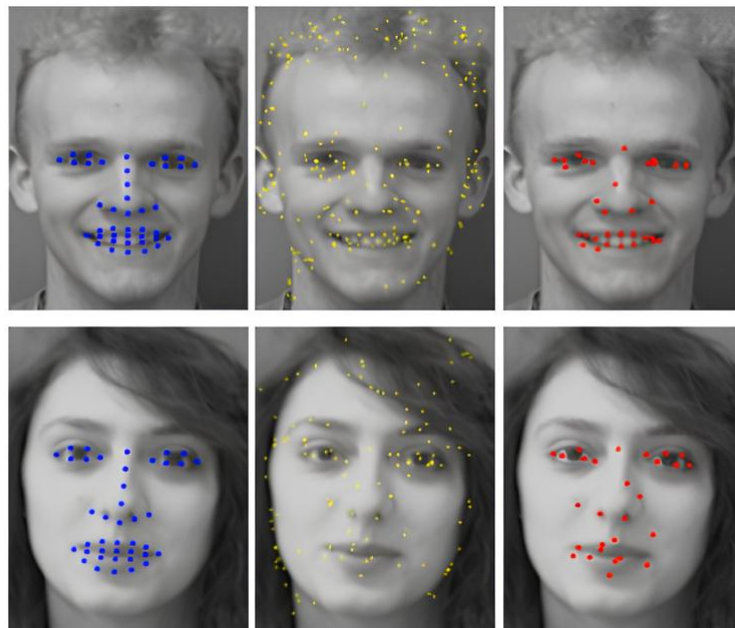
Les auteurs ont présenté une méthode automatique pour la reconnaissance des expressions faciales à partir d'images de visage presque de face. Ils ont proposé deux contributions majeures : d'une part, une méthode adaptative pour générer le modèle initial pour l'ajustement des Modèles d'Apparence Active (AAM), permettant une détection plus précise des traits du visage malgré de grandes variations ; d'autre part, l'introduction d'un ensemble de caractéristiques hybrides, comprenant des caractéristiques géométriques, des paramètres de forme et d'apparence d'AAM. En optimisant ces caractéristiques à l'aide de la méthode de sélection basée sur l'information mutuelle quadratique (QMI), ils ont obtenu des

résultats prometteurs. Les tests effectués sur la base de données CAS-PEAL montrent que, grâce à un classificateur SVM, la méthode atteint un taux de reconnaissance de 87,33 %.

Zhang et al. [48], ont proposé une approche automatique pour la reconnaissance des expressions faciales à partir d'images de visage. Ils ont apporté deux contributions principales : premièrement, une méthode pour générer de manière adaptative le modèle initial pour l'ajustement des Modèles d'Apparence Active (AAM), permettant une détection plus précise des caractéristiques faciales sous de grandes variations ; deuxièmement, l'introduction d'un ensemble de caractéristiques hybrides pour l'expression, comprenant des caractéristiques géométriques, des paramètres de forme et d'apparence d'AAM. En utilisant la méthode de sélection de caractéristiques basée sur l'information mutuelle quadratique (QMI), ces caractéristiques peuvent être optimisées. Les résultats expérimentaux sur la base de données d'expressions faciales CAS-PEAL montrent que la méthode proposée atteint un taux de reconnaissance de 87,33 % avec un classificateur de machine à vecteurs de support (SVM).

#### - Scale Invariant Feature Transform (SIFT)

Le SIFT (Scale-invariant feature transform), proposé par David Lowe [49], est un descripteur d'image local destiné à la correspondance d'images. Il offre une invariance aux changements d'échelle, de translation et de rotation, et une résistance partielle aux variations d'illumination et aux projections affines ou tridimensionnelles (Figure I.7). Il existe plusieurs variantes du SIFT qui améliorent ses performances dans divers contextes. Soyel et Demirel [50], ont développé une méthode de discrimination appelée D-SIFT, qui renforce l'efficacité de la prise de décision en se basant sur les caractéristiques d'apparence globale. Li et al. [51], ont présenté le GA-SIFT, une version du SIFT adaptée aux images multispectrales. Cette approche utilise l'algèbre géométrique pour intégrer les informations spectrales et spatiales, créer l'espace d'échelle et identifier les points caractéristiques.



**Figure I.7: Scale Invariant Feature Transform (les positions initiales de la caractéristique faciale, la caractéristique SIFT et les positions du SIFT amélioré) [52].**

Parmi les travaux de littérature dans le domaine de la reconnaissance des expressions faciales utilisant la méthode SIFT (Scale-Invariant Feature Transform), on trouve :

Berretti et al. [53], ont proposé une méthode pour reconnaître les expressions faciales en se basant sur la géométrie 3D du visage. Cette approche automatique identifie des points clés du visage, extrait des



descripteurs de caractéristiques SIFT à partir des images de profondeur, et sélectionne les caractéristiques les plus pertinentes. En entraînant des machines à vecteurs de support (SVM) pour chaque expression faciale, ils ont atteint un taux de reconnaissance de 78,43 % sur la base de données BU-3DFE.

Dahmane et Meunier [54], ont développé un système automatisé pour analyser les expressions faciales, visant à améliorer l'interaction homme-machine, l'analyse des émotions et les soins médicaux. Le principal défi de ces systèmes réside dans l'alignement des visages en raison des variations spécifiques à chaque individu. Pour pallier ces problèmes, les auteurs ont introduit un modèle basé sur des prototypes à l'aide de la méthode SIFT-flow. Ils ont créé un ensemble de modèles d'expressions faciales prototypes servant de référence pour les émotions, sur lesquels les images des visages sont projetées afin de générer des visages alignés. Des gradients orientés sont ensuite extraits de chaque image alignée pour caractériser les expressions faciales. Ce système a démontré une efficacité avec des taux de reconnaissance de 87 % sur la base de données JAFFE et de 83 % sur la base de données GEMEP-FERA. L'alignement obtenu grâce à cette méthode a montré une performance supérieure aux méthodes existantes, avec un taux de généralisation de 99,52 % en utilisant la stratégie leave-one-out et de 98,73 % avec la validation croisée.

Zhang et al. [55], ont développé une méthode utilisant des réseaux de neurones profonds (DNN) pour améliorer la reconnaissance des expressions faciales sous divers angles. Cette méthode commence par l'extraction des caractéristiques SIFT à des points de repère spécifiques des images faciales, puis ces caractéristiques sont traitées par un modèle DNN pour apprendre des caractéristiques discriminantes optimales pour la classification des expressions. Des tests sur les bases de données BU-3DFE et Multi-PIE montrent que cette approche est efficace et extraient efficacement des informations spatiales discriminantes.

Kalsum et al. [56], ont développé une méthode hybride pour la reconnaissance des émotions faciales, en combinant des descripteurs SBoFs avec SIFT et SURF spatiaux (SBoF-SSIFT et SBoF-SSURF). Cette méthode, robuste aux variations d'échelle, de rotation et d'illumination, utilise des algorithmes de classification comme les K-plus proches voisins et les SVM. Testée sur les bases de données CK+ et JAFFE, elle a atteint des précisions de 98,5% et 98,3% respectivement. En intégrant les informations spatiales et en éliminant les zones statiques du visage, la méthode améliore la reconnaissance des expressions faciales, démontrant ainsi son efficacité et sa robustesse.

#### **I.4.1.3.1.2. Méthodes basées sur l'apparence**

Les méthodes basées sur l'apparence pour l'extraction de caractéristiques en reconnaissance des expressions faciales se concentrent sur l'analyse visuelle des images du visage pour identifier des éléments distinctifs liés aux émotions. Elles exploitent l'ensemble du visage ou des régions spécifiques pour détecter des informations subtiles telles que les rides et autres modifications faciales. Ces techniques capturent les variations de texture et de forme, et les caractéristiques extraites sont ensuite utilisées dans des modèles d'apprentissage automatique pour classifier les expressions faciales. Ces approches sont efficaces pour gérer les variations interpersonnelles et contextuelles des expressions faciales[32]. Les principales méthodes d'extraction de caractéristiques basées sur l'apparence sont les motifs binaires locaux (LBP) et Histogramme des Gradients Orientés - Histogram of Oriented Gradient- (HOG).

##### **- Motifs binaires locaux (LBP)**

Le LBP (Local Binary Patterns) [57], est un descripteur local simple, efficace et robuste, introduit pour la première fois par Ojala et al. En 1996 comme un opérateur binaire de base. En 2004, il a été utilisé pour l'analyse des visages [58]. L'opérateur LBP examine un voisinage local autour de chaque pixel, compare les valeurs des pixels voisins à celle du pixel central, et utilise l'image binaire résultante comme descripteur local de l'image[59]. Le descripteur LBP pour chaque pixel d'une image est donné par :



$$LBP(P, R) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (I.1)$$

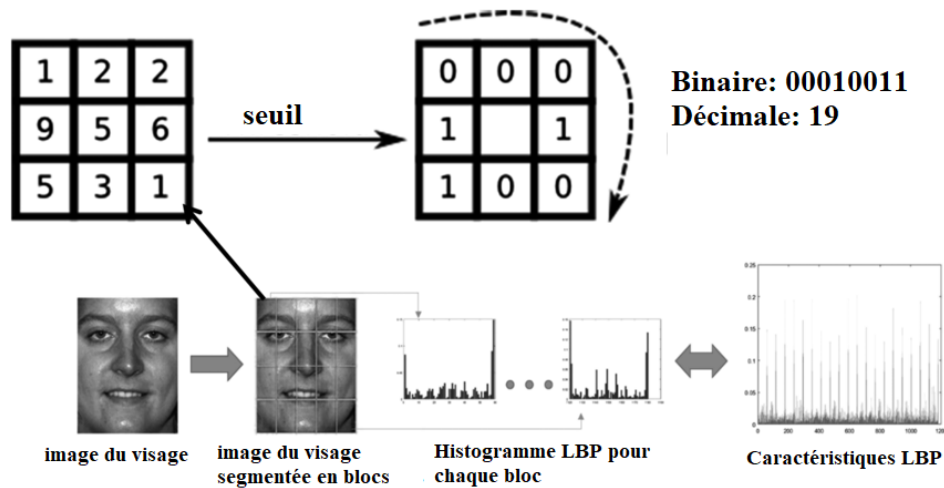
Où :

$s(x)$  est la fonction signe définie par:

$$s(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (I.2)$$

$g_c$  et  $g_p$  sont respectivement les niveaux de gris d'un pixel central et de ses pixels voisins.

Une image originale peut être transformée en une image LBP en assignant à chacun de ses pixels une valeur parmi les  $2^p = 2^8 = 256$  valeurs possibles. Ensuite, l'image résultante sera transformée en un vecteur de caractéristiques en calculant un histogramme indiquant la fréquence d'occurrence correspondant à chaque valeur LBP. Le vecteur final est donc de longueur 256 (Figure I.8).



**Figure I.8: Méthode d'extraction de caractéristiques de modèle binaire local (LBP) [60]**

Dans le domaine de reconnaissance des expressions faciales, He et al. [61], ont proposé une version améliorée de la méthode LBP classique, en se concentrant sur trois aspects principaux : les données d'image, l'extraction des caractéristiques, et la combinaison de ces caractéristiques. Ils ont d'abord utilisé une décomposition par ondelettes pour extraire quatre types de fréquences à partir des images, afin d'exploiter pleinement les données disponibles. Ensuite, ils ont amélioré la robustesse des caractéristiques LBP en les extrayant de manière à la fois locale et globale. Enfin, un mécanisme de pondération adaptatif a été mis en place pour combiner ces caractéristiques de manière plus efficace. Les résultats expérimentaux ont montré que ces améliorations augmentent considérablement la performance de la reconnaissance des expressions faciales par rapport à la méthode LBP traditionnelle.

Shan et Gritti [62], ont proposé une approche en affinant l'utilisation des descripteurs LBP. Au lieu d'utiliser les histogrammes LBP complets extraits des régions faciales locales, ils se concentrent sur l'apprentissage des bins discriminants de l'histogramme LBP (LBPH). Les résultats expérimentaux montrent que ces bins sélectionnés offrent une représentation faciale à la fois compacte et discriminante. Ils ont également démontré l'importance de considérer des échelles multiples de LBP et ont constaté que les motifs uniformes contiennent les informations les plus discriminantes. En utilisant un SVM avec les bins LBPH multiscales sélectionnés, ils ont obtenu une performance de reconnaissance de 93,1% sur la base de données Cohn-Kanade.

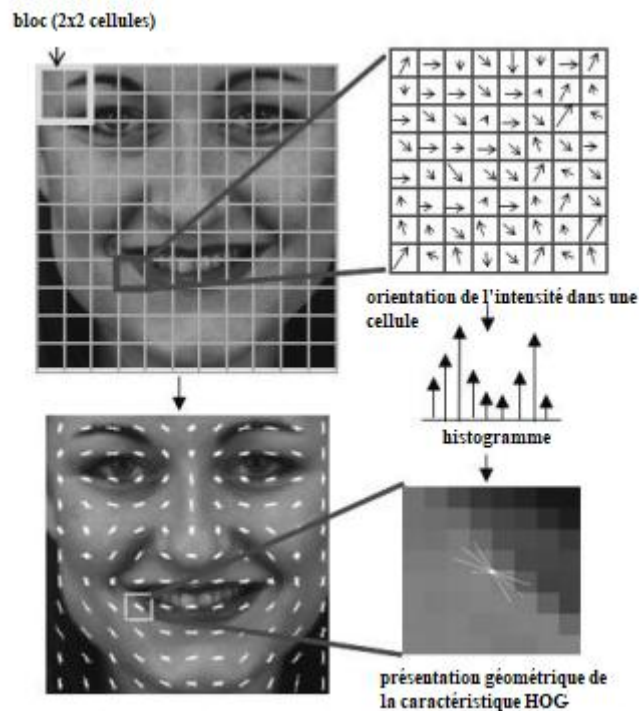
Luo et al. [63], ont proposé une méthode hybride combinant l'analyse en composantes principales (PCA) et les motifs binaires locaux (LBP) pour améliorer la reconnaissance des expressions faciales. Premièrement, ils ont introduit une méthode de segmentation en huit yeux pour extraire les zones pertinentes de l'image faciale, réduisant ainsi les informations inutiles pour l'extraction des caractéristiques. Ensuite, la PCA a été utilisée pour extraire les caractéristiques globales en niveaux de gris de l'image entière, tout en réduisant la taille des données, tandis que le LBP a extrait les caractéristiques locales de texture de la région de la bouche, qui est cruciale pour la reconnaissance des expressions faciales. La fusion de ces caractéristiques globales et locales a permis d'améliorer l'efficacité de la reconnaissance des expressions faciales. Finalement, une machine à vecteurs de support (SVM) a été utilisée pour classifier et reconnaître les expressions à partir des caractéristiques fusionnées. Les résultats expérimentaux montrent que cette méthode peut classifier les différentes expressions de manière plus efficace et atteindre un taux de reconnaissance plus élevé que les méthodes traditionnelles.

Xu et Zhao [64], ont souligné que la rotation des images d'expression faciale constitue un problème difficile à résoudre, entraînant une diminution de la précision de la reconnaissance et ne répondant pas aux exigences réelles. Bien que les réseaux de neurones convolutifs (CNN) puissent extraire des caractéristiques de haute dimension et soient invariants aux variations de niveaux de gris, ils ne présentent pas d'invariance à la rotation. En revanche, le modèle binaire local (LBP) est un algorithme d'extraction de caractéristiques offrant une certaine invariance à la rotation. Pour relever ces défis, les auteurs proposent un algorithme combinant CNN et LBP pour la reconnaissance des expressions faciales et le comparent à d'autres méthodes existantes. Les résultats des simulations montrent que cet algorithme peut améliorer le taux de reconnaissance des expressions malgré la rotation. Les images d'expression faciale utilisées jusqu'à présent sont généralement obtenues dans des conditions contrôlées, avec un fond uniforme, une lumière constante et sans occlusion du visage. Toutefois, dans des situations réelles, la reconnaissance des expressions faciales est souvent perturbée par des arrière-plans complexes, une illumination inégale et des occultations. Améliorer l'algorithme pour une application efficace dans des scènes complexes nécessite encore des recherches supplémentaires.

#### - **Histogramme des Gradients Orientés (HOG)**

Histogramme des Gradients Orientés (HOG) [65], a été développé par Dalal et Triggs pour caractériser l'apparence et la forme locale des objets en se basant sur la distribution des gradients d'intensité locaux ou des directions des contours, indépendamment des positions correspondantes des gradients ou des contours. Son principe de mise en œuvre (Figure I.9), consiste à diviser une image en un certain nombre de petits blocs. Dans chaque bloc, un histogramme local des directions des gradients ou des orientations des contours est calculé. En conséquence, un descripteur HOG est représenté par la combinaison des entrées des histogrammes [65].

Prenons un exemple pratique où le descripteur HOG est utilisé pour extraire des caractéristiques d'une image faciale. Cette dernière est divisée en blocs de  $6 \times 6$ , chaque bloc étant composé de  $2 \times 2$  cellules. L'orientation du gradient est définie entre  $[0-180]$  et est quantifiée en utilisant des histogrammes à 4 bins. Ainsi, le nombre total de caractéristiques HOG pour l'image d'entrée est de  $5 \times 5 \times 2 \times 2 \times 4 = 400$  caractéristiques.



**Figure I.9: Histogramme de l'extraction du gradient orienté à partir du visage [66].**

Greche et ES-Sbai [66], ont développé un système REF automatique en utilisant principalement l'Histogramme des Gradients Orientés (HOG) et la Corrélation Croisée Normalisée. Ils ont évalué les performances du système en variant la résolution des images et l'espace colorimétrique, trouvant que des résultats satisfaisants peuvent être obtenus même avec une faible résolution de 64 par 64 pixels en RGB. Pour améliorer la précision et la robustesse du système, les auteurs recommandent d'élargir les données d'entraînement pour inclure une plus grande diversité de modèles faciaux, et de rechercher une distribution de pics pour estimer la position des expressions plutôt que de se fier à un seul pic, afin de réduire les erreurs de classification.

Chen et al.[67], ont proposé une méthode efficace pour la reconnaissance des expressions faciales en se concentrant sur la détection et l'extraction des composants faciaux tels que les yeux, les sourcils et la bouche, plutôt que sur l'ensemble du visage. Étant donné que les expressions faciales sont causées par les mouvements ou déformations des muscles faciaux, ils ont utilisé l'Histogramme des Gradients Orientés (HOG), sensible aux déformations, pour encoder ces composants en tant que caractéristiques. Un SVM linéaire a ensuite été entraîné pour classer les expressions faciales. Les résultats expérimentaux sur les bases de données JAFFE et Cohn-Kanade étendue montrent des taux de classification de 94,3 % et 88,7 %, respectivement, démontrant ainsi l'efficacité de la méthode.

Orrite et al. [68], abordent le problème de l'identification des émotions humaines à partir de photos fixes prises dans des environnements semi-contrôlés. Ils utilisent les descripteurs Histogramme des Gradients Orientés (HOG) pour décrire l'apparence locale et la forme du visage. Leur méthode propose une formulation bayésienne pour calculer la distribution des contours spécifique à chaque classe et les cartes de vraisemblance sur l'ensemble de données d'entraînement aligné. Un arbre de décision hiérarchique est ensuite construit, utilisant une stratégie ascendante pour regrouper et fusionner les classes à chaque niveau. Pour chaque branche de l'arbre, une liste de caractéristiques HOG potentiellement discriminantes est générée en utilisant les cartes de vraisemblance pour privilégier les emplacements les plus discriminants. Un SVM est utilisé pour le processus décisionnel dans chaque branche. Les résultats expérimentaux sur la base de données Cohn-Kanade AU-Coded montrent que la méthode est efficace pour reconnaître différents états émotionnels à partir de photos uniques de personnes non présentes dans l'ensemble d'entraînement.

Wang et al. [69], proposent une méthode hybride en combinant deux techniques de description de caractéristiques locales : le Weber Local Descriptor (WLD) et le HOG. Le WLD est un descripteur d'image local simple et robuste, tandis que les HOG capturent les informations de densité de gradient et d'orientation des contours. En combinant ces deux approches, la méthode proposée vise à extraire plus efficacement les caractéristiques locales du visage. Les images sont divisées en blocs, chaque bloc est pondéré, puis les caractéristiques WLD et HOG sont extraites et fusionnées. Les histogrammes fusionnés pondérés sont ensuite utilisés pour classifier les expressions faciales en utilisant la distance chi-carré et la méthode des plus proches voisins. Les tests sur les bases de données JAFFE et Cohn-Kanade montrent des taux de reconnaissance de 93,97 % et 95,86 %, respectivement.

#### I.4.1.3.2. Méthodes d'extraction des caractéristiques faciales pour les séquences d'images dynamiques

Les séquences d'images dynamiques capturent le processus continu des mouvements et des expressions du visage. Les caractéristiques faciales dans ces séquences sont principalement représentées par la déformation et les mouvements des muscles du visage. Un modèle temporel de l'expression « sourire » est présenté dans la (Figure I.10) [110]. Bien que ces méthodes soient précises, elles doivent relever des défis comme les variations d'éclairage et d'angles de vue, mais elles restent particulièrement efficaces pour des applications en temps réel. Deux méthodes populaires d'extraction de caractéristiques pour les séquences d'images dynamiques sont les suivantes : le flux optique et le suivi des points caractéristiques [32].

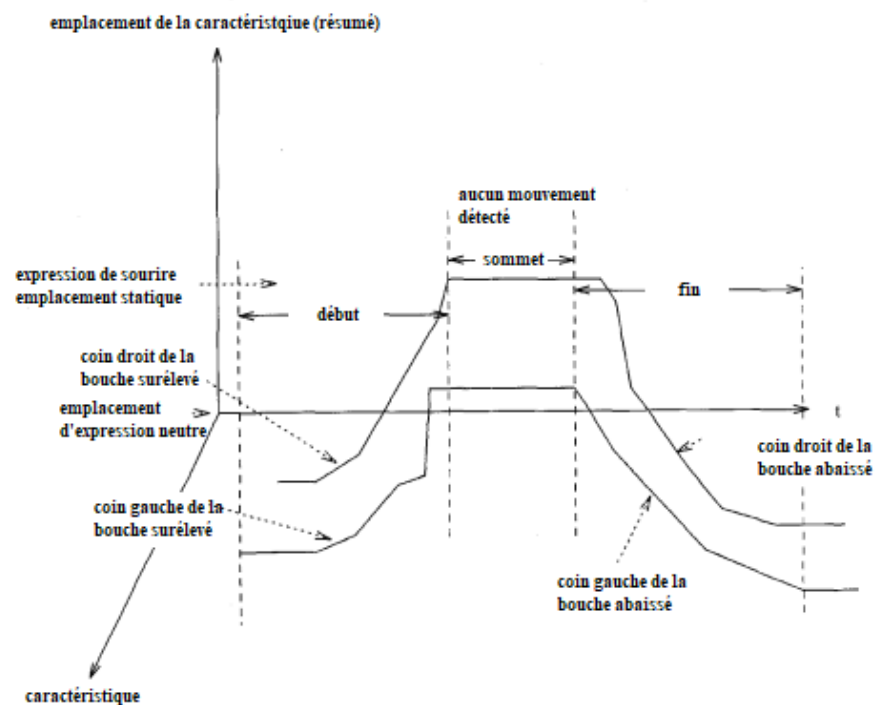
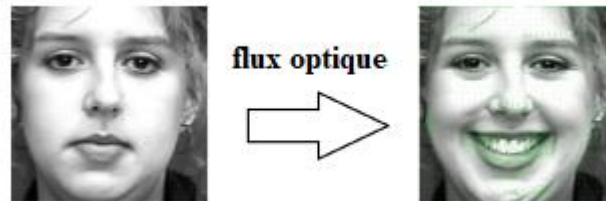


Figure I.10: Le modèle temporel de l'expression « sourire » [70].

##### I.4.1.3.2.1. Flux optique

La méthode du flux optique est une technique employée pour étudier les mouvements dans une séquence d'images, en attribuant un vecteur de vitesse à chaque pixel, ce qui génère un champ de mouvement. Un exemple de flux optique d'une séquence issue de la base de données CK+ est présenté dans la (Figure I.11) [67]. Cette technique permet de suivre les déplacements d'objets ou de caractéristiques au cours du

temps, en fournissant non seulement des informations sur le mouvement, mais aussi des détails concernant la structure tridimensionnelle de l'objet. Negahdaripour [71], considère le flux optique comme des modifications géométriques et radiatives des images dynamiques. Le principe de base de cette technique est d'attribuer à chaque pixel d'une image un vecteur de vitesse, créant ainsi un champ de mouvement. Pendant le mouvement, chaque point de l'image correspond à un point réel de l'objet. Un des principaux avantages de la méthode du flux optique est qu'elle capture non seulement les informations relatives au mouvement de la cible, mais qu'elle fournit également des détails sur sa structure tridimensionnelle.



**Figure I.11: Flux optique d'une séquence issue de la base de données CK+ [27].**

Yacoob et Davis [70], ont développé une méthode innovante pour analyser et reconnaître les expressions faciales à partir de séquences d'images en utilisant le flux optique pour détecter les mouvements du visage. Inspirée par des concepts psychologiques, une représentation symbolique de niveau intermédiaire a été créée pour interpréter ces mouvements. Testée sur un large ensemble de données, cette méthode a démontré son efficacité pour identifier six expressions faciales ainsi que le clignement des yeux.

Cohn et al. [72], ont proposé une méthode basée sur le suivi des points caractéristiques à l'aide du flux optique pour analyser les changements subtils dans les expressions faciales. Contrairement aux approches existantes qui se concentrent sur un nombre limité d'expressions prototypiques, cette méthode permet de détecter des variations fines en analysant des séquences d'images de 100 jeunes adultes. Les unités d'action et leurs combinaisons ont été sélectionnées pour l'analyse et suivies automatiquement, montrant une grande validité par rapport au codage humain avec le Facial Action Coding System (FACS).

Sanchez et al. [73], ont développé et analysé un système semi-automatique de classification des expressions faciales utilisant des séquences d'images pour décrire la dynamique des expressions. Ils ont comparé deux méthodes de suivi basées sur le flux optique : la première, appelée suivi de points caractéristiques, se concentre sur un ensemble réduit de points faciaux distinctifs, tandis que la seconde, nommée suivi dense du flux, utilise un ensemble plus vaste de points répartis uniformément sur la région centrale du visage. Les deux méthodes ont été testées à deux résolutions d'image différentes (originale et réduite de 40%) sur les bases de données Cohn-Kanade et MMI. Les résultats montrent que la méthode de suivi dense du flux a légèrement surpassé la méthode de suivi de points caractéristiques en termes de taux de reconnaissance sur la base de données Cohn-Kanade (95,45% contre 92,42%), mais avec un temps de traitement 68,24% plus élevé. Sur la base de données MMI, les taux de reconnaissance étaient similaires avec les deux méthodes, bien que le suivi dense du flux ait nécessité un prétraitement supplémentaire pour corriger les mouvements globaux de la tête. La méthode de suivi dense du flux présente l'avantage de nécessiter moins de points à marquer manuellement et de maintenir des résultats similaires à différentes résolutions, mais elle est également plus lente, en particulier à la résolution originale.

Allaert et al. [74], ont mené une analyse approfondie des techniques de flux optique pour la reconnaissance des expressions faciales, comparant diverses approches sur plusieurs ensembles de données. Leur objectif était d'évaluer l'efficacité des techniques de flux optique pour encoder les mouvements faciaux. Ils ont découvert que certaines méthodes, notamment celles de Farneback[75]<sup>(1)</sup> et FlowField[76]<sup>(2)</sup>, surpassent les autres en termes de performance. Les résultats soulignent l'importance de l'approximation précise des mouvements pour éviter la propagation d'erreurs, cruciales pour la reconnaissance des expressions faciales. En outre, les auteurs ont montré que l'augmentation artificielle des données en combinant plusieurs approches de flux optique peut améliorer la précision de la classification, avec des gains de performance allant de 2 % à 33 %, selon les approches utilisées et les ensembles de données testés.

<sup>(1)</sup> La méthode de Farneback utilise un modèle de mouvement par translation entre les zones de voisinage de deux images successives dans une décomposition pyramidale. Elle emploie une expansion polynomiale pour estimer les intensités des pixels, en commençant le suivi à basse résolution et en progressant jusqu'à convergence. Cette approche permet de gérer efficacement les grands déplacements de pixels, même au-delà de la taille du voisinage.

<sup>(2)</sup> La méthode FlowFields emploie une technique de correspondance dense qui minimise les valeurs aberrantes. Contrairement aux approches nécessitant une régularisation ou un lissage, cette méthode se concentre sur une recherche orientée par les données pour identifier les correspondances fiables tout en évitant les points erronés.

#### **I.4.1.3.2.2. Suivi des points caractéristiques**

Le suivi des points caractéristiques pour la reconnaissance des expressions faciales dans les séquences d'images implique l'identification et le suivi de points clés du visage, tels que les coins des yeux, la bouche ou le nez, à travers une série d'images. Cette technique permet de saisir les mouvements et les changements du visage au fil du temps, ce qui est crucial pour l'analyse des expressions faciales dynamiques. En observant ces points, il devient possible de détecter des variations subtiles dans les expressions, facilitant ainsi la reconnaissance des émotions[32].

Tong et al. [77], ont développé une approche hiérarchique et multi-états, dépendante de la pose, pour la détection et le suivi des caractéristiques faciales sous des variations d'expressions et de poses du visage. Ils introduisent une représentation hybride, combinant les ondelettes de Gabor et des profils de niveaux de gris, pour une représentation efficace des points caractéristiques. Un modèle statistique hiérarchique de la forme du visage est également développé pour capturer à la fois la forme globale du visage et les détails structurels de chaque composant facial. Des modèles de forme locale multi-états sont également utilisés pour gérer les variations de forme des composants faciaux en fonction des expressions. Lors de la détection et du suivi, les états des composants faciaux et les positions des points caractéristiques sont estimés de manière dynamique à l'aide d'un modèle de mesure hypothétique commuté (SHM)<sup>(3)</sup>. Les tests expérimentaux montrent que la méthode proposée permet un suivi précis et robuste des caractéristiques faciales en temps réel, malgré les variations d'expressions et de poses.

<sup>(3)</sup> Le modèle de changement d'hypothèses de mesures (SHM) est une approche utilisée pour traiter des incertitudes ou des variations dans les mesures au sein de systèmes dynamiques. Il fonctionne en alternant entre différentes hypothèses de mesure, en s'adaptant aux changements des conditions ou des caractéristiques de l'environnement. Ce modèle est souvent utilisé dans des contextes où les mesures peuvent être bruitées, incertaines, ou sujettes à des modifications, permettant ainsi une estimation plus robuste et précise des paramètres ou des états du système [78].

Li et al. [79], ont proposé un cadre hiérarchique fondé sur un Réseau Bayésien Dynamique pour le suivi simultané des caractéristiques faciales et la reconnaissance des expressions faciales. Ce modèle améliore de manière significative la précision du suivi des traits faciaux et la reconnaissance des unités d'action faciales en représentant systématiquement les relations entre les différents niveaux d'activités faciales et

en intégrant les informations temporelles. Cependant, la reconnaissance des six expressions de base n'atteint pas encore les performances des méthodes de pointe, car l'expression globale est déduite des mesures des unités d'action et des points caractéristiques faciaux sans utilisation de mesures spécifiques aux expressions. Les améliorations observées proviennent principalement de la combinaison des modèles d'action faciale avec les mesures d'image, permettant de compenser les erreurs dans les mesures faciales et des unités d'action grâce aux relations intégrées au modèle.

He et al. [80], ont développé une méthode pour la reconnaissance des expressions faciales spontanées en se basant sur le suivi de points caractéristiques. Le processus débute par la normalisation des séquences d'expressions en alignant les coordonnées des pupilles. Ensuite, 23 points sont annotés manuellement dans les images de début et de fin de l'expression. Un filtre de Kalman <sup>(4)</sup> est utilisé pour le suivi de ces points, et deux types de caractéristiques sont extraits : les déplacements des points et les variations de distance entre eux. Un modèle de Markov caché <sup>(5)</sup> sert de classificateur. Les tests réalisés sur la base de données USTC-NVIE démontrent que le suivi des points avec le filtre de Kalman est efficace pour localiser les points, et que les variations de distance entre les points sont plus pertinentes que les déplacements pour la reconnaissance des expressions.

<sup>(4)</sup> Le filtre de Kalman [81], est un algorithme utilisé pour estimer les états d'un système dynamique en combinant des prévisions basées sur un modèle avec des observations réelles, malgré la présence de bruit et d'incertitude. Il fonctionne en deux étapes principales : d'abord, il prédit l'état futur du système et l'incertitude associée à cette prédiction ; ensuite, il ajuste cette estimation en fonction des nouvelles données observées pour affiner les prévisions. Ce processus permet d'obtenir des estimations plus précises et fiables dans des environnements où les mesures peuvent être perturbées ou incomplètes.

<sup>(5)</sup> Le modèle de Markov caché (HMM) [82], est une méthode statistique utilisée pour représenter des processus où l'état du système n'est pas entièrement observable. Il se compose de deux éléments clés : un processus de Markov sous-jacent qui gouverne la séquence des états invisibles, et une série d'observations qui en dépendent.

Yao et al. [83], ont proposé une méthode de reconnaissance des micro-expressions basée sur le suivi et l'alignement des points caractéristiques, utilisant les systèmes FACS et le modèle TLD (Tracking Learning Detection) <sup>(6)</sup>. Le point de départ pour la détection des points caractéristiques dans la première image repose sur Hough Forest <sup>(7)</sup>, et l'extraction des caractéristiques est réalisée avec le Local Binary Pattern (LBP) pour améliorer la précision. Les tests effectués sur les bases de données CASME et SMIC montrent que cette méthode est efficace et performante pour reconnaître certaines micro-expressions spécifiques. En outre, elle se révèle plus précise que les approches basées sur le modèle d'interpolation temporelle (TIM).

<sup>(6)</sup> Le modèle TLD (Tracking-Learning-Detection) [84], est une approche qui permet de suivre un objet dans une séquence vidéo, tout en adaptant dynamiquement son modèle d'apparence. Il se compose de trois éléments clés : le suivi, qui estime la position de l'objet d'une image à l'autre ; l'apprentissage, qui met à jour le modèle en fonction des variations d'apparence de l'objet ; et la détection, qui intervient pour rectifier le suivi en redétectant l'objet en cas de perte. Cette approche est particulièrement efficace dans des situations où l'apparence de l'objet peut changer ou être partiellement obstruée.

<sup>(7)</sup> Le Hough Forest [85], est une technique qui intègre les concepts des forêts d'arbres décisionnels avec la transformée de Hough pour accomplir des tâches de détection et de localisation d'objets dans des images. Chaque arbre dans la forêt participe à la segmentation et à la classification des pixels en s'appuyant sur les données d'entraînement. Collectivement, ces arbres permettent d'accumuler des votes pour déterminer les positions potentielles de l'objet dans l'image, ce qui aide à identifier les zones où l'objet est le plus susceptible d'apparaître. Cette approche est particulièrement adaptée à la détection d'objets dans des scénarios où l'apparence, la pose ou l'échelle varient considérablement.



#### I.4.1.4. Classification des expressions faciales

La classification des expressions faciales consiste à identifier et à catégoriser les différentes expressions du visage humain. Parmi les méthodes de classification les plus utilisées pour la reconnaissance des expressions faciales, on trouve notamment le modèle de Markov caché (HMM), les réseaux de neurones artificiels (ANN), les réseaux bayésiens (BN), l'algorithme des k-plus proches voisins (KNN), les machines à vecteurs de support (SVM), ainsi que la classification basée sur la représentation parcimonieuse (SRC) [32].

Le (Tableau I.9) présente un résumé des diverses méthodes de classification mentionnées dans la revue [32]. Il met en avant les caractéristiques essentielles de chaque méthode de reconnaissance des expressions faciales, tout en fournissant des exemples concrets et récents de leurs applications. Cette présentation permet ainsi une comparaison claire et concise des différentes approches, facilitant la compréhension de leurs avantages et limitations respectifs.

**Tableau I.9: Les modèles de machine learning les plus connus dans la littérature.**

| Méthode                              | Description                                                                                                                                                                                                                                                                                                                                                | Exemples d'Applications                                                                                                                                                                                                                                                                                                                                                              |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Modèle de Markov Caché (HMM)         | Modèle statistique combinant une chaîne de Markov non observable avec des distributions de probabilité pour modéliser des séquences temporelles avec des états cachés. HMM est défini par un triplet $\lambda = (A, B, \pi)$ , où $A$ est la matrice de transition, $B$ la distribution d'observation, et $\pi$ la distribution initiale.                  | <ul style="list-style-type: none"> <li>- Algorithme d'Amélioration de l'Entraînement des Modèles de Markov Cachés et Son Analyse dans la Reconnaissance des Expressions Faciales des Athlètes : Li1 et Bai [86]</li> <li>- Analyse des Expressions Faciales des Nourrissons : Vers un Système de Surveillance Vidéo en Temps Réel Utilisant R-CNN et HMM : Li et al. [87]</li> </ul> |
| Réseau de Neurones Artificiels (ANN) | Modèle computationnel capable de capturer des relations non linéaires complexes entre les entrées et les sorties. Les réseaux les plus couramment utilisés en REF sont le Perceptron Multi-Couche (MLP) et le Réseau de Neurones à Fonction de Base Radiale (RBFNN). Le RBFNN utilise des fonctions gaussiennes pour l'apprentissage des caractéristiques. | <ul style="list-style-type: none"> <li>- Reconnaissance des expressions faciales avec les polynômes de Legendre et un MLP à connexion partielle : Karimi et Heidarian [88]</li> <li>- Reconnaissance des expressions faciales utilisant le descripteur local de Weber et un réseau de neurones F-RBFNN[89]</li> </ul>                                                                |
| Réseau Bayésien (BN)                 | Réseau graphique probabiliste représentant les dépendances entre variables via un graphe acyclique dirigé. Utilisé pour la classification en modélisant l'incertitude et l'incomplétude des données. Les BN classifient les échantillons en fonction des structures de dépendance et des                                                                   | <ul style="list-style-type: none"> <li>- Détection de deux émotions faciales (joie et tristesse) basée sur un classificateur bayésien naïf : Flores-Juarez et al. [90]</li> <li>- Reconnaissance et estimation de l'intensité des expressions faciales à l'aide de classificateurs d'ensemble : Nomiya et al. [91]</li> </ul>                                                        |



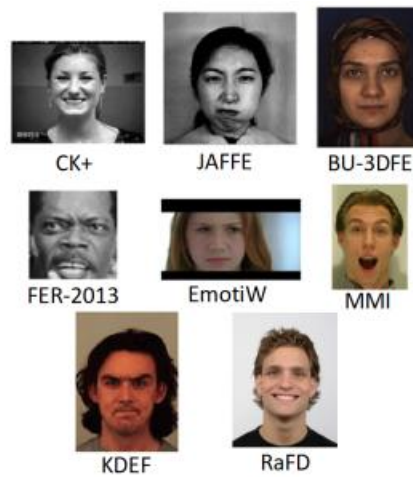
|                              |                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                            |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | distributions des caractéristiques.                                                                                                                                                                                                                                                                                    | - Analyse des expressions faciales basée sur les moments statistiques : Dixit et Gaikwad [92]                                                                                                                                                                                              |
| K-Nearest Neighbour (KNN)    | Algorithme de classification par instance basé sur la proximité dans l'espace des caractéristiques. Un échantillon est classé en fonction des $k$ voisins les plus proches, en général selon la distance euclidienne. Simple et efficace pour des tâches où les frontières de classe sont bien définies.               | - Reconnaissance des expressions faciales à l'aide de l'algorithme de Viola-John modifié et du classificateur KNN : Yadav et Singha [93]<br>- Implémentation d'Efficientnet-Lite et Hybrid CNN-KNN pour la reconnaissance des expressions faciales sur Raspberry Pi : AB WAHAB et al. [94] |
| Support Vector Machine (SVM) | Méthode de classification robuste basée sur la minimisation du risque structurel. Elle projette les données dans un espace de plus grande dimension via des fonctions de noyau pour trouver un hyperplan optimal séparant les classes. Les noyaux couramment utilisés incluent linéaire, polynomial, RBF, et sigmoïde. | - Reconnaissance des expressions faciales basée sur l'apprentissage par transfert et SVM : Yang et al. [95]<br>- Classification des unités d'action pour la reconnaissance des expressions faciales à l'aide de l'apprentissage actif et du SVM : Yao et al. [96]                          |

#### I.4.2. Bases de données REF

Les bases de données d'images pour la reconnaissance des expressions faciales jouent un rôle crucial dans l'évaluation et le développement des systèmes REF. Parmi les plus utilisées[25] [27], la base de données **The Extended Cohn-Kanade (CK+)** [97], est très populaire pour les environnements contrôlés en laboratoire, offrant des séquences qui capturent la transition d'une expression neutre à une expression de pointe. Pour l'évaluation, on sélectionne souvent les dernières images de chaque séquence, en appliquant des expérimentations de validation croisée  $n$ -plis. La base de données **MMI** [98], suit un schéma similaire, mais avec des séquences étiquetées début-apex-fin, où l'expression retourne à la neutralité après avoir atteint son pic. **Oulu-CASIA** [99], se distingue par ses 2 880 séquences d'images capturées sous différentes conditions d'éclairage, souvent utilisées pour la validation croisée à 10 plis.

La base de données **The Japanese Female Facial Expression (JAFPE)**[100], propose 213 images d'expressions faciales posées par des femmes japonaises, et est fréquemment utilisée pour des expériences de validation croisée en laissant un sujet de côté. **FACES** [101], une base de données d'expressions faciales chez les femmes et les hommes jeunes, d'âge moyen et plus âgés. **Facial Expression Recognition 2013 database (FER2013)** [102], quant à elle, est une base de données à grande échelle, non contrainte, collectée automatiquement par Google et contenant des images redimensionnées à 48x48 pixels, idéale pour l'entraînement de modèles de REF. **The Acted Facial Expressions in the Wild (AFEW)** [103]et **The Static Facial Expressions in the Wild (SFEW)** [104], sont des bases de données issues de films, capturant des expressions spontanées sous diverses conditions. **Multi-PIE** [105] et **The Binghamton University 3D Facial Expression (BU-3DFE)** [106]/(**BU-4DFE**) [107], offrent des séquences multivues et 3D, respectivement, pour des analyses plus sophistiquées des expressions faciales.

