SN Computer Science (2023) 4:58
https://doi.org/10.1007/s42979-022-01476-4

ORIGINAL RESEARCH

DNN-Based Knee OA Severity Prediction System: Pathologically Robust Feature Engineering Approach

Darshan Ruikar[1] · Parshuram Kamble[2] · Amruta Ruikar[3] · Kavita Houde[2] · Ravindra Hegadi[2]

Received: 18 April 2022 / Accepted: 24 October 2022 / Published online: 17 November 2022

Abstract

Osteoarthritis (OA) is a common chronic health condition of bone joints. It gradually leads severe pain and disability. Over the last 2 decades, the effect of OA is noticed more in the adult population, as well. People suffering from knee OA struggle a lot to perform daytoday activities, such as walking, climbing, and sitting, etc. The best available treatment to date tries to reduce pain but does not help to cure the chronic condition. The only possible solution is early detection and treatment of OA. There is great scope to develop an automated computer-aided diagnosis (CAD) system which will assist experts to predict the exact knee OA severity grade by analyzing X-ray images. The prediction accuracy of the CAD system is majorly dependent upon feature engineering. This means that the extracted features must truly represent the pathological features that are observed and analyzed by the experts during the diagnosis and prognosis. Hence, it is required that DNN models learn the features accurately. However, due to data scarcity and model design issues, DNN models are not able to extract better representative features. In this research attempt, we propose to combine engineered features (having high similarity with pathological features) into the features learned by the DNN model. We then perform classification on the combined feature space. The proposed system uses NIH-OAI radiographic dataset for the experimentation. The severity prediction follows K &L grades and yields 92.7% accuracy.

Keywords Osteoarthritis · Knee X-ray images · NIH-OAI · K& L grading system · Pathological features · Eigenvalue · Envelope · Deep neural network · Classification · CAD

Introduction

Osteoarthritis (OA) is one of the most common incurable bone joint diseases with devastating consequences for affected individuals and public healthcare. Nowadays, in addition to older people, surprisingly younger audiences are predominantly suffering from knee OA [1]. Lack of exercise, busy daily routine, lifestyle, and wrong food habits are the major reason behind this scenario [2].

In addition to symptoms and medical history, accurate diagnosis and prognosis depend on the appropriate clinical examination process. The examination process starts with the physical examination which includes, but is not limited to observation of joint swelling, difficulty during knee movement and walking, any signs of injury to knee joints or surrounding muscles, and the pain when weight is placed on the knee [3]. In some cases, experts may ask to perform diagnostic tests such as blood tests and urine tests. These tests will be helpful to identify the level of uric acid and other related contents that cause OA. The increase in uric acid is the primary indication of OA. The physical examination and blood tests are helpful to extract primary information about the disease; however, they will not help to identify the exact severity of pains and loss in articular cartilage and bones.

To understand the severity and to do accurate diagnosis, experts make use of various medical imaging modalities like magnetic resonance imaging (MRI) scan, computed tomography (CT) scan, and X-ray scan. These imaging techniques provide detailed information about

D. Ruikar, P. Kamble A. Ruikar and R. Hegadi have contributed equally to this work.

This article is part of the topical collection "Advances in Applied Image Processing and Pattern Recognition" guest edited by K. C. Santosh.

✉ Darshan Ruikar
darshanruikar1986@gmail.com

Extended author information available on the last page of the article

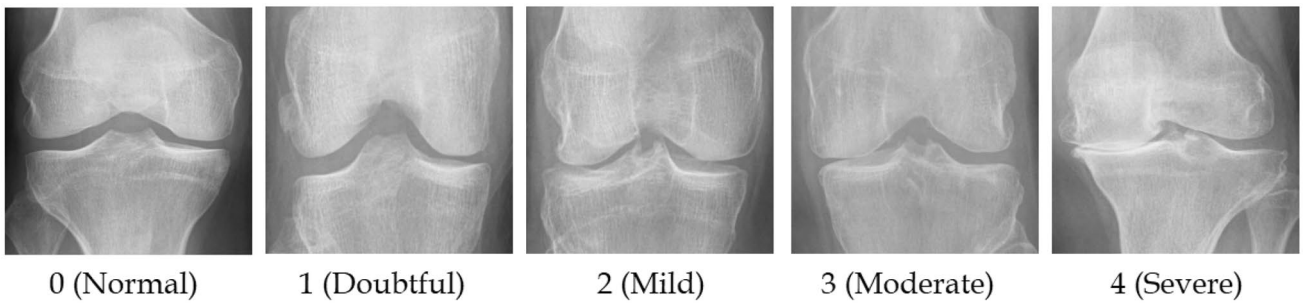

Fig. 1 Sample knee X-ray images representing K &L severity grades

internal body structures like bone anatomy and muscular structure without cutting any specific part of the body [4]. MRI and CT scans provide detailed information regarding bones, muscles, and tissues in three dimensions. Even though MRI and CT scans are highly efficient, they are expensive, slow, and uneasy. In addition to this, patients need to suffer from heavy magnetic radiation and ionizing doses, respectively [5].

Contrary to this, the X-ray images are extensively available, low in cost, less time-consuming, and speedy at providing results. Two-dimensional X-ray images provide all necessary information like bone structure, changes in bone morphology, and distance between bones in detail. In addition to this, the X-ray imaging modality is broadly used not only because it is accessible and moderately economical, but also because it functions as an effective approach in OA diagnosis and surgical procedures. The traditional X-ray images continue to be the "gold standard" for the identification of OA [6]. Though the X-ray images provide in-depth information about knee OA, accurate reading and analysis of X-ray images to decide the exact grade of knee OA is a time-consuming and laborious job. Severity grade may vary from person to person. A lot of experience and practice are required for exact grading and accurate diagnosis. To help experts, in this paper, we proposed a DNN-based knee OA severity prediction system. To improve the overall classification accuracy, hand-crafted features are extracted that truly represent the pathological features in the real sense and combined in feature space.

