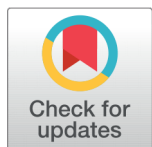


ORIGINAL ARTICLE



OPEN ACCESS

Received: 03/04/2026

Accepted: 01/07/2026

Published: 11/07/2026

Citation: Bruntha S, Chidambaram M (2026) Explainable Deep Hybrid Learning for Early Prediction and Multi-Class Classification of Knee Osteoarthritis Using X-Ray Images. Indian Journal of Science and Technology 19(24): 1648-1657. <https://doi.org/10.17485/IJST/v19i24.525>

* **Corresponding author.**brunthaloga@gmail.com**Funding:** None**Competing Interests:** None

Copyright: © 2026 Bruntha & Chidambaram. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Explainable Deep Hybrid Learning for Early Prediction and Multi-Class Classification of Knee Osteoarthritis Using X-Ray Images

S Bruntha^{1*}, M Chidambaram²

¹ Research Scholar, Department of Computer Science, Rajah Serfoji Government College (Autonomous), Thanjavur-613005, Affiliated to Bharathidasan University, Tiruchirappalli - 620 024, Tamil Nadu, India

² Research Supervisor & Associate Professor, Department of Computer Science, Rajah Serfoji Government College (Autonomous), Thanjavur-613005, Affiliated to Bharathidasan University, Tiruchirappalli - 620 024, Tamil Nadu, India

Abstract

Objectives: To propose a deep hybrid model using the ResNet34 and Vision Transformer to classify the severity of KOA using X-ray images. **Method:** Knee Osteoarthritis (KOA), being one of the most common and progressive conditions in the field of orthopedics, is posing a significant financial challenge to the healthcare sector globally. The study employed the publicly available Osteoarthritis Initiative (OAI) dataset, comprising 8,260 knee X-ray images partitioned into pre-defined training (n=5,778), validation (n=826), and test (n=1,656) splits. CLAHE preprocessing (clip limit 2.0; tile grid 8×8) was applied to all images prior to training. A hybrid ResNet34 and Vision Transformer architecture was trained end-to-end on five-class KL grading under a class-weighted cross-entropy objective, with AdamW optimisation and cosine annealing scheduling. **Findings:** The proposed model, ResNet34-ViT achieved the highest performance with an accuracy of 66.79%. Compared to other models, DenseNet-121 and ResNet-50 V2 reached 66.31% and 65.36% accuracy, respectively. Class-balanced oversampling (10,000 training samples per KL grade) and inverse-frequency loss weighting improved minority-grade recall. Grad-CAM++ heatmaps aligned model attention with joint-space narrowing, osteophyte formation, and subchondral changes across KL grades. However, the model still has limitations, indicating the need for further research to enhance its clinical applicability. **Novelty:** This study introduces a hybrid model combining ResNet34 and Vision Transformer for classifying knee osteoarthritis. A targeted preprocessing pipeline, anchored by Contrast-Limited Adaptive Histogram Equalization (CLAHE; clip limit 2.0, tile grid 8×8) and class-specific oversampling, resolves a 13.21× distributional imbalance present in the OAI dataset, enabling equitable five-category grading across all Kellgren-Lawrence severity tiers. Grad-CAM++ heatmaps further validate model decisions by aligning activation regions with anatomical structures that radiologists routinely examine during KL grading, including the joint space, subchondral bone, and osteophyte margins.

Keywords: Knee Osteoarthritis (KOA); KL Grading; X-Ray Images; Convolutional Neural Network (CNN); Vision Transformer (ViT); GradCAM++

1 Introduction

Knee osteoarthritis, a chronic health problem, is mainly caused by age-related changes in the knee joint. Accurate diagnosis is a critical task, typically performed by healthcare experts through the interpretation of X-ray images, a time-consuming process. Due to the availability of X-ray image data, deep learning methods, particularly CNN-based methods, have shown increased popularity in facilitating the diagnosis of knee osteoarthritis⁽¹⁾. Deep learning-based methods for classifying knee osteoarthritis grades have been discussed in this work. Early diagnosis and classification of the disease have been performed using information learned from X-ray and magnetic resonance images^{(1), (2), (3)}. Classification results have been obtained to assess the severity of the disease, and its progression has been analyzed using different datasets and performance evaluation metrics. Both the multi-class and binary classification approaches are considered using different data sets derived from the original data set to assess the performance of the model. Gradient Weighted Class Activation Map (GradCAM)⁽⁴⁾ is used to improve the assessment of the model interpretability.

