NCCS Research Journal, 4 (1), August 2025, ISSN: 2822-1605

Knee Osteoarthritis Severity Classification Using CNN and Image Enhancement Filters

Krisha Pandey

Email: csit211809_krisha@achsnepal.edu.np

Sabina Rijal

Email: csit211840_sabina@achsnepal.edu.np

Asmita Gaire

Email: csit211838_asmita@achsnpal.edu.np

Received: March 2025; Revised: May 2025; Accepted: July 2025

https://doi.org/10.3126/nccsrj.v4i1.84359

Abstract

CNNs for the classification of knee osteoarthritis is a promising avenue to assist in clinical decision making. Our study examined the influence of several image enhancement filters on CNN performance for KOA severity classification using X-rays. The KOA dataset contained 8,260 labelled X-ray images of five KL grades (0-4) divided into three datasets (training, validation, and test). As we assessed the effects of the different image enhancement preprocessing filters on labelled X-ray images, we applied several typical image enhancement preprocessing filters such as the Sobel, Gaussian, CLAHE, Gabor, and Entropy filters in order to evaluate the effect they had on classification performance. The best classifier for KOA classification, using the CLAHE filter, had the best performance for the 5-layered CNN model with the best accuracy (85%), F1-score (0.72), and AUC score of (0.83). This exemplifies the usefulness of contrast enhancement in medical image classification. Findings from our study can demonstrate that image enhancement techniques contribute to the reproducibility of CNN-based KOA grading systems and offer one more step towards more reliable and automated KOA assessment.

Keywords: convolutional neural networks, image enhancement filters, Knee Osteoarthritis Classification, X-Ray image processing.

1. Introduction: Knee Osteoarthritis Severity Classification Using CNN and Image Enhancement Filters

Knee osteoarthritis (KOA) is a common joint disease that causes pain, stiffness, and difficulty moving. It affects millions of people worldwide and can significantly reduce their quality of life (Antico et al., 2020). Detecting KOA early and correctly assessing its severity is very important for effective treatment. Doctors usually use the Kellgren-Lawrence (KL) grading system, which classifies KOA into five levels based on X-ray images (Schiphof et al., 2008). However, this method is not always reliable because different doctors may interpret X-rays differently, leading to inconsistent results. It is also time-consuming, which can delay treatment (Mohammed et al., 2023). As a result, there is a growing need for automated systems that can assist in accurate and efficient KOA diagnosis.

With the advancement of artificial intelligence, deep learning techniques like Convolutional Neural Networks (CNNs) have been widely used in medical image analysis. CNNs are powerful at detecting patterns in images and can help automate the process of diagnosing and classifying KOA (Guida et al., 2021). However, one major challenge is that the early stages of KOA (KL grades 0 and 1) are difficult to differentiate because the features in X-ray images often overlap. This makes it hard for the model to classify them accurately. To overcome this, researchers have explored the use of image enhancement techniques to improve the quality of X-ray images, making it easier for the model to detect key features (Ahmed & Mstafa, 2022).

In this study, we aim to develop an automated KOA severity classification system using CNNs and image enhancement techniques. Our goal is to improve the model's accuracy in detecting different KOA stages, especially the early ones. By enhancing X-ray images before feeding them into the model, we hope to increase its ability to identify key features and make more reliable predictions. This system has the potential to support doctors by providing quick and consistent KOA diagnosis, reducing errors, and speeding up the decision-making process. The rest of the paper is organized as follows: Section 2 reviews existing studies on KOA classification using deep learning. Section 3 explains the methodology, including how the CNN model works and the image enhancement techniques used. Section 4 presents the results and discusses key findings. Finally, Section 5 concludes the study and provides directions for future research.

2. Literature Review/ Related Works

The diagnosis of Knee Osteoarthritis (KOA) from X-ray images has long been a challenging task due to the complexity of the condition and the need for expert interpretation. Traditional methods rely heavily on manual grading, which is time-consuming and often subjective. Machine learning, particularly Convolutional Neural Networks (CNNs), has emerged as an effective solution for automating KOA detection. Guida et al. (2021) applied 3D CNNs to MRI sequences for classifying KOA severity and demonstrated superior accuracy compared to traditional CNNs used on X-ray images. Their study showed an improvement in multi-class and binary classification accuracy, suggesting that 3D CNNs combined with MR imaging could improve clinical diagnosis. However, their focus was on MRI, and the challenge of applying such methods to X-ray images remains underexplored.

Mohammed et al. (2023) used CNNs to classify KOA severity from X-ray images, with a focus on early-stage detection. Despite achieving a 69% accuracy, their study highlighted significant challenges, including the limited availability of early-stage data and the difficulty in distinguishing between KL grade0, KL grade1 and KL grade2 which exhibit subtle radiographic differences. This points to the need for more robust datasets and improved model architectures that can handle such complexities in X-ray images.

