Mühendislik Bilimleri ve Tasarım Dergisi
13(1), 325-339, 2025
e-ISSN: 1308-6693

Journal of Engineering Sciences and Design
DOI: 10.21923/jesd.1608509

Araştırma Makalesi Research Article

# TRANSFER LEARNING-BASED CLASSIFICATION OF KNEE OSTEOARTHRITIS SEVERITY FROM X-RAY IMAGES

**Miyade MAHFUS[1], Mustafa TOSUN[1], Hanife GÖKER[2*]**

[1] Kütahya Dumlupınar University, Faculty of Simav Technology, Department of Electrical Electronics Engineering, Kütahya, Türkiye
[2] Gazi University, Health Services Vocational College, Ankara, Türkiye

**Keywords**

*Deep Learning, Image Processing, Transfer Learning, ResNet101, Knee Osteoarthritis.*

**Abstract**

Knee osteoarthritis (KOA) a degenerative, long-term joint condition that, more often than not, affects the elderly and is characterized by articular cartilage degradation. Appropriate treatment and early analysis are essential for sickness control. However, traditional diagnostic methods for classifying KOA from X-ray images require laborious expertise and, unfortunately, have a large margin of error. This study presents an image processing-based solution for multi-classification KOA severity from X-ray images using the Bilateral filter, contrast-limited adaptive histogram equalization (CLAHE), and transfer learning models. The CLAHE method improved image quality, while the Bilateral filter enhanced details and minimized blurriness in X-ray images. KOA image dataset consists of 9786 knee images and five class labels. The performances of transfer learning models including AlexNet, ResNet101, DenseNet201 and VGG19 were compared. The ResNet101 model emerged as the most effective, achieving a kappa statistic of 0.970, weighted F1-score of 0.978, and an overall accuracy of 97.85%. This model's high accuracy and precision make it a dependable and objective diagnostic solution.

# DIZ OSTEOARTRITI ŞIDDETININ X-RAY GÖRÜNTÜLERINDEN TRANSFER ÖĞRENME TABANLI SINIFLANDIRILMASI

**Anahtar Kelimeler**

*Derin Öğrenme, Görüntü İşleme, Transfer Öğrenme, ResNet101, Diz Osteoartriti.*

**Öz**

Diz osteoartriti (KOA), çoğunlukla yaşlıları etkileyen ve eklem kıkırdağı bozulmasıyla karakterize dejeneratif, uzun vadeli bir eklem durumudur. Hastalık kontrolü için uygun tedavi ve erken analiz kritiktir. Bununla birlikte, X-ray görüntülerinden KOA sınıflandırması için geleneksel tanı yöntemleri uzmanlık gerektirmektedir, yorucudur ve maalesef büyük bir hata payına sahiptir. Bu çalışma, Bilateral filtresi, kontrast sınırlı adaptif histogram eşitleme (CLAHE) ve transfer öğrenme modelleri kullanarak X-ray görüntülerinden KOA şiddetini sınıflandırmak için görüntü işleme tabanlı bir çözüm sunmaktadır. CLAHE yöntemi görüntü kalitesini iyileştirirken, Bilateral filtresi X-ray görüntülerindeki ayrıntıları iyileştirerek bulanıklığı en aza indirmiştir. KOA görüntü veri seti 9786 diz görüntüsü ve beş sınıf etiketinden oluşmaktadır. AlexNet, ResNet101, DenseNet201 ve VGG19 dahil olmak üzere transfer öğrenme modellerinin performansları karşılaştırıldı. ResNet101 modeli 0,970 kappa istatistiği, 0,978 ağırlıklı F1-skoru ve %97,85 genel doğruluk elde ederek en etkili model olarak ortaya çıkmıştır. Bu modelin yüksek doğruluğu ve kesinliği onu güvenilir ve objektif bir tanı çözümü olduğunu göstermektedir.

**Cite**

Mahfus, M., Tosun, M., Göker, H., (2025). Transfer Learning-Based Classification of Knee Osteoarthritis Severity from X-Ray Images, Journal of Engineering Sciences and Design, 13(1), 325-339.

**Author ID (ORCID Number)**

M. Mahfus, 0009-0002-6358-3680
M. Tosun, 0000-0001-7167-4561
H. Göker, 0000-0003-0396-7885

**Article Process**

| | |
|---|---|
| **Submission Date** | 27.12.2024 |
| **Revision Date** | 11.03.2025 |
| **Accepted Date** | 16.03.2025 |
| **Published Date** | 20.03.2025 |

* Corresponding author: gokerhanife@gazi.edu.tr, +90-312-484 56 35

# TRANSFER LEARNING-BASED CLASSIFICATION OF KNEE OSTEOARTHRITIS SEVERITY FROM X-RAY IMAGES

Miyade MAHFUS [1], Mustafa TOSUN[1], Hanife GÖKER[2†]

[1] Kütahya Dumlupınar University, Faculty of Simav Technology, Department of Electrical Electronics Engineering, Kütahya, Türkiye
[2] Gazi University, Health Services Vocational College, Ankara, Türkiye

## Highlights

- An automatic early diagnosis system for multi-classification of knee osteoarthritis severity from X-ray images was proposed
- Comparative analysis of transfer learning models for the diagnosis of KOA severity was performed
- Advanced image processing methods, including Bilateral filter and CLAHE, were used
- The proposed transfer learning model is a reliable and objective diagnostic solution and shows potential for clinical use.

## Graphical Abstract (If applicable)

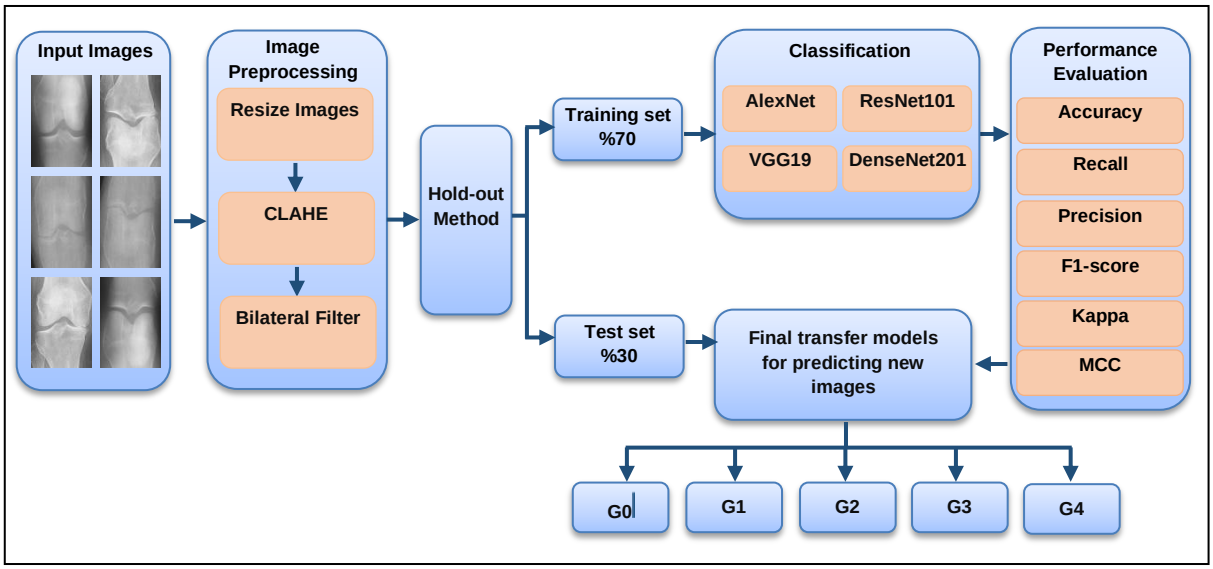


**Figure.** Flow-chart of the Proposed Method

## Purpose and Scope

KOA is a long-term joint disorder caused by the deterioration of the articular cartilage. Although early diagnosis is essential for controlling the disease, which especially affects the elderly, traditional diagnostic methods for KOA are laborious, require expertise, and, unfortunately, have a large margin of error. In this study, we proposed an image processing-based solution for the multi-classification of KOA severity from X-ray images using the Bilateral filter, CLAHE, and transfer learning models.