The remainder of the paper is organized as follows. The section "K &L Grades" briefly explains the K &L knee OA severity grades and the real pathological features that are observed by medical experts. The section "Literature Review" explores the related work. The section "Proposed Methodology" describes the proposed methodology in detail. Experimental results are demonstrated in the section "Experimental Results". The section "Conclusion" provides the conclusion.

K &L Grades

In 1958, Kellgren and Lawrence proposed a Knee OA severity grading scheme based on X-ray images [7]. Now also, it is widely accepted stranded to diagnose and identify exact knee OA severity from X-ray images. Figure 1 shows the sample knee X-ray images of each grade.

The severity of knee OA is based on the combination of four pathological features: joint space narrowing (JSN), osteophytes (i.e., bone spur), sclerosis, and bone deformity generally at the bone end. Table 1 shows the descriptive values of the above-mentioned features for each of the severity grades.

Literature Review

The design and development of automatic knee OA severity prediction can be considered an image classification problem. Patient-specific data collection, region of interest (ROI) extraction, feature engineering, and severity classification are the general steps to be performed to develop a CAD system for automatic knee OA severity prediction.

Well-organized problem-specific dataset is the primary requirement to build a robust machine, learning-based model. CAD system development for knee OA severity classification is also no exception to this analogy. As far as the medical or healthcare domain is concerned hospital and radiology centers are well equipped with patient-specific

Table 1 K&L grading systems for knee OA severity assessment based on X-ray images

Features grades	0	1	2	3	4
JSN	No	Doubtful	Moderate	Definite	Severe
Osteophytes	No	Small	Definite	Multiple	Large
Sclerosis	No	No	No	Definite	Severe
Deformity	No	No	No	Small	Definite

data. However, due to ethical laws like HIPPA[1] and patient-specific data sharing protocols like IRB,[2] data cannot be made data publically available [8]. As a result, most of the existing research is conducted on local datasets. The unfortunate thing is that data available for further experiment and result comparison will be limited. To overcome this constraint NIH–NIA[3] developed a public domain dataset named OAI.[4] Nowadays, most researchers are using the same dataset to conduct the experiments.

After data collection, the next crucial step is the RoI extraction. Before RoI localization, some pre-processing techniques may be applied to remove noise and unwanted artifacts from the input X-ray image moreover to enhance the desired portion (knee joint region) significantly [9]. Various noise removal and image pre-processing methods are applied to reduce the noise and enhance the image quality. Median filter [10], Wiener filter [11], and guided filter [12] are used most of the times to remove noise from input X-ray images. After noise removal, different image enhancement methods are applied to further enhance the quality of desired image. Contrast limited adaptive histogram equalization (CLAHE) methods are used by most researchers to enhance input X-ray images [12].

RoI localization methods aim to extract the knee portion from the input X-ray image. RoI extraction methods can be categorized into three types: manual, semi-automatic, and automatic methods. In manual methods, the ROI localization process is carried out by an expert [13] and can be used to create templates (ground truth) [14]. Semiautomatic methods expect least user interventions [15] either at the beginning to start the segmentation process or at the end to verify output obtained by computerized methods. Generally, these methods expect threshold value, seed point, or initial bounding box from the user to start the segmentation process. Level set-based [16] and active contour-based [17] interactive segmentation methods are used to localize ROI from input X-ray image. Both methods accept an initial bounding box from the user. The third type of ROI localization is automatic methods. These methods do not expect any sort of user intervention to localize ROI. Block-based [18], region-proposal-based [19], neural network-based [20], and fully convolution neural network (CNN)-based methods were largely used in the literature to localize ROI (i.e., knee joint region) automatically.

The overall classification accuracy of the classification system mostly depends on feature engineering. Most of the research attempts extracted hand-crafted features to classify input X-ray images to correct class according to the severity level. Statistical features [18, 21], DWT-based features [22], edge curvature features [17], and textural GLCM (gray-level co-occurrence matrix) features [11] were extracted to predict exact knee OA severity grade. Hand-crafted feature selection and extraction are time-consuming and laborious jobs. To overcome this limitation, nowadays, researchers are using deep features to perform classification. Deep feature learning methods themselves learn and extract effective features from row input data. In recent research, used deep features were to improve classification accuracy. Mostly these techniques are based on the convolution process. That is, a specific mask (filter) is convoluted on the entire image to exact features. Initial few layers extract raw (coarse), whereas the last few layers extract finer features which can truly represent the correct class of input image [20].

In the last step (i.e., knee OA severity prediction), different classifiers are applied to identify correct class of input image. Support vector machine (SVM) [18], maximum likelywood approach [23], feed forward neural network [17, 19], deep convolution neural networks (CNN) [24], and self-organizing map [25]-based classification methods are adapted to predict severity grade of input X-ray image. The methods proposed in [18] demonstrated 80% and 86.7% accuracy for the healthy and affected images, respectively. The method proposed in [21] shows 92% classification accuracy for three-class classification, whereas methods discussed in [6] demonstrated 66.71% accuracy for five-class classification.

By confirming the literature, most of the research attempts to categorize the images into two classes: normal and OA affected. Only a few research attempts are present which results in five-class categorization. However, the classification accuracy is not promising. Therefore, in the near future, several successful research attempts must be made to increase the classification results. Though DNN-based methods produce promising results, their classification accuracy is limited in knee OA severity prediction. Lack of well-classified data and data imbalance are the major reasons behind this. In near future, more research must be made to embed hand-crafted features (which will represent real pathological features) with deep features to improve classification accuracy.

Proposed Methodology

This section provides a detailed description of the proposed methodology devised to develop a CAD system for knee OA classification from knee X-ray images. The proposed CAD system works in stages. The first stage is data acquisition. The second stage is RoI extraction. This stage is responsible

[1] HIPPA: Health Insurance Portability and Accountability Act.

[2] IRB: Institutional Review Board.

[3] NIH–NIA: National Institute of Health–National Institute on Aging.

[4] OAI: osteoarthritis initiative.