En ce qui concerne les grandes bases de données, **EmotioNet** [108] et **The Real-world Affective Face Database (RAF-DB)** [109], se démarquent avec des centaines de milliers d'images annotées pour les expressions de base et composées, tandis qu'**AffectNet** [110], reste la plus grande base de données avec plus d'un million d'images annotées manuellement. Enfin, **The Expression in-the-Wild Database (ExpW)** [111], propose près de 92 000 images d'expressions faciales téléchargées à partir de Google, toutes annotées pour l'une des sept catégories d'expressions de base. **4DFAB** [112], ajoute une dimension temporelle en offrant des vidéos dynamiques 4D de visages capturés sur plusieurs années. Parmi les bases de données moins conventionnelles, **eNTERFACE'05** [113], **Karolinska Directed Emotional Faces database (KDEF)** [114], **Radboud Faces Database (RaFD)** [115], et **The Indian Spontaneous Expression Database (ISED)** [116], fournissent des séquences audio-visuelles et des images capturées sous divers angles et conditions, contribuant ainsi à la diversité et à la richesse des données disponibles pour la recherche en REF.



**Figure I.12: Exemples de bases de données de la reconnaissance des expressions faciales** [27].

Le (Tableau I.10) présente un aperçu de ces ensembles de données, incluant la référence principale, le nombre de sujets, le nombre d'images ou d'échantillons vidéo, l'environnement de collecte, la répartition des expressions et des informations supplémentaires [25].

**Tableau I.10: Aperçu de bases de données de la reconnaissance des expressions faciales.**

| Base de données | Échantillons                 | Sujets | Condi<br>on de<br>collecte | Méthode<br>d'élucidation | Répartition<br>des<br>expressions          | Accès                                                                                                                      |
|-----------------|------------------------------|--------|----------------------------|--------------------------|--------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| CK+             | 593 séquences<br>d'images    | 123    | Laborat<br>oire            | Posé &<br>Spontané       | Sept<br>expressions<br>de base +<br>mépris | <a href="http://www.consortium.ri.cmu.edu/ckagree/">http://w<br/>ww.con<br/>sortium.<br/>ri.cmu.e<br/>du/ckagr<br/>ee/</a> |
| MMI             | 740 images et<br>2900 vidéos | 25     | Laborat<br>oire            | Posé                     | Sept<br>expressions<br>de base             | <a href="https://mmiface.db.eu/">https://<br/>mmiface<br/>db.eu/</a>                                                       |
| JAFFE           | 213 images                   | 10     | Laborat<br>oire            | Posé                     | Sept<br>expressions<br>de base             | <a href="http://www.kasrl.org/jaffe.html">http://w<br/>ww.kasr<br/>l.org/jaff<br/>e.html</a>                               |

|           |                 |     |             |                 |                                                                       |                                                                                                                                                                                                                             |
|-----------|-----------------|-----|-------------|-----------------|-----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| FACES     | 2052 images     | 171 | Laboratoire | Spontané        | Six expressions (colère, dégoût, peur, joie, tristesse et neutralité) | <a href="http://faces.mpi-b-berlin.mpg.de">http://faces.mpi-b-berlin.mpg.de</a>                                                                                                                                             |
| TFD       | 112234 images   | N/A | Laboratoire | Posé            | Sept expressions de base                                              | <a href="mailto:josh@mplab.ucsd.edu">josh@mplab.ucsd.edu</a>                                                                                                                                                                |
| FER-2013  | 35887 images    | N/A | Web         | Posé & Spontané | Sept expressions de base                                              | <a href="https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge">https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge</a> |
| AFEW 7.0  | 1 809 vidéos    | N/A | Film        | Posé & Spontané | Sept expressions de base                                              | <a href="https://sites.google.com/site/emotiwchallenge/">https://sites.google.com/site/emotiwchallenge/</a>                                                                                                                 |
| SFEW 2.0  | 1766 images     | N/A | Film        | Posé & Spontané | Sept expressions de base                                              | <a href="https://cs.anu.edu.au/few/emotiw2015.html">https://cs.anu.edu.au/few/emotiw2015.html</a>                                                                                                                           |
| Multi-PIE | 755370 images   | 337 | Laboratoire | Posé            | Sourire, surprise, clignement, dégoût, cri et neutre                  | <a href="http://www.flintbox.com/public/project/4742/">http://www.flintbox.com/public/project/4742/</a>                                                                                                                     |
| BU-3DFE   | 2,500 3D images | 100 | Laboratoire | Posé            | Sept expressions de base                                              | <a href="http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFEAnalysis.html">http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFEAnalysis.html</a>                                                                       |

|            |                          |     |             |                 |                                                         |                                                                                                                                                       |
|------------|--------------------------|-----|-------------|-----------------|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| BU-4DFE    | 606 3D sequences         | 101 | Laboratoire | Posé            | Sept expressions de base                                | <a href="http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFEAnalysis.html">http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFEAnalysis.html</a> |
| Oulu-CASIA | 2,880 image sequences    | 80  | Laboratoire | Posé            | Six expressions de base (sans neutre)                   | <a href="http://www.cse.oulu.fi/CMV/Downloads/Oulu-CASIA">http://www.cse.oulu.fi/CMV/Downloads/Oulu-CASIA</a>                                         |
| RaFD       | 1,608 images             | 67  | Laboratoire | Posé            | Sept expressions de base plus le mépris                 | <a href="http://www.socsci.ru.nl:8180/RaFD2/RaFD">http://www.socsci.ru.nl:8180/RaFD2/RaFD</a>                                                         |
| KDEF       | 4,900 images             | 70  | Laboratoire | Posé            | Sept expressions de base                                | <a href="http://www.emotionlab.se/kdef/">http://www.emotionlab.se/kdef/</a>                                                                           |
| EmotioNet  | 1,000,000 images         | N/A | Internet    | Posé & Spontané | 23 expressions de base ou expressions composées         | <a href="http://cbl.ece.ohio-state.edu/dbformotionet.html">http://cbl.ece.ohio-state.edu/dbformotionet.html</a>                                       |
| RAF-DB     | 29672 images             | N/A | Internet    | Posé & Spontané | Sept expressions de base et douze expressions composées | <a href="http://www.whdeng.cn/RAF/model1.html">http://www.whdeng.cn/RAF/model1.html</a>                                                               |
| AffectNet  | 450,000 images (labeled) | N/A | Internet    | Posé & Spontané | Sept expressions de base                                | <a href="http://mohammadmahoor.com/databases-codes/">http://mohammadmahoor.com/databases-codes/</a>                                                   |
| ExpW       | 91,793 images            | N/A | Internet    | Posé & Spontané | Sept expressions de base                                | <a href="http://mmlab.ie.cuhk.edu.hk/projects/socialrelation/index.html">http://mmlab.ie.cuhk.edu.hk/projects/socialrelation/index.html</a>           |

|               |                      |     |             |                 |                                                                       |                                                                                                       |
|---------------|----------------------|-----|-------------|-----------------|-----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| 4DFAB         | 1.8 million 3D faces | 180 | Laboratoire | Posé & Spontané | Sept expressions de base                                              | N/A                                                                                                   |
| eINTERFACE'05 | 1166 vidéos          | 42  | Laboratoire | Spontané        | Six expressions de base (sans neutre)                                 | <a href="https://www.INTERFACE.net/INTERFACE05/">https://www.INTERFACE.net/INTERFACE05/</a>           |
| ISED          | 428 vidéos           | 50  | Laboratoire | Spontané        | Quatre expressions : la joie, la surprise, la tristesse et le dégoût. | <a href="https://sites.google.com/site/iseddatabase/">https://sites.google.com/site/iseddatabase/</a> |

Après avoir défini les bases d'images les plus connues pour la reconnaissance des expressions faciales, il est important d'introduire des bases de données spécifiques dédiées à l'étude des micro-expressions.

La base de données **CASME** (Chinese Academy of Sciences Micro-expression) [117], contient 195 micro-expressions capturées à 60 images par seconde. Ces micro-expressions ont été extraites de plus de 1500 mouvements faciaux provoqués. Chaque échantillon est annoté avec des unités d'action (AUs) et des émotions, ainsi que des moments précis de début, d'apogée et de fin. Les participants, au nombre de 35, âgés en moyenne de 22 ans, ont donné leur consentement pour l'étude. La sélection porte sur des micro-expressions ne dépassant pas 500 millisecondes, mais inclut aussi certaines expressions faciales rapides avec un début inférieur à 250 millisecondes. Les enregistrements ont été réalisés avec deux caméras et dans deux environnements différents, conduisant à une classification des données en deux groupes distincts : Classe A et Classe B. Pour la Classe A, les échantillons ont été capturés à l'aide de la caméra BenQ M31, avec une résolution de  $1280 \times 720$  pixels et sous lumière naturelle. Pour la Classe B, les enregistrements ont été réalisés avec la caméra Point Grey GRAS-03K2C, avec une résolution de  $640 \times 480$  pixels, dans une pièce éclairée par deux LED. La Classe B comprend 95 échantillons d'une durée maximale de 500 millisecondes et 30 échantillons avec une phase de déclenchement de moins de 250 millisecondes.

Après le succès de la première version de la base de données CASME, les chercheurs ont cherché à améliorer davantage la qualité et la précision des données pour mieux capturer les micro-expressions faciales. C'est dans ce contexte qu'est née la base CASME II. La base de données **CASME II** [118], présente les caractéristiques suivantes : les micro-expressions sont spontanées et dynamiques, avec des images de référence (neutres) conservées avant et après chaque micro-expression, facilitant ainsi l'évaluation des algorithmes de détection. Les enregistrements offrent une haute résolution temporelle (200 fps) et une résolution faciale relativement élevée ( $280 \times 340$  pixels). De plus, les enregistrements sont réalisés sous un éclairage stable et uniforme, réduisant les reflets sur le visage.

La base de données **SMIC** (Spontaneous Micro-expression Database) [119], contient 164 clips de micro-expressions enregistrés auprès de 16 participants. Elle offre des performances de détection et de reconnaissance de micro-expressions comme références de base, fournissant ainsi une ressource essentielle pour tester les systèmes automatiques d'analyse de micro-expressions, ce qui n'était pas possible avec les bases de données précédentes. La base SMIC, comprend des clips provenant de participants âgés en moyenne de 28,1 ans, incluant six femmes et dix hommes, répartis entre huit Caucasiens et huit Asiatiques.

Le (Tableau I.11) présente un aperçu des bases de données les plus connues dédiées à l'étude des micro-expressions faciales. Ces bases sont spécifiquement conçues pour capturer et analyser des mouvements faciaux subtils et rapides, souvent liés à des émotions dissimulées. Elles se distinguent par leurs

caractéristiques techniques, telles que la résolution temporelle élevée et les méthodes de marquage, et jouent un rôle crucial dans le développement d'algorithmes de détection et de reconnaissance automatique des micro-expressions. Ce tableau met en évidence les principales spécificités de chaque base, permettant ainsi une comparaison directe des bases de données.

**Tableau I.11: Les bases de données les plus connues dédiées à l'étude des micro-expressions faciales.**

|                    | CASME    | CASME II | SMIC     |
|--------------------|----------|----------|----------|
| Sujets             | 19       | 26       | 16       |
| Échantillons       | 195      | 247      | 164      |
| Images par seconde | 60       | 200      | 100      |
| Posé/spontané      | Spontané | Spontané | Spontané |
| Étiquettes AU      | Oui      | Oui      | Non      |
| Classes d'émotions | 7        | 5        | 3        |

#### I.4.3. REF en mode indépendant de la personne

Les expressions faciales varient d'une personne à l'autre en fonction de nombreux facteurs, y compris les différences culturelles, la personnalité, et le contexte social. Selon l'étude classique d'Ekman et Friesen [120], bien que certaines expressions faciales soient universelles, des variations importantes existent en raison de différences culturelles dans l'expression des émotions. Ces différences culturelles sont également soulignées par Russell[121], qui a examiné la reconnaissance des émotions à travers différentes cultures et a mis en évidence des variations individuelles significatives.

Par ailleurs, Keltner et Kring [122], ont montré que les expressions faciales ne servent pas seulement à communiquer des émotions, mais aussi à remplir des fonctions sociales variées, ce qui peut influencer leur manifestation en fonction du contexte et de la personnalité. Matsumoto et Hwang [123], ajoutent que la culture et les normes sociales jouent un rôle crucial dans la façon dont les émotions sont exprimées et interprétées à travers les expressions faciales.

Enfin, une étude de Gross et Levenson [124], a révélé que les individus montrent des différences notables dans leur expressivité émotionnelle, tant au niveau comportemental que physiologique, ce qui se traduit par des variations dans la manière dont les émotions sont exprimées par les expressions faciales. Ces études montrent que, bien que certaines expressions faciales puissent être universelles, elles sont profondément influencées par des facteurs individuels et culturels.

Dans ce qui suit, nous présenterons les différents points essentiels tirés du paragraphe précédent, qui résument les études menées sur la variabilité des expressions faciales d'une personne à l'autre :

- *Variabilité culturelle* : Certaines expressions faciales, bien que souvent considérées comme universelles, montrent des variations significatives selon les cultures, influencées par des normes culturelles spécifiques qui dictent comment et quand les émotions doivent être exprimées.
- *Influence des normes sociales* : Les normes et contextes sociaux affectent à la fois l'expression et la perception des émotions. Cela signifie que l'interprétation d'une même expression faciale peut varier d'une culture à une autre.
- *Impact de la personnalité et du contexte* : La manière dont les émotions sont exprimées varie selon la personnalité de l'individu et le contexte social, ce qui peut entraîner des différences notables dans l'expression de la même émotion.
- *Variabilité individuelle dans l'expressivité émotionnelle* : Chaque individu présente des différences dans la façon dont il exprime ses émotions, que ce soit par des indices comportementaux ou physiologiques, ce qui conduit à une grande diversité d'expressions faciales pour une même émotion.

#### **I.4.3.1. Défis de la REF en mode indépendant de la personne**

Pour la reconnaissance des expressions faciales en mode indépendant de la personne, plusieurs études ont mis en lumière les défis majeurs liés à cette tâche. Ci-après les défis principaux :

- 1- Variabilité inter-individuelle : Les différences anatomiques et musculaires entre les individus influencent considérablement la façon dont les expressions faciales sont perçues et interprétées, ce qui rend difficile l'élaboration de modèles universels pour la reconnaissance des émotions [125] [126].
- 2- Dépendance au contexte : En mode indépendant de la personne, il est particulièrement difficile de tenir compte des variations contextuelles qui influencent la manière dont une émotion est exprimée. Les expressions faciales peuvent être modifiées par l'environnement immédiat, les interactions sociales, et les attentes culturelles. Ces variations contextuelles compliquent la généralisation des modèles de reconnaissance faciale à différents scénarios et individus [121][127].
- 3- Variabilité des données d'entraînement : La reconnaissance des expressions faciales indépendamment de la personne nécessite des ensembles de données d'entraînement diversifiés qui capturent une large gamme de variations faciales. Cependant, de nombreux ensembles de données peuvent être biaisés culturellement ou démographiquement, limitant la capacité des modèles à généraliser à de nouvelles personnes ou à des groupes démographiques différents [128] [110].
- 4- Problèmes de Normalisation : Les différences dans l'angle de vue, l'éclairage et les expressions faciales préalables peuvent affecter la manière dont les expressions sont capturées et interprétées [110] [129]. Pour que les modèles soient efficaces en mode indépendant de la personne, ils doivent être robustes face à ces variations afin de maintenir une performance élevée quelle que soit la configuration de capture.
- 5- Problèmes de généralisation : Les modèles de reconnaissance des expressions faciales formés sur des ensembles de données spécifiques peuvent échouer à se généraliser efficacement à des individus et contextes différents, soulignant la nécessité de données d'entraînement plus variées et représentatives [130].

Pour surmonter les défis de la reconnaissance des expressions faciales et en particulier en mode indépendant de la personne, plusieurs approches et solutions peuvent être mises en œuvre. Voici un aperçu des stratégies possibles pour aborder ces défis :

##### *1- Sélection de Caractéristiques Pertinentes*

La sélection de caractéristiques consiste à identifier et à utiliser uniquement les caractéristiques les plus pertinentes et informatives pour la tâche de reconnaissance des expressions faciales. Cette approche aide à réduire la complexité du modèle et à se concentrer sur les aspects essentiels qui sont robustes face aux variations interindividuelles. Cela améliore la précision en minimisant le bruit et les informations non pertinentes et permet une meilleure généralisation aux expressions faciales de différentes personnes, même avec des variations anatomiques et contextuelles (voir [131]).

##### *2- Augmentation des Données*

Créer des versions modifiées des images d'entraînement par rotation, recadrage, et ajustement de la luminosité pour simuler différentes conditions de vue et d'éclairage permet d'enrichir l'ensemble de données avec une variété d'exemples, améliorant la robustesse du modèle aux variations d'angle de vue et d'éclairage, et aux différences individuelles (voir [132]).

##### *3- Techniques de Normalisation et d'Alignement*

Appliquer des méthodes de normalisation pour corriger les variations d'éclairage et d'alignement des visages, et aligner les images en fonction des points clés du visage permet de réduire l'influence des différences de capture, permettant une interprétation plus cohérente des expressions faciales indépendamment des variations d'angle de vue et d'éclairage (voir [133]).

#### 4- *Modèles Multimodaux*

Intégrer des informations supplémentaires telles que des signaux audio ou des données de mouvement du corps pour compléter les expressions faciales dans le processus de reconnaissance émotionnelle peut améliorer l'interprétation des expressions faciales, rendant les modèles plus résilients aux variations individuelles (voir [134]).

#### 5- *Utilisation de Modèles Profonds Pré entraînés*

Utiliser des modèles de réseaux neuronaux profonds qui ont été préalablement entraînés sur de vastes ensembles de données diversifiés. Ces modèles peuvent ensuite être affinés sur des tâches spécifiques de reconnaissance des expressions faciales et améliore ainsi la capacité du modèle à se généraliser à des individus et des contextes non vus auparavant, en capturant des caractéristiques faciales plus généralisables et robustes.

#### 6- *Ensembles de Modèles (Model Ensembles)*

Combiner plusieurs modèles de reconnaissance des expressions faciales pour obtenir une décision finale basée sur le consensus des prédictions des modèles individuels peut réduire les erreurs potentielles causées par un modèle individuel en tirant parti de la force collective des modèles dans l'ensemble, ce qui améliore la précision et la robustesse (voir [135]).

Pour relever les défis complexes de la reconnaissance des expressions faciales en mode indépendant de la personne, plusieurs solutions techniques ont été proposées, allant de l'utilisation de modèles profonds préentraînés et d'ensembles de modèles à l'augmentation des données et à la normalisation avancée des images. Bien que ces approches offrent des avantages significatifs, elles peuvent souvent être coûteuses en termes de ressources informatiques et de temps de traitement. C'est dans ce contexte que la sélection de caractéristiques se distingue comme une méthode particulièrement efficace. En identifiant et en exploitant les caractéristiques les plus pertinentes, cette approche non seulement simplifie le modèle, mais améliore aussi sa robustesse et sa capacité à généraliser à de nouveaux individus, tout en minimisant le risque de surapprentissage.

Ci-après quelques raisons pour lesquelles la sélection de caractéristiques peut être plus avantageuse dans le contexte de la reconnaissance des expressions faciales en mode indépendant de la personne [131] [136] [137].

- Réduction de la complexité et du temps de calcul :  
En limitant le nombre de variables utilisées, la sélection de caractéristiques simplifie le modèle en ne conservant que les éléments les plus pertinents pour identifier les expressions faciales. Cela diminue à la fois la complexité et le temps nécessaire pour entraîner le modèle et effectuer des prédictions. Comparée à d'autres techniques telles que l'utilisation de modèles profonds pré entraînés ou des ensembles de modèles, la sélection de caractéristiques demande moins de ressources informatiques, ce qui facilite une implémentation rapide et efficace.
- Amélioration de la précision par réduction du bruit :  
La sélection de caractéristiques écarte les variables non pertinentes ou redondantes, ce qui réduit le bruit dans les données et permet au modèle de se concentrer sur les informations essentielles pour la reconnaissance des expressions faciales. À l'inverse, des approches comme l'augmentation des données peuvent introduire une complexité supplémentaire sans



nécessairement filtrer les informations non pertinentes, risquant ainsi d'ajouter du bruit au modèle.

- **Meilleure généralisation aux nouveaux individus :**  
Les modèles utilisant la sélection de caractéristiques sont souvent plus aptes à généraliser à de nouvelles personnes et à des contextes variés, car ils se concentrent sur des traits universels des expressions faciales plutôt que sur des variations individuelles. Contrairement aux techniques de normalisation et d'alignement qui gèrent surtout les variations d'éclairage et d'angle de vue, la sélection de caractéristiques permet de mieux s'adapter aux différences anatomiques et expressives entre les individus.
- **Robustesse contre le surapprentissage (overfitting) :**  
En se focalisant sur les caractéristiques les plus pertinentes, la sélection de caractéristiques limite les risques de surapprentissage, où le modèle se spécialise trop dans les détails de l'ensemble d'entraînement. Les approches comme les modèles multimodaux ou les ensembles de modèles, bien qu'efficaces, sont plus complexes et peuvent être plus susceptibles de surapprentissage, notamment en présence de données déséquilibrées.
- **Facilité d'interprétation :**  
La sélection de caractéristiques offre une meilleure transparence en permettant de comprendre quelles caractéristiques influencent les décisions du modèle pour la reconnaissance des expressions faciales. Cette transparence est souvent plus difficile à atteindre avec des modèles profonds pré entraînés ou des ensembles de modèles, qui fonctionnent fréquemment comme des boîtes noires complexes.

En effet, la sélection de caractéristiques se distingue par sa capacité à simplifier les modèles, à améliorer leur performance, et à favoriser leur généralisation tout en minimisant les besoins en calcul et en réduisant le risque de surapprentissage. Ces atouts en font une solution efficace et pragmatique pour la reconnaissance des expressions faciales en mode indépendant de la personne, particulièrement utile dans les environnements avec des ressources de calcul limitées ou lorsqu'une interprétation claire et transparente du modèle est essentielle.

### **I.5. Conclusion**

Dans ce chapitre, nous avons exploré les concepts fondamentaux liés à la reconnaissance automatique des expressions faciales. Nous avons défini les émotions et leur expression faciale, en mettant en avant leur rôle essentiel dans la communication humaine. Ensuite, nous avons présenté les systèmes de reconnaissance automatique des expressions faciales et les défis qu'ils rencontrent, notamment en mode indépendant de la personne.

L'un des principaux défis du REF en mode indépendant de la personne réside dans la variabilité interindividuelle des expressions faciales, due aux différences morphologiques, culturelles et aux styles d'expression propres à chaque individu. Cette variabilité complexifie la généralisation des modèles et souligne l'importance de la sélection de caractéristiques. En effet, une sélection efficace permet de réduire la dimensionnalité des données, d'améliorer la robustesse du modèle et de minimiser le risque de surapprentissage.

Dans le chapitre suivant, nous nous intéresserons plus en détail aux méthodes de sélection de paramètres pertinents, essentielles pour renforcer la performance et la fiabilité des systèmes de reconnaissance.

# Chapitre II

## Sélection des Paramètres Pertinents et leur Impact sur la Reconnaissance des Expressions Faciales

---

|                                                                                                                                                               |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| II.1. Introduction                                                                                                                                            | 42 |
| II.2. Concepts et définitions de la sélection de caractéristiques                                                                                             | 42 |
| II.2.1. Définition de la sélection de caractéristiques                                                                                                        | 42 |
| II.2.2. Processus de sélection des caractéristiques                                                                                                           | 42 |
| II.2.2.1. Génération de sous-ensembles                                                                                                                        | 43 |
| II.2.2.2. Évaluation des sous-ensembles                                                                                                                       | 44 |
| II.2.2.3. Critères d'Arrêt                                                                                                                                    | 45 |
| II.2.2.4. Validation des Résultats                                                                                                                            | 45 |
| II.2.3. L'influence des variables redondantes et corrélées en sélection de caractéristiques                                                                   | 46 |
| II.2.3.1. Effets de la corrélation sur la redondance                                                                                                          | 46 |
| II.2.3.1.1. Corrélation positive                                                                                                                              | 46 |
| II.2.3.1.2. Corrélation négative                                                                                                                              | 46 |
| II.2.3.2. Utilité des variables individuelles en contexte de combinaison                                                                                      | 46 |
| II.2.3.2.1. Variables apparemment inutiles                                                                                                                    | 46 |
| II.2.3.2.2. Approches de sélection                                                                                                                            | 46 |
| II.2.4. Typologie des méthodes de sélection de caractéristiques                                                                                               | 47 |
| II.2.4.1. Méthodes de filtrage                                                                                                                                | 47 |
| II.2.4.2. Méthodes enveloppantes (wrapper)                                                                                                                    | 49 |
| II.2.4.3. Méthodes Intégrées (Embedded Methods)                                                                                                               | 51 |
| II.2.5. Méthodes d'optimisation en sélection de caractéristiques                                                                                              | 53 |
| II.2.5.1. Méthodes traditionnelles d'optimisation                                                                                                             | 53 |
| II.2.5.1.1. Recherche exhaustive                                                                                                                              | 53 |
| II.2.5.1.2. Méthodes séquentielles                                                                                                                            | 53 |
| II.2.5.2. Méthodes inspirées de la nature                                                                                                                     | 53 |
| II.2.5.2.1. Algorithmes génétiques (Genetic Algorithms, GA)                                                                                                   | 54 |
| II.2.5.2.2. Optimisation par essaim particulaire (Particle Swarm Optimization, PSO)                                                                           | 54 |
| II.2.5.2.3. Optimisation par colonies de fourmis (Ant Colony Optimization, ACO)                                                                               | 54 |
| II.3. Pourquoi la sélection de caractéristiques est cruciale pour la REF                                                                                      | 54 |
| II.4. Application de la sélection de caractéristiques à la reconnaissance des expressions faciales                                                            | 55 |
| II.4.1. Sélection de caractéristiques basées sur des méthodes filtres                                                                                         | 55 |
| II.4.2. Sélection de caractéristiques basées sur des méthodes wrapper                                                                                         | 57 |
| II.4.3. Sélection de caractéristiques basées sur des méthodes embedded                                                                                        | 59 |
| II.4.4. Sélection de caractéristiques basées sur des méthodes hybride                                                                                         | 61 |
| II.5. Sélection de caractéristiques pour la reconnaissance des expressions faciales en mode indépendant de la personne : analyse et comparaison des approches | 61 |
| II.6. Conclusion                                                                                                                                              | 67 |

## II.1. Introduction

Dans le domaine de la reconnaissance des expressions faciales, l'efficacité et la précision des modèles de classification dépendent en grande partie de la qualité des données utilisées pour l'apprentissage. L'une des étapes cruciales pour améliorer la performance de ces systèmes est la sélection de caractéristiques pertinentes, un processus qui consiste à identifier et à retenir uniquement les caractéristiques les plus utiles pour différencier les différentes expressions faciales.

Ce chapitre explore en profondeur la sélection de caractéristiques, en mettant l'accent sur son rôle crucial dans le cadre de la reconnaissance des expressions faciales en mode indépendant de la personne. Nous commencerons par définir le concept de sélection de caractéristiques et son importance dans le contexte de l'apprentissage automatique. Ensuite, nous examinerons les différentes méthodes de sélection de caractéristiques, y compris les approches basées sur les filtres, les wrappers et les techniques intégrées. Nous discuterons également des critères de sélection, des stratégies et des développements récents dans le domaine, soulignant comment la sélection de caractéristiques peut être intégrée dans des systèmes intelligents pour améliorer la robustesse et la précision de la reconnaissance des expressions faciales. À travers cette exploration, ce chapitre mettra en évidence l'importance de choisir les caractéristiques les plus pertinentes pour construire des modèles efficaces et optimiser les performances des systèmes de reconnaissance des expressions faciales.

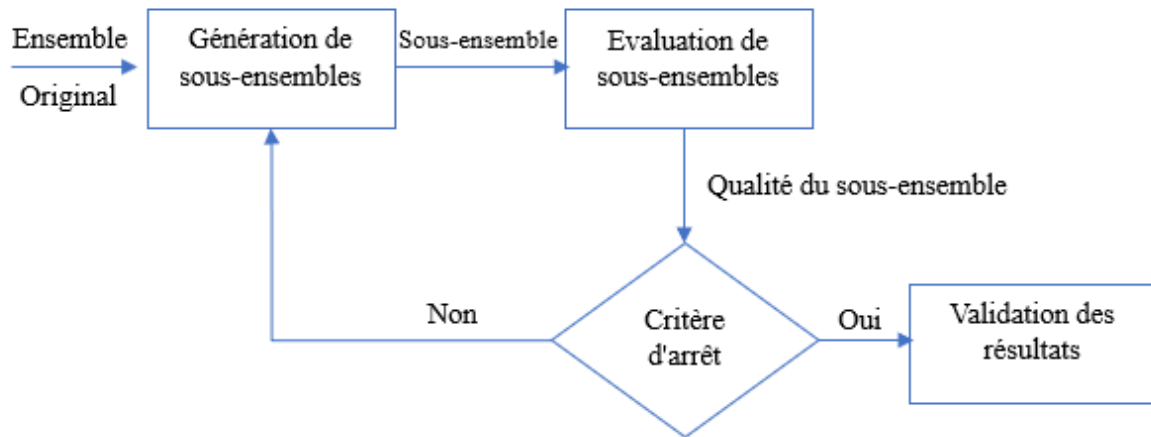
## II.2. Concepts et définitions de la sélection de caractéristiques

### II.2.1. Définition de la sélection de caractéristiques

La sélection de caractéristiques est une technique utilisée en apprentissage automatique et en exploration de données qui consiste à identifier et à conserver uniquement les variables ou caractéristiques les plus pertinentes d'un ensemble de données. Ce processus implique de choisir les caractéristiques les plus pertinentes tout en éliminant celles qui sont non pertinentes, redondantes ou perturbées par du bruit. L'objectif ultime de la sélection de caractéristiques est d'identifier un sous-ensemble optimal des caractéristiques d'origine qui offre les meilleures performances possibles pour une tâche spécifique, sans recourir à des transformations complexes des données [138]. Cette approche permet non seulement d'améliorer la précision des modèles prédictifs, mais aussi de réduire le temps de calcul, d'accroître l'efficacité et de faciliter l'interprétation des modèles en focalisant l'analyse sur les variables les plus influentes.

### II.2.2. Processus de sélection des caractéristiques

Un processus typique de sélection de caractéristiques comprend quatre étapes fondamentales (Figure II.1) : génération de sous-ensembles, évaluation des sous-ensembles, critère d'arrêt, et validation des résultats. La génération de sous-ensembles consiste en une procédure de recherche, qui produit des sous-ensembles de caractéristiques candidats à évaluer, selon une stratégie de recherche spécifique. Chaque sous-ensemble candidat est évalué et comparé au meilleur sous-ensemble trouvé précédemment, en utilisant un certain critère d'évaluation. Si le nouveau sous-ensemble s'avère supérieur, il remplace le précédent. Ce processus de génération et d'évaluation des sous-ensembles se poursuit jusqu'à ce qu'un critère d'arrêt défini soit rempli. Enfin, le meilleur sous-ensemble sélectionné doit généralement être validé par des connaissances préalables ou par divers tests utilisant des ensembles de données synthétiques et/ou réels [137].



**Figure II.1: Les quatre étapes clés de la sélection des caractéristiques [138].**

Dans les sections suivantes, nous allons explorer en détail les quatre étapes principales du processus de sélection de caractéristiques [137].

### II.2.2.1. Génération de sous-ensembles

La génération de sous-ensembles est fondamentalement un processus de recherche heuristique, chaque état dans l'espace de recherche représentant un sous-ensemble candidat à évaluer. Ce processus est influencé par deux facteurs principaux : le point de départ de la recherche et la stratégie de recherche. Premièrement, il faut déterminer le point de départ de la recherche, ce qui impacte la direction de la recherche. Deuxièmement, il faut choisir une stratégie de recherche. Pour un ensemble de données avec  $N$  caractéristiques, il existe  $2^N$  sous-ensembles candidats, rendant l'espace de recherche exponentiellement prohibitif pour une recherche exhaustive même avec un  $N$  modéré. C'est pourquoi différentes stratégies ont été explorées, notamment la recherche complète, séquentielle et aléatoire.

#### **Point de départ de la recherche :**

- Peut commencer avec un ensemble vide et ajouter des caractéristiques progressivement (recherche avant).
- Peut commencer avec un ensemble complet et supprimer des caractéristiques progressivement (recherche arrière).
- Peut débiter des deux côtés en ajoutant et supprimant simultanément (recherche bidirectionnelle).
- Peut commencer avec un sous-ensemble aléatoire pour éviter de tomber dans des optima locaux.

#### **Stratégie de recherche :**

- Recherche Complète : Cette méthode garantit de trouver le meilleur résultat selon le critère d'évaluation utilisé. Bien qu'une recherche exhaustive soit complète, il n'est pas nécessaire qu'une recherche soit exhaustive pour garantir la complétude. Différentes fonctions heuristiques peuvent réduire l'espace de recherche sans compromettre la possibilité de trouver le résultat optimal. Des exemples incluent la recherche par branches et bornes [139] et la recherche par faisceaux [140].
- Recherche Séquentielle : Cette méthode abandonne la complétude et risque donc de manquer des sous-ensembles optimaux. Il existe de nombreuses variations de cette approche, comme la *sélection séquentielle avant*, l'*élimination séquentielle arrière*, et la *sélection bidirectionnelle*. Toutes ces méthodes ajoutent ou suppriment des caractéristiques une à la fois. Une alternative consiste à ajouter (ou supprimer)  $p$  caractéristiques en une étape et à supprimer (ou ajouter)

$q$  caractéristiques dans l'étape suivante ( $p > q$ ). Les algorithmes utilisant la recherche séquentielle sont simples à implémenter et produisent des résultats rapidement, l'ordre de l'espace de recherche étant généralement  $O(N^2)$  ou moins.

- **Recherche Aléatoire** : Cette méthode commence avec un sous-ensemble sélectionné de manière aléatoire et procède de deux manières différentes. La première consiste à suivre une recherche séquentielle, en injectant de la randomisation dans les approches séquentielles classiques, comme l'escalade aléatoire et le recuit simulé[140]. La seconde génère le prochain sous-ensemble de manière complètement aléatoire, sans qu'un sous-ensemble actuel ne croisse ou ne rétrécisse à partir d'un sous-ensemble précédent selon une règle déterministe, une approche également connue sous le nom d'algorithme de Las Vegas. L'utilisation de la randomisation dans ces méthodes aide à échapper aux optima locaux dans l'espace de recherche, et l'optimalité du sous-ensemble sélectionné dépend des ressources disponibles.

### II.2.2.2. Évaluation des sous-ensembles

Un critère d'évaluation est une mesure utilisée pour évaluer la qualité d'un sous-ensemble de caractéristiques généré dans un processus de sélection de caractéristiques. Ce critère détermine l'efficacité du sous-ensemble en fonction de paramètres spécifiques, ce qui signifie qu'un sous-ensemble peut être jugé optimal selon un critère mais pas nécessairement selon un autre. Les critères d'évaluation peuvent être classés en deux grandes catégories : ceux qui sont indépendants des algorithmes de classification appliqués, et ceux qui dépendent directement de la performance de ces algorithmes sur les sous-ensembles de caractéristiques sélectionnés.

- **Critères indépendants** : ces critères évaluent la qualité d'une caractéristique ou d'un sous-ensemble de caractéristiques en exploitant les caractéristiques intrinsèques des données d'entraînement, sans impliquer d'algorithme de classification. Les critères indépendants populaires incluent [141]:
  - *Les mesures de distance* : Ces critères, souvent désignés sous les termes de mesures de séparabilité, de divergence, ou de discrimination, évaluent l'efficacité d'une caractéristique à différencier plusieurs classes. Dans des cas de classification binaire, une caractéristique est jugée plus efficace si elle génère une plus grande distinction entre les probabilités conditionnelles des deux classes. En d'autres termes, si une caractéristique  $X$  parvient à créer une séparation plus marquée entre les classes par rapport à une caractéristique  $Y$ ,  $X$  sera jugée supérieure et préférée.
  - *Les mesures d'information* : Ces mesures évaluent l'apport d'information d'une caractéristique en quantifiant le gain d'information. Ce gain est calculé comme la différence entre l'incertitude initiale (avant l'utilisation de la caractéristique) et l'incertitude restante après avoir pris en compte la caractéristique. Une caractéristique est considérée comme plus efficace si elle permet une réduction significative de l'incertitude en fournissant des informations pertinentes.
  - *Les mesures de dépendance* : Ces mesures, également appelées mesures de corrélation ou de similarité, évaluent la capacité d'une caractéristique à prédire la valeur d'une autre. Dans le contexte de la sélection de caractéristiques pour la classification, l'objectif est de mesurer à quel point une caractéristique est associée à la classe cible. Une caractéristique est préférable si son association avec la classe est plus forte par rapport à une autre caractéristique.
  - *Les mesures de cohérence* : Contrairement aux autres types de mesures, les mesures de cohérence se concentrent principalement sur les informations relatives aux classes. Elles visent à identifier le nombre minimum de caractéristiques nécessaires pour séparer les classes de manière aussi efficace que l'ensemble complet des caractéristiques. Une incohérence est définie comme une situation où deux instances partagent les mêmes valeurs de caractéristiques mais appartiennent à des classes différentes. Les mesures de cohérence cherchent à réduire au maximum ces incohérences.

- **Critères dépendants** : ces critères nécessitent l'utilisation d'un algorithme de classification spécifique pour évaluer les caractéristiques. Ils s'appuient sur la performance de cet algorithme appliqué au sous-ensemble de caractéristiques sélectionné pour décider des caractéristiques à retenir. Bien qu'ils puissent offrir des performances supérieures, ces critères tendent à être plus exigeants en termes de ressources de calcul et peuvent ne pas être adaptés pour d'autres types d'algorithmes de classification.

### II.2.2.3. Critères d'Arrêt

Un critère d'arrêt détermine quand le processus de sélection de caractéristiques doit s'arrêter. Les critères couramment utilisés sont :

- *Complétion de la recherche* : Le processus s'arrête lorsque tous les sous-ensembles possibles ont été examinés, garantissant que le meilleur ensemble de caractéristiques a été trouvé selon les critères d'évaluation choisis.
- *Atteinte d'une limite spécifiée* : La sélection de caractéristiques peut s'arrêter une fois qu'une certaine limite prédéfinie est atteinte, comme un nombre minimal de caractéristiques retenues, un seuil de performance spécifique, ou un nombre maximal d'itérations.
- *Absence d'amélioration* : Si l'ajout ou la suppression de caractéristiques supplémentaires ne conduit plus à une amélioration significative de la performance du modèle, le processus peut être arrêté. Ce critère évite les ajustements inutiles qui n'apportent pas de bénéfices substantiels.
- *Sélection d'un sous-ensemble suffisamment performant* : Si un sous-ensemble de caractéristiques atteint un niveau de performance jugé adéquat pour la tâche en cours (par exemple, une erreur de classification en dessous d'un seuil acceptable), le processus peut s'arrêter, même si des améliorations mineures pourraient encore être possibles.

Ces critères d'arrêt sont essentiels pour équilibrer la qualité des résultats avec l'efficacité computationnelle, en évitant les calculs excessifs tout en assurant que les caractéristiques sélectionnées sont pertinentes et optimales pour les besoins spécifiques de l'application.

### II.2.2.4. Validation des Résultats

Une approche directe pour valider les résultats de la sélection de caractéristiques consiste à comparer les caractéristiques sélectionnées avec des connaissances préalablement établies sur les données. Par exemple, lorsqu'il s'agit de données synthétiques, on peut vérifier si les caractéristiques pertinentes connues sont bien celles qui ont été choisies par le processus de sélection. Cette validation directe permet d'évaluer l'efficacité de l'algorithme de sélection en s'appuyant sur une vérité terrain.

Toutefois, dans les scénarios du monde réel, il est souvent difficile de disposer de connaissances préalables précises sur les caractéristiques pertinentes. Dans de tels cas, des méthodes de validation indirecte sont nécessaires. Une de ces méthodes consiste à observer comment la performance d'un modèle de classification varie en fonction des caractéristiques choisies. En d'autres termes, on peut mesurer l'impact des caractéristiques sélectionnées sur des indicateurs de performance tels que la précision de la classification, le taux de détection, ou l'erreur de prédiction. L'objectif est de vérifier si le sous-ensemble de caractéristiques améliore réellement les performances du modèle comparé à l'utilisation de l'ensemble complet ou d'autres sous-ensembles de caractéristiques.

Ces méthodes de validation, qu'elles soient directes ou indirectes, sont essentielles pour garantir que la sélection de caractéristiques contribue efficacement à la qualité et à la pertinence du modèle construit.

### **II.2.3. L'influence des variables redondantes et corrélées en sélection de caractéristiques**

La sélection de caractéristiques joue un rôle clé dans la création de modèles de machine learning, car elle permet d'identifier les variables les plus pertinentes tout en écartant celles qui sont redondantes ou inutiles. Toutefois, pour effectuer cette sélection efficacement, il est essentiel de considérer non seulement l'importance de chaque variable prise individuellement, mais aussi les interactions potentielles entre elles. Guyon et Elisseeff [131], ont étudié en détail l'influence des variables et de leurs interactions dans le processus de sélection de caractéristiques. Ils ont soulevé trois questions fondamentales : 1) Les variables redondantes peuvent-elles s'entraider ? ; 2) Comment la corrélation influence-t-elle la redondance des variables ? ; 3) Une variable sans valeur apparente peut-elle devenir utile en combinaison avec d'autres ? Dans la suite, nous tenterons d'apporter des éclaircissements sur ces questions.

#### **II.2.3.1. Effets de la corrélation sur la redondance**

##### **II.2.3.1.1. Corrélation positive**

Une forte corrélation positive entre deux variables indique qu'elles varient ensemble, soit en augmentant, soit en diminuant. Cette corrélation peut suggérer une redondance, mais elle peut également signaler une relation complémentaire entre les variables. Par exemple, dans des ensembles de données où les variables reflètent divers aspects d'un même phénomène, une forte corrélation peut en réalité améliorer la capacité du modèle à saisir les détails des données. Ainsi, il est essentiel que les méthodes de sélection distinguent les redondances problématiques des corrélations avantageuses.

