

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



UNIVERSITE M'HAMED BOUGARA DE BOUMERDES

FACULTE DES SCIENCES

DEPARTEMENT DE PHYSIQUE

STRUCTURE INFOTRONIQUE

Domaine : **Sciences et techniques**

Filière : **Génie électrique**

Spécialité : **Imagerie et appareillage biomédical**

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du diplôme de

MASTER

Thème

***Détection précoce du Cancer de larynx à partir
de la voix***

Présenté par :

Melle. DROUZI Meriem,

Melle. DAHMANI Manel,

Dr. F. YKHLEF, Maitre de Recherche, CDTA

Promoteur

Dr. M. AMMAR, Maitre de conférences, Université de Boumerdes,

Co-promoteur

Boumerdes

Année universitaire 2015-2016

Remerciements

*Nous remercions avant tout **Dieu** le tout puissant de nous avoir donné courage, volonté et santé pour effectuer ce travail.*

Nos remerciements sont adressés à notre promoteur, Dr. YKHLEF Fayçal, Chef de l'équipe Traitement de Signal et Applications Médicales du C.D.T. A, qui nous a proposé ce thème et suivi au cours de notre stage au C.D.T.A. Nous le remercions pour ses conseils précieux et son soutien et pour le temps qu'il nous a consacré.

Nous tenons à exprimer notre profonde gratitude à Mr. DERBAL Abdeladhime pour son aide et surtout pour son accompagnement au cours de la réalisation de ce travail.

Nous remercions la Direction du Centre de Développement des Technologies Avancées, C.D.T.A et en particulier les dirigeants de la Division Architecture des Systèmes et Multimédias de nous avoir accueilli et d'avoir mis à notre disposition les conditions favorables pour la réalisation de ce présent travail.

Nos remerciements vont également à notre Co-promoteur Dr. AMMAR Mohammed pour sa disponibilité et son aide.

Nous souhaitons adresser nos remerciements les plus sincères à Mr. SAAD ASOUD Elyes pour son aide technique.

Nos sincères remerciements s'adressent également aux membres du jury, président et examinateurs, de nous avoir fait l'honneur d'évaluer ce modeste travail.

Nos sincères remerciements s'adressent aussi à tous les enseignants que nous avons rencontrés tout au long de notre formation au niveau de la faculté des sciences de l'université de M'Hamed BOUGARRA surtout ceux du département "INFOTRONIQUE".

Enfin, nous tenons à exprimer nos remerciements envers toute personne qui a contribué de près ou de loin à la réalisation de ce modeste travail.

DEDICACE

A celui qui fut le plus brave des hommes, m'ouvrant ses bras dans les sombres moments et m'aidant à aller de l'avant vers le meilleur, et qui m'a tant soutenu moralement et matériellement :

Mon père

A celle qui a été toujours à côté de moi avec ses prières tout au long de ma vie et avoir mis à ma disposition les conditions favorables pour réussir dans mes études.

Ma mère

*A mes sœurs **Insaf, Amina** et ma petite sœur **Ikram** et mon frère **Mohammed** qui ont m'aide*

*A ma tante **Nacira** et sa fille **Amina** qui ont été toujours présentes pour m'encourage*

*A ma tante **Fatiha** qui a été présente avec ses conseils et son encouragement*

*A ma meilleure amie et préférée **Amira** qui a été toujours présente surtout dans les moments difficiles à mes côtés à cet effet je l'ai remercié beaucoup*

*A mon amie **Nesrine** qui a m'aide au cours de la réalisation de ce projet avec sa patience et son soutien morale et physique*

A ceux qui m'ont aidé à réaliser ce travail.

A tous ceux qui j'aime et qui m'aime.

*A mon binôme **Manel** pour tous les bons moments qu'on a passé ensemble*

*Sans oublier la promotion **IABIO2016***

Meriem

Dédicace

*Ce mémoire de fin d'études universitaire fruit de faveur de bon dieu est dédié
principalement :*

*A celle qui a attendu éducation et ses encouragements, à ma tendresse mère **Nacira***

*A celui qui m'a indiqué la bonne volonté, mon père **Rabah***

A mes sœurs et mon frère unique, je lui souhaite la bonne chance et que du bonheur

*A ma très chère tante : **Wahiba** et sa famille*

*A toute ma famille chacun son nom et spécialement mon oncle : **Farhat***

Que j'aime beaucoup

*A mon cher mari **Ahmed***

*Ames amies : **IKram, Asma, Amina, Housseem, Brahim***

Et tous les enseignants qui m'ont enseigné chacun son nom

A tous ceux j'aime et qui m'aime

*A mon meilleur binôme **Meriem** et tout sa famille*

*Sans oublier la promotion **IABIO2016***

Manel

Résumé

L'objectif de ce projet est de concevoir un système d'aide au diagnostic en vue de la détection précoce du Cancer de larynx à partir de la voix. Principalement, il s'agit de différencier entre les pathologies précancéreux « Leucoplasie » et bénignes des patients ayant des troubles vocaliques. Le système proposé est basé sur deux étapes principales : (i) l'extraction d'attributs et (ii) la classification. Les attributs discriminatifs que nous avons choisis englobent trois types de paramètres : (i) paramètres du conduit vocal, (ii) signature de la source d'excitation, et (iii) les paramètres de bruit. Le schéma de classification est basé sur les Machines à Vecteur de Supports « SVMs : *Support Vector Machines* ».

L'extraction des attributs est réalisée par deux logiciels de simulation : le logiciel d'analyse de la parole « PRAAT » et MATLAB de Mathworks. L'outil de datamining : « WEKA » a été utilisé pour réaliser la classification.

La base de données SVD a été utilisée dans nos expérimentations. Les mesures de performances utilisées dans cette étude sont : le taux de bonne classification (TBC), le taux de vrais positifs (TVP), le taux de vrais négatifs (TVN), la courbe ROC et l'AUC (Area Under roc Curve). Les résultats obtenus sont satisfaisants.

Mots clés : Cancer de larynx, Leucoplasie, attributs de la voix, machines à vecteurs de supports.

ملخص

الهدف من هذا العمل هو تصميم نظام دعم التشخيص للكشف المبكر عن سرطان الحنجرة عن طريق الموجات الصوتية. أساس هذا نظام هو التفريق بين بداية سرطانية (leucoplasie) والامراض الحميدة للمرضى الذين يعانون من اضطرابات في الاحبال الصوتية .

النظام المقترح يتركز على مرحلتين اساسيتين: (ا) استخراج سمات (ب) التصنيف. السمات المتميزة التي اخترناها تتضمن ثلاث هيئات: (ا) عوامل المسالك الصوتية (ب) التوقيع على مصدر التأثير (ج) عوامل الضوضاء. ويتركز مخطط التصنيف على طريقة شعاع الدعم الالى . عملية استخراج السمات اجرية بواسطة برنامج محاكاة : برنامج التحليل والتصنيف الصوتي Praat و MATLAB اما البرنامج WEKA استخدم كاداة لاستخراج البيانات في مرحلة التصنيف

استخدمت في هذه التجربة قاعدة البيانات SVD. المقاييس المستخرجة في هذه الدراسة هي: معدل التصنيف الصحيح (TBC) معدل إيجابي الحقيقي (TVP) ، ومعدل سلبي الحقيقي (TNP) منحنى ROC و L'AUC .

كلمات المفاتيح : سرطان الحنجرة-السمات الصوت-شعاع الدعم الالى

Liste d'abréviations et acronymes

APQ3 : *Third Point Amplitudes Perturbation Quotient*

APQ5 : *Fifth Point Amplitudes Perturbation Quotient*

APQ11 : *Eleventh Point Amplitudes Perturbation Quotient*

AUC: *Area under Curve*

DDA: *Difference of Difference of Amplitudes*

DDP: *Difference of Difference of Periods*

EGG: *Electroglottograph*

F0: *Fréquence fondamentale*

HNR: *Harmonic to Noise Ratio*

LPC: *Linear Predictive Coding*

MFCC: *Mel Frequency Cepstral Coefficients*

NHR: *Noise to Harmonics Ratio*

NNE: *Normalized Noise Energy*

PA: *Pitch Amplitude*

PPQ5: *Fifth Point Periods Perturbation Quotient*

RAL : *Reconnaissance Automatique du Locuteur*

RAP: *Relative Average Perturbation*

ROC: *Receiver Operating Characteristic*

SNR: *Signal to Noise Ratio*

SVD: *Saarbrücken Voice Database*

SVM: *Support Vector Machines*

TBC: *Taux de Bonne Classification*

TF: *Transformée de Fourier*

TFD: *Transformée de Fourier Discrète*

TFN: *Taux de Faux Négatifs*

TFP: Taux de Faux Positifs

TVN: Taux de Vrais Négatifs

Table des matieres

Remerciements	2
Dedicace	3
Résumé	5
Liste d'abréviations et acronymes	6
Table de matieres	9
Liste des figures	11
Liste des tableaux	11
Chapitre 1 : Introduction	12
1.1 Motivation	12
1.2 Contributions	12
1.3 Organisation du mémoire	13
Chapitre 2 : Généralités sur les troubles de la voix	14
2.1 Introduction	14
2.2 Catégories de troubles de la voix	14
2.2.1 Pathologies bénignes	15
2.2.2 Pathologies malignes (Cancer du larynx)	15
2.2.3 Pathologies précancéreuses	15
2.3 Paramètres acoustiques de la voix	16
2.3.1 Intensité	16
2.3.2 Hauteur	17
2.3.3 Timbre	17
2.4 Bases de données des voix pathologiques	17
2.4.1 Présentation de la Saarbrücken Voice Database « SVD »	18
2.4.2 Pathologies étudiées	18
2.5 Diagnostic des troubles vocaliques	21
2.5.1 Méthodes subjectives	21
2.5.2 Méthodes objectives	22
2.6 Conclusion	23

Chapitre 3 : Détection précoce du Cancer de larynx : Analyse acoustique de la voix et classification	24
3.1 Introduction	24
3.2 Indices acoustiques du Cancer de larynx	24
3.2.1 Paramètres du conduit vocal	25
3.2.2 Paramètres de la source d'excitation.....	27
3.2.3 Paramètres liés au bruit.....	30
3.2.4 Méthodes d'estimation des paramètres acoustiques	31
3.3 Machines à vecteurs de supports	32
3.4 Méthode de détection du Cancer de larynx	39
3.4.1 Données acoustiques.....	39
3.4.2 Extraction d'attributs	39
3.4.3 Classification.....	41
3.5 Métriques d'évaluation.....	41
3.5.1 Matrice de confusion.....	41
3.5.2 Courbe ROC.....	43
3.5.3 Aire sous la courbe (AUC)	43
3.5.4 Validation croisée	44
3.6 Conclusion.....	44
Chapitre 4 : Expérimentations, résultats et interprétations	45
4.1 Introduction	45
4.2 Logiciels de développement	45
4.2.1 Environnement MATLAB	45
4.2.2 PRAAT	45
4.2.3 WEKA (Waikato Environment for Knowledge Analysis)	46
4.3 Résultats expérimentaux	46
4.3.1 Corpus sonore	46
4.3.2 Extraction d'attributs	48
4.3.3 Classification.....	49
4.4 Evaluation de la méthode	49

4.4.1 Matrice de confusion.....	49
4.4.2 Courbe ROC :	50
4.4 Conclusion.....	51
Chapitre 5 : Conclusions et travaux futures	52
5.1 Conclusions	52
5.2 Perspectives.....	52
Bibliographies	53
Annexe A : Sons pathologiques de la base de données SVD.....	58

Liste des figures

Figure 2.1 : Catégories des pathologies de la voix

Figure 2.2 : Différence entre un larynx normal et un cancéreux

Figure 2.3 : Vue de tumeur précancéreuse

Figure 2.4 : Intensité d'un son

Figure 2.5 : Fréquence d'un son

Figure 3.1 : Organigramme des paramètres exploités

Figure 3.2 : Organigramme de l'extraction des coefficients MFCCs

Figure 3.3 : Recherche de l'hyperplan optimal

Figure 3.4 : Hyperplan H optimal, vecteurs supports et marge maximale. x_1 et x_2 représentent les attributs des classes A et B

Figure 3.5 : Séparation linéaire et non linéaire

Figure 3.6 : Exemple graphique des données linéairement séparables

Figure 3.7 : Transformation non linéaire de données

Figure 3.8 : Organigramme de système d'aide au diagnostique

Figure 3.9 : Organigramme d'extraction d'attribut

Figure 3.9 : Principe de la validation croisée

Figure 4.2 : Courbe ROC sur la SVD

Liste des tableaux

Tableau 3.1 : Matrice de confusion

Tableau 4.1 : Groupes des pathologies utilisées

Tableau 4.2 : Matrice d'attributs

Tableau 4.3 : Corpus de test

Tableau 4.4 : Matrice de confusion

Tableau 4.5 : Performances & résultats

Chapitre 1 : Introduction

1.1 Motivation

Le cancer du larynx est fréquent dans le monde et son traitement est devenu un souci du corps médical [34].

Ainsi, la détection de cette pathologie est devenue une nécessité afin de procéder à des traitements médicaux et sauver la vie des patients. Il existe plusieurs méthodes de diagnostic de cette maladie (Techniques médicales spécialisées : méthodes subjectives). Néanmoins, de nos jours, l'expertise médicale acquise par les spécialistes dans ce domaine ne permette pas l'identification précoce de cette pathologie avant qu'elle atteigne un stade avancé d'évolution [41].

Afin de résoudre ce problème, une nouvelle approche objective en vue de la détection précoce du Cancer de larynx a été proposée par des chercheurs [21, 20]. Cette approche est basée sur les techniques de traitement de signal. Les données analysées sont de nature différente. Elles peuvent être des images endoscopiques de larynx, des enregistrements des voix des patients, des acquisitions des signaux cérébraux ou une combinaison de plusieurs données.

1.2 Contributions

Nous contribuons à concevoir un système automatique d'aide au diagnostic qui permet la détection précoce du Cancer de larynx à partir de la voix. Nous nous intéressons à la distinction entre les voix bénignes et précancéreuses (Leucoplasie). À notre connaissance, la distinction entre ces deux catégories en se basant sur l'onde acoustique de la voix des patients n'a pas été encore abordée dans la littérature.

L'approche que nous proposons est basée sur deux niveaux essentiels : (i) l'extraction d'attributs et (ii) la classification. Les attributs discriminatifs que nous avons choisis englobent trois types de paramètres : (i) paramètres du conduit vocal (*Mel Frequency Cepstral Coefficients: MFCCs*), (ii) signature de la source d'excitation (pitch, Jitter,

Shimmer), et (iii) les paramètres de bruit (*Harmonic to Noise Ration* : HNR et *Noise to Harmonic Ration* : NHR).

Le schéma de classification est basé sur les Machines à Vecteur de Supports « SVMs : *Support Vector Machines* ».

L'extraction des paramètres est réalisée par deux logiciels de simulation : le logiciel d'analyse de la parole « PRAAT » et MATLAB de Mathworks. L'outil de datamining : « WEKA » a été utilisé pour réaliser la classification.