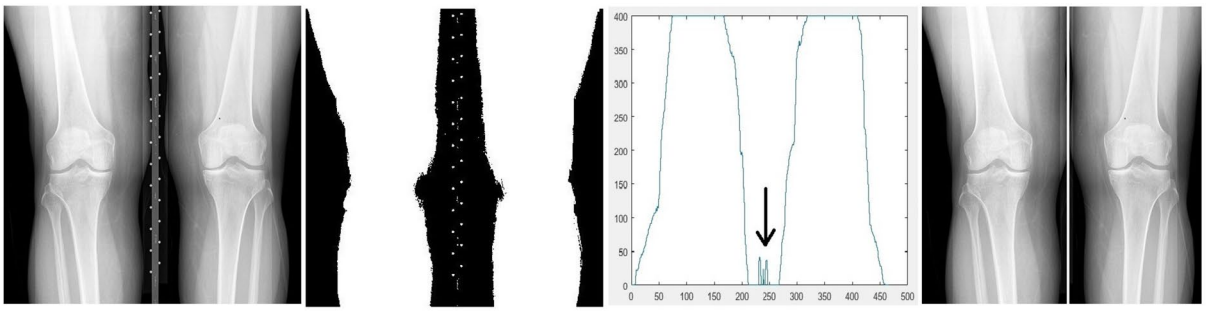

Fig. 2 Left and right knee separation process

for the accurate segmentation of the knee joint region from the X-ray image. After RoI extraction, the contrast stretching-based pre-processing technique is applied to enhance the desired region nicely. The next step is feature engineering. In this step, hand-crafted features are extracted which will truly represent the desired pathological features. In the last stage, engineered features are integrated into the DNN-based classifier to improve knee OA severity prediction accuracy.

Data Preparation

Dataset

To promote the research in knee OA severity prediction, a knee X-ray image database named Osteoarthritis Initiative (OAI) is made available by NIH–NIA. The baseline cohort of the OAI database contains MRI and X-ray images of 4746 participants. In the proposed method, all 4746 X-ray images are used for experimentation.

Left and Right Knee Separation

Each X-ray image in the OAI database has both left as well as right knees. Sample X-ray image is shown in Fig. 2a. Before further processing, the input image must be separated into two sub-images, that is, one image having a left knee and the other having a right knee. To achieve this, initially, the input gray-scale image is converted to a binary image (Fig. 2b). The image-specific threshold value is obtained by applying Otsu's method. Otsu's method chooses the threshold value to minimize the intra-class variance of the threshold black and white pixels [26]. Later, column sum is calculated for each column. The histogram is plotted for the calculated sum (Fig. 2c). By observing the histogram, it is clear that there is a valley near the middle of the image. The lowest point of the valley represents the boundary between the left knee and right knee. The separated images are shown in Fig. 2d. After the left and right knee separation, further, these images are segmented to extract the Knee joint region only.

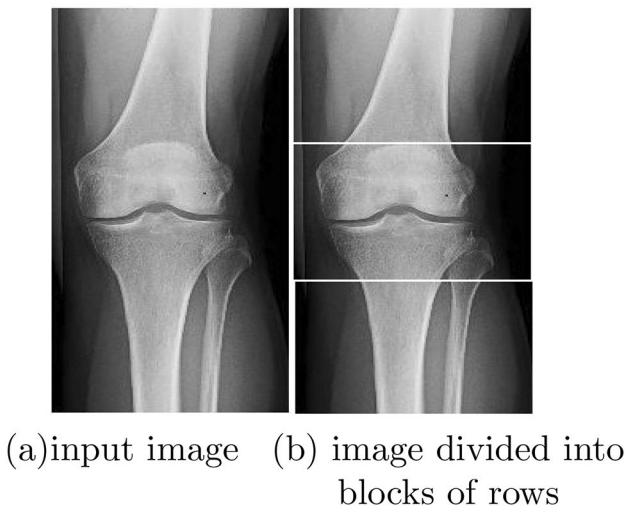

Fig. 3 ROI localization process

Knee Joint Extraction

To segment the knee joint region from an X-ray image, a simple block-based segmentation method is used. In this method, input image (Fig. 3a) is divided into three equal rows (Fig. 3b) and only middle part is considered for the further experimentation.

Further, these segmented images are segregated into five classes based on severity level. This segregation is done by consulting expert radiographers and orthopedic physicians. Normal, minor, mild, moderate, and severe are the labels of the five classes.

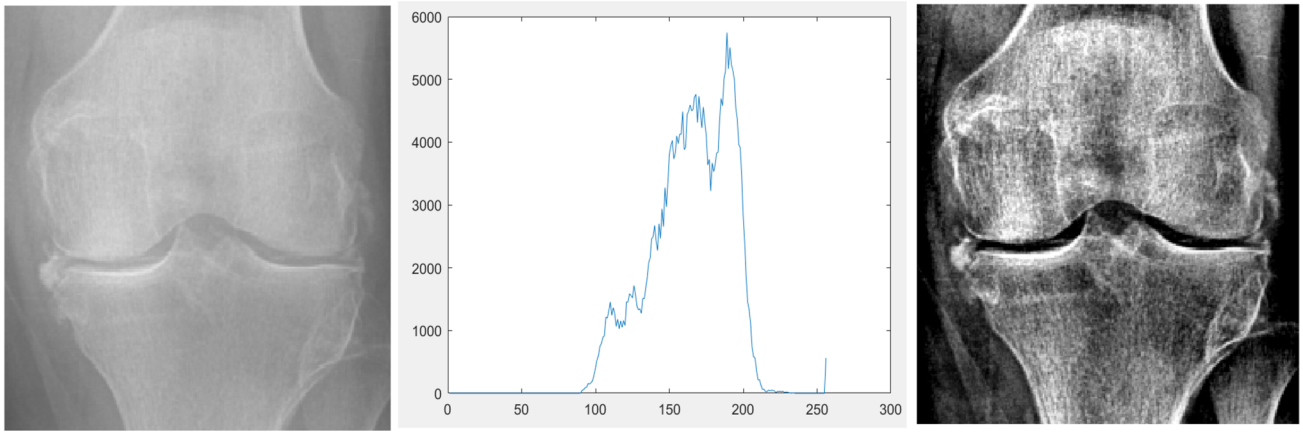

Fig. 4 Result of contrast stretching pre-processing method

Image Enhancement

Histogram of input X-ray image is plotted, to gain more insights and contrast-related information. After observing histogram beans, it is observed that the input images are low-contrast images. Maximum pixels fall under small intensity range (100–190) (Fig. 4b). To improve the dynamic range, contrast stretching-based image enhancement technique is developed. It aims to uplift the contrast of an image by extending the input intensity range to the desired range [27, 28]. Equation 1 is used to perform contrast stretching