To address these challenges, recent studies have explored image enhancement techniques to improve the quality of X-ray images. Image enhancement plays a vital role in medical image analysis, particularly in improving the contrast and visibility of key features in X-ray and ultrasound images. One study critically evaluated the effects of Histogram Equalization, Contrast Limited Adaptive Histogram Equalization, and Fuzzy Enhancement on X-ray images from the MURA dataset, highlighting the potential of these techniques to enhance diagnostic accuracy in clinical practice (Htun & Tun, 2024). These methods significantly improved contrast and edge detection, which are critical for successful image analysis and classification. By making hidden details more visible, such enhancement techniques can improve the performance of downstream machine learning models. However, the integration of such enhancement methods with CNNs for KOA detection remains underexplored, particularly regarding their impact on classification accuracy.

In this study, we propose a novel approach that integrates image enhancement filters with CNNs to improve KOA classification accuracy from X-ray images. By addressing the challenge of poor-quality images and enhancing critical features, our approach aims to improve the model's ability to differentiate between different KOA grades more effectively, especially in the early stages of the disease where subtle differences are often difficult to distinguish. This will ultimately enhance diagnostic accuracy and efficiency in clinical settings, providing a reliable and automated tool that can assist medical professionals in making quicker and more accurate diagnoses. Through this approach, we aim to contribute to the ongoing development of deep-learning-based diagnostic tools that can be used in clinical practice to support physicians and healthcare providers in making informed decisions regarding KOA treatment and management.

3. Methodology

This section describes data collection, feature engineering and selection, data preprocessing, model implementation, training and optimization, performance evaluation and experimental setup.

Data Collection

The knee X-ray images utilized in this study were obtained from the publicly available "Knee Osteoarthritis Dataset with Severity Grading" curated by Falah Gatea and hosted on Kaggle (https://www.kaggle.com/datasets/falahgatea/knee-osteoarthritis-dataset-with-severity-grading). The dataset comprises a total of 8,260 grayscale anterior-posterior knee X-ray images, each labeled according to the Kellgren-Lawrence (KL) grading scale, ranging from Grade 0 (normal) to Grade 4 (severe osteoarthritis). All images were resized to a standardized resolution of 224x224 pixels to maintain uniformity during preprocessing. The dataset was divided into training (70%), validation (10%), and testing (20%) subsets to facilitate model development and performance evaluation. Acknowledging the source of the dataset not only credits the original contributor but also supports reproducibility and transparency in research, allowing future studies to build upon the same data foundation.

Data Preprocessing

All X-ray images were first resized to a standardized resolution of 224x224 pixels to ensure uniform input dimensions for the CNN model. As part of the image enhancement process, the CLAHE (Contrast Limited Adaptive Histogram Equalization) filter was initially applied during preprocessing to improve the local contrast of grayscale images, making critical anatomical structures more distinguishable. Although multiple image enhancement filters were later applied and compared individually during experimentation (as discussed in Section 4), CLAHE was selected as the base filter for early-stage preprocessing to evaluate its baseline impact before introducing additional transformations. This initial use of CLAHE was not mandatory for all experiments, hence earlier noted as “optional,” but served as a controlled starting point for performance benchmarking.

In addition, data augmentation techniques were employed to increase the diversity and size of the training dataset and reduce overfitting. Augmentation steps included random horizontal flipping, brightness and contrast adjustments, and slight rotations. These transformations aimed to simulate variability in real-world clinical imaging conditions and enhance the model’s generalization ability. Finally, pixel values were normalized to ensure numerical stability during training and to speed up convergence of the optimization process.

Model Implementation

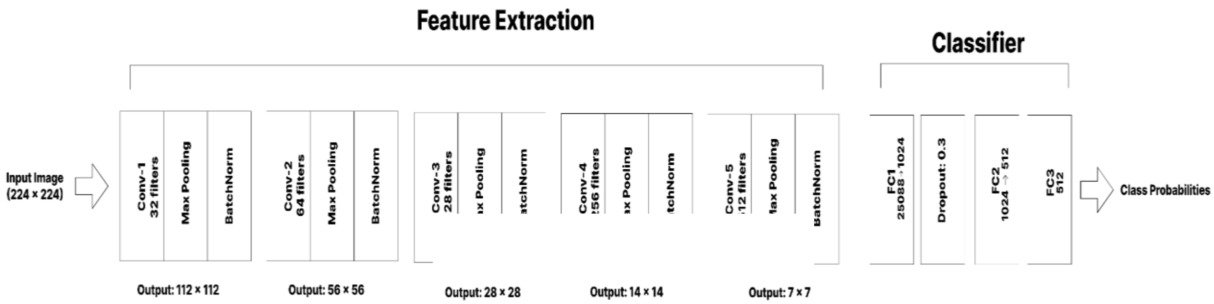
The proposed CNN model is shown in figure 1. The CNN model consists of convolutional and fully connected layers. The architecture begins with five convolutional layers, each followed by batch normalization and max-pooling operations to progressively reduce the spatial dimensions of the input images. The convolutional filters start at 32 and gradually increase to 512 as the layers deepen, effectively capturing complex features at various levels of abstraction. These operations reduce the input image size from 224x224 pixels to a smaller 7x7 feature map. After passing through the final convolutional layer, the resulting 7x7x512 feature map is flattened into a 25088-dimensional vector.