This research is focused on Knee Osteoarthritis (KOA), which is the most common type of arthritis. It is characterized by the progressive degradation of the cartilages and the underlying bones. It is seen among the elderly population. As it is one of the major causes of disability among the global population, its chronic and progressive nature affects the movement of the patients, thereby increasing the disability level and the cost of treatment. Early diagnosis of knee osteoarthritis (KOA) is important to effectively manage the symptoms. Furthermore, early diagnosis enables the initiation of treatment to halt the progression of the disease. It also prevents permanent damage to the knee joint. The Kellgren Lawrence grading scale is the most frequently used scale to measure the severity of knee osteoarthritis (KOA). This grading scale, which is frequently used by doctors, relies on the routine knee X-rays to diagnose osteoarthritis (OA).

The Kellgren-Lawrence scale is used for classifying the severity of OA. The scale ranges from 0 to 4. If the OA is classified as "0," it means there are no signs of OA. If it is classified as "4," it means OA is very advanced. There are three major problems in the current method of detecting bone spurs and the degree of narrowing of the joint space.

- The assessment of the subtle changes observed on the X-ray images is subjective in nature, which can vary from time to time, especially during the early stages of osteoarthritis, such as Kellgren Lawrence grade 1, which is also referred to as "Doubtful OA." The assessment is also based on personal judgment, which results in considerable variations in the results obtained from different radiologists as well as the radiologists' personal opinions. This is where the lack of consistency in the results can create problems in the diagnosis process.
- In the case where the clinics are overwhelmed with patients, the time required for the manual assessment of the X-ray images by experts can greatly escalate the issue.
- Manual classification is not sensitive enough to detect the subtle changes associated with KL grade 1. As a result, the diagnosis is made only when the damage is irreversible, i.e., when the KL grade is 2 or higher.

Deep Learning (DL) techniques have witnessed tremendous developments, especially with the implementation of Convolutional Neural Networks (CNNs). Such techniques are a promising solution to overcome existing challenges by providing a platform to analyze medical image data objectively and in a consistent manner. Deep Learning techniques have the potential to accelerate diagnostic processes, enhance consistency in observations by different individuals, and increase the accuracy of opinions associated with the severity of knee osteoarthritis (KOA). This research proposes a comprehensive deep learning framework for complex multi-class classification of knee osteoarthritis (KOA) severity levels based on X-ray images, along with a potential tool to display images in a manner that would help doctors become confident and comprehend the situation better. Figure 1 represents the KL grading system.

Many researchers have been focusing on the identification, classification, and analysis of knee osteoarthritis through the implementation of advanced algorithm techniques, especially in the context of classification with the help of X-ray images. The performance of the current state-of-the-art deep learning techniques is validated through the comparative analysis of the performance of the techniques with the clinical assessment performed by medical experts. In the present study, the severity of the knee osteoarthritis condition of an individual is assessed through the implementation of the hybrid CNN-based approach. The explainability of the artificial intelligence approach is also used to make sense of the results and judge the outcome. The present study is performed on the following aspects:

- The research involves curation and preprocessing of the data, balancing the data, and utilizing the publicly available Osteoarthritis Initiative (OAI) data set. The deep learning models DenseNet-152, ResNet-34 V2, and a hybrid model of ViT are trained and then tested for multi-class classification of KOA severity.
- Additionally, class imbalance is addressed through the use of techniques such as weighted random sampling and data augmentation. Grad-CAM ++ is used for localization of relevant regions of interest by visualizing the areas of interest



Fig 1. KL-graded X-ray images from the dataset used for this study

for the model. Performance evaluation is carried out through accuracy, precision, recall, F1-score, and confusion matrix analyses.

2 Methodology

In the proposed study, the applicability of the hybrid model in the classification of knee osteoarthritis will be evaluated, while the efficacy of the ResNet-34 model will be compared to other pre-trained CNN-based models. In the proposed model, the ResNet-34 model will be integrated with the Vision Transformer, which will focus on the global contextual information in the images, in the classification of knee osteoarthritis, while the identification of subtle structural variations in the knee joints will be improved. The entire workflow will include data preprocessing, training, evaluation, and visualization of the region of interest. The training process will be performed using the PyTorch platform.

- The ResNet-34 V2 model is highly popular in image classification, including medical image analysis, due to the optimal balance of depth, accuracy, and speed in the model architecture. The ResNet-18 model, although highly efficient in image classification, is not as effective as the other variants, i.e., ResNet-50, ResNet-101, and ResNet-152, due to the lesser depth in the model architecture, which results in the inability of the model to learn the complex features in the images. In addition, the lesser depth in the architecture results in the model being less efficient in image classification compared to the other variants in the ResNet family, while the model is faster in the training process compared to the other variants in the ResNet family.
- ViT, or Vision Transformer, works differently in image classification tasks. It divides an image into smaller segments called patches. It then applies a self-attention mechanism to recognize how different segments relate to one another. Unlike CNNs, Vision Transformer considers the overall picture, not just the minute details. Even though Vision Transformers can handle long-range relationships more effectively than CNNs, they sometimes require more data and computational power.
- GradCAM++ helps to get a better view of how an AI model works. It emphasizes areas in an image that have more impact on the overall prediction made by a model. It creates a heatmap, pointing out crucial areas in an image based on its gradient information, specifically in the last convolutional layers of a neural network. It is an improvement over GradCAM, offering a more detailed view of feature importance. It can be quite useful in medical images.

2.1 Dataset

(<https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity>)

In the proposed system, the publicly available Osteoarthritis Initiative (OAI) dataset is used, which contains X-ray images of the knee joint classified according to the Kellgren and Lawrence grading system, which is based on the severity of the osteoarthritis condition. The Kellgren and Lawrence grading system is based on the progression of the osteoarthritis condition, and the grades are as follows:

- Grade 0 indicates the absence of the condition, and the X-ray image shows no signs of the condition.

- Grade 1 indicates the presence of the condition, and the X-ray image shows some changes, but the changes are not clear.
- Grade 2 indicates the presence of the condition, and the X-ray image shows clear changes, which are mild.
- Grade 3 indicates the presence of the condition, and the X-ray image shows clear changes, which are moderate.
- Grade 4 indicates the presence of the condition, and the X-ray image shows clear changes, which are severe.

2.2 Preprocessing

The diagram below illustrates the different steps that need to be followed to preprocess and augment the data for the classification of knee osteoarthritis. Five different images related to knee X-ray are used for input to this classification system. The images are then subjected to a preprocessing step, where different image enhancements are carried out to make the images better and different. The contrast is improved through the application of the CLAHE filter, and different images are also subjected to salt and pepper noise, tilting, blurring, and pixel changes to make the model even stronger and to simulate real-world scenarios. Then, the system also deals with the problem of class imbalance, a common problem that is seen in medical image datasets. The approach ensures that a balanced class representation is maintained through the application of a random weighted sampler, and also an oversampling approach is followed to ensure that each class is represented with 10,000 images. Then, the images are augmented into an augmented dataset, which is characterized by its varied composition and uniform distribution. The augmented images form the final step of the system, and these are used to train the deep learning model, making the model even more accurate and capable of generalizing [Figure 2].

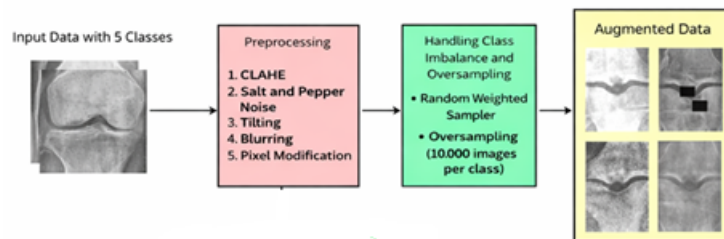


Fig 2. Preprocessing and Data Augmentation Model

2.3 Proposed Hybrid (Resnet-Vit) Deep Learning Architecture

The diagram above depicts a hybrid deep learning approach for differentiating between osteoarthritis and other possible conditions, utilizing enhanced X-ray images. In this approach, enhanced images are first passed through a Vision Transformer (ViT) model. In this approach, an overall understanding of image-level characteristics is achieved through an analysis of all possible interrelationships between patches of images. Then, feature representations of images, based on this approach, are further passed through a ResNet-34 V2 model. This model is highly effective in extracting detailed local information, such as edges and textures, based on its residual learning approach. The feature representations developed in this approach have an overall size of 2048 bytes, encapsulating all detailed information regarding images.

Then, all individual feature representations are further concatenated to develop a unified feature vector. Further, this aggregated feature set is passed through a classification module. In this module, dropout layers have been incorporated to avoid overfitting, and fully connected layers have been utilized for generating the ultimate prediction for osteoarthritis severity. Additionally, Grad-CAM++ is incorporated to develop heatmaps, representing areas within X-ray images having maximum impact on overall determinations made by this approach. This helps in developing an overall comprehension of this approach among medical professionals.

2.4 Experimental Setup

The experimental setup is developed using Python programming and the Keras library for deep learning. The training of the model is done on a machine with an NVIDIA GeForce RTX 2080 GPU with CUDA support. The data is preprocessed and data augmentation techniques are applied to handle class imbalance for model training. To avoid overtraining, early stopping is implemented based on validation loss with a patience of 15. The performance of the model is measured based on various parameters like accuracy, precision, recall, F1 score, and confusion matrix. Table 1 depicts the details of model training.

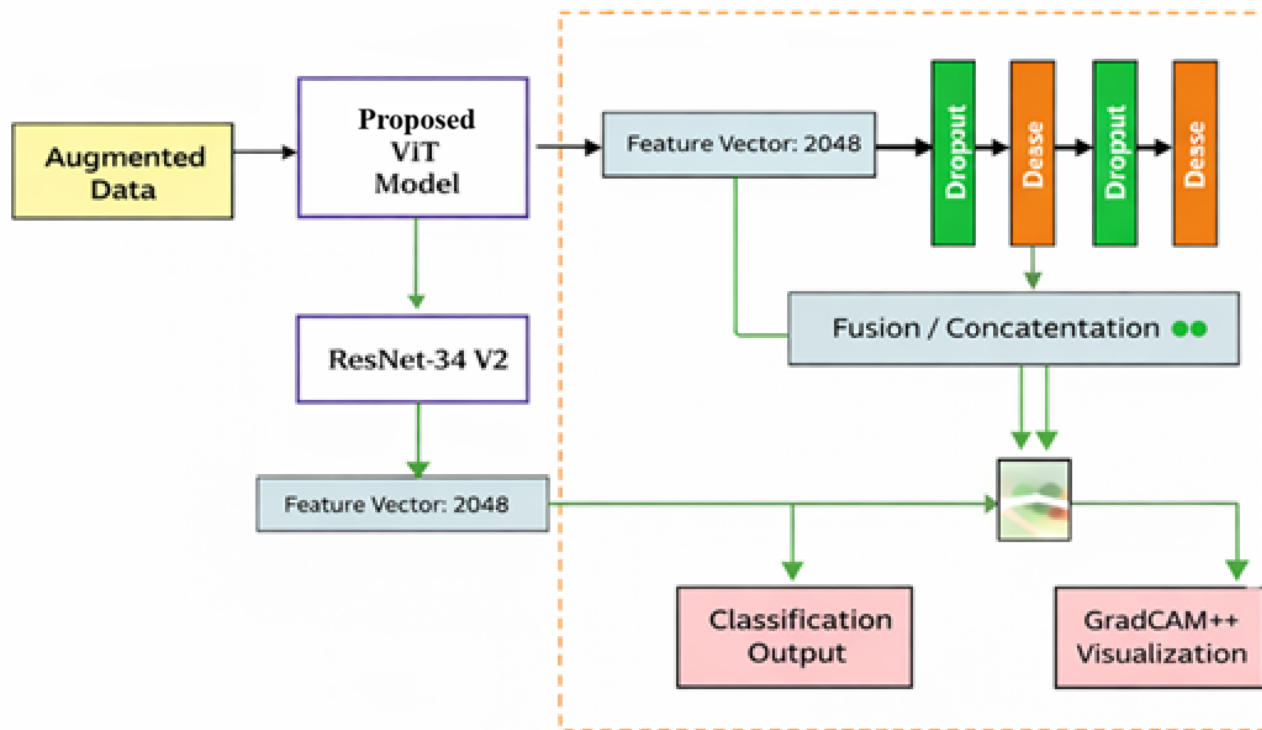


Fig 3. Proposed ResNet-ViT Model Architecture

Table 1. The parameters used for training all the models

Parameter	Name / Value
Platform	NVIDIA RTX 2080 Ti, 4GB VRAM
Framework	PyTorch
Optimizer	Adam
Loss function	CrossEntropyLoss
Batch size	16
Training epochs	30

2.5 Model Training

In the classification part of the suggested model, the severity of the knee osteoarthritis will be classified into five classes. This is comparable to the Kellgren-Lawrence grading system, which has a range of 0 to 4. The categorization will be performed using a fully connected layer, which will consist of five neurons in the output layer. After that, the softmax activation function will be used to calculate the chances of the classes. The CrossEntropyLoss function can be used to train the model. This function is very helpful when there are more than one class, as it will show the difference in the classes. One of the challenges in the medical image dataset is the imbalance in the classes, i.e., the classes will have a varying number of samples, with a large difference in the number of samples in the classes. As can be seen from Table 2, the dataset has an uneven number of samples in the classes, i.e., the number of samples in Class 4 is only 173, while the number of samples in Class 0 is 2286. This would, therefore, result in a form of biased learning, where the model would tend to lean more towards learning from the majority classes, thereby compromising its ability to predict the minority classes. This, therefore, results in a reduction in sensitivity towards severe cases of osteoarthritis, which are of critical importance. To overcome this, we use a weighted random sampling approach, which would help address the issue of class imbalance before the model is trained. By doing this, we ensure that all classes are adequately represented during the creation of the batch, thereby allowing the model to learn features from all classes without bias. Additionally, in order to balance the data, oversampling is employed, where each class is augmented to 10,000

images. To augment the data and improve the model's robustness, various augmentation techniques are employed, including contrast, noise, geometric, and pixel augmentation. After applying all these balancing techniques, uniform distribution of data is observed in all classes, as shown in Table 2.

This is because the model's ability to generalize is significantly improved by the use of this well-balanced data set, thus leading to an improvement in the model's performance over all levels of severity. The predictions are, therefore, more accurate and less biased, especially for classes that are previously underrepresented. Moreover, the need to make the model more interpretable is satisfied by the inclusion of the Grad-CAM++ method. The method enables the generation of heat maps that highlight the most important regions of the knee X-rays, thus illustrating the importance of the regions to the model's predictions. The importance of visual aids cannot be overstated, especially in the medical field.

Table 2. Image Count Before v/s After Handling Class Imbalance

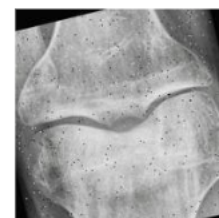
Class	Before Handling Class Imbalance	After Handling Class Imbalance
0	2286	10000
1	1046	10000
2	1516	10000
3	757	10000
4	173	10000



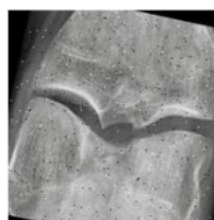
CLAHE



Pixel Modification



Tilting



Salt and Pepper Noise



Blurring

Fig 4. Augmented Data

3 Results and Discussion

3.1 Test report of the DenseNet-121 Model Multiclass Classification

The model demonstrates excellent efficacy with regard to more serious classification levels, namely Class 3 and Class 4. Conversely, the model's efficacy is reduced with regard to Class 1, as validated by an F1-score of 39.51. This is due to the difficulties encountered in the accurate detection of early-stage osteoarthritis. This is due to the nuanced and intermingled nature of the characteristics of the neighboring classes. Moreover, the focus on extracting local features also limits the ability to

accurately distinguish global contextual relationships in the image. To address these limitations, the subsequent model is used to improve the accuracy of classification [Table 3].

Table 3. Test report of the DenseNet-121 Model Multiclass Classification

Class	Precision	Recall	F1 - score
0	74.28	72.54	73.26
1	35.36	43.25	39.51
2	69.74	59.32	64.20
3	73.26	78.64	76.34
4	85.64	81.72	83.46

3.2 Test Report of the Resnet-50 V2 Model Multiclass Classification

The model has performed adequately in most of the classes, but it has faced challenges in the first class, as indicated by the F1 score, which is merely 39.51. This indicates a challenge in the identification of early stages of osteoarthritis. Classes 3 and 4 have performed better compared to other classes, but the performance of the model is hampered due to inadequate modeling of the complex relationships among the intricate features. The model has a tendency to focus on the local features, which are not sufficient to capture the contextual features within the images [Table 4].

Table 4. Test report of the Resnet-50 V2 Model Multiclass Classification

Class	Precision	Recall	F1 - score
0	70.78	72.54	72.63
1	35.36	43.25	39.51
2	64.44	58.32	62.20
3	73.26	78.64	76.34
4	80.64	81.72	80.46

3.3 Test Report for ResNet34-ViT Model Multiclass Classification

The improvement in the performance of the model is clear, especially in most of the categories, such as Class 0 and Class 3, with higher F1-scores, implying consistency in overall performance during classification. The model also has a higher recall rate, 88.24, for Class 4, which is critical in the accurate diagnosis of advanced stages of osteoarthritis. Although the performance for Class 1 remains relatively lower, a slight improvement can be seen compared to previous versions, implying a relatively higher ability to learn complex classifications. Therefore, this model can be considered more reliable for multi-class classification, providing a fair balance in terms of precision, recall, and overall ability to generalize [Table 5].

Table 5. Test report of the ResNet34-ViT Model Multiclass Classification

Class	Precision	Recall	F1 - score
0	75.26	78.56	76.88
1	35.58	46.28	40.23
2	74.69	54.14	62.78
3	81.08	80.72	80.90
4	77.59	88.24	82.57

3.4 Classification report of the best performing model for Multiclass Classification

Finally, the performance of the three deep learning models, DenseNet-152, Resnet-50 V2, and the proposed hybrid model, was evaluated and compared. Out of the three, the proposed hybrid model performed better in terms of accuracy, achieving an accuracy of approximately 66.79%. Moreover, the GradCAM++ visualisation of the image successfully highlighted the most relevant region of the knee joint, which is associated with the severity of the disease, thereby providing an insightful understanding of the decision-making process of the model [Table 6, Table 7].

