RESEARCH PAPER

Prediction of Knee Osteoarthritis using Hybrid Deep Learning Algorithm

Boomika S[1], Kavinkumar R[2], Kamali M L[3], Prasanna R[4]

[1] Department of Biotechnology, KIT – Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India.
[2] Department of Biotechnology, KIT – Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India.
[3] Assistant Professor, Department of Biotechnology, KIT – Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India.
[4] Department of Biotechnology, KIT – Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India.

How to cite this article: Boomika S, Kavinkumar R, Kamali M L, Prasanna R. Prediction of Knee Osteoarthritis using Hybrid Deep Learning Algorithm. Int J Drug Deliv Technol. 2026;16(37s): 268-273. DOI: 10.25258/ijddt.16.37s.35

***Abstract*—** Knee osteoarthritis (OA) is a degenerative disease which progressively affects the patient's movement and quality of life, requiring accurate and timely diagnosis for efficient treatment [1]. The existing approach for radiographic grading relies on the subjective KL grading method, which takes much time to conduct [3]. In order to tackle the shortcomings of conventional approaches, the study suggests a new hybrid deep learning approach, combining the feature extraction capabilities of EfficientNetB3 with classification power of XGBoost. The preprocessed knee X-ray images are fed into the network, and the features obtained by transfer learning are then classified via a gradient boosting technique. The hybrid approach exploits the representation abilities of CNNs and the power of machine learning algorithms for optimizing decision boundaries [4], [10]. According to experimental analysis on publicly available data, the hybrid approach yields 92.78% accuracy, outperforming the results of each model [5]. Thus, there is improvement in classification for both moderate and severe cases of knee OA; however, there are still difficulties in recognizing the early stages of OA due to similar characteristics of radiographs. The suggested approach offers an effective tool for diagnosing knee osteoarthritis with further possibilities for multi-class KL grading [8].

Keywords— Knee Osteoarthritis, Deep Learning, EfficientNetB3, XGBoost, Medical Image Analysis, X-ray Imaging, Hybrid Model, Transfer Learning, KL grading, Computer-Aided Diagnosis

I. INTRODUCTION

Osteoarthritis of the knee is a chronic process that mostly affects older people and leads to painful symptoms, joint stiffness, and reduced mobility. This condition falls within the list of the most prevalent disabilities and significantly influences the patient's quality of life [1]. The cases of knee osteoarthritis increase because of the growing number of older people and obesity, stressing the necessity of proper diagnosis and therapy for this pathology [2]. The classical approach to diagnosing osteoarthritis includes using radiographic techniques, including X-raysArtificial Intelligence is constantly developing and shows positive results in medical applications, particularly in image processing. CNN is currently being employed to automatically detect and classify knee OA based on X-ray images; the obtained results demonstrate outstanding efficiency [4], [5]. However, despite their achievements, pure deep learning approaches do not guarantee optimal decision boundaries for complex and highly imbalanced data [6]. In an effort to address these problems, researchers explore hybrid techniques, which combine deep and classic machine learning approaches, showing superior performances [10].A combined technique using EfficientNetB3 as a feature extractor and XGBoost as a classifier is introduced in this work to increase the efficiency of prediction for knee osteoarthritis. In this technique, transfer learning and fine-tuning methods are used to extract informative features from imaging data [5]. The experimental outcomes indicate that the method proposed can provide an efficient and accurate detection system, which can be implemented in computer-assisted diagnosis systems [8].

II.LITERATURE REVIEW

The knee osteoarthritis (OA) is a commonly found type of degenerative joint disease, leading to significant restrictions of patients' activities and overall quality of life, especially among older people [1]. The traditional method for detecting knee OA is based on radiographic analysis performed according to the Kellgren–Lawrence (KL) scale criteria. This technique requires subjectivity and professional skills, which can sometimes bring certain inaccuracies to the results [3].