This vector is then passed through three fully connected layers designed to process and extract high-level features. The first fully connected layer reduces the feature dimensions to 1024 units, which helps retain key information while reducing computational complexity. To mitigate overfitting, several dropouts specifically 0.25, 0.3, 0.4, and 0.5 are applied after the first fully connected layer and 0.3 consistently yields the best validation performance, providing an effective balance between model generalization and learning capacity.

The second fully connected layer further reduces the dimension to 512 units. Finally, the last layer outputs the classification probabilities for each class, effectively generating the final predictions for the input images. This architecture enables the model to learn meaningful patterns while controlling overfitting, thereby improving prediction accuracy on complex image classification tasks.

Figure 1:

Architectural Diagram of the CNN Model



The image enhancement filters are applied to the input images to highlight important features, such as edges and textures, that are crucial for accurate classification. These filters help the model better capture critical details and improve its ability to learn relevant patterns, leading to enhanced performance in the classification task. These filters were chosen based on their individual strengths in handling noise reduction, edge detection, contrast enhancement, and texture extraction. Each filter contributes differently to emphasizing certain aspects of the image that are crucial for accurate feature learning and classification. For instance, some filters focus on enhancing boundaries, while others improve contrast or preserve structural textures.

a) **Sobel Filters:**

   The Sobel filter algorithm works by detecting edges in an image through convolution with two predefined Sobel kernels: one for detecting horizontal edges (Sobel X) and another for vertical edges (Sobel Y). The image is first converted into a tensor, and a batch dimension is added to prepare it for convolution operations.

b) **Gaussian Filter:** The Gaussian blur algorithm smooths an image by averaging pixel values with their neighbors using a Gaussian kernel, reducing noise and detail. The degree of blurring is controlled by the kernel size, with larger kernels producing stronger blurs. This process results in a softened image, useful for noise reduction or pre-processing in tasks like edge detection.

c) **Bilateral Filter:** The bilateral filter is a widely used edge-preserving smoothing technique that reduces noise in images while maintaining sharp edges. Unlike traditional filters that consider only spatial proximity, the bilateral filter also accounts for differences in pixel intensity, which helps preserve edge information. The filtering process involves a weighted average of nearby pixels, where weights are determined by both the spatial distance and the intensity difference between neighboring pixels. The behavior of the filter is controlled by three key parameters: the neighborhood diameter, which defines the spatial extent of the filter; the intensity sigma, which controls how strongly intensity differences influence filtering; and the spatial sigma, which determines how much influence neighboring pixels have based on their distance. This combination allows the bilateral filter to smooth homogeneous regions while retaining important edge details.

d) **Entropy Filter:** Entropy Filter is implemented as a custom transformation for preprocessing images in the dataset. This filter is designed to enhance the features of the images by highlighting regions of high complexity, which can be particularly useful for distinguishing between different classes of bone conditions in knee osteoarthritis. The Entropy Filter class inherits from object and overrides the call method, which allows it to be used in a pipeline of image transformations.

e) **CLAHE Filter:** CLAHE (Contrast Limited Adaptive Histogram Equalization) is applied to enhance the contrast of the images in the dataset. After reading each image , the CLAHE filter is initialized with specific parameters to limit contrast amplification and define the grid size for processing. The filter is then applied to the grayscale image, improving its visual quality by enhancing local contrast while avoiding over-amplification of noise.

f) **Bidirectional Filter:** The bidirectional filter applies smoothing in two orthogonal directions, such as horizontal and vertical. This helps reduce noise while preserving important features in an image by considering both axes of variation.

g) **Laplacian Filter:** The Laplacian filter is used for edge detection and image enhancement by computing the second spatial derivative of an image. It highlights regions of rapid intensity change.

h) **Gabor Filter:** Gabor filters are used for texture analysis, combining a Gaussian envelope with a sinusoidal carrier. They effectively detect edges and patterns at different orientations and scales, making them useful for feature extraction.

i) **Histogram Equalization:** Histogram equalization improves image contrast by redistributing intensity values. It enhances features that are too dark or light, making the image more visually balanced and revealing more details.