La base de données Saarbrücken Voice Database (SVD) a été utilisée dans la partie expérimentale pour mesurer : le TBC (Taux de Bonne Classification), TVP (Taux de Vrai Positif), TVN (Taux de Vrai Négatif), la courbe ROC (Receive Operating Characteristic) et l'AUC (Area Under Curve).

1.3 Organisation du mémoire

Le mémoire est organisé comme suit :

Le deuxième chapitre : présent les types des trouble de la voix et les paramètres acoustiques utilisées dans la littérature. Il présent aussi une étude sur l'état de l'art des méthodes de détection du cancer de larynx ainsi que les bases de données des voix pathologiques.

Le troisième chapitre : est consacré à la méthode proposée.

Le quatrième chapitre : comprend les résultats expérimentaux.

Le cinquième chapitre : comprend une conclusion générale et des perspectives.

Chapitre 2 : Généralités sur les troubles de la voix

2.1 Introduction

La voix est le son produit par la vibration des cordes vocales, dans le larynx, également connu sous le nom de la boîte vocale. Un trouble de la voix se produit lorsque les cordes vocales ne produisent pas un son clair. Parfois, ces conditions sont appelées troubles du larynx ou troubles de la boîte vocale.

Dans ce chapitre nous présentons dans une première partie les catégories des troubles vocaliques ainsi que les différents paramètres acoustiques de la voix. Par la suite, une étude sur les bases de données des voix pathologiques est présentée. Les méthodes de diagnostic des troubles vocaliques sont aussi présentées. Nous terminerons ce chapitre par une conclusion.

2.2 Catégories de troubles de la voix

Les troubles de la voix (voix pathologiques) se répartissent en trois grandes familles : organiques, fonctionnels ou une combinaison des deux [31]. Ces troubles peuvent être des pathologies malignes, bénignes ou précancéreuses [19].

Les niveaux fonctionnels de pathologies que nous adoptons dans notre travail sont représentés sur la figure 2.1. Nous nous intéressons dans notre travail sur le 2^{ème} niveau. Plus précisément, nous nous intéressons à la distinction entre les voix bénignes et précancéreuses.

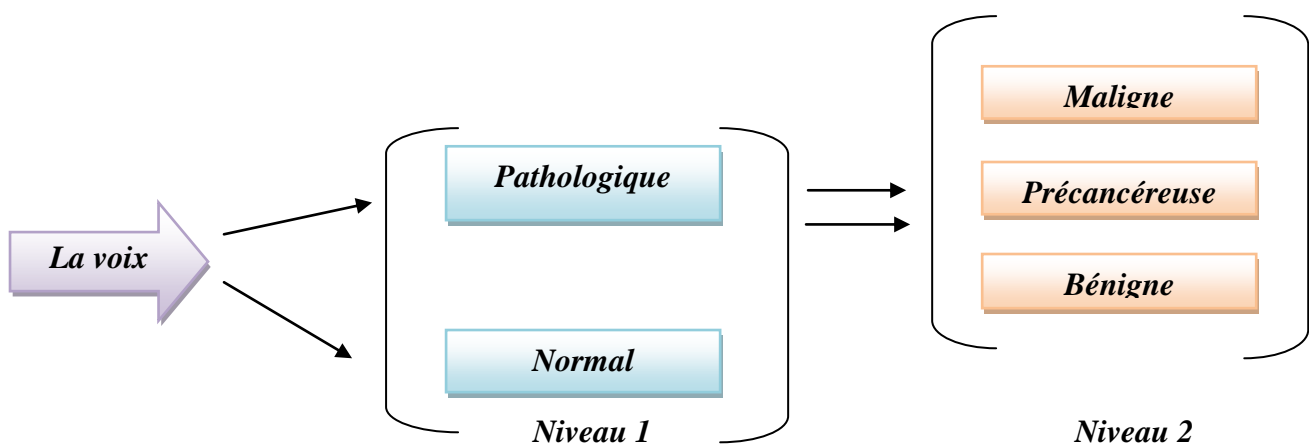


Figure 2.1 : Catégories des pathologies de la voix

2.2.1 Pathologies bénignes

Une affection bénigne est un changement non cancéreux de la surface, ou du revêtement, des structures du larynx [1].

2.2.2 Pathologies malignes (Cancer du larynx)

Une tumeur¹ maligne du larynx est une masse cancéreuse qui peut se propager vers d'autres parties du corps (métastases). Le type de cancer du larynx le plus courant est le carcinome épidermoïde. Il représente 95 % de tous les types du cancer de larynx [1].

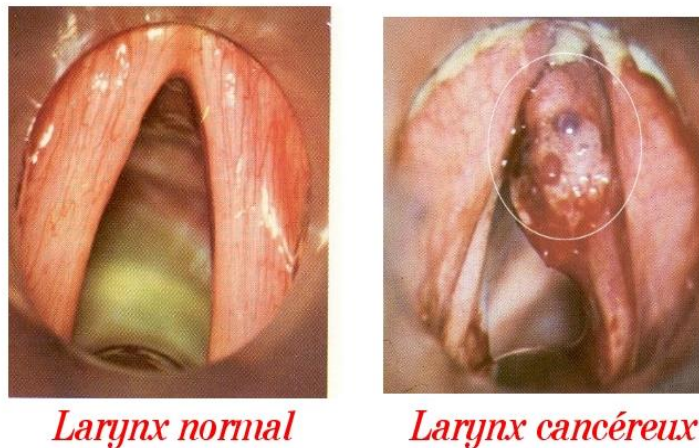


Figure 2.2 : Différence entre un larynx normal et un cancéreux [39]

2.2.3 Pathologies précancéreuses

Un état précancéreux est la forme que prend un tissu, organe ou partie d'organe, où est susceptible de se développer un cancer propre à ce tissu ou cet organe. Il peut se reconnaître par une lésion, dite précancéreuse, au niveau d'un épithélium [33].

¹ Tumeur : trouble de la voix



Figure2.3 : Vue de tumeur précancéreuse [40].

2.3 Paramètres acoustiques de la voix

L'analyse acoustique complète l'évaluation perceptive de la fonction vocale, et vise une meilleure approche sur le plan de diagnostic. De plus, l'analyse acoustique de la parole, constitue une méthode non invasive permettant de quantifier la fonction vocale. Elle comprend l'étude des paramètres acoustiques suivants :

2.3.1 Intensité

L'intensité est la sensation auditive qui permet de faire la différence entre un son fort et un son faible. Elle est exprimée en décibels, qui elle aussi varie dans la parole Elle est réglée par la pression d'air sous glottique [42, 44].

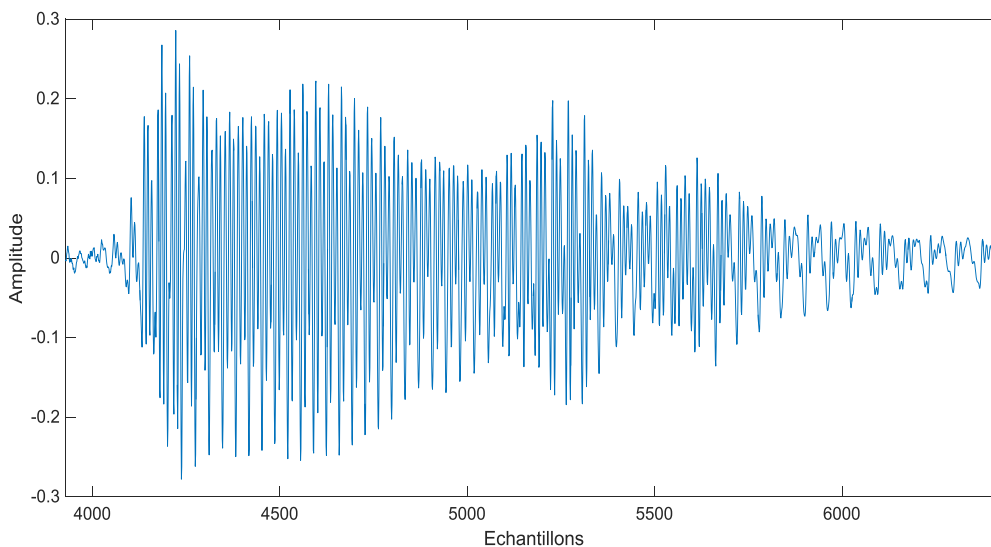


Figure2.4 : Intensité d'un son.

2.3.2 Hauteur

La hauteur ou fréquence ou tonalité est la sensation auditive liée à la fréquence de vibration d'un son. Elle est exprimée en Hertz.

La fréquence fondamentale est le nombre de cycles d'ouverture-fermeture des cordes vocales par seconde, rendant compte d'une voix grave, médium ou aigue naturellement. Dans la parole, la hauteur de la voix varie en permanence, c'est l'intonation, qui lorsqu'elle est riche confère à la voix des qualités esthétiques et de communication [42, 43].

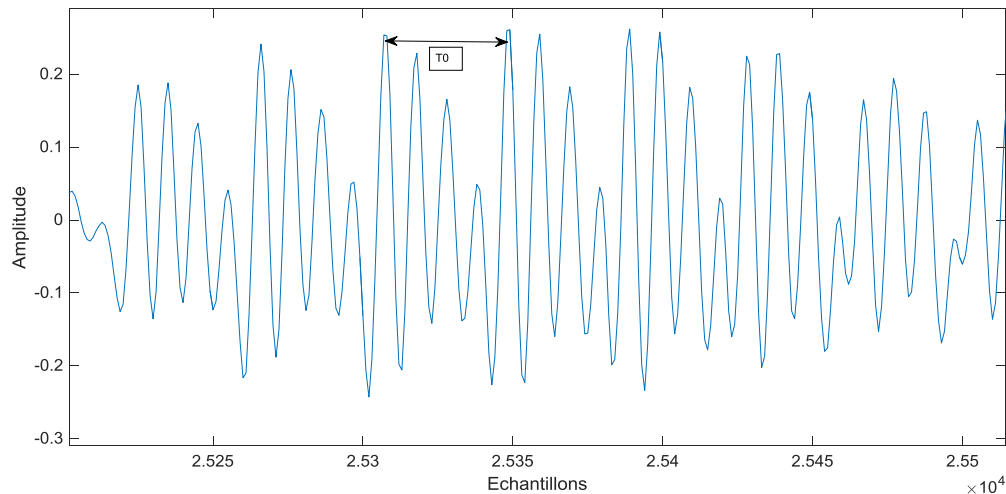


Figure2.5 : Période d'un son $F_0=1/T_0$.

3.3.3 Timbre

Le timbre est un indice qui permet d'identifier une voix et de la reconnaître. La mobilité du voile du palais et de la langue, la tonicité des lèvres et des joues jouent un rôle important dans sa richesse, de même que la qualité de la muqueuse qui tapisse les cavités de résonance [42].

2.4 Bases de données des voix pathologiques

Plusieurs bases de données ont été utilisées dans la littérature pour la tâche de classification des voix saines et pathologiques. Nous pouvons les classer en données publiques et privées. Les bases de données publiques peuvent être librement ou commercialement distribuées pour la communauté scientifique. Cependant, les bases de données privées ne sont pas disponibles puisqu'elles ont été enregistrées localement. Comme il est impossible de comparer les résultats de la recherche lorsque les expériences sont

réalisées avec des données privées, nous pouvons affirmer que les données publiques sont plus utilisées car elles permettent des comparaisons rapides [13].

Il existe plusieurs base de données dans la littérature, nous citons : MAPACI [37], MEEI : Massachusetts Eye & Ear Infirmary [37] et la SVD [27].

Nous avons travaillé avec la SVD car elle est disponible et publiée en ligne gratuitement.

2.4.1 Présentation de la Saarbrücken Voice Database « SVD »

Saarbrücken Voice Database (SVD) est une base de données Allemande qui a été récemment mise gratuitement à disposition en ligne. Elle est une collection des enregistrements vocaux à partir de plus de 2000 personnes, où une session est définie comme une collection de :

- Enregistrements des voyelles / a /, / i /, / u / produit à la hauteur normale, haut, bas et haut -bas.
- Enregistrement de phrase "Guten Morgen, wie geht es Ihnen?" ("Bonjour Comment allez-vous?").

La taille des fichiers qui contiennent les voyelles soutenues est comprise entre 1 et 3 secondes. Tous les enregistrements sont échantillonnés à 50 kHz et leur résolution est de 16 bits. 71 différentes pathologies sont contenues, y compris à la fois fonctionnelle et organique [20].

Nous nous intéressons dans notre travail aux pathologies décrites dans la section suivante.

Pathologies étudiées

a) Pathologies bénignes

a.1) Paralysie des cordes vocales

La paralysie des cordes vocales (Vocal fold paralysis) est causée par une paralysie du nerf récurrent (nerf laryngé inférieur) et apparaît souvent comme une complication dans les opérations de la thyroïde. On peut avoir une paralysie unilatérale ou bilatérale des cordes vocales. Puisque les muscles laryngés intérieurs ne fonctionnent pas, les cordes vocales

paralysées ne peuvent pas être amenées en position de phonation. Au lieu de cela, les muscles laryngés extérieurs (muscle crico-thyroïdien), qui ne sont pas affectés par la paralysie, attirent les cordes vocales en position médiane ou paramédiane par tension vers la ligne du milieu. Dans le cas de paralysie des deux muscles laryngés inférieurs et supérieurs, les cordes vocales paralysées restent en position intermédiaire, avec ou sans excavation. Les conditions de la phonation se détériorent avec l'ampleur de la paralysie. Si seulement le muscle laryngé supérieur est paralysé, la phonation est possible. Les cordes vocales paralysées sont habituellement relâchées et montrent une légère excavation de leurs bords libres [19].

a.2) Syndrome de Parkinson

Typiquement de la maladie de Parkinson, c'est l'incapacité à produire une adduction suffisante des cordes vocales, ce qui conduit à une voix soufflée, douce et monotone. Un tremblement des cordes vocales et, éventuellement, de l'épiglotte, le voile du palais et la langue est souvent présent, outre un déficit de l'adduction. La durée de la phonation est courte [19].

a.3) Laryngite

La laryngite chronique (Laryngitis) présente habituellement une modification ou extinction de la voix, qui apparaît soit trop aiguë, soit trop grave. L'inflammation des cordes vocales (une couche d'épithélium nuisant à la capacité des cordes à vibrer) diminue la tonalité moyenne et réduit les fréquences les plus élevées. Le gonflement des cordes vocales et leur épaisseur irrégulière réduisent la mobilité des bords des cordes, conduisant à une vibration irrégulière des cordes vocales ou même une aphonie [19].

a.4) Oedème de Reinke

L'oedème de Reinke (Reinke odema) des cordes vocales est une forme particulière d'une laryngite chronique hypertrophique, caractérisée par une accumulation de liquide sous l'épithélium des cordes vocales, causant une dysphonie avec une voix d'une tonalité exceptionnellement basse. Cette tonalité est causée par les mouvements de battement de la muqueuse hypertrophique. La fermeture irrégulière est causée par le contact de la muqueuse hypertrophique plutôt que les masses réelles des cordes vocales. La phase de fermeture est relativement longue [19].

a.5) Nodules des cordes vocales

Les Nodules des cordes vocales (Vocal fold nodules) sont causés par la phonation continue avec un niveau de volume très élevé (crier par les enfants par exemple) indiquant un abus d'utilisation des cordes vocales.

Dans ce cas, la symptomatologie d'une dysphonie hyperfonctionnelle existe. La présence de nodules est liée à une voix rauque, éraillée. Les nodules de grande taille, surtout s'ils se trouvent exactement en face de l'autre sur les cordes vocales, provoquent une fermeture incomplète des cordes devant et derrière les nodules (« glotte sablier »). Ils réduisent la mobilité des bords des cordes vocales et conduisent à une petite amplitude des vibrations [19].

a.6) Chondrome

Un chondrome (Chondoma) est une tumeur dans la partie cartilagineuse du larynx. Elle se développe préférentiellement au niveau du cricoïde et se manifeste par une dyspnée d'installation progressive. Avec l'augmentation de la taille de la tumeur, la vibration des cordes vocales peut être affectée, soit par les changements mécaniques dans les cordes, soit par la provocation d'une lésion du nerf récurrent [19].

a.7) Laryngocele

Un laryngocele (Laryngocele) est une expansion du Sinus Morgagni (une couche fibreuse qui tapisse la surface interne des muscles du pharynx et est particulièrement épais dans la partie supérieure - entre constricteur supérieur et la base du crâne[44]) qui peut être dirigée vers l'intérieur (vers le larynx) ou vers l'extérieur (entre le bord du cartilage thyroïde et l'os de la langue). Si le laryngocèle se prolonge dans le larynx, il pousse les fausses cordes vocales ou le pli de l'épiglotte en avant, provoquant une rudesse, des problèmes respiratoires et des problèmes de déglutition [19].

b) Pathologies précancéreuses

b.1) Leucoplasie

Dans les deux cas, il s'agit d'une hyperplasie de l'épithélium des cordes vocales, conduisant à un durcissement dans les couches supérieures de l'épithélium. La capacité des cordes vocales de vibrer est directement liée à l'épaisseur de l'épithélium : si l'épaisseur augmente, la mobilité diminue conduisant à une réduction de la capacité des cordes vocales à vibrer. La voix est caractérisée par une faible tonalité [19].

2.4 Diagnostic des troubles vocaliques

Le diagnostic des troubles vocaliques est basé sur deux catégories de méthodes principales :

- Méthodes subjectives (procédures médicales),
- Méthodes objectives (Algorithmes de traitement de signal).

2.4.1 Méthodes subjectives

La présence d'un cancer du larynx peut être soupçonnée par le médecin après avoir vérifié les antécédents médicaux et une profonde examination. Pour confirmer le diagnostic, le médecin aura recours à certaines analyses, qui pourront également permettre d'établir le « stade » du cancer afin de planifier le traitement. Le médecin pourra avoir recours aux examens suivants [18]:

Laryngoscopie : Cette technique permet d'examiner la gorge et le larynx à l'aide d'un tube étroit et flexible (laryngoscope), muni d'une lumière à son extrémité, que le médecin insère par le nez ou la bouche. La laryngoscopie se fait généralement sous anesthésie locale ; vous recevrez peut-être aussi un léger sédatif pour vous aider à vous détendre. Vous aurez sans doute mal à la gorge après l'examen, mais il s'agit d'un effet normal qui disparaîtra au bout d'un jour ou deux

Biopsie : Si le médecin remarque quelque chose d'anormal au cours d'une laryngoscopie, il pourra prélever plusieurs échantillons de tissus à l'aide du laryngoscope, pour ensuite les examiner au microscope. Cette procédure porte le nom de biopsie. Une biopsie est habituellement nécessaire pour établir avec certitude un diagnostic de cancer. Si les cellules sont cancéreuses, il faudra ensuite déterminer leur rapidité à se multiplier. Si des tissus doivent être prélevés, l'intervention se déroulera peut-être sous anesthésie générale.

Techniques d'imagerie : Ces techniques permettent de procéder à un examen approfondi des tissus, des organes et des os. La radiographie, la tomодensitométrie, la tomographie par émission de positrons (TEP) et l'imagerie par résonance magnétique (IRM) sont autant de moyens pour l'équipe soignante d'obtenir une image de la tumeur et de vérifier si elle s'est étendue. Ces tests sont généralement sans douleur et ne nécessitent aucune anesthésie.

2.4.2 Méthodes objectives

Pour assurer un diagnostic fiable d'une maladie, un logiciel de diagnostic aide le médecin à réduire au minimum les erreurs possibles qui peuvent survenir pendant l'étape de diagnostic. Le logiciel analyse les différentes valeurs des paramètres pris en considération pour le diagnostic des maladies.

Les études de la littérature montrent que la détection précoce de cancer de larynx représenté par la leucoplasie peut être effectuée par la combinaison des paramètres acoustiques et les paramètres reliés au bruit.

Z. Chen et al [22] ont montré que le Jitter, Shimmer, *Noise Normalized Energy* (NNE) et *Harmonic to Noise Ration* (HNR) sont des indicateurs sensibles pour l'évaluation de la voix des patients du leucoplasie.

Le principe de cette étude est d'utiliser l'analyse acoustique en comparant les valeurs de Jitter, Shimmer, NNE et HNR des 63 patients atteints de la leucoplasie (47 hommes, 16 femmes) avec des locuteurs normaux (30 femmes et 30 hommes) ainsi que ceux des patients avec polype (30 femmes, 30 homme) en exploitant le logiciel Dr. Speech. Les résultats trouvés sont les suivants :

- Les valeurs de jitter et de shimmer des patients males avec leucoplasie étaient évidemment plus élevées que ceux de cas normal et des patients avec polype.
- Les valeurs de l'HNR des patients males avec leucoplasie étaient inférieures à ceux des témoins normaux et supérieures à ceux des patients ayant un polype.
- Les valeurs de NNE des patients males avec leucoplasie étaient inférieures à ceux des témoins normaux et des patients avec polype.
- Les valeurs de F_0 des patients males avec leucoplasie étaient similaires à ceux des témoins normaux et les patients avec polype.
- Des résultats similaires ont été trouvés pour les femmes atteintes de leucoplasie.

Par ailleurs, d'autres chercheurs [21] ont exploité les paramètres à long terme (jitter, shimmer) et les paramètres de bruit : NNE, HNR et SNR (Signal to Noise Ratio) avec les paramètres acoustiques de EGG (electroglottograph) : EGG-NNE, EGG-HNR, EGG-SNR, EGG-jitter, EGG-shimmer afin de fournir des indicateurs objective pour évaluer le degré d'enrouement et de fournir des preuves sonores pour la prévention, la récurrence, l'évaluation et la correction de prononciation des cas de la leucoplasie de larynx.

Le principe de cette étude est d'utiliser l'analyse acoustique en comparant les valeurs de Jitter, Shimmer, NNE, HNR et SNR. De plus, une comparaison du signal EEG des 48 patients atteints de la leucoplasie, avec 50 locuteurs normaux a été faite en exploitant le logiciel Dr. Speech. Les résultats trouvés sont les suivants :

- Pour l'analyse acoustique de la voix : Jitter, Shimmer, NNE dans le cas de groupe de leucoplasie était plus élevé que ceux du groupe normal.
- Les valeurs de HNR et SNR dans le groupe de leucoplasie étaient inférieures à ceux du groupe normal.
- Pour l'analyse de EGG : EGG- jitter, EGG- shimmer, EGG-NNE étaient plus élevées dans le cas de larynx leucoplasie que ceux du groupe normal avec une différence significative.
- EGG- HNR, EGG- SNR, étaient inférieures dans le cas de larynx leucoplasie que ceux du groupe normal

2.5 Conclusion

Dans ce chapitre, nous avons présenté les catégories des troubles de la voix ainsi que les paramètres acoustiques de la voix. Par la suite, nous avons établi une étude sur les bases de données des voix pathologiques. Nous avons aussi abordé le diagnostic de ces troubles. Dans le prochain chapitre, nous présentons la méthode de détection précoce du cancer de larynx que nous avons proposée.

Chapitre 3 : Détection précoce du Cancer de larynx : Analyse acoustique de la voix et classification

3.1 Introduction

La détection du Cancer à une phase précoce de son développement augmente considérablement les chances de réussite du traitement. Elle repose sur le dépistage, mais aussi sur l'éducation des patients au diagnostic précoce : les « signes d'alerte ».

Certains types de Cancers sont accessibles à des tests simples permettant de recenser, dans une population en bonne santé, les sujets qui ont un cancer mais ne présentent pas encore de symptômes. C'est le cas, en particulier, du Cancer de larynx.

Dans ce chapitre, nous présentons la méthode d'aide au diagnostic que nous avons proposé. Nous nous intéressons dans une première partie aux indices acoustiques du Cancer de larynx qui sont divisés en paramètres du conduit vocal, paramètres de bruit et de la source d'excitation. Dans la deuxième partie, les machines à vecteur de supports sont présentées. Par la suite, nous présentons la méthode d'aide au diagnostic qui permet de différencier entre les pathologies précancéreux « Leucoplasie » et les pathologies bénignes. Nous terminons ce chapitre par une conclusion.

3.2 Indices acoustiques du Cancer de larynx

L'analyse de la parole du patient est aujourd'hui une technique très utilisée pour la détection des pathologies de la voix, ces pathologies peuvent être identifiées par l'analyse de plusieurs paramètres de signaux acoustiques.

Dans notre étude nous exploiterons quelques paramètres du conduit vocal, d'autres paramètres de la source d'excitation et d'autres de bruit afin de construire de vecteur d'attributs acoustiques. La figure 3.1 présente ces paramètres :

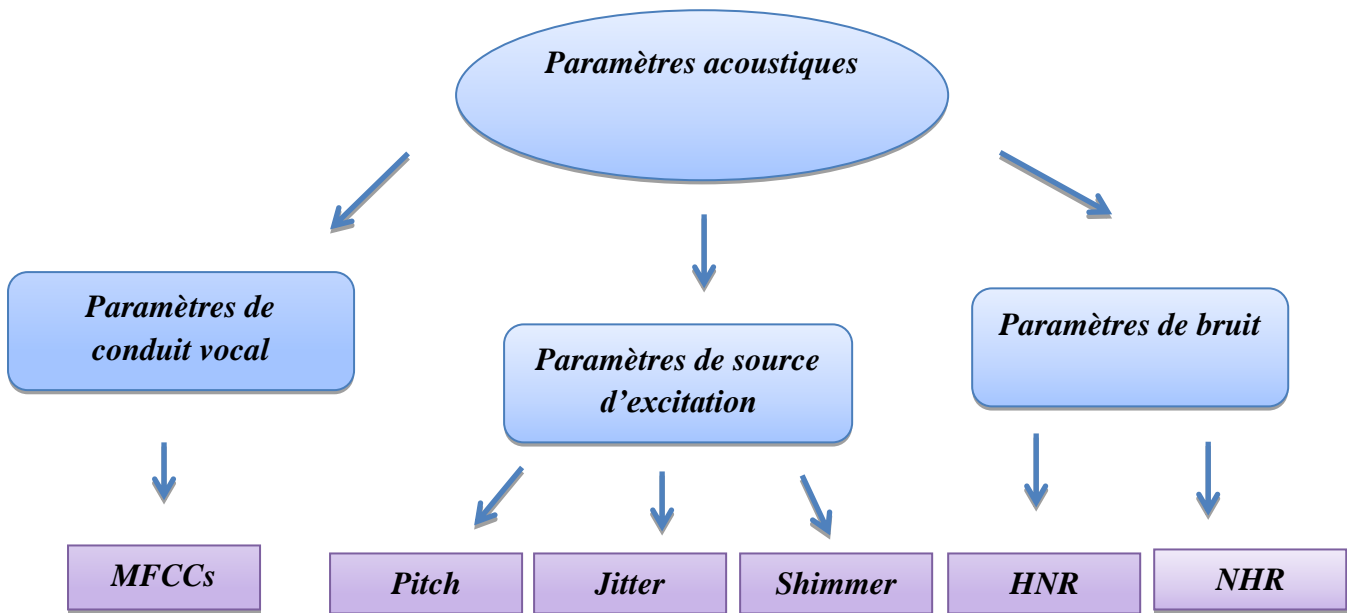


Figure3.1 : Organigramme des paramètres exploités

3.2.1 Paramètres du conduit vocal

a) MFCCs (Mel-scaled Frequency Cepstral Coefficients)

Les MFCCs (Mel-scaled Frequency Cepstral Coefficients) représentent les traits les plus utilisés dans les systèmes de reconnaissance du locuteur. Leur principe de calcul est issu des recherches psycho-acoustiques sur la perception des différentes bandes de fréquences par l'oreille humaine.

Le calcul des MFCCs est en effet basé sur une échelle de Mel qui est linéaire pour les fréquences inférieures à 1000 Hz et logarithmique pour les fréquences supérieures. L'échelle de Mel redistribue les fréquences en fonction de la fréquence perçue [3]. Elle peut être approximée par la formule de conversion suivante :

$$f_{mel} = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \quad (3.1)$$

Avec : f_{mel} fréquence de Mel

f Fréquence perçue

Les étapes de l'extraction des coefficients MFCC sont représentées par l'organigramme suivant :

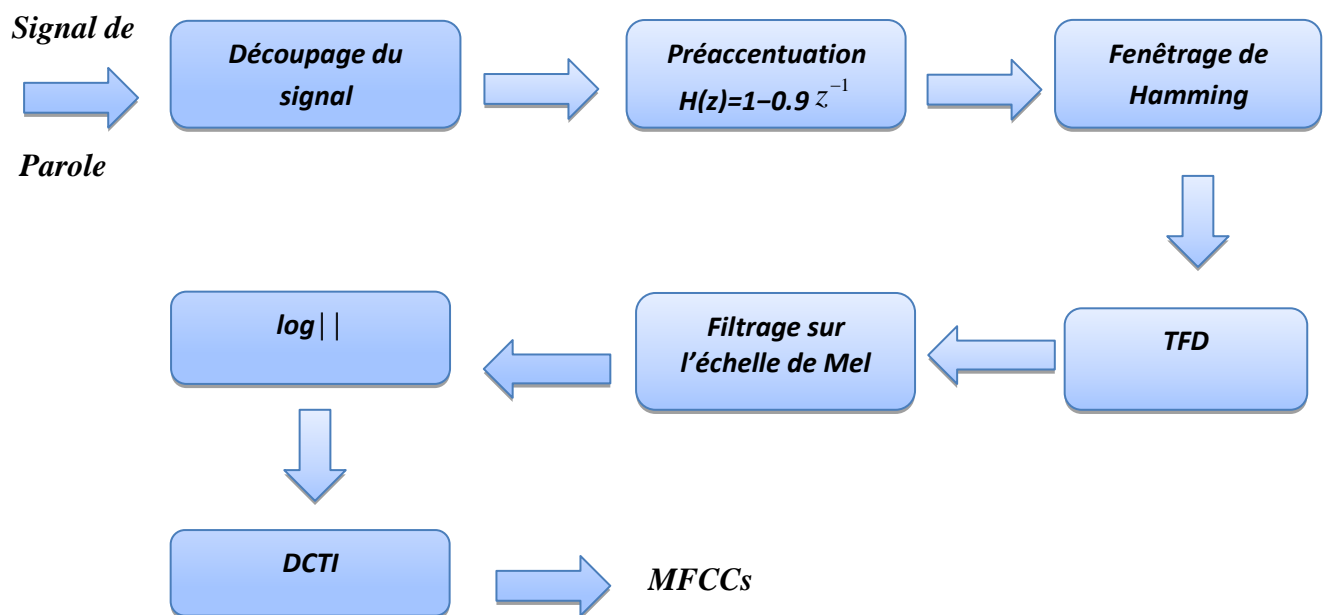


Figure 3.2 : Organigramme de l'extraction des coefficients MFCCs

La procédure de calcul pas à pas des MFCCs est la suivante :

- Découpage du signal en trames chevauchées pour assurer la stationnarité et éviter les transitions brusques de trame en trame,
- Préaccentuation qui permet de renforcer la contribution des hautes fréquences avec un filtre passe-haut de la forme $(1-0.9z^{-1})$,
- Fenêtrage de Hamming qui permet d'assurer la continuité aux bords,
- Calcul de la TFD sur chaque trame,
- Filtrage par un banc de filtres triangulaires répartis le long de l'échelle de Mel,
- Calcul du logarithme du module de l'énergie en sortie du banc de filtres,
- Application de la Transformée en Cosinus Discrète Inverse (DCTI) qui joue le rôle d'une TFD inverse [3],

3.2.2 Paramètres de la source d'excitation

a) Pitch

La fréquence fondamentale F_0 ou pitch : est une mesurée en Hertz (Hz), elle détermine la hauteur de la voix. Elle dépend du nombre de vibrations par seconde des cordes vocales [5]

$$F_0 = \frac{1}{T} \quad (3.2)$$

T : La période fondamentale.

b) Jitter

C'est la mesure la plus connue en termes d'évaluation de l'instabilité à court terme de la fréquence fondamentale F_0 . On entend dire par "court terme" le passage d'un cycle au cycle suivant. Le jitter caractérise la dispersion temporelle du signal vocal. En effet, quel que soit la voix, normale ou pathologique, la F_0 est caractérisée par des variations. Dans le cas normal ces variations sont légères et reste toujours autour du pitch. Dans le cas pathologique, ces variations ne sont plus cernables et c'est pour cela que pour les voix pathologiques le pitch ne varie pas dans un seul sens, c'est-à-dire qu'il peut augmenter et diminuer également [5].

Il existe différentes mesures de ce paramètre. On peut citer :

- **Jitter absolu**

La différence absolue moyenne entre deux périodes consécutives peut être exprimée comme suit [16]:

$$jitter(abs) = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}| \quad (3.3)$$

Avec T_i : c'est la période du i^{eme} l'intervalle

N : C'est le nombre d'intervalle

i : l'intervalle de la période calculée

- **Jitter (%)**

Le Jitter est exprimée par la différence absolue moyenne entre deux périodes consécutives divisée par la durée moyenne [16]:

$$jitter(\%) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (3.4)$$

- **Jitter RAP (relative average perturbation)**

Il est défini comme la différence absolue moyenne entre une période et la moyenne de celui-ci et ses deux voisins divisée par la période moyenne [16] :

$$jitter(rap) = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} |T_i - (T_{i-1} + T_i + T_{i+1}) / 3|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (3.5)$$

- **Jitter PPQ (five point Periods perturbation quotient)**

Il est défini comme la différence absolue moyenne entre une période et la moyenne de celui-ci et ses quatre voisins les plus proches divisée par la durée moyenne [26]:

$$jitter(ppq) = \frac{\frac{1}{N-2} \sum_{i=3}^{N-2} |T_i - (T_{i-2} + T_{i-1} + T_i + T_{i+1} + T_{i+2}) / 5|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (3.6)$$

- **Jitter DDP (difference of difference of periods)**

Le Jitter DDP est défini comme la différence absolue moyenne, entre les différences de deux périodes consécutives, le tout est divisée par la durée moyenne de ces périodes [26] :

$$jitter(ddp) = \frac{1}{N-2} \sum_{i=2}^{N-1} (T_{i-1} - T_i) - (T_i - T_{i+1}) \quad (3.7)$$

c) Shimmer

C'est la mesure équivalente au Jitter mais en termes d'amplitude. Le shimmer reflète le taux d'instabilité à court terme de la variation d'amplitude [3].

Plusieurs mesures de ce paramètre peuvent être calculées. Nous citons :

- **Shimmer relative**

La différence absolue moyenne entre les amplitudes des périodes consécutives divisée par l'amplitude moyenne [10] :

$$shimmer(relative) = \frac{\frac{1}{N} \sum_{i=1}^{N-1} |A_i - A_{i-1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (3.8)$$

A_i : C'est amplitude du i^{eme} l'intervalle

- **Shimmer (dB)**

C'est le logarithme absolu en base 10 de la différence entre les amplitudes des périodes consécutives multiplié par 20 [10] :

$$shimmer(dB) = \sum_{i=1}^{N-1} 20 * \log(A_{i+1} - A_i) \quad (3.9)$$

- **Shimmer APQ3 (3 points amplitudes perturbation quotient)**

Ce paramètre est défini comme l'écart moyen entre l'amplitude d'une période et la moyenne des amplitudes de ses voisins divisé par l'amplitude moyenne [10] :

$$shimmer(apq3) = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} |A_i - (A_{i-1} + A_i + A_{i+1}) / 3|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (3.10)$$

- **Shimmer APQ5 (5 points amplitudes perturbation quotient)**

Ce paramètre est défini comme l'écart moyen entre l'amplitude d'une période et la moyenne des amplitudes de celui-ci et ses 4 voisins les plus proches divisé par l'amplitude moyenne [10] :

$$shimmer(apq5) = \frac{\frac{1}{N-4} \sum_{i=4}^{N-1} |A_i - (A_{i-2} + A_{i-1} + A_i + A_{i+1} + A_{i+2}) / 5|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (3.11)$$

- **Shimmer APQ1 (11 points amplitudes perturbation quotient)**

Il est défini comme l'écart moyen entre l'amplitude d'une période et la moyenne des amplitudes de celui-ci et ses 10 voisins les plus proches divisé par l'amplitude moyenne [10]:

$$shimmer(apq11) = \frac{\frac{1}{N-10} \sum_{i=6}^{N-5} |A_i - (A_{i-5} + \dots + A_{i-1} + A_i + A_{i+1} + \dots + A_{i+5}) / 11|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (3.12)$$

- **Shimmer DDA (difference of difference of amplitudes)**

Il est défini comme la différence absolue moyenne entre les différences des amplitudes des périodes consécutives exprime par [26] :

$$shimmer(dda) = \frac{1}{N-2} \sum_{i=2}^{N-1} |(A_{i+1} - A_i) - (A_i - A_{i-1})| \quad (3.13)$$

3.2.3 Paramètres liés au bruit

a) HNR (Harmonic to Noise Ratio)

Le HNR est une mesure explorant la présence du bruit au cours de la phonation. Il peut être calculé selon plusieurs méthodes. Quel que soit la source de provenance du bruit dans le signal de la parole, le HNR permet de l'évaluer. En effet, plus le HNR est élevé et plus la voix est jugée comme étant une voix normale [3].

$$HNR = \frac{H}{N} \quad (3.14)$$

En attribuant H : à l'énergie de la forme d'onde moyenne

N : à l'énergie de bruit

Ainsi, en considérant tous les cycles inclus dans la formation du signal, on obtient la valeur finale du HNR en la convertissant en dB.

b) NHR (Noise to Harmonic Ratio)

NHR est un rapport moyen de l'énergie des composants inharmoniques. Il représente une évaluation générale de la présence de bruit dans le signal analysé [10].

$$NHR = \frac{N}{H} \quad (3.15)$$

Dans le cas de NHR, lorsque ce dernier tend vers zéro la voix est considérée normale.

En attribuant H : à l'énergie de la forme d'onde moyenne,
 N : à l'énergie de bruit.

3.2.4 Méthodes d'estimation des paramètres acoustiques

a) Méthode d'auto corrélation pour l'estimation du Pitch

L'auto-corrélation est une méthode d'analyse des signaux qui calcule la corrélation entre un signal et une copie retardée du même signal à des retards correspondant au minimum et le maximum prévu des périodes fondamentales [15].

$$r_x(\tau) = \int x(t) * x(t+\tau) dt \quad (3.16)$$

$x(t)$: Signal temporel stationnaire,

$r_x(\tau)$: L'auto-corrélation en fonction de τ .

L'idée de base derrière la technique est que la corrélation entre le signal et la copie retardée atteindra un maximum lorsque la copie est compensée par exactement une période fondamentale.

Le retard au cours de laquelle se produit le pic sert d'estimation de la période fondamentale et l'importance du pic de la fonction d'autocorrélation peut servir en tant que mesure du degré de périodicité du signal.

Les signaux périodiques élevés devraient montrer des pics plus importants que signaux légèrement périodiques. Davis [26] a développé une mesure appelée pitch Amplitude (PA) qui a été basée sur l'analyse d'autocorrélation du signal résidu LPC. Des différences significatives dans les mesures PA ont été trouvées entre les locuteurs normaux et pathologiques, bien que le chevauchement entre les groupes fût suffisamment grand qu'un outil de dépistage basé sur cette mesure ne semble pas réalisable.

b) HNR

La méthode de Boersma [23] a pour but de quantifier la présence de bruit dans la parole. Cette méthode utilise la fonction d'auto-corrélation à court terme de la parole pour déterminer la hauteur, le HNR. L'autocorrélation d'un signal est définie précédemment (équation 3.20) [4].

La période fondamentale T_0 est définie comme étant la valeur de τ qui correspond au maximum le plus élevé de la fonction d'autocorrélation à court terme (expliquer précédemment).

L'énergie du signal de parole fenêtré est la valeur de la fonction d'auto-corrélation à court terme à son indice zéro:

$$r_x(0) = r_p(0) + r_{ap}(0) \quad (3.17)$$

r_p Et r_{ap} étant les énergies respectives des composantes périodiques et apériodiques.

L'auto-corrélation normalisée est définie comme suit :

$$r'_x(\tau) = \frac{r_x(\tau)}{r_x(0)} \quad (3.18)$$

Elle donne la périodicité de la fonction d'autocorrélation de composante périodique et en supposant un bruit blanc additif (non corrélé avec lui-même), l'énergie de la composante périodique est donnée par :

$$r'_p(0) = r'_p(T_0) = r'_x(T_0) \quad (3.19)$$

Puis l'énergie de la composante apériodique :

$$r'_{ap}(0) = 1 - r'_p(T_0) = 1 - r'_x(T_0) \quad (3.20)$$

HNR est défini comme suit :

$$HNR = \frac{r'_p(0)}{r'_{ap}(0)} \quad (3.21)$$

Bien que cet algorithme soit basé sur une composante apériodique de bruit blanc additif, il a été montré qu'il est correctement sensible également au Jitter [4].

3.3 Machines à vecteurs de supports

Les machines à vecteurs de support ou séparateurs à vaste marge (en anglais *Support Vector Machine, SVM*) sont un ensemble de techniques d'apprentissage supervisé destinées à

résoudre des problèmes de discrimination et de régression. Les SVM sont une généralisation des classifieurs linéaires.

Les SVM ont été développés dans les années 1990 à partir des considérations théoriques de Vladimir Vapnik sur le développement d'une théorie statistique de l'apprentissage : la théorie de Vapnik-Chervonenkis. Les SVM ont rapidement été adoptés pour leur capacité à travailler avec des données de grandes dimensions, le faible nombre d'hyper paramètres, leurs garanties théoriques, et leurs bons résultats en pratique [13].

Les SVM ont été appliqués à de très nombreux domaines (bio-informatique, recherche d'information, vision par ordinateur, finance.....). Selon les données, la performance des machines à vecteurs de support est de même ordre, ou même supérieure, à celle d'un réseau de neurones ou d'un modèle de mélanges gaussiens [12].

Le principe fondateur des SVM est justement d'intégrer à l'estimation le contrôle de la complexité, c'est-à-dire le nombre de paramètres qui est associé dans ce cas au nombre de vecteurs supports [12].

• Principe de fonctionnement [14]

Le principe des SVMs est de trouver l'hyperplan optimal parmi l'ensemble des hyperplans possibles (Figure 2.1) permettant de classier correctement les données et maximiser sa distance entre les vecteurs supports les plus proches

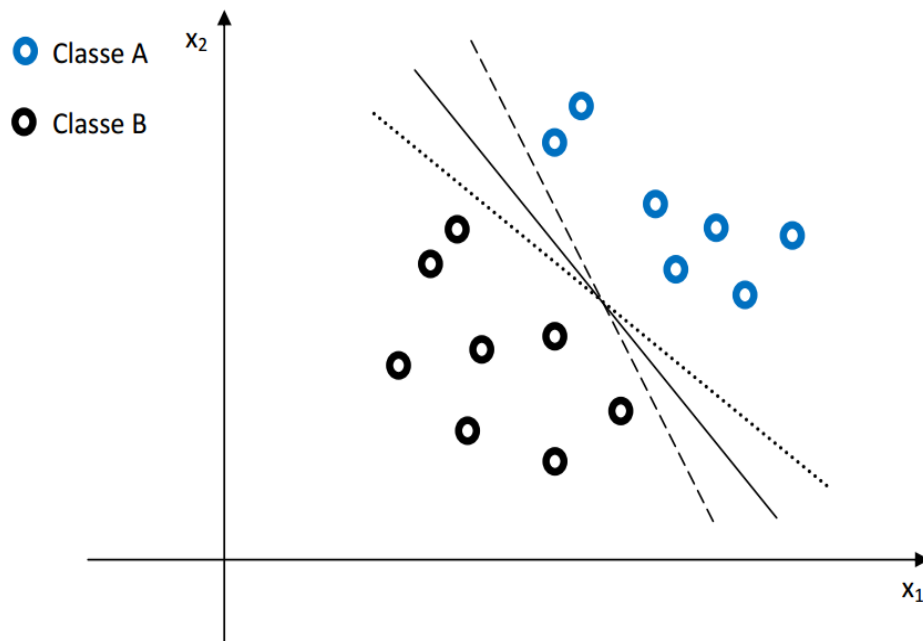


Figure 3.3: Recherche de l'hyper plan optimal

- **Hyperplan**

On appelle hyperplan séparateur un hyperplan qui sépare les points d'apprentissage de deux classes données (A et B). Les deux hyperplans (H1) et (H2) sont appelés hyperplans canoniques.

- **Vecteurs supports**

La détermination de l'hyperplan est basée seulement sur les points situés sur la frontière entre les deux classes de données (parmi l'ensemble total d'apprentissage). Ces points sont appelés vecteurs supports (Figure 3.3)

- **Marge**

Il existe une infinité d'hyperplans capable de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (i.e. la distance entre l'hyperplan et les vecteurs supports), cette distance est appelée la marge.

Le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple :

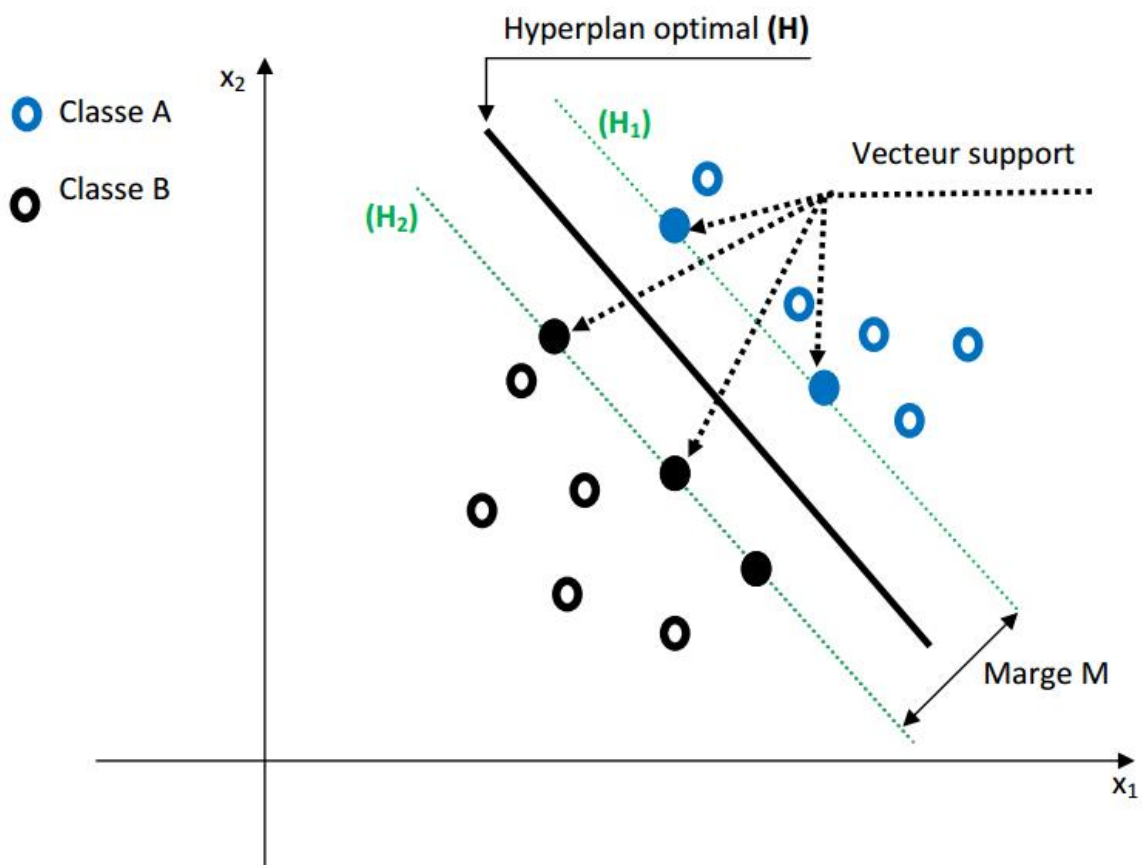


Figure 3.4: L'hyperplan H optimal, vecteurs supports et marge maximale. x_1 et x_2 représentent les attributs des classes A et B

- **Objectif de maximisation de la marge**

De manière intuitive, le fait d'avoir une marge plus large procure plus de sécurité lorsqu'un nouveau exemple est classé. De plus, si un classificateur trouvé qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples.

La bonne classification correspond à la plus large marge car si on a une petite marge l'exemple sera mal classé mais si on a une large marge l'exemple sera bien classé.

- **Linéarité et non Linéarité**

Parmi les modèles des SVMs, on constate les cas linéairement séparables et les cas non linéairement séparables (Figure 2.3). Les premiers sont les plus simples, car ils permettent de trouver facilement le classificateur linéaire. Dans la plupart des problèmes réels, il n'y a pas de séparation linéaire possible entre les données, le classificateur de marge maximale ne peut pas être utilisé, car il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables.

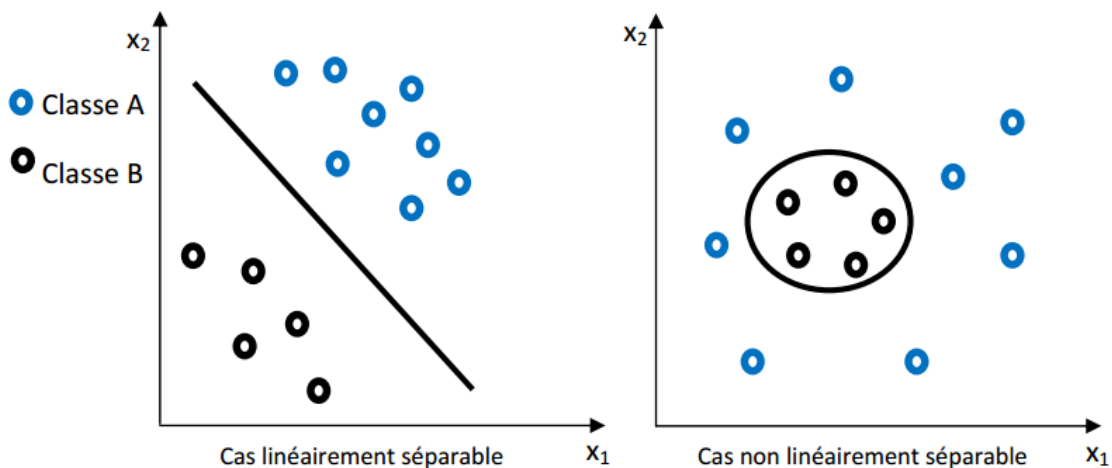


Figure 3.5: Séparation linéaire et non linéaire

- ❖ **Cas linéairement séparable**

Un classificateur est dit linéaire lorsqu'il est possible d'exprimer sa fonction de décision par une fonction linéaire en x (Figure 2.3). Dans la suite, nous supposons que les exemples sont donnés dans un format vectoriel. Notre espace d'entrée x correspond donc à \mathbb{R}^n ou n est le nombre de composantes des vecteurs contenant les données.

$$h(x) = \langle w, x \rangle + b = \sum_{i=1}^n w_i x_i + b \quad (3.22)$$

Si les données sont linéairement séparables, alors il existe un hyperplan d'équation :

$$\langle w.x \rangle + b = 0 \quad \text{Et tel que} \quad (3.24)$$

$$\langle w.x \rangle + b \geq 1 \quad \text{Si } y_i = +1 \quad (3.25)$$

$$\langle w.x \rangle + b \leq -1 \quad \text{Si } y_i = -1 \quad (3.26)$$

On peut combiner ces deux inéquations en une seule :

$$y_i (w.x + b) \geq 1 \quad \text{Pour } i = 1, \dots, n \quad (3.27)$$

Ou x_i : Vecteur des attributs $(x_1, x_2, x_3, \dots, x_n)$.

w_i : Vecteur des poids $(w_1, w_2, w_3, \dots, w_n)$.

b : Seuil du séparateur linéaire.

y_i : Label de classe pouvant prendre la valeur +1 ou -1.

h : Fonction de décision.

i : La dimension des vecteurs d'entrée.

n : La taille de l'ensemble d'apprentissage.

❖ **La distance de l'hyperplan de séparation**

$$d(x) = \frac{|\langle w.x \rangle + b|}{\|w\|} \geq \frac{1}{\|w\|} \quad (3.28)$$

❖ **La distance de l'hyperplan à l'origine**

$$\frac{b}{\|w\|} \quad (3.29)$$

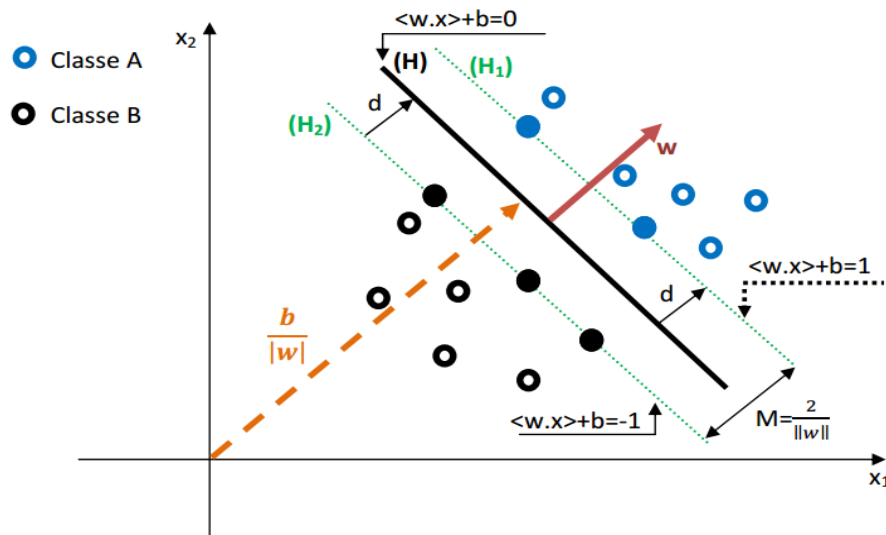


Figure 3.6: Exemple graphique des données linéairement séparables

La marge géométrique représente la distance euclidienne prise perpendiculairement entre l'hyperplan et l'exemple x_i

En prenant un point quelconque x_p se trouvant sur l'hyperplan, la marge géométrique peut s'exprimer par :

$$\frac{w}{\|w\|} \cdot (x_i - x_p) \quad (3.30)$$

La marge est au moins égale à la distance entre les deux hyperplans $H1$ et $H2$ soit $\frac{2}{\|w\|}$ ou $\|w\|$ fait référence à la norme du vecteur w . Maximiser cette marge revient donc à minimiser $\|w\|$

- **Cas linéairement non séparable**

Pour surmonter les inconvénients des cas non linéairement séparables, l'idée des SVM est de changer l'espace des données. La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouveau espace : il est nécessaire de faire un changement de dimension .cette nouvelle dimension est appelle « espace de ré-description »

En effet, intuitivement, plus la dimension de l'espace de ré-description est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée, ceci est illustre par la figure suivante :

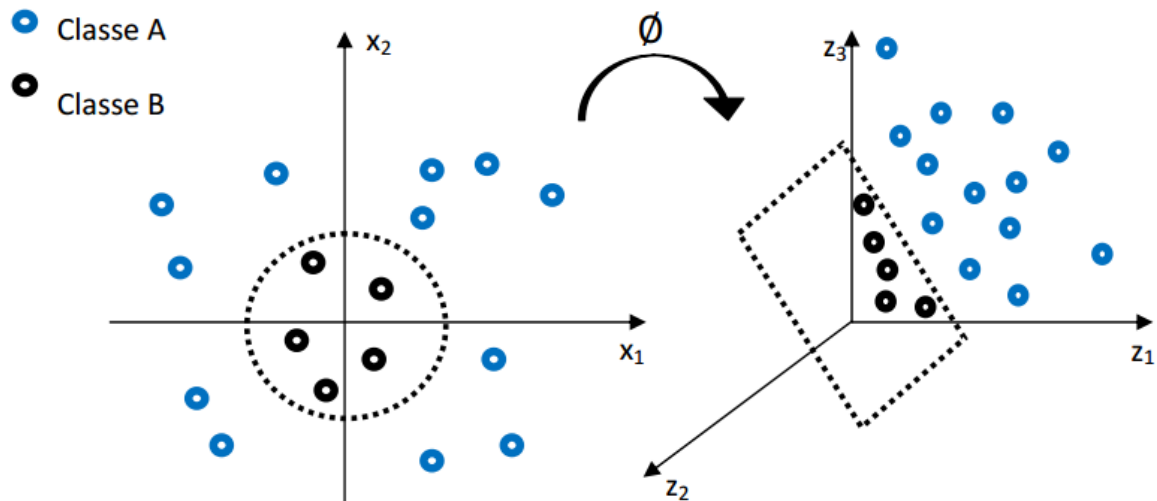


Figure 3.7 : Transformation non linéaire de données.

Il y a donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de ré-description de plus grande dimension. Cette transformation non linéaire est réalisée via une fonction noyau [2].

Le noyau linéaire : $K(x, x) = \langle x, x \rangle$

Le noyau polynomial : $K(x, x) = (a^* \langle x, x \rangle + b)^d$

Le noyau Gaussien : $K(x, x) = \exp\left(-\frac{(x, x')^2}{2\sigma^2}\right)$

❖ **Avantage :**

- Absence d'optimum local,
- Contrôle explicite du compromis entre la complexité du classifieur et l'erreur,
- Possibilité d'utilisation de structure de données comme les chaînes de caractères et arbres comme des entrées,
- Traitement des données à grandes dimensions [38].

❖ **Inconvénients :**

- Demande des données négatives & positives en même temps,
- Besoin d'une bonne fonction Kernel,
- Problèmes de stabilité des calculs dans la résolution de certains programme quadratique a contraintes [38].

3.4 Méthode de détection du Cancer de larynx

La méthode proposée est représentée par l'organigramme suivant :

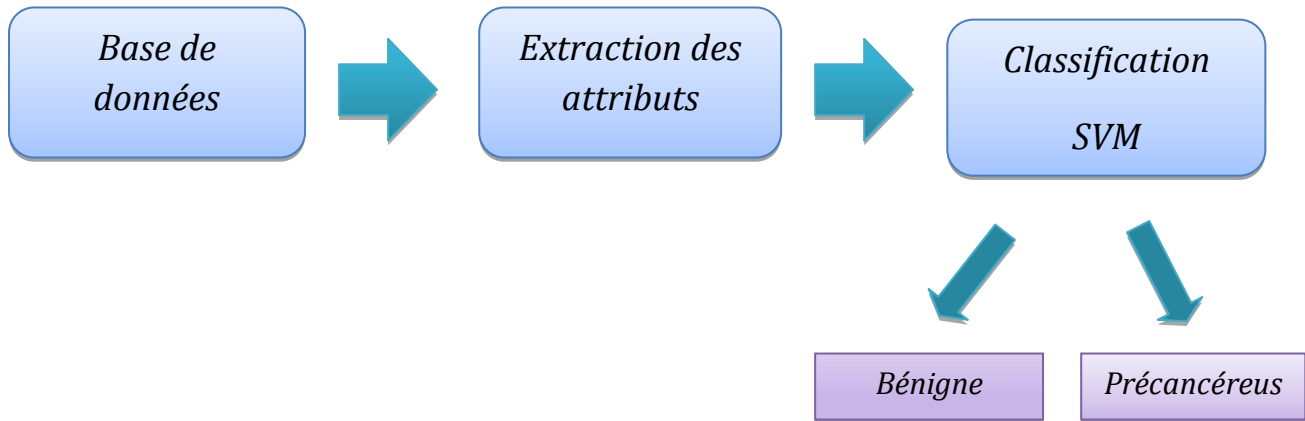


Figure 3.8 : Organigramme de système d'aide au diagnostique

Afin de différencier entre les pathologies bénignes et précancéreuses nous avons proposé cette méthode, les étapes de cette méthode sont les suivants :

3.4.1 Données acoustiques

Notre système se base sur les machines d'apprentissage supervisées nécessitant l'utilisation des bases de données labélisées pour extraire les attributs pertinents et pour estimer les modèles de classification. Dans notre cas, nous avons utilisés la base de données SVD et plus précisément les sons précancéreux et bénignes.

Plus de détail sur cette base sera reporté dans le prochain chapitre.

3.4.2 Extraction d'attributs

Après la collection de données, nous avons fait une analyse acoustique de ces données pour extraire les attributs, l'organigramme suivant explique les étapes suivis :

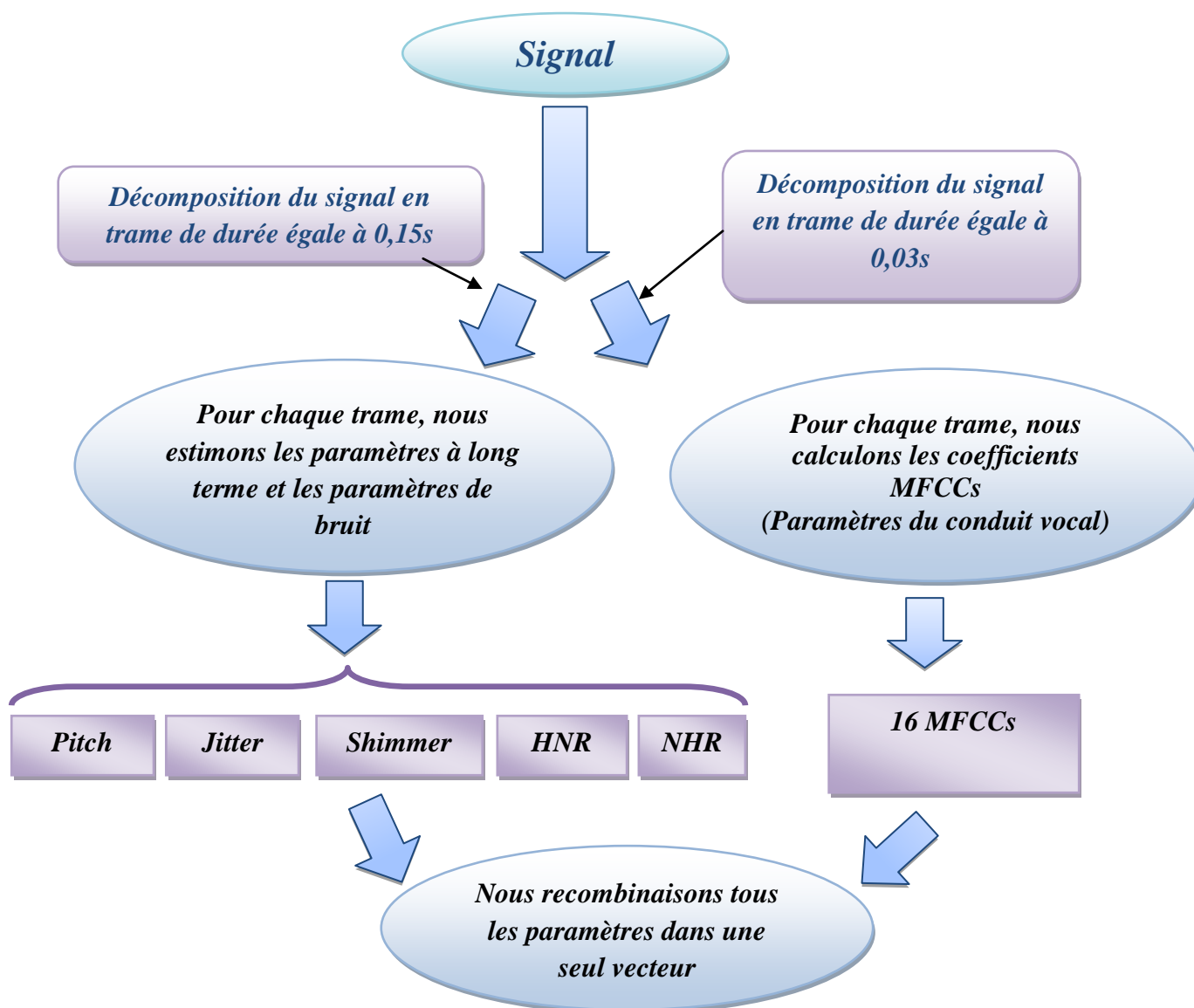


Figure3.9 : Organigramme d'extraction d'attributs

1^{er} étape : décomposition de signal

Nous décomposons le signal en trames de 0.15 s sans décalage pour estimer les paramètres à long terme (jitter, shimmer, HNR, NHR et le pitch)

Puis nous faisons une autre décomposition du signal décomposé en trame de 0.03 s sans décalage pour calculer les coefficients MFCCs (paramètre à court terme)

2^{ème} étape : calcule des paramètres

Après la décomposition de signal .nous avons estimé les paramètres à long terme (pitch, shimmer et jitter), les paramètres de bruit (HNR, NHR) et les paramètres à court terme (les coefficients de MFCCs)

3^{ème} étape : recombinaison des paramètres

Nous avons recombinaisé ces paramétrés et les mettre dans une seule matrice qui sera la matrice d'entrée pour notre classifieur

3.4.3 Classification

Pour classifier les sons en deux classes :

- Classe (1) présente des patients qui sont atteints de la leucoplasie (pathologie précancéreuse),
- Classe (0) présente des patients qui sont atteints des pathologies bénignes (liste des pathologies : Chapitre 2, section 2.4.2)

Nous avons utilisé les SVMs

Les étapes suivies pour la classification sont :

- Construction de la matrice d'attributs à partir de l'analyse acoustique des sons,
- Pour l'optimisation des paramètres C, γ . Nous avons utilisé leurs valeurs par défaut ($C=1, \gamma=0$),
- Construction du model de prédiction.

3.5 Métriques d'évaluation

3.5.1 Matrice de confusion

L'évaluation de la classification est basée sur un tableau à deux dimensions, appelé matrice de confusion. Cette matrice trie tous les cas du modèle en catégories, en déterminant si la valeur prédite correspondait à la valeur réelle. Tous les cas dans chaque catégorie sont ensuite comptés et les totaux sont affichés dans la matrice. Les mesures que nous allons évoquer dans cette section utilisent la matrice de confusion [13].

Tableau 3.1 : matrice de confusion

	Décision Positive	Décision négative	
Etiquette positive	Vrais Positifs (VP)	Faux Négatifs (FN)	Pos
Etiquette négative	Faux Positifs(FP)	Vrais Négatifs (VN)	Neg
Total (T)	Ppos	Pneg	N

Etiquette positive : Patients ont une maladie précancéreuse,

Etiquette négative : Patients ont une maladie bénigne,

Décision Positive : Test était positif,

Décision Négative : Test négatif,

Vrais Positifs (VP) : individus ont une maladie précancéreuse réagissent positivement au test,

Vrais Négatifs (VN) : individus ont une maladie bénigne réagissent négativement au test.

Faux Positifs (FP) : individus ont une maladie bénigne réagissent positivement au test.

Faux Négatifs (FN) : individus ont une maladie précancéreuse réagissent négativement au test

Dans ce contexte, plusieurs mesures sont utilisées pour l'évaluation des performances de classification avec différentes combinaisons de méthodes de réduction de la fonctionnalité et des classificateurs [8].

Taux de Vrais Positifs (TVP) :

$$TVP = \frac{VP}{Pos} = \frac{VP}{VP + FN} \quad (3.31)$$

Taux de vrais Négatifs (TVN) :

$$TVN = \frac{VN}{Neg} = \frac{VN}{VN + FN} \quad (3.32)$$

Taux de Faux Positifs(TFP) :

$$TFP = \frac{FP}{Neg} = \frac{FP}{FP + VN} \quad (3.33)$$

Taux de Faux Négatifs (TFN) :

$$TFN = \frac{FN}{Pos} = \frac{FN}{FN + VP} \quad (3.34)$$

Taux de bonne classification ou l'exactitude (TBC) :

$$TBC = \frac{VP + VN}{VP + FN + VN + FP} \quad (3.35)$$

3.5.2 Courbe ROC

La courbe ROC (Receiver Operating Characteristic) est une méthode de représentation graphique des performances d'un classifieur à deux classes. Elle peut fournir une mesure des performances plus riche que l'exactitude (accuracy) ou le taux d'erreur. Depuis plusieurs années, son utilisation est devenue incontournable dans les méthodes d'évaluations.

Elle est généralement utilisée dans la prise de décision médicale où les problèmes de diagnostic à deux classes sont courants [13].

La courbe ROC représente la sensibilité en ordonnée en fonction, de la quantité (Spécificité) en abscisse. La sensibilité qui est en effet le taux des vrais positifs et la spécificité est le taux de vrais négatifs.

3.5.3 Air sous la courbe (AUC)

L'AUC (Area Under curve) est un outil populaire dans la prise de décision médicale. Elle est la mesure de l'aire de la surface située sous le tracé d'une fonction mathématique

dessinée dans un repère [13]. Formellement, cette valeur correspond à l'intégral de cette fonction. Dans le meilleur des cas $AUC = 1$.

$$AUC = \sum_i S_i \quad (3.36)$$

Si : est la surface des trapèzes.

3.5.4 Validation croisée

La validation croisée désigne le processus qui permet de tester la précision prédictive d'un modèle [32]. Il y a au moins trois techniques de validation croisée l'une entre eux est la méthode « kfold ».

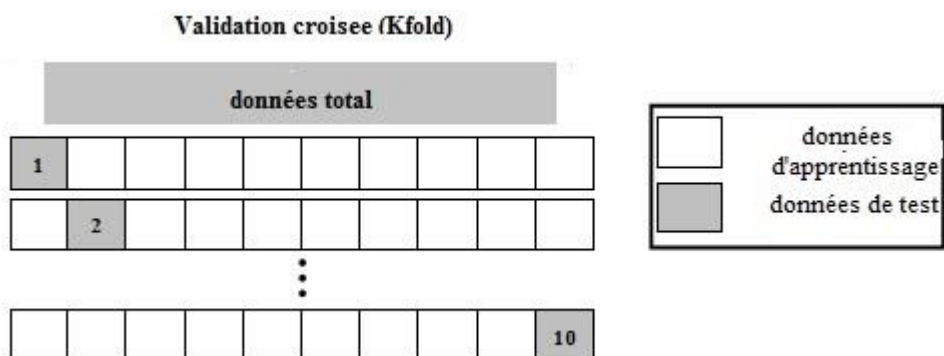


Figure 3.10 : Principe de la validation croisée [12].

La validation croisée est utilisée dans la résolution d'un grand nombre de problèmes : la sélection de modèle, l'additivité et l'identification [17].

3.6 Conclusion

Dans ce chapitre, nous avons étudié les indices acoustiques du Cancer de larynx suivi par une étude sur les *Supports Vector Machines* (SVMs). Par la suite, nous avons présenté une méthode d'aide au diagnostic du Cancer du larynx qui permet de différencier entre les pathologies précancéreuses et les pathologies bénignes.

La dernière partie a été consacrée aux méthodes d'évaluation des classifieurs. Le prochain chapitre sera consacré aux résultats expérimentaux.

Chapitre 4 : Expérimentations, résultats et interprétations

4.1 Introduction

Dans ce chapitre, nous exposerons les résultats obtenus par l'application de la méthode proposée en vue la détection précoce du Cancer de larynx. La base de données qui a été utilisée est la SVD.

La première partie de notre travail consiste à faire une analyse acoustique des sons du corpus sonore pour extraire les attributs. La deuxième partie du travail utilise les vecteurs d'attributs pour construire le model de prédiction à base des SVMs.

Nous évaluerons la méthode proposée par le calcul de : TBC, TVP, TVN la courbe ROC et l'AUC. Enfin, Nous interprétons par les résultats.

4.2 Logiciels de développement

4.2.1 Environnement MATLAB

Le logiciel MATLAB est un langage de développement informatique particulièrement dédié aux applications scientifiques. Il est utilisé pour développer des solutions qui nécessitent une très grande charge de calcul. C'est la version 2013 qui a été utilisée (Matlab R2013a).

Le choix de MATLAB a été fait en raison de sa rapidité, sa richesse en librairie, et la facilité de son environnement de programmation. Les Toolboxes qui ont été utilisées sont :

- ❖ **Voice box** « boîte vocale » : c'est une boîte à outils pour le traitement de la parole, le codage, la synthèse et la reconnaissance [28],
- ❖ **Libsvm** « Library for SVM » : c'est une bibliothèque pour les machines à vecteurs de Support [29].

4.2.2 PRAAT

La méthode que nous avons utilisée pour l'étape d'extraction d'attributs est basée sur le logiciel PRAAT. Praat est un logiciel libre pour l'analyse, la manipulation et l'annotation de sons. Ces fonctionnalités en font un outil complet en particulier pour l'étude de parole. Il permet également de tracer des graphiques. Il a été créé par Paul Boersma [30] et David Weenink [31, 32] au département de l'anglais Phonétique de l'Université d'Amsterdam en 1992.

Il existe différents paramètres acoustiques dans PRAAT. Parmi eux les paramètres que nous avons utilisés, à savoir : Pitch, Jitter avec ses types (Jitter local, Jitter (absolue), Jitter (rap), Jitter(ppq5), Jitter(ddp)), Shimmer avec ses types (Shimmer (local), Shimmer (dB), Shimmer (apq3), Shimmer (apq5), Shimmer (dda)), NHR, et HNR.

Il fournit un environnement informatique ergonomique et facilite des analyses phonétiques d'une grande précision portant sur les divers phénomènes phonétiques.

4.2.3 WEKA (Waikato Environment for Knowledge Analysis)

Weka (Environnement Waikato pour l'analyse de connaissances) est une suite de logiciels d'apprentissage automatique et d'exploration de données écrite en Java. Il est développé à l'université de Waikato en Nouvelle-Zélande [38]. Le format d'entrée par défaut de Weka est l'ARFF (Attribute Relation File Format). D'autres formats peuvent être importés : CSV, binaire, à partir d'une URL, etc.

WEKA met en œuvre des algorithmes de prétraitement des données, la classification, la régression, et les règles d'association. Il comprend également un outil de visualisation [39].

4.3 Résultats expérimentaux

4.3.1 Corpus sonore

La base de données SVD contient plus de 70 pathologies classées en organiques et non organiques. Notre but dans ce projet est de classer ces maladies en 2 catégories : bénignes et précancéreuses. Les pathologies malignes n'ont pas été considérées dans notre application. La prise en charge de ce type de pathologie doit être pris en considération dans le future (continuité de notre travail).

La classe précancéreuse est représentée par la leucoplasie (le début du Cancer de larynx). La classe bénigne est représentée par un ensemble de 7 pathologies disponibles dans la base de données SVD de la voyelle /a/. Ces pathologies sont :

- Paralysie des cordes vocales (vocal fold paralysis),
- Syndrome de parkinson (Parkinson's disease),
- Laryngites (laryngitis),
- Reinke œdème (Reinke oedema),
- Nodules des cordes vocales (Vocal fold nodules),
- Chondrome (Chondroma),
- Laryngocèle (Laryngocele).

Vu le manque des sons dans la base de données SVD, nous avons arrivé à exploiter 5 pathologies seulement de la voyelle /a/ :

- Paralysie des cordes vocales
- Laryngites
- Reinke œdème
- Nodules des cordes vocales
- Laryngocèle

Nous avons partagé le totale des sons en plusieurs groupes. Le groupe 1 (G1) contient les sons précancéreux. Les autres groupes (de G2 à G5) contiennent une ou plusieurs pathologies bénignes (Tableau 4.A).

Tableau 4.1 : Groupes des pathologies utilisées

Les groupes	G1	G2	G3	G4	G5
Les maladies	Leucoplasie Laryngites	Laryngites	Paralysie des cordes vocales	Reinke œdème	-Paralysie des cordes vocales -Laryngites -Reinke œdème -Nodules des cordes vocales -Laryngocèle

Vu que nous n'avons pas eu suffisamment de temps pour tester tous les cas possibles, nous nous sommes limités à une seule expérience :

- Différencier entre le G1 et G2.

Le nombre des sons du G1 (femelle : 12 et male : 24) est égal au nombre de sons du G2. Les statistiques de ce corpus sont citées sur le tableau suivant :

Tableau 4.2 : Corpus de test

		Voix pathologique	Moyenne d'âge	Ecart type
Laryngitis	Hommes	24	53,67	16,45
	Femmes	12	54,83	13,52
Leucoplasie	Hommes	24	52,12	10,438
	Femmes	12	52	15,064

4.3.2 Extraction d'attributs

La première partie d'extraction d'attributs consiste à utiliser le logiciel PRAAT pour calculer les traits acoustiques. Chaque signal est divisé en trames ; la durée de ces trames est égale à 0.15s afin d'estimer les paramètres de la source d'excitation à savoir : pitch, Jitter (Jitter (local), Jitter (absolue), Jitter (rap), Jitter (ppq5), Jitter (ddp)), Shimmer (Shimmer (local), Shimmer (dB), Shimmer (apq3), Shimmer (apq5), Shimmer (dda)) et les paramètres de bruit :HNR, NHR qui sont en total 13 paramètres.

Par la suite, nous avons divisé chaque trame (0.15 s de durée) en 0.03s pour calculer les MFCCs (les paramètres de conduit vocal) par le biais du logiciel MATLAB. Nous avons obtenu 16 coefficients. Enfin Les paramètres à long terme et à court terme sont recombinaés dans une matrice globale pour obtenir la matrice d'attributs.

pitch	Jitter Local	Jitter (abs)	Jitter (rap)	Jitter (ppq5)	Jitter (ddp)	Shimmer r Local	Shimmer (abs)	Shimmer (apq3)	Shimmer (apq5)	Shimmer (apq11)	HNR	NHR	MFCC [1-16]
99.1	0.77	5.07 ^E -05	0.7	0.34	0.55	7.98	0.65	8.98	13.6	14.6	12.6	0.87	0.232
99.9	0.54	9.8 ^E -05	0.8	0.98	0.74	9.87	0.76	6.87	11.3	15.6	19.4	0.9	0.234
109.4	0.87	8.9 ^E -05	0.13	0.24	0.45	5.98	1.76	4.98	15.3	17.4	19.9	0.63	0.5432
115.4	0.55	9.6 ^E -05	0.65	0.13	1.04	3.97	1.56	3.84	9.63	18.5	17.9	0.87	0.2354
124.6	0.43	6.9 ^E -05	0.75	0.24	1.5	1.07	0.43	3.74	10.5	13.5	15.4	0.45	0.2986
134.4	0.52	8.9 ^E -05	0.86	0.45	2.87	6.98	0.76	5.87	11.53	17.4	18.4	0.22	0.7543
97.4	0.67	7.9 ^E -05	0.87	0.34	0.97	3.98	0.34	9.65	8.53	18.4	19.5	0.97	0.8643
103.6	0.98	9.8 ^E -05	0.98	0.76	0.45	5.78	1.56	9.63	9.53	13.5	18.5	0.45	0.9863
124.6	0.75	6.9 ^E -05	0.54	0.87	0.95	6.98	1.43	9.53	6.62	16.4	17.4	0.96	0.2465
100.9	0.64	8.9 ^E -05	0.98	0.54	1.05	7.45	0.87	5.87	9.52	18.4	16.9	0.56	0.3265
134.9	0.95	9.7 ^E -05	0.45	0.74	0.75	9.65	0.34	3.76	10.43	19.4	19.6	0.43	0.2753
98.9	0.64	9.7 ^E -05	0.86	0.34	2.05	6.76	0.98	7.63	11.43	11.4	17.4	0.56	0.2863
164.8	0.85	9.6 ^E -05	0.45	0.85	2.87	7.85	1.56	8.53	14.6	13.8	17.9	0.75	0.2583
156.8	0.57	8.9 ^E -05	0.87	0.45	2.97	3.98	1.87	2.43	11.4	17.4	19.5	0.45	0.2484

Tableau 4.3 : Matrice d'attributs

4.3.3 Classification

Pour la classification, nous avons exploité l'outil WEKA. Nous avons comme entrée la matrice d'attributs. Puis, nous avons choisi le clasifieur SVM (après l'installation de LIBSVM). Les paramètres par défaut de Libsvm ont été utilisés : $C=1$, $\gamma=0$

4.4 Evaluation de la méthode

Dans cette section nous utilisons la matrice de confusion, TVP, TVN, courbe ROC et l'AUR pour évaluer la performance de notre système sur le corpus élaboré.

4.4.1 Matrice de confusion

Tableau 4.4 : Matrice de confusion

	Test positif	Test négatif
Leucoplasie(G1)	1367	2
Laryngites(G2)	0	1430

D'après les résultats de tableau, nous avons obtenu une TBC=99.9%.

Le tableau 4.3 qui suit représente les performances de la méthode proposée.

Tableau 4.5 : Performances & résultats

Les performances	Résultats
TBC	99,9
TVP	0,998
TVN	0,998

4.4.2 Courbe ROC

La courbe ROC est représenté sur la figure 4.2. La valeur du TBC mentionnée précédemment montre la performance de notre système.



Figure 4 .2 : Courbe ROC sur la SVD

Nous avons obtenu ce résultat en raison du problème de variabilité rencontré dans l'étape de la validation croisée. En fait, l'utilisation des trames différentes mais qui proviennent du même fichier sonore pour l'apprentissage et le test génère un problème d'estimation biaisée. Vu que le Weka ne permet pas de séparer les données de test et d'apprentissage d'une

manière adéquate, nous n'avons pas pu résoudre ce problème. L'utilisation d'un langage de programmation évolué permettra de résoudre ce problème.

4.5 Conclusion

Nous avons proposé une méthode pour la détection précoce de cancer de larynx par l'analyse acoustique et les machines à vecteurs de support.

La première partie de notre travail consiste à faire une analyse acoustique des sons du corpus sonore afin d'extraire les attributs. La deuxième partie du travail utilise les vecteurs d'attributs pour construire le modèle de prédiction à base des SVMs.

Nous avons évalué la méthode proposée par le calcul de : TBC=99.9, TVP=0.998, TVN=0.998

Chapitre 5 : Conclusions et travaux futures

5.1 Conclusions

L'objectif initial de ce travail est de concevoir un système d'aide au diagnostic pour la détection précoce de Cancer de larynx à partir de la voix. Principalement, il s'agit de différencier entre les pathologies précancéreux « Leucoplasie » et bénignes des patients ayant des troubles vocaliques. La méthode proposée est composée de deux parties :

La première partie : est l'extraction d'attributs en exploitant le logiciel de l'analyse acoustique de la parole PRAAT. Nous avons travaillé avec 29 attributs partagés entre les paramètres de la source d'excitation : pitch, Jitter (Jitter (local), Jitter (absolue), Jitter (rap), Jitter (ppq5), Jitter (ddp)), Shimmer (Shimmer(local), Shimmer (dB), Shimmer (apq3) Shimmer (apq5), Shimmer (dda)), les paramètres de bruit :HNR, NHR et les paramètres de conduit vocal (les 16 MFCCs).

La deuxième partie : concerne la classification des troubles en deux catégories : (1) précancéreuse et (2) bénigne. Nous avons utilisé les SVMs. L'outil de datamining WEKA a été utilisé pour la simulation. Un corpus de sons a été extrait à partir de la SVD.

Un taux de bonne classification de 99,9%. Le taux élevé est obtenu en raison du problème de variabilité rencontré dans l'étape de la validation croisée.

5.2 Perspectives

Les travaux futurs seront focalisés sur :

- (i) La résolution du problème de variabilité,
- (ii) Une étude complémentaire en ce qui concerne (a) les autres groupes de pathologies bénignes mentionnée dans les annexes, (b) Niveau des pathologies malignes afin d'obtenir un système faible qui contient tous les cas possibles, et (c) l'exploitation des autres phonèmes de la base de données (non seulement la voyelle /a/ mais ajouter les autre voyelles disponibles)

Bibliographies

- [1] «societe canadienne du cancer,» [En ligne]. Available: <http://www.cancer.ca/fr-ca/cancer-information/cancer-type/laryngeal/laryngeal-cancer/malignant-tumours/?region=qc..> [Accès le 17 8 2016].
- [2] D. Panek, A. Skalski and J. Gajda, "Quantification of Linear and Non-linear Acoustic Analysis Applied to Voice Pathology Detection," *Advences In Intellegent Systems And Computing*, vol. 284, pp. 355-364, 2014.
- [3] Y. Kheddache, « Carcterisation Des Cris Des Nourissons en Vue du Diagnoqtique Precoce de Differentes Pathologies, » Thèse de Doctorat, Université de Quebec, Quebec, 2014.
- [4] F. Severin, B. Bazhurt and T. Dutoit, "*HNR Extraction in Voiced Speech Oreinted Towerd Voice Quality Analysis*," *Faculte Polytechniques De Mons*, article, 2005, Belgique.
- [5] S. Mrad, « Analyse non Lineaire de la Parole pour la Detection de Voix Pathologiques, » Mmeoire d'Ingenieura, Ecole National d'Ingenieur de Tunis, Tunisie, 2013.
- [6] L. Balthasar and D. Valero, « Transcription Avec Praat, » Lyon, 2005.
- [7] L. Shruthi, S. Kumaraswamy and P. Scholar, "*Noramtive Value for Praat in The Age Rang of 5-10 Years Old Typically Developing Children*," *Language In India*, vol. 15:5, May 2005.
- [8] Y. Oufelle, «Evaluation du Cconcept de Front de ROC et Combainaison de Calssifieur, » Memoire De Master, Université de Rouen, France.
- [9] V. Nicola, M. Fiorella, D. Spinelli and R. Fiorella, "*Acoustic Analysis of Voice Patients Treated by Reconstructie Subtotal Laryngectomy Evalaution and Critical Review*," *Actaotorthiolaryngol Ital*, vol. 26(2), pp. 59-68, 2006.
- [10] S. Raschka, "*Sebastion Raschka*," 11 7 2014. [Online]. Available: http://sebastianraschka.com/Articles/2014_about_feature_scaling.html. [Accessed 17

- 8 2016].
- [11] C. Kawchinporn, «*knowledge discovery in database and data mining*,» 11 7 2014. [En ligne]. Available: http://sebastianraschka.com/Articles/2014_about_feature_scaling.html.. [Accès le 19 8 2016].
- [12] N. Benammar and A. Boutiche, « Identification Des Troubles De La Voix Par l'Analyse Temps-Frequence Et Les Machine A Vecteur De Support, »Memoire De Master,Universite De Boumerdes ,Boumerdes, 2016.
- [13] R. Abdelaziz and Z. Nab, « Systeme d'Aide à la Decision pour le Diagnostic de la Maladies de Parkinson à Partir de la Voix, » Memoire de Master, Université de Blida, 2014.
- [14] J. M. Hillenbrand, "*Acoustic Analysis Of Voice ;a Tutorial*," Kalamazoo, 2009.
- [15] M. Farrus and J. Hernando, "*Using Jitter And Shimmer In Speaker Verification*," *IET signal process*, vol. 3, pp. 247-257, 2009.
- [16] M. Cornec, « Validation Croisee et Modeles Statistiques, » Paris, 4 juin 2009.
- [17] Societe Canadienne Du Cancer, « Comprendre le diagnostique, » Vols. 311-022, 2013.
- [18] M. Putzer and J. Koreman, "*A German Database of Patterns of Pathological Vocal Fold Vibration*," *Phonus* 3, pp. 143-153, 1997.
- [19] D. Martinez, E. Leida, A. Ortega, A. Miguel and J. Villalba, "*Voice Pathology Detection on the Saarbrucken Voice Database Calibration and Fusion of Scores Using Multifocal Toolkit*," *Iberspeech* , vol. 328, pp. 99-109, 2012.
- [20] Q. Lyu, X. Peng, W. Wang, J. Du, L. Li and S. Zhang, "*Clinic Value of the Vocal Acoustics on Voice Evaluation In Pateints With Larynx Leukoplakia*," *Tianjin Medical Journal* , vol. 12/20014.

- [21] Z.-m. Chen, J.-p. Fan , Y.-m. Yang and S. Zhou, "The Voice Assesment of 63 Patients With Larynx Leukoplkia," *Academic Journal Of Second Military Medical University*, no. 413, 8/2006.
- [22] A. Bendahmane, « Cours de traitement de la parole , », Cours, Oran, 2014
- [23] P. Boersma and D. Weenink, "*Praat Doing Phonitis by Computer*," 2003. [Online]. Available: www.praat.org. [Accessed 1 9 2016].
- [24] P. Manfred and J. B. William, "*Saarbrucken Voice Database*," [Online]. Available: http://www.stimmdatenbank.coli.uni-saarland.de/help_en.php4. [Accessed 14 6 2016].
- [25] "Mathworks," [Online]. Available: <https://www.mathworks.com/matlabcentral/linkexchange/links/797-voiceboxspeechprocessing-toolbox-for-matlab>. [Accessed 6 9 2016].
- [26] W. David and M. Youri, "*Objective Disphonia Measures In The Program Praat:Smoothed Cepstral Peak Prominence And Acoustic Voice Quality Index*," *Journal of Voice*, vol. 29, pp. 35-43, 1 2015.
- [27] "*The UVA Phonetic Sciences Home Page*," [Online]. Available: <http://www.fon.hum.uva.nl/paul/>. [Accessed 8 9 2016].
- [28] "*lion voice clinic*," [Online]. Available: <http://www.lionsvoiceclinic.umn.edu/page3b.htm#top>. [Accessed 9 9 2016].
- [29] "*Dell Statistica*," [Online]. Available: <http://www.statsoft.fr/concepts>. [Accessed 10 9 2016].
- [30] "*Copath*," [Online]. Available: http://campus.cerimes.fr/anatomie-pathologique/enseignement/anopath_7/site/html/3.html. [Accessed 10 9 2016].
- [31] "*Conso Globe*," [Online]. Available: <http://www.planetoscope.com/mortalite/31-nombre-de-deces-dus-au-cancer-dans-le-monde.html>. [Accessed 12 9 2016].

- [32] "Pratique Orthophonique," 2011. [Online]. Available: <http://www.orthophonie.eu/>. [Accessed 16 9 2016].
- [33] "Guitar Tablatures," [Online]. Available: <http://guitare-tablatures.fr/son-bruitmusique-guitare/>. [Accessed 16 9 2016]
- [34] MAPAC, "Voice Disorder Database," [Online]. Available: " 16 9 2016. [Online]. Available: . [Accessed 16 9 2016].
- [35] O. Bekhelifi, "Support Vector Machine," Universite de Sciences et de la Technologie d'Oran,2013, Oran
- [36] S. S. Aksenova," *Machine Learning With WEKA*," Universite de California, 2004.
- [37] F. Boudin, « Machine Learning Avec Weka, » Université de Nantes, 2012 .
- [38] J.-P. Goldman, "Tutoriel Praat," Genève, Décembre 2006.
- [39] SEOInternet, 9 12 2014. [Online]. Available: <https://amiante.guide/maladie-amiante/cancer-du-larynx/le-cancer-du-larynx/>. [Accessed 24 9 2016].
- [40] S. Verma, "Throatdisorder," 2010-2013. [Online]. Available: <http://www.throatdisorder.com/voice-disorders/leukoplakia>. [Accessed 24 9 2016].
- [41] G. Lodi , A. Sardella , C. Bez , F. Demarosi and A. Carrassi, « Interventions pour le traitement de la leucoplasie buccale, » 30 5 2009. [Online]. Available: http://cfebd.ups-tlse.fr/index.php?option=com_content&view=article&id=67%3A-interventions-pour-le-traitement-de-la-leucoplasie-buccale-&catid=61%3Alesions-buccales&directory=1. [Accessed 26 9 2016].
- [42] F. Elizabeth , "Phoniatrie," 1 12 2015. [Online]. Available: <http://laboratoiredevoix.com/phoniatrie/> [Accessed 26 9 2016].

[43] « Pourquoi la Voix est-elle Differentiable d'un Individu à Un Autre?, » [Online].
Available: <http://tpeuh.e-monsite.com/pages/synthese/>. [Accessed 26 9 2016].

[44] J. Francis, "*Sinus of Morgagni*," 2016. [Online]. Available: <https://pgblazer.com/sinus-of-morgagni/>. [Accessed 26 9 2016].

Annexe A : Les sons pathologiques de la base de données SVD

A.1

Groupe1 : Leucoplasie

Patients(ID)	Pathologie1	Pathologie2
119	Leukoplakie	
146	Leukoplakie	
497	Leukoplakie	Laryngitis
715	Leukoplakie	Laryngitis
863	Leukoplakie	
875	Leukoplakie	
884	Leukoplakie	Laryngitis
902	Leukoplakie	Laryngitis
935	Leukoplakie	Laryngitis
1186	Leukoplakie	
1252	Leukoplakie	
1303	Leukoplakie	
1405	Leukoplakie	Laryngitis
1448	Leukoplakie	
1557	Leukoplakie	
1559	Leukoplakie	
1599	Leukoplakie	Laryngitis
1613	Leukoplakie	Laryngitis
1617	Leukoplakie	Laryngitis
1629	Leukoplakie	Laryngitis
1632	Leukoplakie	
1665	Leukoplakie	Laryngitis
1715	Leukoplakie	
1783	Leukoplakie	Laryngitis
1810	Leukoplakie	
1864	Leukoplakie	
1940	Leukoplakie	
1987	Leukoplakie	Laryngitis
2075	Leukoplakie	Laryngitis
2079	Leukoplakie	Laryngitis
2101	Leukoplakie	Laryngitis
2242	Leukoplakie	Laryngitis
2351	Leukoplakie	Laryngitis
2443	Leukoplakie	Laryngitis
2549	Leukoplakie	
2556	Leukoplakie	Laryngitis

Groupe2 : Laryngites

Patients(ID)	Pathologie 1	Pathologie 2
107	Laryngitis	
139	Laryngitis	
141	Laryngitis	
493	Laryngitis	
1259	Laryngitis	
1264	Laryngitis	
1264	Laryngitis	
1269	Laryngitis	
1463	Laryngitis	
1578	Laryngitis	
1614	Laryngitis	
1796	Laryngitis	
1806	Laryngitis	
1809	Laryngitis	
1930	Laryngitis	
1965	Laryngitis	
2128	Laryngitis	
2276	Laryngitis	
2296	Laryngitis	
2315	Laryngitis	
2328	Laryngitis	
2404	Laryngitis	
2424	Laryngitis	
2435	Laryngitis	
2510	Laryngitis	
2511	Laryngitis	
2514	Laryngitis	
2516	Laryngitis	
2541	Laryngitis	
2542	Laryngitis	
2567	Laryngitis	
2574	Laryngitis	
2578	Laryngitis	
2582	Laryngitis	
2582	Laryngitis	
2602	Laryngitis	
2610	Laryngitis	

Groupe3 : Rekurrensparese

<i>Patients(ID)</i>	<i>Pathologie 1</i>	<i>Pathologie 2</i>
105	rekurrensparese	
108	rekurrensparese	
120	rekurrensparese	
128	rekurrensparese	
130	rekurrensparese	
138	rekurrensparese	
142	rekurrensparese	
150	rekurrensparese	
152	rekurrensparese	
155	rekurrensparese	
356	rekurrensparese	
358	rekurrensparese	
362	rekurrensparese	
364	rekurrensparese	
365	rekurrensparese	
629	rekurrensparese	
633	rekurrensparese	
670	rekurrensparese	
728	rekurrensparese	
914	rekurrensparese	
929	rekurrensparese	
1289	rekurrensparese	
1304	rekurrensparese	
1466	rekurrensparese	

1630	rekurrensparese	
1638	rekurrensparese	
1641	rekurrensparese	
1830	rekurrensparese	
2160	rekurrensparese	
2216	rekurrensparese	
2247	rekurrensparese	
2274	rekurrensparese	
2342	rekurrensparese	
2393	rekurrensparese	
2408	rekurrensparese	
2414	rekurrensparese	
2430	rekurrensparese	
2512	rekurrensparese	

Groupe4 : Reinke Oedema

<i>Patients(ID)</i>	<i>Pathologie 1</i>	<i>Pathologie 2</i>
564	reinke oedema	
630	reinke oedema	
821	reinke oedema	
1045	reinke oedema	
1056	reinke oedema	
1253	reinke oedema	
1563	reinke oedema	

1572	reinke oedema	
1656	reinke oedema	
1667	reinke oedema	
1717	reinke oedema	
1721	reinke oedema	
1811	reinke oedema	
1828	reinke oedema	
1889	reinke oedema	
1946	reinke oedema	
1963	reinke oedema	
2031	reinke oedema	
2098	reinke oedema	
2161	reinke oedema	
2188	reinke oedema	
2193	reinke oedema	
2211	reinke oedema	
2271	reinke oedema	
2344	reinke oedema	
2345	reinke oedema	
2358	reinke oedema	
2375	reinke oedema	
2381	reinke oedema	
2411	reinke oedema	
2426	reinke oedema	
2594	reinke oedema	
2595	reinke oedema	

Groupe5 : mélange des pathologies 1

<i>Patients(ID)</i>	<i>Pathologie 1</i>	<i>Pathologie 2</i>
105	Phonationsknötchen	
107	Phonationsknötchen	
108	Phonationsknötchen	
120	Laryngitis	
128	Laryngitis	
131	reinke oedema	
139	reinke oedema	
141	rekurrensparese	
358	rekurrensparese	
362	rekurrensparese	
364	rekurrensparese	
365	rekurrensparese	
568	laryngozele	
670	laryngozele	
726	Phonationsknötchen	
728	Laryngitis	
844	Laryngitis	
900	Laryngitis	
914	Laryngitis	
1120	Laryngitis	
1205	Laryngitis	
1269	Laryngitis	

1294	reinke oedema	
1314	reinke oedema	
1449	reinke oedema	
1464	reinke oedema	
1564	rekurrensparsese	
1572	rekurrensparsese	
1614	rekurrensparsese	
1930	rekurrensparsese	
1963	rekurrensparsese	
2296	rekurrensparsese	
2344	rekurrensparsese	
2345	rekurrensparsese	
2426	rekurrensparsese	
2594	rekurrensparsese	

A.2 (suite d'information)

Groupe1 : Leucoplasie

Patients(ID)	Age	Genre	Phonème	Durée
119	57	H	a	1.973620
146	56	H	a	1.869720
497	39	H	a	1.948680
715	63	F	a	0.789000
863	41	F	a	0.722460
875	41	F	a	0.695760
884	71	H	a	0.873500
902	68	H	a	0.841300
935	32	H	a	1.451220
1186	58	H	a	1.986700
1252	52	H	a	0.798740
1303	64	F	a	0.611980
1405	36	F	a	1.235760

1448	67	F	a	0.712940
1557	65	F	a	1.083660
1559	65	F	a	0.630020
1599	32	H	a	1.991460
1613	36	H	a	1.407800
1617	58	H	a	1.914480
1629	51	H	a	1.980020
1632	56	H	a	1.408740
1665	51	H	a	1.772820
1715	60	H	a	1.055140
1783	55	H	a	1.849160
1810	56	H	a	0.731840
1864	66	F	a	1.675620
1940	55	F	a	0.718560
1987	48	H	a	1.737860
2075	54	H	a	0.646100
2079	26	F	a	0.890500
2101	41	H	a	1.644160
2242	35	F	a	0.988400
2351	59	H	a	1.420320
2443	50	H	a	0.602380
2549	64	H	a	1.711560
2556	63	H	a	1.440620

Groupe2 : Laryngites

Patients (ID)	Age	Genre	Phonème	Durée
107	60	H	a	1,21112
139	18	H	a	1,58326
141	41	H	a	1,55586
493	37	H	a	1,33638
1259	32	H	a	1,38784
1264	47	H	a	1,77758
1269	58	H	a	1,4025
1463	71	H	a	0,97912
1578	45	H	a	1,95344
1614	59	H	a	1,3779
1796	31	H	a	0,74288
1806	27	F	a	0,95536
1809	31	F	a	1,37612
1930	60	H	a	0,77948
1965	68	F	a	1,30222

2128	65	F	a	0,54182
2276	35	F	a	0,95058
2296	61	H	a	1,51976
2315	69	F	a	1,96264
2328	69	H	a	1,13784
2404	64	H	a	0,98518
2424	64	H	a	1,56978
2435	48	H	a	0,90718
2510	70	F	a	0,51556
2511	56	H	a	1,71008
2514	50	F	a	0,49432
2516	70	F	a	1,1961
2541	49	H	a	1,16628
2542	51	H	a	0,92818
2567	70	H	a	1,22626
2574	62	H	a	1,83462
2578	54	F	a	1,22626
2582	70	H	a	0,82702
2602	62	F	a	1,06466
2610	43	F	a	0,90308

Groupe3 : rekurrensparese

Patients (ID)	Age	Genre	Phonème	Durée
105	54	F	a	1.110236
108	74	F	a	1.052408
120	65	F	a	1.991007
128	65	F	a	1.253300
130	65	F	a	1.394709
138	49	F	a	1.234870
142	61	F	a	0.700172
150	60	F	a	1.356800
152	62	F	a	1.972906
155	69	F	a	1.136142

356	68	F	a	0.840576
358	51	F	a	1.570020
362	69	H	a	1.994385
364	64	H	a	1.429854
365	40	H	a	1.979829
629	50	H	a	1.555893
633	69	H	a	1.502133
670	76	H	a	1.864024
728	40	H	a	0.732505
914	41	H	a	0.783287
929	36	H	a	0.910529
1289	48	H	a	0.865837
1304	47	H	a	0.769328
1466	65	H	a	0.584598
1630	63	H	a	0.881903
1638	64	H	a	1.795454
1641	48	H	a	1.039946
1830	78	H	a	1.280501
2160	47	H	a	1.996305
2216	47	H	a	0.882754
2247	48	H	a	0.584033
2274	44	H	a	0.772301
2342	75	H	a	0.856469
2393	56	H	a	0.76455
2408	63	H	a	0.98654
2414	61	H	a	1.095374

2430	61	H	a	0.987643
2512	60	H	a	1.095663

Groupe4 : Reinke Oedema

<i>Patients(ID)</i>	<i>Age</i>	<i>Genre</i>	<i>Phonèmes</i>	<i>Durée</i>
564	45	F	a	1.3456
630	59	F	a	0.98763
821	47	F	a	0.8744
1045	52	F	a	1.425638
1056	45	F	a	0.97773
1253	73	F	a	1.05467
1563	53	F	a	1.45673
1572	42	F	a	0.5632
1656	42	F	a	0.58953
1667	42	F	a	0.76300
1717	43	F	a	0.988743
1721	61	F	a	1.457389
1811	49	H	a	0.875363
1828	73	H	a	0.985373
1889	72	H	a	1.023683
1946	72	H	a	1.097353
1963	72	H	a	0.955360
2031	71	H	a	0.567392
2098	61	H	a	0.97643
2161	53	H	a	0.89752
2188	43	H	a	0.65342

2193	49	H	a	1.453862
2211	43	H	a	1.753392
2271	34	H	a	1.054526
2344	34	H	a	0.985242
2345	58	H	a	1.54209
2358	64	H	a	1.034572
2375	46	H	a	0.98763
2381	52	H	a	0.99552
2411	56	H	a	1.01345
2426	58	H	a	0.76441
2594	65	H	a	1.542729
2595	61	H	a	1.07353

Groupe5 : mélange des pathologie1

<i>Patients(ID)</i>	<i>Age</i>	<i>Genre</i>	<i>phonème</i>	<i>Durée</i>
105	46	F	a	1.11652
107	9	F	a	1.056001
108	39	F	a	1.99563
120	59	F	a	1.096522
128	12	F	a	1.542892
131	45	F	a	1.37658
139	45	F	a	1.397755
141	74	F	a	1.438934
358	65	F	a	0.986363
362	65	F	a	0.742729
364	65	F	a	1.752782

365	49	F	a	0.998522
568	46	H	a	0.562288
670	43	H	a	1.096462
726	48	H	a	1.437383
728	60	H	a	0.863733
844	58	H	a	0.996352
900	59	H	a	1.563883
914	60	H	a	1.996373
1120	61	H	a	0.997373
1205	18	H	a	0.524275
1269	41	H	a	1.024336
1294	52	H	a	1.233896
1314	61	H	a	0.873744
1449	65	H	a	1.437744
1464	58	H	a	1.035744
1564	47	H	a	1.548859
1572	47	H	a	1.559733
1614	48	H	a	1.053484
1930	69	H	a	0.985474
1963	64	H	a	0.567333
2296	65	H	a	0.987233
2344	76	H	a	1.874744
2345	77	H	a	1.573830
2426	78	H	a	1.773933
2594	66	H	a	1.211233

