OSTEOARTHRITIS CLASSIFICATION ALGORITHM USING CNN AND IMAGE EDGE DETECTIONS

G.E.Ziyadullaeva
Acting Associate Professor
R.M.Foziljonov
Master's Student, Tashkent University of Information Technologies,

Abstract. The article presents a comprehensive computer-aided diagnosis (CAD) system developed using knee medical imaging and machine learning algorithms for early detection of knee osteoarthritis (OA). The emergence of deep learning technologies in medical systems has facilitated this transition, allowing for more accurate and efficient processing of large amounts of data.

Keywords: Osteoarthritis, CNN, Xray, Deep Learning, Classification, feature extraction, knee osteoarthritis detection.

Annotatsiya. Maqolada tizza osteoartritini (OA) erta aniqlash uchun tizza tibbiy tasvirlash va mashinani o'rganish algoritmlaridan foydalangan holda ishlab chiqilgan keng qamrovli kompyuter yordamida diagnostika (CAD) tizimi taqdim etilgan. Tibbiy tizimlarda chuqur o'rganish texnologiyalarining paydo bo'lishi bunday o'tishni osonlashtirdi, bu katta hajmdagi ma'lumotlarni aniqroq va samarali qayta ishlash imkonini berdi.

Kalit so'zlar: Osteoartrit, rentgen, chuqur o'rganish, tasniflash, xususiyatlarni ajratib olish, tizza osteoartritini aniqlash.

Аннотация. В статье представлена комплексная система компьютерной диагностики (САПР), разработанная с использованием медицинской визуализации коленного сустава и алгоритмов машинного обучения для раннего выявления остеоартрита коленного сустава (OA). Появление технологий глубокого обучения в медицинских системах облегчило такой переход, позволив более точно и эффективно обрабатывать большие объемы данных.

Ключевые слова. Остеоартроз, сверточные нейронные сети, рентгенография, глубокое обучение, классификация, извлечение признаков, выявление остеоартроза коленного сустава.

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 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,

ISSN 3030-3907

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

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.

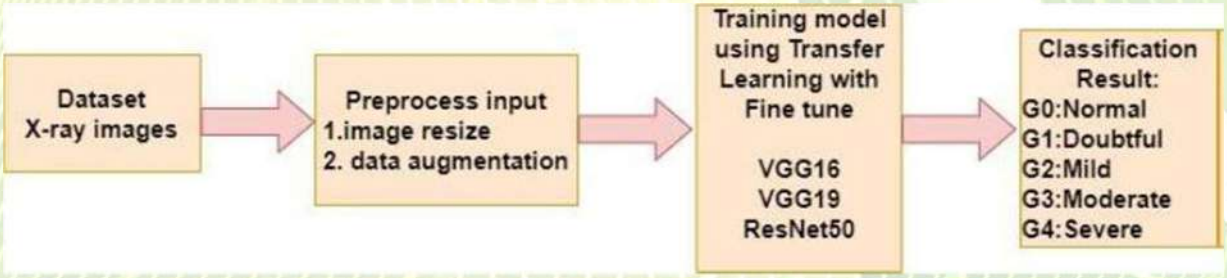


Fig. 1: Computer Aided Medical Image Diagnosis.

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.

ISSN 3030-3907

2. LITERATURE SURVEY

Prior work applies deep learning (especially CNNs) to OA diagnosis and severity assessment using radiographs and MRI, often improving performance through preprocessing (segmentation/ROI), transfer learning, and multi-stage pipelines. These studies indicate that emphasizing structural cues of the knee joint can increase robustness for OA grading and motivates the edge-enhanced CNN approach used in this work.

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 deeper VGG-family CNN that stacks small (3×3) convolutions and pooling blocks. In this study it is used as a comparison model against the proposed VGG16+edge pipeline.

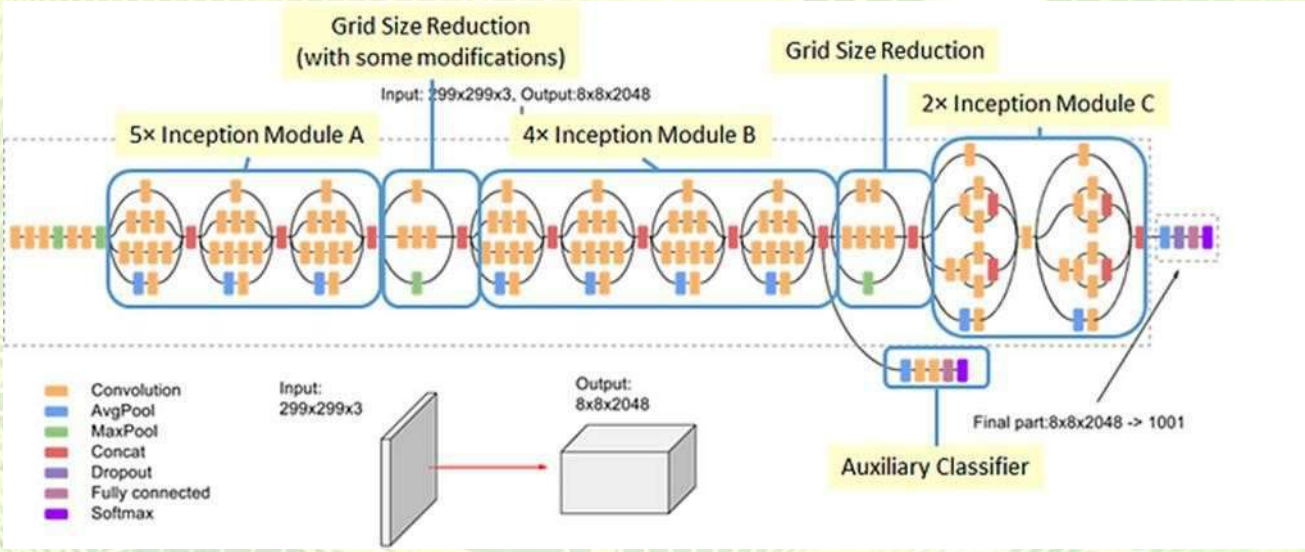


Fig. 2: Inception V3 Architecture.

ISSN 3030-3907



Fig. 3: Customized VGG16 Architecture.

A customized VGG16-based DCNN is adopted for transfer learning, then strengthened by an edge-enhancement preprocessing stage to better capture knee joint boundaries and OA-related structural changes.

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.

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.

ISSN 3030-3907

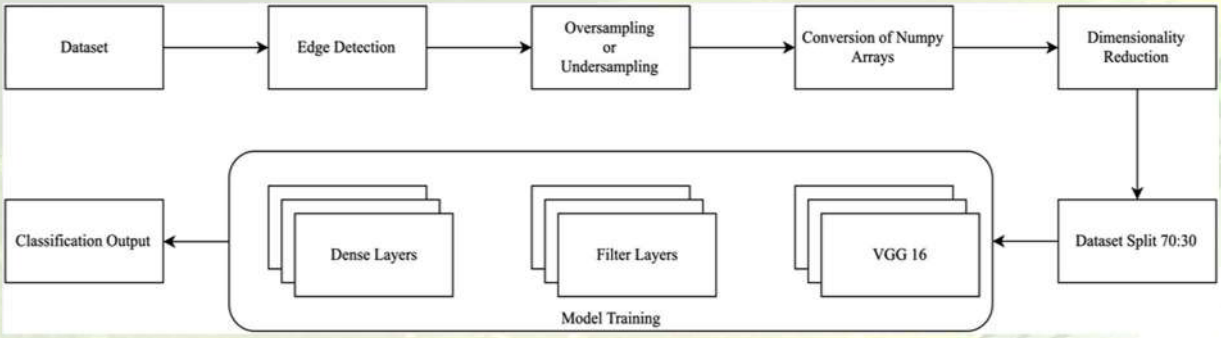


Fig. 4: Proposed Algorithm Architecture for Edge Detection.

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:

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

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:

$$PSNR = 20log_{10}(\frac{MAX_f}{\sqrt{MSE}}) \quad \text{Eq. (2)}$$

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

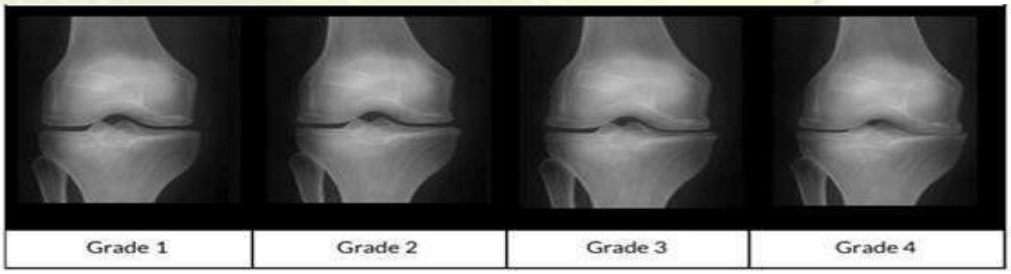
5. Experimental Setup

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

ISSN 3030-3907

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

ISSN 3030-3907

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]:

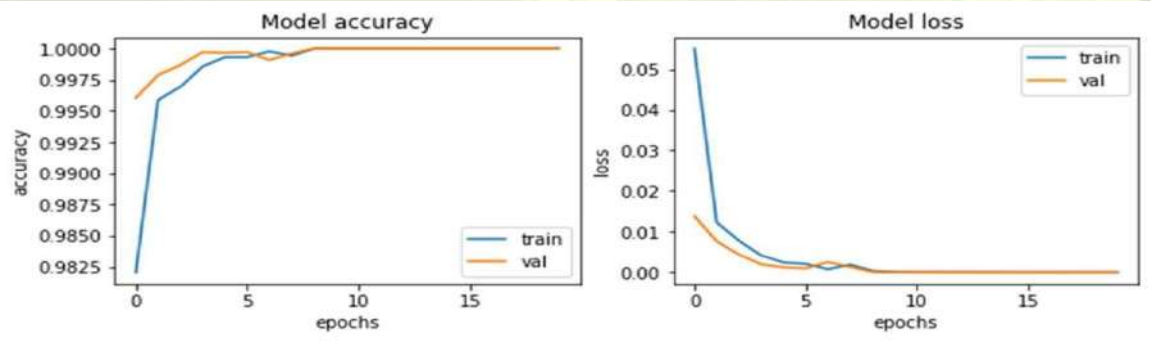
1. Sensitivity (SEN): This metric represents the proportion of true positive cases that were correctly identified by the classifier.
2. Specificity (SPF): This metric represents the proportion of true negative cases that were correctly identified by the classifier.
3. Accuracy (ACC): This metric represents 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.

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

ISSN 3030-3907

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.

ISSN 3030-3907

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, 0.7650.83 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

ISSN 3030-3907

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.

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

ISSN 3030-3907

3. Y. Zhang and J. M. Jordan, “Epidemiology of osteoarthritis,” Clinics in Geriatric Medicine, vol. 26, no. 3. Elsevier, pp. 355–369.
4. 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.
5. 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.
6. K. Taneja et al., “MRI features of the anterolateral ligament of the knee,” Skeletal Radiol., vol. 44, no. 3, pp. 403–410.
7. 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.
8. 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.
9. 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.

ISSN 3030-3907