Training and Optimization

The training process began by addressing class imbalance, which can negatively impact model performance by causing it to favor majority classes. To counter this, class weights were calculated and incorporated into the CrossEntropyLoss function, which measures the dissimilarity between the predicted class probabilities and the true class labels, penalizing incorrect predictions more when the model is confident about the wrong class. This helped in improving prediction accuracy across all categories. For optimization, the Adam optimizer was selected due to its efficiency in handling sparse gradients and its adaptive learning rate. A learning rate of 0.0001 was used to facilitate stable learning, while weight decay of 1e-5 was applied to prevent overfitting by regularizing the model's parameters. The training process was set to run for a maximum of 150 epochs, but to avoid unnecessary computations and overfitting, an early stopping mechanism was implemented. If the validation loss failed to improve for 12 consecutive epochs, training was halted to preserve the best-performing model.

Each epoch processed data in mini batches of 32 images, which improved computational efficiency and model convergence. Throughout training, the model's progress was closely monitored using training and validation loss, along with their corresponding accuracy metrics. These evaluations ensured that the model was learning effectively and not overfitting to the training data. At the end of training, the best-performing model determined by the lowest validation loss was saved for further testing and deployment. This approach ensured that the final model had strong generalization capabilities, making it reliable for real-world applications.

Performance Evaluation

The model was evaluated using several performance metrics including the confusion matrix, classification accuracy, precision, recall, F1-score, and Area Under the Curve (AUC) to assess its effectiveness on validation data.

- **Confusion Matrix**: This tabulates the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), helping visualize misclassification patterns.
- **Accuracy**:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

This measures the overall correctness of the model's predictions.

- **Precision:** Precision evaluates the proportion of correctly predicted positive samples out of all predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall:** Recall measures how well the model identifies actual positive cases.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1-Score:** This harmonic mean of precision and recall ensures a balanced evaluation, especially in imbalanced datasets.

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

- **AUC:** AUC quantifies the model's ability to distinguish between classes by measuring the area under the Receiver Operating Characteristic (ROC) curve. A higher AUC indicates better class separation.

These metrics collectively provide a comprehensive understanding of the model's strengths and weaknesses, particularly in handling class imbalance and distinguishing between KOA severity levels.

Experimental Setup

The experiments were conducted on a system equipped with an Intel Core i5 processor and an MX 550 graphics card, along with 16 GB of RAM. The implementation leveraged the PyTorch framework along with supporting libraries like Seaborn and Scikit-learn for data preprocessing, visualization, and model development.

4. Results & Findings

The performance of the model was analyzed through a series of experiments using various combinations of preprocessing filters and dropout values. The experiments were structured around two primary conditions: those conducted with image preprocessing filters such as Contrast Limited Adaptive Histogram Equalization and histogram equalization and those conducted without any filters. The objective was to evaluate how these preprocessing steps affected the model’s ability to classify the severity of knee osteoarthritis (KOA).

A dropout rate of 0.3 was used in the majority of the experimental combinations. This particular value was chosen based on its consistent and stable performance during preliminary trials. It effectively balanced the risks of overfitting and underfitting, making it a reliable baseline for further comparisons. However, to understand the sensitivity of the model to dropout changes, additional experiments were selectively conducted with dropout values of 0.25, 0.4, and 0.5. These alternative values were not applied across all filter conditions. The reason for this selective testing was twofold: first, combinations with very high dropout rates (such as 0.4 and 0.5) yielded poor performance during initial testing, indicating a loss of important features; and second, the study aimed to maintain computational efficiency and feasibility, so only promising configurations were explored further.

To ensure diversity while maintaining experimental control, specific combinations were chosen. For the filtered input condition, experiments were performed with dropout rates of 0.25, 0.3, and 0.4. In contrast, for the unfiltered input condition, only dropout values of 0.3 and 0.5 were tested. This decision was made because extending all dropout values across both conditions was not practical due to time and resource limitations. Moreover, earlier results suggested that certain combinations were unlikely to perform well, so they were intentionally omitted. Although a complete matrix of experiments would be ideal, the selected combinations were sufficient to demonstrate the general trend and impact of both filters and dropout on model performance.

All the trained models were evaluated on a separate test dataset. This ensured that the reported performance metrics accuracy, precision, recall, F1-score, and AUC reflected the model's ability to generalize beyond the training and validation data. The use of this test set provided a robust and unbiased assessment of each experimental setup. The results from these combinations were then compared to identify the most effective configuration for classifying KOA severity. By focusing on meaningful experimental designs and justifying the exclusion of unpromising combinations, the study maintained a balance between comprehensive analysis and practical constraints.

Table 1:

Performance of CNN Model with Different Dropout and Image Enhancement Filters

Model	Dropouts	Filters	Classification Accuracy for Test Data	Metrics
5-layer Convolutional Neural Network	0.25	-	Class 0:0.79 Class 1:0.77 Class 2:0.80 Class 3:0.93 Class 4:0.97	Precision:0.62 F1 score:0.65 AUC:0.78
5-layer Convolutional Neural Network	0.3	-	Class 0:0.81 Class 1:0.76 Class 2:0.84 Class 3:0.95 Class 4:0.98	Precision:0.72 F1 score:0.71 AUC:0.81
5-layer Convolutional Neural Network	0.3	ENTROPY	Class 0:0.79 Class 1:0.73 Class 2:0.82 Class 3:0.95 Class 4:0.97	Precision:0.69 F1 score:0.66 AUC:0.77

5-layer Convolutional Neural Network	0.3	HISTOGRAM EQUALIZATION	Class 0:0.76 Class 1:0.74 Class 2:0.80 Class 3:0.95 Class 4:0.97	Precision:0.69 F1 score:0.70 AUC:0.79
5-layer Convolutional Neural Network	0.3	BILATERAL	Class 0:0.77 Class 1:0.72 Class 2:0.80 Class 3:0.93 Class 4:0.96	Precision:0.64 F1 score:0.66 AUC:0.76
Model	Dropouts	Filters	Classification Accuracy for Test Data	Metrics
5-layer Convolutional Neural Network	0.3	SOBEL	Class 0:0.78 Class 1:0.80 Class 2:0.81 Class 3:0.94 Class 4:0.98	Precision:0.68 F1 score:0.68 AUC:0.80
5-layer Convolutional Neural Network	0.3	DIRECTIONAL	Class 0:0.75 Class 1:0.77 Class 2:0.84 Class 3:0.93 Class 4:0.98	Precision:0.70 F1 score:0.66 AUC:0.77

5-layer Convo-lutional Neural Network	0.3	LAPLACIAN	Class 0:0.76 Class 1:0.77 Class 2:0.79 Class 3:0.91 Class 4:0.99	Precision:0.65 F1 score:0.63 AUC:0.77
5-layer Convo-lutional Neural Network	0.3	GABOR	Class 0:0.75 Class 1:0.66 Class 2:0.83 Class 3:0.94 Class 4:0.99	Precision:0.70 F1 score:0.69 AUC:0.80
5-layer Convo-lutional Neural Network	0.3	GAUSSIAN	Class 0:0.78 Class 1:0.73 Class 2:0.78 Class 3:0.92 Class 4:0.99	Precision:0.68 F1 score:0.69 AUC:0.78
5-layer Convo-lutional Neural Network	0.3	CLAHE	Class 0:0.82 Class 1:0.76 Class 2:0.85 Class 3:0.95 Class 4:0.99	Precision:0.72 F1 score:0.72 AUC: 0.83
Model	Dropouts	Filters	Classification Accuracy for Test Data	Metrics
5-layer Convo-lutional Neural Network	0.4	-	Class 0:0.81 Class 1:0.72 Class 2:0.83 Class 3:0.95 Class 4:0.99	Precision:0.71 F1 score:0.72 AUC:0.80

5-layer Convolutional Neural Network	0.5	-	Class 0:0.78 Class 1:0.71 Class 2:0.83 Class 3:0.94 Class 4:0.99	Precision:0.73 F1 score:0.70 AUC:0.80

The confusion matrices for different dropout rates and image enhancement filters are provided below, illustrating the detailed classification performance across all categories.

Figure 2:

Confusion Matrix for 0.25 Dropout

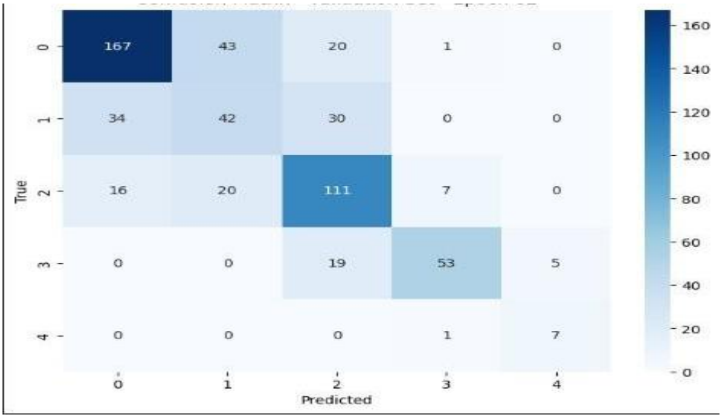


Figure 3:
Confusion Matrix for 0.30 Dropout

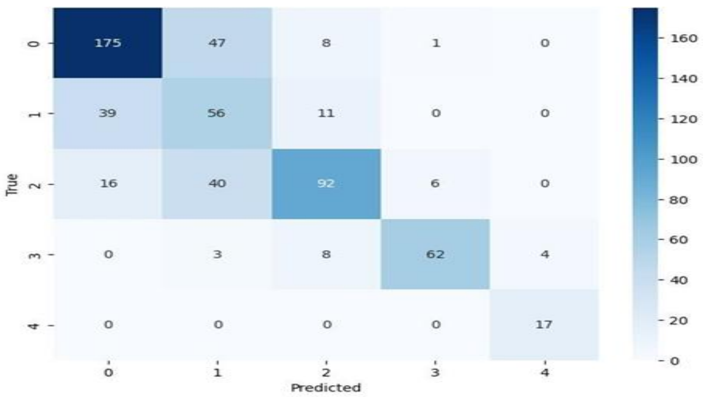


Figure 4:

Confusion Matrix for 0.40 Dropout

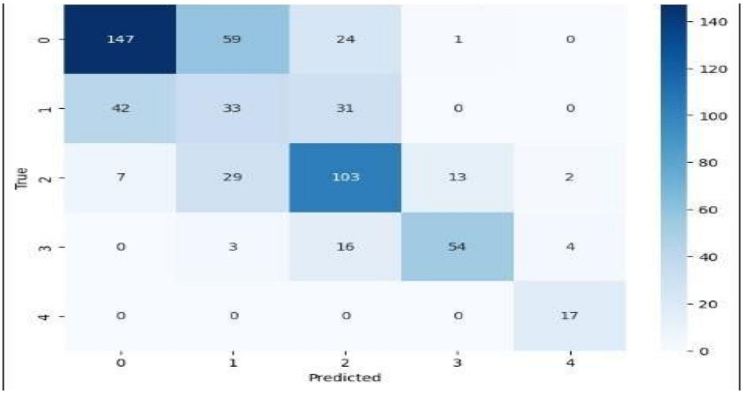


Figure 5:

Confusion Matrix for Sobel Filter

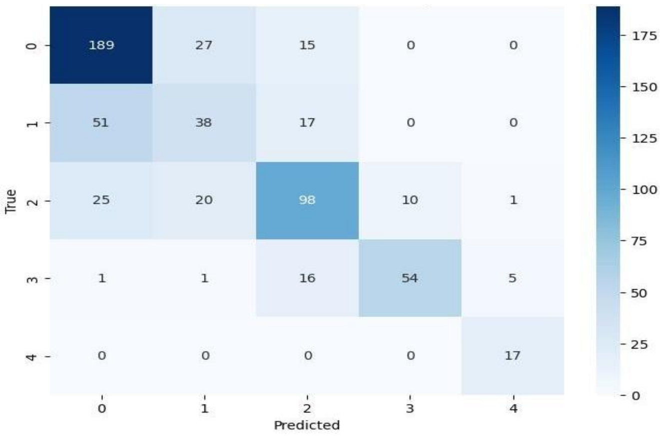


Figure 6:

Confusion Matrix for Gaussian Filter

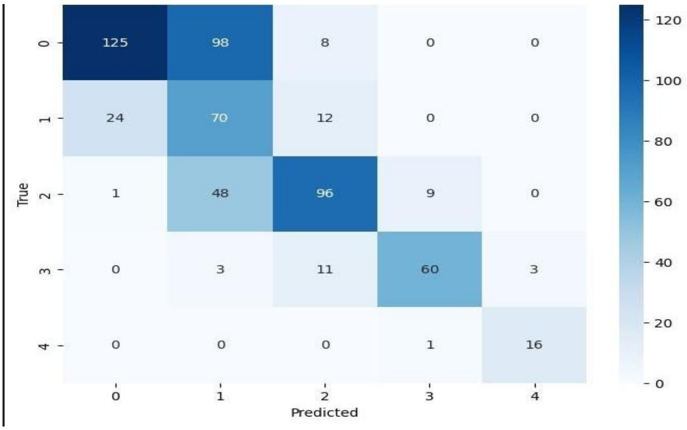


Figure 7:

Confusion Matrix for Entropy Filter

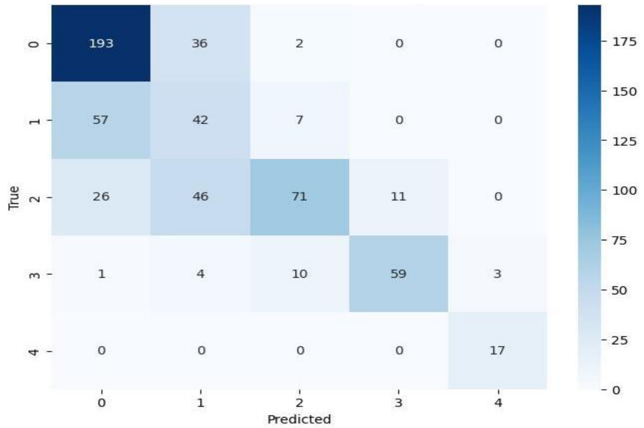


Figure 8:

Confusion Matrix for Directional Filter

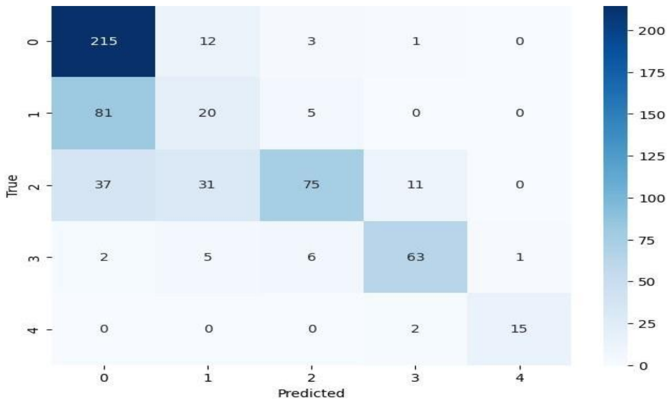


Figure 9:

Confusion Matrix for Histogram Equalization Filter

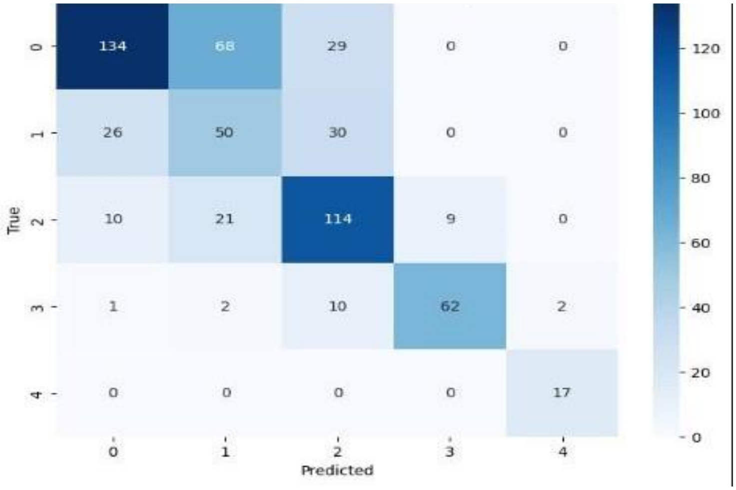


Figure 10:

Confusion Matrix for Laplacian Filter

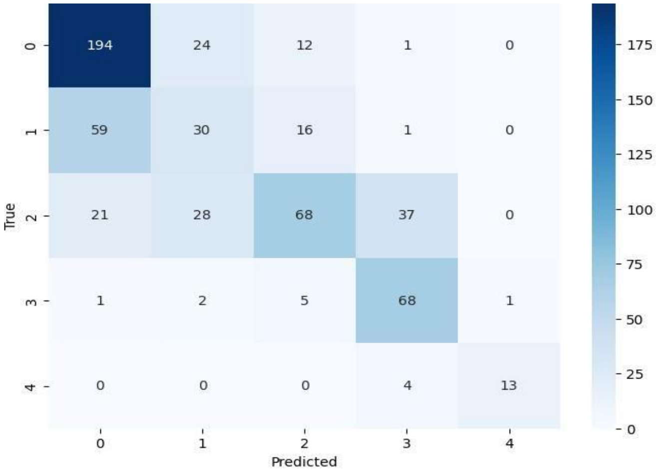


Figure 11:

Confusion Matrix for Gabor Filter

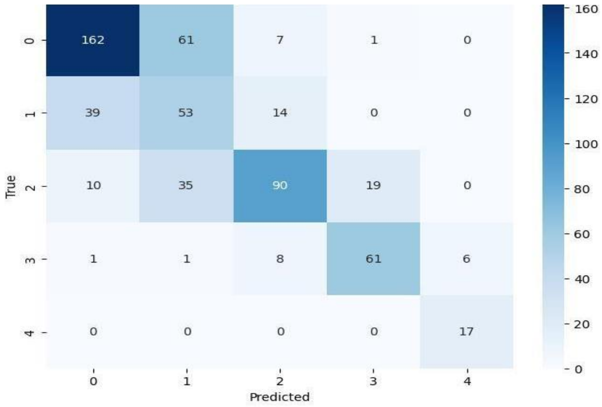


Figure 12:

Confusion Matrix for Bilateral Filter

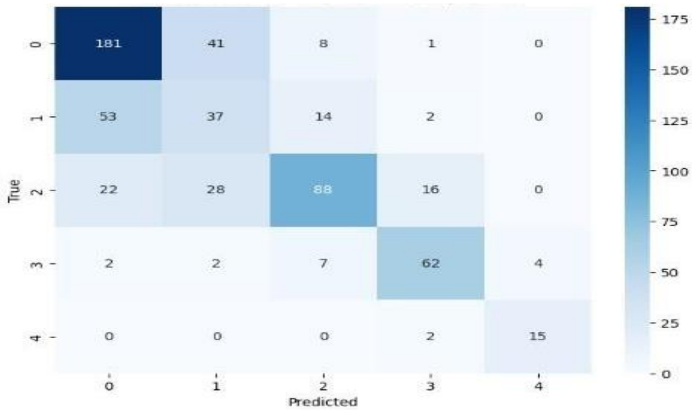
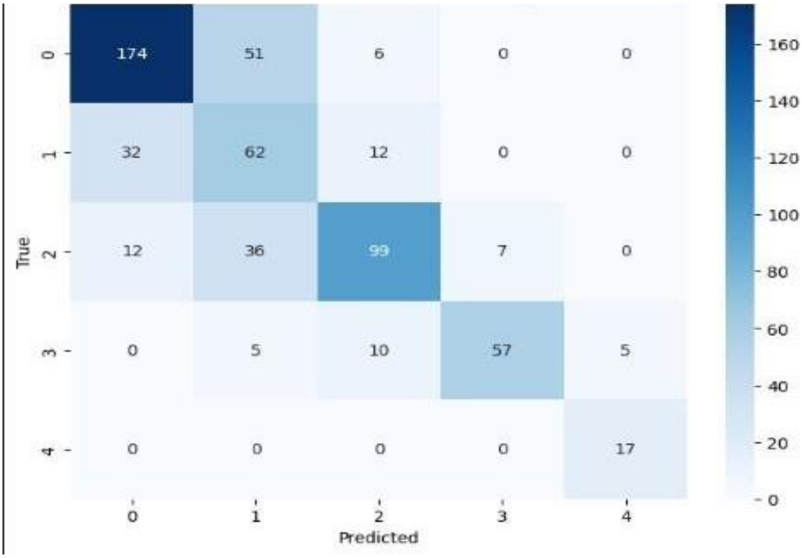


Figure 13 :

Confusion Matrix for Clahe Filter



Based on the analysis of the confusion matrices, the CLAHE (Contrast Limited Adaptive Histogram Equalization) filter demonstrated improved performance, particularly in distinguishing between overlapping classes in Knee Osteoarthritis (KOA) detection. The model showed significant improvements in accuracy, especially for Class 4 (moderate KOA) with an accuracy of 0.95 and Class 5 (severe KOA) with an accuracy of 0.99. These results suggest that CLAHE effectively enhanced the contrast of key features, enabling the model to better identify and classify the more advanced stages of KOA.

The model's overall performance, measured by the AUC score of 0.83, indicates a strong ability to differentiate between the stages of KOA. This demonstrates that CLAHE not only improved the accuracy for later stages but also contributed to better classification overall. The results confirm that CLAHE is a valuable preprocessing technique, particularly for enhancing the detection of advanced KOA stages.

5. Conclusion

This study implemented a CNN-based Knee Osteoarthritis (KOA) classification system, integrating various enhancement filters to improve classification accuracy, particularly for advanced KOA stages (KL grades 3 and 4). The model was rigorously evaluated based on its ability to address class imbalance, which is a common issue in medical image datasets, and to enhance feature extraction, with Contrast Limited Adaptive Histogram Equalization (CLAHE) emerging as the most effective preprocessing technique. The results demonstrate that image enhancement filters significantly improve KOA classification, particularly for later-stage cases (Moderate (3) and Severe (4)), leading to more reliable and accurate predictions. By refining the quality of X-ray images, the system is able to more effectively capture key features that distinguish between different severity levels. However, distinguishing between early KOA stages (KL grades 0, 1, and 2) remains a challenge due to overlapping features, which makes it difficult for the model to differentiate subtle differences in the early stages of the disease.

The study underscores the importance of image preprocessing in medical image classification, emphasizing that proper data enhancement is crucial for improving the performance of deep learning models. It also highlights how enhancement techniques, such as CLAHE, influence deep learning model performance by allowing the model to better recognize and interpret features that would otherwise be overlooked. The findings suggest that incorporating advanced preprocessing methods and optimizing feature extraction strategies can significantly enhance the model's ability to differentiate between various KOA stages and thus further refine KOA severity grading. Moreover, this study contributes to the ongoing development of deep-learning-based diagnostic tools, which hold great potential for transforming clinical practice by providing more accurate and automated assessments.

References

Antico, M., Sasazawa, F., Dunnhofer, M., Camps, S. M., Jaiprakash, A. T., Pandey, A. K., Crawford, R., Carneiro, G., Fontanarosa, D. (2020). Deep learning-based femoral cartilage automatic segmentation in ultrasound imaging for guidance in robotic knee arthroscopy. *Ultrasound in Medicine & Biology, 46*(2), 422-435.

Guida, C., Zhang, M., & Shan, J. (2021). Knee osteoarthritis classification using 3D CNN and MRI. *Applied Sciences, 11*(11), 5196.

Htun, N. T., & Tun, K. M. M. (2024). Evaluating the impact of image enhancement on X-ray diagnostics using HE, CLAHE, and fuzzy enhancement techniques. *Iconic Research And Engineering Journals, 8*(1), 376–387.

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.

Schiphof, D., Boers, M., & Bierma-Zeinstra, S. M. (2008). Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Annals of the Rheumatic Diseases, 67*(7), 1034-1036.