## Methodology

The transfer learning-based method consists of a) image preprocessing stages, b) dividing into training and test sets, and c) comparing transfer learning algorithms according to performance evaluation criteria for the multi-classification of KOA severity. Firstly, the X-ray images were resized based on the default input sizes for transfer learning models in the image preprocessing stage. The KOA severity grading dataset contains 9786 knee images. The images were resized to 224x224 for RestNet101, VGG19, DenseNet201, and 227x227 for AlexNet. The CLAHE method was employed to enhance contrast and the Bilateral filter reduced blurriness while sharpening

† Corresponding author: gokerhanife@gazi.edu.tr, +90-312-484 56 35

image details. Then, the KOA dataset was split into training and test sets using the hold-out method. The performances of DenseNet201, AlexNet, ResNet101, and VGG19 models were compared. Finally, the performance of these transfer learning models was assessed using model performance metrics.

**Experimental Results**

Transfer learning models; AlexNet, ResNet101, DenseNet201, and VGG19 were compared according to model performance metrics. Experimental results show that ResNet101 model combined with CLAHE and Bilateral filter can be used to accurately classify KOA severity from X-ray images. The transfer learning based model achieved 2872 correctly labeled X-ray images. This outperformed the other models, with DenseNet201, AlexNet, and VGG19 following. The correctly labeled X-ray images were 2872 for Resnet101, the correctly labeled images were 2754 for DenseNet201, the correctly labeled images were 2649 for AlexNet, and the correctly labeled images were 2636 for VGG19. The transfer learning model combined with the CLAHE, Bilateral filter, and ResNet101 attained the maximum performance with an overall accuracy of 97.85%. The transfer learning model was followed by DenseNet201 (93.83%), AlexNet (90.26%), and VGG19 (89.81%) models, respectively. ResNet101 achieved 0.970 kappa statistic, 0.978 weighted-F1 score, and 97.85% overall accuracy. In addition, the performance results of CLAHE, Bilateral filter, and ResNet101 were 0.990 recall, 0.988 precision, 0.989 F1-score, 0.993 specificity, and 982 MCC for the G0 label; 0.981 recall, 0.984 precision, 0.983 F1-score, 0.995 specificity, and 979 MCC for the G1 label; 0.975 recall, 0.975 precision, 0.975 F1-score, 0.991 specificity, and 966 MCC for the G2 label; 0.973 recall, 0.966 precision, 0.970 F1-score, 0.996 specificity, and 965 MCC for the G3 label; 0.857 recall, 0.886 precision, 0.871 F1-score, 0.995 specificity, and 867 MCC for the G4 label. This model, which we propose in our study, provides an unbiased and reliable diagnostic tool.

**Research limitations**

This study has several limitations. Firstly, the KOA severity grading dataset consists of 9786 knee images. Although the dataset is comprehensive, images are obtained using the X-ray imaging technique. Using different image techniques, such as magnetic resonance imaging, ultrasound, or computed tomography, will increase the robustness and accuracy of the KOA severity classification. Secondly, even though using transfer learning models, including AlexNet, ResNet101, DenseNet201, and VGG19, creates a robust structure, comparisons can also be made using different transfer learning algorithms. For future studies, the power and robustness of the proposed model can be investigated using different medical images.

**Practical implications**

In this study, image processing-based model was presented to support the clinicians using CLAHE, bilateral filter, and transfer learning models for the multi-classification of KOA severity. The study has practical contributions to developing decision support systems that analyze medical images. Using transfer learning algorithms facilitates the workload of experts and can provide objectivity and rapid decisions based on experimental results. Unlike previous studies, this study conducted a comprehensive and comparative analysis procedure on image processing and transfer learning models. The practical contributions of the image-processing model are delineated below: (i) The ResNet101 model combined with CLAHE and Bilateral filter for multi-classification KOA severity showed the best performance. The experiment results demonstrated that the transfer learning model achieved a promising performance compared to other methods in the literature with a higher accuracy of 97.85%. (ii) The transfer learning model for multi-classification of KOA severity can speed up the diagnostic procedure and provide time efficiency for clinicians. (iii) Early detection of KOA severity can facilitate well-timed interventions, decelerate disease progression, and enhance patient outcomes in clinical practice.

**Originality**

The Bilateral filter, CLAHE and ResNet101 transfer learning model had a very high accuracy of 97.85%. The advanced image preprocessing techniques such as CLAHE and Bilateral filter were implemented. In addition to both high accuracy and advanced preprocessing, the model is clinically impactful. The transfer learning-based model was proposed to improve patient outcomes significantly and enhance healthcare efficacy. According to the experimental results, it is an accurate method of early detection of KOA. From a public health viewpoint, it could reduce the disease burden by better detecting KOA severity earlier.

## 1. Introduction

KOA, a chronic and progressive joint disease, is characterized by the deterioration of joint cartilage tissue. It affects millions of people worldwide, particularly the elderly population (Geng et al., 2023). With the growing older population, the superiority of KOA is anticipated to increase significantly, with more than 20% of people predicted to be at risk by 2030 (Ortman et al., 2014). Moreover, it is estimated that 130 million people worldwide will suffer from KOA by 2050 (Wang et al., 2021). In developed countries, KOA is the most common cause of disability related to joint disorders, especially in the knees and hips (Grazio & Balen, 2009). The disease is marked by the deterioration of articular cartilage, bone remodeling, and joint inflammation, involving complex molecular, anatomical, and physiological changes that extend beyond simple "wear and tear". Patients with KOA face a higher risk of mortality compared to the general population, with significant contributing factors such as diabetes, cardiovascular disease, and walking disabilities (Haidari, 2011). Given its widespread prevalence and debilitating impact, KOA remains a significant public health challenge globally. The disease reduces the quality of life for millions and places a considerable economic burden on healthcare systems due to increased medical expenses and loss of productivity (Islam & Rony, 2024). Early intervention can slow the progression of KOA, reduce pain, and enhance joint functionality, emphasizing the importance of timely and accurate diagnosis.

KOA is diagnosed with the progression of symptoms, which leads to a delay in disease management. Early detection of the disease becomes difficult, leading to late treatment initiation and worsening symptoms. Patients frequently put off getting care until they experience distressing symptoms. Therefore, healthcare providers might not see KOA patients until they exhibit severe stiffness, pain, and functional limitations (Langworthy et al., 2024). The increasing need for automated grading of KOA severity stems from the imperative for swifter and early detection, driven by a lack of radiologists and the laborious process of analyzing knee X-ray images, especially in remote locations (Kishore et al., 2024). KOA, which cannot be diagnosed early, reduces the quality of life of individuals and can cause mobility restrictions and dependency in daily activities in later stages. Traditional diagnostic methods are based on clinical assessment and radiological imaging techniques, and modern approaches include biomarkers and advanced imaging technologies. However, deep learning is widely used in modern diagnostic methods for KOA severity detection (Zeng et al., 2023). These models also produce more accurate, faster, and more objective results for analyzing complex data. In particular, transfer learning models, which are a version of deep learning, offer an innovative and attractive solution for KOA severity classification by enabling the application of knowledge from a pre-trained network to new data. Transfer learning models can directly work through challenging and large-scale image data, keeping the understanding learned from solving a task and later utilizing it once encountering another task (Göker, 2024). Transfer learning can be repurposed, where features and weights are developed from pre-trained models. So, using prior knowledge enables the development of models that achieve higher performance and less training with faster learning.

