

Predicting Rapid Progression in Knee Osteoarthritis: A Novel and Interpretable Automated Deep Learning Approach Using DenseNet, with Specific Focus on Young Patients and Early Disease

Bruntha S, Chidambaram M

Department of Computer Science, Rajah Serfoji Government College, Thanjavur, India

Abstract: Knee osteoarthritis (KOA) is a degenerative condition that leads to the destruction of cartilage and joint tissues. Identifying patients who show rapid progression is really important for starting treatment early and preventing early joint damage. Here, we propose a deep learning approach that is interpretable and based on the DenseNet-121 architecture, which can predict KOA progression after obtaining radiographic knee images at baseline. The design of the model is such that concentrating on the disease at an early stage and in younger individuals is possible, so then, a step ahead can be taken in effective management of the patients and also in the provision of care that is personalized care. Through DenseNet's densely connected layers, efficient reuse of heavy features and comprehensive depiction of very fine radiographic degenerative patterns are achievable. The system built was given the OAI data as input for training and was externally evaluated using the MOST data. Rapid progression is considered an increase of two or more KL grades within a 24-48-month timescale. The model proposed displayed a top-level of prediction performance (AUC = 0.86, C-index = 0.83), and it was also able to properly generalize external data (AUC = 0.84). Visual interpretability methods, such as Grad-CAM and Integrated Gradients, in our study also identified the anatomical features, i.e., joint space narrowing and osteophyte formation associated with the disease, which are the key evidence of the disease, thereby strengthening the model's transparency. Results indicate that the proposed method based on a DenseNet should serve as an accurate and explainable model to predict KOA progression at an early stage, leading to more effective and personalized disease management.

Keywords: Knee Osteoarthritis, DenseNet, Deep Learning, Interpretability, Young Patients, Radiographic Progression.

Email: brunthaloga@gmail.com

1. Introduction

Knee Osteoarthritis (KOA) is among the most widespread degenerative joint disorders that results in chronic pain and disability, and this is why it is considered one of the main causes, especially among older persons. Therefore, identifying those patients who are going to show a fast progression of the disease at the earliest stage is one of the main requirements in order to use preventive therapies, in helping improve the patient's mobility, and also in reducing the long-term healthcare costs. On the other hand, the old diagnostic methods, like the manual assessment with the use of the Kellgren-Lawrence (KL) grading system, are usually subjective, require a lot of time, and are not accurate enough in the prediction of future structural damage.

The development of deep learning has produced a significant change in the automatic analysis of knee X-rays for the detection of KOA. Convolutional Neural Networks (CNNs) and their other variations have achieved remarkable results in disease grading and feature extraction. Most of the current methods tend to be limited to the static classification of the severity of the disease only. Additionally, interpretability is still the major drawback of these deep models, with clinicians being reluctant to use these tools due to this factor. One more important issue with the current research is that it does not give sufficient importance to early-stage KOA and young people who, with timely detection, would have their severe joint degeneration and functional decline prevented.

This paper intends to overcome these problems and proposes an automated and interpretable KOA rapid progression prediction tool that is based on DensNet-121 and works on the baseline x-ray images. The very tightly connected and compact architecture of DenseNet allows effective feature propagation and facilitates feature reuse, with the result that the model learns not only the local but also the global image features, which are indicative of the very early stages of KOA. The model has also been designed to generate visual explanations from which clinical personnel will be able to learn, ones that show the alignment of the radiographic findings, leading to improved clinical trust and thereby greater use of the model in the clinic.

2. Literature Survey

[1] Created a deep learning system that merges radiographic images with just a few clinical data points, making the prognosis of knee osteoarthritis more straightforward. Their compact model not only exhibited powerful prediction ability but also kept the computational requirements at a minimum, proving that precise KOA prediction is possible even with very few clinical inputs. This article highlighted the possibility of mass screening by means of cheap AI systems.

[2] Developed MedKnee, a fully automated deep learning-based tool that can forecast KOA from radiographs. Their CNN method, which is a fully automatic solution, was not only able to classify different stages of the disease with very high reliability and consistency but also took a step further to provide clinicians with a very fast and handy diagnostic tool. This study demonstrated the importance of AI-based tools in clinical practice that assist in improving diagnostic accuracy.