$$I_{\text{out}} = (b-a)/(d-c) \times (I_{\text{in}} - c) + a, \tag{1}$$

where I_{in} and I_{out} are input and resultant images, respectively. The variable a and b represents the low and high values of output intensity range, respectively, whereas the variable c and d represent the low and high values of input intensity range, respectively. Equation $(b-a)/(d-c)$ is the slope of straight line. For 8 bit gray-scale input X-ray image, the values of a and b are set 0 and 255, respectively, to cover entire dynamic range. The values of c and d are set 102 and 229, respectively, by conducting experiments empirically. The simplified Eq. 2 is obtained after putting the values of a, b, c, d in Eq. 1. Equation 2 linearly scales the intensity values that results less harsh enhancement than histogram equalization. Figure 4a and c shows input and enhanced X-ray image, respectively. In the enhanced image, all the pathological feature are highlighted significantly

$$I_{\text{out}} = (I_{\text{in}} - 102) \times 2. \tag{2}$$

Hand-Crafted Feature Engineering

The primary aim of hand-crafted feature engineering is to extract features having high similarity pathological features that are observed and analyzed by experts. Joint space narrowing, bone spur, sclerosis, and deformation are the main features considered by the expert to predict the severity grade. To represent these feature values, hand-crafted feature engineering is performed. Histogram modeling-based statistical features and eigenvalues of upper and lower envelopes were extracted.

Histogram Modeling-Based Statistical Features

To represent the joint space narrowing (JSN) histogram, modeling-based statistical features, such as quartiles, were calculated. At a first column-wise number of black pixels in the knee joint region were identified. Then, histogram of column-wise pixels is plotted (Fig. 5) and a lower quartile ($Q1$), median ($Q2$), and upper quartile ($Q3$) are identified.

To represent the observation from the histogram in the numerical format, the following formula is used:

$$D = \|(Q2 - Q1) - (Q3 - Q2)\|, \tag{3}$$

where $Q1$, $Q2$, $Q3$ are the values of a lower quartile, median, and upper quartile. The interesting property of quartiles is if $(Q2 - Q1) = (Q3 - Q2)$, then the all the quartiles are equidistant. This means that all the histogram beans have nearly equal values. For the normal knee, the joint value of the D is lesser (near to 0) than the value for the knee joint with severe OA. In the normal knee, the joint space between the femur and tibia bone is more. In terms of histogram at each