Transfer learning-based medical image processing has achieved promising results in improving diagnostic accuracy for KOA severity in recent years. Kokkotis et al. (2020) presented a machine learning-based model for the KOA classification using support vector machine (SVM) and k-nearest neighbors (K-NN) algorithms. The SVM algorithm achieved 74.07% accuracy (Kokkotis et al., 2020). Also, Abedin et al. (2019) used the convolutional neural network (CNN) algorithm for KOA severity classification using X-ray images. The root mean squared error for the CNN was 0.77 (Abedin et al., 2019). Guan et al. (2022) proposed a combination of deep learning and traditional machine learning algorithms. The traditional machine learning models, random forests (RF), logistic regression (LR), and artificial neural networks (ANN), were employed. In contrast, the CNN deep learning algorithm was used to classify KOA. The combined model achieved 80.9% specificity, 72.3% sensitivity, and 0.807 AUC (Guan et al., 2022). Similarly, Brahim et al. (2019) proposed a machine learning-based model for the KOA classification using Fourier filter, independent component analysis (ICA), and machine learning algorithms. The Fourier filter was used for image preprocessing, and ICA was employed to decrease dimensionality. Then, the performances of Naive Bayes and RF algorithms were compared. The RF algorithm, which achieved the highest success, had 80.65% specificity, 87.15% sensitivity, and 82.98% accuracy (Brahim et al., 2019). Jain et al. (2024) presented an attentive multi-scale deep CNN for KOA severity classification. They used the high-resolution network (HRNet), and the HRNet algorithm relies on deep learning models to capture multi-scale capabilities of images. The method achieved the best accuracy of 71.74% (Jain et al., 2024). Solak (2024) compared VGG16, Xception, EfficientNetB0, DenseNet20, and ResNet-50 transfer learning models for KOA severity classification. The DenseNet201 model, which has the highest performance, obtained 87.7% accuracy, 87.2% F1-score, and 0.75

kappa statistics (Solak, 2024). Although transfer learning for image processing is faster and more objective than traditional methods, much more work needs to be done before transfer learning-based models can be used in the clinic. Moreover, improvement and validation of transfer learning-based models through real-world applications are important for the robustness, reliability, and effectiveness of the models.

In this study, we proposed an image processing-based solution for the multi-classification of KOA severity from X-ray images using the Bilateral filter, CLAHE, and transfer learning models. The main contributions of this study to the literature can be summarized as follows:

i. A transfer learning-based model was proposed to classify KOA severity using image processing methods.
ii. Image contrast was increased with CLAHE, and the use of Bilateral filter improved the detail clarity and reduced blur in X-ray images.
iii. Transfer learning models; AlexNet, ResNet101, DenseNet201, and VGG19 were compared. ResNet101 achieved the best performance with 0.970 kappa statistic, 0.978 weighted F1 score, and 97.85% overall accuracy.
iv. Experimental results show that the ResNet101 model combined with CLAHE and Bilateral filter can accurately classify KOA severity from X-ray images. This model provides an unbiased and reliable diagnostic tool.

## 2. Materials and Methods

### 2.1. Proposed Method

The proposed transfer learning-based method consists of a) image preprocessing stages, b) dividing into training and test sets, and c) comparing transfer learning algorithms according to performance evaluation criteria for the multi-classification of KOA severity. Firstly, the X-ray images were resized based on the default input sizes for transfer learning models in the image preprocessing stage. The KOA severity grading dataset contains 9786 knee images. The images were resized to 224x224 for RestNet101, VGG19, DenseNet201, and 227x227 for AlexNet. The CLAHE method was employed to enhance contrast, and the Bilateral filter reduced blurriness while sharpening image details. Then, the KOA dataset was split into training and test sets using the hold-out method. The performances of DenseNet201, AlexNet, ResNet101, and VGG19 models were compared. Finally, the performance of these transfer learning models was assessed using model performance metrics. Figure 1 shows the flowchart of the proposed method.

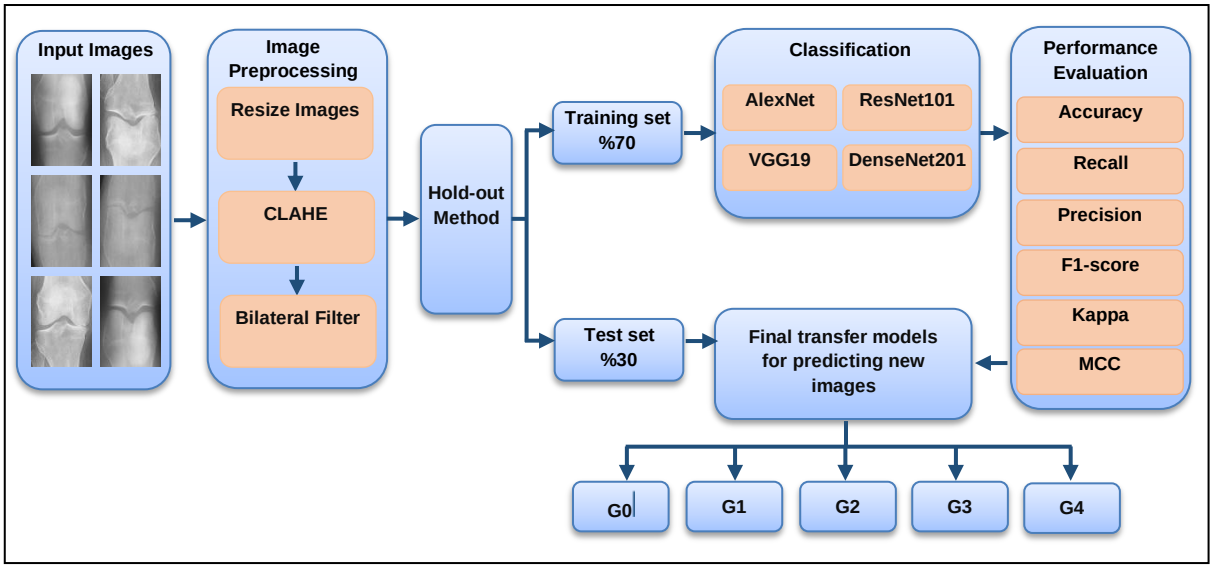


**Figure 1.** The flowchart of the proposed method

### 2.2. Dataset

The knee images were taken from the KOA severity dataset (Tiwari et al., 2022). The KOA dataset of 9786 X-ray images, including the left and right knee, was graded according to Kellgren-Lawrence (KL). The original dataset was categorized into five grades: normal, doubtful, minimal, moderate, and severe. There were 3857 images in normal grades (G0), 1770 images in doubtful (G1), 2578 images in minimal (G2), 1286 images in moderate (G3),

and 295 images severe (G4). The size of each image is 224 × 224. Table 1 shows the characteristics of the KOA severity dataset.

**Table 1.** Characteristics of the KOA Severity Dataset

| GRADE | NUMBER OF X-RAYS | KL CLASSES DESCRIPTION |
|---|---|---|
| **G0** | 3857 | No signs of disease are seen on the knee joint X-ray. The joint appears healthy |
| **G1** | 1770 | There may be bone spurs called osteophytes and slight narrowing of the joint space. However, these findings do not definitively indicate osteoarthritis |
| **G2** | 2578 | Osteophytes are definitely seen and there may be narrowing of the joint space. |
| **G3** | 1286 | Steophytes, significant narrowing of the joint space and slight hardening of the bones (sclerosis) are seen |
| **G4** | 295 | Osteophytes, severe narrowing of the joint and widespread hardening of the bones are seen. This indicates advanced osteoarthritis. |
| **Total** | 9786 | |

The KOA dataset was split into training set (6851 knee images) and test set (2935 knee images) in the ratio of 70:30. Table 2 shows the distribution of the dataset.

**Table 2.** Distribution of the Dataset

| SET\ KL GRADES | G0 | G1 | G2 | G3 | G4 | Total |
|---|---|---|---|---|---|---|
| **Training** | 2700 | 1239 | 1805 | 900 | 207 | 6851 |
| **Test** | 1157 | 531 | 773 | 386 | 88 | 2935 |
| **Total** | 3857 | 1770 | 2578 | 1286 | 295 | 9786 |

Figure 2 shows samples of knee images representing each label (normal, doubtful, minimal, moderate, and severe) in the KOA dataset.

