ICACTCE'21
MOROCCO

INTERNATIONAL CONFERENCE ON ADVANCES IN COMMUNICATION TECHNOLOGY, COMPUTING AND ENGINEERING
(ICACTCE 2021)

Certificate of Participation

We hereby certify that:

Khadidja Messaoudene

University M'hamed Bougara of Boumerdes, Algeria

Has participated and presented a paper entitled:

Detection of Knee Osteoarthritis based on Wavelet and Random Forest Model

Authored by: Khadidja Messaoudene, Khaled Harrar

During the International Conference on Advances in Communication Technology, Computing and
Engineering (ICACTCE'21), which was held virtually from Morocco on March 24-26, 2021.

ICACTCE'21
March 24-26, 2021

Dr. Mariya Ouaissa
General Chair

Dr. Zakaria Boulouard
General Co-Chair

Recent Advances in Communication Technology, Computing and Engineering

Article Title	Pages
Gender Prediction and Age Estimation using CNN from Face Images Mohammed Kamel Benkaddour	1-10
Covid-19 Situation in Hungary Using Time Series Analysis Aadil Gani Ganie, Mohd Aaqib Lone, Owais Mujtaba Khanday, Samad Dadvandipour	11-19
Hybrid Localization Approach Based on DV-Distance and Fruit Fly Optimization Algorithm for WSNs Seddik Rabhi, Fouzi Semchedine, Nader Mbarek	20-31
A Deep Learning Solution for Learning Style Detection using Cognitive - Affective Features Farouk Lawan Gambo, Gregory Maksha Wajiga, Etemi Joshua Garba, Aminu Aliyu Abdullahi, Desmond Bala Bisandu	32-43

Detection of Knee Osteoarthritis based on Wavelet and Random Forest Model

Khadidja Messaoudene[1], Khaled Harrar[2]

[1]*LIMOSE Laboratory, University M'Hamed Bougara of Boumerdes, Algeria*
(E-mail: k.messaoudene@univ-boumerdes.dz)
[2]*LIST Laboratory, University M'Hamed Bougara of Boumerdes, Algeria*
(E-mail: khaled.harrar@univ-boumerdes.dz)

Abstract. The most recurrent kind of osteoarthritis is Knee osteoarthritis (KOA). Doctors encounter difficulties for a precise diagnosis through its features and to the naked eye. In this paper, we propose a new approach for the classification of KOA by combining the discrete wavelet decomposition (DWT) and random forest classifier from knee X-ray images. A total of 50 images from patients suffering or not from osteoarthritis were used in this study.

The suggested technique includes image enhancement using the Gaussian filter followed by Haar wavelet transform. Five texture features namely, contrast, entropy, correlation, energy, and homogeneity were extracted from the transformed image, and these attributes were used to differentiate the radiographs into two groups: normal (KL 0) or affected with osteoarthritis (KL2). Four classifiers including random forest, SVM, RNN, and Naïve Bayes were tested and compared. The results obtained reveal that random forest achieved the highest performance in terms of accuracy (ACC = 88%) on X-Ray images of the Osteoarthritis Initiative (OAI) dataset.

Keywords and phrases: Knee osteoarthritis, X-ray images, DWT, Random forest.

1 Introduction

Osteoarthritis (OA) is the most common chronic disease that affects patients older than 70 years [1]. It is estimated that up to 6 % of adults over the age of 30 are symptomatic of knee arthritis [2]. OA is considered a disease of articular cartilage. It is usually due to several clinical and pathological disorders that result in structural and functional failure of synovial joints [3]. The risk factors for OA usually include age, gender, prior joint injury, obesity, genetic predisposition, or even mechanical factors such as misalignment and abnormal joint Shape [4]. KOA is classified into five grades based on Kellgren and Lawrence system [5]. Table 1 shows the different grades of OA disease. The changes in the bone structure caused by OA are illustrated in Figure 1.
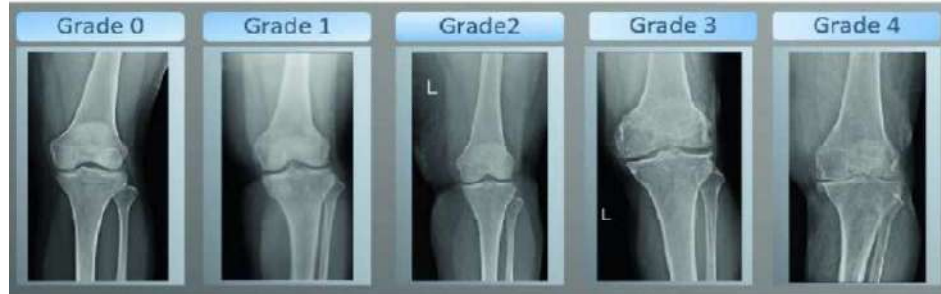
Figure 1: KL classification system for the evaluation of KOA severity [6]

KL Grades	OA Analysis
Grade 0	No changes in the features of the X-Ray image (none)
Grade 1	Join space narrowing (doubtful OA)
Grade 2	Definite join space narrowing (mild OA)
Grade 3	Presence of multiple osteophytes, and some sclerosis (Moderate OA)
Grade 4	Marked joint space narrowing, large osteophyte (severe OA)

Table 1: KOA severity [7]

Multiple imaging modalities may be used to visualize OA. In daily clinical practice, the first most used imaging modality for initial diagnosis and routine follow-up is conventional radiography despite its known limitations [8,9]. Hence, the necessity to develop automatic methods to increase the accuracy of the diagnosis.

Texture analysis is a suitable tool for characterizing images. It has been used in assessing bone fragility as osteoporosis [10,11]. KOA can also be detected using texture analysis from X-ray images. Several works have been done on texture analysis for the diagnosis of KOA. Bayramoglu et al [12] have developed a fully automated system for locating the informative subchondral bone region using adaptive segmentation. To divide radiographs into small regions which meet boundaries of natural texture, they used an over segmentation technique. In the standard ROI and within the proposed adaptive ROIs, Fractal Dimension (FD), Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), Haralick features, and Shannon entropy approaches have been measured, then logistic regression models are used. The results showed 84% of accuracy and 80.4% of overage precision.

To predict OA, Bertlan et al [13] combined fractal and entropic bone texture analysis with Joint Space Width (JSW) and Joint Space Area (JSA). They concluded that the combination of simple subchondral bone radiography with clinical parameters and JSW/A for fractal texture analysis and entropy analysis was better than JSW/A and clinical parameters alone in predicting the

incidence of OA. Tiulpin et al [14], proposed an approach for automatic localization of joint area in knee radiograph, they used the HOG method and SVM classifier. An accuracy of 80% is reached by the proposed technique.

The objective of our paper is to propose a new system for the detection and classification of KOA in knee radiographs. The method consists of the discrete wavelet decomposition and texture features combined with the random forest model, to distinguish two populations composed of 25 subjects with the normal knee (grade KL 0) and 25 pathological cases with minimal OA (grade KL 2). Different classifiers are also tested and compared to our approach. The purpose of our work is to improve the performance of the screening of the disease.

This paper is organized as follows: In Section 2, the dataset and the method are presented, including the DWT, features extraction, and classification. In section 3, the results obtained on the knee images are presented. In section 4, the analysis is discussed, and section 5 offers a conclusion.

2 Material and methods

2.1 Dataset

The dataset consists of 50 X-ray images of the knees which are labeled according to the Kellgren and Lawrence (KL) grading system (KL 0, KL 2). We took grade 0 (no pathology) and compare it to an average grade of the disease KL 2.
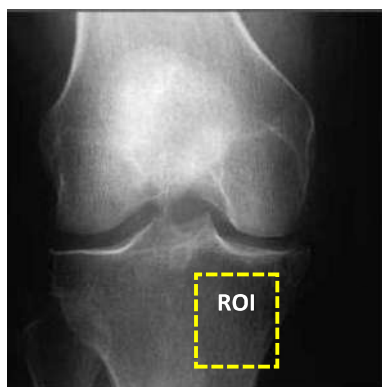

Figure 2: Knee radiographic image from [16]

The database used in our experiment was obtained from the publicly accessible Osteoarthritis Initiative (OAI) [15]. OAI data is a multicenter, longitudinal, and prospective observational study of KOA, composed of X-rays of the knee in fixed flexion [15]. In our study, we worked with the lateral region of the knee radiographic image. Figure 2 shows the region of interest (ROI) used in our work.

2.2 Methods

In this section, we describe a methodology adopted in our approach for the classification of KOA. The proposed system consists of four steps which include a pre-processing stage for artifact removing, features extraction using discrete wavelet transform, and the classification into KL 0 and KL 2. Figure 3 shows the flow diagram of the suggested approach.
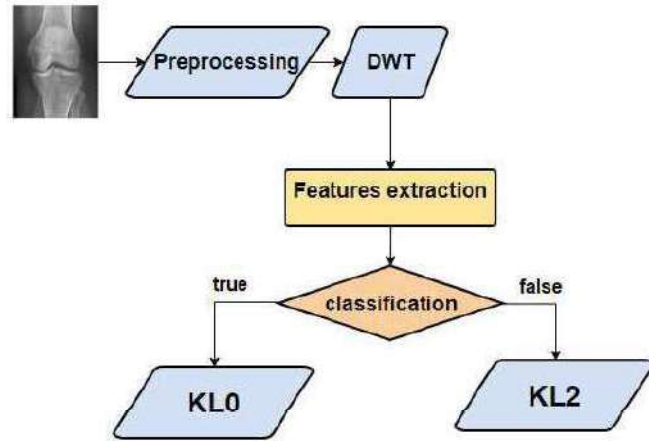

Figure 3: Block diagram of the proposed methodology.

2.2.1 Preprocessing

Major sources of noise in an X-ray imaging system are inherent noise and X-ray quantum noise, which follows Gaussian distribution and the distribution of Poisson, respectively. The need for robust ways to eliminate noise and unwanted particles has therefore arisen [17]. In this work, the Gaussian filter is used for removing acquisition noise with kernels of 3×3 and 5×5. The choice of this filter is justified by the nature of the noise present in the images. Figure 4 shows the results of filtering using the Gaussian filter.
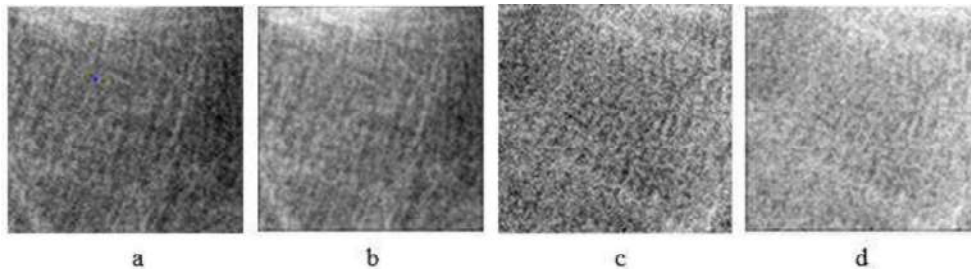

Figure 4: Filtering. Original ROI from KL 0 knee radiograph (a), filtered ROI from KL 0 knee radiograph (b), original ROI from KL 2 knee radiograph (c), filtered ROI from knee radiograph of KL 2 (d).