##### **II.2.3.1.2. Corrélation négative**

Une corrélation négative, où une variable augmente pendant que l'autre diminue, peut également être significative. Elle peut indiquer que les variables offrent des perspectives divergentes mais complémentaires sur les données. Par exemple, dans les systèmes de détection d'anomalies, des variables présentant une corrélation négative peuvent fournir des indications précieuses sur les écarts par rapport à un comportement normal.

#### **II.2.3.2. Utilité des variables individuelles en contexte de combinaison**

##### **II.2.3.2.1. Variables apparemment inutiles**

Certaines variables peuvent apparaître peu importantes lorsqu'elles sont examinées isolément, mais elles peuvent acquérir une grande importance lorsqu'elles sont combinées avec d'autres. Cela est souvent vrai dans les situations où les variables interagissent de manière complexe. Par exemple, une variable qui semble avoir une faible capacité discriminante seule peut, en interaction avec d'autres variables, créer des combinaisons offrant des signaux puissants pour la classification.

##### **II.2.3.2.2. Approches de sélection**

Pour capturer ces interactions complexes, des méthodes de sélection de caractéristiques, telles que les techniques de filtrage, d'enveloppement (wrapper) et intégrées, ainsi que les algorithmes inspirés de la nature comme les algorithmes génétiques ou l'optimisation par essaim de particules, peuvent être employées. Ces approches permettent d'identifier des sous-ensembles de caractéristiques qui optimisent la performance du modèle en tenant compte des interactions entre les variables.

En résumé, la sélection de caractéristiques ne devrait pas seulement se concentrer sur les mérites individuels des variables, mais aussi sur leurs interactions et la manière dont elles influencent collectivement la performance du modèle. Comprendre la redondance, les corrélations, et les interactions complexes entre les variables est essentiel pour créer des modèles de machine learning efficaces et robustes.

### II.2.4. Typologie des méthodes de sélection de caractéristiques

Les méthodes de sélection de caractéristiques se divisent généralement en plusieurs catégories, chacune possédant ses propres avantages et inconvénients. Ces catégories incluent les méthodes basées sur des filtres, les méthodes d'enveloppement (Wrappers Methods) et des méthodes Intégrées (Embedded Methods).

#### II.2.4.1. Méthodes de filtrage

Les méthodes de filtrage évaluent la pertinence des caractéristiques en fonction des propriétés internes des données, indépendamment de l'algorithme d'apprentissage (Figure II.2). Elles se déroulent en deux étapes : d'abord, les caractéristiques sont classées individuellement selon un critère spécifique, tel que la distance ou l'entropie ; ensuite, les caractéristiques les mieux classées sont sélectionnées en appliquant un seuil, tandis que les autres sont considérées comme non pertinentes. Ces méthodes, rapides et adaptées aux ensembles de données de haute dimension, présentent néanmoins la limite de ne pas tenir compte des interactions entre variables indépendantes, ce qui peut entraîner la sélection de caractéristiques redondantes [138].

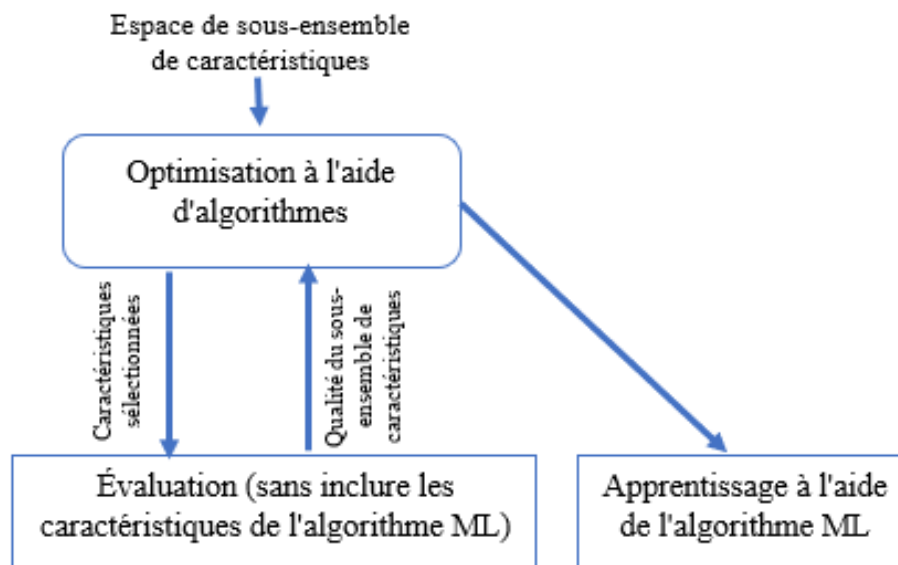


Figure II.2: Approche basée sur les filtres [142].

#### Algorithme de Filtrage

Les algorithmes du modèle de filtrage peuvent être illustrés par un algorithme de filtrage généralisé (comme indiqué dans le Tableau II.1)). Pour un ensemble de données donné  $D$ , l'algorithme commence la recherche à partir d'un sous-ensemble initial  $S_0$  (qui peut être un ensemble vide, un ensemble complet ou un sous-ensemble sélectionné aléatoirement) et explore l'espace des caractéristiques en utilisant une stratégie de recherche spécifique. Chaque sous-ensemble  $S$  généré est évalué par une mesure indépendante  $M$  et comparé avec le meilleur sous-ensemble précédent. S'il est jugé supérieur, il devient le nouveau meilleur sous-ensemble actuel. La recherche se poursuit jusqu'à ce qu'un critère d'arrêt prédéfini (comme décrit dans ci-dessus) soit atteint. L'algorithme retourne alors le dernier sous-ensemble  $S_{best}$  comme résultat final. En variant les stratégies de recherche et les mesures d'évaluation utilisées aux étapes 5 et 6 de l'algorithme, on peut concevoir différents algorithmes individuels au sein du modèle de filtrage. Comme ce modèle applique des critères d'évaluation indépendants sans impliquer



d'algorithme d'exploration de données, il ne subit aucun biais lié à un algorithme d'exploration spécifique et offre une efficacité computationnelle élevée [137].

**Tableau II.1: Un algorithme de filtrage généralisé [138].**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><u>Entrées :</u><br/> <math>D (F_0, F_1, \dots, F_{n-1})</math> : un ensemble de données d'entraînement avec <math>N</math> caractéristiques<br/> <math>S_0</math> : un sous-ensemble à partir duquel commencer la recherche<br/> <math>\delta</math> : un critère d'arrêt</p> <p><u>Sortie :</u><br/> <math>S_{best}</math> : un sous-ensemble optimal</p> <ol style="list-style-type: none"> <li>01. Début</li> <li>02. Initialiser : <math>S_{best} = S_0</math>;</li> <li>03. <math>\gamma_{best} = eval (S_0, D, M)</math> : évaluer <math>S_0</math> par une mesure indépendante <math>M</math></li> <li>04. Faire</li> <li>05. <math>S = générer D</math> : générer un sous-ensemble pour l'évaluation</li> <li>06. <math>\gamma = eval (S, D, M)</math> : évaluer le sous-ensemble actuel <math>S</math> par <math>M</math></li> <li>07. Si <math>\gamma</math> est meilleur que <math>\gamma_{best}</math> alors</li> <li>08. <math>\gamma_{best} = \gamma</math>;</li> <li>09. <math>S_{best} = S</math>;</li> <li>10. Fin si (<math>\delta</math> soit atteint) ;</li> <li>11. Retourner <math>S_{best}</math> ;</li> <li>12. Fin</li> </ol> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Ci-dessous quelques exemples de méthodes de sélection de caractéristiques de type filtre couramment utilisées :

#### 1- Information Mutuelle (Mutual Information) :

L'information mutuelle mesure la quantité d'information que deux variables partagent, en évaluant la dépendance entre une caractéristique et la classe cible. Une valeur élevée d'information mutuelle indique une forte relation entre la caractéristique et la classe cible, suggérant que la caractéristique est utile pour la prédiction [143].

#### 2- L'algorithme ReliefF (ReliefF Algorithm) :

L'algorithme ReliefF est une extension de l'algorithme Relief. Il fonctionne en évaluant l'importance des caractéristiques en tenant compte de la capacité de chaque caractéristique à différencier les instances proches qui appartiennent à des classes différentes. ReliefF prend en charge les interactions entre caractéristiques et les corrélations, et est particulièrement efficace pour les problèmes multiclasse et les ensembles de données bruités. Il attribue un poids à chaque caractéristique basée sur sa capacité à séparer des instances de classes différentes[144].

#### 3- Gain d'Information (Information Gain) :

Le gain d'information mesure la réduction d'incertitude ou d'entropie de la classe cible après avoir observé une caractéristique spécifique. Basée sur la théorie de l'information de Shannon, cette méthode quantifie l'information supplémentaire qu'une caractéristique apporte à la prédiction de la classe cible. Plus le gain d'information est élevé, plus la caractéristique est jugée pertinente pour la tâche de classification. Cette méthode est couramment utilisée pour construire des arbres de décision et sélectionner des caractéristiques dans des contextes de classification [145].

#### 4- Corrélation de Pearson (Pearson correlation) :

Le coefficient de corrélation de Pearson évalue la relation linéaire entre une caractéristique continue et la classe cible (qui peut être numériquement codée). Un coefficient de corrélation élevé (positif ou négatif) indique une forte association, signifiant que la caractéristique est pertinente pour prédire la classe cible. Cette méthode est simple à calculer et est efficace pour éliminer les caractéristiques qui ne montrent aucune corrélation significative avec la classe [146].

### 5- Variance Threshold:

La méthode du seuil de variance élimine les caractéristiques dont la variance (dispersion des valeurs) est inférieure à un seuil prédéfini. Les caractéristiques avec une faible variance ne contiennent pas suffisamment d'information discriminative pour être utiles dans la classification. Cette méthode est simple et rapide à mettre en œuvre, et elle est souvent utilisée comme étape de prétraitement pour réduire la dimensionnalité des données, surtout dans des ensembles de données avec de nombreuses caractéristiques [131].

#### II.2.4.2. Méthodes enveloppantes (wrapper)

Contrairement aux méthodes de filtrage, l'approche wrapper implique la construction d'un nouveau modèle pour chaque sous-ensemble de caractéristiques généré (Figure II.3). Après la création de chaque modèle, la performance de prédiction est utilisée comme critère d'évaluation pour déterminer l'efficacité du sous-ensemble en question. Cette approche permet de prendre en compte les interactions entre les différentes caractéristiques, offrant ainsi une évaluation plus complète et potentiellement plus précise. En général, les méthodes wrapper ont tendance à surpasser les méthodes de filtrage en termes de performance, car elles sont spécifiquement optimisées pour la tâche de classification ou de prédiction. Cependant, elles présentent un inconvénient majeur : elles sont souvent beaucoup plus coûteuses en termes de temps de calcul et de ressources nécessaires, en raison de la complexité inhérente à la construction répétée de modèles. Malgré ces défis, les wrappers demeurent un choix populaire et efficace, surtout lorsque la précision est prioritaire par rapport à l'efficacité computationnelle [138].

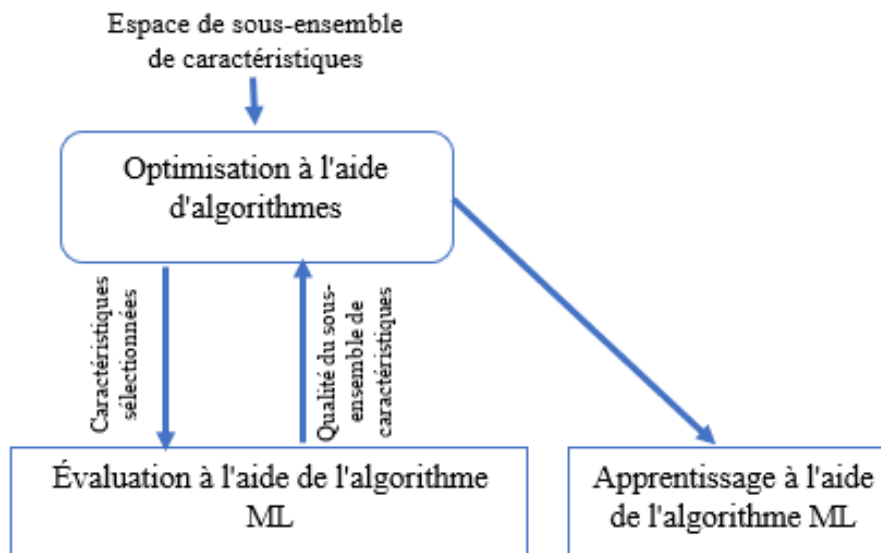


Figure II.3: Approche basée sur les wrappers [142].

### Algorithme Wrapper

Un algorithme wrapper généralisé (Tableau II.2), fonctionne de manière similaire à l'algorithme de filtrage généralisé, mais avec une différence clé : il utilise un algorithme de prédiction de données prédéfini, noté  $A$ , à la place d'une mesure indépendante  $M$  pour évaluer les sous-ensembles. Pour chaque sous-ensemble généré  $S$ , la qualité de celui-ci est évaluée en appliquant l'algorithme de prédiction de données aux données correspondantes et en mesurant la qualité des résultats obtenus. Cela signifie que différents algorithmes de prédiction peuvent conduire à des sélections de caractéristiques distinctes. En modifiant les stratégies de recherche via la fonction generate ( $D$ ) ainsi que les algorithmes de prédiction utilisés ( $A$ ), on peut développer différents algorithmes wrapper. Le modèle wrapper a tendance à offrir des performances supérieures, car les sous-ensembles de caractéristiques identifiés sont spécifiquement optimisés pour l'algorithme de prédiction choisi. Cependant, cette méthode est également plus exigeante en termes de calcul par rapport au modèle de filtrage [137].

**Tableau II.2: Un algorithme wrapper généralisé [138].**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><u>Entrées :</u><br/> <math>D (F_0, F_1, \dots, F_{n-1})</math> : un ensemble de données d'entraînement avec <math>N</math> caractéristiques<br/> <math>S_0</math> : un sous-ensemble à partir duquel commencer la recherche<br/> <math>\delta</math> : un critère d'arrêt<br/> <u>Sortie :</u><br/> <math>S_{best}</math> : un sous-ensemble optimal</p>                                                                                                                                                                                                                                                                                                                                                            |
| <p>01. début<br/> 02. Initialiser : <math>S_{best} = S_0</math>;<br/> 03. <math>\gamma_{best} = eval (S_0, D, A)</math> : évaluer <math>S_0</math> à l'aide d'un algorithme de fouille de données <math>A</math><br/> 04. Faire<br/> 05. <math>S = générer D</math> : générer un sous-ensemble pour l'évaluation<br/> 06. <math>\gamma = eval (S, D, A)</math> : évaluer le sous-ensemble actuel <math>S</math> par <math>A</math><br/> 07. Si <math>\gamma</math> est meilleur que <math>\gamma_{best}</math> alors<br/> 08. <math>\gamma_{best} = \gamma</math>;<br/> 09. <math>S_{best} = S</math>;<br/> 10. Fin si (<math>\delta</math> soit atteint) ;<br/> 11. Retourner <math>S_{best}</math> ;<br/> 12. Fin</p> |

De nombreuses méthodes de type wrapper ont été proposées dans la littérature. Dans ce qui suit, nous présentons les plus connues, conformément à ce qui est indiqué dans l'étude[138].

#### 1- Élimination Récursive de Caractéristiques par Machine à Vecteurs de Support (SVM-RFE) :

Guyon [147] a introduit une méthode de sélection de caractéristiques appelée élimination récursive de caractéristiques basée sur une machine à vecteurs de support (SVM-RFE). Cette méthode commence par utiliser toutes les caractéristiques disponibles, puis elle les élimine de manière itérative sans possibilité de revenir en arrière. Cette stratégie gloutonne, descendante, calcule le vecteur normal d'un hyperplan de décision créé par un classificateur SVM linéaire. Les projections de ce vecteur normal sur le système de coordonnées sont ensuite utilisées pour évaluer le classement des caractéristiques individuelles. Les caractéristiques avec les classements les plus élevés sont conservées, tandis que celles avec les classements les plus faibles sont éliminées. Ce processus est répété jusqu'à ce que le nombre désiré de caractéristiques soit atteint ou que la performance ne puisse plus être améliorée.

#### 2- Sélection Séquentielle Avant (SFS)

La méthode SFS est une stratégie de recherche gloutonne[148]. Elle commence par un ensemble vide et ajoute des caractéristiques de manière itérative et opportuniste jusqu'à ce que les performances ne puissent plus être améliorées. Cette approche fonctionne bien lorsque le nombre de caractéristiques représentatives (sous-ensemble optimal) est minimal. Cependant, l'inconvénient de cette technique est que, une fois qu'une caractéristique est ajoutée au sous-ensemble, elle ne peut plus être retirée, ce qui peut mener à des résultats sous-optimaux.

#### II.2.4.3. Méthodes Intégrées (Embedded Methods)

À la différence de la stratégie wrapper, qui repose sur une recherche heuristique guidée par les résultats du classificateur, la méthode intégrée se positionne comme un compromis entre les approches de filtrage et de wrapper. Elle combine la sélection de caractéristiques et l'optimisation du classificateur pendant le processus d'apprentissage lui-même. En tant qu'approche intermédiaire, la méthode intégrée identifie les caractéristiques pertinentes au fur et à mesure de l'apprentissage, en s'appuyant sur les critères d'évaluation du classificateur. Cette approche offre des avantages en termes de réduction des coûts de calcul par rapport aux wrappers. Les techniques de régularisation figurent parmi les méthodes intégrées les plus fréquemment employées [138].

##### *Algorithme de sélection intégré*

L'algorithme de sélection de caractéristiques intégré, appelé méthode embedded (Tableau II.3), commence par évaluer un sous-ensemble initial de caractéristiques, noté  $S_0$ , à l'aide d'un modèle d'apprentissage  $A$ , qui intègre une technique de régularisation, comme la régression Lasso ou Ridge. Ces techniques de régularisation aident à contrôler la complexité du modèle en réduisant ou en éliminant les caractéristiques non pertinentes. À chaque itération, l'algorithme génère de nouveaux sous-ensembles en ajustant les poids des caractéristiques ou en modifiant les paramètres de régularisation. Ces sous-ensembles sont ensuite évalués sur leur capacité à améliorer la performance du modèle tout en pénalisant les caractéristiques superflues. Le processus se poursuit jusqu'à ce qu'un critère d'arrêt prédéfini soit atteint, par exemple un nombre fixe d'itérations, ou lorsqu'aucune amélioration notable de la performance n'est observée lors de l'ajout ou de la suppression de caractéristiques. Enfin, l'algorithme identifie et retourne le sous-ensemble de caractéristiques qui a permis d'obtenir les meilleures performances durant l'entraînement.

**Tableau II.3: Un algorithme généralisé de type intégré (embedded).**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><u>Entrées :</u><br/> <math>D (F_0, F_1, \dots, F_{n-1})</math> : un ensemble de données d'entraînement avec <math>N</math> caractéristiques<br/> <math>S_0</math> : un sous-ensemble à partir duquel commencer la recherche<br/> <math>\delta</math> : un critère d'arrêt<br/> <math>\lambda</math> : un paramètre de régularisation (pour contrôler la pénalisation des caractéristiques)</p> <p><u>Sortie :</u><br/> <math>S_{best}</math> : un sous-ensemble optimal</p> <ol style="list-style-type: none"> <li>01. Début</li> <li>02. Initialiser : <math>S_{best} = S_0</math>;</li> <li>03. <math>\gamma_{best} = eval (S_0, D, A, \lambda)</math> : évaluer <math>S_0</math> à l'aide d'un algorithme d'apprentissage <math>A</math> avec un terme de régularisation <math>\lambda</math></li> <li>04. Faire</li> <li>05. Pour chaque itération :</li> <li>06. <math>S = générer D</math> : générer un nouveau sous-ensemble de caractéristiques pour l'évaluation (souvent en ajustant les poids des caractéristiques ou en modifiant la régularisation)</li> <li>07. <math>\gamma = eval (S, D, A, \lambda)</math> : évaluer le sous-ensemble actuel <math>S</math> en utilisant <math>A</math> et <math>\lambda</math></li> <li>08. Si <math>\gamma</math> est meilleur que <math>\gamma_{best}</math> alors</li> <li>09. <math>\gamma_{best} = \gamma</math>;</li> <li>10. <math>S_{best} = S</math>;</li> </ol> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

11. Fin si ( $\delta$  soit atteint) ;
12. Retourner  $S_{best}$  ;
13. Fin

### 1- LASSO (Régularisation L1)

LASSO, développé par Robert Tibshirani en 1996 [149], est une méthode de sélection de caractéristiques qui utilise la régularisation pour limiter les valeurs absolues des coefficients des paramètres d'un modèle d'apprentissage automatique. Cette contrainte supplémentaire permet de réduire certains coefficients à zéro, excluant ainsi les caractéristiques non pertinentes. Un paramètre de régularisation détermine l'intensité de cette réduction : plus il est élevé, plus de coefficients seront annulés, ce qui diminue la complexité du modèle. En procédant ainsi, LASSO contribue à limiter la variance, à éviter le surapprentissage et à améliorer la capacité de généralisation du modèle. De plus, il favorise l'interprétation du modèle en éliminant les caractéristiques non essentielles.

### 2- RIDGE (Régularisation L2)

La régularisation RIDGE, ou régularisation L2, consiste à ajouter une pénalité proportionnelle au carré des coefficients des paramètres du modèle dans la fonction de coût. Contrairement à la méthode LASSO, qui annule certains coefficients pour sélectionner un petit sous-ensemble de caractéristiques, RIDGE diminue les coefficients vers zéro sans les rendre exactement nuls. L'avantage principal de RIDGE par rapport à LASSO est sa capacité à conserver des coefficients non nuls, ce qui permet une interprétation plus nuancée des caractéristiques du modèle, même si elles ne sont pas fortement significatives.

Le (Tableau II.4) offre un récapitulatif détaillé des propriétés des trois types de méthodes de sélection de caractéristiques, en soulignant leurs principaux avantages et inconvénients, comme le montre la revue[138].

**Tableau II.4: Avantages et inconvénients des approches de sélection de caractéristiques** [138].

|                           | Avantages                                                                                                                                                                                                                                   | Inconvénients                                                                                                                                                                                                                                 |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Méthodes Filtres</b>   | <ul style="list-style-type: none"> <li>• Indépendants du modèle d'apprentissage</li> <li>• Exécution rapide</li> <li>• Adaptés aux données de haute dimension</li> <li>• Bonne généralisabilité</li> </ul>                                  | <ul style="list-style-type: none"> <li>• Les interactions entre les caractéristiques sont ignorées</li> <li>• Incapacité à gérer le problème de redondance</li> <li>• Aucune interaction avec l'algorithme d'apprentissage</li> </ul>         |
| <b>Méthodes Wrappers</b>  | <ul style="list-style-type: none"> <li>• Meilleure performance atteignable</li> <li>• Prend en compte les interactions entre les caractéristiques</li> <li>• Identifie les interactions de caractéristiques de plus haut niveau</li> </ul>  | <ul style="list-style-type: none"> <li>• Coûteux en termes de temps d'exécution</li> <li>• Sujet à un surajustement (overfitting)</li> <li>• L'algorithme d'apprentissage est construit à partir de zéro pour chaque sous-ensemble</li> </ul> |
| <b>Méthodes intégrées</b> | <ul style="list-style-type: none"> <li>• Plus rapide que les wrappers,</li> <li>• Précis</li> <li>• Prend en compte les interactions entre les caractéristiques</li> <li>• Identifie les dépendances entre les caractéristiques.</li> </ul> | <ul style="list-style-type: none"> <li>• Spécifique à l'algorithme d'apprentissage</li> <li>• Sélection dépendante du classificateur.</li> </ul>                                                                                              |

**Remarque :** Il existe des approches hybrides de sélection de caractéristiques qui combinent les avantages des méthodes de filtrage et des méthodes wrapper pour optimiser la performance de la sélection tout en atténuant les inconvénients associés à chaque approche. Ces méthodes hybrides commencent généralement par une phase de pré-filtrage, où des mesures indépendantes permettent de réduire l'ensemble initial de caractéristiques, suivie d'une validation par un algorithme d'apprentissage pour affiner cette sélection. Elles sont particulièrement efficaces pour gérer des jeux de données volumineux tout en prenant en compte les interactions entre les caractéristiques. Les méthodes hybrides offrent un bon équilibre entre rapidité d'exécution et précision du modèle, réduisant ainsi le risque de sur-ajustement et les coûts computationnels, et démontrent des performances supérieures par rapport aux approches de filtrage ou wrapper utilisées seules.

### **II.2.5. Méthodes d'optimisation en sélection de caractéristiques**

L'optimisation en sélection de caractéristiques est une étape cruciale dans de nombreux systèmes de machine learning. La sélection de caractéristiques consiste à identifier les sous-ensembles de données les plus pertinents qui améliorent la performance du modèle tout en réduisant sa complexité. L'optimisation intervient pour affiner ce processus, en cherchant à maximiser l'efficacité du modèle selon un ou plusieurs objectifs. Elle peut être abordée sous deux perspectives principales : l'optimisation mono-objectif, qui se concentre sur un seul critère à la fois, et l'optimisation multi-objectifs, qui vise à trouver un équilibre entre plusieurs critères souvent conflictuels [131]. L'optimisation en sélection de caractéristiques joue ainsi un rôle déterminant dans l'amélioration des performances globales des systèmes d'apprentissage automatique.

Les méthodes d'optimisation pour la sélection de caractéristiques se classent généralement en deux catégories : les méthodes traditionnelles d'optimisation et les méthodes inspirées de la nature.

#### **II.2.5.1. Méthodes traditionnelles d'optimisation**

Les approches classiques comprennent des algorithmes de recherche tels que les méthodes exhaustives, les recherches séquentielles (avant et arrière), ainsi que les algorithmes heuristiques. Cependant, ces méthodes rencontrent souvent des limitations en raison de la taille importante des espaces de recherche dans les problèmes à haute dimension [150].

##### **II.2.5.1.1. Recherche exhaustive**

Cette approche évalue tous les sous-ensembles possibles de caractéristiques pour identifier la meilleure combinaison. Bien qu'elle assure un résultat optimal, elle devient rapidement impraticable avec un grand nombre de caractéristiques en raison de son coût computationnel élevé [150].

##### **II.2.5.1.2. Méthodes séquentielles**

Ces méthodes ajoutent ou suppriment progressivement des caractéristiques dans une séquence déterminée. La sélection avant (Sequential Forward Selection, SFS) commence avec un sous-ensemble vide et ajoute progressivement des caractéristiques, tandis que la sélection arrière (Sequential Backward Selection, SBS) fonctionne en supprimant des caractéristiques d'un ensemble complet. Ces approches sont plus efficaces que la recherche exhaustive, mais elles peuvent encore souffrir du phénomène de "nesting", où des décisions prises au début ne peuvent pas être annulées plus tard [131] [150].

##### **II.2.5.2. Méthodes inspirées de la nature**

Pour surmonter les limites des méthodes traditionnelles, les algorithmes d'optimisation inspirés de la nature ont gagné en popularité pour traiter des problèmes complexes de sélection de caractéristiques. Ces techniques comprennent des algorithmes évolutionnaires et des métaheuristiques, qui imitent des processus naturels pour explorer l'espace des solutions de façon plus efficace et économique [150].

#### **II.2.5.2.1. Algorithmes génétiques (Genetic Algorithms, GA)**

Les algorithmes génétiques [151], sont inspirés par le processus de sélection naturelle et l'évolution biologique. Ils commencent par une population de solutions possibles (sous-ensembles de caractéristiques), et, à chaque génération, des opérateurs comme la mutation et le croisement sont appliqués pour produire de nouvelles solutions. Les meilleures solutions sont sélectionnées pour former la prochaine génération. Ce processus itératif conduit à une amélioration progressive des solutions.

#### **II.2.5.2.2. Optimisation par essaim particulaire (Particle Swarm Optimization, PSO)**

L'optimisation par essaim de particules (PSO) [152], simule le comportement collectif des essaims d'oiseaux ou des bancs de poissons. Chaque particule représente une solution candidate, et elle se déplace dans l'espace de recherche en fonction de sa meilleure position connue et de celle de ses voisins. L'objectif est d'atteindre la meilleure position (solution optimale) à travers un mouvement coordonné.

#### **II.2.5.2.3. Optimisation par colonies de fourmis (Ant Colony Optimization, ACO)**

L'optimisation par colonie de fourmis (ACO) [153], est inspirée par le comportement des fourmis lorsqu'elles recherchent de la nourriture. Dans le cadre de la sélection de caractéristiques, chaque fourmi explore un chemin (un sous-ensemble de caractéristiques) et laisse des phéromones sur les chemins menant aux meilleures solutions. À chaque itération, les fourmis sont plus susceptibles de suivre les chemins les plus prometteurs, conduisant ainsi à une convergence vers un sous-ensemble optimal de caractéristiques.

### **II.3. Pourquoi la sélection de caractéristiques est cruciale pour la REF**

La sélection de caractéristiques joue un rôle essentiel dans la reconnaissance des expressions faciales pour plusieurs raisons, notamment l'amélioration des performances des modèles, la réduction de la complexité computationnelle, et l'augmentation de la robustesse et de la généralisation des modèles[154] [155].

- Amélioration des Performances des Modèles

Dans la REF, les images du visage peuvent contenir beaucoup d'informations redondantes et non pertinentes. La sélection de caractéristiques permet de focaliser l'attention sur les traits du visage les plus pertinents pour identifier les diverses expressions, comme les mouvements des sourcils, des yeux ou de la bouche. En éliminant les éléments inutiles et en mettant en avant les caractéristiques clés, les modèles peuvent apprendre plus efficacement, ce qui se traduit par une meilleure précision et une capacité accrue à détecter les expressions, même subtiles, telles que la surprise ou le sourire.

- Réduction de la Complexité Computationnelle

L'analyse d'images ou de séquences vidéo haute résolution, souvent utilisée en REF, peut être très exigeante en termes de temps de calcul et de ressources matérielles. La sélection des caractéristiques contribue à réduire la dimensionnalité des données d'entrée, allégeant ainsi la charge computationnelle. Cela est particulièrement avantageux pour les systèmes de reconnaissance des expressions faciales utilisés en temps réel dans des applications pratiques, telles que les interfaces homme-machine ou les systèmes de surveillance.

- Renforcement de la Robustesse et de la Généralisation

Pour les systèmes de REF, la robustesse est une qualité essentielle, surtout lorsqu'ils sont utilisés dans des environnements où des facteurs comme l'éclairage, l'angle de vue, ou les bruits de fond varient fortement. En sélectionnant des caractéristiques peu sensibles à ces variations, les modèles deviennent plus robustes et généralisent mieux sur des données non vues auparavant, ce qui les rend plus efficaces pour traiter des expressions faciales naturelles et spontanées.

- Meilleure Identification des Caractéristiques Pertinentes

La sélection de caractéristiques offre également une meilleure compréhension des aspects clés du visage qui jouent un rôle dans la reconnaissance des expressions. Cela non seulement améliore les performances des modèles, mais aide aussi à mieux comprendre les mécanismes sous-jacents aux expressions faciales, ce qui est crucial pour concevoir des systèmes d'interaction homme-machine plus naturels.

## **II.4. Application de la sélection de caractéristiques à la reconnaissance des expressions faciales**

L'application de la sélection de caractéristiques dans la reconnaissance des expressions faciales est cruciale pour améliorer les performances des systèmes, réduire la complexité des modèles, et renforcer leur robustesse. En utilisant des techniques adaptées, il est possible d'optimiser la détection des émotions en se concentrant uniquement sur les caractéristiques les plus pertinentes. Cette section aborde les différentes approches et méthodes appliquées à la reconnaissance des expressions faciales, en mettant en lumière leur rôle dans l'amélioration de la précision et de l'efficacité des systèmes.

### **II.4.1. Sélection de caractéristiques basées sur des méthodes filtres**

La sélection de caractéristiques basées sur les filtres est largement utilisée dans le domaine de la reconnaissance des expressions faciales en raison de sa simplicité et de son efficacité. Cette méthode évalue chaque caractéristique individuellement à l'aide de critères statistiques ou de mesures de pertinence, sans nécessiter de modèle d'apprentissage supervisé. Elle permet d'identifier les caractéristiques les plus pertinentes tout en éliminant celles qui sont redondantes ou peu utiles, contribuant ainsi à réduire la dimensionnalité et à améliorer les performances du modèle. Un large éventail de techniques de sélection de caractéristiques basées sur les filtres a été développé et exploré.

Lajevardi et Hussain, dans [156], ont étudié différentes méthodes de sélection de caractéristiques dans le cadre d'un système REF. Après l'extraction des caractéristiques à l'aide de techniques telles que les filtres de Gabor, l'opérateur LBP, et la méthode HLACLF, ils ont appliqué des approches de sélection de caractéristiques basées sur les méthodes filter et wrapper, en mettant l'accent sur l'algorithme mRMR (Minimum Redundancy Maximum Relevance), qui utilise l'information mutuelle pour sélectionner les sous-ensembles de caractéristiques les plus discriminants. Les résultats obtenus montrent que l'approche mRMR offre des avantages significatifs en termes de simplicité de calcul et d'invariance aux transformations des données, tout en améliorant la précision du système. Combinée à l'extraction de caractéristiques via la méthode HLACLF, cette sélection de caractéristiques a permis d'atteindre une précision plus élevée dans la classification des expressions faciales, surpassant les autres méthodes testées. Les auteurs concluent que la sélection basée sur l'information mutuelle (MI) est particulièrement efficace pour améliorer les performances globales du système de reconnaissance des expressions faciales.

Lajevardi et Hussain, dans un autre travail [157], ont appliqué l'algorithme de sélection de caractéristiques par information mutuelle (MIFS) pour un système REF. En effet, après avoir utilisé des filtres Log-Gabor pour extraire des caractéristiques à partir des images du visage entier et des régions faciales, les caractéristiques les plus informatives sont classifiées à l'aide du classificateur NB afin de reconnaître six expressions : la colère, la peur, le dégoût, la surprise, la tristesse et la joie. Les résultats expérimentaux obtenus à l'aide des bases de données CK et JAFFE ont montré l'efficacité de la méthode de sélection de caractéristiques proposée pour améliorer les performances globales du système REF.

Abdulrazaq et al. [158], ont utilisé la méthode de sélection de caractéristiques Relief-F pour sélectionner les caractéristiques les mieux classées pour chaque expression faciale. Dans cette méthode, les caractéristiques sont classées dans un ordre séquentiel en fonction de leurs valeurs de variance. Ensuite, les caractéristiques sélectionnées sont utilisées pour entraîner et tester six classificateurs, à savoir le



Perceptron multicouche (MLP), la Forêt aléatoire (RF), l'Arbre de décision (DT), la Machine à vecteurs de support (SVM), le K-plus proche voisin (KNN) et la Fonction de base radiale (RBF). Les résultats ont montré que le classificateur KNN est le plus précis parmi les autres avec un taux de précision total de 94,93 % en utilisant la base de données Cohn-Kanade (CK+).

Dino and Abdulrazzaq [159], Ils ont comparé différentes méthodes de sélection de caractéristiques pour la reconnaissance des expressions faciales en utilisant des algorithmes de classification multiples tels que KNN, MLP, DT (J48) et NB. L'objectif principal de cette comparaison, qui est de déterminer le meilleur classificateur capable d'atteindre un taux de précision acceptable, a été réalisé sur la base de trois techniques de sélection de caractéristiques, à savoir CFS, le Gain Ratio (GR) et IG. Les résultats expérimentaux, basés sur la base de données CK+, ont montré que le classificateur KNN était le meilleur, avec une précision de 91 % en utilisant seulement 30 caractéristiques.

Wang et al. [160], ont proposé un modèle de fusion de caractéristiques basé sur la sélection multiple de caractéristiques, réalisé par la mesure du coefficient de corrélation RV, afin d'augmenter les performances du système de reconnaissance des expressions faciales. Quatre types de caractéristiques, à savoir les Local Binary Patterns (LBP), la Pyramide d'Histogrammes de Gradients Orientés (PHOG), les Histogrammes de Gradients Orientés (HOG) et les caractéristiques de Gabor, sont sélectionnés pour la fusion. De plus, les sous-espaces de l'Analyse Canonique des Corrélations (CCA) et de l'Analyse en Composantes Principales (PCA) sont utilisés pour fusionner les caractéristiques sélectionnées. Les résultats expérimentaux, basés sur les bases de données JAFFE et CK (en utilisant un classificateur SVM), ont montré que le modèle proposé améliore les performances et élimine la redondance des caractéristiques.

Li et al. [161], ont proposé une méthode innovante pour extraire et sélectionner des caractéristiques saillantes à partir d'images statiques. Ils utilisent une combinaison de filtres fixes et de filtres adaptatifs non linéaires en cascade, inspirés des mécanismes biologiques d'inhibition. Les filtres fixes extraient des caractéristiques primitives (comme les contours), tandis que les filtres adaptatifs sont entraînés pour identifier des caractéristiques plus complexes, lesquelles sont ensuite classifiées par des machines à vecteurs de support (SVM). Une approche de sélection des caractéristiques, basée sur la réduction de l'information mutuelle, est appliquée pour optimiser les performances du système. Les résultats, obtenus sur la base de données JAFFE, montrent un taux de classification de 96,7 % en utilisant seulement deux tiers des caractéristiques extraites, surpassant plusieurs méthodes existantes.

Siddiqi et al. [162], ont proposé une méthode de sélection de caractéristiques basée sur la normalisation de l'information mutuelle. Cette approche améliore la méthode max-relevance min-redundancy (mRMR) en équilibrant la pertinence et la redondance des caractéristiques. Pour l'extraction des caractéristiques, ils utilisent la transformée en courbelettes, ce qui permet de conserver les informations relatives aux lignes, courbes et bords dans les images faciales. Ensuite, l'analyse discriminante linéaire (LDA) est appliquée pour réduire la dimension de l'espace des caractéristiques, avant que les expressions ne soient classées à l'aide d'un modèle de Markov caché (HMM). Le système REF proposé, a été validé sur quatre bases de données publiques : JAFFE, Yale B, Cohen-Kanade, et USTC-NVIE, avec une validation croisée à 10 volets. Le système a surpassé les méthodes existantes, atteignant un taux de reconnaissance moyen pondéré de 99 % sur tous les ensembles de données. L'approche se distingue par son efficacité en termes de reconnaissance, tout en étant moins coûteuse en termes de calcul.

Saha et al. [163], ont proposé une technique de sélection de caractéristiques appelée Supervised Filter Harmony Search Algorithm (SFHSA), basée sur la similarité cosinus et la méthode minimal-redundancy maximal-relevance (mRMR), pour améliorer la reconnaissance des émotions faciales (REF). La similarité cosinus permet d'éliminer les caractéristiques redondantes, tandis que mRMR, via le coefficient de corrélation de Pearson (PCC), identifie les sous-ensembles de caractéristiques optimales. Les caractéristiques ayant une faible corrélation avec d'autres et une forte corrélation avec les classes d'expressions faciales sont favorisées. Cette approche a été évaluée sur deux ensembles de données de

référence en REF : la base de données des visages Radboud (RaFD) et la base de données des expressions faciales des femmes japonaises (JAFFE). Cinq descripteurs de caractéristiques, incluant le motif binaire local uniforme (uLBP), le motif binaire local du voisinage horizontal-vertical (hvnLBP), les filtres de Gabor, l'histogramme des gradients orientés (HOG) et le pyramidal HOG (PHOG), ont été utilisés pour l'extraction des caractéristiques. Les résultats montrent que la technique SFHSA optimise efficacement les vecteurs de caractéristiques, réduisant la redondance et améliorant les performances de classification par rapport aux méthodes de sélection de caractéristiques existantes, telles que SA (Simulated Annealing), GA (Genetic Algorithms) et MA (Memetic Algorithms).

Dong et Sayama [164], ont proposé une méthode de sélection de caractéristiques visant à réduire le nombre de repères faciaux nécessaires. Ils analysent les corrélations d'information entre les mouvements des repères faciaux en mesurant l'information mutuelle, puis utilisent un clustering hiérarchique agglomératif (HAC) pour regrouper les repères selon des motifs de mouvement similaires. Pour chaque cluster, ils sélectionnent une caractéristique représentative, ce qui permet de simplifier le modèle tout en conservant les informations essentielles. Cette sélection de caractéristiques est essentielle pour diminuer la charge computationnelle, permettant au classificateur basé sur un support vectoriel (SVM) d'atteindre une précision comparable à celle utilisant l'ensemble de données complet, tout en réduisant le temps de calcul de 63,5 %. Les résultats montrent que la sélection efficace des caractéristiques contribue à une performance robuste tout en étant adaptée aux dispositifs légers.

Yurtkan et Demirel [165], ont proposé une méthode de sélection de caractéristiques pour la reconnaissance des expressions faciales en utilisant des points géométriques tridimensionnels (3D) du visage. Chaque expression (colère, dégoût, peur, bonheur, tristesse et surprise) bénéficie d'une sélection de caractéristiques distincte basée sur l'entropie. Cette approche permet d'extraire les points les plus informatifs et discriminants pour chaque émotion. Testée sur la base de données BU-3DFE, la méthode utilisant un classificateur SVM dépasse les techniques existantes en termes de taux de reconnaissance, démontrant une précision accrue dans la classification des six expressions de base.

Khan et al. [166], ont proposé une nouvelle stratégie de sélection de caractéristiques pour la reconnaissance des expressions faciales, en utilisant la décomposition en valeurs singulières (SVD) et le co-clustering afin d'identifier les régions faciales les plus saillantes et discriminantes. Cette approche, qui est la première à appliquer explicitement le co-clustering dans ce domaine, utilise des filtres de Gabor pour extraire des caractéristiques locales, puis sélectionne les plus pertinentes en fonction de leur appartenance à des co-clusters basés sur les classes d'expressions. Les expériences réalisées sur les bases de données CK+, JAFFE et MMI montrent que cette méthode localise efficacement les régions du visage autour des yeux, de la bouche et du nez, réduisant ainsi la dimensionnalité des données tout en améliorant la précision de reconnaissance. Les résultats montrent des taux de reconnaissance élevés, atteignant 96,05 % avec CK+, en utilisant seulement 20 % des caractéristiques. L'algorithme réduit également la complexité computationnelle et surpasse d'autres méthodes de sélection de caractéristiques.