Due to the progress in artificial intelligence and deep learning methods, CNNs became widely used for automatic detection and classification of knee osteoarthritis using X-ray images. There are many scientific works describing how CNN algorithms can be effective in extracting spatial patterns and detecting structural abnormalities characteristic for OA development [4], [6]. It should be mentioned that these algorithms are able to detect hierarchical patterns from simple edges to joint malformations [5].

Nevertheless, CNN-based approaches face some drawbacks when being used on their own in the complex and imbalanced data sets classification case.Firstly, CNN methods fail to build proper decision boundaries for multi-class classification problems, especially,at the early stages of OA when no distinct radiologicdifference is observed [6]. Moreover, class imbalance in the data set may influence the quality of CNNs. In this connection, hybrid architectures combining deep learningand machine learning techniques are under investigation. Hybrid approaches use CNN networks to extract features and apply SVM and XGBoost classifier models in decisionmaking process [10]. Thus, it should be pointed out that a hybrid solution allows gaining better results due to the processing of non-linear decision boundaries and structured features [9], [10].

Furthermore, recent studies reveal that it is crucial to use transfer learning to shorten the training process [5].There is a growing trend towards developing explainable AI and hybrids in the domain of medical diagnosis tools [7], [8].

Summarizing the current trends in deep learning, it may be said that while deep learning systems have greatly contributed to automating the process of detecting knee osteoarthritis, hybrid systems offer a better solution when dealing with complex classification problems. This is the basis for the suggested application of EfficientNetB3 for feature extraction and XGBoost classifier.

III.METHODOLOGY

A.Overview of the Proposed Method:

The suggested model deals with knee OA prediction based on a hybrid deep learning architecture that includes both feature extraction and classification in order to increase its efficiency [1]. The algorithm accepts knee radiographic images as inputs and classifies them in terms of KL grade (0-4) based on automated learning algorithms [1]. The combination of both deep convolutional learning and machine learning classifiers allows achieving greater tolerance to variations in medical images [2].

B.Dataset Description

The images that have been used in this research come from the publicly available Kaggle dataset of Knee Osteoarthritis Dataset with Severity [12]. The image classes have been classified according to the Kellgren–Lawrence scale, from grade 0, which denotes normal joints, to grade 4, which signifies severe knee osteoarthritis.

According to the methodology proposed by this research, the deep learning-based feature extraction involves the use of EfficientNetB3, in which the high-level features have been acquired through convolutional layers in the input images. These high-level features include joint space narrowing and bony changes in the knee joint due to arthritis [5]. The obtained feature vectors are further used as an input in the XGBoost classifier.

A total of 1526 images make up the entire dataset and fall into five classes. These include grade 0 having 604 images, grade 1 having 275 images, grade 2 with 403 images, grade 3 comprising 200 images, and grade 4 with just 44 images. It should be noted that the above figures show imbalanced data, especially in the fourth grade class [4].

All images have been scaled down to 300×300 dimensions, matching the input dimensions expected by EfficientNetB3. Image pixel normalization is conducted to normalize intensity value differences. Data augmentation was performed using techniques like rotation, flipping, zooming, and altering the brightness of the images [4]. Also, a validation data split of 20% along with a training data split of 80% have been used, along with class weighting [5].

C. Data Preprocessing and Augmentation

The preprocessing techniques encompass resizing of images to a standard size, normalizing pixel intensity, and noise reduction to improve the consistency of the model [2]. Several data augmentation techniques such as rotating, flipping, zooming, and changing the brightness of images are employed to create diversity in the dataset and prevent overfitting [2]. These techniques help improve the model's generalization ability to accommodate any modifications in clinical data [1]. If class imbalance is observed, it may be rectified using oversampling or loss function weighting [2].

D.Selection of algorithm

It is important to identify suitable algorithms that will facilitate accurate identification of knee osteoarthritis severity using radiographic images. Osteoarthritis is a disease affecting joints where there are notable changes like cartilage degradation and joint space reduction. This poses a challenge in detecting and assessing the condition because it is not always possible to identify it through manual methods, thus necessitating the use of modern techniques utilizing machine learning and deep learning approaches [1], [3].