2.2.2 Discrete Wavelet Transform (DWT)

Wavelets are widely used in image processing for denoising, edge detection, segmentation, compression, coding and decoding, texture analysis, etc. In this work, Haar wavelet transform was used for texture analysis. The Haar transformation can be defined as a sampling process in which an input data sequence with finer and finer resolution increase in powers of 2 is sampled by rows of the transform matrix. In image processing applications, the Haar transform provides a transform domain in which a type of differential energy is concentered in localized regions [18]. Haar [19] defined a complete orthogonal function system in L^p ([0,1]), p [1, ∞[which take values from the set $\{0,2^j: j\epsilon N\}$. This function system has the property that any continuous function at intervals [0,1] can be presented by a uniform and convergent series in terms of this system's elements. The Haar formulation is defined as follows [20,21]:

$$haar(0,t), for \ t \in \big[0,1\big), \ haar(1,t) = \begin{cases} 1 & t \in \left[0, \dfrac{1}{2}\right) \\[2ex] -1 & t \in \left[\dfrac{1}{2}, 1\right) \end{cases} \tag{1}$$

with $haar(k,0) = \lim\limits_{t \to 0} haar(k,t), \quad haar(k,1) = \lim\limits_{t \to 1} haar(k,t)$ and at the point of discontinuity within the interior (0,1) $haar(k,t) = \dfrac{1}{2}(haar(k,t-0) + haar(k,t+0))$.

2.2.3 Features extraction

The five texture features are extracted: entropy, contrast, energy, correlation, and homogeneity derived from the co-occurrence matrix of the grey level (GLCM).
The approach GLCM is based on pixel intensity distribution statistical studies [22]. Haralick [23] proposed the gray levels matrices (GLCM) that have become one of the most commonly used and well-known texture characteristics.
For a replacement P_d, the matrix of the G*G gray stage is defined as follows: the entry(i,j) of P_d is the number of occurrence of the pair of gray levels i and j that are separated by a distance d. It is formally described as [24]:

$$p_d(i,j) = \big|\{(r,s),(t,v): I(r,s) = i, I(t,v) = j\}\big| \tag{2}$$

Where $(t,v) = (r + dx, s + dy)$, $(r,s),(t,v)$ $\in N*N$ and $|\ \ |$ is a set's cardinality.
Important features are taken out from the matrix as the texture representation. The texture features retained for this study are:

$$\text{Energy: } \sum_{i,j} p(i,j)^2 \tag{3}$$

Contrast: $\sum |i\text{-}j|^2 p(i,j)$ (4)

Correlation: $\sum_{i,j} \dfrac{(i-\mu_i)(j-\mu_j)}{\partial_i \partial_j}$ (5)

Homogeneity: $\sum_{i,j} \dfrac{p(i,j)}{1+|i-j|}$ (6)

Entropy: $\sum_{i,j} p(i,j) \log p(i,j)$ (7)

2.2.4 Classification

A random forest model is used to classify the input image into one of two grades of OA: KL 0 and KL 2. Random forest is a model based on a set of defined tree classifiers $\{T(x,\Theta k)$, with $k=1, ...\}$. The set $\{\Theta k\}$ are identically distributed independent random vectors. A tree emits one vote for the most common category at input x [26]. In the image classes, each tree's leaf node is labeled with posterior distribution estimates. Each internal node contains a test that best splits the space of data to be classified.

An image is defined by sending any tree down and aggregating the leaf distributions that have been reached. During the training, it is possible to introduce randomness at two items: by sub-sampling the training data so that each tree is designed using a different subset of the tree; and by determining the node tests [27]. Random forest is a multi-purpose method that is relevant to problems of regression and classification, including multi-class classification. They offer an internal generalization error estimate because cross-validation is unnecessary.

The random forest approach has many benefits. In particular, it predicts which features are relevant in the classification. It can process large data sets efficiently. It can also be used as an important tool for calculating missing data.

2.2.5 Statistical analysis

The performance of each parameter to distinguish between the classes (KL 0, KL 2) was assessed using the correlation coefficient. The relationship (or correlation) between the two variables is indicated by the symbol (r) and quantified with a number ranging from -1 to +1. Zero means no correlation, while 1 means complete or ideal correlation. The r sign shows the direction of the correspondence. A negative r means the variable is inversely related to each other [28].

Several metrics were calculated from the matrix and this study included:

TP: is the correctly defined number of true positive, pathological patients (KL 0).

FN: the number of false negatives healthy patients (KL 2) incorrectly identified.

TN: the number of true negatives, healthy patients (KL 0) correctly identified.

FP: the number of false positives, pathological patients (KL 2) incorrectly identified.

FPrate: is the proportion of negative cases wrongly defined as positive in the data (i.e. the probability that false alerts will be raised):

$$FPrate = \frac{FP}{FP + TN} \qquad (8)$$

TPrate: is a calculation of the proportion of positive cases in the data correctly defined as such:

$$TPrate = \frac{TP}{TP + FN} \qquad (9)$$

Accuracy (ACC): A measurement metric that allows a model to quantify the total number of accurate predictions. This metric is expressed by the formula (10):

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \qquad (10)$$

Precision: The precision assesses how accurate the model is in predicting positive labels. The formula for Precision is given by the expression (11):

$$Precision = \frac{TP}{TP + FP} \qquad (11)$$

Recall: measurement of the percentage of positive actuals of a correctly defined model (True Positive). The formula for the recall is given by the expression (12):

$$Recall = \frac{TP}{TP + FN} \qquad (12)$$

F1-Score: This metric provides precision and recall weighted average. This score takes both false negatives and false positives into consideration. F1-Score can be used to provide a more honest and useful evaluation of the learning model on a highly unbalanced dataset. The formula for F1-Score is given by the expression (13):

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (13)$$
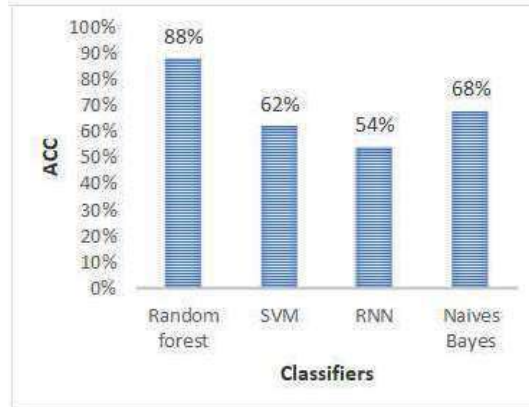
Figure 5: Comparison of classification performance of four different algorithms: Random forest, SVM, RNN, and Naive Bayes.

3   Experimental results

The proposed system is performed on various knee radiographs of different ages. The experiment is carried out on 50 Knee radiographs. Various characteristics were calculated such as entropy, homogeneity, energy, contrast, and correlation. Texture parameters are derived from radiograph images and applied to several classifiers.

The performance of the classification of the four algorithms above is carried out and the random forest provided an accuracy rate of 88%, SVM was up to 62%, RNN to 54% and Naive Bayes provided 68% of accuracy (Figure 5).

Based on the results obtained from the proposed methodology, the precision was better for the classification of random forest. The data is trained and tested to obtain the classification rate for each classifier discussed above. The following tables summarize the accuracy of the classification of the four classifiers tested and the parameters were also calculated.

| Parameter | KL 0 | KL 2 | |
	Mean ± SD	Mean ± SD	*r*
Entropy	1.0398 ± 0.0053	1.0198 ± 0.0951	0.2935
Contrast	24.4166 ± 0.3652	24.4056 ± 0.4188	-0.2647
Correlation	-9.0336e-04± 0.0150	4.5600e-04 ± 0.0175	-0.2833
Energy	0.2458± 6.4031e-04	0.2459± 6.7454e-04	0.0598
Homogeneity	0.5592 ± 0.0066	0.5595 ± 0.0079	-0.3088

Table 2: A comparison between the features to discriminate KL 0 and KL 2.

Model	TP	FP	FN	TN	TP rate	FP rate
Random forest	24	2	4	20	0.86	0.09
SVM	17	8	11	14	0,61	0.36
RNN	17	8	15	10	0.53	0.44
Naive Bayes	18	6	10	16	0.64	0.27

Table 3: Confusion matrix comparing different models.

Model	ACC	Precision	Recall	F1-Score
Random forest	88 %	0.923	0.857	0.14
SVM	62%	0.68	0.61	0.39
RNN	54%	0.68	0.53	0.39
Naive Bayes	68%	0.75	0.64	0.32

Table 4: Evaluation of the accuracy of classification for all models.

Authors	Year	Method	Classifier	Results	Dataset
Thomson et al [28]	2015	Texture information (Fractal signature) Shape information	Random forest	AUC = 84.9%	OAI
Janvier et al [29]	2016	Fractal analysis	Logistic regression	AUC = 71%	OAI
Kaggie et al [30]	2017	Texture parameters	Neural network	AUC = 74%	OAI
Brahim et al [31]	2019	Multivariate linear Regression (MLR)	Naive Bayes random forest	AUC = 82.98%	OAI
Proposed method	2021	DWT	Random forest	AUC = 88.9%	OAI

Table 5: Comparison of proposed system classification rates with existing studies in the literature

4 Discussion

Experiments show that the combination of wavelets and random forest classifier leads to a considerable improvement of the overall classification performance, with an AUC of 88.9% compared to the other classifiers (SVM, RNN, Naive Bayes). The parameters of the classifiers are tuned and optimized to provide the best performance.

Various parameters were included in this analysis to distinguish two groups composed of 25 subjects in KL 0 and 25 patients in KL 2. Table 2 shows the distribution (mean and standard deviation (SD)) for the normal and pathological groups of the parameters collected. We can notice that there is a negligible positive relationship ($r = 0.0598$) of energy from Table 2 and a weak relationship of all other parameters (entropy, homogeneity, contrast, and correlation).

Table 3 illustrates a comparison of the four classifiers. As we can see random forest provided the best performance with a 0.86 TP rate and lowest FP rate (0.09). It outperforms the second-ranked method which is Naive Bayes with a remarkable margin. The worst performing method would be RNN with a 0.53 TP rate and 0.44 FP rate. Table 4 shows the comparison between the parameters involved in this study. We can notice that random forest revealed the best performance in terms of the discrimination of the two populations based on every metric possible. This model achieved an accuracy of 88 % compared to the second method with 68%. A precision of 0.92 in comparison with RNN with a much lower precision of 0.68. These results are highlighted by bar plots presented in Figure 5.

Various studies have already shown the value of texture analysis in predicting the evolution of KOA. However, these works are difficult to compare since they use different techniques on small and non-free populations. Table 5 summarizes the comparison of overall automated classification performances between the proposed method and other techniques for detecting KOA. The metric for the performance evaluation is Area Under Curve (AUC). Authors in [30] showed that combining the results of the shape and texture-based classifiers leads to considerable improvement in overall classification performance, with an AUC of 0.849. In [31] the AUC of the discrimination between grade KL 0 (non-OA) and grade KL 2 (minimum OA) was 0.71. Authors in [32] combined texture analysis with Neural Network classifier to predict radiographic disease progression over 3 years, and achieved high sensitivity (86%), a specificity of 64%, and an AUC of 74% for the prediction of OA progression. Finally, in [33], the authors proposed a computer-assisted method and achieved an overall AUC of 82.98%. Table 5 reveals that our proposed method achieved the highest rate of AUC comparing to the other methods, with a score of 88.9%. This high rate of classification is reached with the combination of the discrete wavelet decomposition and the random forest classifier which is suitable for medical applications. This work could add value to the screening of KOA in clinical routine.

5 Conclusion

In this paper, effective detection of KOA is proposed. Five features were chosen and trained using a random forest classifier. Experimental results suggest that more than 88% of KOA was diagnosed. The results achieved are promising when compared to methods existing in the

literature. In future work, we will focus on increasing the number of selected features to improve the accuracy. It is also suitable to combine several modalities to provide better results.

Moreover, in this paper, we focused on basic diagnosis (KL0 vs KL 2). In the future, we will extend the classification to differentiate between the other KL classes, to assess the consistency of each grade of osteoarthritis.

References

[1] J. Chodosh, S. C. Morton, W. Mojica, M. Maglione, M J. Suttorp, L. Hilton¸ S. Rhodes, P. Shekelle, Meta-analysis: chronic disease self-management programs for older adults, Annals of Internal Medicine, 143(2005), 427–I32, DOI: 10.7326/0003-4819-143-6-200509200-00007.

[2] S. Jevsevar, Treatment of Osteoarthritis of the Knee: Evidence-Based Guideline, Journal of the American Academy of Orthopaedic Surgeons, 21(2013), 571-576, DOI: 10.5435/JAAOS-21-09-571.

[3] R. F. Loeser, S. R. Goldring, C. R. Scanzello, M. B. Golding, Osteoarthritis: A disease of the joint as an organ, Arthritis & Rheumatology, 64(2012), 1697–1707, DOI: 10.1002/art.34453.

[4] J. Hunter, D. T. Felson, Osteoarthritis, British Medical Journal, 332(2006), 639–420, DOI: 10.1136/bmj.332.7542.639.

[5] H. Kellgren, J. S. Lawrence Radiological Assessment of Osteoarthritis, Annals of the rheumatic diseases., 16(1957), 494–502, DOI:10.1136/ard.16.4.494.

[6] S.Chan, K.Dittakan, S. Hegadi, osteoarthritis stages classification to human joint imagery using texture analysis: a comparative study on ten texture descriptors, Communications in Computer and Information Science, 1036(2019), 209–225, DOI: 10.1007/978-981-13-9184-2_19.

[7] S. S. Gornale, P. U. Patravali, R.R. Manza, Detection of Osteoarthritis using Knee X-Ray Image Analyses: A Machine Vision based Approach, International Journal of Computer Applications, 145(2016), 0975–8887, DOI: 10.5120/ijca2016910544.

[8] D. Hayashi, F.W. Roemer, A. Guermazi, Imaging of Osteoarthritis by Conventional Radiography, MR Imaging, PET–Computed Tomography, and PET–MR Imaging, PET clinics, 14(2019), 17–29, DOI: 10.1016/j.cpet.2018.08.004.

[9] K. Harrar, M. Khider, Texture analysis using multifractal spectrum, International Journal of Modeling and Optimization, 4(2014), 336–341, DOI: 10.7763/IJMO.2014.V4.396.

[10] K. Harrar, L. Hamami, S. Akkoul, E. Lespessailles, R. Jennane, Osteoporosis Assessment Using Multilayer Perceptron Neural Network, 3rd International Conference on Image Processing Theory, Tools and Applications IPTA'12, Istanbul, Turkey, (2012), 217–221, DOI: 10.1109/IPTA.2012.6469528.

[11] K. Harrar, R. Jennane, Quantification of Trabecular Bone Porosity on X-Ray Images, Journal of Industrial and Intelligent Information, 3(2015), 280–285, DOI:10.12720/jiii.3.4.280-285.

[12] N. Bayramoglu, A. Tiulpin, J. Hirvasniemi, M.T. Nieminen, S. Saarakkala, Adaptive Segmentation of Knee Radiographs for Selecting the Optimal ROI in Texture Analysis, arXiv, 1(2019), DOI: 10.1016/j.joca.2020.03.006.

[13] Z. Bertalan, R. Ljuhar, B. Norman, D. Ljuhar, A. Fahrleitner-Pammer, H.-P. Dimai, S. Nehrer, Combining fractal- and entropy-based bone texture analysis for the prediction of osteoarthritis: data from the multicenter osteoarthritis study (most), Osteoarthritis and Cartilage, 26(2018), 10–59, DOI: 10.1016/j.joca.2018.02.112.

[14] A.Tiulpin, J. Thevenot, E.Rahtu, S. Saarakkala, A novel method for automatic localization of joint area on knee plain radiographs, Scandinavian Conference on Image Analysis, 3(2017), 290–301, DOI: 10.1007/978-3-319-59129-2_25.

[15]   G. Lester, The osteoarthritis initiative: a NIH public–private partnership, Hospital for Special Surgery Journal, 8(2012), 62–3, DOI: 10.1007/s11420-011-9235-y.

[16]   A. Tiulpin, J. Thevenot, E.rahtu, P. Lehenkari, S.Saarakkala, Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach, Scientific reports, 8(2018),1–10, DOI: 10.1038/s41598-018-20132-7.

[17]   S. Yuewen, L. Ximing, C. Peng, L. Litao, Z. Zhongwei, Digital radiography image denoising using a generative adversarial network, Journal of X-ray science and technology, 26(2018), 523–534, DOI: 10.3233/XST-17356.

[18]   P. Porwik, A. Lisowska, The Haar–Wavelet Transform in Digital Image Processing: Its Status and Achievements, Machine Graphics & Vision, 13(2004), 79–98, DOI: 10.1016/S0045-7906(01)00011-8.

[19]   A.Haar, Zur Theorie der orthogonalen Funktionensysteme, Mathematische Annalen, 69(1911), 331–371, DOI: 10.1007/BF01456927.

[20]   H. F. Harmuth, Sequency Theory, FAAP, New York, 1978.

[21]   K. Grochenig, W. R. Madych, Multiresolution Analysis, Haar bases and self–similar tilings of Rn, IEEE Transactions on Information Theory, 38(1992), 556–568, DOI: 10.1109/18.119723.

[22]   M. Mirmehdi, X. Xie, J. Suri, Handbook of texture analysis, Imperial College Press, 2008, 424, DOI: 10.1142/p547.

[23]   B. Julesz., E. N. Gilbert, L. A. Shepp, H. L. Frisch, Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited, Perception, 2(1973), 391–405, DOI: 10.1068/p020391.

[24]   M. Tuceryan, A.K. Jain, The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang, World Scientific Publishing, 1998, 207–248, DOI: 10.1142/3414.

[25]   B. Leo, Random forests, Machine Learning, 45(2001), 5–32, DOI: 10.1023/A:1010933404324.

[26]   A. Bosch, A. Zisserman, X. Munoz, Image classification using random forests and ferns, IEEE 11th international conference on computer vision (ICCV), Rio de Janeiro, Brazil, 2007,1–8, DOI: 10.1109/ICCV.2007.4409066.

[27]   A.haldun, user's guide to correlation coefficients, Turkish journal of emergency medicine, 18(2018), 91–93, DOI: 10.1016/j.tjem.2018.08.001.

[28]   J. Thomson, T. O'Neill, D. Felson, T. Cootes, Automated Shape & Texture Analysis for Detection of Osteoarthritis from Radiographs of the Knee", International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015, 127–134, DOI: 10.1007/978-3-319-24571-3_16.

[29]   T. Janvier, R. Jennane, A. Valery, K. Harrar, M. Delplanque, C. Lelong, D. Loeuille, H. Toumi, E. Lespessailles, Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative, Tibial bone texture & knee OA progression, Osteoarthritis and Cartilage, 25(2017), 259–266, DOI: 10.1016/j.joca.2016.10.005.

[30]   J. Kaggie, R.Tovey, J. MacKay, F. Gilbert, F. Gallagher, A. McCaskie, M. Graves, Automated Textural Classification of Osteoarthritis Magnetic Resonance Images, International Society for Magnetic Resonance in Medicine, (2017), https://submissions.mirasmart.com/ISMRM2018

[31]   A. Brahim, R. Jennane, R. Riad, T. Janvier, L. Khedher, H. Toumi, E. Lespessailles,A Decision Support Tool for Early Detection of Knee OsteoArthritis using X-ray Imaging and Machine Learning: Data from the OsteoArthritis Initiative, Computerized Medical Imaging and Graphics, 73(2019), 11–18, DOI: 10.1016/j.compmedimag.2019.01.007

Second International Conference on Artificial Intelligence and its Applications (AIAP'2022)-Online

JANUARY 24-26, 2022 - EL OUED, ALGERIA.

ATTESTATION OF PARTICIPATION

This is to certify that

Khadidja Messaoudene

Has presented a paper:

Title: A hybrid LBP-HOG model and naive Bayes classifier for knee osteoarthritis detection: data from the osteoarthritis initiative

Authors: Khadidja Messaoudene, Khaled Harrar

at the Second International Conference on Artificial Intelligence and its Applications, AIAP'22
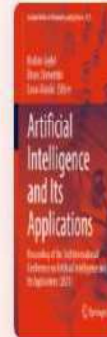
held in El Oued, Algeria, JANUARY 24-26, 2022.

Prof. Brahim Lejdel
Chairman
Conference chair

A Hybrid LBP-HOG Model and Naive Bayes Classifier for Knee Osteoarthritis Detection: Data from the Osteoarthritis Initiative

Artificial Intelligence and Its

Applications

(AIAP 2021)

Khadidja Messaoudene ✉ & Khaled Harrar

Sections	Figures	References

Abstract

Keywords

Introduction

Material and Methods

Results and Discussion

Conclusion

References

Author information

Abstract

A hybrid LBP-HOG model and naive Bayes classifier for knee osteoarthritis detection: data from the osteoarthritis initiative

Khadidja Messaoudene[1][0000-0003-4009-4868] and Khaled Harrar[2][0000-0003-4559-0940]

[1] LIMOSE Laboratory, University M'Hamed Bougara of Boumerdes, Algeria
k.messaoudene@univ-boumerdes.dz
[2] LIST Laboratory, University M'Hamed Bougara of Boumerdes, Algeria
khaled.harrar@univ-boumerdes.dz

Abstract. Knee OsteoArthritis (KOA) is a disease characterized by a degeneration of cartilage and the underlying bone. It does not evolve uniformly; it can stay silent for a long time and can quickly intensify for several months or weeks. For this reason, it is necessary to develop an automatic system for diagnosis and reduce the subjectivity in the detection of the disease. In this paper, we present a method for detecting knee osteoarthritis based on the combination of histograms of oriented gradient (HOG) and local binary pattern (LBP). Four classifiers including KNN, SVM, Adaboost, and Naïve Bayes were tested and compared for the prediction of the illness. A total of 620 X-Ray images were analyzed, composed of 310 images from healthy subjects (Grade 0), and 310 images from pathological patients (Grade 2). The results obtained reveal that Naïve Bayes achieved the highest performance in terms of accuracy (ACC = 91%) on the Osteoarthritis Initiative (OAI) dataset. The fusion of HOG and LBP features in KOA classification outperforms the use of either feature alone and the existing methods in the literature.

Keywords: Knee osteoarthritis, X-ray images, LBP, HOG, Naïve Bayes.

1 Introduction

Knee osteoarthritis (KOA) is a chronic disease of the joint which progressively destroys the cartilage. It is often mistakenly thought to be associated with aging against which little can be done, whereas it is a real disease that causes disability in about 40% of adults over the age of 70 [1]. As for osteoporosis [2, 3], KOA is a highly prevalent health problem. KOA is typically diagnosed by radiography (X-ray imaging) as well as other imaging modalities like MRI and CT scan. Despite many limitations, conventional radiography (X-ray imaging) remains the first option and most widely utilized for OA because it is more inexpensive and accessible than other diagnostic modalities. The Kellgren and Lawrence (KL) scale is the most frequently used for defining the level of knee OA [4]. The grade in the KL classification system ranges from 0 to 4, according the intensity of OA. Figure 1 depicts the illness phases according to the KL categorization system.
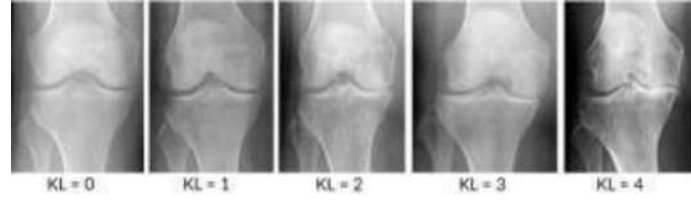
Fig. 1. Knee OA severity [5]

The treatment of knee OA depends on the quality of diagnosis that is why many researchers propose automatic systems aid diagnosis in rheumatology. In [6] the researchers proposed an approach for automatic localization of joint area in knee radiograph. They used the HOG and SVM classifier, the proposed methodology achieved an accuracy of 80%. Haftner et al. [7] describe a method of collecting additional information on the texture of the lateral and medial condyles of the distal femur. Shannon entropy and six other indicative features describing texture roughness and anisotropy were applied. Their framework selected an optimal combination of different texture parameters from six different regions for evaluation with various classifiers. They achieved an accuracy of 72%. Akter et al. [8] described an approach to extract texture features in radiographic images for osteoarthritis detection. The proposed method is based on Zernike orthogonal features and group method of data handling (GMDH) Neural Networks. This technique improved the detection accuracy by 82.8% for lateral images. In [9] the authors combined different texture descriptors (LBP and GLCM) with different classifiers (KNN, SVM, neural network) to determine the intact stage of knee osteoarthritis in radiographic images. The highest performance was obtained with a multilayer perceptron (MLP) classifier, with an overall accuracy of 90.2%.

In this paper, LBP and HOG methods are combined with the naive Bayes classifier on the OAI dataset to detect knee osteoarthritis in two stages of the disease: KL0 (normal case), KL2 (pathological case). First, LBP parameters are extracted from the images, then the HOG parameters are estimated, finally, several classifiers (Naive Bayes, SVM, Adaboost, and KNN) are carried out for the prediction of the disease. In the first stage, each model (LBP, or HOG) is tested and evaluated alone, then a combination of the two models is performed to improve the ability of the prediction. This is the first study to combine LBP and HOG for KOA detection.

This paper is structured as follows: Section 2 covers the material and approach and its extensions; Section 3 illustrates the results and discussion, and Section 4 summarizes the findings.

2. Materiel and Methods

2.1 Dataset

In this study, the data from the OAI was used. The OAI covers persons at risk of developing clinical tibiofemoral osteoarthritis. A total of 4,796 participants aged 45-79

years took part in the study between 2004 and 2006. The images were analyzed using the Kellgren-Lawrence (KL) grading method [10]. The present study focuses on the early detection of knee OA. Therefore, only radiographs with a KL grade 0 (no OA) and a KL grade 2 (minimal OA) were considered. We used 620 radiographs of the knee in the lateral region. Fig. 3 shows the ROI used in our study.

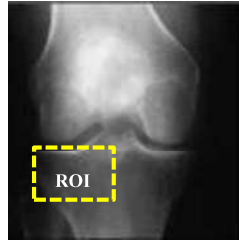**Fig. 2.** Knee radiographic image.

2.2 Methods

The major goal of this study is to present a texture feature extraction technique that performs well in this situation. In our tests, we employed the LBP descriptor, the HOG, and a combination of them. Figure 3 depicts the design of our system. A brief overview of each phase of our method is provided below.
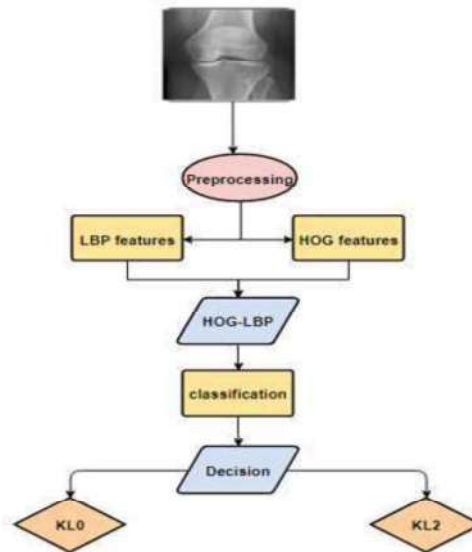

Fig. 3. Proposed classification system

Preprocessing

The anisotropic diffusion filter (ADF) has been effectively used in image processing to eliminate high frequencies while preserving the major existing objects without deleting substantial elements of the image content, often edges, lines, or other features crucial for image interpretation [11]. ADF is defined as:

$$\frac{\partial I}{\partial t} = div(c(x,y,t)\nabla I) = \nabla c.\nabla I + c(x,y,t)\Delta I$$

(1)

I is the input image, Δ represents the Laplacian, ∇ is the gradient, $c(x,y,t)$ denotes the diffusion coefficient, div() is the divergence operators. Fig. 4 shows the results of filtering.
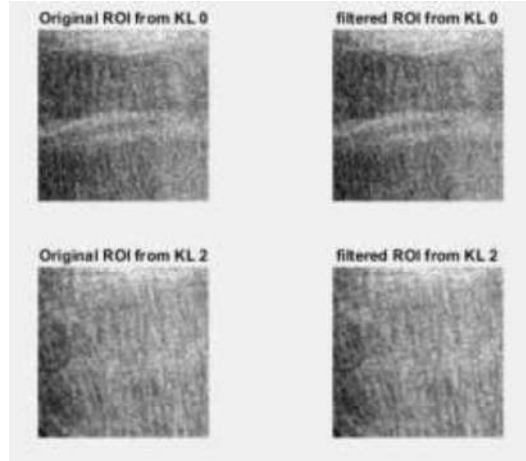
Fig. 4. Results of filtering

Histogram of Oriented Gradient (HOG)

The HOG method was suggested by Dalal and Trrigs in 2005 [12]. The original idea of this descriptor is that the local structure of the object is described by calculating the gradient distribution of the local intensities or the direction of the contours without knowing the localization of the gradient or the position of the contours in the image [13]. The HOG descriptors are the main features that encode object features into a sequence of specific numbers, which can be used to distinguish items from each other [14]. Gradients are the rate of modifications in local intensity at a specific pixel position. A gradient is a quantity of the vector that has both direction and magnitude. The

pixel gradient magnitude $V(x,y)$ and direction $\alpha(x,y)$ are indicated in equations (2) and (3) respectively:

$$V(x,y) = \sqrt{Vx(x,y)^2 + Vy(x,y)^2} \qquad (2)$$

$$\alpha(x,y) = \arctan[Vx(x,y)/Vy(x,y)] \qquad (3)$$

Figure 5 shows the HOG features extracted from an image using three different cell sizes. This figure shows the visualization of cell sizes [2 2], [4 4], and [8 8]. The size cell [2 2] contains more shape information than the size cell [8 8] in their visualization. In the latter case, the dimensionality of the feature vector using HOG increases compared to the former. A good choice is the cell size [8 8]. By using this size, the number of dimensions is limited, which speeds up the training process. It also contains enough information to visualize the shape of the mode image.
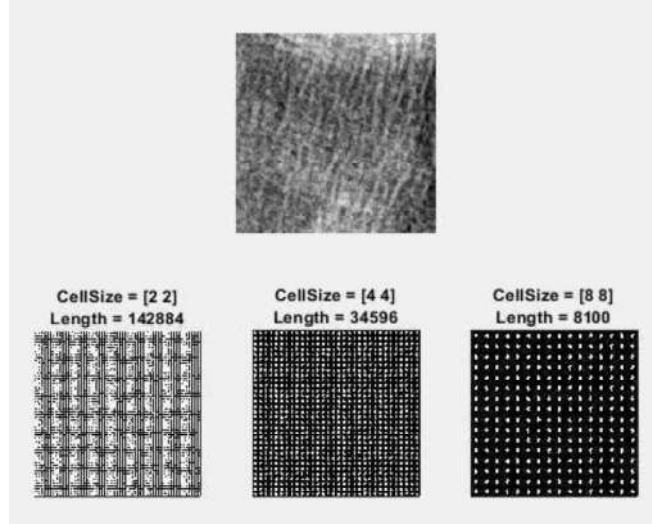
Fig. 5. HOG features of an X-ray image with different cell sizes

Local Binary Pattern (LBP)

Ojala et al. developed the LBP approach to measure texture patterns [15]. The LBP approach compares each neighboring pixel in neighborhood 3×3 against the center pixel to determine if it is 0 or 1. Each binary value is then multiplied by the corresponding weight. An LBP number for a unit of texture is obtained by adding up all the multiplications. LBP can generate up to 256 patterns.

$$LBP(x_c, y_c) = \sum S(i_a - i_c)2^n \qquad (4)$$

6

Where i_a is the gray level of the pixel (x_c, y_c), i_c is the gray level of the circular neighborhood of the pixel (x_c, y_c), and S is the Heaviside function.

Classification

After the feature extraction step (HOG, LBP), we applied the Bayes model due to its speed and efficiency in the prediction of knee osteoarthritis. The naive Bayes system is a highly simplified Bayesian probabilities model. The naive Bayes classifier is considered one of the strongest independence assumptions [16]. This indicates that the probability of one characteristic has no influence on the results of the other. First, we tested several classifiers on the LBP parameters alone and then on the HOG. Then we combined the parameters (LBP-HOG) and tested the different classification models (Naive Bayes, SVM, Adaboost, and KNN).

Model Evaluation

Knowing a model's accuracy is necessary, but it is not sufficient to provide a full understanding of a model's level of efficiency. So, there are other measurement criteria that will help understand how performative the model is? The other metrics used in this study are: Precision, recall, ROC curve, MCC, etc.

Accuracy (ACC): A metric that allows a model to quantify the number of total accurate predictions.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Precision (Pr): is defined as the ratio of correct positive predictions to all positive predictions.

$$Pr = \frac{TP}{TP + FP} \tag{6}$$

Sensitivity (True Positive Rate (TPR)) : is a measurement of the proportion of positives that are correctly identified.

$$TPR = \frac{TP}{TP + FN} \tag{7}$$

Specificity (True Negative Rate (TNR)) is the proportion of negatives that are correctly identified.

$$TNR = \frac{TN}{TN + FP} \tag{8}$$

FPrate (False Positive Rate (FPR)): is the percentage of negative values wrongly defined as positive in the data.

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

F1-Score: is the weighted average between precision and sensitivity.

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \tag{10}$$

Where TP is true positive, TN true negative, FP false positive, and FN false negative.

3. Results and discussion

In the first test, we evaluated the performance of the LBP parameters, the HOG parameters, and the LBP-HOG system. To perform a comparison, we tested radiographic knee images taken from different subjects for the two stages of osteoarthritis (0 and 2). For each stage, 310 images were involved.

For each case (LBP, HOG, LBP-HOG), we tested four classifiers (Adaboost, Naive Bayes, SVM, and KNN). The results are presented in the following tables.

Table 1. Classification performance for LBP features

Classifier	TP	FP	FN	TN	Pr	FPR	TNR	TPR	F1-Score	ACC
Naïve Bayes	**213**	**91**	**95**	**221**	**0.70**	**0.29**	**0.71**	**0.69**	**0.37**	**0.70**
SVM	195	115	121	189	0,63	0,38	0.62	0.62	0.43	0.62
Adaboost	155	145	153	167	0.52	0.46	0.54	0.50	0.48	0.52
KNN	150	158	165	147	0,49	0,52	0.48	0.48	0.50	0.48

Table 2. Classification performance for HOG features

Classifier	TP	FP	FN	TN	Pr	FPR	TNR	TPR	F1-Score	ACC
Naïve Bayes	216	93	105	206	0.70	0.31	0.69	0.67	0.37	0.68
SVM	**237**	**85**	**77**	**221**	**0.74**	**0.28**	**0.72**	**0.75**	**0.36**	**0.74**
Adaboost	195	142	118	165	0.58	0.46	0.54	0.62	0.50	0.58
KNN	158	148	162	152	0.52	0.49	0.51	0.49	0.49	0.50

Table 3. Classification performance for the combination of LBP-HOG features

Classifier	TP	FP	FN	TN	Pr	FPR	TNR	TPR	F1-Score	ACC
Naive Bayes	**286**	**20**	**36**	**278**	**0.93**	**0.07**	**0.93**	**0.89**	**0.11**	**0.91**
SVM	257	36	48	279	0.88	0.11	0.89	0.84	0.18	0.86
Adaboost	192	114	134	180	0.63	0.39	0.61	0.59	0.42	0.60
KNN	248	62	69	241	0.80	0.20	0.80	0.78	0.29	0.79

Table 1 depicts a comparison of the four classifiers for the LBP parameters. As we can observe, Naive Bayes provided the best performance with a TPR of 0.69 and the lowest FPR (0.29). It outperformed the second-ranked method (SVM), by a significant margin. The worst performing method would be KNN with a low TNR (0.48), a high FPR (0.52), and a TPR of 0.48. The LBP model is shown to perform better with the Naive Bayes classifier.

The results of the classification using the HOG method are shown in Table 2. We can see that the combination of HOG parameters with the SVM model gave excellent results with an accuracy of 74% and a low FPR (0.28). The KNN classifier gave bad results in terms of FPR (0.49).

The combined performance of LBP-HOG with four classifiers is shown in Table 3. As can be seen, Naive Bayes provided the best performance with a TPR of 0.89 and the lowest FPR (0.07). It outperformed the second-highest ranked method (SVM), by a remarkable margin. On the other hand, Adaboost, with a TPR of 0.59 and an FPR of 0.39, is the worst-performing method in this case.

Regarding the F1-score, the same findings are noticed. The combination of LBP and HOG models provided the lowest rate (0.11), where LBP gave 0.37, and HOG achieved 0.36 of F1-score.

Through the results described in the previous Tables, it is clear that the combination of characteristics of the two models (LBP-HOG) achieved the best detection rate. The results obtained with the combination show a better performance than the systems based on each method alone.

Table 4 illustrates a comparison of our proposed method with the state of the art. Tiulpin et al [6] used HOG and SVM to detect osteoarthritis and provided an accuracy of 80%. Haftner et al [7] achieved a lower accuracy (72%) with entropy and LDA technique. Akter et al [8] achieved an accuracy of 82.8% using Zernike and GMDH classifier. Peuna et al [9] combined LBP and GLCM with MLP classifier and provided good results in terms of accuracy (90.2%). Examining Table 4, our proposed method achieved the highest accuracy with the combination of LBP and HOG descriptors and Naïve Bayes as classifier, where the rate achieved 91%.

Table 4. A comparison study with the state of the art

Author	Year	Method	Classifier	ACC (%)
Tiulpin et al [6]	2017	HOG	SVM	80
Haftner et al [7]	2017	Entropy	LDA	72
Akter et al [8]	2019	Zernike	GMDH	82.8
Peuna et al [9]	2021	LBP-GLCM	MLP	90.2
Proposed method	**2021**	**LBP-HOG**	**Naive Bayes**	**91**

4. Conclusion

This study offers an efficient and precise approach for the classification and identification of knee OA. The present work was carried out on a dataset composed of 620 radiographs of patients divided into 310 images of healthy subjects (Grade 0), and 310 images from patients suffering from KOA (Grade 2). Following the successful implementation of the proposed classification system using HOG and LBP methods with Naive Bayes classifier, we have demonstrated that the proposed system provided promising results in terms of classification of patients suffering from Knee OA with high accuracy (ACC = 91%).

We believe that our system can help and assists doctors in osteoarthritis diagnosis. In the future, we are planning to improve the feature extraction stage and the classification using other techniques. We are exploring other types of features to train classifiers and analyze the effects of other machine learning algorithms for the classification of knee OA images. Moreover, we are testing more images and we are working to assess other stages of OA (KL1, KL3, and KL4) to provide a reliable classification system.

References

1. Attur, M., Krasnokutsky-Samuels, S., Samuels, J., Abramson, S.B.: Prognostic biomarkers in osteoarthritis. Current Opinion in Rheumatology 25, 136–144 (2013).
2. Harrar, K., Jennane, R.: Quantification of trabecular bone porosity on X-ray images, Journal of Industrial and Intelligent Information 3(4), 280–285 (2015).
3. Harrar, K., Jennane, R.: Trabecular texture analysis using fractal metrics for bone fragility assessment, International Journal of Biomedical and Biological Engineering 9, 683–688 (2015).
4. Kellgren, J., Lawrence.: Radiological assessment of osteo-arthrosis. Annals of the rheumatic diseases 16(4), 494–502 (1957).

5. Bayramoglu, N., Nieminen, M. T., Saarakkala, S.: A lightweight CNN and joint shape-joint space (JS2) descriptor for radiological osteoarthritis detection. In annual conference on medical image understanding and analysis, pp. 331–345. Springer, Cham (2020).

6. Tiulpin, A., Thevenot, J., Rahtu, E., Saarakkala, S.: A novel method for automatic localization of joint area on knee plain radiographs. In : Scandinavian conference on image analysis. Springer, Cham, pp. 290–301 (2017).

7. Haftner, T.S., Ljuhar, R., Dimai, H.P.: Combining radiographic texture parameters increases tibiofemoral osteoarthritis detection accuracy: Data from the osteoarthritis initiative. Osteoarthritis and Cartilage 25, S261 (2017).

8. Akter, M., Jakaite, L.: Extraction of texture features from x-ray images: case of osteoarthritis detection. In: Third International Congress on Information and Communication Technology. Springer, Singapore, pp. 143–150 (2019).

9. Peuna, A., Thevenot, J., Saarakkala, S., Nieminen, M.T., Lammentausta, E.: Machine learning classification on texture analyzed T2 maps of osteoarthritic cartilage: oulu knee osteoarthritis study, Osteoarthritis and Cartilage. Osteoarthritis and Cartilage 29(6), pp. 859–869 (2021).

10. Eckstein, F., Wirth, W., Nevitt, M.: Recent advances in osteoarthritis imaging--the osteoarthritis initiative. Nature Reviews Rheumatology 8(12), 622–30 (2012).

11. Perona,P., Malik,J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(7), 629–639 (1990).

12. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. Conference on Computer Vision & Pattern Recognition 1, pp. 886–893 (2005).

13. Pauly, L., Sankar, D.: Non-Intrusive eye blink detection from low resolution images using HOG-SVM classifier. International Journal of Image, Graphics and Signal Processing 8(10), 2016.

14. Bhende, P., Cheeran, A.: A novel feature extraction scheme for medical X-ray images. International journal of engineering research and applications 6(2), 53–60 (2016).

15. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Recognition and Machine Intelligence 24(7), 971–987 (2002).

16. Al-Sharafat, W., Naoum, R.: Development of genetic-based machine learning for network intrusion detection. World Academy of Science, Engineering and Technology 55, 2009.

The 1st National Workshop on Wireless Network, Cloud Computing and Cryptography

WWN3C'23, April 26, 2023, Boumerdes, Algeria

Faculty of Technology

University M'Hamed Bougara of Boumerdes

WWN3C 2023

Certificate of participation

Is awarded to

Khadidja Messaoudene

For Oral presentation of a paper entitled:

Hybrid classification system using GoogLeNet and support vector regression for the diagnosis of knee osteoarthritis

Authors: Khadidja Messaoudene, Khaled Harrar and Dehia Abdiche

WWN3C'23 General Chair

Pr. Karim BAICHE

الجمهـوريـة الجزائريــة الديمقراطيــة الشعبيــة

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

University M'Hamed Bougara of Boumerdes

كليــة التكنولوجيـا

*F*aculty of *Technology*

1ˢᵗ National Workshop on

The 1st National Workshop on Wireless Network, Cloud Computing and Cryptography (WWN3C'23)

April 26, 2023, Boumerdes

<u>*Preface*</u>

*The M'Hamed Bougara university is very pleased to welcome you in Boumerdes at the first National Workshop on Wireless Network, Cloud Computing and Cryptography (**WWN3C'2023**) held in the Faculty of Technology.*

*The **WWN3C'2023** aims for presenting new advances and research results in the fields of Wireless Networks, Cloud Computing, Cryptography, Telecommunications and Signal Processing.*

The Workshop is expected to provide researchers, lecturers, engineers, and scientists around the world the opportunity to interact and present their latest advanced research.

A wide range of topics is addressed, including Wireless Communications, 5G Communication Networks, Internet of Things, Cloud Computing, Medical Imaging, Pattern Recognition and Machine Learning, Image and Video Processing, Artificial Intelligence and Deep learning.

*We deeply thank the organizing committee and all those who by their help, devotion, competence, and good mood, allowed the good progress of the day of the **WWN3C'2023**.*

We would like to thank the invited speaker for her excellent track record in her own field of research.

We also thank all the participants who insured the animation of the Workshop, and allowed fruitful exchanges and discussions profitable to everybody.

WWN3C'2023

Organizing Committee

Honorary General Chairs

Pr. Mostepha YAHI, Rector of Univ. Boumerdes, Algeria
Pr. Mohamed SAIDI, Dean of Faculty of Technology, Univ. Boumerdes,Algeria

Conference General Chairs

Pr. Dalila ACHELI, Univ. Boumerdes, Algeria
Dr. Karim BAICHE, Univ. Boumerdes, Algeria

Organizing Committee Chair

Dr. Yassine MERAIHI, Univ. Boumerdes, Algeria

Organizing Committee

Miss. Sihem TIRECHE, Univ. Boumerdes, AlgeriaDr.
Youcef TABET, Univ. Boumerdes, Algeria
Dr. Amel BOUSTIL, Univ. Boumerdes, Algeria
Dr. Med Amine RIAHLA, Univ. Boumerdes, Algeria
Dr. Khaled HARRAR, Univ. Boumerdes, Algeria
Dr. Mohammed AMMAR, Univ. Boumerdes, Algeria Dr.
Hamza AKROUM, Univ. Boumerdes, Algeria
Dr. Hadjira BELAIDI, Univ. Boumerdes, Algeria
Miss. Faiza IBAOUNI, Univ. Boumerdes, Algeria
Miss. Selma YAHIA, Univ. Boumerdes, Algeria
Mrs. Sylia MEKHMOUKH, Univ. Boumerdes, AlgeriaMiss.
Amylia AIT SAADI, Univ. Boumerdes, Algeria
Miss. Souad REFAS, Univ. Boumerdes, Algeria
Dr. Abdelkrim AMMAR, Univ. Boumerdes, Algeria
Dr. Fadhila LACHEKHAB, Univ. Boumerdes, Algeria

Scientific Committee

Pr. ACHELI Dalila, Univ. Boumerdes, Algeria
Pr. AZRAR Arab, Univ. Boumerdes, Algeria
Pr. BENTARZI Hamid, Univ. Boumerdes, Algeria
Pr. BOUDEN Toufik, Univ Jijel, Algeria
Pr. BOUDRAH Rabah Univ. Msila, Algeria
Pr. BOUKABOU Abdelkrim, Univ Jijel, Algeria
Pr. BOUKRA Abdelmadjid, Univ. Bab Ezzouar, Algeria
Pr. BOUSHAKI Razika, Univ. Boumerdes, Algeria Pr.
CALLAL Mouloud, Univ. Boumerdes, Algeria
Pr. DAAMOUCHE Abdelhamid, Univ. Boumerdes, Algeria
Pr. FELLAG Sid Ali, Univ. Boumerdes, Algeria
Pr. FERGUENE Farid, Univ. Bab Ezzouar, Algeria Pr.
GHARBI Nawel, Univ. Bab Ezzouar, Algeria
Pr. GROUNI Saïd, Univ. Boumerdes, Algeria
Pr. HAMADOUCHE M'Hamed, Univ. Boumerdes, Algeria

Hybrid classification system using GoogLeNet and support vector regression for the diagnosis of knee osteoarthritis

Khadidja Messaoudene
LIMOSE Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
k.messaoudene@univ-boumerdes.dz

Khaled Harrar
LIST Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
khaled.harrar@univ-boumerdes.dz

Dehia Abdiche
LIST Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
d.abdiche@univ-boumerdes.dz

Abstract—**Osteoarthritis (OA) of the knee is a prevalent and chronic degenerative joint disease, affecting a substantial portion of the global population. Diagnosis of OA can be challenging due to the need for extensive analysis of medical imaging data, such as X-rays and Magnetic Resonance Imaging (MRI). In this study, we present a hybrid system that combines deep feature-based and Machine Learning (ML) approaches for detecting knee OA. The proposed method employs Gabor filter-based preprocessing, data augmentation with translation and rotation, feature extraction using the GoogLeNet model, feature selection via F-Score, and classification using the Support Vector Regression (SVR) model. Our experimental results demonstrate that the proposed approach outperforms existing state-of-the-art methods, achieving an accuracy rate of 83.6%. These findings suggest that the hybrid system has the potential to improve the accuracy and efficiency of diagnosing knee OA.**

Index Terms—**KOA, Gabor filter, GoogLeNet, F-Sore, SVR.**

I. INTRODUCTION

Osteoarthritis of the knee is a degenerative joint disease that is characterized by the loss of cartilage and the deterioration of joint function [1]. It is a major cause of pain, disability, and reduced quality of life for millions of people worldwide. The use of imaging techniques, such as X-rays and Magnetic Resonance Imaging (MRI) [2], has long been a valuable tool in the diagnosis of knee OA. However, the interpretation of these images can be subjective and can vary between radiologists. To address this challenge, researchers have turned to Artificial Intelligence (AI) and Deep Learning (DL) techniques to assist in the diagnosis of knee OA. Machine Learning (ML) techniques have the potential to automate this analysis, enabling faster and more accurate diagnosis of knee OA. However, the performance of such techniques depends heavily on the preprocessing and feature extraction methods used. Preprocessing techniques such as edge detection and image enhancement can improve the quality of the images and facilitate the identification of key features. Features extraction techniques can capture the most informative features in the images, enabling accurate classification of knee OA. Several studies have explored the use of Convolutional Neural Networks (CNNs) for the detection of knee OA from radiographs.

For example, Jakaite et al. [3] used different texture features and ML methods to analyze high-resolution X-ray images. The Haralick features and Zernike moments were used to optimize the performances of the ML techniques, including Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Network (ANN), and the proposed GMDH-type network. Zernike moments with DMDH provided better accuracy (77.5%) than Haralick features. Wang et al. [4] proposed an automated approach for knee OA classification using Deep Neural Network (DNN). The approach involves preprocessing the knee X-ray images using frequency-domain filtering and histogram normalization, and a two-step classification strategy is used to extract the joint center and classify OA grades. The method reported a classification accuracy of 81.41%. The proposed techniques appear effective, but further validation is needed on larger and more diverse datasets before implementing them in the clinical setting. Cueva et al. [5] proposed a semi-automatic CADx model based on Deep Siamese CNN and a fine-tuned ResNet-34 for detecting OA lesions in the two knees according to the KL scale. The model was trained using a public dataset, and the validation was performed using a private dataset. The imbalanced dataset problem was addressed using transfer learning. The model achieved an average multi-class accuracy of 61%, with better performance for classes KL0, KL3, and KL4 than KL1 and KL2. The model results were compared and validated using the classification of experienced radiologists. Lim et al. [6] presented a DL approach to predict the presence of OA in subjects aged 50 years and older using statistical data. The study uses a DNN and Principal Component Analysis (PCA) to automatically generate features from the data and identify risk factors for OA prevalence. They achieved an accuracy of 71.97%.

In this study, we propose a novel ML-based approach for the detection of knee OA that includes several innovative techniques, including Gabor filter-based preprocessing, data augmentation with translation and rotation, feature extraction with GoogLeNet model, feature selection with F-Score, and classification with Support Vector Regression (SVR) machine.

The proposed approach aims to improve the accuracy and interpretability of knee OA detection, making it more suitable for clinical use.

The remainder of this paper is organized as follows: Section II presents materials and methods where the dataset and the methods used are described. In Section III the obtained results are reported and discussed. A conclusion is provided in Section IV.

II. Materials & Methods

A. Dataset

The dataset used in our experiment was obtained from the publicly accessible Osteoarthritis Initiative (OAI) database [7]. The data contains 688 radiographs of the knee that have been categorized using the Kellgren and Lawrence (KL) rating system (KL0, KL2). We compared grade (no OA) to disease overall grade KL2 (mild OA). In our study, we worked with the medial region of the knee radiographic image. Fig. 1 shows the Region Of Interest (ROI) used in our work. Table I demonstrates the data distribution.

TABLE I
Data distribution

Data	Training data	Test data	Validation data
KL0	262	58	24
KL2	262	58	24

Fig. 1. Knee radiographic image

B. Methods

Our approach to classifying KOA involves several techniques, which we will describe in this section. The method we propose consists of four main steps: firstly, we perform artifact removal through preprocessing, followed by deep feature extraction. Then, we apply feature selection using F-Score

and finally classify the results into KL0 and KL2 categories. Fig. 2 provides a visual representation of the flowchart of our proposed approach.

1) Preprocessing: The Gabor descriptor is widely used for textured images and has been shown to be robust to changes in illumination and scale in several works in the literature. This descriptor allows the extraction of microstructure and macrostructure information that describes the texture in an image. Linear Gabor filters are used for various applications such as iris recognition, object detection, and medical imaging [8]–[10].

The Gabor filter is a complex wavelet filter created by multiplying a modulated sine and cosine wave by a two-dimensional Gaussian window. The Gaussian function is used to define the spatial extent and the bandwidth of the filter, while the sinusoidal wave is used to define the frequency and orientation of the filter. By varying the parameters of the Gaussian and sinusoidal functions, Gabor filters can be designed to selectively highlight different features of an image, such as edges, textures, or blobs. The 2D Gabor filter is described mathematically in equations (1) and (2):

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\frac{(\bar{x}^2 + \bar{y}^2)}{\sigma_x^2 + \sigma_y^2} + 2\pi jW\bar{x}\right) \quad (1)$$

$$\bar{x} = x\cos\theta + y\sin\theta \quad \bar{y} = -x\sin\theta + y\cos\theta \quad (2)$$

where : σ_x and y indicate the spread of the current pixel in the neighborhood. W is the center frequency of the complex sinusoid. $\theta \in [0, \pi]$ denotes the orientation of the bands. The number of oscillations and the angle of the Gabor filter are determined by the frequency and orientation. These two main parameters allow a representation of multi-scale and multi-orientation textures, the adjustment of which is given by the following formulas (3) and (4):

$$\frac{dy}{dx} = f(x,y) \quad (3)$$

$$y(x_0) = y_0 \quad (4)$$

In summary, in order to extract the Gabor descriptor, the image is convoluted with a bank of Gabor filters at different frequencies and orientations, and the resulting histograms are used as feature vectors for texture classification. Fig. 3 shows the result of the preprocessing stepp.
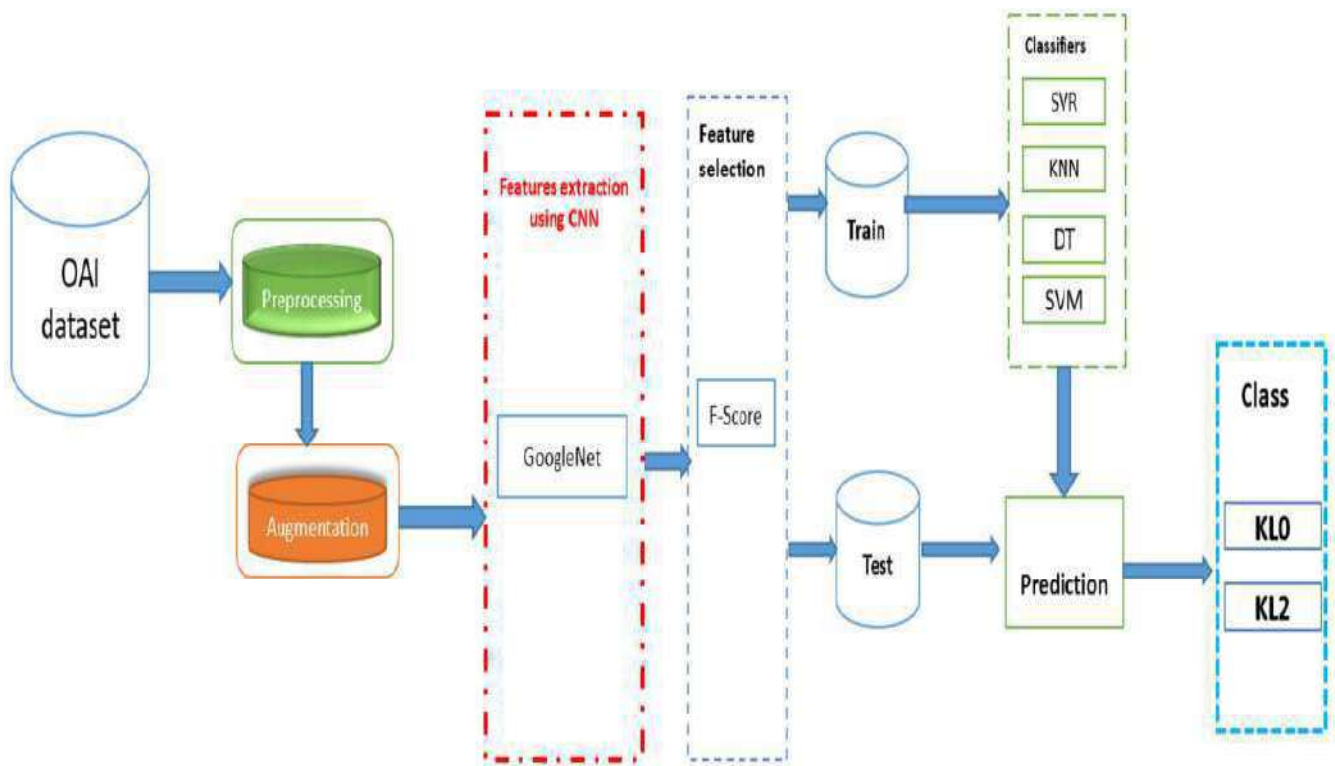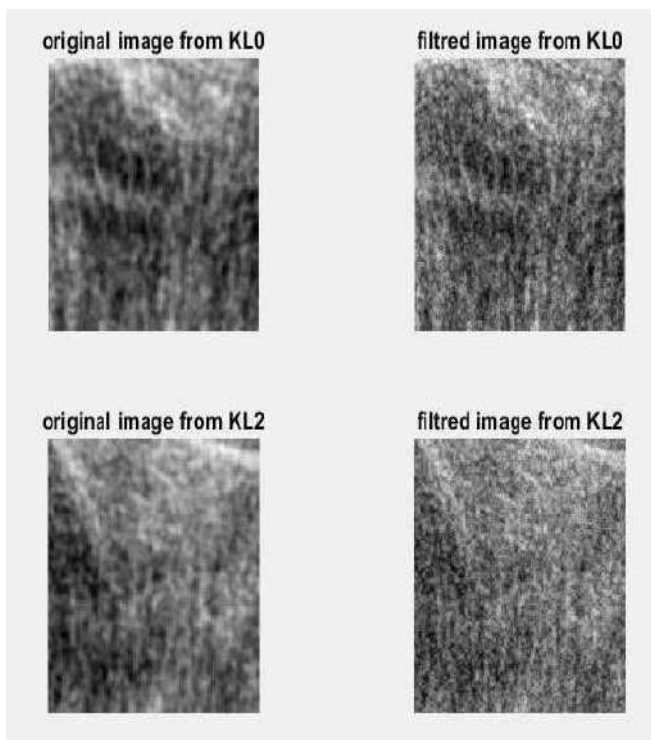
Fig. 2. Proposed method

Fig. 3. The result of preprocessing

2) Data augmentation: To increase the size and diversity of the training dataset, we applied data augmentation techniques, including translation and rotation. The translation is performed by shifting the image horizontally and vertically, while rotation is performed by rotating the image around its center. This approach creates new training examples that are similar to the original images but with slightly different perspectives. This helps to improve the generalization ability of the ML model and reduce overfitting.

3) Features extraction: DL is a subset of ML that uses ANN to solve tasks in various application domains such as object detection, speech recognition, and image classification. Over time, several variants of ANN have been developed, such as Recurrent Neural Network (RNN), Long-Short-Term Memory (LSTM), and CNN. The latter is the subject of our study for a texture classification task in medical images revealing AO.

CNNs use a mathematical operation called convolution to extract relevant and deep features from images, then apply pooling to reduce the dimensionality of these features. These two techniques are organized into a layer system. Different arrangements and settings of these parameters and layers have given rise to several CNN variants. The different CNN variants that have won the ImageNet Large-Scale Visual Recognition

Challenge (ILSVRC) each in a specific year are AlexNet, ResNet, VGG, GoogLeNet, etc. GoogLeNet [11], which we will discuss in more detail, is the CNN variant we used in this article.

GoogLeNet [11], also known as Inception v1, is a deep convolutional neural network architecture developed by Google researchers in 2014. It was designed to improve the accuracy and efficiency of image classification tasks by reducing the number of parameters in the network while increasing its depth [12]. The main innovation of the GoogleNet architecture is the use of a module called the Inception module, which combines filters of different sizes (1x1, 3x3, and 5x5) in parallel to capture different levels of spatial information within an image. By using filters of different sizes, the Inception module can capture both fine-grained and coarse-grained features of an image. Another key feature of the GoogleNet architecture is the use of the auxiliary classifier, which is designed to mitigate the vanishing gradient problem that occurs in deep neural networks. The auxiliary classifier consists of a small classifier attached to an intermediate layer of the network, which encourages the network to learn more robust features. Overall, the GoogleNet architecture achieved state-of-the-art performance on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014, and has since become a popular choice for image classification tasks. Its success has inspired further development of the Inception family of models, including Inception v2, Inception v3, and Inception-ResNet. Fig. 4, illustrates the architecture of the Inception model with several parallel convolutional layers of different filter sizes (1x1, 3x3, 5x5), allowing it to detect features at different scales. In addition, it also uses pooling convolutions to reduce the dimensionality of the data and thus reduce computational costs. The outputs of each branch are concatenated and passed to the next layer [13].
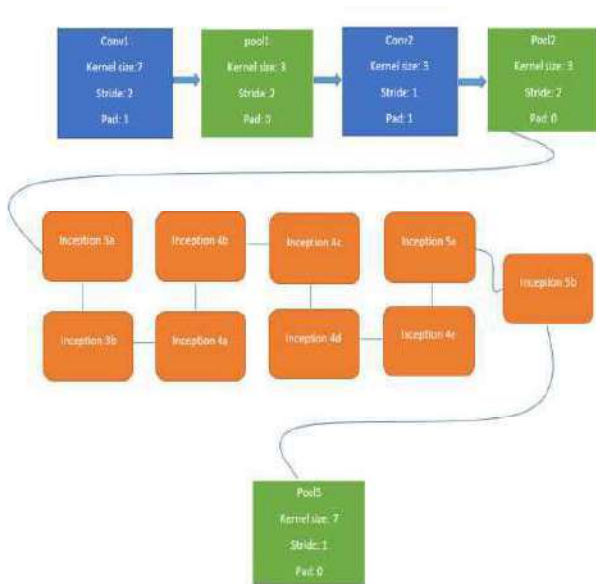
GoogLeNet consists of 22 layers, including 9 Inception modules connected to the global average pooling layer. The final output of the network is a softmax layer, which produces the predicted class probabilities. The overall architecture of GoogLeNet is illustrated in Fig. 5.
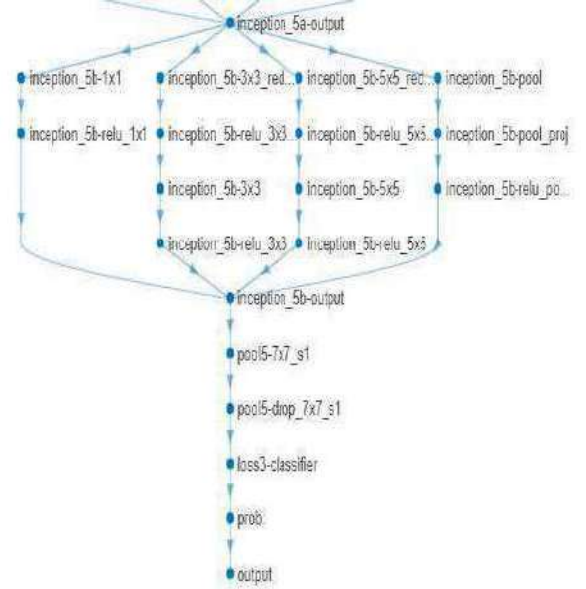

Fig. 5. Last layers of the GoogLeNet before finetuning

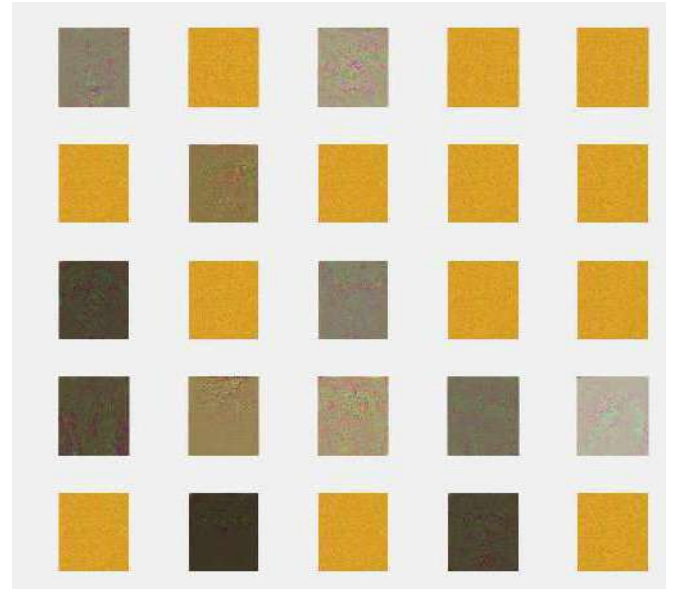Fig. 6 shows the features extracted by GoogLeNet architecture.

Fig. 6. Features of a Convolutional Neural Network

4) Features selection: To identify the most informative features for classification, we used the F-Score feature selection [14]. The F-score is a statistical method for selecting features that evaluate each feature individually and rank them

Fig. 4. A simplified architecture of the CNN used in GoogLeNet

accordingly. The higher the F-score value, the more relevant the feature is considered to be [15]. Equation (5) explains the formula of the F-Score.

$$F\text{-}Score(f_i) = \frac{\sum_j (n_j/(c-1))(\mu_j - \mu)^2}{\frac{1}{n-c}\sum_j (n_j - 1)\sigma_j^2} \qquad (5)$$

x represents the i feature, x denotes the total number of instances belonging to class j, and n represents the total number of instances across all classes [16]. The variable μ refers to the mean value of the features across all classes, while μ_j indicates the mean value of the features in the j class. The standard deviation of class j is denoted by σ.

5) Classification: We used an SVR for the classification of knee OA. SVR is a type of ML algorithm that is used to predict a continuous output variable. It is a variation of the SVM algorithm, which is commonly used to solve classification problems [17].

SVR's main idea is to find a hyperplane in a high-dimensional space with the greatest possible margin to the data points. The hyperplane is used in SVR to make predictions for new data points. SVR works by reducing the difference between predicted and actual values. A loss function is used to calculate the error, which penalizes the algorithm when it makes large errors. The epsilon-insensitive loss function is the most commonly used loss function in SVR [18].

In this study, we tested other classifiers and compared them to the proposed method. This includes decision Tree (DT), Support Vector Machine (SVM), and KNN.

III. RESULTS AND DISCUSSION

The following section presents the key findings from the experiments conducted on high-resolution X-ray images, using various features and ML techniques outlined in the Data section. Specifically, GoogLeNet features were employed, and the experiments were carried out using DT, SVM, KNN, and the SVR model. Fig. 7illustrates the main outcomes. The results indicate that the SVR model utilizing GoogLeNet parameters achieved superior accuracy compared to other classification techniques, including DT, KNN, and SVM.

A confusion matrix is a table that is often used to describe the performance of a classification model. In this case, the rows represent the actual class of the data and the columns represent the predicted class. Fig. 8 shows the confusion matrix. Four different models combining deep learning and machine learning classifiers were used for the classification of knee images: GoogLeNet with SVR, GoogLeNet with KNN, GoogLeNet with DT, and GoogLeNet with SVM. According to the confusion matrix, the GoogLeNet model with the SVR model correctly predicted 44 positive samples (True Positives) and misclassified 5 negative samples as positive (False Positives). On the other hand, it incorrectly predicted 14 positive samples as negative (False Negatives) and correctly predicted 53 negative samples (True Negatives). Overall, this model seems to have performed well with a relatively high true positive rate compared to the false negative rate. GoogLeNet

model with KNN correctly predicted 41 positive samples and misclassified 13 negative samples as positive. It also incorrectly predicted 17 positive samples as negative and correctly predicted 45 negative samples. Based on the confusion matrix, this model has a lower true positive rate and a higher false negative rate compared to the previous model.
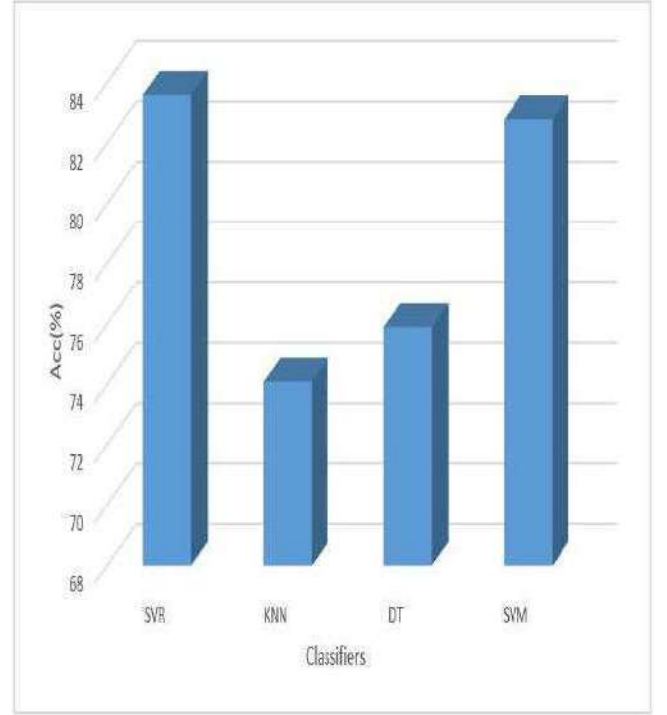

Fig. 7. Comparison of classification performance of the four different classifiers

The GoogLeNet model with DT had similar results to the KNN model. It correctly predicted 42 positive samples and misclassified 12 negative samples as positive. It also incorrectly predicted 16 positive samples as negative and correctly predicted 46 negative samples.

Lastly, the GoogLeNet model with SVM correctly predicted 43 positive samples and misclassified 5 negative samples as positive. It incorrectly predicted 15 positive samples as negative and correctly predicted 53 negative samples. This model appears to have a similar performance to the first model we discussed with a high true positive rate and a relatively low false negative rate.

Overall, it seems that the GoogLeNet model with SVR and the GoogLeNet model with SVM performed the best in terms of correctly predicting positive samples.

The results presented in Table II show the performance of different classification models for knee OA, with and without feature selection using an F-Score. The evaluation metric used is accuracy (Acc), which measures the proportion of correctly classified instances.

The results demonstrate that the feature selection process improves the accuracy of most models. For instance, the
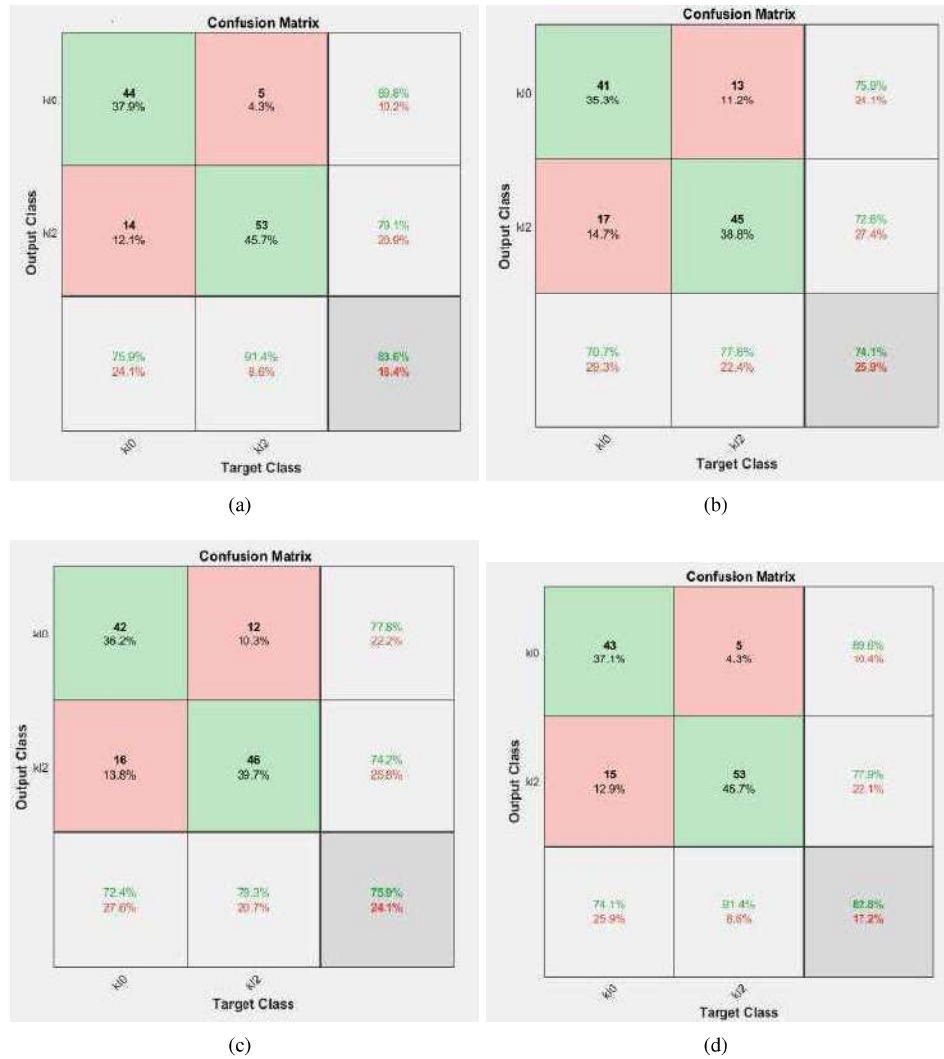
Fig. 8. Confusion matrix for KOA classification: (a) GoogLeNet+SVR, (b) GoogLeNet+KNN,(c) GoogLeNet+DT, and (d) GoogLeNet+SVM

TABLE II

COMPARISON OF CLASSIFICATION RESULTS WITH AND WITHOUT FEATURE SELECTION.

	With features selection	Without selection features
GoogLeNet+SVR	83.6	78.45
GoogLeNet+KNN	74.1	70.5
GoogLeNet+DT	75.9	71.3
GoogLeNet+SVM	82.8	76.65

GoogLeNet+SVR model achieved an accuracy of 83.6% with feature selection, compared to 78.45% without it. Similarly, the GoogLeNet+SVM model achieved an accuracy of 82.8% with feature selection and 76.65% without it.

However, it is worth noting that not all models benefit equally from feature selection. For instance, the GoogLeNet+KNN model only saw a small improvement in accuracy with feature selection (from 70.5% to 74.1%), while the GoogLeNet+DT model saw a more significant improve-ment (from 71.3% to 75.9%). The results suggest that feature selection using an F-Score can be a useful technique for improving the accuracy of classification models for knee OA. However, the effectiveness of feature selection may depend on the specific model used.

Comparing our results with the most recent methods(Table III), our proposed approach outperforms several other ML-based methods for knee OA detection. For example, the method proposed by Lim et al. achieved an accuracy of 71.97% [6], while our approach achieved an accuracy of 83.6%. Similarly, the method proposed by Jakaite et al [3]. achieved an accuracy of 77.5%. These results demonstrate that our proposed approach is highly effective for detecting knee OA, and represents a significant improvement over the existing methods. In conclusion, the proposed approach combines several innovative techniques and achieves high performance for knee OA diagnosis. The high accuracy of the proposed approach indicates that it can be used as an effective tool

TABLE III
COMPARISON OF CLASSIFICATION RESULTS WITH AND WITHOUT FEATURE SELECTION

Authors	Year	Features extraction method	Classifiers	Acc (%)
Lim et al. [6]	2019	DNN with PCA	/	71.97S
Jakaite et al. [3]	2021	Zernike moment	GMDH	77.5
Cuiva et al. [5]	2022	ResNet-34	/	61
Proposed method	2023	GoogLeNet	SVR	83.6

for the early detection and diagnosis of knee OA, which may contribute to improved patient outcomes and reduced healthcare costs.

IV. CONLUSION

The present study proposes a hybrid system that combines deep feature-based methods and Machine Learning techniques for detecting knee osteoarthritis. The proposed approach comprises Gabor filter-based preprocessing, data augmentation, feature extraction utilizing the GoogLeNet model, feature selection via F-Score, and classification using SVR. Empirical findings reveal that the proposed approach surpasses other methods in detecting knee OA. This research presents a promising solution to the challenging task of diagnosing knee OA through the analysis of medical imaging data. Furthermore, it highlights the potential of ML techniques in the field of healthcare to enhance diagnosis and treatment outcomes. Further investigation may explore the feasibility of applying this approach to other stages of OA and imaging modalities, as well as its integration into clinical practice for personalized medicine.

REFERENCES

[1] A. Cui, H. Li, D.Wang, J. Zhang, Y. Chen, H. Liu, "Global, regional prevalence, incidence, and risk factors of knee osteoarthritis in population-based studies", Clin Med, vol. 29, 100587, 2020.

[2] M. Chalian, F. Roemer, A. Guermazi, "Advances in osteoarthritis imaging", Curr. Opin. Rheumatol., vol. 35, no. 1, pp. 44-54, 2023.

[3] L. Jakaite, V. Schetinin, J. Hladůvka, S. Minaev, A. Ambia, W. Krzanowski, "Deep learning for early detection of pathological changes in X-ray bone microstructures: case of osteoarthritis", Sci. Rep., vol. 11, no. 1, pp. 1-9, 2021.

[4] Y. Wang, X. Wang, T. Gao, L. Du, W. Liu, "An Automatic Knee Osteoarthritis Diagnosis Method Based on Deep Learning: Data from the Osteoarthritis Initiative", J. Healthc Eng., pp. 1-10, 2021.

[5] J.H. Cueva, D. Castillo, H. Espinós-Morató, D. Durán, P. Díaz, V. Lakshminarayanan, "Detection and Classification of Knee Osteoarthritis", Diagnostics (Basel), vol. 12, no. 10, pp.2362, 2022.

[6] J. Lim, J. Kim, S. Cheon, "A Deep Neural Network-Based Method for Early Detection of Osteoarthritis Using Statistical Data", Int. J. Environ. Res. Public Health, vol. 16, no. 7, pp.1281, 2019.

[7] G. Lester, "The osteoarthritis initiative: an NIH public-private partnership", HSSJ, vol. 8, pp. 62-3, 2012.

[8] J. Han, K. Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval", Image Vis. Comput., vol. 25, pp. 1474–1481, 2007.

[9] B.S. Manjunath, W.Y. Ma, "Texture features for browsing and retrieval of image data", IEEE Trans. Pattern Anal. Mach. Intell. vol. 18, no. 8, pp. 837–842, 1996.

[10] H. Hadizadeh, "Multi-resolution local Gabor wavelets binary patterns for gray-scale texture description", Pattern Recognit. Lett., vol. 65, pp. 163–169, 2015.

[11] S. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions", In Proceedings of the IEEE conference on Computer Vision and pattern recognition 2015, pp. 1-9.

[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going Deeper with Convolutions", In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015, pp. 1–9.

[13] C. Tataru, D. Yi, A. Shenoyas, A. Ma, "Deep learning for abnormality detection in chest x-ray images", In IEEE Conference on deep learning, 2017.

[14] B. Manavalan, S. Basith, T.H. Shin, L. Wei, G. Lee, mAHTPred, "A sequence-based meta-predictor for improving the prediction of antihypertensive peptides using effective feature representation", Bioinform., vol. 35, no. 16, pp. 2757-2765, 2019.

[15] Y.W. Chen, C.J. Lin, "Combining SVMs with various feature selection strategies", NIPS 2003 feature selection challenge, pp. 1-10, 2003.

[16] L. Wei, S. Luan, L.A.E. Nagai, R. Su, Q. Zou, "Exploring sequence-based features for the improved prediction of DNA N4-methylcytosine sites in multiple species", BioinforM., vol. 35, pp. 1326–1333, 2019.

[17] A.J. Smola, B. Schölkopf, "A tutorial on support vector regression", Stat Comput, vol. 14, pp. 199–222, 2004.

[18] F. Zhang, L.J. O'Donnell, "Support vector regression", Machine learning, pp. 123-140, 2020.