Journal of
Engineering Sciences
Faculty of Engineering
Assiut University
JES
مجلة
العلوم الهندسية
كلية الهندسة
جامعة أسيوط
journal homepage: http://jesaun.journals.ekb.eg

Osteoarthritis Classification Algorithm Using CNN and Image Edge Detections

Received 29 January 2024; Revised 21 December 2024; Accepted 21 December 2024

Hanafy M. Ali[1]

Keywords
Osteoarthritis, CNN, X-ray, Deep Learning, Classification, feature extraction, knee osteoarthritis detection.

Abstract: The paper presents a comprehensive computer-aided diagnosis (CAD) system designed for early detection of knee Osteoarthritis (OA) utilizing knee medical imaging and machine learning algorithms. Osteoarthritis is a prevalent chronic disease affecting various joints, primarily the fingers, thumbs, spine, hips, knees, and big toes, with secondary occurrences linked to pre-existing joint anomalies. Although more common among older individuals, OA can develop in adults of any age, characterized by degenerative changes in joints. Traditional diagnosis involves examining joint scans, typically through X-ray analyses being conducted by trained radiologists and orthopaedists, which can be time-consuming and subject to precision loss due to manual segmentation. Automatic segmentation and interpretation of joint X-ray scans are thus necessary to enhance clinical outcomes and bone calculation precision. The advent of deep learning technologies in medical systems has facilitated such transition, enabling efficient processing of large data volumes with improved accuracy. In particular, Convolutional Neural Networks (CNNs) being among the deep learning methods, have proven effectiveness in automating X-ray scan segmentation. The paper provides an overview of various deep learning and image processing techniques employed for automatic segmentation and interpretation of X-ray scans, facilitating disease diagnosis based on image data along with a proposed improved model of Visual Geometry Group VGG-16 with edge detection using X-ray images. A classification algorithm based on CNN and image edge detection is proposed demonstrating promising results, achieving predictive accuracies exceeding 90% across all suggested models. Particularly is the performance of the proposed VGG-16 after training with edge detection, which attained a training accuracy of 100% and a testing accuracy of 98.2%. This highlights the efficacy of deep learning approaches in enhancing diagnostic accuracy and efficiency in knee OA detection. A comparative evaluation of the proposed algorithm against other techniques based on performance metrics is reported.

1. Introduction

Recent technological advancements highlight the burgeoning field of Deep Learning, poised for extensive application in health and across industries, from low-tech to high-tech sectors

[1] Computers and Systems engineering Depart. , Faculty of Engineering, Minia University, El Minia, Egypt. **hmali@mu.edu.eg**

https://doi.org/10.21608/jesaun.2024.266440.1306

with massive datasets [1]. By 2025, the volume of data is projected to triple, driving a trillion-dollar global market. Yann LeCun [2] defines Deep Learning as enabling computational models with multiple layers of processing to learn data representations with varying levels of abstraction, revolutionizing current computing paradigms to meet the diverse needs of health, society, and education. Osteoarthritis (OA) [3], predominantly affecting females, the elderly, and overweight individuals, stands out as one of the most severe forms of arthritis. With over 27 million projected cases in the United States alone [4], it primarily afflicts individuals aged 60 and older. Referring to as "wear-and-tear" arthritis, OA typically develops gradually with age, stemming from prolonged joint use. Knee OA, in particular, targets cartilage, the protective connective tissue covering bone ends within joints. As OA progresses, cartilage erosion leads to bone-on-bone contact, resulting in severe discomfort. The Kellgren-Lawrence (KL) scheme offers a standardized classification method, dividing individual joints into five groups [5, 6].

Symptoms of OA vary among individuals, with joint damage often occurring gradually over years, accompanied by increasing pain. However, progression can also be rapid. While some individuals experience mild symptoms that do not significantly disrupt daily life, others endure severe disability and pain. OA manifests in two main types: primary OA, prevalent among the elderly due to genetic or age-related factors, and secondary OA, which may arise earlier in life due to conditions such as diabetes, obesity, athletic injuries, or rheumatoid arthritis. Key OA symptoms include joint pain, motion limitations, reduced mobility, stiffness following rest, and diminished participation in activities [7]. Current OA assessment relies on clinical examination, distinctive radiographic evaluations such as MRI techniques [8-11], enhancing image quality and acquisition speed. Recent biomedical imaging methods further improve image acquisition efficiency.

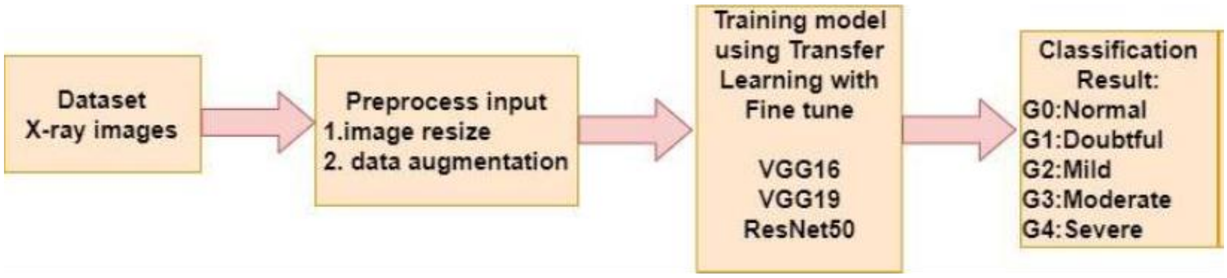


Fig. 1: Computer Aided Medical Image Diagnosis.

Figure 1 illustrates the operation scheme of a computer-aided system designed for OA diagnosis, integrating advancements in deep learning and imaging technologies to enhance diagnostic accuracy and efficiency. Further the structure of the paper is as follows. Section 2 is devoted to the literature survey. Section 3 reviews the CNN algorithms. The proposed model based on VGG16 algorithm with edge detection is described in section 4. Experimental setup and dataset details are provided in section 5. Results and analysis are discussed in section 6. Finally, section 7 concludes the paper.