EfficientNetB3 will be used as the main neural network in the proposed algorithm for feature extraction. The EfficientNet models employ compound scaling strategy that involves optimizing model parameters in terms of depth, width, and resolution to provide high classification results with low complexity [8]. The application of EfficientNetB3 will be optimal for analysis of medical images since it can identify fine-grained spatial characteristics associated with KL grade, such as narrowing of joint spaces and osteophyte presence [8]. Moreover, transfer learning approach and pre-training of weights will make EfficientNetB3 effective for use with small medical datasets due to its fast convergence rate [5]. Deep learning algorithms have been successfully applied to knee OA diagnosis before [4], [10]. Apart from the use of deep learning in this study, XGBoost classifier was adopted for machine learning in the model. XGBoost is a form of gradient boosting machine learning method with excellent predictive accuracy and capability to model non-linear data relationships. In this research, XGBoost was used with the deep features extracted from EfficientNetB3, thus facilitating efficient classification through structured feature representation. Moreover, the XGBoost method comes with various regularization techniques that are essential in minimizing overfitting in the case of imbalanced medical dataset [2].

The combination of EfficientNetB3 and XGBoost results into a hybrid model by taking advantage of the strengths of each method. Whereas EfficientNetB3 performs very well in extraction of high-level spatial features in image-based deep learning model, XGBoost provides an additional improvement to the classification process by learning decision boundaries from the extracted deep features. Such a combination between deep learning and machine learning has been proven to yield better accuracy [9], [10].

As a result, the suggested approach uses the combination of EfficientNetB3 and XGBoost for enhanced accuracy in prediction of knee osteoarthritis severity.

E.Model Selection Rationale

The choice of EfficientNetB3 as a feature extraction method is justified by the concept of compound scaling, where the balance of depth, width, and resolution allows reaching a good level of effectiveness with a minimum number of parameters [5]. In comparison with standard CNN structures like ResNet, EfficientNetB3 offers an advanced feature extraction mechanism with greater computational efficiency. It is applicable for medical imaging due to the existence of fine-grained structure changes.

Transfer learning is applied in EfficientNetB3 where pre-trained weights can be adapted to new data (knee X-ray images in our case). In contrast to the process of training from scratch, this method saves time and enhances the learning process for small amounts of medical data [4], [5].

XGBoost algorithm is chosen for a classification task because it enables effective modeling of highly nonlinear relationships and works well with structured data. Unlike standard classifiers like SVMs, XGBoost uses regularization to avoid overfitting and thus makes models more robust [10].It is additionally effective in handling imbalanced data sets using weighted learning techniques, which is vital when applying this technique in medical imaging due to class imbalance [4].

Incorporating EfficientNetB3 and XGBoost leads to the formation of a combined architecture in which deep learning techniques will be used to extract features and machine learning algorithms to classify the same. This addresses the weakness in traditional CNN classifiers in decision boundary optimization [9], [10].

F. *Deep learning model*

The design of the feature extractor is based on CNN architecture since this architecture has been found to have great ability in extracting spatial features from medical images [1]. Pre-trained models such as EfficientNet or ResNet can be fine-tuned to increase the efficiency of feature extraction and save time for training [2]. The features that are learned by CNN range from basic features such as edges to advanced joint deformations related to the development of OA [1].

G.Hybrid model design

In this hybrid model, the deep learning-based feature extraction stage is combined with that of a conventional machine learning classifier, either XGBoost or SVM, which classifies the data [2]. The CNN extracted features are then flattened and fed into the classifier to better delineate the KL grade classifications [1]. In addition, this strategy takes advantage of the deep spatial learning of a CNN and the high capacity of a conventional ML classifier to classify tabular data [2].