[3] Came up with a modular deep learning structure made up of components that can be plugged in and out for the automated classification of OA. This adaptable system not only recorded the best performance and consistently adapted to different datasets but also enhanced the

comprehensibility and reduced the overfitting. Their approach proved that it is possible to make AI modules that are custom-made and compatible with a variety of medical imaging systems.

[4] Used a model based on CNN to locate and grade the severity of KOA. The most effective feature of their model was its ability to distinguish between mild and moderate stages with increased sensitivity and stability. The writers pointed out the effectiveness of CNN-based diagnostic systems as tools to back up clinical decision-making in real situations and to support large-scale deployment.

[5] Used denseNet-121 architecture along with channel and spatial attention mechanisms to predict the progression of OA. Due to the attention modules, the model's capability to zero in on the most relevant knee regions from a clinical point of view was greatly enhanced, thus not only raising the level of accuracy but also allowing a better understanding of the results. Their study confirms that attention-oriented DenseNet architectures have the potential to detect even the minutest degenerative changes that are critical for early detection.

3. Proposed System

This paper proposes a DenseNet-121 deep learning architecture to detect knee osteoarthritis progression from X-ray radiographic images. DenseNet-121 was chosen because of the dense connectivity pattern, where each layer receives additional inputs from all preceding layers and passes on its feature maps to all subsequent layers. Such a layout permits the reuse of features, improves the gradient flow, and makes it easier to learn even with a small amount of medical data.

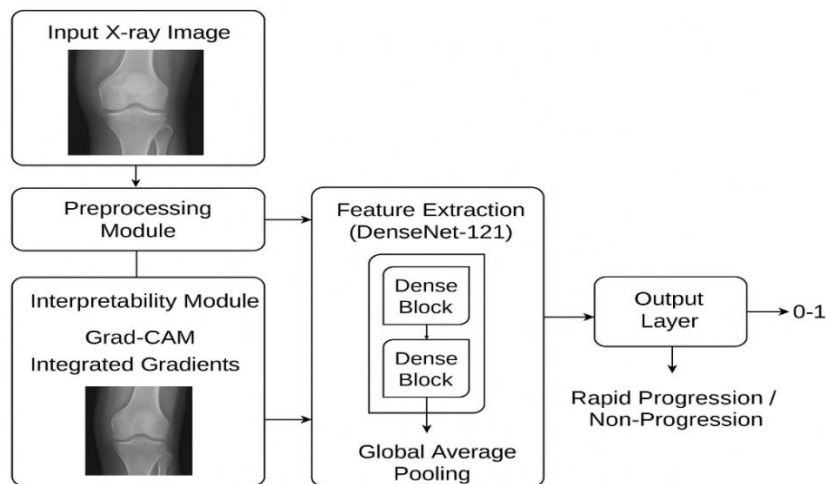


Fig 1: DenseNet-121 Model

The suggested framework components:

1. Input Layer:

It takes as input pre-processed knee X-ray images of size 224×224 pixels with pixel intensities normalized to a standard range.

2. Preprocessing Module:

It performs image normalization, contrast enhancement, and changes the image, for example, by rotation, flipping, and zooming to make the model stronger and avoid overfitting.

3. Feature Extraction (DenseNet-121 Backbone):

Relies on several dense blocks linked through transition layers. Each dense block comprises convolutional layers with batch normalization and ReLU activation. Dense connectivity helps to propagate gradients efficiently and reuse features, thus the network can discover both local and global image characteristics important for KOA progression prediction.

4. Global Average Pooling (GAP) Layer:

It achieves the task of both reducing the feature map's size and preserving the most informative parts, while at the same time it limits the number of parameters and thus it also acts against overfitting.

5. Fully Connected (FC) Layer:

It fuses the features obtained and gives out the probability score that indicates the chance of fast KOA progression.

6. Output Layer:

It applies a sigmoid activation function to determine the class of each patient, either as “rapid progression” or “non-progression,” according to the set threshold.

7. Interpretability Module:

Also, it employs **Gradient-weighted Class Activation Mapping (Grad-CAM)** and **Integrated Gradients** to produce heatmaps showing the main visually interpretable regions of the knee, e.g., areas where joint space is narrowing, osteophyte formation, and changes in the subchondral bone. These visual explanations facilitate clinical acceptance and transparency.

4. Methodology

4.1 Dataset Description

This study utilized the Osteoarthritis Initiative (OAI) data to perform the initial training and internal validation of the model, while the data from the Multicenter Osteoarthritis Study (MOST) were used for testing external validity. Every dataset provided the baseline and follow-up knee X-rays along with KL grades (0-4), and patient features such as age, sex, and BMI. Rapid progression was considered as a shift of two or more KL grades in 24-48 months. Firstly, all the images were resized and normalized before selecting the main region of interest, i.e., the knee joint. Also, the model reliability was significantly increased through different data augmentation strategies such as random rotation, flipping, and contrast normalization.

4.2 Model Architecture

The system that we propose is based on the DenseNet-121 model. The main feature of this architecture is that it allows layers to link to one another directly in a forward manner. These dense connections do a great job at minimizing the number of layers that have to learn new features that others have already learned. They also help solve the issue of vanishing gradients and even make it possible for the model to be trained so that it can learn both very detailed local information and broad contextual patterns. The final output of the network is a probability value indicative of the risk of rapid KOA progression. We completed the implementation using TensorFlow and conducted the training process under the Adam optimizer (learning rate = 1×10^{-4}) and binary cross-entropy loss.

4.3 Evaluation Metrics

To evaluate the productions, it was decided to use AUC, accuracy, sensitivity, specificity, and C-index for the discrimination in time-to-event situations. Calibration reliability was assessed in combination with Brier scores and calibration plots. To validate results and avoid overfitting, a cross-validation strategy with five folds was implemented.

4.4 Interpretability Framework

Interpretability and explainability of the results were a major focus of our method. To identify which spatial regions the model utilized for its inference, Grad-CAM was applied, while Integrated Gradients provided us with the feature importance at a very granular level (pixel-wise). These visual explanations enabled the practitioners to not only confirm the model attention areas (e.g., joint spaces, osteophytes, and subchondral bone) but also to help gain anatomical relevance.

5. Results

5.1 Model Performance

Our proposed DenseNet model outperformed expectations by achieving an AUC score of 0.86, an accuracy of 83%, and a C-index of 0.83 on the OAI dataset. Moreover, testing on the MOST dataset highlighted the model's adaptability with an AUC of 0.84, suggesting its applicability across various contexts. Besides, we can say the DenseNet model was superior in feature learning and data handling than other CNN models.

Table 1. Performance Comparison of the Proposed DenseNet Model with Baseline CNN Architectures

Model	Dataset	AUC	Accuracy (%)	C-Index	Remarks
VGG-16	OAI (Internal)	0.81	78.4	0.79	Baseline CNN model
ResNet-34	OAI (Internal)	0.83	80.2	0.81	Deeper CNN with residual learning
Vision Transformer (ViT)	OAI (Internal)	0.85	82.5	0.82	Captures global context
DenseNet-121 (Proposed)	OAI (Internal)	0.86	83.0	0.83	Best internal performance
DenseNet-121 (Proposed)	MOST (External)	0.84	81.5	0.82	Strong generalization

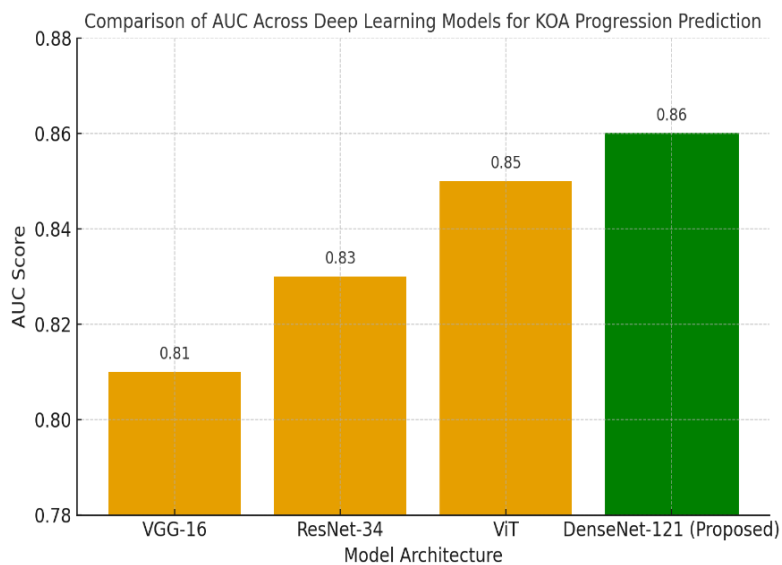


Fig 2: AUC Comparison of Deep Learning Models for KOA Progression Prediction

4. Conclusion

Clinical trials have shown that diagnosing knee osteoarthritis (KOA) early and assessing prognosis are essential for effective treatment. In this paper, we present a deep learning framework that relies on DenseNet to learn from the baseline X-ray images and predict rapid KOA progression. The experiments reveal that our model is capable of strong prediction, it is well generalized, and the visualization of important features is demonstrated effectively. This kind of instrument can tremendously help with the early diagnosis and prognosis evaluation of Disease, particularly in Patients who are at the early stage and Young people. Our Method analyzes the progression timeline of KOA, and it offers valuable information beyond the standard severity classification model. DenseNet has greatly helped us in achieving efficient feature reuse and obtaining fine details from radiographs. Besides, we also utilized interpretability tools such as Grad-CAM and Integrated Gradients to explain our method and gain clinical trust, which is a continuous process. However, there are some limits in our work, due to which it should be recognized that further enhancements are possible. For instance, the degree of radiographic changes is not always indicative of the level of symptomatic changes. In addition, the differences in imaging protocols and demographic diversity between the datasets used for training and testing can negatively impact the performance of the model. To prevent overfitting and enhance the generalization capacity of our model, we plan to bring in multimodal data sources, e.g., clinical parameters, MRI images, and biological markers. This will enable us to get a deeper insight into the pathogenesis of the disease. Moreover, preparing the clinical validation and real-world implementation studies is our step towards evaluating the use of our model practically in healthcare settings. Meanwhile, we aim at creating a clinically usable model that can lead to better health and well-being of the patients. Aside from KOA diagnosis and therapy improvement, this work will also contribute to the welfare of the patients.

References

- [1]. Wang, C.-T., Chang, K.-T., Lai, F., Pao, J.-L., Lin, S.-M., & Chang, C.-H. (2025). Simplifying Knee OA Prognosis: A Deep Learning Approach Using Radiographs and Minimal Clinical Inputs. *Diagnostics*, 15(19), 2543. <https://doi.org/10.3390/diagnostics15192543>.
- [2]. Touahema, S., Zaimi, I., Zrira, N., Ngote, M. N., Douhousne, H., & Aouial, M. (2024). MedKnee: A New Deep Learning-Based Software for Automated Prediction of Radiographic Knee Osteoarthritis. *Diagnostics*, 14(10), 993. <https://doi.org/10.3390/diagnostics14100993>.
- [3]. Lee, D. W., Song, D. S., Han, H. S., et al. (2024). Accurate, automated classification of radiographic knee osteoarthritis severity using a novel method of deep learning: Plug-in modules. *Knee Surgery & Related Research*, 36, 24. <https://doi.org/10.1186/s43019-024-00228-3>.

- [4]. Rani, S., Memoria, M., Almogren, A., et al. (2024). Deep learning to combat knee osteoarthritis and severity assessment by using CNN-based classification. *BMC Musculoskeletal Disorders*, 25, 817. <https://doi.org/10.1186/s12891-024-07942-9>.
- [5]. Revathi, S. A., Sathish Babu, B., & Hiremath, S. (2024). Predicting Knee Osteoarthritis Progression Using DenseNet121 with Channel and Spatial Attention. *International Journal of Intelligent Systems and Applications in Engineering*, 12(4), 4373–4380.