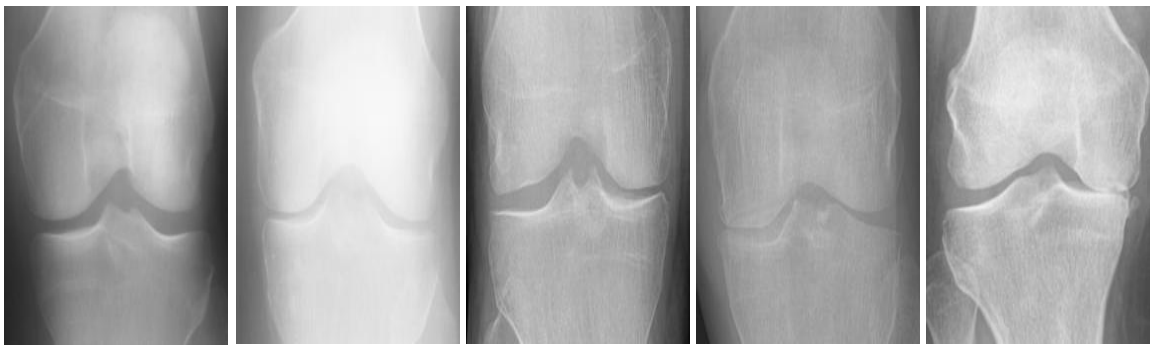

**Figure 2.** Samples of X-ray knee images per label, (i) normal, (ii) doubtful, (iii) minimal, (iv) moderate, (v) severe

### 2.3. Image Preprocessing

X-ray images often have various quality problems such as noise, low contrast, and motion distortions; this situation negatively affects the performance of transfer learning models. Image preprocessing techniques are essential to address these challenges and enhance the visibility of crucial anatomical structures like the knee joint. During the image preprocessing stage, the images were resized based on the default input sizes used by transfer learning models. The knee images were resized to 224x224 for RestNet101, VGG19, DenseNet201, and 227x227 for AlexNet. Following resizing, CLAHE equalization was applied to enhance the images, thereby improving the visibility of key features (Ahmed et al., 2022). This method divides the image into smaller blocks, performs equalization separately on each block, prevents excessive contrast enhancement, and better preserves image details.

Then, the Bilateral filter was performed to decrease noise in the images. The bilateral filter can smooth images, especially while preserving the edges. This dual capability is achieved through the combination of spatial and range kernels, which weigh neighboring pixels based on their geometric proximity and intensity similarity (Li & Duan, 2022). The filter works by substituting a weighted average of a pixel's neighbors for the pixel's value; the weights are established by the pixels' spatial separation and intensity difference (Yang et al., 2024). The bilateral filter is formulated as below (Singh et al., 2023):

$$\text{Bilateral filter}\,(x,y) = \frac{\sum_{x'}\sum_{y'} I(x',y')\,g_{\sigma_s}(x-x',y-y')\,g_{\sigma_r}(I(x,y)-I(x'-y'))}{\sum_{x'}\sum_{y'} g_{\sigma_s}(x-x',y-y')\,g_{\sigma_r}(I(x,y)-I(x'-y'))} \quad (1)$$

where the pixels neighboring the position (x, y) in an image are referred to as (x′, y′).

$$g_{\sigma_s}(x,y) = exp\left(-\frac{(x^2+y^2)}{2\sigma_s^2}\right) \quad (2)$$

$$g_{\sigma_r}(\hat{a}) = exp\left(\frac{\hat{a}^2}{2\sigma_r^2}\right) \quad (3)$$

In Equations 2 and 3, σr denotes the minimum amplitude, while σs represents the size of the spatial kernel.

### 2.4. Transfer Learning Models

Transfer learning is a pre-training method that uses information from trained data to create new models. Transfer learning is based on the premise that knowledge in the form of a model from a source task can be transferred to support learning in related target tasks, which is particularly beneficial when labeled information is scarce or costly to acquire (Mahmoud et al., 2024). Transfer learning algorithms use supervised pre-trained models with large datasets to allow activation mapping to other problems. Transfer learning algorithms such as AlexNet, ResNet, VGG19, and DenseNet are commonly used in medical image analysis applications (Kim et al., 2022). The optimum parameters of the transfer learning models are summarized in Table 3.

**Table 3.** The Optimum Parameters of the Transfer Learning Models

| HYPER-PARAMETERS | ALEXNET | RESNET101 | VGG19 | DENSENET201 |
|---|---|---|---|---|
| **Input size** | 227x227 | 224x224 | 224x224 | 224x224 |
| **MaxEpochs** | 30 | 30 | 30 | 30 |
| **MiniBatchSize** | 16 | 16 | 16 | 16 |
| **Optimizer** | sgdm | sgdm | sgdm | sgdm |
| **ValidationFrequency** | 3 | 3 | 3 | 3 |
| **InitialLearnRate** | 1e-4 | 1e-4 | 1e-4 | 1e-4 |
| **Momentum** | 0.9 | 0.9 | 0.9 | 0.9 |
| **WeightLearnRateFactor** | 20 | 20 | 20 | 20 |
| **BiasLearnRateFactor** | 20 | 20 | 20 | 20 |
| **L2Regularization** | 0.005 | 0.005 | 0.005 | 0.005 |

The optimum hyper-parameters for transfer learning models were selected by comparing the results of several experiments that were carried out. For the “MaxEpochs” parameter, the values “30” and “60” were compared, and the value “30” was chosen, giving the best result. The “sgdm”, “adam”, “rmsprop” values of the “optimizer” parameter were compared, and "sgdm", was selected. The “16” and “32” values were compared for the “MiniBatchSize” parameter, and the “16” value was selected.

#### 2.4.1. AlexNet

The AlexNet consists of five “convolutional layers” and three “fully connected layers”. It employs Rectified Linear Unit (ReLU) activation and dropout to avoid overfitting and enhance performance (Tang et al., 2023). Compared to the traditional CNN architecture, AlexNet was designed with a deeper architecture, having more filter layers including stacked convolutional layers. The AlexNet is given in Figure 3 (Karim et al., 2022).

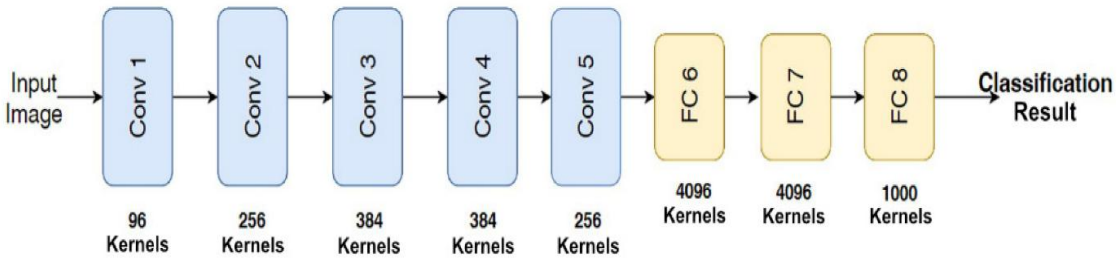


**Figure 3.** The AlexNet neural network

AlexNet employs ReLU as the activation function for its non-linear elements, in contrast to earlier neural networks that primarily utilized tanh or sigmoid activations. Figure 4 shows the definitions of activation functions (Karlik & Olgac, 2011).

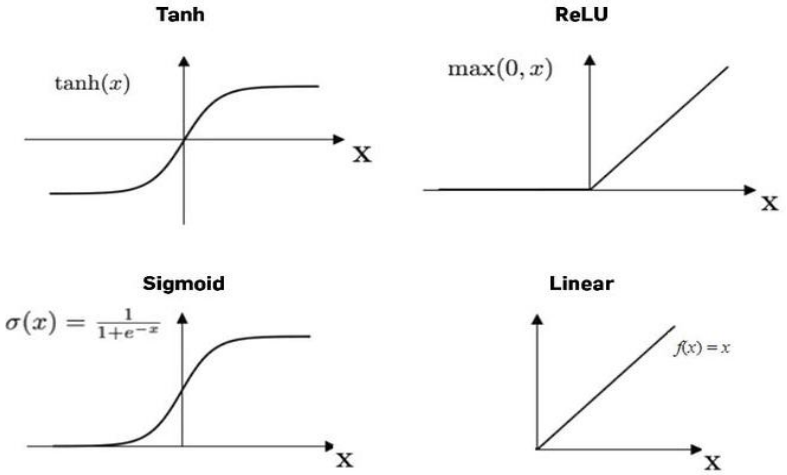


**Figure 4**. Functions description

The first, second, and fifth convolutional layers are also succeeded by max-pooling layers. One of the notable advantages of AlexNet is its use of local response normalization, which helps prevent the saturation of activation functions.

### 2.4.2. ResNet101

The ResNet101 model, a variant of the Residual Network architecture, includes 101 layers and uses skip connections during backpropagation, which allows gradients to flow through the network more effectively. ResNet101 has gained prominence because of its ability to facilitate the training of deep neural networks while reducing the vanishing gradient problem. This innovation has led to significant improvements in various image classification tasks, making ResNet101 a popular choice in practical applications. The ResNet101 performance of imaging analysis is quite strong. The input image sizes for the ResNet101 model are 224×224. The ResNet101 is given in Figure 5 (Tong et al., 2020):

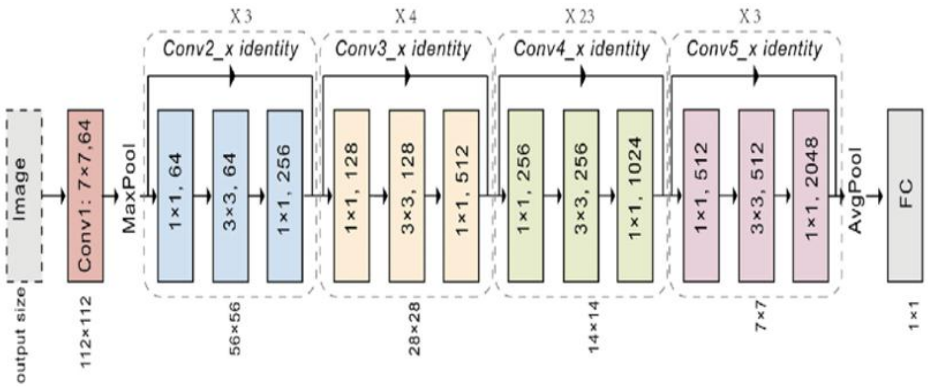


**Figure 5.** The ResNet101 neural network

### 2.4.3. VGG19

VGG19 is a deep CNNs architecture consisting of 19 layers, including 16 convolutional layers, 3 fully connected layers, and a softmax classification layer. VGG19 maintains a simple and consistent structure using only 3x3 filters in all convolutional layers. These small filters allow building a deeper network and learning more complex features. The ReLU activation function is used in every convolutional and fully connected layer and facilitates faster learning and reduces computational costs, thus enhancing model performance. Max-pooling layers applied after each block reduce spatial dimensions and computational load while preserving prominent features. With pre-trained weights on large datasets, only the upper layers need to be retrained for new tasks, which provides a significant advantage in limited data conditions. The VGG19 is given in Figure 6 (Khattar & Quadri, 2022):

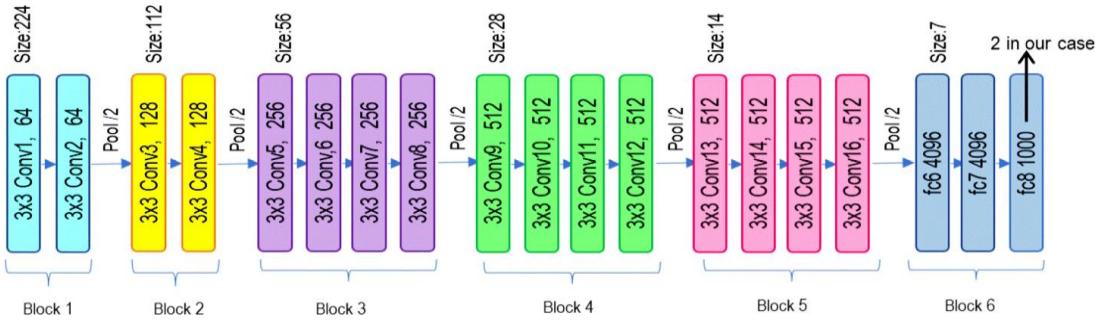


**Figure 6.** The VGG19 neural network

### 2.4.4. DenseNet201

Dense Convolutional Network (DenseNet) is a transfer learning model based on the principle that each layer receives the feature maps from every layer before it. This form of connection increases the flow of information. Dense connections decrease the vanishing gradient problem, which is frequently seen in deep networks. This contributes to a more stable training process, especially in very deep networks (Jung et al., 2024). DenseNet201 consists of 201 layers in total, and dense connections between these layers enable the network to learn deeper and more complex features. This characteristic of DenseNet contributes to more efficient use of parameters, increases the model performance, and also reduces overfitting. The input image sizes for the DenseNet201 are 224×224. The DenseNet201 is given in Figure 7 (Kumar et al., 2021):

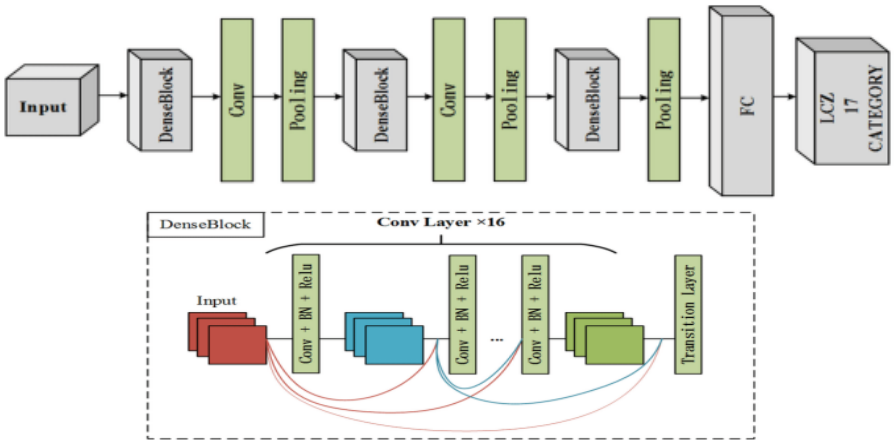


**Figure 7.** The DenseNet201 neural network

## 2.5. Performance Evaluation Metrics

The performance of transfer learning models was assessed utilizing performance evaluation metrics, including overall accuracy, precision, specificity, recall, Matthew's correlation coefficient (MCC), F1-score, weighted F1-score, and the kappa statistic. True Positive (TP) gives the number of data correctly classified by the classifier model from the positive class; False Positive (FP) gives the number of data actually labeled as positive by the classifier model as a result of the classification of data belonging to the negative class. True Negative (TN) gives the number of data correctly classified by the classifier model from the negative class; False Negative (FN) gives the number of data actually labeled as negative by the classifier model as a result of the classification of data belonging to the positive class (Çelik et al., 2023). The equations used to derive these performance metrics are delineated in equations 4-11.

$$\text{Overall Accuracy} = \text{Number of accurately classified samples/ Total number of samples} \quad (4)$$

$$\text{Precision} = TP / (FP + TP) \quad (5)$$

$$\text{Recall} = TP / (FN + TP) \quad (6)$$

$$\text{Specificity} = TN/(TN + FP) \quad (7)$$

$$F1-score = 2 * Precision * Recall / (Precision + Recall) \quad (8)$$

$$Weighted\text{-}F1\ scores = Weighted\text{-}averaged\ of\ F1\text{-}scores \quad (9)$$

$$MCC = (TP * TN - FN * FP)/\sqrt{(FP + TP) * (TN + FN) * (FN + TP) * (TN + FP)} \quad (10)$$

The kappa statistic is performed to assess inter-rater reliability. It measures the degree of agreement between two or more raters by correcting for any agreement that might occur by chance. The kappa value ranges from -1 to 1, where 1 demonstrates perfect agreement, 0 demonstrates no agreement beyond chance, and negative values demonstrate agreement less than would be expected by chance. The formula for the kappa statistic is provided in Equation 11 (Cantor, 1996).

$$Kappa = (P\ o - P\ e) / (1 - P\ e) \quad (11)$$

$$Po = \sum_{i=1}^{k} CM_{ii} / N$$

$$Pe = \sum_{i=1}^{k} \left( \sum_{i=1}^{k} CM_{ij}\ x\ \sum_{j=1}^{k} CM_{ji} / N^2 \right)$$

Pe denotes the expected agreement and Po denotes the observed agreement. $CM_{ii}$ are the diagonal values of the confusion matrix, representing correct classifications, k denotes the number of labels, and N denotes the total number of samples.

## 3. Experimental Results and Discussion

In this study, the transfer learning-based models were compared to classify the KOA severity. The KOA severity grading dataset contains 9786 knee images. The images were resized to 224x224 for RestNet101, VGG19, DenseNet201, and 227x227 for AlexNet. The CLAHE was performed for image quality, while the bilateral filter was performed to reduce the noise of images and enhance details. Then, the model performance was evaluated using the holdout method. The KOA severity grading dataset was split into 6851 images training set (%70) and 2935 images test set (%30). Finally, the performances of RestNet101, VGG19, DenseNet201, and AlexNet transfer learning models were compared based on model performance metrics. The confusion charts of transfer learning models are shown Figure 8.

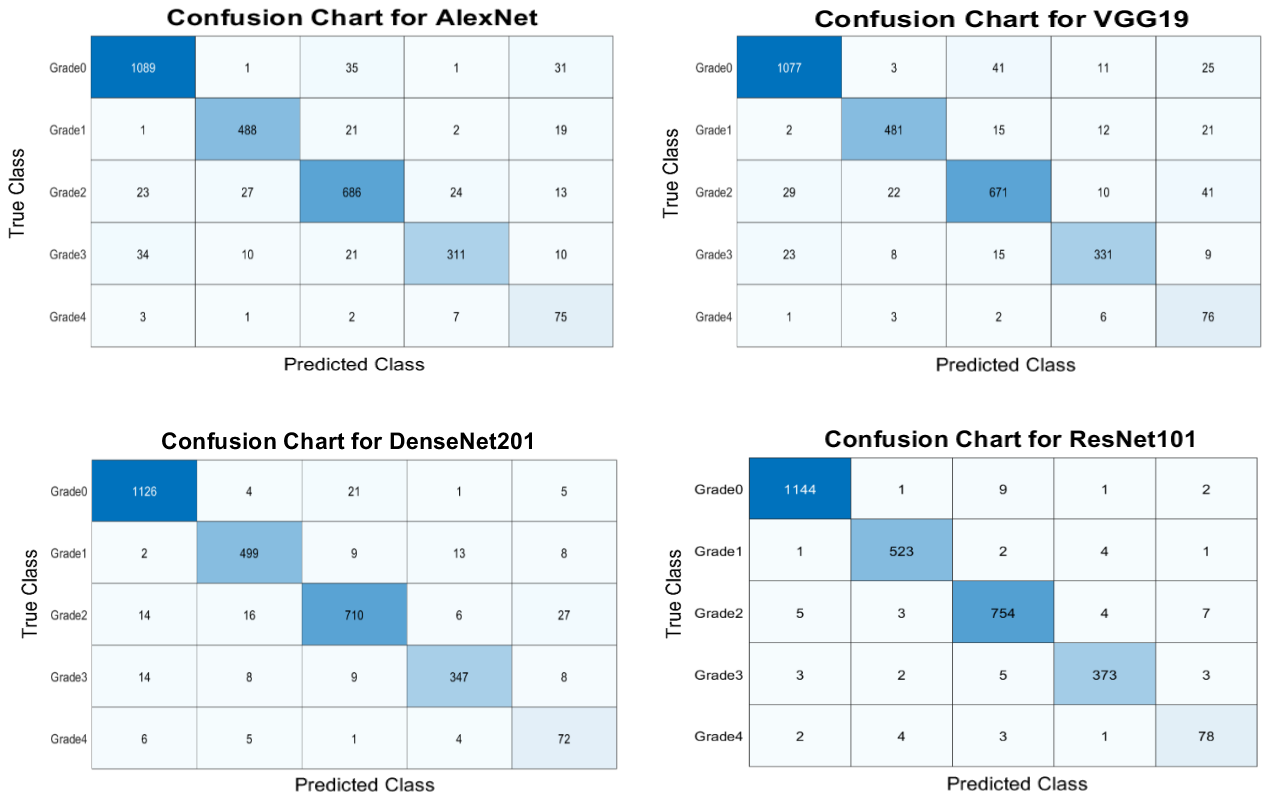


**Figure 8.** The confusion charts of transfer learning models

When examining Figure 8, the ResNet101 transfer learning model achieved 2872 correctly labeled X-ray images. This outperformed the other models, with DenseNet201, AlexNet, and VGG19 following. The correctly labeled X-ray images were 2872 for Resnet101, the correctly labeled images were 2754 for DenseNet201, the correctly labeled images were 2649 for AlexNet, and the correctly labeled images were 2636 for VGG19. Performance metrics, namely recall, precision, specificity, F1-score, kappa statistic, MCC, weighted-F1 score, and overall accuracy, were calculated to evaluate model performance comprehensively. The performance results of the transfer learning models are shown in Table 3.

**Table 3.** The Performance Results of the Transfer Learning Models

| MODELS | LABELS | RECALL | PRECISION | SPECIFICITY | F1-SCORE | MCC |
|---|---|---|---|---|---|---|
| **CLAHE + Bilateral Filter + AlexNet** | **G0** | 0.947 | 0.941 | 0.965 | 0.944 | 0.907 |
| | **G1** | 0.926 | 0.919 | 0.983 | 0.922 | 0.905 |
| | **G2** | 0.896 | 0.887 | 0.963 | 0.892 | 0.853 |
| | **G3** | 0.901 | 0.805 | 0.986 | 0.850 | 0.831 |
| | **G4** | 0.506 | 0.852 | 0.974 | 0.635 | 0.644 |
| **CLAHE + Bilateral Filter + VGG19** | **G0** | 0.951 | 0.930 | 0.969 | 0.941 | 0.903 |
| | **G1** | 0.930 | 0.905 | 0.985 | 0.917 | 0.900 |
| | **G2** | 0.901 | 0.868 | 0.966 | 0.884 | 0.844 |
| | **G3** | 0.894 | 0.857 | 0.984 | 0.875 | 0.857 |
| | **G4** | 0.441 | 0.863 | 0.966 | 0.584 | 0.602 |
| **CLAHE + Bilateral Filter + DenseNet201** | **G0** | 0.969 | 0.973 | 0.979 | 0.971 | 0.952 |
| | **G1** | 0.938 | 0.939 | 0.986 | 0.938 | 0.925 |
| | **G2** | 0.946 | 0.918 | 0.981 | 0.932 | 0.908 |
| | **G3** | 0.935 | 0.899 | 0.990 | 0.916 | 0.904 |
| | **G4** | 0.600 | 0.818 | 0.983 | 0.692 | 0.690 |
| **CLAHE + Bilateral Filter + ResNet101** | **G0** | 0.990 | 0.988 | 0.993 | 0.989 | 0.982 |
| | **G1** | 0.981 | 0.984 | 0.995 | 0.983 | 0.979 |
| | **G2** | 0.975 | 0.975 | 0.991 | 0.975 | 0.966 |
| | **G3** | 0.973 | 0.966 | 0.996 | 0.970 | 0.965 |
| | **G4** | 0.857 | 0.886 | 0.995 | 0.871 | 0.867 |

When the models' performances are examined in Table 3 and Table 4, the transfer learning model combined with the CLAHE, Bilateral filter, and ResNet101 attained the maximum performance with an overall accuracy of 97.85%. The transfer learning model was followed by DenseNet201 (93.83%), AlexNet (90.26%), and VGG19 (89.81%) models, respectively. In addition, the performance results of CLAHE, Bilateral filter, and ResNet101 were 0.990 recall, 0.988 precision, 0.989 F1-score, 0.993 specificity, and 982 MCC for the G0 label; 0.981 recall, 0.984 precision, 0.983 F1-score, 0.995 specificity, and 979 MCC for the G1 label; 0.975 recall, 0.975 precision, 0.975 F1-score, 0.991 specificity, and 966 MCC for the G2 label; 0.973 recall, 0.966 precision, 0.970 F1-score, 0.996 specificity, and 965 MCC for the G3 label; 0.857 recall, 0.886 precision, 0.871 F1-score, 0.995 specificity, and 867 MCC for the G4 label. The performance metrics are expected to be close to one because when the metrics are near one, it indicates that the model's performance is robust and high. When evaluating model performance, it is important to examine accuracy, kappa statistics, and weighted-F1 score. Table 5 presents the weighted-F1 score, kappa, and overall accuracy results.