#### **II.4.2. Sélection de caractéristiques basées sur des méthodes wrapper**

Ewees et al. [167], ont proposé une méthode d'optimisation, appelée Algorithme Sinus-Cosinus (SCA), pour la sélection de caractéristiques. Cette méthode, qui est un algorithme métaheuristique utilisant les formes mathématiques des fonctions sinus et cosinus pour résoudre des problèmes d'optimisation, a été utilisée afin d'améliorer les performances d'un système en temps réel destiné à reconnaître les expressions faciales des étudiants. La performance de cette méthode a été évaluée sur une base de données des expressions faciales des étudiants, en utilisant divers classificateurs, notamment KNN, RF et SVM. Les résultats expérimentaux ont montré que le classificateur KNN a obtenu les meilleures performances, avec une réduction remarquable du nombre de caractéristiques.

Lajevardi et Hussain [168], ont adopté un processus d'optimisation globale basé sur l'algorithme génétique (GA) pour sélectionner les caractéristiques optimales, générées à l'aide de filtres log-Gabor.

Ce processus a permis de réduire le nombre de caractéristiques de 36 000 caractéristiques/image à 100 caractéristiques/image et a atteint une précision de 78 % en utilisant la base de données CK avec le classificateur Naïf Bayésien (NB).

Siddiqi et al. [169], ont développé une approche wrapper basée sur la forêt aléatoire pour la sélection de caractéristiques. La performance de cette méthode a été évaluée sur la base des caractéristiques extraites à l'aide d'une grande banque de filtres Gabor, dans laquelle un classificateur Support Vector Machine (SVM) multi-classes a été utilisé pour la reconnaissance des expressions faciales. Les résultats obtenus à partir de la base de données CK+ ont montré que cette méthode améliore le taux de reconnaissance pour les expressions de colère, de tristesse et de neutralité.

Soyel et al. [154], ont proposé une méthodologie de sélection de caractéristiques pour la reconnaissance des expressions faciales basée sur l'algorithme génétique de tri non dominé II (NSGA-II). Cette méthodologie est l'un des derniers algorithmes génétiques développés pour résoudre des problèmes avec une grande précision. Les résultats obtenus, en utilisant la base de données d'expressions faciales 3D BU-3DFE, ont démontré l'efficacité et la flexibilité de la méthodologie basée sur NSGA-II.

Mlakar et al. [170], ont introduit une stratégie de sélection de caractéristiques efficace basée sur l'approche wrapper pour la reconnaissance des expressions faciales, utilisant un algorithme d'évolution différentielle multi-objectif modifié. Dans ce cadre, deux stratégies de sélection ont été développées et validées sur des bases de données d'évaluation couramment utilisées. Les résultats obtenus ont montré des taux de reconnaissance élevés avec des ensembles de caractéristiques considérablement réduits, démontrant que cette approche est très prometteuse pour améliorer les performances des systèmes de REF et réduire les coûts computationnels dans diverses applications.

Paharia et al. [171], ont proposé une méthode de sélection de caractéristiques basée sur une approche wrapper. Cette méthode, nommée MOCBGWO amélioré, consiste en une amélioration de l'optimiseur binaire de loup gris compétitif. Les caractéristiques sélectionnées sont ensuite évaluées sur les ensembles de données CK+ et JAFFE en utilisant des classificateurs SVM et KNN. Les résultats obtenus ont indiqué que la méthode MOCBGWO proposée améliore considérablement la précision de reconnaissance tout en réduisant la taille du vecteur de caractéristiques, soulignant ainsi son efficacité dans les systèmes de reconnaissance des expressions faciales.

Al-Qablan et al. [172], ont proposé une approche de sélection de caractéristiques hybride combinant l'optimiseur binaire de loup gris (BGWO) et la montée de colline adaptative  $\beta$  ( $A\beta CH$ ) pour améliorer l'approche Wrapper. Dans cette approche, le classificateur KNN est utilisé pour évaluer les performances de la méthode hybride proposée. Les résultats obtenus, basés sur 18 ensembles de données de référence UCI pour la sélection de caractéristiques, ont montré que cette approche surpasse, en termes de précision et de taille des caractéristiques sélectionnées, diverses méthodes de sélection de caractéristiques à la pointe de la technologie, démontrant ainsi des performances prometteuses pour diverses applications en apprentissage automatique et en analyse de données.

Perez-Gomez et al. [173], ont étudié la sélection des caractéristiques géométriques optimales pour classer les expressions faciales. En effet, en exploitant des techniques inspirées par le Système de Codage des Actions Faciales (FACS) et le MPEG-4, un premier ensemble de 89 caractéristiques, dérivées de la géométrie des repères faciaux 2D et 3D, a été proposé. Deux méthodes de sélection de caractéristiques, à savoir l'Analyse en Composantes Principales (PCA) et l'algorithme génétique (GA) qui fonctionne comme une méthode de sélection de caractéristiques Wrapper (en utilisant une machine à vecteurs de support), sont employées pour réduire l'ensemble des caractéristiques. Les résultats obtenus, sur les ensembles de données Bosphorus et UIBVFED, ont démontré que cette méthode a une précision médiane de 86,62 % et 93,92 %, respectivement, ce qui souligne l'importance de la sélection de caractéristiques pour obtenir une haute précision avec un faible coût computationnel.

Dornaika et al. [174], ont proposé une méthode pour la reconnaissance dynamique des expressions faciales dans des vidéos en utilisant la sélection de sous-ensembles de caractéristiques. Leur approche applique la sélection de caractéristiques aux signatures temporelles représentant les dynamiques faciales, ce qui améliore la reconnaissance par rapport aux méthodes basées sur des caractéristiques brutes. Ils utilisent plusieurs classificateurs (SVM, KNN, Naïve Bayes, Réseaux Bayésiens, et Arbres de décision) combinés à la technique "wrapper" pour optimiser à la fois les caractéristiques et les paramètres du classificateur. Les résultats expérimentaux montrent une amélioration des performances en termes de précision et de complexité, prouvant que cette approche élimine les caractéristiques redondantes et non pertinentes.

Danisman et al. [175], ont proposé une méthode de sélection de caractéristiques basée sur la création d'un masque d'émotion, en utilisant une recherche exhaustive pour identifier les meilleures positions et tailles de fenêtres de caractéristiques dans les images faciales. Cette approche, axée sur l'élimination des pixels non pertinents, applique une méthode de sélection par wrapper. Grâce à la technique de suppression de caractéristiques non pertinentes, les fenêtres de caractéristiques optimales sont identifiées, ce qui améliore la reconnaissance des expressions faciales. Bien que la phase de création du masque soit coûteuse en temps, la reconnaissance avec ce masque fonctionne en temps réel.

Vignolo et al. [176], ont proposé une méthode de sélection de caractéristiques basée sur un algorithme génétique multi-objectif pour les tâches de reconnaissance faciale. L'objectif est de sélectionner le sous-ensemble de caractéristiques le plus pertinent, tout en minimisant la taille de ce sous-ensemble et en maximisant sa capacité discriminante. La méthode explore différentes combinaisons possibles de caractéristiques afin d'optimiser la performance de classification tout en réduisant la complexité. Les résultats expérimentaux montrent que cette approche améliore la précision de classification par rapport à d'autres méthodes, tout en réduisant la dimensionnalité des données, ce qui permet d'éviter le sur-apprentissage et de simplifier la tâche de classification.

#### **II.4.3. Sélection de caractéristiques basées sur des méthodes embedded**

Richhariya et al. [177], ont proposé un Machine à Vecteurs de Support Jumeaux Universum Itérative (IUTWSVM) utilisant la méthode de Newton pour la classification des émotions multiclasse par le biais de l'analyse des expressions faciales. Cette approche innovante, qui intègre une méthode itérative, vise à améliorer la sélection des caractéristiques et la performance de classification tout en réduisant les coûts computationnels. Pour améliorer la sélection des données universelles et garantir qu'elles contribuent efficacement à la tâche de classification, les auteurs ont introduit un nouveau système basé sur l'entropie d'information. Les résultats expérimentaux basés sur des bases de données de référence ont démontré que l'approche IUTWSVM surpasse les algorithmes existants, tant en termes de précision de classification que d'efficacité d'apprentissage. Cela est vrai pour les tâches de reconnaissance des émotions faciales, qu'elles soient binaires ou multiclasse.

Talaat et al. [178], ont développé un système d'identification des émotions en temps réel pour détecter les troubles du spectre autistique chez les enfants autistes. Le système proposé intègre une méthode de sélection de caractéristiques basée sur un autoencodeur, mise en œuvre dans une architecture de Réseau de Neurones Convolutifs Profonds (DCNN). Les résultats obtenus ont montré que cet autoencodeur a atteint une précision remarquable de 95,23 % avec le modèle Xception. De plus, il améliore la détection des émotions en temps réel, fournissant ainsi une solution technologique avancée adaptable sur divers appareils intelligents.

Mistry et al. [179], ont proposé un système REF qui met l'accent sur l'optimisation des caractéristiques à l'aide d'une variante améliorée de l'algorithme d'optimisation par essaim particulaire (PSO). L'extraction initiale des caractéristiques est réalisée via un opérateur de motifs binaires locaux modifiés (LBP), tandis que la sélection et l'optimisation des caractéristiques se font à l'aide d'un algorithme PSO intégré à un micro-algorithme génétique (mGA-embedded PSO). Cette approche permet de sélectionner

les caractéristiques les plus discriminantes et significatives pour différencier sept expressions faciales. Le processus de sélection des caractéristiques optimise la diversité des solutions grâce à une stratégie de mémoire non remplaçable et une sous-population secondaire. Cette méthode améliore la recherche locale et globale, tout en évitant la convergence prématurée typique du PSO classique. En optimisant les caractéristiques faciales, le système réduit la complexité tout en maintenant une grande précision. Les résultats expérimentaux montrent que l'algorithme mGA-embedded PSO surpasse les autres méthodes d'optimisation, y compris le PSO classique et l'algorithme génétique (GA), offrant des gains significatifs en termes de sélection de caractéristiques optimales, et permettant une précision de 100 % sur la base CK+ et de 94,66 % sur la base MMI.

Ghazouani [180], a proposé un cadre basé sur la programmation génétique (GP) pour la reconnaissance des expressions faciales, appelé GP-FER, permettant une sélection et une fusion discriminatives des caractéristiques pour chaque paire de classes d'expressions. Ce cadre fonctionne en trois phases : la sélection des caractéristiques, leur fusion, et la classification. La méthode GP optimise des sous-ensembles de caractéristiques géométriques et texturales, évitant ainsi la redondance et la haute dimensionnalité. Testée sur des bases de données comme DISFA et CK+, la méthode GP-FER a surpassé ou égalé les approches de pointe existantes.

Liu et al. [181], ont proposé le Feature Disentangling Machine (FDM), un cadre conçu pour sélectionner et désentrelacer efficacement les caractéristiques des expressions faciales. Ce système permet de distinguer les caractéristiques communes, partagées entre plusieurs expressions, des caractéristiques spécifiques, qui sont discriminantes pour une expression particulière. En intégrant une machine à vecteurs de support sparse et l'apprentissage multitâche, le FDM utilise une nouvelle fonction de perte et des contraintes afin de contrôler la sparsité et d'assurer la séparation entre les groupes de caractéristiques. Les expériences réalisées sur deux bases de données reconnues (CK+ et JAFFE) démontrent que le FDM surpasse les méthodes existantes, tout en affichant une excellente capacité de généralisation lors des validations croisées.

Ghimire et al. [33], ont proposé un système REF qui se concentre sur la sélection de caractéristiques géométriques saillantes. Ils utilisent des représentations en points, lignes et triangles, en initialisant et suivant un ensemble dense de points faciaux grâce à l'algorithme d'appariement graphique à bunch élastique (EBGM) et au traqueur Kanade-Lucas-Tomasi (KLT). La sélection des caractéristiques les plus discriminantes est réalisée via l'algorithme AdaBoost multiclasse et une classification par machine d'apprentissage extrême (ELM). Les résultats montrent que les représentations triangulaires surpassent celles basées sur des lignes, qui sont elles-mêmes plus efficaces que les représentations par points. En effet, les performances de reconnaissance s'améliorent significativement lorsque l'on considère les relations entre les mouvements des caractéristiques faciales au fil du temps, ce qui est essentiel pour discriminer les expressions faciales.

Puthanidam et Moh [182], ont proposé une approche combinant des étapes de prétraitement d'images avec différentes structures de réseaux de neurones convolutifs (CNN). Cette approche vise à améliorer l'extraction et la sélection des caractéristiques, réalisées de manière automatique par les CNN, ce qui correspond à une méthode de type embedded. Cela permet d'obtenir de meilleures performances en termes de précision et de réduction du temps d'entraînement. Les CNN sont utilisés à la fois pour l'extraction des caractéristiques et la classification des émotions faciales, tirant parti de leur capacité à capturer la structure spatiale des visages. Des étapes de prétraitement comme le recadrage et la réduction de la résolution ont également contribué à ces gains de performance. Testée sur plusieurs ensembles de données, l'approche a atteint des taux de précision élevés, dont 100 % sur la base JAFFE et environ 89,58 % sur la base KDEE.

#### **II.4.4. Sélection de caractéristiques basées sur des méthodes hybride**

Sidhom et al.[183], ont développé une méthode hybride en trois étapes pour améliorer la sélection de caractéristiques dans les tâches de reconnaissance des expressions faciales. Leur approche combine les algorithmes de sélection par filtres, wrappers, et méthodes intégrées. Dans la première étape, ils utilisent une méthode de sélection de caractéristiques basée sur la corrélation afin de réduire l'espace de recherche, ce qui permet d'améliorer l'efficacité et de rendre le processus plus rapide. La deuxième étape consiste en une élimination récursive des caractéristiques pour affiner davantage la sélection. Enfin, dans la troisième étape, ils appliquent l'algorithme XGBoost pour déterminer le sous-ensemble optimal de caractéristiques. Cette méthode a été testée sur cinq ensembles de données spécifiques à la reconnaissance des expressions faciales (DISFA, CK+, DISFA+, JAFFE, et FER2013) en utilisant des mesures telles que l'exactitude, la précision, le rappel et le F1-Score. Les résultats obtenus montrent une performance notable avec des taux d'exactitude de 94,68 % pour DISFA, 98,82 % pour CK+, 95,8 % pour DISFA+, 97,7 % pour JAFFE, et 66,1 % pour FER2013, confirmant ainsi l'efficacité de l'approche.

Ghosh et al.[184], ont proposé une approche de sélection de caractéristiques basée sur l'optimisation par colonies de fourmis (ACO), combinant à la fois des méthodes de filtre et de wrapper pour réduire la complexité computationnelle. Contrairement aux méthodes traditionnelles d'ACO qui sont soit basées sur des filtres soit sur des wrappers, cette approche hybride (nommée WFACOFs) utilise une méthode de filtre pour évaluer les sous-ensembles de caractéristiques, tout en intégrant une mise à jour des phéromones dépendante de la dimension des caractéristiques. Cette technique permet une sélection multi-objective des caractéristiques, en prenant en compte à la fois l'exactitude du modèle et le nombre de caractéristiques sélectionnées. Des expérimentations ont été menées sur des jeux de données réels, notamment des bases de données d'émotions faciales, et les résultats ont montré que cette méthode surpasse la plupart des algorithmes de sélection de caractéristiques existants.

Ghosh et al. [185], ont proposé un algorithme de sélection de caractéristiques, nommé LHCMA (Late Hill Climbing based Memetic Algorithm), qui intègre une nouvelle méthode de recherche locale, LAHCRR (Late Acceptance Hill Climbing through Redundancy and Relevancy), pour optimiser la sélection des caractéristiques. En combinant des techniques de Hill-Climbing et la méthode mRMR, LHCMA réduit efficacement la redondance tout en augmentant la pertinence des caractéristiques. Cette approche est testée sur des descripteurs d'images comme uLBP, Gabor, HOG, et PHOG, en utilisant les jeux de données RaFD et JAFFE. Les résultats montrent que LHCMA surpasse d'autres algorithmes, comme GA et SA, dans la majorité des cas, démontrant une meilleure sélection des sous-ensembles de caractéristiques tout en évitant les optima locaux.

Cossetin et al. [186], ont proposé une méthode hybride pour sélectionner des caractéristiques pertinentes et réduire la dimensionnalité ainsi que la complexité computationnelle. Dans cette méthode, les auteurs ont combiné l'approche « Filter », dans laquelle les caractéristiques sont classées par pertinence en utilisant des techniques telles que la sélection de caractéristiques basée sur la corrélation (CFS), Relief-F et le gain d'information (IG), avec l'approche « Wrapper ». Cette dernière permet d'optimiser progressivement le sous-ensemble des caractéristiques les plus discriminantes pour chaque paire d'expressions. Les résultats obtenus dans [186] ont démontré que cette combinaison surmonte les limitations des méthodes existantes tout en évitant de s'appuyer sur des points de référence et en améliorant la précision de classification du SVM.

#### **II.5. Sélection de caractéristiques pour la reconnaissance des expressions faciales en mode indépendant de la personne : analyse et comparaison des approches**

La reconnaissance des expressions faciales indépendante de la personne pose des défis importants en raison des différences interpersonnelles inhérentes, telles que les différences de morphologie faciale,

d'expressions individuelles et de conditions de capture d'images. Ces changements augmentent la complexité du problème et rendent difficile la distinction précise des expressions faciales des différents visages. Pour surmonter ce défi, la sélection de caractéristiques pertinentes constitue une étape importante, qui non seulement réduit la dimensionnalité des données mais améliore également la capacité discriminante du modèle de classification. Cette section explore diverses méthodes de sélection de caractéristiques appliquées à la reconnaissance des expressions faciales dans des modes indépendants de la personne.

Shanthi et Nickolas [187], ont proposé une approche de fusion de caractéristiques locales pour améliorer la reconnaissance des expressions faciales. Cette méthode combine deux descripteurs complémentaires : le Local Binary Pattern (LBP) et le Local Neighborhood Encoded Pattern (LNEP). Le LBP encode la relation entre un pixel central et ses voisins immédiats, tandis que le LNEP représente la relation entre deux pixels voisins les plus proches. Afin de réduire la dimensionnalité des données, les auteurs ont utilisé l'analyse statistique par chi-carré pour sélectionner les caractéristiques les plus pertinentes issues de cette combinaison. L'approche a montré une amélioration significative des performances, avec des taux de reconnaissance atteignant 97,86 % sur le jeu de données CK+ et 97,11 % sur le jeu de données MMI. La méthode s'est également révélée robuste dans des environnements bruités, démontrant son efficacité par rapport aux approches classiques de reconnaissance faciale.

Makhmudkhujayev et al. [188], ont proposé une méthode de sélection de caractéristiques basée sur le descripteur Local Prominent Directional Pattern (LPDP). LPDP est conçu pour capturer les variations directionnelles locales des mouvements musculaires faciaux, en tenant compte des statistiques des informations de bord dans un voisinage local d'un pixel. Cette approche permet d'encoder des informations plus fiables et pertinentes, même en présence de variations de position ou de bruit, tout en évitant les artefacts indésirables dans les régions lisses comme les joues. Pour la sélection de caractéristiques, les auteurs calculent un histogramme des occurrences des codes LPDP sur un ensemble d'images d'expression, ce qui permet d'identifier les motifs les plus récurrents, liés aux expressions faciales, et donc les plus pertinents pour la classification. Les caractéristiques moins fréquentes, souvent associées à du bruit ou à des variations individuelles, sont soit ignorées, soit intégrées de manière contrôlée afin de ne pas dégrader la performance globale. Cette méthode de sélection de caractéristiques permet ainsi de se concentrer sur les motifs d'expression les plus robustes, augmentant la précision de la reconnaissance des expressions tout en réduisant la complexité des données.

Lajevardi et Hussain [189], ont comparé trois méthodes de sélection de caractéristiques pour la reconnaissance des expressions faciales à partir d'images fixes en utilisant des filtres log-Gabor, avec des tests réalisés sur la base d'images Cohn-Kanade. Les approches incluent la sélection de 4 filtres optimaux (log-Gabor/4-OF), l'analyse en composantes principales (log-Gabor/PCA) et la sélection basée sur l'information mutuelle (log-Gabor/MIFS). Les résultats montrent que la méthode log-Gabor/MIFS a obtenu la meilleure performance avec un taux de classification moyen de 75,5 %, surpassant les méthodes 4-OF (68,9 %) et PCA (52,5 %). L'information mutuelle s'est révélée plus efficace que les approches linéaires comme la corrélation, permettant une meilleure discrimination des expressions faciales et une simplicité computationnelle accrue.

Makhmudkhujayev et al. [190], ont proposé un nouveau descripteur, appelé Histograms of Prominent Edge Directions (HPED), pour améliorer la reconnaissance des expressions faciales dans un environnement indépendant de la personne, tout en soulevant le problème de l'erreur d'échantillonnage observée dans les descripteurs existants. Concernant la sélection des caractéristiques, HPED se distingue en utilisant un histogramme des deux directions de bord les plus marquantes (primaire et secondaire) dans de petites régions spatiales du visage, ce qui permet d'éviter l'erreur d'échantillonnage courante lorsque trop de bins sont utilisés avec un nombre limité d'échantillons. En limitant chaque histogramme à huit bins, HPED réduit la probabilité de bins vides et maximise la préservation de l'information spatiale pertinente. Les auteurs ont également optimisé des paramètres tels que la taille des grilles (11x11) pour

trouver un compromis optimal entre l'évitement de l'erreur d'échantillonnage et la préservation des informations spatiales.

Shan et al. [191], ont introduit la technique Boosted-LBP, qui utilise l'algorithme AdaBoost pour sélectionner les caractéristiques les plus discriminantes parmi les Local Binary Patterns (LBP) dans le cadre de la reconnaissance des expressions faciales. Cette approche vise à optimiser la représentation des données en apprenant spécifiquement quelles caractéristiques LBP sont les plus informatives pour la classification des expressions. En intégrant Boosted-LBP avec des classificateurs tels que les machines à vecteurs de support (SVM), ils parviennent à améliorer les performances de reconnaissance. L'utilisation de Boosted-LBP se révèle particulièrement efficace, car elle permet d'extraire des caractéristiques pertinentes tout en maintenant une bonne généralisation sur différents ensembles de données. De plus, les auteurs soulignent que la dérivation d'une représentation faciale efficace à partir des images d'origine est essentielle pour la réussite de la reconnaissance des expressions faciales. Les résultats expérimentaux montrent que les caractéristiques LBP sont non seulement efficaces et efficientes pour cette tâche, mais qu'elle performe également de manière stable et robuste, même dans des conditions de basse résolution.

Zarbakhsh et Demirel [192], ont abordé la détection des expressions faciales en trois dimensions (3D) en proposant un algorithme combinant la sélection et la classification des caractéristiques. Ils ont extrait un grand nombre de caractéristiques basées sur les distances entre les points du modèle facial, puis appliqué l'algorithme de sélection séquentielle avant (SFFS) pour choisir les caractéristiques les plus discriminantes. L'algorithme SFFS, couplé à une validation croisée à 5 volets, permet de minimiser l'erreur de classification du classificateur Naïve Bayes. Comparée à la méthode du t-test, cette approche prend en compte les interactions entre les caractéristiques, évitant ainsi la redondance et améliorant la qualité des sous-ensembles sélectionnés. Les expériences sur le jeu de données BU-3DFE ont montré que l'utilisation de SFFS a permis d'obtenir une meilleure précision par rapport à la sélection par t-test, avec un taux de reconnaissance amélioré de 72,33 % à 87,67 % lorsque couplée à un classificateur flou SVM (FSVM).

Yurtkan et Demirel [165], ont proposé une méthode de sélection de caractéristiques pour améliorer la reconnaissance des expressions faciales en utilisant des points géométriques 3D du visage. Cette procédure repose sur l'analyse d'entropie, qui permet de sélectionner les points les plus discriminants en fonction des déformations faciales liées aux six émotions de base (colère, dégoût, peur, joie, tristesse, surprise). En transformant les données 3D en valeurs de magnitude, les points avec la plus forte entropie sont retenus, maximisant ainsi l'information utile pour la classification. Cette sélection de caractéristiques est intégrée dans un classificateur SVM à deux niveaux, où différents points sont sélectionnés à chaque niveau. Les résultats expérimentaux, réalisés sur la base de données BU-3DFE, montrent une amélioration notable des taux de reconnaissance, passant de 80% à 88,28% après l'application de la sélection de caractéristiques, démontrant ainsi son efficacité dans un contexte de reconnaissance indépendante de la personne.

Tekgüç et al. [193], ont proposé une méthode optimale de sélection de caractéristiques pour la reconnaissance des expressions faciales en 3D, qui utilise l'algorithme génétique multi-objectifs Non-dominated Sorted Genetic Algorithm II (NSGA II). Cette méthodologie vise à réduire le nombre de caractéristiques tout en maximisant la séparabilité entre les classes et en minimisant la séparabilité au sein des classes. Pour ce faire, les auteurs se basent sur le critère de Fisher, qui mesure la distance entre les moyennes des classes et la variance intra-classe, permettant ainsi de quantifier la discriminabilité des caractéristiques sélectionnées. En évaluant leur approche sur la base de données BU-3DFE, composée de 420 modèles faciaux 3D, les auteurs ont pu extraire un ensemble de 25 caractéristiques discriminantes. L'algorithme NSGA II a généré une courbe de compromis entre le nombre de caractéristiques et le critère de Fisher, permettant de sélectionner les caractéristiques les plus pertinentes tout en optimisant la performance de classification. Le système a montré un taux de reconnaissance moyen de 88,18 % pour des expressions faciales variées telles que la colère, la joie, et la tristesse. En



s'assurant que les sujets de test ne figuraient pas dans l'ensemble d'entraînement, cette méthode a permis d'atteindre une classification indépendante des individus, ce qui représente un défi supplémentaire. Les résultats indiquent que cette méthode de sélection de caractéristiques optimisée améliore significativement la reconnaissance des expressions faciales par rapport aux méthodes existantes.

Xue et al. [194], ont proposé une méthode entièrement automatique de reconnaissance des expressions faciales en 3D en se basant sur des caractéristiques de profondeur extraites de patches locaux autour de points heuristiques. Concernant la sélection de caractéristiques, ils ont utilisé la méthode mRMR pour choisir les caractéristiques les plus représentatives des différentes expressions faciales. Cette méthode maximise la pertinence des caractéristiques tout en minimisant leur redondance, ce qui permet d'améliorer la précision de la classification. Ensuite, une validation croisée en 10 volets a été réalisée pour déterminer le nombre optimal de caractéristiques. Le modèle de classification final est un SVM multi-classes entraîné sur les caractéristiques sélectionnées.

Dans [195], Happy et Routray abordent le défi de la reconnaissance automatique des micro-expressions, des mouvements faciaux subtils et de courte durée révélant les émotions. Ils proposent d'utiliser les descripteurs FHOFO (Fuzzy Histogram of Optical Flow Orientation) pour capturer les orientations des mouvements sans tenir compte de leur magnitude, car la direction du mouvement est jugée plus informative que son ampleur dans ce contexte. FHOFO construit un histogramme des orientations de mouvement en utilisant uniquement les vecteurs de direction extraits du flux optique. Ensuite, différentes méthodes de sélection de caractéristiques sont appliquées pour réduire la dimensionnalité des données, incluant PWFP (Pair-wise Feature Proximity), ReliefF, et FSV (Feature Selection via concave minimization). PWFP sélectionne les caractéristiques en évaluant la proximité des échantillons de même classe et leur éloignement par rapport aux autres classes, facilitant ainsi la sélection des attributs les plus pertinents. ReliefF est une méthode basée sur les instances qui met à jour les poids des caractéristiques en comparant des paires d'instances : une appartenant à la même classe (near-hit) et l'autre à une classe différente (near-miss), en ajustant le score d'une caractéristique selon la différence entre ces paires. FSV réduit la complexité du modèle en supprimant autant que possible les composantes du vecteur normal du plan séparateur obtenu par SVM tout en maintenant une performance acceptable. Les performances des modèles KNN, SVM et LDA sont ensuite évaluées sur des bases de données telles que CASME, CASME2, et SMIC. Les résultats montrent que PWFP, combinée avec un classificateur SVM, offre des performances globales stables et supérieures, notamment avec les descripteurs FHOFO.

Zhou et al. [196], ont exploré l'utilisation de réseaux neuronaux convolutifs (CNN) pour la reconnaissance des expressions faciales, en mettant l'accent sur la sélection de caractéristiques liées aux unités d'action (AU). En appliquant un transfert d'apprentissage d'un CNN pré-entraîné sur ImageNet aux bases de données KDEF, RaFD et CK+, ils ont montré que les couches convolutives supérieures des CNN deviennent sélectives aux AUs, ce qui est crucial pour la reconnaissance des expressions. Ils ont proposé deux méthodes de sélection de caractéristiques : l'une utilisant un réseau déconvolutionnel pour projeter les activations des cartes de caractéristiques dans les régions faciales, et l'autre en classant les cartes selon leur sélectivité aux AUs. Leurs résultats démontrent que les cartes de caractéristiques sélectives aux AUs jouent un rôle plus important dans la tâche de reconnaissance des expressions, validées par des expériences d'ablation des caractéristiques.

Xie et al. [197] ont proposé un modèle transformant la reconnaissance multiclassique des expressions faciales en un problème de triplet d'expressions, afin de mieux exploiter la spécificité de chaque variation d'expression. Ce modèle repose sur l'optimisation des unités d'action (AUs) et des patches faciaux, avec un processus en deux étapes. D'abord, les poids des AUs et des patches sont optimisés hors ligne à partir d'une matrice de probabilité conditionnelle. Ensuite, en phase de test, les AUs actifs des échantillons sont détectés en ligne pour une meilleure précision de reconnaissance. L'algorithme obtient des taux de reconnaissance compétitifs de 89,67 % sur la base de données JAFFE et 94,09 % sur Cohn-Kanade (CK+), surpassant les approches existantes telles que AdaBoost et LDA. Le modèle se distingue par trois contributions majeures : l'utilisation du mode triplet pour représenter la diversité des expressions,

l'optimisation des poids des AUs et des patches pour chaque triplet, et la détection en ligne des AUs actifs pour chaque échantillon de test.

Mlakar et Potocnik [198], ont proposé une méthode automatisée pour la reconnaissance des expressions faciales basée sur le descripteur Histogram of Oriented Gradients (HOG) et optimisée par un algorithme génétique (GA). Cette méthode indépendante des sujets vise à reconnaître six émotions prototypiques en comparant les descripteurs HOG d'une expression neutre et d'une expression émotionnelle maximale. L'algorithme génétique a été utilisé pour sélectionner les caractéristiques optimales du descripteur HOG, influençant directement la précision du système. Chaque solution générée par l'algorithme évolue à travers des opérateurs génétiques, tels que le croisement et la sélection par roulette, afin d'affiner les caractéristiques du HOG, constituant ainsi une forme de sélection de caractéristiques. Le modèle utilise des machines à vecteurs de support (SVM) avec un noyau à fonction de base radiale (RBF) pour classifier les émotions. Chaque classifieur est formé en utilisant les vecteurs de différence HOG entre les images neutres et émotionnelles. La méthode a été validée sur 1232 images de la base Cohn-Kanade et 192 images de la base JAFFE, obtenant un taux de reconnaissance moyen de 95,64% sur Cohn-Kanade, surpassant de nombreuses méthodes existantes basées sur une seule image ou vidéo. La sélection des caractéristiques est réalisée indirectement, en optimisant les caractéristiques du HOG via l'algorithme génétique, qui évalue chaque configuration sur un sous-ensemble de données de validation. Ce processus permet de sélectionner les caractéristiques les plus pertinents du descripteur HOG, capturant ainsi les informations les plus discriminantes des expressions faciales. Les résultats ont montré que cette approche améliore la reconnaissance par rapport à l'utilisation d'une seule image émotionnelle, avec une amélioration d'environ 4% sur Cohn-Kanade et 20% sur JAFFE.

Shojaeilangari et al. [199], ont proposé une méthode de sélection de caractéristiques basée sur un algorithme génétique (GA) pour l'analyse dynamique des expressions faciales à partir de séquences vidéo. En utilisant des filtres de Gabor pour extraire les caractéristiques, ils ont identifié que la dimensionnalité élevée des vecteurs de caractéristiques entraînait des coûts de calcul importants et une redondance. Pour surmonter ces défis, le GA a été employé pour sélectionner un sous-ensemble optimal de filtres de Gabor, permettant ainsi de réduire le nombre de caractéristiques tout en améliorant la performance de classification. Les résultats expérimentaux, en utilisant un classificateur SVM, ont montré que cette méthode de sélection a permis d'atteindre un taux de reconnaissance de 92,97 % pour les six émotions de base sur la base de données Cohn-Kanade (CK+), tout en n'utilisant qu'un sous-ensemble de huit filtres parmi quarante. Ainsi, l'application du GA pour la sélection de caractéristiques s'avère être une stratégie prometteuse pour améliorer l'efficacité des systèmes de reconnaissance des expressions faciales en se concentrant sur les filtres les plus pertinents.

Soyel et al. [154], ont proposé une méthode de sélection de caractéristiques pour la reconnaissance des expressions faciales 3D en utilisant l'algorithme génétique de tri non-dominé II (NSGA-II). Ce dernier optimise un vecteur de poids de caractéristiques afin de maximiser la séparabilité des classes en fonction du critère de Fisher, tout en minimisant le nombre de caractéristiques sélectionnées. Le réseau neuronal probabiliste (PNN) est utilisé pour la classification des expressions faciales à partir des caractéristiques optimisées. Testée sur la base de données BU-3DFE, la méthode atteint un taux de classification de 88,2 %. NSGA-II est appliqué ici pour trouver des solutions optimales dites Pareto-front, garantissant un équilibre entre le nombre de caractéristiques et la séparabilité des classes. Ce processus repose sur une approche multi-objective avec des étapes de sélection, de croisement et de mutation.

Mlakar et al. [170], ont proposé un système de sélection de caractéristiques performant pour la reconnaissance des expressions faciales, capable de distinguer sept émotions, y compris l'expression neutre, en utilisant des descripteurs HOG et des vecteurs de différence. Ils ont modifié l'algorithme d'évolution différentielle multi-objectifs (DEMO) pour minimiser le nombre de caractéristiques tout en maximisant la précision des classificateurs SVM. Deux stratégies de sélection ont été introduites : l'une axée sur les émotions spécifiques et l'autre sur les caractéristiques discriminantes globales, cette dernière s'étant révélée plus efficace. La méthode a été validée sur les bases de données Cohn-Kanade, JAFFE,

et MMI, atteignant des taux de reconnaissance moyens de 98,37 %, 92,75 % et 84,07 % respectivement, tout en réduisant de 89 % le nombre de caractéristiques utilisées par rapport aux vecteurs de départ. Cette réduction significative diminue ainsi les coûts de calcul, tout en offrant des résultats comparables, voire supérieurs aux méthodes de pointe existantes.

Le (Tableau II.5) récapitule et présente une variété de méthodes de sélection de caractéristiques ainsi que leurs impacts sur les systèmes de reconnaissance des expressions faciales en mode indépendant de la personne. Ces méthodes jouent un rôle crucial dans la capacité des systèmes à généraliser les caractéristiques pertinentes des expressions faciales tout en réduisant le bruit et la redondance des données.

En examinant les différentes méthodes de sélection, on remarque que l'utilisation de techniques comme le filtre Chi-carré pour la fusion de LBP et LNEP a permis d'obtenir des taux de reconnaissance élevés de 97,86 % sur CK+ et 97,11 % sur MMI. Cette approche de sélection de caractéristiques est efficace car elle évalue la pertinence des caractéristiques de manière indépendante par rapport à l'algorithme de classification. Cela souligne l'importance de choisir des caractéristiques qui apportent réellement des informations discriminantes pour la tâche de classification.

À l'opposé, la méthode LPDP, qui repose sur le code LPDP et l'information d'occurrence, montre des résultats compétitifs, atteignant 94,50 % sur CK+, 70,63 % sur MMI, 73.4% sur BU-3DFE, 78.32 % sur ISED et 94.72% sur FACES. Cela démontre que des méthodes basées sur des techniques d'optimisation spécifiques peuvent également être efficaces, même si elles ne sont pas toujours aussi performantes que celles basées sur des filtres. L'importance de la sélection de caractéristiques ici réside dans sa capacité à extraire des informations significatives qui peuvent améliorer la performance globale du modèle de classification.

La méthode de sélection séquentielle avant (SFFS) appliquée à la distance entre les points du modèle facial 3D, a atteint 87,67 % sur BU-3DFE. Cette approche itérative permet d'ajouter des caractéristiques progressivement, en évaluant leur impact sur la performance à chaque étape. Cela démontre que, même si certaines caractéristiques peuvent ne pas sembler pertinentes au départ, leur combinaison peut produire des résultats significatifs. Cela souligne l'importance d'une stratégie de sélection de caractéristiques qui tient compte non seulement des caractéristiques individuelles, mais également de leurs interactions.

En revanche, des méthodes comme l'algorithme de poids pour l'optimisation qui a produit des résultats de 89,67 % sur JAFFE et 94,09 % sur CK+, montrent que des techniques de sélection basées sur des heuristiques peuvent également s'avérer efficaces. Cependant, ces méthodes peuvent être sensibles au choix des paramètres, ce qui peut affecter leur robustesse et leur généralisation sur d'autres jeux de données.

L'utilisation de méthodes telles que NSGA II, qui se base sur des techniques d'optimisation multi-objectifs, a permis d'atteindre un taux de reconnaissance de 88,2 % sur BU-3DFE. Cela suggère que l'optimisation de plusieurs critères simultanément, tels que la précision et la complexité du modèle, peut être bénéfique pour développer des systèmes plus robustes et efficaces.

En outre, l'impact de méthodes de sélection comme MIFS (Mutual Information Feature Selection) appliquées aux filtres log-Gabor, qui ont obtenu un taux de reconnaissance de 75,5 %, souligne l'importance de la corrélation entre les caractéristiques et leur capacité à capturer les variations des expressions faciales. Bien que cette méthode ait affiché des résultats moins impressionnants, elle est révélatrice des défis liés à la sélection de caractéristiques dans des scénarios où la diversité des données est élevée.

Il est également intéressant de noter que les méthodes basées sur l'apprentissage profond, comme celles utilisant des cartes de caractéristiques extraites par des CNN pré-entraînés, montrent des résultats

remarquables (97,75 % sur RAFD). Ces méthodes exploitent des architectures complexes qui automatisent la sélection de caractéristiques, permettant d'atteindre des performances de pointe. Cela indique que les méthodes traditionnelles de sélection de caractéristiques, bien qu'efficaces, peuvent parfois être surpassées par des approches plus sophistiquées qui apprennent à extraire les caractéristiques les plus pertinentes directement à partir des données brutes.

En somme, le tableau illustre que les méthodes de sélection de caractéristiques sont essentielles pour optimiser les systèmes de reconnaissance des expressions faciales en mode indépendant de la personne. Chacune de ces méthodes présente des avantages et des inconvénients qui peuvent influencer la performance en fonction des jeux de données utilisés. Une sélection appropriée des caractéristiques peut non seulement améliorer la précision de la reconnaissance, mais également réduire la complexité du modèle, augmentant ainsi son efficacité et sa capacité à généraliser. L'avenir de la reconnaissance des expressions faciales réside dans la recherche de méthodes de sélection de caractéristiques toujours plus robustes et adaptatives, capables de s'ajuster aux variations des données et aux exigences des applications pratiques.

## **II.6. Conclusion**

Ce chapitre a exploré en profondeur le domaine de la sélection de caractéristiques dans le contexte de la reconnaissance des expressions faciales. Nous avons commencé par établir des concepts fondamentaux et des définitions, avant de souligner l'importance cruciale de cette étape pour améliorer la performance des systèmes de REF. En examinant diverses méthodes de sélection de caractéristiques, nous avons mis en lumière leur impact sur la capacité des modèles à capturer des informations pertinentes tout en minimisant le bruit et la redondance. Nous avons également analysé l'application de ces méthodes à la reconnaissance des expressions faciales, en particulier dans un mode indépendant de la personne. La comparaison des différentes approches de sélection de caractéristiques souligne la diversité des méthodes disponibles et leur impact sur les performances des systèmes de reconnaissance des expressions faciales. La recherche continue dans ce domaine est essentielle pour développer des systèmes REF plus robustes et adaptatifs, capables de s'ajuster aux variations des données et aux exigences des applications pratiques.

**Tableau II.5: Tableau récapitulatif des travaux de recherche adoptant des méthodes de sélection de caractéristiques pour les systèmes de reconnaissance des expressions faciales en mode indépendant de la personne.**

| Référence | Méthode d'Extraction de Caractéristiques (Nature des caractéristiques) | Méthode de Sélection de Caractéristiques                          | Modèle de Classification | Jeux de Données                          | Taux de Reconnaissance                                                                       | Temps de calcul                         |
|-----------|------------------------------------------------------------------------|-------------------------------------------------------------------|--------------------------|------------------------------------------|----------------------------------------------------------------------------------------------|-----------------------------------------|
| [187]     | Fusion de LBP et LNEP                                                  | Filter (Chi-carré)                                                | Multiclass SVM           | CK+, MMI                                 | 97,86% (CK+), 97,11% (MMI)                                                                   | Non précisé (CK+), 29.94 secondes (MMI) |
| [188]     | LPDP                                                                   | LPDP code Occurrence information                                  | Linear SVM               | CK+, MMI, BU-3DFE, ISED, FACES           | 94.50% (CK+), 70.63% (MMI), 73.4% (BU-3DFE), 78.32 % (ISED), 94.72% (FACES)                  | 4.2 milliseconds                        |
| [189]     | Filtres log-Gabor                                                      | MIFS                                                              | KNN                      | CK                                       | 75,5 %                                                                                       | Non précisé                             |
| [190]     | Histograms of Prominent Edge Directions (HPED)                         | sélection basée sur une recherche exhaustive des caractéristiques | SVM                      | CK+, FACES, BU-3DFE, RaFD                | 93.74% (CK+), 93.62% (FACES), 73.39 % (BU-3DFE), 95.52 % (RaFD)                              | Non précisé                             |
| [191]     | LBP                                                                    | Boosted-LBP                                                       | SVM (linear)             | CK-6 émotions, CK-7 émotions, MMI, JAFFE | 95.0% $\pm$ 3.2 (CK-6 émotions), 91.1% $\pm$ 4.0 (CK-7 émotions), 86.7% (MMI), 79.8% (JAFFE) | Non précisé                             |
|           |                                                                        |                                                                   | SVM (polynomial)         |                                          | 95.0% $\pm$ 3.2 (CK-6 émotions), 91.1% $\pm$ 4.0 (CK-7 émotions), 86.7% (MMI), 79.8% (JAFFE) |                                         |

|       |                                                                                                  |                                                                |                                                    |                                              |                                                                                              |             |
|-------|--------------------------------------------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------|----------------------------------------------|----------------------------------------------------------------------------------------------|-------------|
|       |                                                                                                  |                                                                | SVM (RBF)                                          |                                              | 95.1% $\pm$ 3.4 (CK-6 émotions), 91.4% $\pm$ 3.8 (CK-7 émotions), 86.9% (MMI), 81.0% (JAFPE) |             |
| [192] | Distances entre points du modèle facial 3D                                                       | Sélection séquentielle avant (SFFS)                            | fuzzy SVM (FSVM) à deux niveaux                    | BU-3DFE                                      | 87,67 %                                                                                      | Non précisé |
| [165] | Positions géométriques des points faciaux en 3D                                                  | Basée sur l'entropie des points lors des déformations faciales | SVM à deux niveaux                                 | BU-3DFE                                      | 88.28%                                                                                       | Non précisé |
| [193] | 83 points caractéristiques du visage de la base de données BU-3DFE                               | NSGA II                                                        | Probabilistic Neural Network (PNN)                 | BU-3DFE                                      | 88.18%                                                                                       | Non précisé |
| [194] | Profondeur des patches autour des points heuristiques (30 points à partir de 5 points de repère) | mRMR                                                           | SVM multi-classe                                   | BU-3DFE                                      | 83.0%                                                                                        | Non précisé |
| [195] | FHOFO                                                                                            | PWFP                                                           | Linear SVM                                         | CASME, CASME2, SMIC                          | 67.24% (CASME), 56.78% (CASME2), 52.01% (SMIC)                                               | Non précisé |
| [196] | Cartes de caractéristiques (feature maps) extraites par un CNN pré-entraîné sur ImageNet         | Face-Occupancy-based Feature Selection (FOS)                   | Réseau de neurones fully-connected à trois couches | KDEF, RAFD, CK+ (6 classes), CK+ (8 classes) | 88.27% (KDEF), 97.75% (RAFD), 97.35% (CK+ (6 classes)), 96.37% (CK+ (8 classes))             | Non précisé |
|       |                                                                                                  | Action Unit Selectivity based Feature Selection (AUS)          |                                                    | CK+ (6 classes), CK+ (8 classes)             | 97.50% (CK+ (6 classes)), 96.95% (CK+ (8 classes))                                           |             |

|       |                                                                                          |                               |                                                    |                |                                            |             |
|-------|------------------------------------------------------------------------------------------|-------------------------------|----------------------------------------------------|----------------|--------------------------------------------|-------------|
| [197] | Extraction de caractéristiques basée sur des correctifs (patch based feature extraction) | Weight optimization algorithm | SVM                                                | JAFFE, CK+     | 89.67% (JAFFE), 94.09% (CK+)               | Non précisé |
| [165] | 83 points caractéristiques du visage de la base de données BU-3DFE                       | Analyse d'entropie            | SVM (à deux niveaux)                               | BU-3DFE        | 88.28%                                     | Non précisé |
| [198] | HOG                                                                                      | Algorithme génétique (GA)     | SVM avec un noyau à fonction de base radiale (RBF) | CK, JAFFE      | 95.64% (CK), 87.82% (JAFFE)                | Non précisé |
| [199] | Filtres de Gabor                                                                         | Algorithme génétique (GA)     | SVM                                                | CK+            | 92,97%                                     | Non précisé |
| [154] | Calcul des caractéristiques de distance faciale 3D                                       | NSGA-II                       | PNN                                                | BU-3DFE        | 88,2 %                                     | Non précisé |
| [170] | HOG                                                                                      | DEMO                          | linear SVM                                         | CK, JAFFE, MMI | 98,37 % (CK), 92,75 % (JAFFE), 84,07 % MMI | Non précisé |

# Chapitre III

## Implémentation et Résultats

---

|                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------|-----|
| III.1. Introduction                                                                                                                      | 72  |
| III.2. Sélection de caractéristiques locales à l'aide de l'approche wrapper pour la reconnaissance des expressions faciales              | 72  |
| III.2.1. Principes du système REF proposé                                                                                                | 75  |
| III.2.2. Résultats expérimentaux et discussion                                                                                           | 79  |
| III.2.2.1. L'impact de l'ajout d'une approche de sélection de caractéristiques sur la performance du système REF                         | 79  |
| III.2.2.2. Analyse du phénomène de la malédiction de la dimensionnalité                                                                  | 82  |
| III.2.2.3. Pertinence des régions faciales locales sur la performance du système REF                                                     | 84  |
| III.2.3. Comparaison avec les méthodes d'état de l'art                                                                                   | 85  |
| III.3. Un nouvel algorithme de reconnaissance des expressions faciales basé sur l'extraction et la sélection de caractéristiques par DWT | 86  |
| III.3.1. Le modèle proposé                                                                                                               | 87  |
| III.3.1.1. Prétraitement des images                                                                                                      | 88  |
| III.3.1.2. Extraction de caractéristiques                                                                                                | 88  |
| III.3.1.3. Sélection de caractéristiques                                                                                                 | 93  |
| III.3.1.4. Reconnaissance émotionnelle                                                                                                   | 95  |
| III.3.2. Résultats expérimentaux et discussion                                                                                           | 96  |
| III.3.2.1. Préparation des données                                                                                                       | 96  |
| III.3.2.2. Expériences                                                                                                                   | 97  |
| III.3.2.3. Résultats et discussion                                                                                                       | 98  |
| III.3.2.4. Évaluation des performances                                                                                                   | 99  |
| III.3.2.5. Comparaison des performances                                                                                                  | 100 |
| III.4. Conclusion                                                                                                                        | 102 |



### III.1. Introduction

Dans ce chapitre, nous allons explorer deux approches distinctes que nous avons développées pour la reconnaissance des expressions faciales, chacune visant à améliorer la sélection de caractéristiques et l'efficacité des algorithmes de classification.

Le premier travail, intitulé "Local Feature Selection Using the Wrapper Approach for Facial Expression Recognition", repose sur l'utilisation de caractéristiques géométriques locales pour améliorer la reconnaissance des expressions faciales. Les caractéristiques géométriques, qui capturent la structure spatiale et la configuration des points clés du visage, sont combinées avec une méthode de sélection par wrapper. Cette approche permet de sélectionner dynamiquement les caractéristiques les plus pertinentes en fonction des performances d'un modèle de classification, optimisant ainsi la reconnaissance.

Le second travail, intitulé "A New Facial Expression Recognition Algorithm Based on DWT Feature Extraction and Selection", introduit une nouvelle méthode d'extraction et de sélection de caractéristiques basée sur la transformation en ondelettes discrètes (DWT). Cette approche exploite la capacité de la DWT à représenter des informations fréquentielles et spatiales, améliorant ainsi la représentation des expressions faciales pour une reconnaissance plus précise.

Les sections suivantes décriront en détail la méthodologie, la mise en œuvre et les résultats expérimentaux de chaque approche, mettant en lumière les avantages et les limites de chaque méthode.

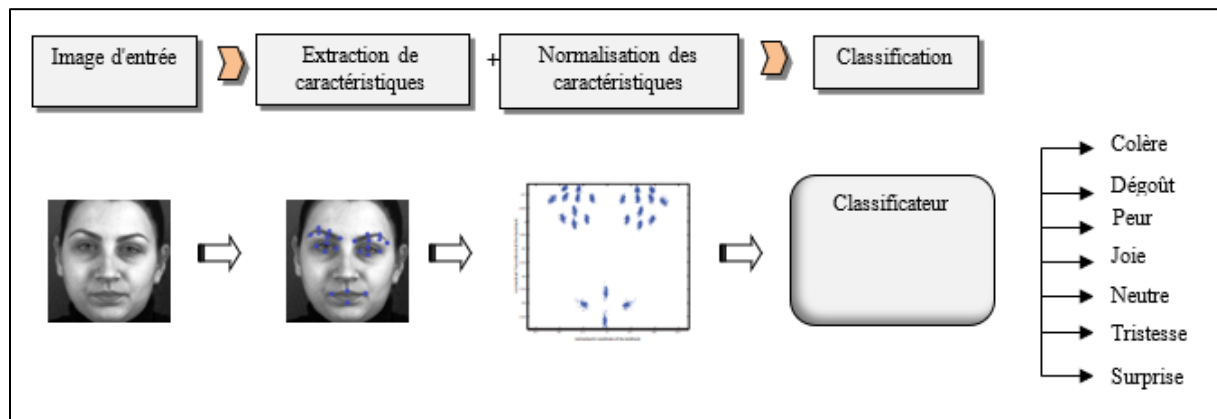
### III.2. Sélection de caractéristiques locales à l'aide de l'approche wrapper pour la reconnaissance des expressions faciales

Les systèmes de reconnaissance automatique des expressions faciales jouent un rôle essentiel dans l'expression et l'interprétation des états émotionnels et mentaux des individus. Ces systèmes transforment les images faciales en ensembles de caractéristiques pour entraîner des classificateurs capables de distinguer diverses émotions. Toutefois, un défi courant réside dans le fait que les vecteurs de caractéristiques extraits peuvent contenir des éléments non pertinents ou redondants, ce qui compromet la précision du classificateur et alourdit les temps de calcul. Pour résoudre ce problème, il est crucial de réduire la dimensionnalité en sélectionnant uniquement les caractéristiques les plus pertinentes. Dans le domaine de la reconnaissance des expressions faciales, une étape essentielle est la sélection des caractéristiques, qui permet d'extraire les informations les plus significatives des images. Cela améliore les performances du modèle à travers différentes modalités, telles que les données bidimensionnelles (2D), tridimensionnelles (3D) et temporelles. Pour la REF en 2D, la sélection des caractéristiques implique l'identification des éléments faciaux essentiels qui capturent les informations nécessaires à la reconnaissance des expressions tout en éliminant les données non pertinentes.

Dans ce qui suit, l'objectif est de démontrer l'efficacité et l'impact de l'ajout d'une étape de sélection de caractéristiques au processus REF, ainsi que de mener une étude sur l'effet des régions faciales locales, telles que la bouche, les yeux et les sourcils, sur les performances du système REF. Cette étude vise à déterminer les régions faciales locales contenant les caractéristiques les plus importantes pour assurer une bonne reconnaissance des expressions émotionnelles faciales. Pour ce faire, nous proposons de modifier et d'améliorer un système REF présenté dans [200]. Nous étudions l'impact de l'ajout de l'approche de sélection par "wrapper" et de l'utilisation des informations fournies par différentes régions locales du visage, telles que la bouche, les yeux et les sourcils, sur les performances d'un système REF basé sur une méthode d'extraction de caractéristiques géométriques locales.

Une vue schématique du système de reconnaissance des expressions faciales proposé par Silva et al. [200] est montrée dans la (Figure III.1). Comme l'illustre cette figure, le système en question comprend trois blocs fonctionnels. Le premier bloc caractérise la fonction de détection du visage, le deuxième

concerne l'extraction et la normalisation des caractéristiques faciales, et enfin, le dernier bloc fonctionnel représente la classification des expressions faciales.



**Figure III.1: Un aperçu schématique du système REF proposé dans [200].**

Le principe de chacun de ces blocs est décrit ci-dessous.

- Bloc de Détection du Visage :

Dans [200], les visages sont localisés à l'aide de l'algorithme proposé par Viola et Jones [29], qui se compose de deux étapes principales : l'extraction de caractéristiques de type Haar et la classification par AdaBoost.

- Bloc d'Extraction des Caractéristiques :

Après l'étape de détection du visage, les caractéristiques sont extraites en utilisant la méthode ASM proposée par Cootes et al. [41]. L'ASM est une méthode de correspondance de caractéristiques basée sur un modèle statistique [32] comprenant deux étapes. Dans la première étape, chaque objet structuré est représenté par un ensemble de points de repère placés manuellement sur chaque image. Dans la seconde étape, les points de repère sont automatiquement alignés pour minimiser la distance entre leurs points correspondants. Ensuite, l'ASM crée un modèle statistique de la forme faciale qui se déforme itérativement pour adapter le modèle à une nouvelle image [200]. Le nombre optimal de points de repère faciaux garantissant une meilleure performance est de 24 [200]. Ce nombre, qui assure le rapport de précision maximum, a été choisi après un test d'évaluation basé sur 9 ensembles de points de repère faciaux contenant respectivement 4, 12, 20, 24, 27, 35, 39, 55 et 76 points de repère. Le (Tableau III.1) décrit le nombre de points de repère extraits pour chaque région faciale utilisée.

**Tableau III.1: Nombre de points de repère extraits pour chaque région du visage.**

| Région faciale | Nombre de points de repère |
|----------------|----------------------------|
| Bouche         | 4                          |
| Œil gauche     | 4                          |
| Œil droit      | 4                          |
| Sourcil gauche | 6                          |
| Sourcil droit  | 6                          |
| <b>Totale</b>  | <b>24</b>                  |

Le choix des régions faciales est basé sur le fait que les expressions faciales sont des déformations des principales zones permanentes du visage (bouche, yeux, etc.). Ces zones locales contiennent la plupart

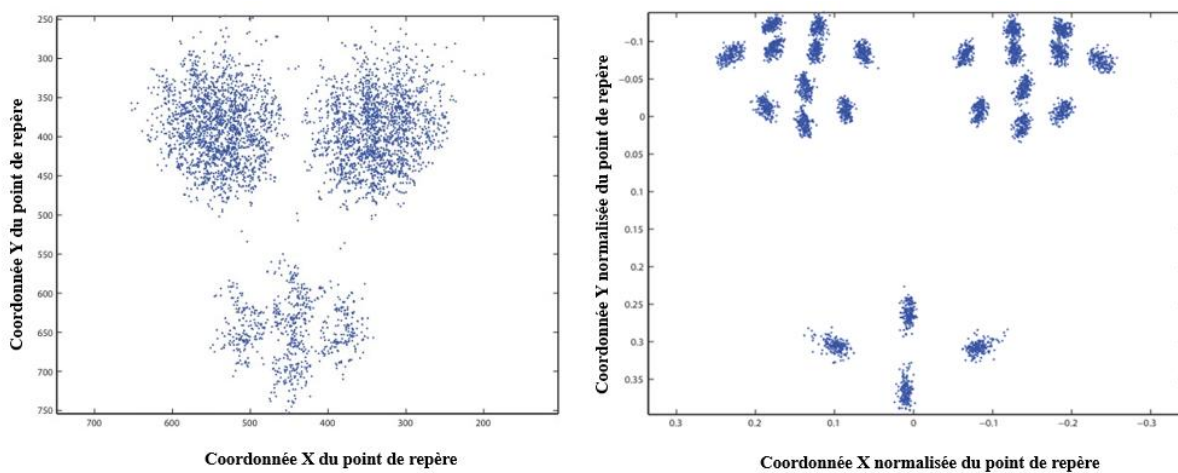
des informations émotionnelles [201]-[202], ce qui permet de reconnaître clairement différentes émotions en fonction des déformations spécifiques.

- Bloc de Normalisation des Caractéristiques

Dans cette étape, Generalized Procrustes Analysis (GPA) est utilisée pour accomplir les tâches suivantes :

- Normaliser les points de repère extraits ;
  - Éliminer les effets de l'échelle, de la rotation et de la translation dans l'ensemble des points de repère ;
- Réduire les variations entre les points de repère correspondants et rechercher le meilleur ajustement.

La GPA est une méthode statistique multivariée standard largement appliquée dans l'analyse des formes. Pour trouver la superposition optimale de deux ou plusieurs configurations, la GPA implique trois transformations : la translation, la rotation et la mise à l'échelle [203]. La (Figure III.2) montre les caractéristiques extraites de l'expression de surprise en utilisant la base de données MUG. La (Figure III.2 (a)) présente les données brutes et la (Figure III.2 (b)) montre les données normalisées avec la GPA.



**Figure III.2: Normalisation des repères.**

**(a) Données brutes (b) Normalisation avec GPA.**

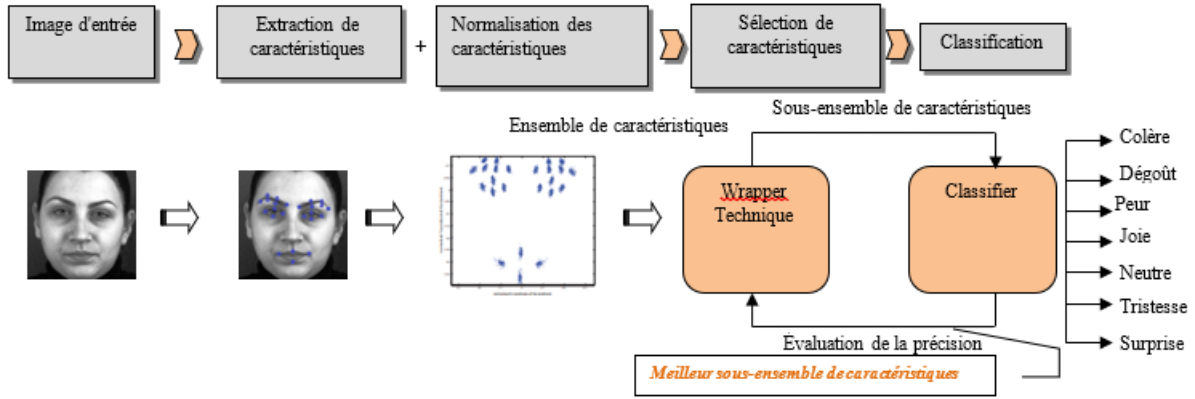
- Bloc de Classification des Expressions Faciales :

La dernière étape de ce système REF est la classification des expressions faciales, réalisée dans [200] en utilisant plusieurs classificateurs, à savoir les classificateurs KNN [204], NB [205], LDA [206] et DT [145].

Tout d'abord, l'algorithme KNN, l'un des algorithmes les plus populaires pour la reconnaissance de motifs, prédit la catégorie de l'échantillon de test en fonction des  $K$  échantillons d'entraînement qui lui sont les plus proches voisins, puis détermine la catégorie en sélectionnant celle avec la probabilité la plus élevée [207]. Ensuite, NB est un classificateur statistique basé sur le théorème de Bayes, attribuant la classe la plus probable à un exemple donné décrit par son vecteur de caractéristiques. Troisièmement, LDA consiste à trouver le plan de projection qui minimise la variance interclasses et maximise la distance entre les moyennes projetées des classes [208]. Enfin, l'algorithme de classification DT établit la relation entre les attributs d'entrée et les attributs de sortie pour construire un modèle qui prédit la classe désirée avec la plus haute précision [209].

### III.2.1. Principes du système REF proposé

La (Figure III.3) montre le schéma proposé du système REF. Comme illustré dans cette figure, notre principale contribution est l'adoption d'une phase de sélection de caractéristiques pour le système REF initial afin de déterminer dans quelle mesure sa performance globale peut être améliorée.



**Figure III.3: Aperçu schématique de notre système REF proposé.**

Selon ce système REF, en utilisant les méthodes ASM et GPA, 48 caractéristiques ( $48 = 24 \times 2$  : représentant les coordonnées cartésiennes de 24 points de repère) sont extraites de l'image faciale et normalisées. Le vecteur de caractéristiques est ensuite construit selon l'équation suivante :

$$F_{\overline{ij}} = (f_1, f_2, \dots, f_n) \quad (\text{III.1})$$

Où :

$i$  et  $j$  désignent l'ensemble des points de repère  $i^{\text{ème}}$  pour l'expression faciale  $j^{\text{ème}}$ ;  $n = 24$  est le nombre de points de repère extraits ;  $f = (x, y)$  sont les coordonnées cartésiennes du point de repère  $f$ .

Les 48 caractéristiques locales, résultant de l'extraction des caractéristiques, sont données comme suit :

$$\begin{aligned} & m_{x1}, m_{y1}, m_{x2}, m_{y2}, m_{x3}, m_{y3}, m_{x4}, m_{y4} \\ & le_{x1}, le_{y1}, le_{x2}, le_{y2}, le_{x3}, le_{y3}, le_{x4}, le_{y4} \\ & re_{x1}, re_{y1}, re_{x2}, re_{y2}, re_{x3}, re_{y3}, re_{x4}, re_{y4} \\ & lw_{x1}, lw_{y1}, lw_{x2}, lw_{y2}, lw_{x3}, lw_{y3}, lw_{x4}, lw_{y4}, lw_{x5}, lw_{y5}, lw_{x6}, lw_{y6} \\ & rw_{x1}, rw_{y1}, rw_{x2}, rw_{y2}, rw_{x3}, rw_{y3}, rw_{x4}, rw_{y4}, rw_{x5}, rw_{y5}, rw_{x6}, rw_{y6} \end{aligned}$$

Avec :

$(m_{xi}, m_{yi})$  sont les coordonnées cartésiennes des 4 points de repère positionnés sur « la bouche » ;  $(le_{xi}, le_{yi})$  sont les coordonnées cartésiennes des 4 points de repère positionnés sur « l'œil gauche » ;  $(re_{xi}, re_{yi})$  sont les coordonnées cartésiennes des 4 points de repère positionnés sur « l'œil droit » ;  $(lw_{xi}, lw_{yi})$  sont les coordonnées cartésiennes des 6 points de repère positionnés sur « le sourcil gauche » ;  $(rw_{xi}, rw_{yi})$  sont les coordonnées cartésiennes des 6 points de repère positionnés sur « le sourcil droit ».

Après l'extraction et la normalisation des caractéristiques, dans notre système REF proposé, nous avons intégré une approche de sélection pour choisir les caractéristiques les plus pertinentes. Il existe trois principales approches de sélection de caractéristiques couramment discutées dans la littérature : les approches « Filter », les approches « Wrapper » et les approches « Embedded » [210][211] [212]. Les approches basées sur le filtre, qui sélectionnent les caractéristiques indépendamment de tout

classificateur, évaluent les caractéristiques en fonction de leurs propriétés statistiques ou de leur corrélation avec les variables de sortie sans tenir compte de l'interaction avec le classificateur. Ces approches fournissent un moyen rapide et facile d'éliminer les caractéristiques non pertinentes ou redondantes avant d'appliquer des techniques de modélisation plus complexes.

Les approches basées sur le Wrapper consistent à sélectionner un sous-ensemble optimal de caractéristiques en minimisant directement l'erreur de classification. Contrairement aux approches basées sur le filtre, les approches Wrapper évaluent les sous-ensembles de caractéristiques en utilisant un classificateur spécifique, prenant ainsi en compte l'interaction entre les sous-ensembles de caractéristiques et les performances du classificateur. Ce processus itératif permet une sélection plus précise des caractéristiques qui contribuent directement à la performance du classificateur. Les approches Embedded intègrent la sélection des caractéristiques directement dans le processus d'entraînement du modèle, optimisant simultanément les paramètres du modèle et la sélection des caractéristiques [167] -[213]. Dans ce travail, nous proposons l'utilisation de l'approche « Wrapper » qui peut sélectionner les caractéristiques les plus pertinentes, plus robustes aux variations telles que les changements d'angle, d'illumination et d'occultation présents dans les données réelles [214], ainsi qu'aux différences inter-personnes dans l'expression émotionnelle [215] [216]. En effet, la pertinence des caractéristiques peut aider à résoudre des problèmes tels que l'extraction insuffisante des caractéristiques et la représentation inexacte des classes. Elle peut également atténuer le problème de la dépendance à une seule mesure de similarité en permettant une exploration plus exhaustive de l'espace des caractéristiques et en sélectionnant les ensembles de caractéristiques les mieux adaptés à la tâche de classification spécifique [217].

L'approche wrapper peut être utilisée de manière dynamique pour sélectionner les meilleures caractéristiques à chaque étape de l'entraînement du modèle, permettant ainsi au modèle de s'adapter continuellement aux nouvelles données et aux variations inter-domaines qui peuvent surgir, augmentant ainsi son efficacité [218]. Dans le cadre d'une stratégie de recherche, cette approche implique la sélection d'un algorithme qui recherche le sous-ensemble optimal de caractéristiques en fonction d'une fonction objectif donnée. Cela transforme le problème de sélection de caractéristiques en un problème d'optimisation, où l'objectif est généralement de maximiser la précision de classification tout en réduisant la taille du sous-ensemble de caractéristiques correspondant. Il existe de nombreuses stratégies de recherche pouvant être utilisées dans l'approche de sélection wrapper. Celles-ci peuvent être regroupées en trois catégories : (i) les stratégies de complexité exponentielle, (ii) les stratégies basées sur une approche populationnelle et (iii) les stratégies de sélection séquentielle [219]. Dans notre étude, nous avons introduit une approche de sélection qui combine le principe des stratégies de complexité exponentielle avec celui des stratégies de sélection séquentielle. Plus précisément, nous évaluons chaque sous-ensemble probable de caractéristiques pour déterminer le sous-ensemble optimal. En même temps, nous utilisons la Sélection Avant Séquentielle (SFS), une méthode itérative qui intègre systématiquement des caractéristiques d'un ensemble candidat. Cette double stratégie exploite la robustesse des méthodes exponentielles pour atteindre des solutions optimales et tire parti de l'efficacité des stratégies séquentielles pour améliorer le processus de sélection des caractéristiques. Ce qui suit décrira le principe de notre algorithme de sélection proposé. Le diagramme de flux illustré dans la (Figure III.4) montre les trois premières itérations du processus de sélection des caractéristiques intégré dans le système REF proposé.

Comme le montre cette figure, le processus commence avec un ensemble vide dépourvu de toute caractéristique. À chaque itération, chaque caractéristique de l'ensemble initial est testée pour évaluer son impact sur la performance globale du système. Ensuite, la caractéristique la mieux performante est ajoutée à l'ensemble vide et le processus se poursuit en testant les combinaisons de cet ensemble résultant avec les caractéristiques restantes. Cette méthode a déjà été utilisée dans d'autres travaux [220], [221], [222] et a donné de bons résultats pour les applications de reconnaissance de formes. La procédure générale de cette méthode peut être décrite comme suit :

**Étape 1 : Initialisation**

**Ensemble de caractéristiques :** commencer avec un ensemble vide de caractéristiques, 'S', à sélectionner dans l'ensemble initial de caractéristiques, 'F' ;

Ensemble  $F \leftarrow$  "Ensemble initial de n caractéristiques" ;

$S \leftarrow$  "Sous-ensemble vide" ;

**Modèles de base :** KNN, DT, NB, LDA ;

**Critère de performance :** maximisation de la précision ;

**Étape 2 : première caractéristique à sélectionner**

1. Testez chaque caractéristique individuellement :
  - Entraîner le modèle avec chaque caractéristique individuellement ;
  - Évaluer les performances du modèle pour chaque cas :  
 $\forall f_i \in F$ , calculer la précision en utilisant  $f_i$  ;
  - Sélectionnez la caractéristique, ' $f_{s1}$ ', qui donne les meilleures performances ;
  - Ajoutez la meilleure caractéristique ' $f_{s1}$ ' au sous-ensemble 'S' et retirez-le de l'ensemble 'F' ;  
 $F \leftarrow F - \{f_{s1}\}$ ;  
 $S \leftarrow \{f_{s1}\}$ .

**Étape 3 : Ajout progressif de caractéristiques**

1. Combinaison avec la meilleure caractéristique :
  - Pour chaque caractéristique  $f_i$  de l'ensemble  $F$ , combinez-la avec le sous-ensemble  $S$  ;
  - Entraînez le modèle sur le nouvel ensemble  $S \cup \{f_i\}$  ;
  - Évaluer les performances et ajouter la caractéristique qui apporte la plus grande amélioration ;
2. Répétition du processus de sélection des caractéristiques jusqu'à ce que toutes les caractéristiques soient testées :
  - Continuez à ajouter des caractéristiques une par une, en testant chaque combinaison, jusqu'à ce que toutes les caractéristiques aient été testées.

Notez ici que notre approche basée sur le wrapper sélectionne soigneusement les caractéristiques pendant la phase d'entraînement et applique ces mêmes colonnes de caractéristiques à l'ensemble de test sans re-sélectionner les caractéristiques pour ce dernier. Cette approche garantit que les métriques de performance obtenues pendant les tests reflètent la véritable généralisabilité des caractéristiques sélectionnées et du modèle entraîné. La cohérence dans la sélection et l'application des colonnes de caractéristiques vise à fournir des prédictions fiables et précises pour les données non vues, renforçant ainsi la validité de notre méthode pour des applications en temps réel et des bases de données variées.

Pour la tâche de classification, nous avons utilisé les classificateurs KNN, NB, LDA et DT afin de trouver la performance de reconnaissance optimale. Bien que l'étude de [200] rapporte que LDA donne de meilleurs résultats, nous avons décidé de tester divers classificateurs, car l'efficacité de notre approche de sélection de caractéristiques est directement liée à la performance du classificateur. En effet, en éliminant systématiquement les caractéristiques non pertinentes ou redondantes, notre approche pourrait atteindre une meilleure performance en termes de temps de réponse du système tout en maintenant une précision de reconnaissance relativement acceptable lorsqu'on utilise des classificateurs autres que LDA. Par conséquent, l'évaluation de l'approche proposée en utilisant tous les classificateurs présente deux avantages :

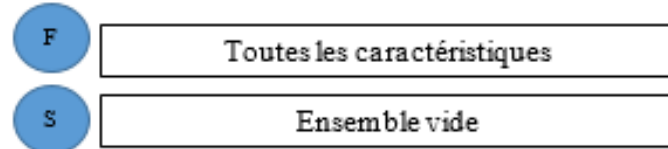
1. Il valide les résultats rapportés dans [200] ;
2. Il assure une évaluation approfondie des classificateurs pour déterminer la meilleure combinaison, ce qui peut potentiellement conduire à un système REF plus efficace et réactif.

F: ensemble de fonctionnalités " f "

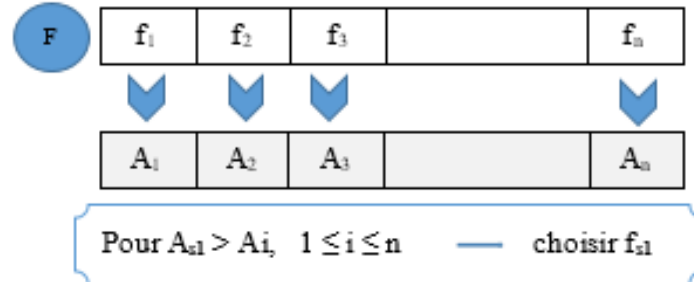
A : Précision du classificateur

S: sous-ensemble de caractéristiques pertinentes

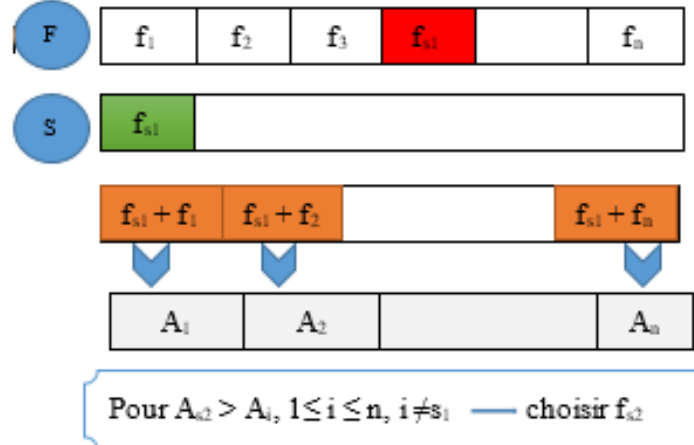
State i=0 (État initial):



État i=1:



État i=2:



État i=3

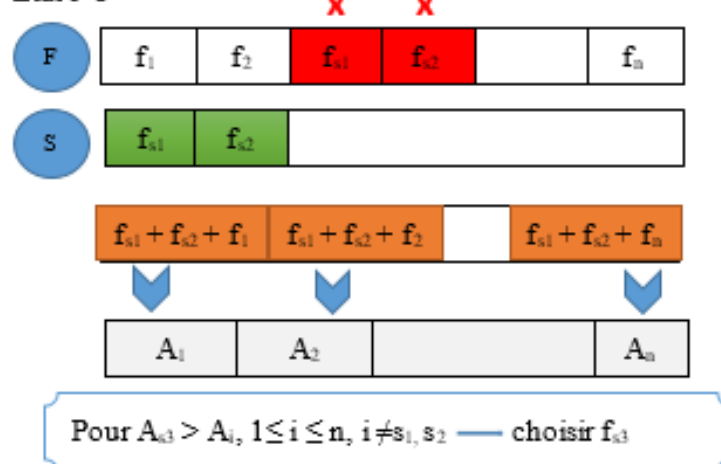


Figure III.4: Les trois premières itérations du processus de sélection des caractéristiques intégrées dans notre système REF proposé.

Bien que la plupart des travaux récents dans le domaine de la reconnaissance des expressions faciales se concentrent sur les approches basées sur l'apprentissage profond, notre travail se concentre sur les modèles de classification traditionnels. L'objectif principal est d'évaluer dans quelle mesure notre approche de sélection de caractéristiques peut améliorer la performance du système de reconnaissance tout en explorant différents types de modèles. Une fois que nous aurons démontré l'efficacité de cette approche avec divers classificateurs, nous pourrions envisager de généraliser son utilisation aux approches basées sur l'apprentissage profond.

Dans la section suivante, nous montrerons la contribution de cette méthode au système REF proposé en comparant les résultats obtenus avant et après son ajout.

### III.2.2. Résultats expérimentaux et discussion

Pour tester la performance du système REF proposé, il est évalué sur la base de données MUG. Cette dernière, proposée par [223], se compose de séquences d'images de 86 sujets (35 femmes et 51 hommes), âgés de 20 à 35 ans et d'origine caucasienne, effectuant des expressions faciales. Notre paradigme expérimental est réalisé en mode indépendant des personnes, où aucun des sujets utilisés pour les tests n'apparaît dans l'ensemble d'entraînement. Les expériences ont été menées sur la base de 1260 images (180 images pour chaque expression faciale), avec 840 images pour la phase d'entraînement (560 images d'hommes et 280 images de femmes) et 420 images pour les tests (280 images d'hommes et 140 images de femmes).

Concernant l'environnement de travail, le (Tableau III.2) présente les principales spécifications et la configuration logicielle du système utilisé pour notre travail.

**Tableau III.2: Spécifications techniques et environnement.**

| Logiciel / Matériel            | Détails des spécifications                        |
|--------------------------------|---------------------------------------------------|
| <b>Système</b>                 | Windows 10 Professionnel, 64 bits                 |
| <b>Processeur</b>              | Intel (R) Core (TM) i5-4300U à 2,49 GHz           |
| <b>RAM</b>                     | 4 GB                                              |
| <b>Environnement</b>           | MATLAB R2023a                                     |
| <b>Boîte à outils utilisée</b> | Statistics and Machine Learning Toolbox de MATLAB |

L'analyse et la discussion des résultats obtenus sont réparties en trois sous-sections, présentées ci-dessous.

#### III.2.2.1. L'impact de l'ajout d'une approche de sélection de caractéristiques sur la performance du système REF

Comme expliqué précédemment dans le principe du système REF proposé, après avoir appliqué l'approche de sélection de caractéristiques wrapper à chaque vecteur de caractéristiques (résultant de l'extraction des caractéristiques locales pour chaque image de la base de données en utilisant la méthode ASM), quatre modèles de classification sont utilisés pour classer les images de test dans l'une des sept classes universelles d'expressions faciales. Les résultats, obtenus en utilisant les caractéristiques les plus pertinentes sélectionnées par les classificateurs KNN, DT, NB et LDA, sont présentés dans les Tableaux (III.3), (III.4), (III.5) et (III.6), respectivement.

Ici, la métrique de performance utilisée pour évaluer la performance de la classification est la certitude, qui représente le nombre d'éléments de données correctement prédits parmi toutes les données. Elle est donnée par :

$$Accuracy = \frac{\sum_{i=1}^n \sum_{j=1}^n [prediction(i,j)]_{li=j}}{\sum_{i=1}^n \sum_{j=1}^n [prediction(i,j)]} \times 100 \quad (III.2)$$



Avec  $i$  : la classe prédite ;  $j$  : la classe correcte ;  $n$  : le nombre de classes.

La classification, en utilisant le classificateur KNN, a été réalisée en mesurant la distance Euclidienne avec différentes valeurs du plus proche voisin  $K$ .

Comme le montre le (Tableau III.3), pour le classificateur KNN, la meilleure la certitude, avec un pourcentage de 85,48 %, est obtenue pour  $K = 1$ . Les résultats de la classification, obtenus en utilisant le classificateur DT (selon différents nombres minimaux d'observations des nœuds feuilles), ont atteint une certitude maximale de 70,48 % pour une taille minimale de feuille égale à 3 (voir (Tableau III.4)). De même, une précision de 69,29 % est obtenue en utilisant le classificateur NB avec distribution Normale (Gaussienne) (voir (Tableau III.5)). Enfin, les résultats de la classification utilisant LDA (pour trois types discriminants ; à savoir, linéaire, diagLinear et diagQuadratic) sont présentés dans le (Tableau III.6). Comme le montre ce tableau, le meilleur résultat, avec une certitude de 100 %, est obtenu en utilisant une fonction discriminante linéaire.

**Tableau III.3: Résultats de classification en utilisant le classificateur KNN.**

| Valeur de $K$ | Certitude maximale (%)            |                           |
|---------------|-----------------------------------|---------------------------|
|               | Système REF de Silva et al. [200] | Notre système REF propose |
| 1             | 75.23%                            | <b>85. 48%</b>            |
| 3             | 76.42%                            | <b>83. 57%</b>            |
| 5             | 75.71%                            | <b>83. 57%</b>            |
| 7             | 75.23%                            | <b>84. 29%</b>            |
| 9             | 75.95%                            | <b>83. 33%</b>            |

**Tableau III.4: Résultats de classification utilisant le classificateur DT.**

| Nombre minimum d'observations du nœud de branche | Certitude maximale (%)            |                           |
|--------------------------------------------------|-----------------------------------|---------------------------|
|                                                  | Système REF de Silva et al. [200] | Notre système REF propose |
| 3                                                | 58.09%                            | <b>70. 48%</b>            |
| 5                                                | 58.80%                            | <b>68. 09%</b>            |
| 8                                                | 58.33%                            | <b>66. 43%</b>            |
| 10                                               | 58.57%                            | <b>69. 52%</b>            |
| 15                                               | 58.10%                            | <b>70%</b>                |

**Tableau III.5: Résultats de classification utilisant le classificateur NB.**

| Distribution | Certitude maximale (%)            |                                   |
|--------------|-----------------------------------|-----------------------------------|
|              | Système REF de Silva et al. [200] | Notre système REF propose         |
| Normal       | 58.33%                            | <b>69. 29%</b> (with 16 features) |
| Kernel       | 58.57%                            | <b>69. 28%</b> (with 18 features) |

**Tableau III.6: Résultats de classification utilisant le classificateur LDA.**

| Type de discriminant | Certitude maximale (%)            |                           |
|----------------------|-----------------------------------|---------------------------|
|                      | Système REF de Silva et al. [200] | Notre système REF proposé |
| Linear               | 99.71%                            | <b>100%</b>               |
| diagLinear           | 62.34%                            | <b>70. 48%</b>            |
| diagQuadratic        | 60.07%                            | <b>69. 29%</b>            |

Un résumé de la comparaison des meilleurs résultats entre le système REF de Silva et al. [200] et le système REF proposé dans ce travail pour les quatre classificateurs mentionnés est présenté dans le (Tableau III.7). Comme le montre ce tableau, le système REF proposé offre de meilleures performances dans tous les cas. En effet, comme illustré dans le même tableau, en intégrant l'approche de sélection de caractéristiques "Wrapper" dans le système traditionnel, le système proposé réalise une amélioration de la précision comprise entre 0,3 % et 12 %. De plus, il réduit la longueur du sous-ensemble de caractéristiques sélectionnées, ce qui se traduit par une réduction du nombre de caractéristiques entre 50 % et 77,08 %, par rapport au système original, pour une certitude presque identique ou même meilleure.

Des simulations utilisant l'outil MATLAB ont été réalisées pour évaluer le temps de réponse des systèmes basés sur les quatre classificateurs REF avant et après l'intégration de notre approche de sélection de caractéristiques. Comme le montre le (Tableau III.8), sachant que le classificateur NB a le plus grand temps de réponse (qui représente 100 %), les résultats démontrent une amélioration considérable du temps de réponse pour tous les classificateurs évalués. Plus précisément, le temps de réponse du classificateur KNN a diminué de 58,82 % à 4,31 %, suivi par le DT avec un temps de réponse de 4,73 % après l'intégration de notre approche de sélection de caractéristiques. Des temps de réponse similaires sont observés pour NB (5,66 %) et LDA (4,63 %), illustrant que la réduction de la complexité des données grâce à la sélection de caractéristiques peut considérablement accélérer les processus de classification tout en maintenant des niveaux de performance comparables.

**Tableau III.7: Comparaison des performances entre le système de REF de Silva et al. [200] et le système proposé.**

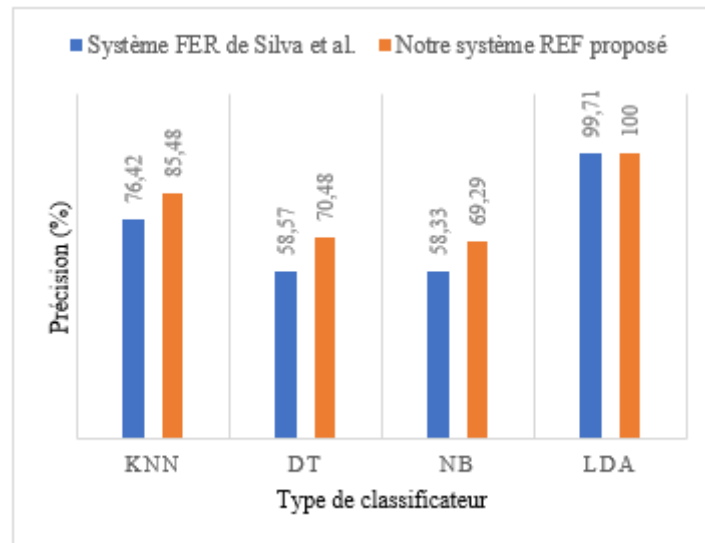
| Classificateur | Système REF de Silva et al. [200] |                                | Notre système REF propose |                                | Amélioration de la certitude | Taux de réduction du nombre de caractéristiques |
|----------------|-----------------------------------|--------------------------------|---------------------------|--------------------------------|------------------------------|-------------------------------------------------|
|                | Certitude (%)                     | Dimension des caractéristiques | Certitude (%)             | Dimension des caractéristiques |                              |                                                 |
| KNN            | 76.42%                            | 48                             | 85.48%                    | 16                             | 9%                           | 66.67%                                          |
| DT             | 58.57%                            |                                | 70.48%                    | 11                             | 11.91%                       | 77.08%                                          |
| NB             | 58.33%                            |                                | 69.29%                    | 16                             | 10.96%                       | 66.67%                                          |
| LDA            | 99.71%                            |                                | 100%                      | 24                             | 0.29%                        | 50%                                             |

**Tableau III.8: Comparaison des temps de réponse entre le système REF de Silva et al. [200] et le système proposé.**

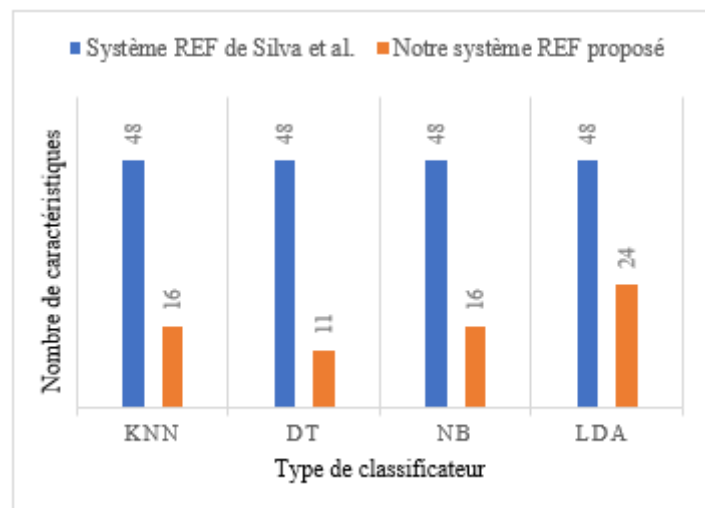
| Classificateur | Temps de réponse en %             |                           |
|----------------|-----------------------------------|---------------------------|
|                | Système REF de Silva et al. [200] | Notre système REF propose |
| KNN            | 58.82%                            | 4.31%                     |
| DT             | 67.49%                            | 4.73%                     |
| NB             | 100%                              | 5.66%                     |
| LDA            | 74.15%                            | 4.63%                     |

La (Figure III.5) révèle comment notre système REF proposé surperforme, en termes de certitude et de nombre de caractéristiques traitées, par rapport au système original [200]. En effet, l'augmentation de la certitude est clairement observable avec une réduction du nombre de caractéristiques utilisées dans le processus de reconnaissance. Par conséquent, nous pouvons conclure que la sélection de caractéristiques

contribue à améliorer la précision de la reconnaissance grâce à l'élimination des caractéristiques redondantes et non pertinentes, tout en réduisant le temps d'entraînement du classificateur.



**Figure III.5: (a) Comparaison de la précision entre le système REF de Silva et al. [200] et le système proposé.**



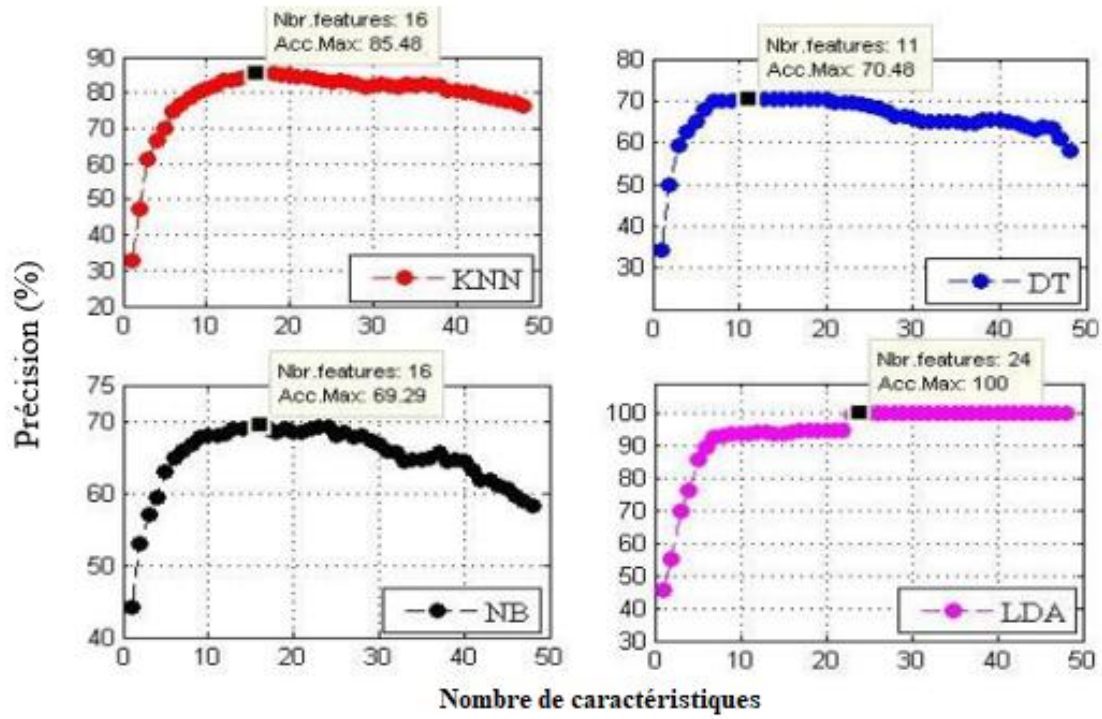
**Figure III.5: (b) Comparaison des dimensions des caractéristiques entre le système REF de Silva et al. [200] et le système proposé.**

### III.2.2.2. Analyse du phénomène de la malédiction de la dimensionnalité

Le terme "malédiction de la dimensionnalité" a été introduit pour la première fois par Richard E. Bellman [224] pour décrire le phénomène de l'augmentation exponentielle de la quantité de données avec la dimensionnalité. D'un point de vue théorique, augmenter la dimension des données ajoute plus d'informations et améliore donc la performance. Cependant, dans la pratique, ce principe n'est pas toujours valide. En effet, dans la plupart des cas, l'augmentation de la dimension entraîne une augmentation du bruit et de la redondance pendant l'opération d'analyse.

Dans ce qui suit, nous étudierons l'impact de ce phénomène sur les résultats du système REF original. De plus, nous examinerons comment l'approche de sélection "Wrapper" peut éliminer ce phénomène et

ainsi améliorer la performance globale. Les variations de la certitude résultantes, en fonction du nombre de caractéristiques, pour les quatre classificateurs sont montrées dans la (Figure III.6). Dans cette figure, nous représentons la réponse du système REF, basée sur chaque sous-ensemble de caractéristiques sélectionné  $S_j$  (où  $j$  est le numéro d'itération variant de 1 à 48), en utilisant chacun des quatre classificateurs.



**Figure III.6: Variation de la certitude en fonction du nombre de caractéristiques**

**En haut à gauche :** classificateur KNN ( $K=1$ ), **En haut à droite :** classificateur DT (nombre d'observations de nœuds de branche égal à 3), **En bas à gauche :** classificateur NB (distribution : normale), **En bas à droite :** classificateur LDA (type de discriminant : linéaire).

Comme le montre cette figure, pour les classificateurs KNN, DT et NB, chacune des courbes de la certitude correspondantes atteint sa valeur maximale pour un nombre spécifique de caractéristiques, puis diminue à nouveau. Ce comportement est connu sous le nom de phénomène de malédiction de la dimensionnalité ou phénomène de pic. L'approche de sélection des caractéristiques pertinentes permet de surmonter ce problème et, par conséquent, d'améliorer la performance du système REF. En effet, comme le montre cette figure, le système REF, basé sur le classificateur KNN, atteint son certitude maximale (égale à 85,48 %) en exploitant seulement 16 caractéristiques. Pour le classificateur DT, ce même système a atteint une certitude maximale de 70,48 % en utilisant seulement 11 caractéristiques. Enfin, pour les classificateurs NB et LDA, le système REF a atteint des valeurs de 69,29 % et 100 %, respectivement, en utilisant 16 et 24 caractéristiques. Pour le classificateur LDA, comme illustré dans cette figure, la courbe atteint son plateau à 24 caractéristiques sélectionnées avec une certitude de 100 %. Par conséquent, ce résultat démontre comment le système REF présente une meilleure performance après l'utilisation de l'approche de sélection "Wrapper" et du classificateur LDA.

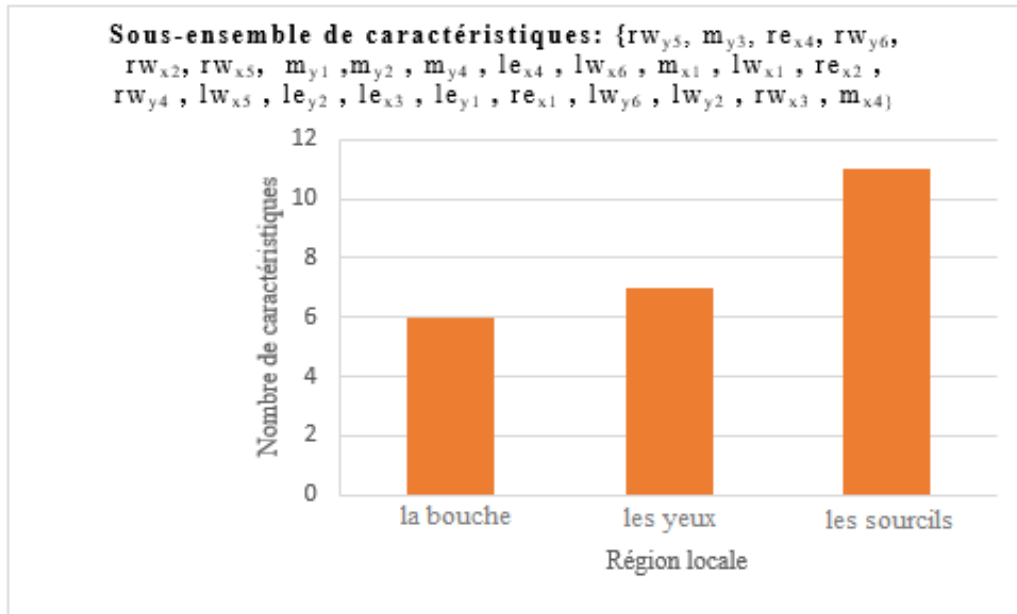
Le (Tableau III.9) fournit plus de détails sur l'opération de sélection des caractéristiques. En particulier, il présente les dix premières caractéristiques pertinentes ainsi que les valeurs de certitude de reconnaissance, correspondant aux différents sous-ensembles  $S_j$ , obtenues en utilisant le classificateur LDA. Comme le montre ce tableau, l'approche de sélection a atteint une certitude d'environ 93 % avec une réduction dimensionnelle du vecteur de caractéristiques de 83 % (8/48), ce qui prouve son efficacité pour le processus de reconnaissance des expressions faciales.

**Tableau III.9: Les dix premières caractéristiques sélectionnées “ $f_{si}$ ” et leurs valeurs de certitude obtenues en utilisant le classificateur LDA pour le cas optimal.**

| Certitude maximale = 100% obtenue en utilisant un nombre optimal de caractéristiques sélectionnées égal à 24. |           |              |               |               |               |               |              |              |              |               |
|---------------------------------------------------------------------------------------------------------------|-----------|--------------|---------------|---------------|---------------|---------------|--------------|--------------|--------------|---------------|
| $i$                                                                                                           | 1         | 2            | 3             | 4             | 5             | 6             | 7            | 8            | 9            | 10            |
| $f_{si}$                                                                                                      | $rw_{y5}$ | $m_{y3}$     | $re_{x4}$     | $rw_{y6}$     | $rw_{x2}$     | $rw_{x5}$     | $m_{y1}$     | $m_{y2}$     | $m_{y4}$     | $le_{x4}$     |
| Sub-set ( $S_i$ )                                                                                             | $rw_{y5}$ | $S_1+m_{y3}$ | $S_2+re_{x4}$ | $S_3+rw_{y6}$ | $S_4+rw_{x2}$ | $S_5+rw_{x5}$ | $S_6+m_{y1}$ | $S_7+m_{y2}$ | $S_8+m_{y4}$ | $S_9+le_{x4}$ |
| Certitude (%)                                                                                                 | 45.48%    | 55.24%       | 69.52%        | 75.95%        | 85.71%        | 89.05%        | 92.38%       | 93.10%       | 93.57%       | 93.57%        |

### III.2.2.3. Pertinence des régions faciales locales sur la performance du système REF

Pour déterminer la région locale la plus importante du visage, il est nécessaire d'étudier les caractéristiques locales les plus pertinentes, permettant au système d'atteindre une précision maximale.



**Figure III.7: Caractéristiques les plus pertinentes en fonction de leur position sur le visage en utilisant (classificateur LDA).**

Comme le montre la (Figure III.7), après une analyse des caractéristiques les plus pertinentes (en fonction de leur position sur le visage), nous constatons que la majorité d'entre elles appartiennent à la région locale des « sourcils » (avec 11 caractéristiques sur 24), suivie de la région des « yeux » (avec 7 caractéristiques sur 24) et enfin de la région de la « bouche » (avec 6 caractéristiques sur 24).

Ce résultat met en évidence l'importance de la région des « sourcils » dans le domaine de la reconnaissance des expressions faciales. En effet, les sourcils sont extrêmement mobiles et leur mouvement contribue de manière significative à l'expression de diverses émotions. La position et le mouvement des sourcils fournissent des indices subtils sur les intentions et l'état intérieur d'une personne, souvent avant les changements dans d'autres aspects du visage. Cette expressivité nuancée souligne le rôle crucial des sourcils dans la communication non verbale, en faisant un élément essentiel pour comprendre et reconnaître avec précision les expressions faciales. Malgré leur importance, ce concept a été négligé par les chercheurs dans ce domaine. En général, les études se concentrent principalement sur les trois principales régions faciales : la bouche, le nez et les yeux. Cependant, des

recherches futures pourraient améliorer les algorithmes en intégrant la dynamique des sourcils, ce qui pourrait conduire à une reconnaissance des expressions faciales plus précise.

### III.2.3. Comparaison avec les méthodes d'état de l'art

Lorsque nous comparons notre approche avec celles proposées dans [145], [225] [226][227], [228][31] [229], nous constatons que chaque méthode de sélection de caractéristiques a ses propres avantages et inconvénients. Le choix d'une approche spécifique dépend souvent du contexte d'application et des contraintes pratiques, telles que le compromis entre la précision et la complexité computationnelle acceptable.

- Notre approche : Elle montre une performance optimale en termes de reconnaissance en capturant les interactions complexes entre les caractéristiques. Cela est particulièrement efficace lorsque le nombre de caractéristiques n'est pas trop élevé, car le coût computationnel reste modéré. Par conséquent, notre méthode est particulièrement adaptée lorsque les ressources computationnelles sont disponibles et que l'atteinte de la précision maximale est cruciale.

-Relief-F et Minimum Redundancy Maximum Relevance (mRMR) : Ces méthodes sont bien adaptées pour gérer la redondance et le bruit des caractéristiques tout en étant moins coûteuses en termes de calcul. Elles trouvent un équilibre entre pertinence et redondance, ce qui les rend efficaces pour les grandes bases de données avec de nombreuses caractéristiques.

- Chi-square, GR et IG : Ces méthodes sont rapides et simples à mettre en œuvre. Cependant, elles peuvent manquer de sophistication pour capturer les interactions complexes entre les caractéristiques, ce qui pourrait entraîner une performance sous-optimale dans les scénarios où ces interactions sont critiques.

- SCA et Autoencodeur: Ces méthodes offrent des capacités d'exploration globale et la capacité de capturer des relations non linéaires, respectivement. Malgré leurs capacités avancées, ces méthodes nécessitent un réglage fin et sont intensives en calcul, ce qui pourrait limiter leur praticité pour des problèmes à grande échelle sans ressources computationnelles suffisantes.

Le (Tableau III.10) compare la performance de notre système REF proposé à celle des systèmes les plus pertinents rapportés dans la littérature.

**Tableau III.10: Comparaison entre les performances du système REF proposé et celles des systèmes les plus pertinents rapportés dans la littérature.**

| Ref.          | Base de données | Région d'intérêt                                            | Nbre d'émotions | Méthode FS                 | Type d'approche | Classificateur |
|---------------|-----------------|-------------------------------------------------------------|-----------------|----------------------------|-----------------|----------------|
| Notre travail | MUG             | Bouche, Yeux et Sourcils                                    | 7               | Notre approche FS proposée | Wrapper         | LDA            |
| [158]         | CK+             | Visage complet                                              | 8               | Relief-F                   | Filter          | KNN            |
| [167]         | SFE             | Visage complet                                              | 8               | SCA                        | Wrapper         | KNN            |
| [186]         | TFEID           | Visage complet divisé en régions régulières                 | 8               | Pairwise FS                | Hybrid          | SVM            |
| [159]         | CK+             | Régions de l'œil gauche et moitié de la région de la bouche | 8               | GR                         | Filter          | KNN            |

|       |          |                |   |                   |          |      |
|-------|----------|----------------|---|-------------------|----------|------|
| [160] | JAFFE    | Visage complet | 7 | RV<br>Correlation | Filter   | SVM  |
| [177] | JAFFE    | Visage complet | 7 | IUTWSVM           | Embedded | SVM  |
| [178] | Xception | Visage complet | 6 | Autoencoder       | Embedded | DCNN |

L'analyse réalisée ici montre que le système proposé, qui sélectionne des régions d'intérêt (ROIs) telles que la bouche, les yeux et les sourcils, atteint une certitude remarquable de 100 % en utilisant la base de données MUG. Contrairement à l'analyse de l'ensemble du visage, se concentrer sur ces régions clés permet de mieux capter les nuances émotionnelles, améliorant ainsi la robustesse et l'efficacité des modèles. Enfin, cette comparaison souligne l'efficacité prometteuse de la méthode proposée, même face à la diversité des approches existantes dans ce domaine.

### III.3. Un nouvel algorithme de reconnaissance des expressions faciales basé sur l'extraction et la sélection de caractéristiques par DWT

Dans cette deuxième partie, nous proposons un modèle efficace pour améliorer la précision et le coût computationnel d'un système REF. Ce modèle se déroule en trois étapes. La première, correspondant à l'extraction des caractéristiques, introduit trois descripteurs dérivés de la Transformation en Ondelette Discrète (DWT) pour extraire différents types de caractéristiques. Dans la deuxième étape, axée sur la sélection des caractéristiques, une approche Wrapper est adoptée pour sélectionner avec soin les caractéristiques les plus pertinentes parmi le pool extrait précédemment. Après la sélection des caractéristiques, le classificateur SVM est utilisé dans la dernière étape pour déterminer l'état affectif d'un individu.

En effet, plusieurs applications de la tâche REF ont pour objectif d'identifier l'état émotionnel d'une personne en classant son image faciale dans une classe d'émotion ciblée  $C_i$ : {colère, dégoût, mépris, peur, joie, tristesse, surprise, neutre} ; sinon, elle est classée dans une classe  $\bar{C}_i$  (dans le cas contraire). Par exemple, prendre en compte la peur dans un lieu public (y compris les aéroports, les banques, les hôpitaux, etc.) est un élément essentiel pour détecter le stress et ainsi assurer la sécurité (sécurité civile et surveillance). De même, la détection d'une émotion de joie permet de mesurer la satisfaction du client, ce qui aide les individus à s'adapter à leurs besoins (marketing), tandis que la détection d'une émotion de dégoût lors d'une session d'apprentissage à distance peut améliorer la qualité de la performance éducative (éducation). En outre, détecter le mépris ou la tristesse aide à identifier certaines maladies psychologiques (médecine) [230], [231], [232].

Bien qu'un système REF se compose généralement de trois processus principaux : détection du visage, extraction des caractéristiques et classification [233]. Selon la littérature, l'extraction des caractéristiques, qui est le processus le plus important [234] [235], consiste à transformer l'image en un ensemble de caractéristiques permettant de séparer différentes classes. Dans ce contexte, diverses méthodes ont été proposées dans la littérature scientifique. Les méthodes les plus couramment utilisées pour l'analyse des expressions faciales incluent la transformation de caractéristiques invariantes à l'échelle (SIFT) [236], les filtres de Gabor [237], les filtres log-Gabor [238], les motifs binaires locaux (LBP) [57] et l'histogramme des gradients orientés (HOG) [65]. Toutes ces méthodes ont prouvé leur efficacité pour la représentation d'images. Cependant, en raison de la grande dimensionnalité des caractéristiques extraites, l'utilisation directe de ces méthodes augmentera le temps de calcul et la mémoire nécessaire pour le stockage [59], [239]. Par exemple, la méthode de représentation par ondelettes de Gabor est capable de détecter les changements de texture multi-échelle et multi-direction [32]. Néanmoins, elle génère une matrice de coefficients Gabor de grande dimension composée de caractéristiques redondantes. Cette redondance des caractéristiques est responsable de l'augmentation de la confusion et, par conséquent, de la diminution de la précision du système REF [240].



En ce qui concerne l'utilisation des filtres log-Gabor pour l'extraction des caractéristiques, les mêmes limitations, en termes de complexité, peuvent être observées. En effet, selon l'étude réalisée dans [238], l'utilisation de ce type de filtres dans les systèmes REF a démontré que, malgré leur haute précision de reconnaissance, les filtres log-Gabor ont un inconvénient majeur qui réside dans la complexité et le temps de calcul. Cette charge computationnelle élevée les rend moins efficaces par rapport à d'autres méthodes, notamment pour certaines expressions faciales [238]. De même, la méthode LBP, qui attribue un code binaire à un pixel en fonction de son voisinage, surmonte les problèmes de déplacement, de rotation et d'illumination déséquilibrés dans une image. En outre, elle extrait suffisamment d'informations de texture permettant au classificateur de bien gérer le problème REF [241]. Cependant, selon [59], un petit voisinage ne peut pas capturer les caractéristiques dominantes avec des structures à grande échelle, tandis qu'un grand voisinage augmente la taille de l'histogramme de caractéristiques et augmente par la suite le coût computationnel de chaque image LBP [241]. Dans un autre exemple, la méthode HOG s'est avérée être un descripteur efficace pour préserver l'information locale en utilisant la distribution de la densité d'orientation et le gradient du contour [239]. Néanmoins, la méthode HOG est limitée par son coût computationnel élevé et la grande dimension de ses caractéristiques.

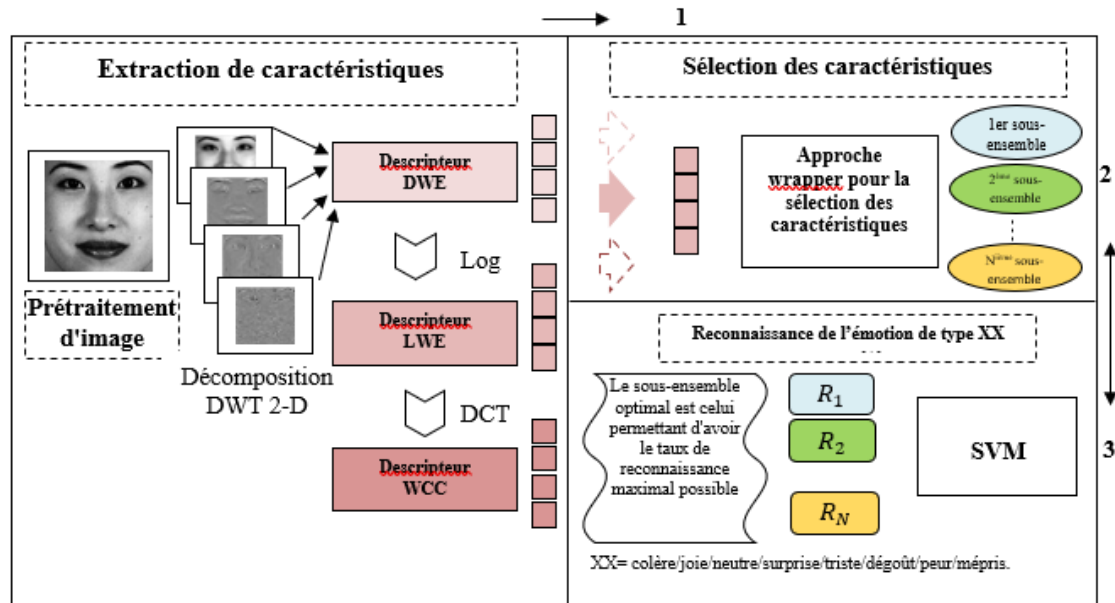
Pour remédier à tous ces problèmes, dans ce travail, nous proposons une nouvelle méthode d'extraction de caractéristiques plus compacte que de nombreuses méthodes existantes. Notre méthode, capable de réduire le temps de calcul et d'augmenter la précision de reconnaissance, consiste en l'utilisation de trois caractéristiques extraites par la Transformation en Ondelette Discrète (DWT). La procédure utilisée ici a été appliquée dans d'autres domaines de traitement de signal unidimensionnel (1-D), tels que le traitement de la parole [242], [243], l'identification d'appareils électriques [222] et l'identification de sujets sains et parkinsoniens [244], [245], mais n'a jamais été utilisée comme solution pour les problèmes de traitement d'images bidimensionnelles (2D), notamment le problème REF.

Pour améliorer l'efficacité, en termes de temps de réponse et de performance d'apprentissage de notre méthode proposée, elle est utilisée en conjonction avec le processus de sélection des caractéristiques. Ce processus consiste à sélectionner le sous-ensemble de caractéristiques qui discrimine le mieux entre les différentes classes d'émotion. La procédure utilisée ici a deux objectifs principaux. Le premier est de ne conserver que les caractéristiques les plus pertinentes et d'éviter les redondances afin d'augmenter la vitesse, en termes de temps de réponse, du système REF. Le second est de remédier à la malédiction de la dimensionnalité, qui a souvent des conséquences négatives sur le comportement et les performances des algorithmes d'apprentissage.

### **III.3.1. Le modèle proposé**

Une vue d'ensemble du modèle proposé, qui se compose d'un ensemble de processus et de tâches, est présentée dans la (Figure III.8). Dans ce qui suit, nous exposerons le principe de notre système de reconnaissance des expressions faciales proposé. Nous expliquerons donc en détail le principe de chacune de ces tâches.





**Figure III.8: Organigramme du système proposé.**

### III.3.1.1. Prétraitement des images

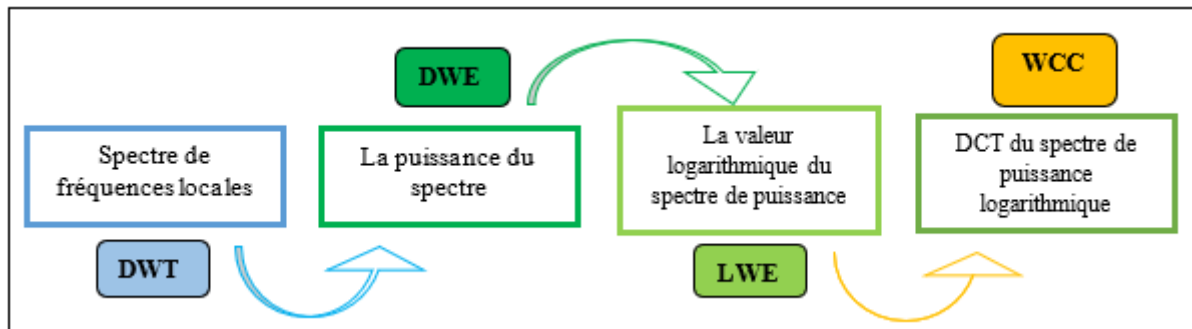
L'objectif de la tâche de prétraitement est d'améliorer la qualité d'une image et de la préparer pour une analyse ultérieure. Cette étape est cruciale pour réduire le bruit et les variations d'illumination. Dans ce travail, nous convertissons d'abord l'image originale en niveaux de gris pour faire face à la mauvaise qualité d'acquisition. Cette opération est souvent utilisée dans le prétraitement des images pour plusieurs raisons, notamment :

- L'élimination des variations de couleur peut simplifier l'analyse et rendre les caractéristiques pertinentes plus visibles.
- Dans une image en niveaux de gris, chaque pixel représente uniquement la luminance ou l'intensité lumineuse à cet endroit, permettant ainsi de conserver les informations sur la structure et les détails de l'image sans être affecté par les variations de couleur.

Ensuite, comme deuxième étape, nous découpons l'image originale en fonction de la position des deux yeux et la normalisons à une taille de  $200 \times 200$  pixels. Dans la littérature, plusieurs méthodes peuvent être appliquées pour la détection des visages, telles que la méthode de Viola-Jones ou d'autres basées sur des algorithmes de détection pré-entraînés. Dans notre travail, étant donné que l'objectif était principalement axé sur le développement de méthodes d'extraction et de sélection de caractéristiques pertinentes, nous avons opté pour une récupération manuelle à partir d'images sans bruit.

### III.3.1.2. Extraction de caractéristiques

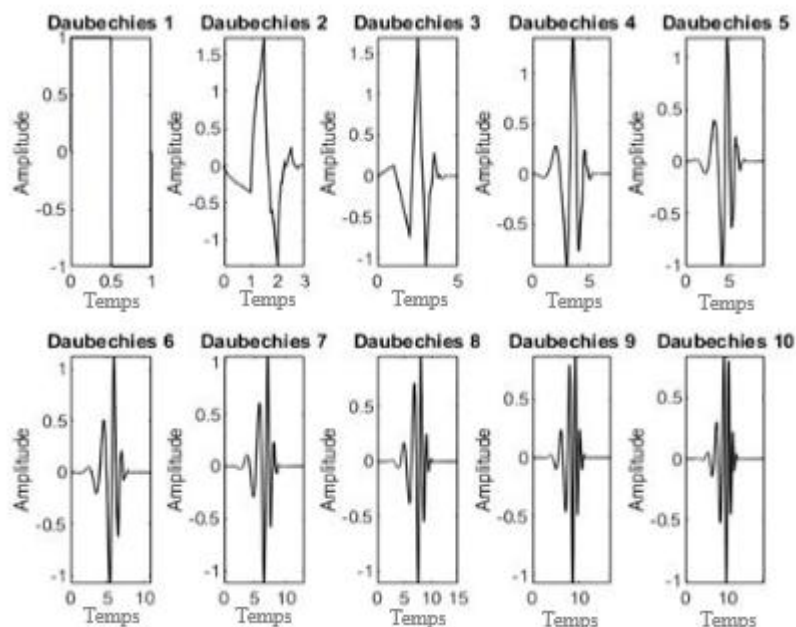
Un diagramme en blocs de notre méthode d'extraction de caractéristiques proposée est illustré dans la (Figure III.9).



**Figure III.9: Schéma fonctionnel de la méthode d'extraction de caractéristiques proposée.**

Comme illustré dans cette figure, cette méthode comprend quatre phases. La première consiste à effectuer une décomposition en ondelettes discrètes bidimensionnelle (2D-DWT) de l'image faciale, en utilisant une ondelette mère spécifique. Cette dernière est une fonction fenêtrée qui se déplace/décalle le long du signal de la série temporelle [242]. Les familles d'ondelettes varient en fonction de plusieurs propriétés importantes. On peut citer par exemple Daubechies, Biorthogonale, Coiflets, Symlets, et Morlet. Dans ce travail, trois fonctions d'ondelettes, à savoir Daubechies, Coiflet et Symlet, ont été utilisées pour réaliser cette méthode d'extraction de caractéristiques.

Comme le montrent les Figures (III.10), (III.11) et (III.12), les différences entre ces familles d'ondelettes résident principalement dans leurs propriétés de régularité, de symétrie et d'asymétrie, qui les rendent adaptées à différents types d'applications en traitement du signal et en analyse de données [246]. En réalité, les ondelettes Daubechies ont des moments de haute régularité, ce qui signifie qu'elles sont bien localisées à la fois dans le temps et la fréquence. Les ondelettes Coiflet, quant à elles, ont été développées comme une alternative aux ondelettes Daubechies, avec une certaine asymétrie et une régularité améliorée. Les ondelettes Symlet sont une extension symétrique des ondelettes Coiflet.



**Figure III.10: Famille des ondelettes de Daubechies.**

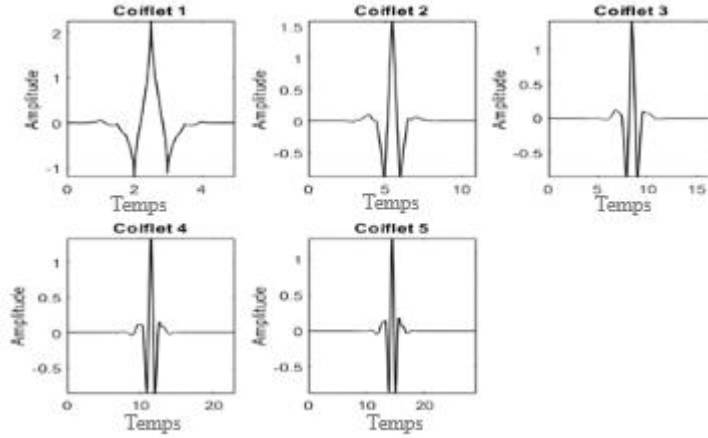


Figure III.11: Famille des ondelettes de Coiflet.

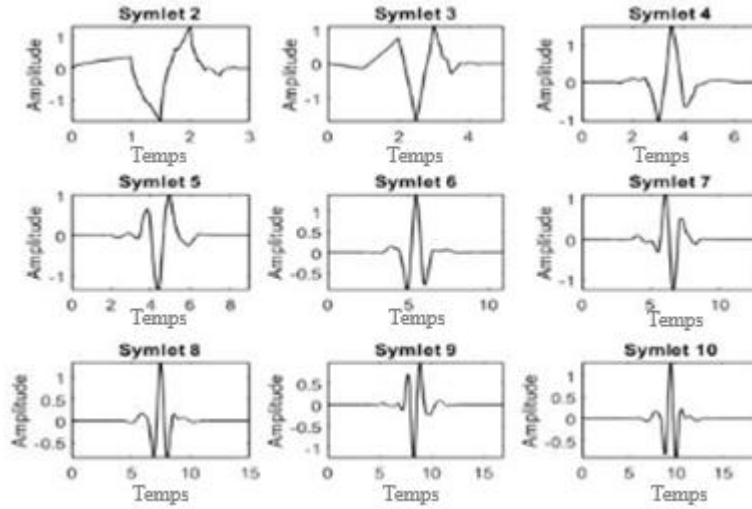


Figure III.12: Famille des ondelettes de Symlet.

Rappelons que la transformation en ondelettes discrètes (DWT) divise l'information en ses composantes de basse et haute fréquence. En effet, la composante de basse fréquence est une approximation passe-bas de l'image originale (variations douces) qui contient la plupart de l'énergie de l'image. Cependant, la composante de haute fréquence représente les composants de bord (qui fournissent les détails) et présente trois directions unidimensionnelles distinctes : horizontale, verticale et diagonale. De plus, la DWT peut être calculée à diverses résolutions ou échelles [247], [248].

Dans la configuration 2D-DWT, nous devons utiliser une fonction d'échelle  $\varphi(x, y)$  et trois ondelettes  $\psi^H(x, y)$ ,  $\psi^V(x, y)$  et  $\psi^D(x, y)$ . Chaque fonction d'échelle ou ondelette est le produit d'une fonction d'échelle unidimensionnelle  $\varphi$  et de l'ondelette correspondante  $\psi$ .

Les quatre produits en 2D produisent la fonction d'échelle donnée par l'Équation (III.3) et des ondelettes directionnelles séparables sensibles aux directions, données par les Équations (III.4, III.5 et III.6) [249].

$$\varphi(x, y) = \varphi(x)\varphi(y) \quad (III.3)$$

$$\psi^H(x, y) = \varphi(y)\psi(x) \quad (\text{horizontale}) \quad (III.4)$$

$$\psi^V(x, y) = \varphi(x)\psi(y) \quad (\text{verticale}) \quad (III.5)$$

$$\psi^D(x, y) = \psi(x)\psi(y) \quad (\text{diagonale}) \quad (\text{III.6})$$

Avec :

$\psi^H(x, y)$  : mesure les variations de niveaux de gris le long des colonnes (bords horizontaux)

$\psi^V(x, y)$  : mesure les variations de niveaux de gris le long des lignes (bords verticaux)

$\psi^D(x, y)$  : correspond aux variations le long des diagonales.

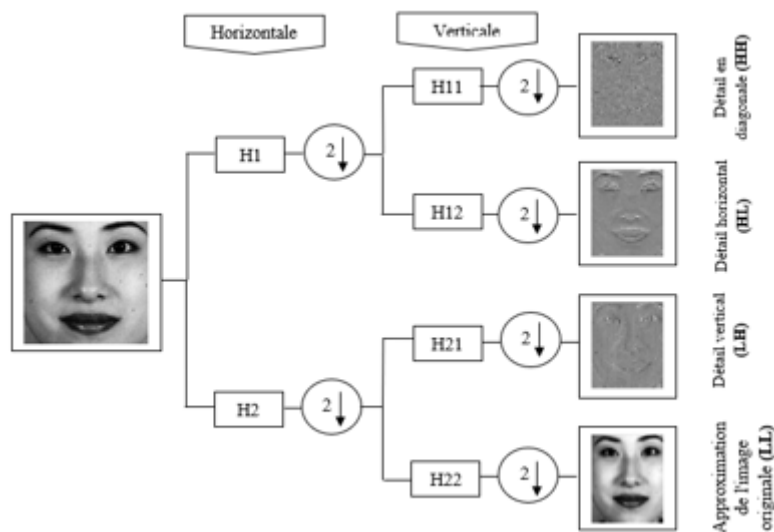
Il est à noter que la fonction ondelette agit comme un filtre passe-bande dont la largeur de bande est réduite de moitié après chaque échelle. De plus, pour une image en 2D, les lignes et les colonnes sont traitées de manière distincte pour extraire les différentes informations de fréquence.

Les colonnes sont traitées comme des signaux 1D. Le diagramme en blocs de la (Figure III.13) illustre la décomposition en ondelettes 2D sur une image d'entrée. Selon ce diagramme, lors de la première étape de la décomposition de niveau 1, chaque ligne de l'image est mise à l'échelle et transformée par ondelettes (transformation horizontale). Les résultats de cette étape sont deux demi-images, l'une avec les coefficients d'échelle (L) et l'autre avec les coefficients d'ondelettes (H). À l'étape suivante, les transformations d'échelle (L) et d'ondelettes (H) sont appliquées sur chaque colonne des deux demi-images de l'étape précédente (transformation verticale). Les résultats de cette étape sont quatre images de taille quart, à savoir : les coefficients d'approximation (LL), les détails verticaux (LH), les détails horizontaux (HL) et les détails diagonaux (HH). Au niveau suivant, lorsque la transformation par ondelettes est appliquée, les étapes ci-dessus seront répétées sur l'image LL [250]. Les fonctions de base mises à l'échelle et translatées sont définies par les Équations (III.7) et (III.8) données comme suit :

$$\Phi_{j,m,n}(x, y) = 2^{j/2} \varphi(2^j x - m, 2^j y - n) \quad (\text{III.7})$$

$$\psi^i_{j,m,n}(x, y) = 2^{j/2} \psi^i(2^j x - m, 2^j y - n) \quad (\text{III.8})$$

Où : l'indice  $i = \{H, V, D\}$  identifie les ondelettes directionnelles en termes de valeurs Horizontales, Verticales et Diagonales (H, V, D).



**Figure III.13: Décomposition 2D-DWT à un niveau sur une image d'entrée.**

Les coefficients d'approximation et de détail, correspondant à la fonction  $f(x, y)$  de taille  $M$  par  $N$ , sont formulés comme suit :

Coefficients d'approximation :

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j_0, m, n}(x, y) \quad (\text{III.9})$$

Coefficients de détail :

$$W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \psi_{j, m, n}^i(x, y) \quad (\text{III.10})$$

Où :

$j_0$  est la valeur de départ de l'échelle  $j$  ( $j \geq j_0$ );

$j = 0, 1, 2, \dots, J - 1$ ;

$2^J$  est la taille des données sélectionnées ;

$n = m = 0, 1, 2, \dots, 2^j - 1$ ;

Les coefficients donnés par les Équations (III.9 et III.10) sont ensuite utilisés pour proposer, à travers les trois phases suivantes de notre méthode d'extraction de caractéristiques, trois types de descripteurs. Le premier, appelé Discret Wavelet Energy (DWE), consiste à convertir chaque image faciale en un vecteur de caractéristiques. Le DWE est composé de l'énergie des coefficients d'ondelettes calculée à chaque niveau de décomposition de l'image. Le deuxième, appelé Log Wavelet Energy (LWE), consiste à appliquer le logarithme aux caractéristiques DWE. Le dernier descripteur, appelé Wavelet Cepstral Coefficient (WCC), consiste à appliquer la Transformée en Cosinus Discrète (DCT) au vecteur de caractéristiques LWE.

Comme nous le savons, dans le traitement du signal, le signal de sortie d'un système linéaire invariant dans le temps est le résultat de la convolution entre le signal d'entrée et la réponse impulsionnelle du système. La convolution dans le domaine temporel correspond au produit normal dans le domaine fréquentiel et vice versa. Dans la méthode d'analyse cepstrale, le logarithme du spectre est utilisé pour convertir la multiplication dans le domaine fréquentiel en addition (somme), simplifiant ainsi la séparation du spectre du signal d'entrée et la réponse en fréquence du système. Ensuite, en appliquant la transformée en cosinus, nous pouvons séparer le signal d'entrée de la réponse impulsionnelle du système. Cette transformation, qui est uniquement le cepstre réel représentant la reconversion du log-spectre dans le domaine temporel, a un effet de décorrélation entre les caractéristiques du logarithme du spectre énergétique. Par conséquent, dans notre modèle proposé, nous avons appliqué la méthode cepstrale à la banque de filtres de la décomposition dyadique en ondelettes afin d'obtenir les caractéristiques pour classifier l'état émotionnel d'une personne.

Les trois descripteurs proposés sont les suivants :

- Descripteur DWE :

Le premier descripteur consiste à convertir les coefficients d'approximation et de détail, extraits à chaque niveau  $j$  de la décomposition DWT, en coefficients d'énergie. Ici,  $E_\varphi(j_0)$ , qui est l'énergie des coefficients d'approximation à l'échelle  $j_0$ , est donné par :

$$E_\varphi(j_0) = \sum_m \sum_n |W_\varphi(j_0, m, n)|^2 \quad (\text{III.11})$$

De même, l'énergie des coefficients de détail, à l'échelle  $j$  et à l'orientation  $i$ , est donnée par :

$$E_{\psi}^i(j) = \sum_m \sum_n |W_{\psi}^i(j, m, n)|^2 \quad (\text{III.12})$$

- Descripteur LWE :

Le deuxième descripteur consiste à calculer le logarithme des coefficients d'énergie  $E_{\varphi}(j_0)$  et  $E_{\psi}^i(j)$  à chaque niveau de décomposition en ondelettes. Les coefficients résultants sont donnés comme suit :

$$LWE(W_{\varphi}(j_0, m, n)_K) = \log[E_{\varphi}(j_0)] \quad (\text{III.13})$$

$$LWE(W_{\psi}^i(j, m, n)_L) = \log[E_{\psi}^i(j)] \quad (\text{III.14})$$

- Descripteur WCC :

Le dernier descripteur est obtenu en appliquant la Transformée en Cosinus Discrète (DCT) au logarithme des valeurs d'énergie résultantes. Cette transformation décompose une séquence finie de points de données en une somme de fonctions cosinus oscillant à différentes fréquences [251]. Dans notre travail, la DCT est appliquée pour obtenir des coefficients faiblement corrélés. Ainsi, les coefficients d'approximation WCC sont donnés comme suit [252] :

$$WCC_s(W_{\varphi}(j_0, m, n)) = \sqrt{\frac{2}{K}} \sum_{q=1}^K (\log[E_{\varphi}(j_0)])_q \cos\left(\frac{\pi s}{K} (q - 0.5)\right) \quad (\text{III.15})$$

Les coefficients de détail WCC sont donnés comme suit :

$$WCC_s(W_{\psi}^i(j, m, n)) = \sqrt{\frac{2}{K}} \sum_{q=1}^K (\log[E_{\psi}^i(j)])_q \cos\left(\frac{\pi s}{K} (q - 0.5)\right) \quad (\text{III.16})$$

Où :  $K$  est le nombre de canaux de la banque de filtres.

$$s = 1, \dots, L \quad (L = 3j + 1) \quad (\text{III.17})$$

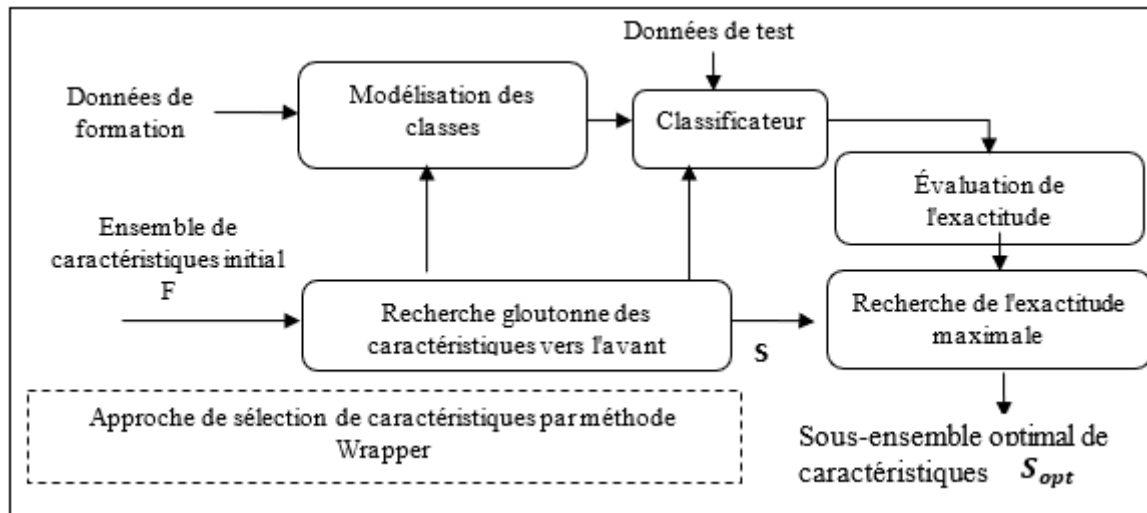
En plus des caractéristiques résultant de l'utilisation de ces trois descripteurs, nous avons également intégré, dans notre étude, une caractéristique d'énergie supplémentaire  $Et$  donnée par l'Équation (III.18). Cette caractéristique est également soumise à l'application des différentes transformations mentionnées précédemment (logarithme et DCT).

$$Et = \sum (a(n), h(n), v(n), d(n), h(n-1), v(n-1), d(n-1), \dots, h(1), v(1), d(1))^2 \quad (\text{III.18})$$

Après avoir extrait les caractéristiques appropriées, nous pouvons effectuer une sélection de caractéristiques pour améliorer la performance du système.

### III.3.1.3. Sélection de caractéristiques

Pour trouver un compromis entre l'amélioration des performances et la réduction de la taille des sous-ensembles de caractéristiques, nous avons adopté un algorithme de sélection de caractéristiques dont le principe est illustré dans la (Figure III.14). Cet algorithme permet d'identifier les caractéristiques les plus importantes dans un ensemble de données et de supprimer les informations non pertinentes.



**Figure III.14: Concept général de l'algorithme de sélection de caractéristiques adopté dans notre système proposé.**

Comme illustré dans cette figure, l'algorithme de sélection de caractéristiques adopté est basé sur l'approche Wrapper. Cette approche commence avec un ensemble vide de caractéristiques, et à chaque itération, la caractéristique dont l'ajout améliore les performances est ajoutée (en utilisant une recherche avancée gloutonne 'greedy forward search'). En d'autres termes, les différents sous-ensembles de caractéristiques sont générés successivement en ajoutant les caractéristiques une par une. Le classificateur, illustré dans la même figure, est ensuite utilisé pour évaluer la pertinence des différents sous-ensembles générés. Le processus s'arrête lorsqu'il n'y a plus de variables supplémentaires à ajouter. Le sous-ensemble optimal est alors celui qui donne la meilleure précision de reconnaissance. Nous expliquerons par la suite le principe de l'application de cet algorithme pour notre système proposé.

Soit  $F$  l'ensemble de départ et  $S$  le sous-ensemble de caractéristiques pertinentes. Initialement, le sous-ensemble de caractéristiques pertinentes  $S$  est vide, tandis que l'ensemble de départ  $F$  contient toutes les caractéristiques. L'algorithme de sélection procède à déterminer le taux de reconnaissance en considérant chacune des caractéristiques  $f_i$  de l'ensemble de départ, et après avoir déterminé les différents taux de reconnaissance correspondant aux différentes caractéristiques, l'algorithme retire la caractéristique  $f_{s1}$  (de l'ensemble de départ  $F$ ), permettant ainsi d'obtenir la reconnaissance maximale de l'ensemble de départ  $F$  et l'ajoute au sous-ensemble de caractéristiques pertinentes  $S$ . L'algorithme répète la même procédure sur les caractéristiques restantes dans l'ensemble de départ  $F$ . En effet, en comparant les taux de reconnaissance obtenus en utilisant les caractéristiques du sous-ensemble  $S$  déjà sélectionné et combiné à chaque fois avec l'une des caractéristiques restantes dans l'ensemble de départ  $F$ , l'algorithme retire (de l'ensemble  $F$ ) à chaque itération la caractéristique qui permet d'obtenir le taux de reconnaissance maximal de l'ensemble  $F$  et l'ajoute à l'ensemble  $S$  (jusqu'à ce qu'il complète le test sur toutes les caractéristiques disponibles). Plusieurs sous-ensembles candidats avec différents nombres de caractéristiques sont donc formés. Enfin, le sous-ensemble optimal  $S_{opt}$  représente le sous-ensemble permettant le meilleur taux de reconnaissance parmi ceux obtenus par les différents sous-ensembles testés (candidats) [222], [220], [243].

L'algorithme proposé peut être résumé comme suit :

Algorithme 1 : Algorithme de sélection de caractéristiques proposé.

*Etape 1 :* Initialiser l'ensemble de départ  $F$  avec toutes les caractéristiques et le sous-ensemble de caractéristiques pertinentes  $S$  sans caractéristiques (sous-ensemble vide).

*Etape 2 :* Calculer le taux de reconnaissance en utilisant chaque caractéristique  $f_i \in F$ .

*Etape 3 :* Trouver la première caractéristique  $f_{s1}$  qui maximise le taux de reconnaissance.

*Etape 4 :* Retirer la caractéristique  $f_{s1}$  de l'ensemble  $F$  et l'ajouter au sous-ensemble de caractéristiques pertinentes  $S$ .

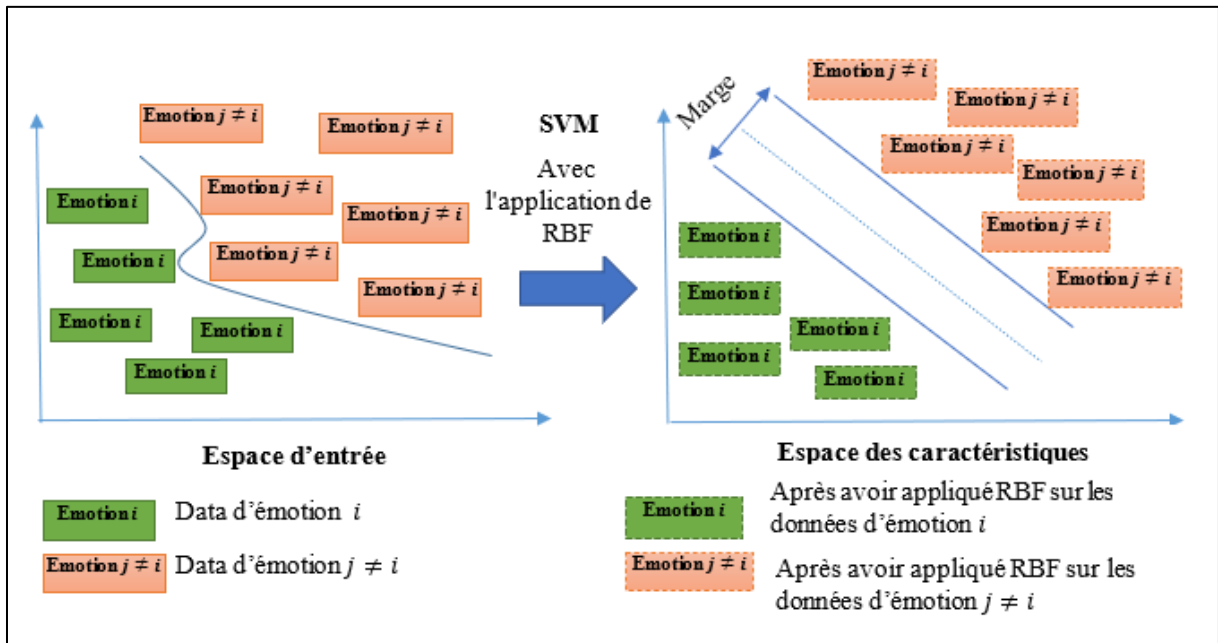
*Etape 5 :* Répéter le processus (de l'étape 2 à l'étape 4) jusqu'à ce que tous les tests sur les caractéristiques soient complétés. Cela se fait comme suit :

- 1)  $\forall f_i \in F$ , déterminer le taux de reconnaissance en utilisant  $\{S, f_i\}$ .
- 2) Choisir la caractéristique  $f_i \in F$  qui maximise le taux de reconnaissance en utilisant  $\{S, f_i\}$  à l'étape  $j$ .
- 3) Affecter  $F \leftarrow F - \{f_{sj}\}$ ;  $S \leftarrow S \cup \{f_{sj}\}$ .

*Etape 6 :* Retirer le sous-ensemble  $S$  des caractéristiques sélectionnées qui maximise le taux de reconnaissance.

### III.3.1.4. Reconnaissance émotionnelle

Rappelons que l'objectif de notre système de reconnaissance des expressions faciales est de reconnaître l'état émotionnel d'une personne en vérifiant si son image faciale correspond à une classe émotionnelle ciblée. La validation de la méthode proposée est réalisée en considérant les différents types d'émotions (colère, joie, mépris, neutre, surprise, tristesse, dégoût et peur) en utilisant un SVM à deux classes. Comme illustré dans (Figure III.15), le SVM cherche à trouver un hyperplan optimal pour séparer linéairement les deux classes.



**Figure III.15: Représentation du concept du noyau RBF dans le SVM.**

Lorsque les données ne sont pas linéairement séparables, différents noyaux, tels que les noyaux polynomial, sigmoïde et RBF, peuvent être utilisés pour construire une fonction de décision linéaire dans l'espace des caractéristiques afin que la base de données devienne séparable avec la marge maximale. Dans nos expériences, nous utilisons le noyau RBF. La fonction de décision du SVM utilisant le noyau RBF est donnée par :

$$f(x) = \text{sgn}\left(\sum_{i=1}^N \alpha_i y_i K(x, x_i) + b\right) \quad (\text{III.19})$$

Où :



$sgn(\cdot)$ : est la fonction de signe ;

$x$ : les données d'entrée ;

$\alpha_i$  : sont les multiplicateurs de Lagrange obtenus lors de la phase d'entraînement ;

$y_i$  : sont les étiquettes de classe des points de données d'entraînement ;

$b$  : est le terme de biais ;

$K(x, x_i)$ : est la fonction noyau RBF, définie comme suit :

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \quad (\text{III.20})$$

Avec :

$\gamma$ : est un hyperparamètre qui contrôle la largeur du noyau ;

$\| \cdot \|$ : est la distance euclidienne.

$f(x)$  prédit l'étiquette de classe du point de données d'entrée  $x$ . Pendant l'entraînement, le SVM apprend les valeurs optimales de  $\alpha_i$  et  $b$  en résolvant le problème d'optimisation pour maximiser la marge entre les vecteurs de support. La fonction de décision utilise ensuite ces paramètres pour classer les nouveaux points de données en fonction de leur proximité avec la frontière de décision définie par l'hyperplan [253].

### III.3.2. Résultats expérimentaux et discussion

Des expériences ont été menées pour tester les performances du système proposé. À cette fin, plusieurs situations ont été envisagées.

#### III.3.2.1. Préparation des données

Pour évaluer notre modèle proposé, deux bases de données ont été utilisées. Elles sont décrites comme suit :

- *Base de données JAFFE [100]* : la base de données JAFFE se compose d'un ensemble d'images faciales, chacune avec une résolution de 256×256 pixels. Les images représentent sept expressions faciales (joie, tristesse, colère, peur, surprise, dégoût, neutre) de 10 sujets féminins japonais différents, simulant 3 à 4 exemples pour chaque émotion.
- *Base de données Extended Cohn–Kanade (CK+) [97]* : cette base de données est l'une des plus largement utilisées pour la détection des émotions. Elle comprend un total de 593 séquences d'images provenant de 123 sujets différents. Tous les participants étaient âgés de 18 à 50 ans ; 69 % étaient des femmes, 81 % étaient des Américains d'origine européenne, 13 % étaient Afro-Américains et 6 % étaient d'autres groupes. En général, le dernier cadre des séquences est utilisé pour la reconnaissance des expressions faciales basée sur les images. Dans notre étude, la version CK+48 a été utilisée. Celle-ci consiste en un ensemble de 981 images faciales, chacune avec une résolution de 48×48 pixels. Les images représentent sept expressions faciales, à savoir : colère, mépris, dégoût, peur, joie, tristesse et surprise.

Le (Tableau III.11) résume la distribution des émotions de chaque base de données.

**Tableau III.11: Nombre d'images (par émotion) utilisées dans nos expériences à partir des bases de données CK+ et JAFFE.**

| Base de données | Colère | Mépris | Peur | Dégout | Triste | Heureux | Surprise | Neutre | $\Sigma$ |
|-----------------|--------|--------|------|--------|--------|---------|----------|--------|----------|
| CK+             | 135    | 54     | 75   | 177    | 84     | 207     | 249      | /      | 981      |
| JAFFE           | 30     | /      | 32   | 29     | 30     | 31      | 31       | 30     | 213      |

Dans nos expériences (pour les deux bases de données), le nombre d'images pour chaque émotion a été divisé en deux sous-ensembles : 50 % pour la phase d'entraînement et 50 % pour les tests. Il convient de noter que le choix de diviser les données de manière égale entre l'entraînement et les tests présente plusieurs avantages et justifications. En effet, cette répartition permet une évaluation équilibrée des performances du modèle. Cela signifie que le modèle est évalué sur une quantité substantielle de données qui n'a pas été utilisée pendant l'entraînement, ce qui permet une estimation plus fiable de sa capacité à généraliser à de nouvelles données. D'autre part, allouer une portion égale de données aux tests aide à éviter le sur-apprentissage, où le modèle apprend trop spécifiquement à partir des données d'entraînement et ne généralise pas bien aux nouvelles données.

### III.3.2.2. Expériences

Comme expliqué précédemment, pour l'extraction des caractéristiques, la DWT a été appliquée à chaque image pour la décomposer en ses coefficients d'approximation et de détail respectifs. En pratique, deux critères sont nécessaires pour assurer une meilleure décomposition. Ceux-ci se concentrent sur la fonction mère d'ondelette optimale et le niveau de décomposition en ondelettes. Comme nous l'avons discuté précédemment, dans nos expériences, trois familles différentes d'ondelettes ont été utilisées pour effectuer l'analyse DWT sur chaque image. Celles-ci sont les suivantes :

- La famille Daubechies avec un ordre variant de 1 à 8 (Db1, Db2, ..., Db8).
- La famille Coiflet avec un ordre variant de 1 à 5 (Coif1, Coif2..., Coif5).
- La famille Symlet avec un ordre variant de 1 à 8 (Sym1, Sym2, ..., Sym8).

La décomposition en ondelettes d'une image au niveau  $n$  renvoie un vecteur  $C$  contenant les coefficients d'approximation et de détail organisés par niveau ( $a(n)$ ,  $h(n)$ ,  $v(n)$ ,  $d(n)$ ,  $h(n-1)$ ,  $v(n-1)$ ,  $d(n-1)$ , ...,  $h(1)$ ,  $v(1)$ ,  $d(1)$ ). Ici,  $a$ ,  $h$ ,  $v$  et  $d$  sont des vecteurs contenant les coefficients d'approximation, les coefficients de détail horizontal, les coefficients de détail vertical et les coefficients de détail diagonal, respectivement.

Après la décomposition en ondelettes 2D, les pourcentages d'énergie, correspondant à l'approximation  $Ea$ , aux détails horizontaux  $Eh$  aux détails verticaux  $Ev$  et aux détails diagonaux  $Ed$ , sont déterminés à chaque niveau de décomposition. L'ensemble de caractéristiques résultant, combiné avec la caractéristique d'énergie  $Et$ , forme le premier descripteur DWE donné comme suit :

$$X1 = [Ea(n), Eh(n), Ev(n), Ed(n), Eh(n-1), Ev(n-1), Ed(n-1), \dots, Eh(1), Ev(1), Ed(1), Et] \quad (\text{III.21})$$

Ensuite, le second descripteur LWE est obtenu en calculant le logarithme des valeurs appartenant au vecteur  $X1$ . Cela est donné par :

$$X2 = \log(X1) \quad (\text{III.22})$$

Enfin, le descripteur WCC est obtenu en appliquant la DCT au vecteur de caractéristiques appartenant au deuxième descripteur :

$$X3 = DCT(X2) \quad (III.23)$$

Après la phase d'extraction des caractéristiques, un algorithme de sélection de caractéristiques est utilisé pour tester les sous-ensembles de caractéristiques résultants et ainsi sélectionner le plus pertinent. Le choix du sous-ensemble optimal est effectué par le classificateur, qui évalue tous les sous-ensembles candidats pour choisir celui garantissant la meilleure reconnaissance pour chaque émotion.

### III.3.2.3. Résultats et discussion

Dans ce qui suit, nous présentons les résultats et les discussions du système REF proposé. Tout d'abord, nous examinerons l'effet de l'ondelette mère et des niveaux de décomposition associés sur les performances.

Selon les différents ordres de l'ondelette mère et le niveau maximal de décomposition associé, le classificateur SVM est entraîné et testé en utilisant les caractéristiques sélectionnées. Les résultats de performance, pour chaque émotion, ont été évalués en calculant le taux de reconnaissance  $R$  donné comme suit [254] :

$$R = \frac{\text{Number of correctly recognised images}}{\text{Number of test images}} \times 100 \quad (III.24)$$

Les meilleurs résultats, en termes de taux de reconnaissance des différentes émotions, correspondant à chacun des trois descripteurs (DWE, LWE et WCC), sont présentés dans le (Tableau III.12). Ces résultats ont été obtenus après avoir testé les différents cas possibles selon les trois types de caractéristiques résultant de l'utilisation des trois descripteurs.

L'analyse des résultats obtenus montre que le système proposé atteint un bon taux de reconnaissance dépassant 81 % pour tous les types d'expressions. En effet, notre système proposé, pour les caractéristiques extraites du descripteur DWE, donne un taux de reconnaissance plus élevé pour les expressions de surprise et de peur. Pour les expressions de colère et neutre, ce taux est plus élevé lorsqu'on utilise les caractéristiques extraites du descripteur LWE. Enfin, pour les émotions de mépris, dégoût, bonheur et tristesse, le même taux est plus élevé lorsqu'on utilise les caractéristiques extraites du descripteur WCC.

**Tableau III.12: Résultats de performance obtenus pour chaque émotion.**

| Émotion  | Meilleur descripteur | Ondelette mère optimale | Niveau de décomposition | Taux de reconnaissance (JAFEE) | Taux de reconnaissance (CK+) |
|----------|----------------------|-------------------------|-------------------------|--------------------------------|------------------------------|
| Colère   | LWE                  | Sym1                    | 7                       | 85.71%                         | 86.34%                       |
| Mépris   | WCC                  | Db8                     | 3                       | /                              | 94.41%                       |
| Peur     | DWE                  | Db2                     | 6                       | 90.48%                         | 90.06%                       |
| Dégoût   | WCC                  | Db1                     | 7                       | 91.43%                         | 81.37%                       |
| Triste   | WCC                  | Sym7                    | 3                       | 87.62%                         | 90.68%                       |
| Heureux  | WCC                  | Sym2                    | 6                       | 90.47%                         | 85.51%                       |
| Surprise | DWE                  | Sym6                    | 4                       | 96.19%                         | 85.92%                       |
| Neutre   | LWE                  | Db6                     | 4                       | 85.71%                         | /                            |

### III.3.2.4. Évaluation des performances

Les résultats présentés dans les Tableaux (III.13) et (III.14) montrent les contributions de l'approche de sélection de caractéristiques proposée pour améliorer la performance globale du système de reconnaissance. En effet, comme le montrent ces deux tableaux, l'approche de sélection proposée augmente de manière significative le taux de reconnaissance en utilisant uniquement un nombre réduit de caractéristiques pertinentes. Cela est dû au fait que la méthode de sélection résout le problème de la malédiction de la dimensionnalité. Par exemple, pour le cas de l'émotion de la colère, l'utilisation de l'approche de sélection proposée a permis d'exploiter uniquement 4,35 % (une réduction de 95,65 %) du nombre total initial de caractéristiques (utilisées avant l'opération de sélection) pour garantir la tâche de reconnaissance avec une augmentation de 17,14 % du taux de reconnaissance.

**Tableau III.13: Taux de reconnaissance et nombre de caractéristiques obtenues pour chaque émotion de la base de données JAFFE avant et après l'adoption de l'approche de sélection.**

| Émotion  | Taux de reconnaissance       |                              |                                     | Nombre de caractéristiques      |                                 | Taux de réduction du nombre de caractéristiques |
|----------|------------------------------|------------------------------|-------------------------------------|---------------------------------|---------------------------------|-------------------------------------------------|
|          | Avant la sélection ( $R_b$ ) | Après la sélection ( $R_a$ ) | Taux d'amélioration ( $R_a - R_b$ ) | Avant la sélection ( $N_{fb}$ ) | Après la sélection ( $N_{fa}$ ) | $(1 - (N_{fa}/N_{fb})) \times 100$              |
| Colère   | 68.57%                       | <b>85.71%</b>                | <b>17.14%</b>                       | 23                              | <b>1</b>                        | <b>95.65%</b>                                   |
| Mépris   | /                            | /                            | /                                   | /                               | /                               | /                                               |
| Peur     | 74.29%                       | <b>90.48%</b>                | <b>16.19%</b>                       | 20                              | <b>9</b>                        | <b>55%</b>                                      |
| Dégoût   | 81.90%                       | <b>91.43%</b>                | <b>9.53%</b>                        | 23                              | <b>8</b>                        | <b>65.22%</b>                                   |
| Triste   | 72.38%                       | <b>87.62%</b>                | <b>15.24%</b>                       | 11                              | <b>1</b>                        | <b>90.91%</b>                                   |
| Heureux  | 84.76%                       | <b>90.47%</b>                | <b>5.71%</b>                        | 20                              | <b>7</b>                        | <b>65%</b>                                      |
| Surprise | 89.52%                       | <b>96.19%</b>                | <b>6.67%</b>                        | 14                              | <b>9</b>                        | <b>35.71%</b>                                   |
| Neutre   | 45.71%                       | <b>85.71%</b>                | <b>40%</b>                          | 14                              | <b>1</b>                        | <b>92.86%</b>                                   |

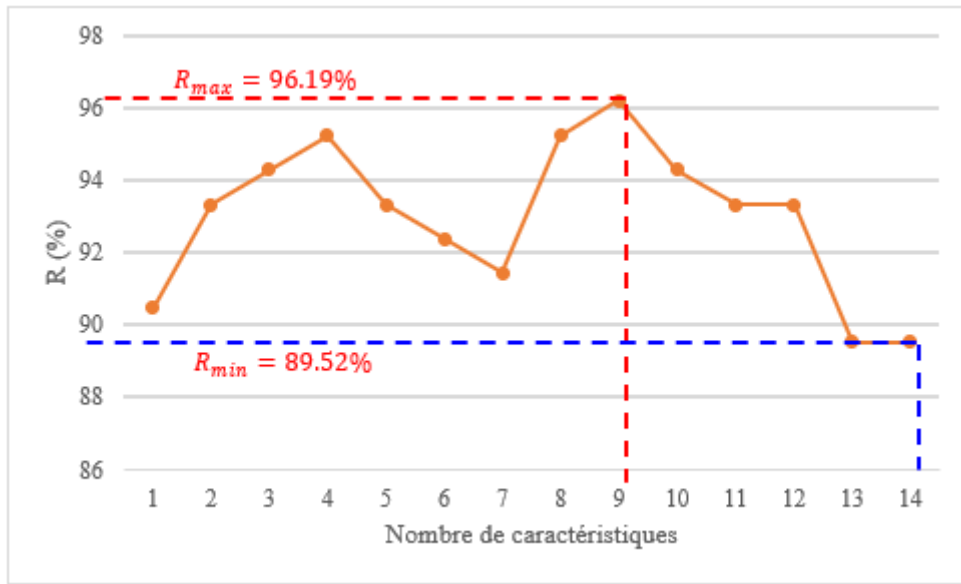
**Tableau III.14: Taux de reconnaissance et nombre de caractéristiques obtenues pour chaque émotion dans la base de données CK+ avant et après l'adoption de l'approche de sélection.**

| Émotion         | Taux de reconnaissance       |                              |                                     | Nombre de caractéristiques      |                                 | Taux de réduction du nombre de caractéristiques |
|-----------------|------------------------------|------------------------------|-------------------------------------|---------------------------------|---------------------------------|-------------------------------------------------|
|                 | Avant la sélection ( $R_b$ ) | Après la sélection ( $R_a$ ) | Taux d'amélioration ( $R_a - R_b$ ) | Avant la sélection ( $N_{fb}$ ) | Après la sélection ( $N_{fa}$ ) | $(1 - (N_{fa}/N_{fb})) \times 100$              |
| <b>Colère</b>   | 75.57%                       | <b>86.34%</b>                | <b>10.77%</b>                       | 23                              | <b>1</b>                        | <b>95.65%</b>                                   |
| <b>Mépris</b>   | 82.61%                       | <b>94.41%</b>                | <b>11.8%</b>                        | 11                              | <b>1</b>                        | <b>90.91%</b>                                   |
| <b>Peur</b>     | 87.99%                       | <b>90.06%</b>                | <b>2.07%</b>                        | 20                              | <b>17</b>                       | <b>15%</b>                                      |
| <b>Dégoût</b>   | 73.08%                       | <b>81.37%</b>                | <b>8.29%</b>                        | 23                              | <b>1</b>                        | <b>95.65%</b>                                   |
| <b>Triste</b>   | 70.39%                       | <b>90.68%</b>                | <b>20.29%</b>                       | 11                              | <b>1</b>                        | <b>90.91%</b>                                   |
| <b>Heureux</b>  | 82.19%                       | <b>85.51%</b>                | <b>3.32%</b>                        | 20                              | <b>12</b>                       | <b>40%</b>                                      |
| <b>Surprise</b> | 75.98%                       | <b>85.92%</b>                | <b>9.94%</b>                        | 14                              | <b>5</b>                        | <b>64.29%</b>                                   |
| <b>Neutre</b>   | /                            | /                            | /                                   | /                               | /                               | /                                               |

En général, pour toutes les autres émotions, l'approche de sélection de caractéristiques proposée réduit le nombre de caractéristiques et augmente le taux de reconnaissance, ce qui améliore les performances en termes de temps de réponse. En effet, l'augmentation du taux de reconnaissance varie entre 5,71 %

et 40 % pour la base de données JAFFE et entre 2,07 % et 20,29 % pour la base de données CK+. De même, le taux de réduction du nombre de caractéristiques était compris entre 35,71 % et 95,65 % pour la base de données JAFFE et entre 15 % et 95,65 % pour la base de données CK+.

Pour mieux expliquer ce phénomène, prenons l'exemple de la reconnaissance de l'émotion de surprise en utilisant la base de données JAFFE. La variation résultante du taux de reconnaissance en fonction du nombre de caractéristiques est illustrée dans la (Figure III.16). Comme le montre cette figure, si le classificateur utilise l'ensemble complet des caractéristiques extraites, alors le niveau maximal de décomposition, correspondant à l'ondelette Sym6, est égal à 4. Ce processus a donné un total de 14 caractéristiques (1 coefficient d'approximation, 4 coefficients de détail horizontal, 4 coefficients de détail vertical, 4 coefficients de détail diagonal, plus 1 coefficient d'énergie supplémentaire). Dans ce cas, le taux de reconnaissance atteint une valeur de 89,52 %.



**Figure III.16: Variation du taux de reconnaissance (émotion de surprise) en fonction du nombre de caractéristiques.**

Si le classificateur évalue maintenant la pertinence de chacune des caractéristiques et sélectionne uniquement celles qui améliorent la performance, le taux de reconnaissance atteindra 96,19 % (une amélioration de 6,67 %) en utilisant seulement 9 caractéristiques (soit une réduction du nombre de caractéristiques de 35,71 %).

### III.3.2.5. Comparaison des performances

Pour démontrer l'efficacité de notre modèle REF proposé, nous le comparons à la méthode LBP. Cette dernière s'est révélée être une méthode simple et efficace pour la représentation des expressions faciales [255]. Elle consiste en un descripteur pouvant être utilisé en traitement d'image et en vision par ordinateur pour représenter la texture d'une image [57]. L'opérateur LBP prend un voisinage local autour de chaque pixel, limite les pixels de ce voisinage à la valeur du pixel central, et utilise le patch d'image binaire résultant comme un descripteur local d'image [59].

Le descripteur LBP pour chaque pixel d'une image est donné comme suit [57] :

$$LBP(P, R) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (III.25)$$

Où :  $s(x)$  représente la fonction de signe correspondant au LBP.  $g_p$  et  $g_c$  sont respectivement les niveaux de gris d'un pixel voisin et du pixel central.

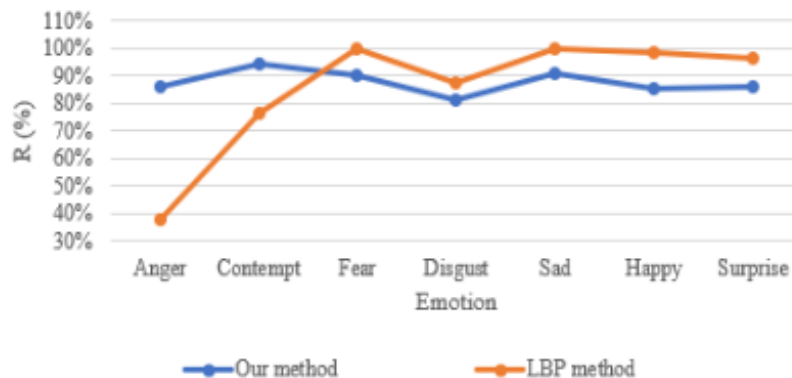
Le code binaire résultant est ensuite converti en un nombre décimal, qui représente le motif de texture de ce pixel. En appliquant le LBP à chaque pixel d'une image, un histogramme de ces motifs peut être créé, fournissant une représentation compacte de la texture de l'image.

Dans ce travail, la comparaison des performances de notre système proposé avec celui basé sur le LBP est effectuée dans les mêmes conditions en utilisant les mêmes données. Les résultats, en termes de taux de reconnaissance des expressions faciales, sont présentés dans les (Figure III.17) et (Figure III.18) pour les bases de données JAFFE et CK+, respectivement. Comme le montre la (Figure III.17), pour la base de données JAFFE, notre cadre surpasse le cadre basé sur le LBP pour la majorité des émotions (peur, dégoût, joie, surprise, tristesse et neutre).

Cependant, le taux de reconnaissance est similaire à celui obtenu par le système basé sur le LBP pour l'émotion de colère. Pour la base de données CK+, comme le montre la (Figure III.18), le système proposé, comparé au système basé sur le LBP, obtient les meilleures performances pour les émotions de colère et de mépris. Pour les autres émotions, le système basé sur le LBP surpasse légèrement notre système. Malgré ces derniers résultats concernant le taux de reconnaissance de notre cadre, il reste plus efficace en termes de temps de réponse que le cadre basé sur le LBP, comme nous le verrons dans la section suivante.



**Figure III.17: Comparaison entre les taux de reconnaissance des expressions faciales (en %) obtenus par notre modèle proposé et ceux obtenus par le modèle basé sur le LBP dans la base de données JAFFE.**



**Figure III.18: Comparaison entre les taux de reconnaissance des expressions faciales (en %) obtenus par notre modèle proposé et ceux obtenus par le modèle basé sur le LBP dans la base de données CK+.**

### Comparaison des Temps de Réponse

Pour comparer le temps de réponse de notre système proposé avec celui du système basé sur le LBP, une étude a été réalisée pour déterminer le nombre de caractéristiques nécessaires pour entraîner chaque modèle. Les résultats sont les suivants :

- Pour le système REF basé sur le LBP : comme indiqué dans le (Tableau III.15), chaque image originale est transformée en une image LBP en attribuant une valeur parmi  $2^8 = 256$  valeurs possibles à chacun de ses pixels. Ensuite, l'image résultante est convertie en un vecteur de caractéristiques en calculant un histogramme montrant la fréquence d'apparition correspondant à chaque valeur LBP. Le vecteur final a donc une longueur de 256.
- Pour notre système REF proposé : comme le montre le même tableau, le nombre de caractéristiques ne dépasse pas 9 dans la base de données JAFFE (pour la peur et la surprise) ou 17 caractéristiques dans la base de données CK+ (pour la peur).

Par conséquent, comparé à celui de notre vecteur de caractéristiques, la taille du vecteur de caractéristiques résultant du système basé sur le LBP est grande, ce qui prouve que notre système proposé surpasse celui basé sur le LBP en termes de complexité computationnelle. Pour confirmer ce résultat, des simulations utilisant l'outil MATLAB ont été réalisées pour tester le temps de calcul consommé par les deux systèmes mentionnés ci-dessus. Les résultats, fournis et illustrés dans le (Tableau III.15), démontrent l'efficacité en temps de réponse de notre méthode proposée par rapport à celle de la méthode LBP. En effet, selon ce tableau, notre système proposé a un temps de calcul très faible pour les deux bases de données (25 % pour JAFFE et 10,63 % pour CK+).

**Tableau III.15: Résultats de la simulation du critère de temps de calcul pour les bases de données JAFFE et CK+.**

| Base de données       | Méthode              | Nombre de caractéristiques requises | Temps consommé en % |
|-----------------------|----------------------|-------------------------------------|---------------------|
| Base de données JAFFE | <b>Notre méthode</b> | <b>9</b>                            | <b>25%</b>          |
|                       | Méthode LBP          | 256                                 | 100%                |
| Base de données CK+   | <b>Notre méthode</b> | <b>17</b>                           | <b>10.63%</b>       |
|                       | Méthode LBP          | 256                                 | 100%                |

## III.4. Conclusion

Ce chapitre a présenté deux approches pour améliorer les systèmes de reconnaissance des expressions faciales. Le premier travail a introduit un système REF amélioré en intégrant une approche de sélection "wrapper" et en exploitant les caractéristiques locales du visage. Grâce à l'utilisation de l'ASM pour l'extraction des caractéristiques et du GPA pour leur normalisation, le système proposé a montré des performances améliorées par rapport aux systèmes traditionnels, en résolvant efficacement le problème de la malédiction de la dimensionnalité. L'analyse a révélé que les caractéristiques des sourcils jouent un rôle crucial dans la reconnaissance des émotions, ce qui pourrait influencer les futurs algorithmes de reconnaissance et leurs applications dans divers domaines.

Le deuxième travail a développé un système REF basé sur une nouvelle méthode d'extraction de caractéristiques combinée à une sélection par "Wrapper". Utilisant trois descripteurs basés sur la DWT, cette approche a été évaluée à l'aide des bases de données JAFFE et CK+. Les résultats ont montré que le système proposé surpassait le système basé sur le LBP en termes de taux de reconnaissance global et de temps de réponse, offrant une meilleure performance pour certaines émotions et une réponse plus rapide, grâce à une complexité réduite. Ce système démontre un compromis efficace entre précision et complexité

## CONCLUSION GENERALE

---

Dans ce travail, nous avons exploré la reconnaissance automatique des expressions faciales, un domaine clé à l'intersection de l'intelligence artificielle, de la vision par ordinateur et de l'interaction homme-machine. Nous avons étudié les fondements théoriques des émotions, leur expression faciale ainsi que les différentes approches utilisées pour leur reconnaissance automatique, en mettant en évidence les défis et les solutions existantes.

L'un des défis majeurs dans ce domaine réside dans la généralisation du modèle en mode indépendant de la personne. Cette capacité permet à un système de reconnaissance des expressions faciales d'identifier correctement les émotions d'individus qu'il n'a jamais vus lors de l'apprentissage, malgré les variabilités interindividuelles liées aux différences morphologiques, culturelles et expressives. Pour relever ce défi, nous avons mis en évidence l'importance cruciale de la sélection de caractéristiques. Une sélection efficace permet non seulement de réduire la dimensionnalité des données, mais aussi d'améliorer la robustesse et la précision des modèles de classification, tout en optimisant le temps de calcul et les ressources.

Dans un premier volet, nous avons exploré l'impact de l'ajout d'une approche wrapper pour la sélection des caractéristiques, en utilisant des informations extraites de régions locales spécifiques du visage, telles que la bouche, les yeux et les sourcils. Les résultats, obtenus à partir de la base de données MUG, ont révélé que le système REF, associé à cette stratégie de sélection, offrait des performances supérieures à celles du système initial, pour plusieurs modèles de classification comme KNN, arbre de décision, Naive Bayes et LDA. Une réduction de 50 % du nombre de caractéristiques utilisées, ainsi qu'une précision de 100 % avec LDA, ont été obtenues, démontrant une amélioration considérable en termes de temps de calcul, d'efficacité et d'optimisation de la mémoire. De plus, les sourcils se sont avérés être une région clé pour la reconnaissance des émotions.

Dans un second volet, nous avons proposé un modèle efficace pour améliorer la précision et réduire le coût de calcul d'un système de reconnaissance des expressions faciales. Ce modèle se compose de trois étapes principales : l'extraction des caractéristiques à l'aide de descripteurs basés sur la Transformée en Ondelette Discrète (DWT), la sélection des caractéristiques par une approche wrapper, et la classification à l'aide d'une machine à vecteurs de support (SVM). Les expérimentations, réalisées sur les bases de données JAFFE et CK+, ont montré des gains significatifs, surpassant les performances de la méthode basée sur les motifs binaires locaux (LBP).

Les résultats obtenus tout au long de cette thèse confirment l'importance cruciale de la sélection de caractéristiques pour relever les défis posés par la reconnaissance des expressions faciales en mode indépendant des personnes. En plus d'améliorer la précision des classifieurs, ces travaux offrent des gains notables en termes d'efficacité computationnelle. Cette thèse ouvre ainsi des perspectives intéressantes pour l'avenir, notamment en explorant l'application de nouvelles méthodes de sélection et d'optimisation dans le cadre de la reconnaissance d'expressions faciales.



## **PERSPECTIVES FUTURES**

Pour renforcer encore les systèmes de reconnaissance des expressions faciales, les travaux futurs pourraient se concentrer sur les axes suivants :

- 1- Intégration des techniques d'apprentissage profond : incorporer des réseaux de neurones convolutionnels (CNN) et d'autres architectures d'apprentissage profond pour une extraction plus fine et complète des caractéristiques faciales.
- 2- Enrichissement et diversification des bases de données : élaborer des bases de données plus variées en termes d'ethnicité, d'âge, et de conditions environnementales afin d'améliorer la généralisation des modèles de REF dans des contextes réels et diversifiés.
- 3- Fusion Multimodale : Explorer la fusion de données multimodales, telles que les expressions faciales combinées avec les données vocales ou gestuelles, pour une reconnaissance plus robuste et précise des émotions.

Ces orientations futures permettront non seulement d'améliorer la performance des systèmes REF existants mais aussi de les adapter à des applications plus larges et diversifiées, ouvrant ainsi de nouvelles avenues pour la recherche et le développement dans le domaine de la reconnaissance des émotions.

## REFERENCES BIBLIOGRAPHIQUES

---

- [1] Merinfeld, E. (2002). Introduction: La question des émotions. *Cahiers critiques de thérapie familiale et de pratiques de réseaux*, 29(2), 7-12.
- [2] Lange, C. G. (1885). The mechanism of the emotions. *The classical psychologists*, 672-684.
- [3] Rimé, B. (2009). Chapitre 2. Qu'est-ce qu'une émotion?. *Quadrige*, 43-66.
- [4] Kleinginna Jr, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and emotion*, 5(4), 345-379.
- [5] Courgeon, M. (2011). *Marc: modèles informatiques des émotions et de leurs expressions faciales pour l'interaction Homme-machine affective temps réel* (Doctoral dissertation, Université Paris Sud-Paris XI).
- [6] Christophe, V. (1998). *Les émotions: tour d'horizon des principales théories* (Vol. 3). Presses Univ. Septentrion.
- [7] Planalp, S. (1999). *Communicating emotion: Social, moral, and cultural processes*. Cambridge University Press.
- [8] Belin, L. (2016). *Impact des stimuli visuels sur la réactivité émotionnelle* (Doctoral dissertation, Rennes 1).
- [9] Servant, D., Lebeau, J. C., Mouster, Y., Hautekeete, M., Logier, R., & Goudemand, M. (2008). La variabilité cardiaque: un bon indicateur de la régulation des émotions. *Journal de thérapie comportementale et cognitive*, 18(2), 45-48.
- [10] Faltacas, A. M. (2018). *Tendances à l'action communiquées par les expressions faciales des émotions* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- [11] Calkins, S. D., & Hill, A. (2007). Caregiver influences on emerging emotion regulation: Biological and environmental transactions in early development.
- [12] Roth, G., Vansteenkiste, M., & Ryan, R. M. (2019). Integrative emotion regulation: Process and development from a self-determination theory perspective. *Development and psychopathology*, 31(3), 945-956.
- [13] Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications.
- [14] Cuisinier, F., Tornare, E., & Pons, F. (2015). Les émotions dans les apprentissages scolaires: un domaine de recherche en émergence. *Revue de l'ANAE*, 139, 527-536.
- [15] Noyer, S. (2018). *L'observation des émotions au travers des communications non-verbales*, Université de Toulouse.
- [16] Benleghib, N. (2023). *LES MUSCLES PEAUCIERS*, Faculté de médecine de Constantine.
- [17] Lakouichmi, P. & Marghadi, A. (2023). *SMAS et MUSCLES PEAUCIERS DE LA FACE*, Laboratoire d'anatomie - Faculté de Médecine de Marrakech.
- [18] Elmoataz, A. (2007). *Analyse de visage et catégorisation des expressions faciales*.

- [19] Nugier, A. (2009). Histoire et grands courants de recherche sur les émotions. *Revue électronique de psychologie sociale*, 4(4), 8-14.
- [20] Izard, C. E. (2013). *Human emotions*. Springer Science & Business Media.
- [21] Tomkins, S. S. (1980). Affect as amplification: Some modifications in theory. In *Theories of emotion* (pp. 141-164). Academic Press.
- [22] Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion* (pp. 3-33). Academic press.
- [23] Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, 6(3-4), 169-200.
- [24] Ouakkaf, H., & Berkane, M. (2017). Reconnaissance automatique des expressions faciales par support vector machine.
- [25] Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE transactions on affective computing*, 13(3), 1195-1215.
- [26] Ryan, A., Cohn, J. F., Lucey, S., Saragih, J., Lucey, P., De la Torre, F., & Rossi, A. (2009, October). Automated facial expression recognition system. In *43rd annual 2009 international Carnahan conference on security technology* (pp. 172-177). IEEE.
- [27] Canedo, D., & Neves, A. J. (2019). Facial expression recognition using computer vision: A systematic review. *Applied Sciences*, 9(21), 4678.
- [28] Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. *CVonline: On-Line Compendium of Computer Vision*, 9, 11.
- [29] Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57, 137-154.
- [30] Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1867-1874).
- [31] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters*, 23(10), 1499-1503.
- [32] Zhao, X., & Zhang, S. (2016). A review on facial expression recognition: feature extraction and classification. *IETE Technical Review*, 33(5), 505-517.
- [33] Ghimire, D., Lee, J., Li, Z. N., & Jeong, S. (2017). Recognition of facial expressions based on salient geometric features and support vector machines. *Multimedia Tools and Applications*, 76, 7921-7946.
- [34] Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, 61(1), 38-59.
- [35] Cament, L. A., Galdames, F. J., Bowyer, K. W., & Perez, C. A. (2015). Face recognition under pose variation with local Gabor features enhanced by active shape and statistical models. *Pattern Recognition*, 48(11), 3371-3384.
- [36] Lee, Y. H., Kim, C. G., Kim, Y., & Whangbo, T. K. (2015). Facial landmarks detection using improved active shape model on android platform. *Multimedia Tools and Applications*, 74, 8821-8830.

- [37] Shbib, R., & Zhou, S. (2015). Facial expression analysis using active shape model. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 8(1), 9-22.
- [38] Kim, D. J. (2016). Facial expression recognition using ASM-based post-processing technique. *Pattern Recognition and Image Analysis*, 26, 576-581.
- [39] Si, J., Wan, Y., & Zhang, T. (2016, May). Facial expression recognition based on ASM and Finite Automata Machine. In *2016 2nd Workshop on Advanced Research and Technology in Industry Applications (WARTIA-16)* (pp. 1306-1310). Atlantis Press.
- [40] Suja, P., & Tripathi, S. (2016, February). Real-time emotion recognition from facial images using Raspberry Pi II. In *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 666-670). IEEE.
- [41] Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, 23(6), 681-685.
- [42] Antonakos, E., Alabort-i-Medina, J., Tzimiropoulos, G., & Zafeiriou, S. (2014, October). Hog active appearance models. In *2014 IEEE International Conference on Image Processing (ICIP)* (pp. 224-228). IEEE.
- [43] Anderson, R., Stenger, B., & Cipolla, R. (2014). Using bounded diameter minimum spanning trees to build dense active appearance models. *International journal of computer vision*, 110, 48-57.
- [44] Chen, Y., Hua, C., & Bai, R. (2014). Regression-based active appearance model initialization for facial feature tracking with missing frames. *Pattern Recognition Letters*, 38, 113-119.
- [45] Zheng, W., & Liu, C. (2016, May). Facial expression recognition based on texture and shape. In *2016 25th Wireless and Optical Communication Conference (WOCC)* (pp. 1-5). IEEE.
- [46] Wang, L., Li, R., & Wang, K. (2014). A Novel Automatic Facial Expression Recognition Method Based on AAM. *J. Comput.*, 9(3), 608-617.
- [47] Ko, K. E., & Sim, K. B. (2010, October). Development of a Facial Emotion Recognition Method based on combining AAM with DBN. In *2010 International Conference on Cyberworlds* (pp. 87-91). IEEE.
- [48] Zhang, L., Chen, S., Wang, T., & Liu, Z. (2012). Automatic facial expression recognition based on hybrid features. *Energy Procedia*, 17, 1817-1823.
- [49] Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150-1157). Ieee.
- [50] Soyel, H., & Demirel, H. (2010). Facial expression recognition based on discriminative scale invariant feature transform. *Electronics letters*, 46(5), 343-345.
- [51] Li, Y., Liu, W., Li, X., Huang, Q., & Li, X. (2014). GA-SIFT: A new scale invariant feature transform for multispectral image using geometric algebra. *Information Sciences*, 281, 559-572.
- [52] Feng, Y., An, X., & Liu, X. (2016, October). The application of scale invariant feature transform fused with shape model in the human face recognition. In *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (pp. 1716-1720). IEEE.

- [53] Berretti, S., Ben Amor, B., Daoudi, M., & Del Bimbo, A. (2011). 3D facial expression recognition using SIFT descriptors of automatically detected keypoints. *The Visual Computer*, 27, 1021-1036.
- [54] Dahmane, M., & Meunier, J. (2014). Prototype-based modeling for facial expression analysis. *IEEE Transactions on Multimedia*, 16(6), 1574-1584.
- [55] Zhang, T., Zheng, W., Cui, Z., Zong, Y., Yan, J., & Yan, K. (2016). A deep neural network-driven feature learning method for multi-view facial expression recognition. *IEEE Transactions on Multimedia*, 18(12), 2528-2536.
- [56] Kalsum, T., Anwar, S. M., Majid, M., Khan, B., & Ali, S. M. (2018). Emotion recognition from facial expressions using hybrid feature descriptors. *IET Image Processing*, 12(6), 1004-1012.
- [57] Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7), 971-987.
- [58] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face recognition with local binary patterns. In *Computer Vision-ECCV 2004: 8th European Conference on Computer Vision*, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I 8 (pp. 469-481). Springer Berlin Heidelberg.
- [59] Abdulrahman, M., Gwadabe, T. R., Abdu, F. J., & Eleyan, A. (2014, April). Gabor wavelet transform based facial expression recognition using PCA and LBP. In *2014 22nd signal processing and communications applications conference (SIU)* (pp. 2265-2268). IEEE.
- [60] Dagher, I., & Azar, F. (2019). Improving the SVM gender classification accuracy using clustering and incremental learning. *Expert Systems*, 36(3), e12372.
- [61] He, L., Zou, C., Zhao, L., & Hu, D. (2006, January). An enhanced LBP feature based on facial expression recognition. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 3300-3303). IEEE.
- [62] Shan, C., & Gritti, T. (2008, September). Learning Discriminative LBP-Histogram Bins for Facial Expression Recognition. In *BMVC* (pp. 1-10).
- [63] Yuan, L. U. O., & ZHANG, Y. (2013). Facial expression feature extraction using hybrid PCA and LBP. *The Journal of China Universities of Posts and Telecommunications*, 20(2), 120-124.
- [64] Xu, Q., & Zhao, N. (2020, June). A facial expression recognition algorithm based on CNN and LBP feature. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (Vol. 1, pp. 2304-2308). IEEE.
- [65] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). Ieee.
- [66] Greche, L., & Es-Sbai, N. (2016, March). Automatic system for facial expression recognition based histogram of oriented gradient and normalized cross correlation. In *2016 International Conference on Information Technology for Organizations Development (IT4OD)* (pp. 1-5). IEEE.
- [67] Chen, J., Chen, Z., Chi, Z., & Fu, H. (2014, August). Facial expression recognition based on facial components detection and hog features. In *International workshops on electrical and computer engineering subfields* (pp. 884-888).

- [68] Orrite, C., Ganán, A., & Rogez, G. (2009). Hog-based decision tree for facial expression classification. In *Pattern Recognition and Image Analysis: 4th Iberian Conference, IbPRIA 2009 Póvoa de Varzim, Portugal, June 10-12, 2009 Proceedings 4* (pp. 176-183). Springer Berlin Heidelberg.
- [69] Wang, X., Jin, C., Liu, W., Hu, M., Xu, L., & Ren, F. (2013, December). Feature fusion of HOG and WLD for facial expression recognition. In *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration* (pp. 227-232). IEEE.
- [70] Yacoob, Y., & Davis, L. S. (1996). Recognizing human facial expressions from long image sequences using optical flow. *IEEE Transactions on pattern analysis and machine intelligence*, 18(6), 636-642.
- [71] Negahdaripour, S. (1998). Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9), 961-979.
- [72] Cohn, J. F., Zlochower, A. J., Lien, J. J., & Kanade, T. (1998, April). Feature-point tracking by optical flow discriminates subtle differences in facial expression. In *Proceedings third IEEE international conference on automatic face and gesture recognition* (pp. 396-401). IEEE.
- [73] Sánchez, A., Ruiz, J. V., Moreno, A. B., Montemayor, A. S., Hernández, J., & Pantrigo, J. J. (2011). Differential optical flow applied to automatic facial expression recognition. *Neurocomputing*, 74(8), 1272-1282.
- [74] Allaert, B., Ward, I. R., Bilasco, I. M., Djeraba, C., & Bennamoun, M. (2022). A comparative study on optical flow for facial expression analysis. *Neurocomputing*, 500, 434-448.
- [75] Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Image Analysis: 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29–July 2, 2003 Proceedings 13* (pp. 363-370). Springer Berlin Heidelberg.
- [76] Bailer, C., Taetz, B., & Stricker, D. (2015). Flow fields: Dense correspondence fields for highly accurate large displacement optical flow estimation. In *Proceedings of the IEEE international conference on computer vision* (pp. 4015-4023).
- [77] Tong, Y., Wang, Y., Zhu, Z., & Ji, Q. (2007). Robust facial feature tracking under varying face pose and facial expression. *Pattern Recognition*, 40(11), 3195-3208.
- [78] Wang, Y. (2003, October). Joint region tracking with switching hypothesized measurements. In *Proceedings Ninth IEEE International Conference on Computer Vision* (pp. 75-82). IEEE.
- [79] Li, Y., Wang, S., Zhao, Y., & Ji, Q. (2013). Simultaneous facial feature tracking and facial expression recognition. *IEEE Transactions on image processing*, 22(7), 2559-2573.
- [80] He, S., Wang, S., & Lv, Y. (2011, August). Spontaneous facial expression recognition based on feature point tracking. In *2011 Sixth International Conference on Image and Graphics* (pp. 760-765). IEEE.
- [81] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.
- [82] Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *ieee assp magazine*, 3(1), 4-16.
- [83] Yao, S., He, N., Zhang, H., & Yoshie, O. (2014, May). Micro-expression recognition by feature points tracking. In *2014 10th International Conference on Communications (COMM)* (pp. 1-4). IEEE.

- [84] Kalal, Z., Mikolajczyk, K., & Matas, J. (2011). Tracking-learning-detection. *IEEE transactions on pattern analysis and machine intelligence*, 34(7), 1409-1422.
- [85] Gall, J., Yao, A., Razavi, N., Van Gool, L., & Lempitsky, V. (2011). Hough forests for object detection, tracking, and action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 33(11), 2188-2202.
- [86] Intelligence, C. (2023). Retracted: Deep Learning and Improved HMM Training Algorithm and Its Analysis in Facial Expression Recognition of Sports Athletes. *Computational Intelligence and Neuroscience*, 2023, 9817376.
- [87] Li, C., Pourtaherian, A., van Onzenoort, L., a Ten, W. T., & De With, P. H. N. (2020). Infant facial expression analysis: towards a real-time video monitoring system using r-cnn and hmm. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1429-1440.
- [88] Karimi, G., & Heidarian, M. (2021). Facial expression recognition with polynomial Legendre and partial connection MLP. *Neurocomputing*, 434, 33-44.
- [89] Revina, I. M., & Emmanuel, W. S. (2018, June). Face expression recognition using weber local descriptor and F-RBFNN. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 196-199). IEEE.
- [90] Flores-Juarez, U. A., Álvarez-Cedillo, J. A., & Álvarez-Sánchez, T. E. O. D. O. R. O. (2021). Two facial emotion detection based on naive Bayesian classifier. *Journal of Theoretical and Applied Information Technology*, 99(23).
- [91] Nomiya, H., Sakaue, S., & Hochin, T. (2016). Recognition and intensity estimation of facial expression using ensemble classifiers. *International Journal of Networked and Distributed Computing*, 4(4), 203-211.
- [92] Dixit, B. A., & Gaikwad, A. N. (2015, June). Statistical moments based facial expression analysis. In *2015 IEEE International Advance Computing Conference (IACC)* (pp. 552-557). IEEE.
- [93] Yadav, K. S., & Singha, J. (2020). Facial expression recognition using modified Viola-John's algorithm and KNN classifier. *Multimedia Tools and Applications*, 79(19), 13089-13107.
- [94] Ab Wahab, M. N., Nazir, A., Ren, A. T. Z., Noor, M. H. M., Akbar, M. F., & Mohamed, A. S. A. (2021). Efficientnet-lite and hybrid CNN-KNN implementation for facial expression recognition on raspberry pi. *IEEE Access*, 9, 134065-134080.
- [95] Yang, L., Zhang, H., Li, D., Xiao, F., & Yang, S. (2021, September). Facial expression recognition based on transfer learning and SVM. In *Journal of Physics: Conference Series* (Vol. 2025, No. 1, p. 012015). IOP Publishing.
- [96] Yao, L., Wan, Y., Ni, H., & Xu, B. (2021). Action unit classification for facial expression recognition using active learning and SVM. *Multimedia Tools and Applications*, 80, 24287-24301.
- [97] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 IEEE computer society conference on computer vision and pattern recognition-workshops* (pp. 94-101). IEEE.
- [98] Valstar, M., & Pantic, M. (2010, May). Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In *Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect* (Vol. 10, pp. 65-70).

- [99] Zhao, G., Huang, X., Taini, M., Li, S. Z., & Pietikäinen, M. (2011). Facial expression recognition from near-infrared videos. *Image and vision computing*, 29(9), 607-619.
- [100] Lyons, M., Akamatsu, S., Kamachi, M., & Gyoba, J. (1998, April). Coding facial expressions with gabor wavelets. In *Proceedings Third IEEE international conference on automatic face and gesture recognition* (pp. 200-205). IEEE.
- [101] Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior research methods*, 42, 351-362.
- [102] Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., ... & Bengio, Y. (2013). Challenges in representation learning: A report on three machine learning contests. In *Neural information processing: 20th international conference, ICONIP 2013, daegu, korea, november 3-7, 2013. Proceedings, Part III* 20 (pp. 117-124). Springer berlin heidelberg.
- [103] Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. (2012). Collecting large, richly annotated facial-expression databases from movies. *IEEE multimedia*, 19(3), 34-41.
- [104] Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. (2011, November). Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. In *2011 IEEE international conference on computer vision workshops (ICCV workshops)* (pp. 2106-2112). IEEE.
- [105] Ralph, G., Iain, M., Jeffrey, C., Takeo, K., & Simon, B. (2010). Multi-pie. *Image and vision computing*, 28(5), 807-813.
- [106] Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M. J. (2006, April). A 3D facial expression database for facial behavior research. In *7th international conference on automatic face and gesture recognition (FGR06)* (pp. 211-216). IEEE.
- [107] Zhang, X., Yin, L., Cohn, J. F., Canavan, S., Reale, M., Horowitz, A., & Liu, P. (2013, April). A high-resolution spontaneous 3d dynamic facial expression database. In *2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG)* (pp. 1-6). IEEE.
- [108] Fabian Benitez-Quiroz, C., Srinivasan, R., & Martinez, A. M. (2016). Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5562-5570).
- [109] Li, S., Deng, W., & Du, J. (2017). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2852-2861).
- [110] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1), 18-31.
- [111] Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2018). From facial expression recognition to interpersonal relation prediction. *International Journal of Computer Vision*, 126, 550-569.
- [112] Cheng, S., Kotsia, I., Pantic, M., & Zafeiriou, S. (2018). 4dfab: A large scale 4d database for facial expression analysis and biometric applications. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5117-5126).



- [113] Martin, O., Kotsia, I., Macq, B., & Pitas, I. (2006, April). The eNTERFACE'05 audio-visual emotion database. In 22nd international conference on data engineering workshops (ICDEW'06) (pp. 8-8). IEEE.
- [114] Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior research methods*, 40(1), 109-115.
- [115] Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion*, 24(8), 1377-1388.
- [116] Happy, S. L., Patnaik, P., Routray, A., & Guha, R. (2015). The Indian spontaneous expression database for emotion recognition. *IEEE Transactions on Affective Computing*, 8(1), 131-142.
- [117] Yan, W. J., Wu, Q., Liu, Y. J., Wang, S. J., & Fu, X. (2013, April). CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces. In 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG) (pp. 1-7). IEEE.
- [118] Yan, W. J., Li, X., Wang, S. J., Zhao, G., Liu, Y. J., Chen, Y. H., & Fu, X. (2014). CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PloS one*, 9(1), e86041.
- [119] Li, X., Pfister, T., Huang, X., Zhao, G., & Pietikäinen, M. (2013, April). A spontaneous micro-expression database: Inducement, collection and baseline. In 2013 10th IEEE International Conference and Workshops on Automatic face and gesture recognition (fg) (pp. 1-6). IEEE.
- [120] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- [121] Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological bulletin*, 115(1), 102.
- [122] Keltner, D., & Kring, A. M. (1998). Emotion, social function, and psychopathology. *Review of General Psychology*, 2(3), 320-342.
- [123] Matsumoto, D. (2011). Reading facial expressions of emotion.
- [124] Gross, J. J., & John, O. P. (1997). Revealing feelings: facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of personality and social psychology*, 72(2), 435.
- [125] Ekman, P., & Friesen, W. V. (1978). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.
- [126] Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, 141(1), 19.
- [127] Hess, U., & Fischer, A. (2013). Emotional mimicry as social regulation. *Personality and social psychology review*, 17(2), 142-157.
- [128] Zafeiriou, S., Zhang, C., & Zhang, Z. (2015). A survey on face detection in the wild: past, present and future. *Computer Vision and Image Understanding*, 138, 1-24.
- [129] Zhao, G., & Pietikainen, M. (2007). Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE transactions on pattern analysis and machine intelligence*, 29(6), 915-928.

- [130] Dhall, A., Goecke, R., Joshi, J., Wagner, M., & Gedeon, T. (2013, December). Emotion recognition in the wild challenge 2013. In *Proceedings of the 15th ACM on International conference on multimodal interaction* (pp. 509-516).
- [131] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.
- [132] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- [133] Zhu, X., & Ramanan, D. (2012, June). Face detection, pose estimation, and landmark localization in the wild. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 2879-2886). IEEE.
- [134] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 423-443.
- [135] Gunes, H., & Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, 31(2), 120-136.
- [136] Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1), 4-37.
- [137] Liu, H., & Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering*, 17(4), 491-502.
- [138] Bouchlaghem, Y., Akhiat, Y., & Amjad, S. (2022). Feature selection: a review and comparative study. In *E3S web of conferences* (Vol. 351, p. 01046). EDP Sciences.
- [139] Narendra, & Fukunaga. (1977). A branch and bound algorithm for feature subset selection. *IEEE Transactions on computers*, 100(9), 917-922.
- [140] Doak, J. (1992). An evaluation of feature selection methods and their application to computer security. *Techninal Report CSE-92-18*.
- [141] Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, 1(1-4), 131-156.
- [142] Visalakshi, S., & Radha, V. (2014, December). A literature review of feature selection techniques and applications: Review of feature selection in data mining. In *2014 IEEE international conference on computational intelligence and computing research* (pp. 1-6). IEEE.
- [143] Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.
- [144] Kononenko, I. (1994, April). Estimating attributes: Analysis and extensions of RELIEF. In *European conference on machine learning* (pp. 171-182). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [145] Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1, 81-106.
- [146] Hall, M. A. (1999). *Correlation-based feature selection for machine learning* (Doctoral dissertation, The University of Waikato).
- [147] Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46, 389-422.
- [148] Raman, B., & Ioerger, T. R. (2002). Instance-based filter for feature selection. *Journal of Machine Learning Research*, 1(3), 1-23.

- [149] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267-288.
- [150] Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2015). A survey on evolutionary computation approaches to feature selection. *IEEE Transactions on evolutionary computation*, 20(4), 606-626.
- [151] Deb, K. (1998). Genetic algorithm in search and optimization: the technique and applications. In *Proceedings of International Workshop on Soft Computing and Intelligent Systems*, (ISI, Calcutta, India) (pp. 58-87). *Proceedings of International Workshop on Soft Computing and Intelligent Systems*, (ISI, Calcutta, India).
- [152] Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp. 1942-1948). *IEEE*.
- [153] Dorigo, M., Birattari, M., & Stutzle, T. (2007). Ant colony optimization. *IEEE computational intelligence magazine*, 1(4), 28-39.
- [154] Soyel, H., Tekguc, U., & Demirel, H. (2011). Application of NSGA-II to feature selection for facial expression recognition. *Computers & Electrical Engineering*, 37(6), 1232-1240.
- [155] Gunes, T., & Polat, E. (2008, October). Feature selection for multi-SVM classifiers in facial expression classification. In *2008 23rd International Symposium on Computer and Information Sciences* (pp. 1-5). *IEEE*.
- [156] Lajevardi, S. M., & Hussain, Z. M. (2012). Automatic facial expression recognition: feature extraction and selection. *Signal, Image and video processing*, 6, 159-169.
- [157] Lajevardi, S. M., & Hussain, Z. M. (2009). Feature extraction for facial expression recognition based on hybrid face regions. *Advances in Electrical and Computer Engineering*, 9(3), 63-67.
- [158] Abdulrazaq, M. B., Mahmood, M. R., Zeebaree, S. R., Abdulwahab, M. H., Zebari, R. R., & Sallow, A. B. (2021, February). An analytical appraisal for supervised classifiers' performance on facial expression recognition based on relief-F feature selection. In *Journal of Physics: Conference Series* (Vol. 1804, No. 1, p. 012055). *IOP Publishing*.
- [159] Dino, H. I., & Abdulrazzaq, M. B. (2020). A comparison of four classification algorithms for facial expression recognition. *Polytechnic Journal*, 10(1), 13.
- [160] Wang, Y., Lu, Y., & Wan, X. (2023, August). Multi-feature fusion based on RV correlation coefficient for facial expression recognition. In *Journal of Physics: Conference Series* (Vol. 2562, No. 1, p. 012027). *IOP Publishing*.
- [161] Li, P., Phung, S. L., Bouzerdom, A., & Tivive, F. H. C. (2010, July). Feature selection for facial expression recognition. In *2010 2nd European Workshop on Visual Information Processing (EUVIP)* (pp. 35-40). *IEEE*.
- [162] Siddiqi, M. H., Ali, R., Idris, M., Khan, A. M., Kim, E. S., Whang, M. C., & Lee, S. (2016). Human facial expression recognition using curvelet feature extraction and normalized mutual information feature selection. *Multimedia Tools and Applications*, 75, 935-959.
- [163] Saha, S., Ghosh, M., Ghosh, S., Sen, S., Singh, P. K., Geem, Z. W., & Sarkar, R. (2020). Feature selection for facial emotion recognition using cosine similarity-based harmony search algorithm. *Applied Sciences*, 10(8), 2816.

- [164] Dong, Y., & Sayama, H. (2019, December). Mutual-information-based feature selection for facial emotion recognition on light-weight devices. In 2019 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 2455-2461). IEEE.
- [165] Yurtkan, K., & Demirel, H. (2014). Entropy-based feature selection for improved 3D facial expression recognition. *Signal, Image and Video Processing*, 8, 267-277.
- [166] Khan, S., Chen, L., & Yan, H. (2017). Co-clustering to reveal salient facial features for expression recognition. *IEEE Transactions on Affective Computing*, 11(2), 348-360.
- [167] Ewees, A. A., ElLaban, H. A., & ElEraky, R. M. (2019, July). Features selection for facial expression recognition. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.
- [168] Lajevardi, S. M., & Hussain, Z. M. (2009, July). Feature selection for facial expression recognition based on optimization algorithm. In 2009 2nd International Workshop on Nonlinear Dynamics and Synchronization (pp. 182-185). IEEE.
- [169] Siddiqi, M. H., Alruwaili, M., & Ali, A. (2019). A novel feature selection method for video-based human activity recognition systems. *IEEE Access*, 7, 119593-119602.
- [170] Mlakar, U., Fister, I., Brest, J., & Potočnik, B. (2017). Multi-objective differential evolution for feature selection in facial expression recognition systems. *Expert Systems with Applications*, 89, 129-137.
- [171] Paharia, N., Jadon, R. S., & Gupta, S. K. (2022). Feature selection using improved multiobjective and opposition-based competitive binary gray wolf optimizer for facial expression recognition. *Journal of Electronic Imaging*, 31(3), 033039-033039.
- [172] Al-Qablan, T. A., Noor, M. H. M., Al-Betar, M. A., & Khader, A. T. (2023). Improved binary gray wolf optimizer based on adaptive  $\beta$ -hill climbing for feature selection. *IEEE Access*, 11, 59866-59881.
- [173] V. Perez-Gomez, H. V. Rios-Figueroa, E. J. Rechy-Ramirez, E. Mezura-Montes, and A. Marin-Hernandez, "Feature selection on 2d and 3d geometric features to improve facial expression recognition," *Sensors* (Switzerland), vol. 20, no. 17, pp. 1–20, Sep. 2020, doi: 10.3390/s20174847.
- [174] Dornaika, F., Lazkano, E., & Sierra, B. (2011). Improving dynamic facial expression recognition with feature subset selection. *Pattern Recognition Letters*, 32(5), 740-748.
- [175] Danisman, T., Bilasco, I. M., Martinet, J., & Djeraba, C. (2013). Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. *Signal Processing*, 93(6), 1547-1556.
- [176] Vignolo, L. D., Milone, D. H., & Scharcanski, J. (2013). Feature selection for face recognition based on multi-objective evolutionary wrappers. *Expert Systems with Applications*, 40(13), 5077-5084.
- [177] Richhariya, B., & Gupta, D. (2019). Facial expression recognition using iterative universum twin support vector machine. *Applied Soft Computing*, 76, 53-67.
- [178] Talaat, F. M., Ali, Z. H., Mostafa, R. R., & El-Rashidy, N. (2024). Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children. *Soft Computing*, 28(9), 6695-6708.

- [179] Mistry, K., Zhang, L., Neoh, S. C., Lim, C. P., & Fielding, B. (2016). A micro-GA embedded PSO feature selection approach to intelligent facial emotion recognition. *IEEE transactions on cybernetics*, 47(6), 1496-1509.
- [180] Ghazouani, H. (2021). A genetic programming-based feature selection and fusion for facial expression recognition. *Applied Soft Computing*, 103, 107173.
- [181] Liu, P., Zhou, J. T., Tsang, I. W. H., Meng, Z., Han, S., & Tong, Y. (2014). Feature disentangling machine-a novel approach of feature selection and disentangling in facial expression analysis. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13* (pp. 151-166). Springer International Publishing.
- [182] Puthanidam, R. V., & Moh, T. S. (2018, January). A hybrid approach for facial expression recognition. In *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication* (pp. 1-8).
- [183] Sidhom, O., Ghazouani, H., & Barhoumi, W. (2024). Three-phases hybrid feature selection for facial expression recognition. *The Journal of Supercomputing*, 80(6), 8094-8128.
- [184] Ghosh, M., Guha, R., Sarkar, R., & Abraham, A. (2020). A wrapper-filter feature selection technique based on ant colony optimization. *Neural Computing and Applications*, 32, 7839-7857.
- [185] Ghosh, M., Kundu, T., Ghosh, D., & Sarkar, R. (2019). Feature selection for facial emotion recognition using late hill-climbing based memetic algorithm. *Multimedia Tools and Applications*, 78, 25753-25779.
- [186] Cossetin, M. J., Nievola, J. C., & Koerich, A. L. (2016, July). Facial expression recognition using a pairwise feature selection and classification approach. In *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 5149-5155). IEEE.
- [187] Shanthi, P., & Nickolas, S. (2021). An efficient automatic facial expression recognition using local neighborhood feature fusion. *Multimedia Tools and Applications*, 80, 10187-10212.
- [188] Makhmudkhujaev, F., Abdullah-Al-Wadud, M., Iqbal, M. T. B., Ryu, B., & Chae, O. (2019). Facial expression recognition with local prominent directional pattern. *Signal Processing: Image Communication*, 74, 1-12.
- [189] Lajevardi, S. M., & Hussain, Z. M. (2009, March). Feature selection for facial expression recognition based on mutual information. In *2009 5th IEEE GCC Conference & Exhibition* (pp. 1-5). IEEE.
- [190] Makhmudkhujaev, F., Iqbal, M. T. B., Arefin, M. R., Ryu, B., & Chae, O. (2018). Person-Independent Facial Expression Recognition with Histograms of Prominent Edge Directions. *KSII Transactions on Internet and Information Systems (TIIS)*, 12(12), 6000-6017.
- [191] Shan, C., Gong, S., & McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. *Image and vision Computing*, 27(6), 803-816.
- [192] Zarbakhsh, P., & Demirel, H. (2017, September). Fuzzy SVM for 3D facial expression classification using sequential forward feature selection. In *2017 9th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 131-134). IEEE.
- [193] Tekguc, U., Soyel, H., & Demirel, H. (2009, September). Feature selection for person-independent 3D facial expression recognition using NSGA-II. In *2009 24th International Symposium on Computer and Information Sciences* (pp. 35-38). IEEE.

- [194] Xue, M., Mian, A., Liu, W., & Li, L. (2014, March). Fully automatic 3D facial expression recognition using local depth features. In IEEE Winter Conference on Applications of Computer Vision (pp. 1096-1103). IEEE.
- [195] Happy, S. L., & Routray, A. (2018). Recognizing subtle micro-facial expressions using fuzzy histogram of optical flow orientations and feature selection methods. *Computational Intelligence for Pattern Recognition*, 341-368.
- [196] Zhou, Y., & Shi, B. E. (2017, May). Action unit selective feature maps in deep networks for facial expression recognition. In 2017 international joint conference on neural networks (IJCNN) (pp. 2031-2038). IEEE.
- [197] Xie, W., Shen, L., Yang, M., & Lai, Z. (2017). Active AU based patch weighting for facial expression recognition. *Sensors*, 17(2), 275.
- [198] Mlakar, U., & Potočnik, B. (2015). Automated facial expression recognition based on histograms of oriented gradient feature vector differences. *Signal, Image and Video Processing*, 9, 245-253.
- [199] Shojaeilangari, S., Yun, Y. W., & Khwang, T. E. (2011, December). Person independent facial expression analysis using Gabor features and Genetic Algorithm. In 2011 8th International Conference on Information, Communications & Signal Processing (pp. 1-5). IEEE.
- [200] Silva, C., Sobral, A., & Vieira, R. T. (2014, October). An automatic facial expression recognition system evaluated by different classifiers. In X Workshop de Visão Computacional, at Uberlândia, Minas Gerais, Brazil (pp. 208-212).
- [201] Ekman, P. (1999). Facial expressions. *Handbook of cognition and emotion*, 16(301), e320.
- [202] Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of experimental psychology: human perception and performance*, 4(3), 373.
- [203] Xiong, H., Zhang, D., Martyniuk, C. J., Trudeau, V. L., & Xia, X. (2008). Using generalized procrustes analysis (GPA) for normalization of cDNA microarray data. *Bmc Bioinformatics*, 9, 1-13.
- [204] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.
- [205] Lewis, D. D. (1998, April). Naive (Bayes) at forty: The independence assumption in information retrieval. In *European conference on machine learning* (pp. 4-15). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [206] Hastie, T. (2009). *The elements of statistical learning: data mining, inference, and prediction*.
- [207] Suguna, N., & Thanushkodi, K. (2010). An improved k-nearest neighbor classification using genetic algorithm. *International Journal of Computer Science Issues*, 7(2), 18-21.
- [208] Xanthopoulos, P., Pardalos, P. M., & Trafalis, T. B. (2012). *Robust data mining*. Springer Science & Business Media.
- [209] Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of applied science and technology trends*, 2(01), 20-28.
- [210] Khalid, S., Khalil, T., & Nasreen, S. (2014, August). A survey of feature selection and feature extraction techniques in machine learning. In 2014 science and information conference (pp. 372-378). IEEE.

- [211] Miao, J., & Niu, L. (2016). A survey on feature selection. *Procedia computer science*, 91, 919-926.
- [212] Liu, S., & Motani, M. (2025). Improving Mutual Information based Feature Selection by Boosting Unique Relevance. *Journal of Artificial Intelligence Research*, 82, 1267-1292.
- [213] Choudhary, D., & Shukla, J. (2020, September). Feature extraction and feature selection for emotion recognition using facial expression. In *2020 IEEE sixth international conference on multimedia big data (BigMM)* (pp. 125-133). IEEE.
- [214] Mortezaie, Z., & Hassanpour, H. (2019). A survey on age-invariant face recognition methods. *Jordanian Journal of Computers and Information Technology*, 5(2).
- [215] Jiang, M., & Yin, S. (2023). Facial expression recognition based on convolutional block attention module and multi-feature fusion. *International journal of computational vision and robotics*, 13(1), 21-37.
- [216] Jiang, Y., & Yin, S. (2023). Heterogenous-view occluded expression data recognition based on cycle-consistent adversarial network and K-SVD dictionary learning under intelligent cooperative robot environment. *Computer Science and Information Systems*, 20(4), 1869-1883.
- [217] Meng, X., Wang, X., Yin, S., & Li, H. (2023). Few-shot image classification algorithm based on attention mechanism and weight fusion. *Journal of Engineering and Applied Science*, 70(1), 14.
- [218] Teng, L. (2023). Brief Review of Medical Image Segmentation Based on Deep Learning [J]. *IJLAI Transactions on Science and Engineering*, 1(02), 01-08.
- [219] El Aboudi, N., & Benhlima, L. (2016, September). Review on wrapper feature selection approaches. In *2016 international conference on engineering & MIS (ICEMIS)* (pp. 1-5). IEEE.
- [220] Hacine-Gharbi, A., Petit, M., Ravier, P., & N mo, F. (2015, January). Prosody based Automatic Classification of the Uses of French ‘Oui’ as Convinced or Unconvinced Uses. In *International Conference on Pattern Recognition Applications and Methods (Vol. 2, pp. 349-354)*. SCITEPRESS.
- [221] Meziane, M. N., Hacine-Gharbi, A., Ravier, P., Lamarque, G., Le Bunetel, J. C., & Raingeaud, Y. (2017, February). Electrical appliances identification and clustering using novel turn-on transient features. In *6th International Conference on Pattern Recognition Applications and Methods* (pp. 647-654). SCITEPRESS-Science and Technology Publications.
- [222] Hacine-Gharbi, A., & Ravier, P. (2018, January). Wavelet Cepstral Coefficients for Electrical Appliances Identification using Hidden Markov Models. In *ICPRAM* (pp. 541-549).
- [223] Aifanti, N., Papachristou, C., & Delopoulos, A. (2010, April). The MUG facial expression database. In *11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10* (pp. 1-4). IEEE.
- [224] Bellman, R. (1958). Dynamic programming and stochastic control processes. *Information and control*, 1(3), 228-239.
- [225] Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, 7, 39-55.
- [226] Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8), 1226-1238.

- [227] Mirjalili, S. (2016). SCA: a sine cosine algorithm for solving optimization problems. *Knowledge-based systems*, 96, 120-133.
- [228] Klein, M. W., Enkrich, C., Wegener, M., & Linden, S. (2006). Second-harmonic generation from magnetic metamaterials. *Science*, 313(5786), 502-504.
- [229] Liu, H., & Setiono, R. (1995, November). Chi2: Feature selection and discretization of numeric attributes. In *Proceedings of 7th IEEE international conference on tools with artificial intelligence* (pp. 388-391). Ieee.
- [230] Clavel, C., & Ehrette, T. (2008). Fear-type emotion recognition and abnormal events detection for an audio-based surveillance system. *WIT Transactions on Information and Communication Technologies*, 39, 471-479.
- [231] Kołakowska, A., Landowska, A., Szwoch, M., Szwoch, W., & Wrobel, M. R. (2014). Emotion recognition and its applications. *Human-Computer Systems Interaction: Backgrounds and Applications* 3, 51-62.
- [232] Martin, D., O'Neill, M., Hubbard, S., & Palmer, A. (2008). The role of emotion in explaining consumer satisfaction and future behavioural intention. *Journal of Services Marketing*, 22(3), 224-236.
- [233] Huang, Y., Yan, Y., Chen, S., & Wang, H. (2018). Expression-targeted feature learning for effective facial expression recognition. *Journal of Visual Communication and Image Representation*, 55, 677-687.
- [234] Dino, H., Abdulrazzaq, M. B., Zeebaree, S. R., Sallow, A. B., Zebari, R. R., Shukur, H. M., & Haji, L. M. (2020). Facial expression recognition based on hybrid feature extraction techniques with different classifiers. *TEST Engineering & Management*, 83, 22319-22329.
- [235] Youssif, A. A., & Asker, W. A. (2011). Automatic facial expression recognition system based on geometric and appearance features. *Computer and Information Science*, 4(2), 115.
- [236] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60, 91-110.
- [237] Gabor, D. (1946). Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering*, 93(26), 429-441.
- [238] Lajevardi, S. M., & Hussain, Z. M. (2009, February). Facial expression recognition using log-Gabor filters and local binary pattern operators. In *Proceedings of the International Conference on Communication, Computer and Power* (pp. 349-353).
- [239] Nazir, M., Jan, Z., & Sajjad, M. (2018). Facial expression recognition using histogram of oriented gradients based transformed features. *Cluster Computing*, 21, 539-548.
- [240] Dosodia, P., Poonia, A., Gupta, S. K., & Agrwal, S. L. (2015, April). New Gabor-DCT feature extraction technique for facial expression recognition. In *2015 Fifth International Conference on Communication Systems and Network Technologies* (pp. 546-549). IEEE.
- [241] Moore, S., & Bowden, R. (2011). Local binary patterns for multi-view facial expression recognition. *Computer vision and image understanding*, 115(4), 541-558.
- [242] Karataş, F., Koyuncu, I., Tuna, M., Alçın, M., Avcioglu, E., & Akgul, A. (2022). Design and implementation of arrhythmic ECG signals for biomedical engineering applications on FPGA. *The European Physical Journal Special Topics*, 1-16.



- [243] Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2), 273-324.
- [244] Bengacemi, H., Hacine-Gharbi, A., Ravier, P., Abed-Meraim, K., & Buttelli, O. (2021, February). Surface EMG Signal Classification for Parkinson's Disease using WCC Descriptor and ANN Classifier. In *ICPRAM* (pp. 287-294).
- [245] Xiong, L., Zhao, Z., Pan, J., Yang, H., & Wang, W. (2019). Recognition of heart sound based on wavelet cepstrum coefficient and probabilistic neural network. *Hans Journal of Biomedicine*, 9(1), 10-16.
- [246] Mallat, S. (1999). *A wavelet tour of signal processing*. Elsevier.
- [247] Daren, H., Jiufen, L., Jiwu, H., & Hongmei, L. (2001, August). A DWT-based image watermarking algorithm. In *IEEE International Conference on Multimedia and Expo, 2001. ICME 2001.* (pp. 80-80). IEEE Computer Society.
- [248] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7), 674-693.
- [249] Ye, Z., Mohamadian, H., & Ye, Y. (2007, September). Information measures for biometric identification via 2d discrete wavelet transform. In *2007 IEEE International Conference on Automation Science and Engineering* (pp. 835-840). IEEE.
- [250] Zhang, D. (2019). Wavelet transform. In *Fundamentals of image data mining: Analysis, Features, Classification and Retrieval* (pp. 35-44). Cham: Springer International Publishing.
- [251] Ahmed, N., Natarajan, T., & Rao, K. R. (2006). Discrete cosine transform. *IEEE transactions on Computers*, 100(1), 90-93.
- [252] Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., ... & Woodland, P. (1999). *The HTK book*.
- [253] Liu, Q., Chen, C., Zhang, Y., & Hu, Z. (2011). Feature selection for support vector machines with RBF kernel. *Artificial Intelligence Review*, 36(2), 99-115.
- [254] Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing*, 52(3), 397-399.
- [255] Ahmed, F., Bari, H., & Hossain, E. (2014). Person-independent facial expression recognition based on compound local binary pattern (CLBP). *Int. Arab J. Inf. Technol.*, 11(2), 195-203.