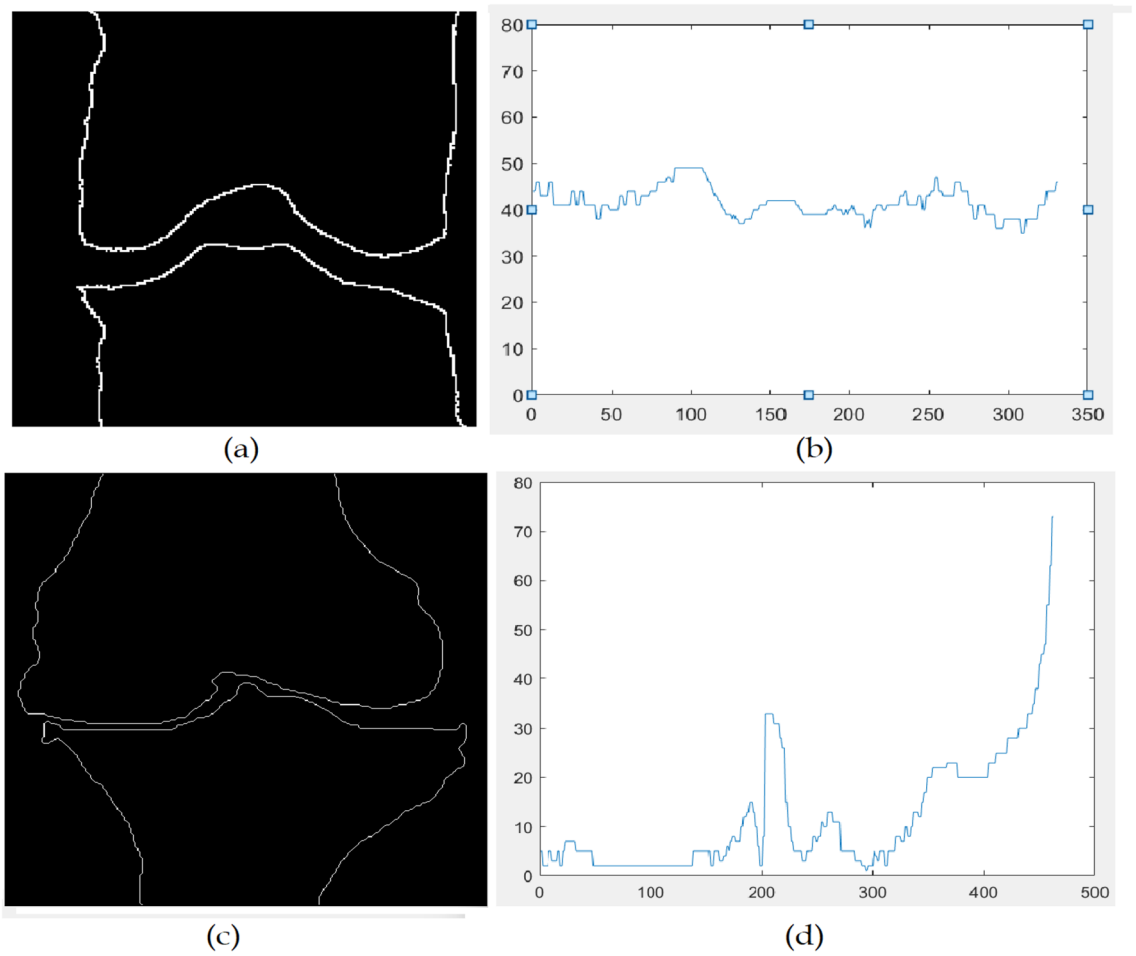

Fig. 5 Histogram modeling-based statistical features

place, there will similar number of bins everywhere, whereas in the knee joint which is affected by OA, the histogram has lesser beans mostly either at the left or right end depending on severity level which represent the reduction in joint space. Figure 5b shows the number of black pixels in knee joints ranging from 40 to 50 where, in severely affected in knee, there is significant variation in the number of black pixels. Figure 5d shows the variation in number of pixels, i.e., less number of pixels are left part and more number of pixels at right part. In addition to this, for most of the region, a number of black pixels are ranging from 0 to 10 which is quite less than the number of pixels in normal knee joint. The values of $Q1$, $Q2$, $Q3$ will be much same in normal one that results lesser value of D where, in severely affected knee, there will much variation in the values that result in greater value of D.

Eigen Values of Upper and Lower Envelope

To represent the bone spur and bone deformation, small and large eigenvalues are computed for the upper and lower envelope.

- Upper and Lower Envelope:
  The upper and lower envelope are curve connecting the uppermost and lowermost pixels of the synovial cavity (gap between femur and tibia bone). To extract the upper and lower envelop, first bone boundaries are extracted using the canny edge detection operator. To extract the upper envelope, each column of the image is traversed from top to bottom. The location of the first white pixel in each column came across is marked as a point of the upper envelope. In the same way for a lower envelope, each column of the image is traversed from bottom to top

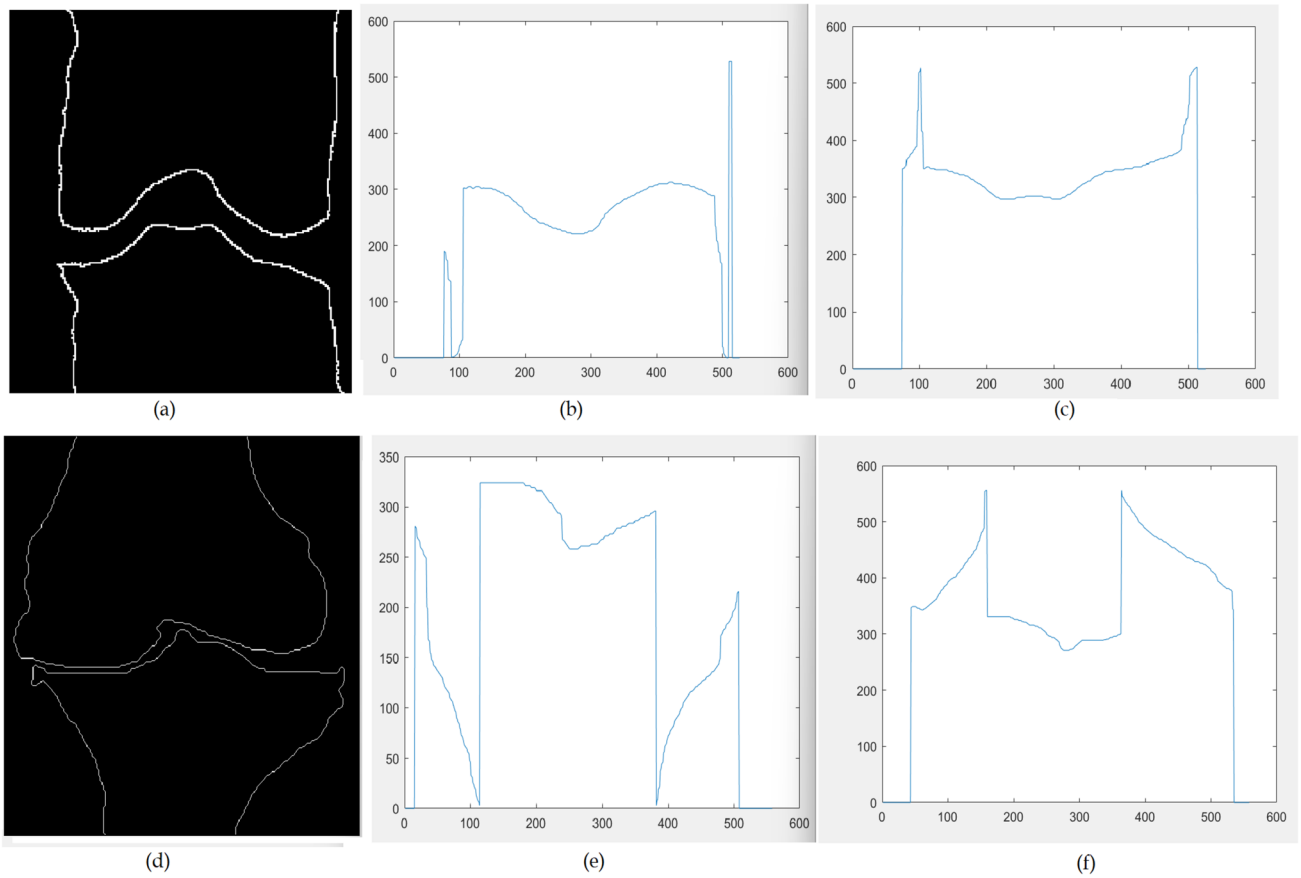

Fig. 6 Upper and lower envelop of normal and severely affected knee joint

recording the first white pixel of the lower envelope [29]. First row in Fig. 6 shows (a) the femur and tibia bone boundaries, (b) upper envelope of the femur bone, and (c) lower envelope of tibia bone for normal knee joint. Whereas the second row in Fig. 6 shows the same things for severely affected knee joint. It is observed that there is less variation in both the envelopes for normal knee joint than the severely affected knee joint.

Upper and lower envelope images were fused by applying union operation among these images and covariance is computed on the resultant image. From the upper and lower envelopes, high and low points were obtained, and further, their difference and ratio are computed.

- Eigen Values:

In real-world dynamic problems, Eigenvalues have great importance and are used in many applications [30]. The basic mathematical equation for computation of Eigen Values is expressed as

$$AX = \lambda X, \tag{4}$$

where A is an $n \times n$ matrix, X is a non-zero $n \times 1$ vector, and λ is a Scalar value (can be real or complex), which is called Eigen Value of $n \times n$ matrix A. There exists a nontrivial solution X of $AX = \lambda X$; such an X is called as Eigenvector corresponding to the Eigenvalue λ. The above Eq. 4 can be rewritten as

$$A \times X - \lambda \times X = 0, \tag{5}$$

$$A \times X - \lambda \times I \times X = 0, \tag{6}$$

$$(A - \lambda \times I) \times X = 0. \tag{7}$$

Here, I is an identity matrix. If X is non-zero, then the equation becomes

$$\det(A - \lambda \times I) = 0. \tag{8}$$

This is called the characteristic equation of A. For each Eigenvalue, there will be an Eigenvector for which the Eigenvalue equation is true.

For calculating large and small Eigenvalues of both the envelope, at first covariance of both envelope were calculated, and then, large and small Eigen values were calculated using the following formulae:

Table 2 OACnet (osteoarthritis network) architecture

Layer		Feature map	Output size	Kernel size	Stride
Input	Image	1	112×112	–	–
1	$2 \times$ Convolution	128	$112 \times 112 \times 128$	3×3	1
	Max Pooling	128	$56 \times 56 \times 128$	3×3	2
3	$3 \times$ Convolution	256	$56 \times 56 \times 256$	3×3	1
	Max Pooling	256	$28 \times 28 \times 256$	3×3	2
5	$3 \times$ Convolution	512	$28 \times 28 \times 512$	3×3	1
	Max Pooling	512	$14 \times 14 \times 512$	3×3	2
7	$2 \times$ Convolution	512	$14 \times 14 \times 512$	3×3	1
	Max Pooling	512	$7 \times 7 \times 512$	3×3	2
9	FC	–	4096	–	
Output	SoftMax	–	5	–	

$$\lambda_L = \frac{1}{2}\left[C_{11} + C_{22} + \sqrt{(C_{11} - C_{22})^2 + 4C_{12}^2}\right], \tag{9}$$

$$\lambda_S = \frac{1}{2}\left[C_{11} + C_{22} - \sqrt{(C_{11} - C_{22})^2 + 4C_{12}^2}\right], \tag{10}$$

where $\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$ is the covariance matrix of λ_L and λ_S. Later, we computed union of large and small Eigen values of upper and lower envelopes.

Figure 6 shows less variation in an upper and lower envelope for a normal image, whereas high variation for a severe one. The same scenario is reflected in feature values in terms of the values of small and large eigenvalues which help the DNN model to learn bone spur and deformation pathological features.

Classification

After problem-specific dataset preparation, the next step is classification. The main aim of this work is to develop a DNN-based classification system with the integration of hand-crafted feature engineering to improve overall classification accuracy.

DNN-Based Classification System

In this work, we developed our DNN-based architecture OACnet (Osteoarthritis network), especially for knee OA severity classification problems. The OACnet is more lightweight than other DNN-based counterpart architecture available in the literature. The complexity of the proposed architecture is reduced by limiting the number of convolution layers followed by the fully connected layers at the end. There are 11,75,08,864 parameters in OACnet. In addition to this, our architecture expects an input size of 112×112 pixels, whereas the rest of the counterpart architectures expect an input size much higher than 112×112. The main reason behind limiting the layers is that the other DNN-based architectures considered here suffer from overfitting due to bigger architecture. Such bigger architectures are not needed at least for knee OA severity classification problems.

The proposed model comprises four blocks of convolution layers followed by max-pooling layers. The activation maps produced by the seventh convolution layer consist of promising features with finer details that are supplied to a fully connected convolution layer having 4096 neurons. At last output of fully connected layers is fed to the softmax layer for five-class severity classification. Table 2 describes the architecture of OACnet in detail.

The overall accuracy (average of per-class accuracy) of the OACnet is 83.74%. The classification accuracy of OACnet is not promising. The primary reason behind this is data imbalance, i.e., number of images per class is not the same. The dataset has more images of class 0 than the remaining classes. In addition to this, the activation map of some intermediate layers was visualized. It is observed that the minute details, such as the amount of joint space, bone deformation, abnormal bone growth, and smoothness of bone boundaries, which improve the classification accuracy were not extracted. To improve the classification accuracy, we integrate engineered features into the model.

DNN Model with the Integration of Hand-Crafted Features

As discussed in "Hand-Crafted Feature Engineering", histogram modeling-based and eigenvalues of envelope values are extracted, which will truly represent the pathological features. Quartile values and its difference based on the histogram of joint space distance were extracted to represent JSN pathological feature. Small and large eigenvalues of upper and lower envelopes, and difference and ratio of the high and low points of the envelopes were considered. This set of features correctly represents bone spur and deformity in the real scene. The hand-crafted features are integrated

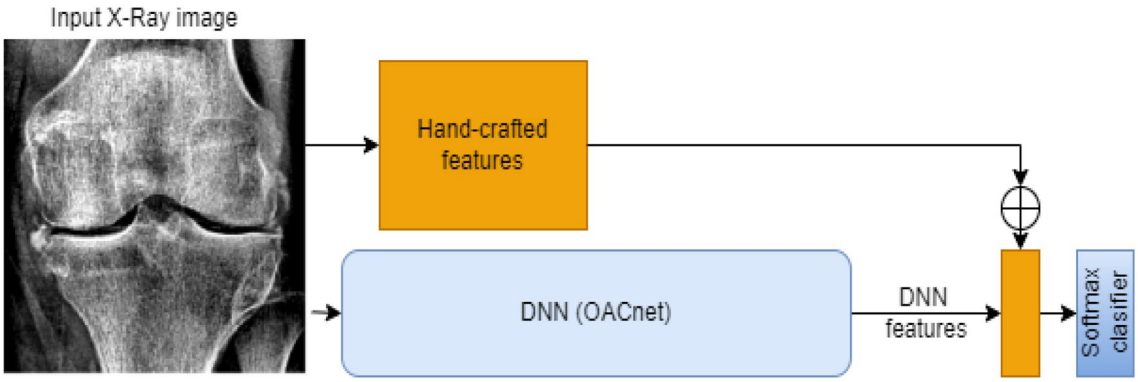

Fig. 7 DNN model with integration of hand-crafted features

Table 3 Classification results of proposed and existing method (on OAI Dataset)

Method	Grade	Accuracy	MSE	Accuracy
DNN (OACnet)	0	83.52	16.48	83.74
	1	84.56	15.44	
	2	89.12	18.88	
	3	80.25	19.75	
	4	81.25	18.75	
DNN (OACnet)+Hand-crafted features	0	92.46	18.24	**92.7**
	1	93.61	17.09	
	2	98.66	20.90	
	3	88.84	21.86	
	4	89.94	20.76	

The result of proposed method is highlighted in bold just to show accuracy is been increased because of proposed method

Table 4 Performance comparison of various DNN model

Network	Precision	Recall	F1
Densenet121	0.67	0.87	0.75
Densenet169	0.39	0.45	0.41
Densenet201	0.52	0.54	0.52
Inceptionv3	0.75	0.72	0.73
Resnet153	0.75	0.78	0.76
Vgg16	0.76	0.68	0.71
OACnet	0.81	0.83	0.82
'OACnet with hand-crafted features'	**0.90**	**0.92**	**0.91**

The result of proposed method is highlighted in bold just to show accuracy is been increased because of proposed method

with the DNN features before the softmax classier. Model and engineered features' combination is shown in Fig. 7.

Experimental Results

For both the experimentation (i.e., DNN-based and DNN with integrated hand-crafted features), machine having Windows 10 operating system with i7 processor, 8GB memory, and Nvidia GeForce 940 M graphical processing unit (GPU) is used. Python 3, TensorFlow 1.12, compute unified data architecture (CUDA) 10.0, and CUDNN 7.4 are used to implement both methods. The DNN training parameters, such as batch size, learning rate, and epoch, are set to 16, 0.001, and 20, respectively.

Among the database, 50% (4746) images were used for training, and 30% (2848) images were used for validation, whereas the rest 20% (1898) were kept for testing in both the approaches. Out of 4746 images, 520 are of grade 0 (normal), 547 are of grade 1 (minor), 726 are of grade 2 (mild), 1255 are of grade 3 (moderate), and 1698 are of grade 4 (severe). By observing the X-ray images belonging to each grade, we can conclude that due to lack of prevalence, patients start examination at a later stage when the severity of pain is intolerable. Therefore, X-ray images having severe grades are more than moderate ones.

Comparison of DNN-Based OACnet with OACnet with Hand-Crafted Features

The results of DDN-based OACnet architecture are compared with OACnet integrated with hand-crafted features. The comparative results of both the methods are tabulated in Table 3. Two evaluation parameters, classification accuracy, and mean square error (MSE) for each grade are considered for result comparison. The overall accuracy (average of per-class accuracy) of the OACnet method is 83.74%, whereas for the OACnet with the integration of the hand-crafted feature method is 92.7%. The improvement in overall accuracy is more for OACnet with

integrated hand-crafted features as compared to DNN-based OACnet, because more robust features are extracted by hand-crafted feature engineering.

Comparison of Proposed Method with Other DNN-Based Models

In addition to comparison with DNN-based OACnet, the performance of OACnet integrated with hand-crafted features is compared with other DNN-based architectures, such as densenet121, densenet169, densenet201, inceptionv3, resnet153, and vgg16. The evaluation parameters precision, recall, and F1 score are calculated for the comparison. The detailed comparison is tabulated in Table 4. Among all the DNN-based architectures, proposed OACnet integrated with hand-crafted features shows highest precision. Since knee OA classification does not require heavy architecture (i.e., architectures with higher number of parameters), the proposed method best fits for the said task.

Conclusion

The classification accuracy of a machine learning-based model is largely dependent on well-organized problem-specific data and feature engineering. As far as the knee OA severity classification system is concerned, the primary aim of feature engineering is to represent pathological parameters (joint space distance, bone spur, and change in bone morphology for instance) precisely. Though the DNN-based model performs better, still the classification accuracy is not promising in OA severity prediction. Lack of well-classified training data and data imbalance are the major reasons for this situation. To improve the classification accuracy, we proposed a DNN-based knee OA severity prediction system that integrates engineered features into the model. Two type of hand-crafted features are extracted: (a) histogram modeling-based statistical features such as quartiles and its difference; (b) small and large eigenvalues of the upper and lower envelope. The engineered features exact features that have high similarity with clinical features. The proposed severity prediction system follows K &L grades and yields 92.7% accuracy on NIH-OAI radiographic image dataset. By confirming the literature, this is one of the unique attempts to improve classification accuracy based on pathologically similar feature engineering.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Roemer FW, Demehri S, Omoumi P, Link TM, Kijowski R, Saarakkala S, Crema MD, Guermazi A. State of the art: imaging of osteoarthritis-revisited 2020. Radiology. 2020;296(1):5–21.
2. Harris R, Strotmeyer ES, Sharma L, Kwoh CK, Brach JS, Boudreau R, Cauley JA. The association between severity of radiographic knee OA and recurrent falls in middle aged and older adults: the osteoarthritis initiative. J Gerontol. 2022:1–7.
3. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review. JAMA. 2021;325(6):568–78.
4. Ruikar D, Sawat D, Santosh K, Hegadi R. 3d imaging in biomedical applications: a systematic review. In: Santosh KC, Antani S, Guru DS, Dey N, editors. Medical Imaging. Boca Raton: CRC Press; 2018. p. 155–81.
5. Ruikar DD, Hegadi RS, Santosh K. A systematic review on orthopedic simulators for psycho-motor skill and surgical procedure training. J Med Syst. 2018;42(9):1–21.
6. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. Sci Rep. 2018;8(1):1–10.
7. Smith R, Egger P, Coggon D, Cawley M, Cooper C. 9. kellgren jh, lawrence js. radiological assessment of osteoarthritis of the hip joint and acetabular dysplasia in osteoarthritis. Ann Rheum Dis. 1957;16:494–502.
8. Ruikar DD, Santosh K, Hegadi RS, Rupnar L, Choudhary VA. 5k+ ct images on fractured limbs: a dataset for medical imaging research. J Med Syst. 2021;45(4):1–11.
9. Ruikar DD, Santosh K, Hegadi RS. Segmentation and analysis of CT images for bone fracture detection and labeling. Med Imaging. 2019;131:131–54.
10. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, Yamamoto S, Nakamura K, Yoshimura N, Kawaguchi H. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008;16(11):1300–6.
11. Hegadi RS, Navale D, Pawar TD, Ruikar DD. Multi feature-based classification of osteoarthritis in knee joint x-ray images. In: Hegadi RS, Navale DN, Pawar TD, Ruikar DD, editors. Medical Imaging. CRC Press; 2019. p. 75–96.
12. Hegadi RS, Chavan UP, Navale DI. Identification of knee osteoarthritis using texture analysis. In: Data analytics and learning, Springer, 2019. p. 121–9.
13. Sun Y, Teo E, Zhang Q. Discussions of knee joint segmentation. In: 2006 International conference on biomedical and pharmaceutical engineering 2006; IEEE.
14. Stammberger T, Eckstein F, Michaelis M, Englmeier K-H, Reiser M. Interobserver reproducibility of quantitative cartilage measurements: comparison of b-spline snakes and manual segmentation. Magn Reson Imaging. 1999;17(7):1033–42.
15. Zhao F, Xie X. An overview of interactive medical image segmentation. Ann BMVA. 2013;2013(7):1–22.
16. Seise M, McKenna SJ, Ricketts IW, Wigderowitz CA. Segmenting tibia and femur from knee x-ray images. In: Proc. of medical image understanding and analysis, 2005. p. 103–6.
17. Hegadi RS, Navale DI, Pawar TD, Ruikar DD. Osteoarthritis detection and classification from knee x-ray images based on artificial neural network. In: International conference on recent trends in image processing and pattern recognition, Springer, 2018. p. 97–105.
18. Navale DI, Hegadi RS, Mendgudli N. Block based texture analysis approach for knee osteoarthritis identification using SVM. In: Electrical and Computer Engineering (WIECON-ECE), 2015 IEEE International WIE Conference On. IEEE, 2015. p. 338–41.

19. Suresha S, Kidziński L, Halilaj E, Gold G, Delp S. Automated staging of knee osteoarthritis severity using deep neural networks. Osteoarthritis Cartilage. 2018;26:441.
20. Antony J, McGuinness K, Moran K, O'Connor NE. Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks. In: International conference on machine learning and data mining in pattern recognition, Springer, 2017. p. 376–90.
21. Kawathekar PP, Karande KJ. Use of textural and statistical features for analyzing severity of radio-graphic osteoarthritis of knee joint. In: Information Processing (ICIP), 2015 International Conference On, 2015. p. 1–4. IEEE.
22. Navale DI, Ruikar DD, Houde KV, Hegadi RS. Dwt textural feature-based classification of osteoarthritis using knee x-ray images. In: International conference on recent trends in image processing and pattern recognition, Springer, 2020. p. 50–9.
23. Riad R, Jennane R, Brahim A, Janvier T, Toumi H, Lespessailles E. Texture analysis using complex wavelet decomposition for knee osteoarthritis detection: data from the osteoarthritis initiative. Comput Electric Eng. 2018;68:181–91.
24. Antony J, McGuinness K, O'Connor NE, Moran K. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In: Pattern Recognition (ICPR), 2016 23rd International Conference On, 2016. p. 1195–200. IEEE.
25. Anifah L, Purnama IKE, Hariadi M, Purnomo MH. Osteoarthritis classification using self organizing map based on Gabor kernel and contrast-limited adaptive histogram equalization. Open Biomed Eng J. 2013;7:18.
26. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. 1979;9(1):62–6.
27. Ruikar DD, Santosh K, Hegadi RS. Contrast stretching-based unwanted artifacts removal from CT images. In: International Conference on Recent Trends in Image Processing and Pattern Recognition, Springer, 2018. p. 3–14.
28. Ruikar DD, Santosh K, Hegadi RS. Automated fractured bone segmentation and labeling from CT images. J Med Syst. 2019;43(3):1–13.
29. Jagtap AB, Hegadi RS. Offline handwritten signature recognition based on upper and lower envelope using eigen values. In: 2017 World Congress on Computing and Communication Technologies (WCCCT), 2017. p. 223–6. IEEE
30. Jagtap AB, Hegadi RS. Eigen value based features for offline handwritten signature verification using neural network approach. In: International Conference on Recent Trends in Image Processing and Pattern Recognition, Springer, 2016. p. 39–48.

Authors and Affiliations

Darshan Ruikar[1] · Parshuram Kamble[2] · Amruta Ruikar[3] · Kavita Houde[2] · Ravindra Hegadi[2]

Parshuram Kamble
parshu1983@gmail.com

Amruta Ruikar
amrutashegadar8887@gmail.com

Kavita Houde
kavitahoude@gmail.com

Ravindra Hegadi
rshegadi@gmail.com

[1] School of Computer Science, MIT World Peace University, Kotharud, Pune, Maharashtra 411038, India

[2] School of Computer Science, Central University of Karnataka, Aland Road, Kadaganchi, Karnataka 585367, India

[3] Department of Computer Science and Engineering, SPM Polytechnic, Kumathe, Solapur, Maharashtra 413224, India