Table 6. Classification report of the best performing model for multiclass classification

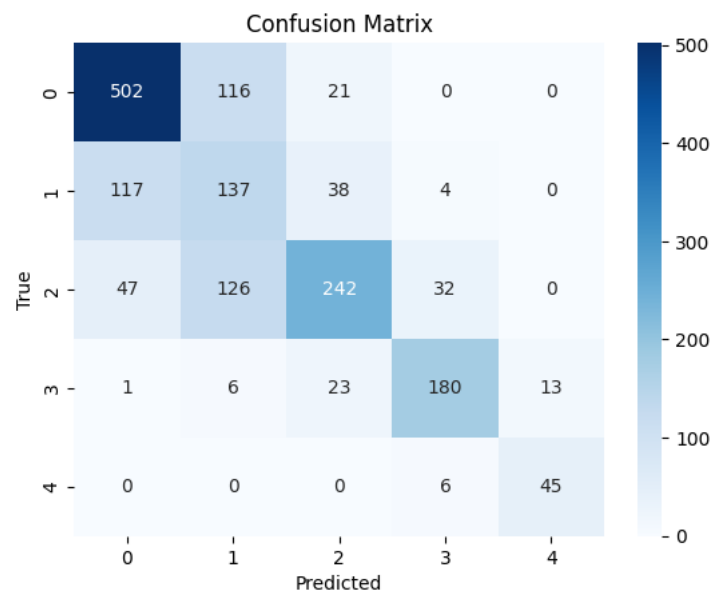
Class	Precision	Recall	F1 - score
0	75.26	78.56	76.88
1	35.58	46.28	40.23
2	74.69	54.14	62.78
3	81.08	80.72	80.90
4	77.59	88.24	82.57

Table 7. Overall Accuracy of the Models

Models	Overall Accuracy
Resnet 50 V2	65.36%
Densenet-121	66.31%
Resnet34 + ViT	66.79%

3.5 Confusion Matrix for the Resnet34 + ViT model

Figure 5 illustrates the confusion matrix for the hybrid ResNet-34 V2 + Vision Transformer model, which provides an explanation of the relationship between the observed classifications and the predicted classifications. The diagonal values represent the correct classifications, which indicate the model's ability to correctly predict the classes corresponding to the higher severity levels. On the other hand, the off-diagonal values represent the misclassifications, which are mainly confined to the early stages due to the feature characteristics of these classes.

**Fig 5.** Confusion Matrix for Resnet34 + ViT mode

3.6 GradCAM++ Visualization

The following diagram demonstrates how XAI can be utilized in explaining how an AI model classifies the disease of knee osteoarthritis based on its five levels of severity, classified under Classes 0 to 4. A color-coded image, referred to as Grad-CAM, is utilized to identify areas of the image that the AI model considers to be of utmost importance in its decision-making process. In the initial stage, where no disease is yet present (Class 0), the model focuses on the joint space, which is obviously unobstructed. In subsequent stages of the disease, specifically in Classes 3 and 4, the diagnostic focus of the AI shifts to clinical markers such

as osteophytes, joint space narrowing, and osseous contact. In this regard, it can be seen in the diagram how the diagnostic process of the AI aligns with clinical abnormalities identified in images of osteoarthritis. The Color Scale Representation are,

- Red/Pink Color: These areas represent the background, where minimal impact on the score is expected.
- Yellow/Green Color: These areas represent secondary features to help guide the AI within the knee joint.
- Blue/Violet Color: These areas represent the "smoking gun" feature that helps identify the Osteoarthritis grade.
- High Intensity: These areas have a stronger "confidence" level for the presence of clinical markers.
- Large Spread: These areas provide a "holistic" view of how the joint looks in terms of wear and tear [Figure 6].

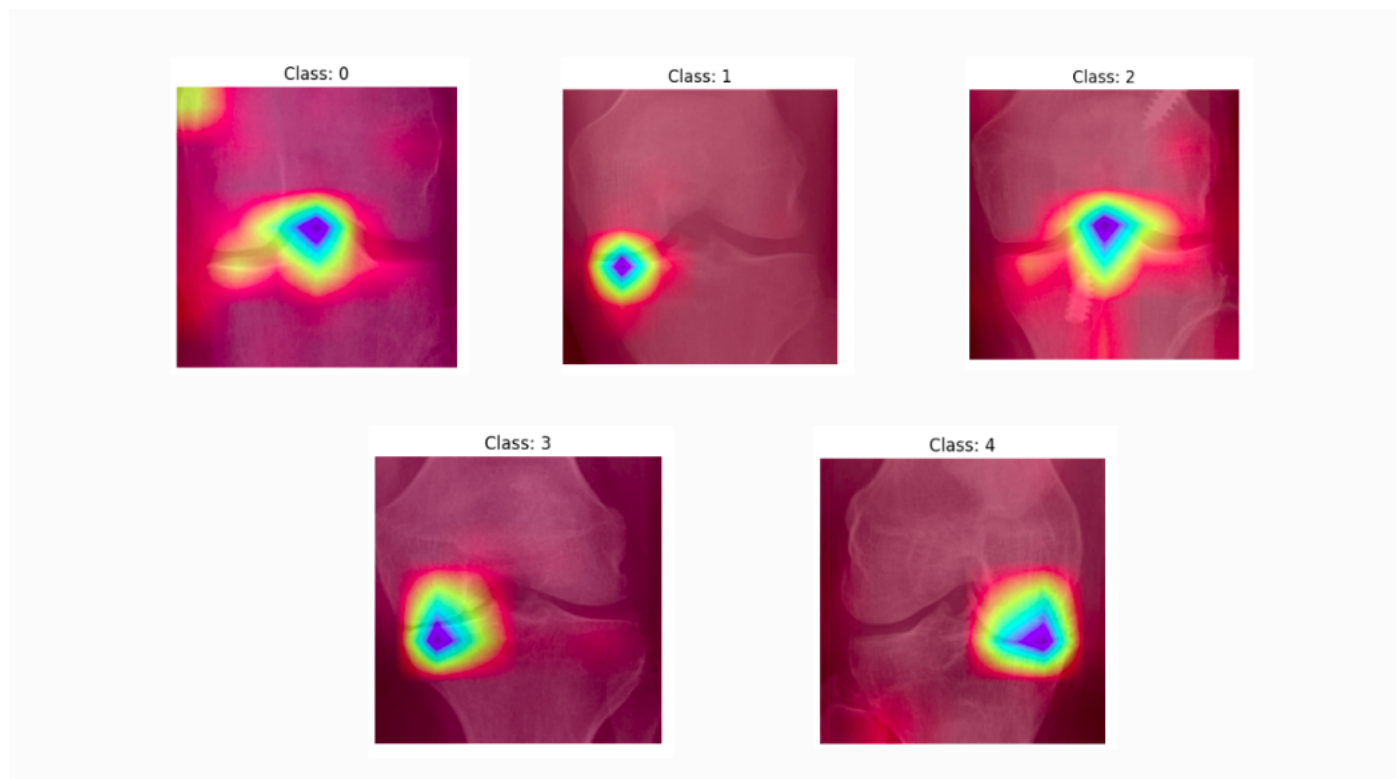


Fig 6. GradCAM++ Visualization to highlight Region of Interest

4 Conclusion

This paper has effectively demonstrated how deep learning hybrid model can be used for the early detection of Knee Osteoarthritis (KOA) from X-ray images. The Kellgren-Lawrence grading system is applied to assess the severity of KOA using advanced models such as the ResNet34-ViT hybrid model, which achieved an accuracy of 66.79%. To overcome the dataset limitations, data augmentation and class balancing techniques have been strategically employed to enhance the capabilities of the model. The data augmentation techniques utilized in this research work include Weighted Random Sampling and data augmentation. The Grad-CAM++ algorithm was utilized for validating the results obtained from the model, highlighting crucial areas in the Knee Joint region. The capabilities of the proposed model, which utilize both local spatial features and global contextual data, have been found extremely useful in distinguishing between subtle differences in Knee Osteoarthritis severity, especially in the case of grades 1 (doubtful) and 2 (mild) KOA, which fall under the intermediate range.

Moreover, this proposed system has been designed in such a manner that future research could be carried out using other joint-related diseases or data, thus making this proposed system more versatile. The proposed research work has introduced a diagnostic tool that uses artificial intelligence, which has been made scalable and interpretable, thus making it easier to be incorporated into various radiology departments. The proposed tool has immense capabilities, including increasing accuracy, eliminating discrepancies between observers, and speeding up the process. The proposed research work has effectively utilized

an integrated approach, thus making it a major step forward in Knee Osteoarthritis diagnosis, where precision, accuracy, and utility have been incorporated.

