# 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.

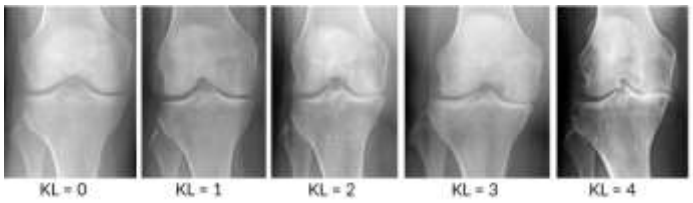


**Fig. 1.** Knee OA severity [5]

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.

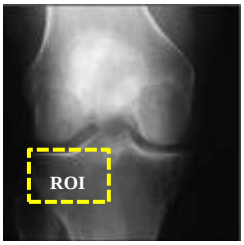


**Fig. 2.** Knee radiographic image.

### 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.

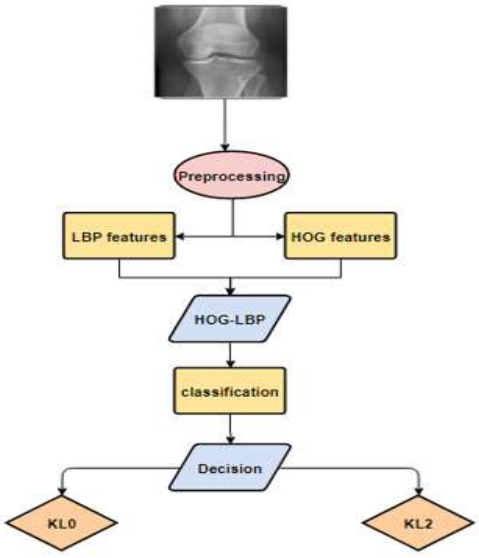


**Fig. 3.** Proposed classification system

**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:

$$\frac{\partial I}{\partial t} = div(c(x, y, t)\nabla I) = \nabla c.\nabla I + c(x, y, t)\Delta I \tag{1}$$

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

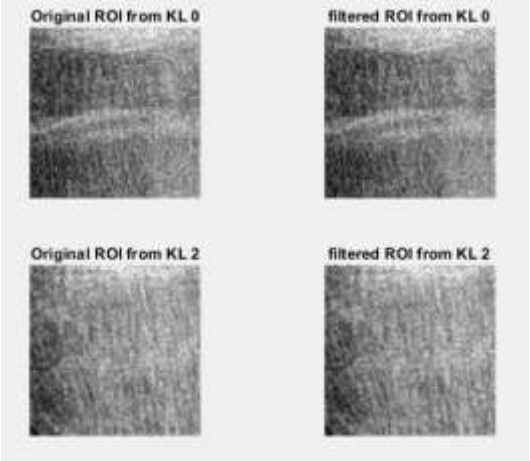


**Fig. 4.** Results of filtering

**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:

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

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

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.

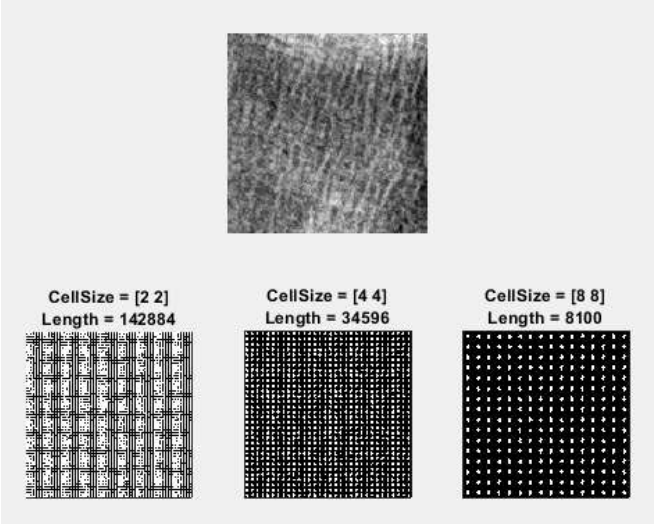


**Fig. 5.** HOG features of an X-ray image with different cell sizes

**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.

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

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 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Naive 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.