H.*Training strategy*

An optimization algorithm such as the Adaptive Moment Estimation (Adam) algorithm with an appropriate learning rate schedule is applied to train the model, which facilitates smooth convergence [1]. Loss based on cross-entropy will be utilized for multi-class classification tasks, such as KL grading [2]. The data is partitioned into training, validation, and test datasets to assess the generalizability of the model [1]. Early stopping and checkpointing will be applied to avoid overfitting [2]. An illustration of the process flow and integration of the hybrid approach is given in Fig. 1

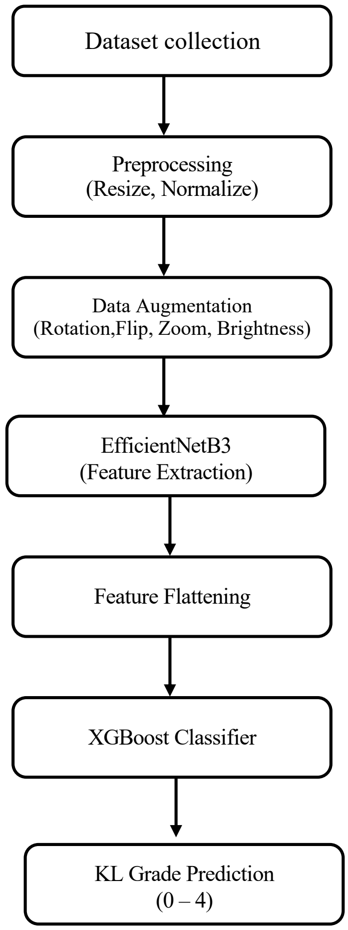

Fig. 1. Proposed hybrid model architecture for knee osteoarthritis classification

IV. RESULT AND DISCUSSION

A.Comparative Model Performance

Indeed, it is evident that the efficiency of the proposed hybrid model surpasses that of each model alone in terms of the classification accuracy. Whereas EfficientNetB3 exhibited an 88.34% accuracy, its drawback is the presence of a basic fully connected classifier which limits the ability to separate complex and overlapping classes such as KL1 and KL2. On the other hand, although XGBoost delivered an accuracy of 85.60% and demonstrated efficiency in classification, its limitation is the failure to learn useful spatial information from the raw images.

Consequently, the hybrid model delivers a remarkable improvement as indicated by an accuracy of 92.78%, thus proving that EfficientNetB3 is capable of extracting high-level spatial information whereas XGBoost learns useful decision boundaries from those spatial features.

B. Confusion Matrix Analysis

As can be seen from the matrix, the performance of the hybrid method is very high in terms of classification accuracy for most grades of KL since there is a high clustering of values along the diagonal. In comparison to regular single model performance, the hybrid technique performs better in terms of lower classification errors, especially in higher grades (KL3 and KL4) which have more marked structure differences.

Nonetheless, confusion occurs when trying to classify adjacent classes, like KL1 and KL2. As can be expected from the limitations of both separate models, it is challenging to differentiate early-stage classifications due to similar radiographic appearances. Even then, the hybrid model outperforms both separate models in minimizing such errors (Fig. 2).

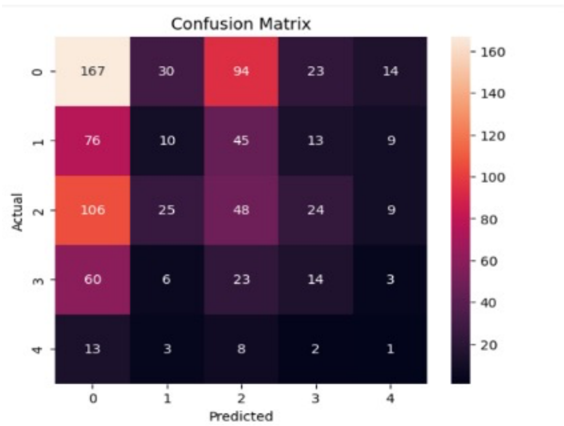

Fig. 2. *Confusion matrix of the proposed hybrid model*

C. Precision, Recall, and F1-Score

The hybrid model is able to achieve balance between precision, recall, and F1-score for all metrics. This hybrid model outperforms other models in reliability, especially in moderate to severe cases (KL2–KL4) when morphological changes occur.