**Table 4.** The Kappa, Weighted-F1 Score, and Overall Accuracy Results of the Transfer Learning Models

| MODELS | KAPPA | WEIGHTED-F1 SCORE | OVERALL ACCURACY |
|---|---|---|---|
| **CLAHE + Bilateral Filter + AlexNet** | 0.866 | 0.900 | 90.26% |
| **CLAHE + Bilateral Filter + VGG19** | 0.861 | 0.893 | 89.81% |
| **CLAHE + Bilateral Filter + DenseNet201** | 0.915 | 0.937 | 93.83% |
| **CLAHE + Bilateral Filter + ResNet101** | 0.970 | 0.978 | 97.85% |

The overall accuracy is the ratio of accurately classified samples to total samples. Therefore, overall accuracy alone is unreliable if there is no balance in the distribution across groups. The F1-score evaluates a trade-off between

recall and precision. The weighted F1 score is the weighted average of F1 scores. Therefore, the weighted score produces reliable results even if there is an imbalance in the distribution across groups. Kappa, which evaluates the agreement between two or more raters and adjusts for the agreement that occurs by chance, is an important measure that increases the reliability of the evaluation. The proposed transfer learning-based medical image processing has achieved promising results for KOA severity classification from X-ray images. Table 5 presents the comparative analysis of studies for KOA severity classification.

**Table 5.** The Comparative Analysis of Studies for KOA Severity Classification

| REFERENCE | IMAGES SIZE | IMAGE PREPROCESSING | CLASS SIZE | CLASSIFIER | ACCURACY (%) |
|---|---|---|---|---|---|
| Wahyuningrum et al. (2020) | 4737 images | Data augmentation, normalization, CLAHE, Region of Interest (ROI) | 5 class | CNN | 77.24% |
| Qali et al. (2021) | 998 images | - | 2 class | VGG16 | 69.23% |
| | | | | SVM | 77% |
| | | | | CNN | 90% |
| Ahmed & Mohammed (2022) | 1650 images | Resize images, data augmentation | 5 class | VGG16 | 87.27% |
| | | | | VGG19 | 89.69% |
| | | | | ResNet50 | 91.51% |
| Alshamrani et al. (2023) | 3836 images | 2D median filter, Gaussian smoothing techniques, sharpening filters, contrast stretching, and histogram equalization technique. | 2 class | ResNet50 | 90.63% |
| | | | | CNN | 90.95% |
| | | | | VGG16 | 92.17% |
| Goswami (2023) | 4130 images | Segmentation, contour detection, contrast enhancement | 5 class | CNN | 91.03% |
| Mohammed et al. (2023) | 9786 images | Segmentation, equalization | 5 class | MobileNetV2 | 67% |
| | | | | ResNet101 | 69% |
| | | | | VGG16 | 66% |
| | | | | VGG19 | 64% |
| | | | | InceptionResNetV2 | 63% |
| | | | | DenseNet121 | 64% |
| Rehman & Gruhn (2024) | 1650 images | Data augmentation (Horizontal flip, Vertical flip, Rotation -45, Rotation 90, Crop 0.1, Crop 0.2, Gaussian noise, Gamma contrast, Sigmoid contrast, Linear contrast, Channel shuffling, and Inverted colors) | 5 class | VGG19 | 84.14% |
| | | | | ResNet50 | 85.28% |
| | | | | VGG16 | 88.45% |
| | | | | CNN | 90.38% |
| | | | | VGG16+CNN | 93.27% |
| Nurmirinta et al. (2024) | 1213 images | - | 3 classes | Balanced Random Forest | 65.9% |
| **The Proposed Method** | 9786 images | Resize images, CLAHE, Bilateral filter | 5 class | VGG19 | 89.81% |
| | | | | Alexnet | 90.26% |
| | | | | Resnet101 | 93.83% |
| | | | | DenseNet201 | 97.85% |

When Table 5 is examined, CNN (Wahyuningrum et al., 2020; Qali et al., 2021; Alshamrani et al., 2023; Goswami, 2023; Rehman & Gruhn, 2024), VGG16 (Qali et al., 2021; Ahmed & Mohammed, 2022), VGG19 (Ahmed & Mohammed, 2022; Mohammed et al., 2023; Rehman & Gruhn, 2024), MobileNetV2 (Mohammed et al., 2023), ResNet50 (Ahmed & Mohammed, 2022; Alshamrani et al., 2023), and DenseNet121 (Mohammed et al., 2023) algorithms are used. While classic machine learning techniques like Balanced Random Forest are also used, the majority of the studies focus on deep learning methodologies (Nurmirinta et al., 2024). Because deep learning methods can automatically extract complicated patterns from pictures, they have demonstrated significant performance increases, making them appropriate for medical imaging applications like as KOA severity assessment. Regarding class distribution, numerous studies have used the Kellgren and Lawrence rating scale for either binary classification (two classes) (Qali et al., 2021; Alshamrani et al., 2023) or multi-classification (three or five classes) (Wahyuningrum et al., 2020; Ahmed & Mohammed, 2022; Goswami, 2023; Rehman & Gruhn, 2024; Nurmirinta et al., 2024). Since differentiating between two classes (healthy and KOA), which is binary classification, is simpler than categorizing the severity of KOA into numerous stages, binary classification tasks typically yield higher accuracy rates. However, multi-classification offers more thorough diagnostic data, which is essential in medical settings. The experimental results support the efficacy of transfer learning models for classifying KOA severity. ResNet101, achieved the highest accuracy (97.85%) in this study. With its high accuracy, ResNet101 can assist radiologists by improving the speed and accuracy of KOA assessments. Furthermore, the model's exceptional performance was further boosted by the application of the bilateral filter and CLAHE, which

greatly improved the quality of the input images. These experimental results highlight the deep learning and image processing methods for medical imaging applications. The proposed transfer learning-based model may change the diagnosis and treatment of KOA severity, enhancing patient outcomes, clinical workflows, and the life quality of people.

## 4. Conclusion

In conclusion, we compared the performance of transfer learning models to classify KOA severity. Firstly, the X-ray images were resized to fit input-size images for transfer learning models. Then, we applied CLAHE for image enhancement, a technique that improves the contrast of the images, and the Bilateral filter was used to sharpen details and reduce blurriness in knee images. The Bilateral filter is a non-linear smoothing filter that reduces noise and preserves edges. Finally, transfer learning models were compared, including AlexNet, ResNet101, DenseNet201, and VGG19. The ResNet101 transfer learning model performed better than the others. The ResNet101 transfer learning achieved a kappa statistic of 0.970, a weighted F1-score of 0.978, and an overall accuracy of 97.85%. The experimental results show that the image processing-based transfer learning model is robust and reliable for diagnosing KOA severity from X-ray images. The study has practical contributions to developing decision support systems that analyze medical images. Using transfer learning algorithms facilitates experts' workload and can provide objectivity and rapid decisions based on experimental results. Unlike previous studies, this study conducted a comprehensive and comparative analysis procedure on image processing and transfer learning models. The practical contributions of the image-processing model are delineated below: (i) The ResNet101 model combined with CLAHE and Bilateral filter for multi-classification KOA severity achieved the best performance. The experiment results demonstrated that the transfer learning model achieved a promising performance compared to other methods in the literature with a higher accuracy of 97.85%. (ii) The transfer learning model for multi-classification of KOA severity can speed up the diagnostic procedure and provide time efficiency for clinicians. (iii) Early detection of KOA severity can facilitate well-timed interventions, decelerate disease progression, and enhance patient outcomes in clinical practice. This study has several limitations. Firstly, the KOA severity grading dataset consists of 9786 images. Although the dataset is comprehensive, images are obtained using the X-ray imaging technique. Using different image techniques, such as magnetic resonance imaging, ultrasound, or computed tomography, will increase the robustness and accuracy of the KOA severity classification. Secondly, even though using transfer learning models, including AlexNet, ResNet101, DenseNet201, and VGG19, creates a robust structure, comparisons can also be made using different transfer learning algorithms. For future studies, the power and robustness of the proposed model can be investigated using different medical images. Consequently, this model offers a dependable and objective diagnostic tool, potentially enabling more prompt interventions.

## Conflict of Interest

No conflict of interest was declared by the authors.

## References


Abedin, J., Antony, J., McGuinness, K., Moran, K., O'Connor, N. E., Rebholz-Schuhmann, D., & Newell, J. 2019. Predicting knee osteoarthritis severity: comparative modeling based on patient's data and plain X-ray images. Scientific Reports, 9(1), 5761.

Ahmed, H. A., & Mohammed, E. A. (2022). Detection and classification of the osteoarthritis in knee joint using transfer learning with convolutional neural networks (CNNs). Iraqi Journal of Science, 5058-5071.

Alshamrani, H. A., Rashid, M., Alshamrani, S. S., & Alshehri, A. H. (2023). Osteo-net: An automated system for predicting knee osteoarthritis from x-ray images using transfer-learning-based neural networks approach. Healthcare. 11 (9), 1-30.

Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H., & Lespessailles, E. 2019. A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. Computerized Medical Imaging and Graphics, 73, 11-18.

Cantor, A. B. (1996). Sample-size calculations for Cohen's kappa. Psychological Methods, 1(2), 150-151. doi: 10.1037/1082-989X.1.2.150

Çelik, Y., Dengiz, B., & Güney, S. (2023). Ahşap ham maddelerde yüzey hatasını belirlemek için görüntü işleme tabanlı kalite kontrol sistemi. Mühendislik Bilimleri ve Tasarım Dergisi, 11(4), 1365-1382.

Geng, R., Li, J., Yu, C., Zhang, C., Chen, F., Chen, J., Haonan, N., Wang, J., Kang, K., Wei, Z., Xu, Y., & Jin, T. (2023). Knee osteoarthritis: Current status and research progress in treatment. Experimental and therapeutic medicine, 26(4), 1-11.

Goswami, A. D. (2023). Automatic classification of the severity of knee osteoarthritis using enhanced image sharpening and CNN. Applied Sciences, 13(3), 1658.

Göker, H. (2024). Detection of cervical cancer from uterine cervix images using transfer learning architectures. Eskişehir Technical University Journal of Science and Technology A-Applied Sciences and Engineering, 25(2), 222-239.

Guan, B., Liu, F., Mizaian, A. H., Demehri, S., Samsonov, A., Guermazi, A., & Kijowski, R. (2022). Deep learning approach to predict pain progression in knee osteoarthritis. Skeletal Radiology, 1-11.

Islam, M. S., & Rony, M. A. T. (2024). CDK: A novel high-performance transfer feature technique for early detection of osteoarthritis. Journal of Pathology Informatics, 15, 100382.

Jain, R. K., Sharma, P. K., Gaj, S., Sur, A., & Ghosh, P. 2024. Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. Multimedia Tools and Applications, 83(3), 6925-6942.

Jung, J., Han, J., Han, J. M., Ko, J., Yoon, J., Hwang, J. S., Park, J. I., Hwang, G., Jung, J. H. & Hwang, D. D. J. (2024). Prediction of neovascular age-related macular degeneration recurrence using optical coherence tomography images with a deep neural network. Scientific Reports, 14(1), 5854. 1-12.

Karim, A. M., Kaya, H., Alcan, V., Sen, B., & Hadimlioglu, I. A. (2022). New optimized deep learning application for COVID-19 detection in chest X-ray images. Symmetry. 14 (5): 1003.

Karlik, B., & Olgac, A. V. (2011). Performance analysis of various activation functions in generalized MLP architectures of neural networks. International Journal of Artificial Intelligence and Expert Systems, 1(4), 111-122.

Khattar, A., & Quadri, S. M. K. (2022). Generalization of convolutional network to domain adaptation network for classification of disaster images on twitter. Multimedia Tools and Applications, 81(21), 30437-30464.

Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., & Ganslandt, T. (2022). Transfer learning for medical image classification: a literature review. BMC medical imaging, 22(1), 69.

Kishore, V. V., Kalpana, V., & Dosapati, U. B. (2024, April). Interpretation of KOA by KL Grading System using Deep Learning. In 2024 10th International Conference on Communication and Signal Processing (ICCSP) (pp. 109-114). IEEE.

Kokkotis, C., Moustakidis, S., Giakas, G., & Tsaopoulos, D. 2020. Identification of risk factors and machine learning-based prediction models for knee osteoarthritis patients. Applied Sciences, 10(19), 6797.

Kumar, H., Virmani, A., Tripathi, S., Agrawal, R., & Kumar, S. (2021). Transfer learning and supervised machine learning approach for detection of skin cancer: performance analysis and comparison. Drugs and Cell Therapies in Hematology, 10(1). 1-16

Langworthy, M., Dasa, V., & Spitzer, A. I. (2024). Knee osteoarthritis: disease burden, available treatments, and emerging options. Therapeutic Advances in Musculoskeletal Disease, 16, 1759720X241273009.

Li, H., & Duan, X. L. (2022). SAR ship image speckle noise suppression algorithm based on adaptive bilateral filter. Wireless Communications and Mobile Computing, 2022(1), 9392648.

Mahmoud, M., Kasem, M. S., & Kang, H. S. (2024). A Comprehensive Survey of Masked Faces: Recognition, Detection, and Unmasking. arXiv preprint arXiv:2405.05900.

Mohammed, A. S., Hasanaath, A. A., Latif, G., & Bashar, A. (2023). Knee osteoarthritis detection and severity classification using residual neural networks on preprocessed x-ray images. Diagnostics, 13(8), 1380.

Nurmirinta, T. A., Turunen, M. J., Korhonen, R. K., Tohka, J., Liukkonen, M. K., & Mononen, M. E. (2024). Two-stage Classification of future knee osteoarthritis severity after 8 Years using MRI: data from the osteoarthritis initiative. Annals of Biomedical Engineering, 1-12.

Qali, A., Selek, M., & Abbas, S. S. (2021). Termal görüntü işleme ile diz osteoartritinin tespit edilmesi. Avrupa Bilim ve Teknoloji Dergisi, (30), 69-72.

Rehman, S. U., & Gruhn, V. (2024). A Sequential VGG16+ CNN based Automated Approach with Adaptive Input for Efficient Detection of Knee Osteoarthritis Stages. IEEE Access.

Solak, F. Z. 2024. Classification of knee osteoarthritis severity by transfer learning from X-ray images. Karaelmas Science and Engineering Journal, 14(2), 119-133.

Tang, W., Sun, J., Wang, S., & Zhang, Y. (2023). Review of alexnet for medical image classification. EAI Endorsed Transactions on E-Learning, 9. 1-13. https://doi.org/10.4108/eetel.4389

Tiwari, A., Poduval, M., & Bagaria, V. (2022). Evaluation of artificial intelligence models for osteoarthritis of the knee using deep learning algorithms for orthopedic radiographs. World Journal of Orthopedics, 13(6), 603.

Tong, Y., Lu, W., Deng, Q. Q., Chen, C., & Shen, Y. (2020). Automated identification of retinopathy of prematurity by image-based deep learning. Eye and Vision, 7, 1-12.

Wahyuningrum, R. T., Yasid, A., & Jacob Verkerke, G. (2020, December). Deep neural networks for automatic classification of knee osteoarthritis severity based on X-ray images. In Proceedings of the 2020 8th International Conference on Information Technology: IoT and Smart City (pp. 110-114).

Wang, Y., Wang, X., Gao, T., Du, L., & Liu, W. (2021). An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative. Journal of Healthcare Engineering, 2021(1), 5586529.

Yang, Z., Nashik, S., Huang, C., Aibin, M., & Coria, L. (2024). Next-Gen Remote Airport Maintenance: UAV-Guided Inspection and Maintenance Using Computer Vision. Drones, 8(6), 225.

Zeng, L., Zhou, G., Yang, W., & Liu, J. (2023). Guidelines for the diagnosis and treatment of knee osteoarthritis with integrative medicine based on traditional Chinese medicine. Frontiers in medicine, 10, 1260943.