References

- 1) Ahmed R, Imran A. Knee osteoarthritis analysis using deep learning and XAI on X-rays. *IEEE Access*. 2024;12:1–12. <https://doi.org/10.1109/ACCESS.2024.3400987>.
- 2) Rehman A, Alghofaily B, Saba T, Raza A, Alamri FS. Transfer learning-based smart features engineering for osteoarthritis diagnosis from knee X-ray images. *IEEE Access*. 2023;11:152040–152053. <https://doi.org/10.1109/ACCESS.2023.3294542>.
- 3) Nisha, Chhabra M. Deep learning based knee osteoarthritis grade classification. *Proceedings of the INDIACom Conference*. 2024. Greater Noida, India. <https://doi.org/10.23919/INDIACom61295.2024.10499112>.
- 4) Jain RK, Sharma PK, Gaj S, Sur A, Ghosh P. Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. *Comput Biol Med*. 2023;137:104799. <https://doi.org/10.1016/j.compbimed.2021.104799>.
- 5) Kitukale G, Shelke NA, Agrawal R, Singh NP, Quamara S. Predicting knee osteoarthritis using deep neural network. *Proceedings of the IEEE I2CT Conference*. 2024. <https://doi.org/10.1109/I2CT61223.2024.10543978>.
- 6) Chetlapalli AK, Bhuvaneswari R. Advanced knee osteoarthritis detection using deep learning. *Proceedings of the IEEE I2CT Conference*. 2024. <https://doi.org/10.1109/I2CT61223.2024.10543594>.
- 7) Zhao H, Ou L, Zhang Z, Zhang L, Liu K, Kuang J. The value of deep learning-based X-ray techniques in detecting and classifying K-L grades of knee osteoarthritis: a systematic review and meta-analysis. *Eur Radiol*. 2024;35(1):327–340. <https://doi.org/10.1007/s00330-024-10928-9>.
- 8) Rawat A, Dhyani A, Panwa D, Bisht G, Vats S, Sharma V. Knee osteoarthritis prediction using transfer learning. *Proceedings of the IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*. 2024. <https://doi.org/10.1109/AIC61668.2024.10731036>.
- 9) Moura FA, Pelegrinelli ARM, Catelli DS, Lamontagne M, Torres RS. Knee osteoarthritis prediction from gait kinematics: exploring the potential of deep neural networks and transfer learning methods for time series classification. *J Biomech*. 2025;p. 112881. <https://doi.org/10.1016/j.jbiomech.2025.112881>.
- 10) Sohail M, Azad MM, Kim HS. Knee osteoarthritis severity detection using deep inception transfer learning. *Comput Biol Med*. 2025;186:109641. <https://doi.org/10.1016/j.compbimed.2024.109641>.
- 11) Nurmirinta TAT, Turunen MJ, Tohka J, Mononen ME, Liukkonen MK. Predicting knee osteoarthritis severity from radiographic predictors: data from the Osteoarthritis Initiative. *Ann Biomed Eng*. 2025. <https://doi.org/10.1007/s10439-025-03740-z>.
- 12) He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. <https://doi.org/10.1109/CVPR.2016.90>.
- 13) Harish H, Patrot A, Bhavan S, Gousiya S, Livitha A. Knee osteoarthritis prediction using deep learning. *Proceedings of the IEEE ICRAIS Conference*. 2023. <https://doi.org/10.1109/ICRAIS59684.2023.10367065>.
- 14) Sharma A, Singh J, Kumar A, Bajaj V, Gupta S. Comparative analysis of deep learning models for early detection of osteoarthritis using knee radiographs: a retrospective study. *J Clin Orthop Trauma*. 2025;p. 103212. <https://doi.org/10.1016/j.jcot.2025.103212>.
- 15) Vaattovaara E, Panfilov E, Tiulpin A, Niinimäki T, Niinimäki J, Saarakkala S, et al. Kellgren-Lawrence grading of knee osteoarthritis using deep learning: diagnostic performance with external dataset and comparison with four readers. *Osteoarthritis Cartilage*. 2025;p. 100580. <https://doi.org/10.1016/j.ocarto.2025.100580>.
- 16) Litjens G, Kooy T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. <https://doi.org/10.1016/j.media.2017.07.005>.
- 17) Pateriya N, Jain P, Niveditha KP, Tiwari V, Vishwakarma S. Deep residual networks for image recognition. *Int J Innov Res Comput Commun Eng*. 2023;11(9). <https://doi.org/10.15680/IJIRCCCE.2023.1109026>.
- 18) Kinger S. Deep learning for automatic knee osteoarthritis severity grading and classification. *Indian J Orthop*. 2024;58(10):1458–1473. <https://doi.org/10.1007/s43465-024-01259-4>.
- 19) Rani S, Memoria M, Almogren A, Bharany S, Joshi K, Altameem A, et al. Deep learning to combat knee osteoarthritis and severity assessment by using CNN-based classification. *BMC Musculoskelet Disord*. 2024;25:79. <https://doi.org/10.1186/s12891-024-07942-9>.
- 20) Kishore VV, Padmini GH, Sudharshan I, Tejaswini K, Roopa M, Inaganti R. Predicting knee osteoarthritis grades using deep learning: an extensive examination. *Proceedings of the Advances in Intelligent Systems and Computing Conference*. 2024. https://doi.org/10.2991/978-94-6463-471-6_71.