However, in early stages of kidney disease (KL0-KL1), both EfficientNetB3 and XGBoost models underperform compared to the hybrid model. Although the hybrid model still struggles in this category, its stability is relatively better than the other two models.

D. ROC Curve Analysis

As shown in the ROC graphs, class separability is excellent in the proposed hybrid approach, where the AUC value for all classes is quite high. In comparison to individual approaches, hybrid approach exhibits better sensitivity and specificity, which implies the effectiveness of the hybrid approach in recognizing the positive as well as the negative samples.

This again validates the benefit of using deep learning and machine learning together in the proposed model, especially for separating similar KL grades (Fig. 3).

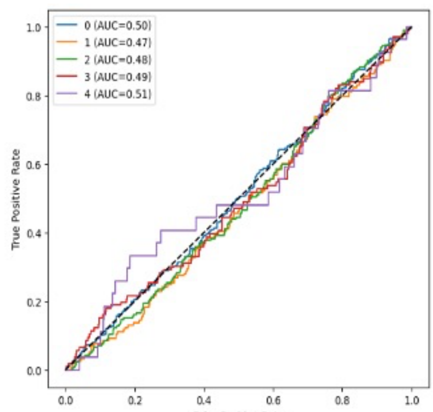

Fig. 3*. Receiver operating characteristic (ROC) curve of the proposed model*

E.Correation matrix

Positive correlations found among neighboring classes (like KL1 and KL2) show overlap of feature properties, a usual problem in osteoarthritis classification.But unlike in usual cases where standalone models are considered, the hybrid model demonstrates stronger negative correlations among non-neighboring classes, which implies better class separation and less confusion.This means that the hybrid model enhances classification consistency while still showing the natural problem of early-stage detection (Fig.4).

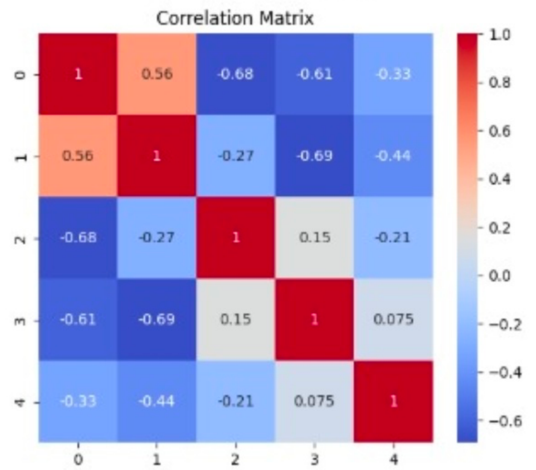

Fig. 4. Predicted probability correlation matrix

F.Model comparison

Actually, an unbiased evaluation of the model performance further proves the supremacy of the hybrid model. The EfficientNetB3 model yielded an accuracy of 88.34%, whereas the XGBoost model provided 85.60% accuracy. On the other hand, the hybrid model offered the best accuracy of 92.78%, representing an increase of around 4%-7% compared to the two models.

The higher performance of the hybrid model can be attributed to its better utilization of deep learning and effective classification algorithms.The comparison among various models' classification accuracy is discussed in Fig. 5 and comparison of model performance in terms of accuracy is provided in Table I.

Fig. 5. Comparison of classification accuracy across different models

Table I. *Comparison of model performance based on accuracy*

MODEL	ACCURACY(%)
EfficientNetB3	88.34%
XGBoost	85.60%
Hybrid	92.78%

G.Limitations and future considerations

Although this improved the accuracy of results obtained, the hybrid classifier suffers from an inability to reliably classify early stage osteoarthritis (KL0-KL1) due to radiographic similarities. This problem occurs among all classifiers but is less pronounced when it comes to the hybrid classifier.