2. Literature Survey

The Multicenter Osteoarthritis Study [12] utilized a Deep Siamese Convolutional Neural Network (CNN) to diagnose osteoarthritis. The study initially trained the model on data from the Multicenter Osteoarthritis Study and then validated it on a randomly selected subset of 3,000 subjects (5,960 knees) from the Osteoarthritis Initiative dataset. The accuracy achieved by this method was reported to be 66%. In another study by Deodar et al. [13], datasets were collected from various hospitals and diagnostic centers, specifically focusing on MRI images. These images underwent processing where different image processing algorithms were used to extract features such as GLCM texture, statistical measures, and shape characteristics. The accuracy of this method in diagnosing osteoarthritis was reported to be 95.24%.

A discrete step algorithm for X-Ray bone image segmentation was presented [14]. A dataset of 3D MRI Scans was generated, and Artificial Neural Network (ANN) was employed to learn the mapping function between the CDI feature space and OA severity with 70.6% accuracy [15]. The data of 33 patients with osteoarthritis in a National Hospital of Indonesia were used for prediction bases on Random Forest algorithm with 86.9% accuracy [16]. Dataset of X-ray images and KL grades from the Osteoarthritis Initiative (OAI) have been analyzed using deep Convolutional Neural Networks (CNN) with 53.9% accuracy [17]. Dataset of X-ray images and KL grades from the Osteoarthritis Initiative (OAI) was investigated through fractal analysis of trabecular bone textures on radiographs [18]. An edge detection algorithm was proposed [19] to be adapted with human knee osteoarthritis Images. Knee X-ray images of 200 patients with different ages, and various socio-economic blood groups were analyzed using effective contour segmentation technique was applied for diagnosing the disease by segmenting a portion of the knee images. Numerous functions such as Horlick, Statistical, First Four Moments, Texture and Shape were calculated and analyzed using Random Forest classifier for diagnosis of the disease with 87% accuracy [21].

Two CNNs were sequentially used to autonomously quantify the severity of knee OA. Given the size of the knee joints scattered with slight changes in the X-ray images, they used the modified one stage YOLOv2 network to detect the knee joints in the first stage. Second, CNN, which is the most popular, has been improved. In the investigation, the VGG-19 model yielded an accuracy of 69.7% [21, 22]. A novel, transparent, computer-aided diagnostic technique based on a Deep Siamese CNN was introduced to automatically grade the severity of OA in the knee using the Kellgren-Lawrence rating system. They verified their strategy in 3000 randomly selected participants (5960 knees) from the OA initiative dataset, using solely data from the Multicenter OA Study. An average multi-class accuracy of 66.71% was obtained using their method [23]. In reference [24], it was stated that automatic evaluation of the severity of knee OA consists of two stages. First, the knee joints were automatically localized. Localized knee joint images were then classified. A CNN-based method was proposed to autonomously identify knee joints. By maximizing the weighted ratio of two loss functions, CNNs were trained to automatically assess the severity of OA in the knee [24, 25].

3. CNN Algorithms.

Deep Learning (DL) techniques, particularly those based on Artificial Neural Networks (ANNs), have indeed shown remarkable success in various fields, including disease diagnosis and detection. ANNs are inspired by the human brain's functioning, with the basic premise being the processing of numerous input signals through nonlinear operations to generate output signals, akin to how neurons process information. In the context of disease diagnosis and detection, DL models utilize large datasets comprising various features related to the disease under consideration. These features could include medical images, clinical data, genetic information, and more. One of the key advantages of DL models is their ability to automatically learn and extract relevant features from raw data without explicit feature engineering, a process often referred to as feature extraction or representation learning.

The process of feature extraction in DL models involves the successive transformation of raw input data through multiple hidden layers, also known as "hidden layers" or "latent layers," within the neural network architecture. Each hidden layer comprises a set of neurons that perform nonlinear transformations on the input data, gradually abstracting and representing higher-level features as the data propagates through the network. The exact mechanisms by which these hidden layers extract meaningful features from the input data can be considered somewhat opaque or "secret," as they are determined by the weights and biases learned during the training phase of the neural network. However, through the iterative optimization process, the network learns to identify and amplify the features that are most relevant for the task at hand, effectively encoding them within the learning capacity of the model.

Once trained, the DL model can effectively classify or predict disease outcomes based on the learned representations of the input data. This capability has led to significant advancements in medical diagnosis, enabling more accurate and timely identification of diseases from various modalities, such as medical imaging, electronic health records, and genomic data. In summary, DL techniques based on ANNs leverage the principles of nonlinear feature extraction to automatically learn representations of input data, including those relevant to disease diagnosis and detection. While the exact process by which these features are extracted may not be fully transparent, the overall effectiveness of DL models in medical applications underscores their potential to augment and improve healthcare outcomes [26, 27].

Usually, feature extraction in deep learning is conducted as a step of the training process of the model. Unlike traditional machine learning techniques where features are extracted separately, deep learning models learn to extract features directly from raw data, such as images or text. Various machine learning techniques, including support vector machines

(SVM), k-means clustering, and neural networks, have been utilized in diagnosing neurological disorders. These techniques leverage data from different sources, such as medical images or clinical records, to aid in disease detection and classification. Some techniques used in diagnosing disorders are inspired by natural phenomena. These approaches may mimic biological processes or systems to achieve better results in disease diagnosis. Such bio-inspired methods have shown promising results in previous studies. Early diagnosis of Arthritis disease **(AD)** is crucial for effective treatment and prevention of disease progression. Magnetic resonance imaging (MRI) of the knee, being one of the medical imaging techniques of the knee, provide valuable information for diagnosing AD. Features extracted from MR images of the knee are evaluated and classified to aid in the detection of Arthritis disease. These features should accurately capture the stages and characteristics of AD, facilitating more accurate diagnosis and treatment planning. Overall, the integration of machine learning techniques, including deep learning models, and advanced imaging technologies holds promise for improving the early detection and management of neurological disorders like Arthritis disease [28, 29]. In this section, three different pre-trained models emerging within the scope of deep learning are examined. Inception V3, VGG16, and VGG19 models are discussed below.

3.1. Inception V3

Inception-v3 as shown in figure 2 is a deep convolutional, 48-layer wide, neural network. The inception V3 function receives images from ImageNet and returns by their classification [30]. The network inception V3 has a 299-by-299 picture scale for input.

3.2. VGG16

As shown in figure 3, VGG16 is a 16-layer VGG layout version representing a convolutional neural network architecture initially designed for image classification tasks. It comprises 13 convolutional layers, 5 pooling layers, and 3 fully connected layers. The architecture is divided into six blocks, with each block having a specific number of convolutional and pooling layers. All convolutional layers in VGG16 use 3x3 filters [31].

3.3. VGG19

VGG19 is a 19-layer VGG layout version (16 layers of convolution, 3 fully connected layers, and 5 layers of MaxPool and 1 layer of SoftMax). Many versions of VGG include VGG11, VGG16, and others. As an input to this network, a fixed size of (224 x 224) RGB image was given, indicating that the matrix is in form (224, 224, and 3) [31, 32]. For the only completed pre-processing image, subtraction of the mean RGB value from each pixel was made for the whole training set. Use of (3 x 3) size kernels with a 1-pixel phase scale allows them to cover the whole notions of image. Spatial wrapping helps to maintain spatial resolution of the image. The total pooling was carried out over a 2 x 2-pixel side frame 2. This is accompanied by rectified linear unit with added non-linearity to allow the model better to discriminate and decrease the computational time to prove much better than commonly-used previous models based on simulations using tanh or sigmoid functions.

Applied three totally connected layers from which the first two are 4096 heights, then a layer with 1000 channels for classification of ILSVRC 1000-way and the final layer is a SoftMax feature.

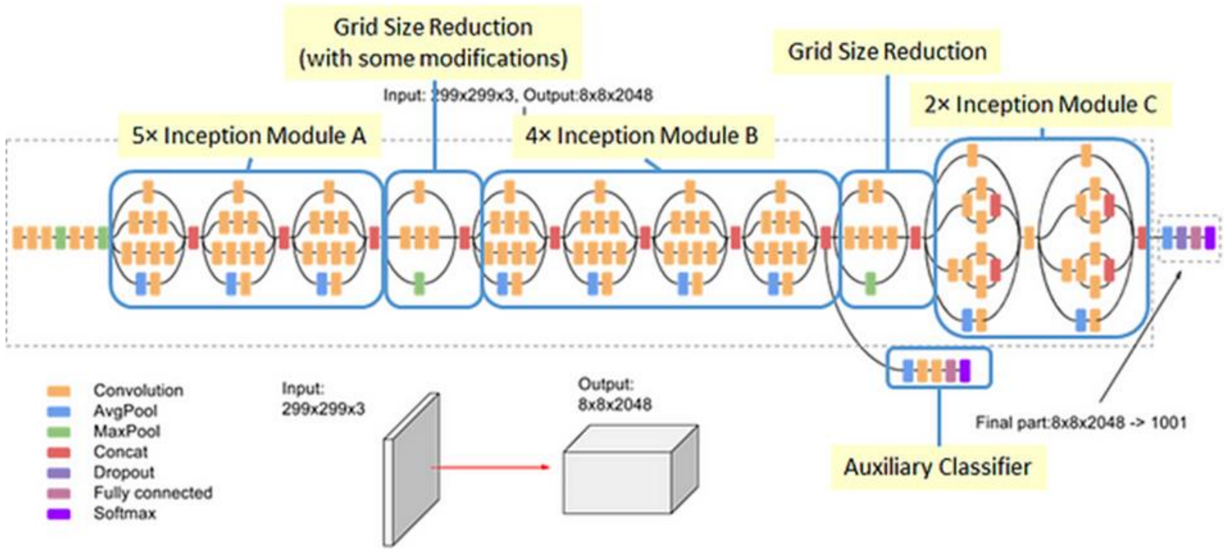


Fig. 2: Inception V3 Architecture.

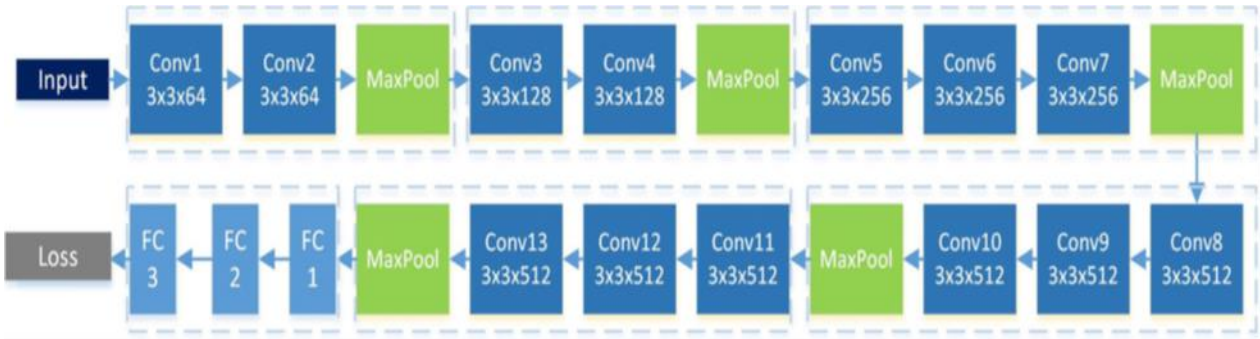


Fig. 3: Customized VGG16 Architecture.

A deep convolutional neural network (DCNNs) based on the VGG-16 architecture, with a focus on a significantly reduced complexity but with a performance that is comparable or superior to those of all the other existing edge detection techniques, was proposed [33]. The objective of significantly reduced complexity of the network was achieved using fire modules so much so that it is possible to increase the depth of the network while keeping its character of low complexity. A chapter delved into the intricate field of image edge detection, a pivotal aspect of computer vision and image processing was published [34]. It provided a comprehensive exploration of the underlying principles, methodologies, and algorithms employed in the identification and extraction of significant contours in digital images. Traditional edge-detection techniques, as well as advanced approaches being based on deep learning, were thoroughly examined.

4. Proposed algorithm VGG16 with Edge Detection.

As shown in Fig. 4, the proposed algorithm is based on VGG16 supported with edge detection. Edge detection includes a variety of mathematical methods, which aim to identify points in a digital image where the brightness of the image changes sharply or has discontinuities, more formally. Typically, the points at which image brightness changes sharply are structured into a set of curved line segments called edges. The same problem of finding discontinuities in one-dimensional signals is known as phase detection, and shift detection is known as the problem of finding discontinuities in signal over time. Edge detection is an important tool for image processing, machine vision and computer vision, particularly in the fields of feature detection and extraction.

To extract the edges of the knee joint from a pre-processed image, one can use various edge detection techniques. One commonly used method is the canny edge detector, which is a multi-stage algorithm involving gradient calculations and non-maximum suppression.

Algorithm 1: algorithm to extract the edges of the knee:

Step 1: Import Libraries dataset plain images as the input of the algorithm.

Step 2: The plain images are processed by automatic segmentation and enhancement through the use of deep nested nets [35].

Step 3: The pre-processed images are resized and normalized.

Step 4: The Canny edge detection algorithm is used to detect edges in the image. This involves specifying appropriate minimum and maximum threshold values for edge detection. The mean squared error (MSE), and peak signal-to-noise ratio (PSNR) are considered the measures for assessing the quality of edges. The MSE and PSNR are calculated with values of 105.1442 and 27.9129, respectively [36].

Step 5: The quality of edges in the images is improved by applying post-processing techniques such as morphological operations (e.g., dilation, erosion) to refine the detected edges [36].

Step 6: The images with improved edges are displayed or saved as the output of the algorithm.

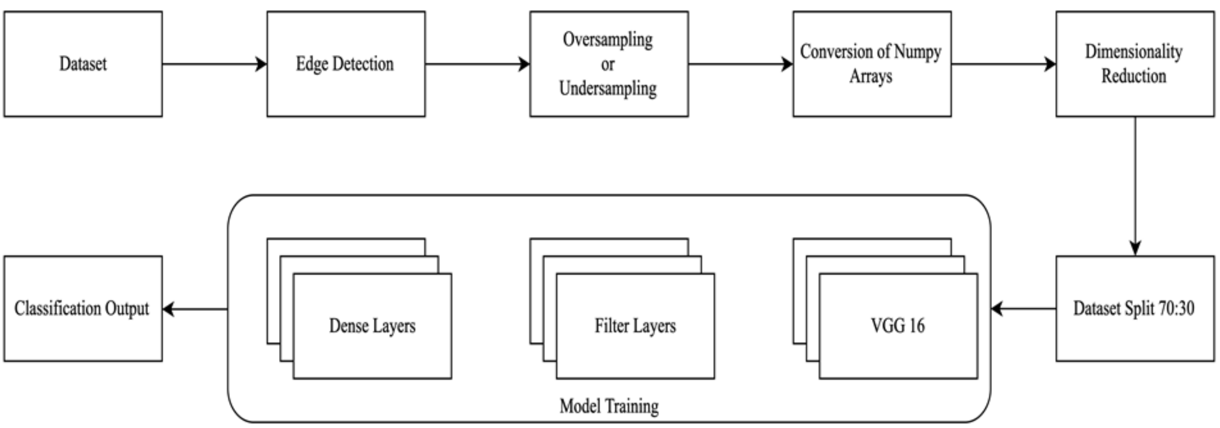


Fig. 4: Proposed Algorithm Architecture for Edge Detection.

The Visual Geometry Group VGG-16 in the proposed algorithm contains 13 convolutional layers (C1 ~ C13) and 3 fully connected layers (extract feature layer) (FC1 ~ FC3). VGG16 has performed much better as compared to other convolutional networks, like Alex Net, ResNet50, InceptionNetV3, because of the use of VGG16 architecture to extract features from the images [1].

Algorithm 2: proposed algorithm VGG16 with edge detection to Classifiers:

Step 1: The images of with improved edges are given as input to the algorithm.

Step 2: The augment and data generator routine functions are created for testing, training, and validation of the input images.

Step 3: The VGG-16 model is built and compiled.

Step 4: The Kellgren-Lawrence (KL) scheme is graded for computing the values of accuracy and loss. This represents the outputs of the algorithm.

Step 5: The output of the algorithm is compared against that obtained from Inception-v3, VGG19, and VGG16 without image edge detection.

4.1 Mean Square Error (MSE), and Peak Signal-to-Noise Ratio (PSNR)

The mean square error (MSE) is calculated to make sure that the original and edge detection images are in variations or not. The lesser the MSE between two images, the better the edge detection; it is defined as:

$$MSE(q_1,q_2) = \frac{1}{MN}\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (q_1(i,j) - q_2(\mathrm{i},\mathrm{j}))^2 \qquad \text{Eq. (1)}$$

Where: q_1 (i, j) and q_2 (i, j) indicate the original and edge detection images respectively, M and N represent the size of the image in both directions.

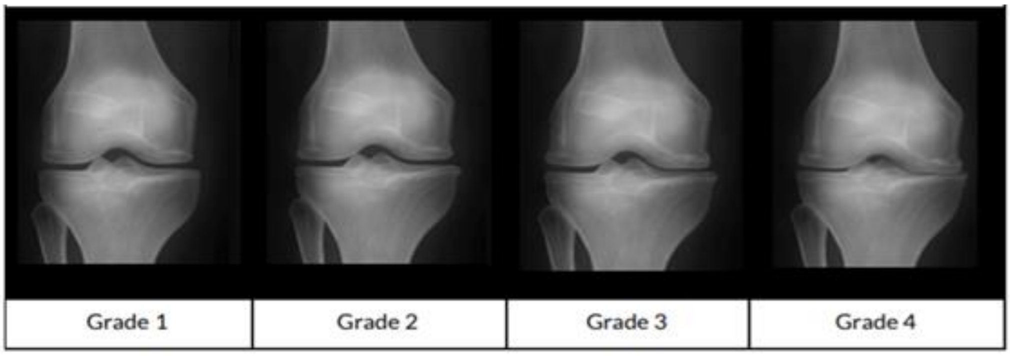
PSNR is calculated from equation 2:

$$PSNR = 20log_{10}(\frac{\mathrm{MAX}_f}{\sqrt{MSE}}) \qquad \text{Eq. (2)}$$

Where: MAX_f is the maximum signal value in the original image.

5. Experimental Setup

In this paper, the grades of OAI Dataset and Kellgren-Lawrence (KL) as ground reality are used to characterize OA X-ray photographs of the knee. The KL classification scheme is now considered the gold standard for initial determination of the extent of knee osteoarthritis in X-rays. This uses five classes to show seriousness of OA x-ray knee. .'Grade 0' is standard, 'Grade 1' is doubtful, 'Grade 2' is minor, 'Grade 3' is mild and 'Grade 4' is extreme. Figure 5 indicates the method for rating the KL. The rating scheme Kellgren Lawrence is a radiological assessment of osteoarthritis of the hip. This progresses from grade 0 to grade IV which is focused on x-rays. There are a range of rating schemes developed for knee osteoarthritis, the rating scheme Kellgren-Lawrence is the most used and accepted method for the treatment of knee osteoarthritis.



Classification	Normal	Doubtful	Mild	Moderate	Severe
Description	No Features of OA	Minute Doubtful	Definite Normal joint space	Moderate joint space	Joint space greatly reduced

Fig. 5: Kellgren-Lawrence grading scheme.

5.1. Edge Detection Using Image Processing

Figure 6 depicts Image Processing for detecting the edges of the image and separating it from the other images based on the Kellgren-Lawrence grading scheme.

(A) (B)

Fig. 6: Edge detection before (A) and after (B) edge detection.

5.2. Osteoarthritis Dataset

A total of 8260 X-ray radiographies were gathered from 4796 participants in a study [35]. The radiographs include the left and right knee. They are divided in Table 1 into three sets: train, test, and validation set in the ratio of 7:2:1 with balanced distribution of all KL grades [37]. The actual (original) value for calculating the MSE was given along with the dataset [37].

Table 1: Dataset

Dataset	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4	Total
Training	2286	1046	1516	757	173	5778
Testing	639	296	447	223	51	1656
Validation	328	153	212	106	27	826
Total	3253	1495	2178	1086	251	8260

Reference is made to Table 1 which shows that the size of training, testing, and validation is high enough to test their respective accuracy with no need to conduct additional experiments. It is known that a model with too little size cannot learn and predict well the Kellgren-Lawrence (KL) grades. On the other hand, a model with much higher size can learn well with overfitting the training dataset. Both cases result in a model that does not predict well the osteoarthritis classification. In the present work, the size of dataset is not too small and not too high in order to predict well osteoarthritis classification.

5.3. Evaluation of Parameters

For testing the capability of suggested features on using different classifies, several experiments are made. Five performance parameters are used for evaluation to include [38-41]:

1. Sensitivity (SEN): This metric represents the proportion of true positive cases that were correctly identified by the classifier.
2. Specificity (SPF): This metric represents the proportion of true negative cases that were correctly identified by the classifier.
3. Accuracy (ACC): This metric represents the overall proportion of correctly classified cases (both positive and negative).
4. Positive Predictive Value (PPV): This metric represents the proportion of positive test results that were truly positive.
5. Negative Predictive Value (NPV): This metric represents the proportion of negative test results that were truly negative.

Sensitivity = (True Positive)/ (True positive + False Negative) * 100 Eq. (3)

Specificity = (True Positive)/ (True Negative + False Positive) * 100 Eq. (4)

Precision = (True Positive)/ (True positive + False Positive) * 100 Eq. (5)

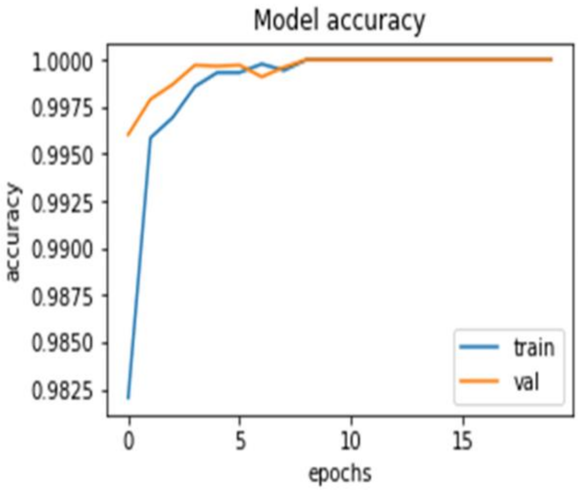
Accuracy = (Correct predication /Total predication) *100 Eq. (6)

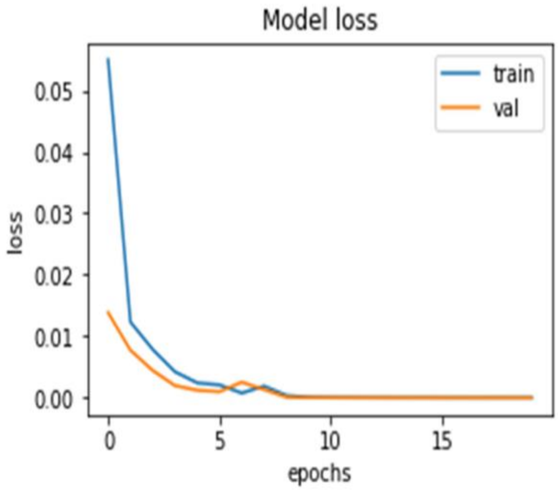
= ((True Positive+ True Negative)/ (True positive + False Negative+ False positive + True Negative)) * 100

F-measure = 2*((Sensitivity+ Precision)/ (Sensitivity * Precision)) *100 Eq. (7)

6. Results and Discussion

In this section, the experimental results and performance analysis of the proposed model are discussed. At first, the performance of the proposed transfer learning model for OS detection is analyzed through model accuracy and model loss. Figure 7 shows the VGG16 with edge detection transfer learning model for the OS detection. Figure 7(a) shows the graph of accuracy versus epochs and Fig. 7(b) shows the loss graph versus epochs. The graph shows that the model loss is minimized through various regularization techniques and overfitting is avoided through dropouts and L1 and L2 regularizations. Subsequently, the accuracy of the model is improved. The difference between the training and testing accuracy and loss shows the model is performing well.





(a) **(b)**

Fig. 7: Proposed Model Performance using Transfer Learning VGG16 with Edge Detection (a) Accuravy versus Epoch (b) Loss versus Epoch

Table 2: Training, test, and validation accuracy of VGG16, VGG19, and VGG16 (Edge Detection).

Deep Neural Network	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
Inception V3	95.7	89.2	92.6
VGG16	96.1	90.3	94.4
VGG19	95.1	87.1	91.5
VGG16(Edge Detection)	100	96.64	98.2

Figure 7 dictates that VGG16 with edge detection scores 100% for both the training accuracy and the validation accuracy using transfer learning against 100% and 96.64% without transfer learning in Table 2. The results confirm the capability of transfer learning to adapt the proposed model to match characteristics of the investigated dataset pointing to robust and accurate classification of the medical images. The sensitivity, specificity, precision, accuracy, and F-measure of the obtained results on applying the different deep NN models are summarized in Table 3. These metrics provide an overview into the performance of the models in terms of correct identification of positive cases, actual capture of positive instances, and achievement of a balance between precision and recall. The proposed model, among these models, the proposed model "VGG16 with edge dedication" exhibited the highest sensitivity, specificity, precision, accuracy, and F-measure, indicating its excellence in classification accuracy.

Table 3: Sensitivity, specificity, precision, accuracy, and F-measure obtained by the different CNN algorithms.

Deep Neural Network		Sensitivity	Specificity	Precision	Accuracy	F-measure
Inception V3	Grade 0	0.8	0.92	0.88	0.9379	0.83
	Grade 1	0.79	0.925	0.74	0.9035	0.77
	Grade 2	0.89	0.90	0.76	0.9154	0.79
	Grade 3	0.79	0.85	0.8	0.9198	0.80
	Grade 4	0.77	0.83	0.83	0.928	0.80
VGG16	Grade 0	0.782	0.948	0.87	0.825	0.825
	Grade 1	0.72	0.93	0.75	0.812	0.765
	Grade 2	0.789	0.90	0.77	0.795	0.8
	Grade 3	0.75	0.85	0.8	0.782	0.805
	Grade 4	0.74	0.82	0.82	0.731	0.83
VGG19	Grade 0	0.80	0.915	0.57	0.75	0.66
	Grade 1	0.651	0.92	0.63	0.74	0.02
	Grade 2	0.49	0.88	0.60	0.71	0.44
	Grade 3	0.39	0.85	0.66	0.65	0.42
	Grade 4	0.37	0.75	0.74	0.69	0.40
VGG16 with Edge Detection	Grade 0	0.9776	0.985	0.9839	0.9818	0.9808
	Grade 1	0.967	0.981	0.981	0.98	0.98
	Grade 2	0.961	0.979	0.98	0.978	0.97
	Grade 3	0.95	0.97	0.96	0.97	0.971
	Grade 4	0.949	0.962	0.966	0.971	0.964

As given in Table 3, the sensitivity, specificity, precision, accuracy, and F-measure assume values for all grades in the ranges 0.72-0.782, 0.82-0.948, 0.75-0.87, 0.731-0.825, 0.765-0.83 on using VGG16 algorithm against 0.949-0.978, 0.962-0.985, 0.96-0.984, 0.97-0.982, and 0.964-0.981on using VGG16 with edge detection. This confirms the improvement resulting from the use of the proposed VGG16 with edge detection.

Table 4: Result of the proposed model (VGG16 with image edge detection).

Parameters	Outcome in %
Accuracy	97.9
Sensitivity	100
Specificity	100
Error rate	2.1

Table 4 lists a few metrics that are calculated for the VGG16 with edge detection binary classification, including accuracy, sensitivity, specificity, and error rate, with values of 97.9, 100, 100, and 2.1%, accordingly. Different from Table 3, Table 4 lists metrics calculated for the proposed model VGG16 with edge detection based on binary classification of patients into patients with or without osteoarthritis disease. These metrics are generated for patients with the disease to end up with respective values of accuracy, sensitivity, specificity, and error rate equal to 96.64, 100, 100, and 2.1%, as given in Table 4. As regards the class imbalance in the dataset, reference is made to the parameters' evaluation in section 5.2, including sensitivity, specificity, precision, accuracy, and F-measures. The F-measure, which is determined by both the sensitivity and the precision, was considered [42] to vary from 1 for perfect classifier to zero with no image was correctly identified. Fortunately, the F-measure assumes values in Table 3 in the range 0.96-0.98, which confirms that the dataset is almost free from any class imbalance.

7. Conclusions

This study aims at presenting an approach to use X-ray images for improving the performance of VGG16- based knee osteoarthritis detection. The approach utilizes an edge detection technique combined with VGG16. The main idea is to lead the model to focus on the edges detection of the bones in the joint. The proposed approach improves the detection accuracy of knee osteoarthritis in comparison with conventional classfication techniques (Inception V3, VGG16, and VGG19) without image edge detection. Therefore, the novelty in the present paper lies in introduction of image edge detection for improving the performance the VGG16. The results confirm the ability of learning transfer to adapt the proposed model to agree with the characteristics of the investigated dataset indicating robust and accurate classification of the medical images. The results showed that the performance of VGG16 with edge detection model is quite good in detecting osteoarthritis with an accuracy value 98.18%, sensitivity value 97.76%, and specificity value 98.5%.

References

[1] K. Kamal, and H. Ez-Zahraouy A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images. Research Square; 2023. DOI: 10.21203/rs.3.rs-2863523/v1.

[2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May-2015.

[3] S. Glyn-Jones et al., "Osteoarthritis," in the Lancet, 2015, vol. 386, no. 9991, pp. 376–387.

[4] Y. Zhang and J. M. Jordan, "Epidemiology of osteoarthritis," Clinics in Geriatric Medicine, vol. 26, no. 3. Elsevier, pp. 355–369, 01-Aug-2010.

[5] R. Mahum, SU. Rehman, T. Meraj, HT. Rauf, A. Irtaza, AM. El-Sherbeeny MA, El-Meligy (2021) a novel hybrid approach based on deep CNN features to detect knee osteoarthritis. Sensors 21(18):6189

[6] M. D. Kohn, A. A. Sassoon, and N. D. Fernando, "Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis," Clin. Orthop. Relat. Res., vol. 474, no. 8, pp. 1886–1893, Aug. 2016.

[7] C. R. James, J. S. Dufek, and B. T. Bates, “Effects of injury proneness and task difficulty on joint kinetic variability,” Med. Sci. Sports Exerc., vol. 32, no. 11, pp. 1833–1844, Nov. 2000.
[8] A. K. Taneja et al., “MRI features of the anterolateral ligament of the knee,” Skeletal Radiol., vol. 44, no. 3, pp. 403–410, 2015.
[9] J. Andrew, M. DIvyavarshini, P. Barjo, and I. Tigga, “Spine Magnetic Resonance Image Segmentation Using Deep Learning Techniques,” in 2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020, 2020, pp. 945–950.
[10] J. Andrew, R. Fiona, and H. Caleb Andrew, “Comparative study of various deep convolutional neural networks in the early prediction of cancer,” in 2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019, 2019, pp. 884–890.
[11] M. R. Karim et al., "DeepKneeExplainer: Explainable Knee Osteoarthritis Diagnosis from Radiographs and Magnetic Resonance Imaging," in IEEE Access, vol. 9, pp. 39757-39780, 2021, doi: 10.1109/ACCESS.2021.3062493.
[12] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, “Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach,” Sci. Rep., vol. 8, no. 1, pp. 1–10, Dec. 2018.
[13] D. D. Deokar and C. G. Patil, “Effective Feature Extraction Based Automatic Knee Osteoarthritis Detection and Classification using Neural Network.” 2015.
[14] P. Sharma and J. M. Singh, “A Novel Approach towards X-Ray Bone Image Segmentation using Discrete Step Algorithm,” undefined, 2013.
[15] M. Izadpanahkakhk, S. M. Razavi, M. Taghipour-Gorjikolaie, S. H. Zahiri, and A. Uncini, “Deep region of interest and feature extraction models for palmprint verification using convolutional neural networks transfer learning,” Appl. Sci., vol. 8, no. 7, Jul. 2018.
[16] A. Jamshidi, J. P. Pelletier, and J. Martel-Pelletier, “Machine-learning-based patient-specific prediction models for knee osteoarthritis,” Nature Reviews Rheumatology, vol. 15, no. 1. Nature Publishing Group, pp. 49–60, Jan-2019.
[17] J. Antony, K. McGuinness, N. E. O. Connor, and K. Moran, “Quantifying Radiographic Knee Osteoarthritis Severity using Deep Convolutional Neural Networks,” Proc. - Int. Conf. Pattern Recognit., vol. 0, pp. 1195–1200, Sep. 2016.
[18] G. W. Stachowiak, M. Wolski, T. Woloszynski, and P. Podsiadlo, “Detection and prediction of osteoarthritis in knee and hand joints based on the X-ray image analysis,” Biosurface and Biotribology, vol. 2, no. 4, pp. 162–172, Dec. 2016.
[19] Y. Du, J. Shan, R. Almajalid, and M. Zhang, “Knee osteoarthritis severity level classification using whole knee cartilage damage index and ANN,” in Proceedings - 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2018, 2019, pp. 19–21.
[20] S. S., P. U., and R. R., “Detection of Osteoarthritis using Knee X-Ray Image Analyses: A Machine Vision based Approach,” Int. J. Comput. Appl., vol. 145, no. 1, pp. 20–26, Jul. 2016.
[21] S. Boukir, O. Regniers, L. Guo, L. Bombrun and C. Germain, "Texture-based Forest cover classification using random forests and ensemble margin," 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 2015, pp. 3072-3075, doi: 10.1109/IGARSS.2015.7326465.
[21] Li, G., Li, S., Xie, J. et al. “Identifying changes in dynamic plantar pressure associated with radiological knee osteoarthritis based on machine learning and wearable devices”. J Neuro-Engineering Rehabil 21, 45 (2024). https://doi.org/10.1186/s12984-024-01337-6
[22] K. S. Raju, V. Amudha and M. N, "Early Detection and Quantification of Osteoarthritis Severity in Knee Using Support Vector Machine with Improved Accuracy Compared to Convolutional Neural Network," 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICONSTEM56934.2023.10142317.
[23] M. Tinhinane, N. Yassine, A. Ahmad, A. Soraya and J. Rachid, "Effects of Region of Interest Location

on Osteoarthritis Detection Using Deep Feature Learning," 2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, France, 2023, pp. 1-6, doi: 10.1109/IPTA59101.2023.10319997.

[24] M. Karaköse, H. Yetış and M. Çeçen, "A New Approach for Effective Medical Deepfake Detection in Medical Images," in IEEE Access, vol. 12, pp. 52205-52214, 2024, doi: 10.1109/ACCESS.2024.3386644.

[25] N. P. Challa, B. Naseeba, G. Vyshnavi, T. Priyanka, N. Jajam and K. S. Prasanna, "Osteoarthritis Disease Detection using Efficient Hyper-Tuning Parameters," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-9, doi: 10.1109/ACCAI58221.2023.10200102

[26] R. Gautam and M. Sharma, “Prevalence and Diagnosis of Neurological Disorders Using Different Deep Learning Techniques: A Meta-Analysis,” J. Med. Syst., vol. 44, no. 2, pp. 1–24, Feb. 2020.

[27] R. Gautam, P. Kaur, and M. Sharma, “A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings,” Progress in Artificial Intelligence, vol. 8, no. 4. Springer, pp. 401–424, Dec-2019.

[28] P. Kaur and M. Sharma, “Diagnosis of Human Psychological Disorders using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis,” Journal of Medical Systems, vol. 43, no. 7. Springer New York LLC, pp. 1–30, Jul-2019.

[29] E. Hosseini-Asl, R. Keynto, and A. El-Baz, “Alzheimer’s Disease Diagnostics by Adaptation of 3D Convolutional Network,” Proc. - Int. Conf. Image Process. ICIP, vol. 2016-Augus, pp. 126–130, Jul. 2016.

[30] Debnath, S., Roy, R. & Changder, S. A novel approach using deep convolutional neural network to classify the photographs based on leading line by fine-tuning the pre-trained VGG16 neural network. Multimed Tools Appl 83, 3189–3214 (2024). https://doi.org/10.1007/s11042-022-13338-5

[31] SR. Shah, S. Qadri, H. Bibi, SMW. Shah, MI. Sharif, F. Marinello, "Comparing Inception V3, VGG 16, VGG 19, CNN, and ResNet 50: A Case Study on Early Detection of Rice Disease". Agronomy. 2023; 13Rice. https://doi.org/10.3390/agronomy13061633

[32] S. Mohsen, A. M. Ali, E. -S. M. El-Rabaie, A. ElKaseer, S. G. Scholz and A. M. A. Hassan, "Brain Tumor Classification Using Hybrid Single Image Super-Resolution Technique With ResNext101_32× 8d and VGG19 Pre-Trained Models," in IEEE Access, vol. 11, pp. 55582-55595, 2023, doi: 10.1109/ACCESS.2023.3281529.

[33] A. Al-Amaren, M. O. Ahmed, and M. S. Swamy, " A low-complexity residual deep neural network for image edge detection," Applied Intelligence, Vol. 53, pp. 11282-11299, 2023 in IEEE Access, vol. 11, pp. 55582-55595, 2023.

[34] A. Benhagyousef, and A; Saidani, "Recent Advances on Image Edge Detection," Al-Amaren, M. O. Ahmed, M. S. Swamy, "A low-complexity residual deep neural network for image edge detection," published in an Edited Volume "Digital Image Processing - Latest Advances and Applications," editors F. Cuevas, P. Mazzeo and A. Bruno, DOI: 10.5772/intechopen.1003763, January 2024.

[35] M. Liu and P. Qian, "Automatic Segmentation and Enhancement of Latent Fingerprints Using Deep Nested UNets," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1709-1719, 2021, doi: 10.1109/TIFS.2020.3039058.

[36] Dipmala Salunke, Pallavi Tekade, Nihar Ranjan, Deepali Ujalambkar, Sunil Sangve, Deepak Mane, "Real-Time Dimension Detection using Customized Canny Edge Detection Algorithm," International Journal of Engineering Trends and Technology, vol. 71, no. 9, pp. 375-384, 2023. Crossref, https://doi.org/10.14445/22315381/IJETT-V71I9P233

[37] P. Chen, "Knee osteoarthritis severity grading dataset", 2018.

[38] N. P. Challa, B. Naseeba, G. Vyshnavi, T. Priyanka, N. Jajam and K. S. Prasanna, "Osteoarthritis Disease Detection using Efficient Hyper-Tuning Parameters," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-9, doi: 10.1109/ACCAI58221.2023.10200102.

[39] R. PV and U. Shanmugam, "Explainable AI for Medical Imaging: Advancing Transparency and Trust in Diagnostic Decision-Making," 2023 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2023, pp. 1-6, doi: 10.1109/i-PACT58649.2023.10434658.

[40] H. A. Alshamrani, M. Rashid, S. S. Alshamrani, A.H.D. Alshehri, "Osteo-NeT: An Automated System for Predicting Knee Osteoarthritis from X-ray Images Using Transfer-Learning-Based Neural Networks Approach". Healthcare, 11, 1206. https://doi.org/10.3390/healthcare11091206, 2023.

[41] Y. X. Teoh, A. Othmani, Kh. W. Lai, S. Li Goh, J. Usman, "Stratifying knee osteoarthritis features through multitask deep hybrid learning: Data from the osteoarthritis initiative", Computer Methods and Programs in Biomedicine, Volume 242,2023,107807, ISSN 01692607, https://doi.org/10.1016/j.cmpb.2023.107807.
(https://www.sciencedirect.com/science/article/pii/S016926072300473X).

[42] EA. Murphy, et al. Machine learning outperforms clinical experts in classification of hip fractures. Sci Rep. 2022 Feb 8; 12(1):2058. Doi: 10.1038/s41598-022-06018-9. PMID: 35136091; PMCID: PMC8825848.