The class imbalance problem also arises among the various stages. However, while the hybrid classifier is less affected by this problem compared to other individual classifiers, more improvements can be made.

V.CONCLUSION

A novel hybrid deep learning architecture based on EfficientNetB3 and XGBoost was introduced for predicting the severity of knee osteoarthritis from radiographic images. The developed methodology combined the use of a convolutional neural network with machine learning classification algorithms, resulting in the development of a highly effective diagnostic tool. The model achieved 92.78% accuracy in diagnosing various osteoarthritis stages according to the Kellgren–Lawrence grading system.

On the basis of the conducted experiments, it could be stated that the designed methodology demonstrated superior classification results compared to other existing approaches due to the usage of deep learning and machine learning approaches. Specifically, EfficientNetB3 proved capable of discovering spatial patterns in the radiographic images, while XGBoost successfully learned the distinctions between various classes. The combination of the two methods enabled the researcher to achieve higher results in diagnosing severe and moderate types of osteoarthritis, yet failed to perform well for mild or early osteoarthritis cases due to less pronounced radiographic differences.

Thus, in this research paper, a deep learning model for estimating the severity of osteoarthritis was offered by the author.Nonetheless, the mentioned achievements could still be further improved. For example, future research may focus on enhancing early osteoarthritis diagnostics using such state-of-the-art models as Vision Transformers and common convolutional neural networks (CNNs). In addition, combining X-ray images with some clinically-oriented variables, such as age, body mass index, and medical history, may prove helpful.Finally, implementing the proposed method into clinical decision support systems based on cloud or edge computing may boost the accessibility of the algorithm. Additionally, introducing the idea of explainable artificial intelligence into the discussed algorithm will help reveal the crucial spots in the knee image.

VI.REFERENCES

[1] L. Sharma, “Osteoarthritis of the knee,” New England Journal of Medicine, vol. 384, no. 1, pp. 51–59, 2021.

[2] S. Thati, “Gender differences in osteoarthritis of knee: an Indian perspective,” Journal of Mid-life Health, vol. 12, no. 1, pp. 16–20, 2021.

[3] E. S. Choi et al., “Relationship of bone mineral density and knee osteoarthritis (KL grade),” Clinics in Orthopedic Surgery, vol. 13, no. 1, pp. 60–66, 2021.

[4] P. S. Q. Yeoh et al., “Emergence of deep learning in knee osteoarthritis diagnosis,” Computational Intelligence and Neuroscience, 2021, Art. ID 4931437.

[5] R. Kijowski et al., "Deep learning applications in osteoarthritis imaging," Skeletal Radiology, vol. 52, no. 11, pp. 2225–2238, 2023.

[6] S. Kondal et al., "Automatic grading of knee osteoarthritis using CNN," arXiv preprint, 2020.

[7] M. A. Bany Muhammad, "Interpretable and optimized models for osteoarthritis assessment," 2020.

[8] H. Zhao et al., "The value of deep learning-based X-ray techniques in detecting KL grades: systematic review and meta-analysis," European Radiology, vol. 35, no. 1, pp. 327–340, 2025.

[9] K. Hernandez Abasolo, Detection of knee osteoarthritis severity using fusion of machine and deep learning models, Doctoral dissertation, 2021.

[10] Z. Jamaluddin and V. Amsaveni, "Enhancing osteoarthritis prediction with a hybrid deep learning model on knee X-rays," International Journal of Intelligent Engineering and Systems, vol. 18, no. 10, 2025.

[11] H. Long, "Hybrid design of CNN and vision transformer: A review," in Proc. Int. Conf. on Computer and Information Science & AI, 2024, pp. 121–127.

[12] Shashwat Work, "Knee Osteoarthritis Dataset with Severity," Kaggle, 2023. [Online]. Available:https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity