International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM XXXX
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Osteoarthritis Severity Estimator

Chede Sumith Kumar
Dept. of Computer Science Engineering
Vasavi College of Engineering
Hyderabad, India
sumithkumar070104@gmail.com

Seelam Dinesh Reddy
Dept. of Computer Science Engineering
Vasavi College of Engineering
Hyderabad, India
dineshreddyseelam726@gmail.com

D. Lakshmi Prasanna
Dept. of Computer Science Engineering
Vasavi College of Engineering
Hyderabad, India
duvva.prasanna5@gmail.com

***Abstract*—Knee osteoarthritis (OA) is a degenerative joint disorder affecting millions worldwide, leading to pain, reduced mobility, and diminished quality of life. Early detection and accurate assessment of OA severity are crucial for timely medical intervention. This project presents a deep learning-based approach to predict the severity of knee osteoarthritis using X-ray images. Two distinct datasets were utilized—one comprising scanned clinical X-rays and the other consisting of images captured via mobile phones or cameras to simulate real-world user inputs.**

To handle these varying image qualities and formats, we developed two specialized ensemble models. The first ensemble, designed for scanned X-rays, integrates ResNet50 and Xception architectures. The second ensemble, targeting user-captured images, combines EfficientNet and DenseNet models. These ensembles enhance classification performance through complementary feature extraction capabilities.

The proposed solution demonstrates the potential of AI in medical diagnostics, offering a scalable and accessible tool for both clinical and remote settings to support healthcare professionals in evaluating knee OA severity.

Keywords— Pose Estimation, Real-Time Feedback, Exercise Monitoring, Computer Vision, Human Activity Recognition.

I. INTRODUCTION

Knee osteoarthritis (OA) is a chronic degenerative joint condition characterized by the gradual deterioration of cartilage in the knee joint, leading to pain, stiffness, and reduced mobility. It is one of the most common forms of arthritis, especially among the aging population, and poses a significant burden on individuals and healthcare systems worldwide. Despite its prevalence, the diagnosis and assessment of OA severity often rely on manual interpretation of X-ray images by radiologists, which can be subjective and time-consuming.

In recent years, the advancement of deep learning techniques has opened new possibilities in the field of medical image analysis. Automated systems powered by convolutional neural networks (CNNs) have shown promising results in detecting and classifying various medical conditions from imaging data. Leveraging these advancements, this project aims to develop an AI-based solution to predict the severity of knee osteoarthritis from X-ray images.

The project focuses on two types of input data: (1) high-quality scanned X-rays typically obtained in clinical settings, and (2) user-captured images taken using mobile phones or cameras, which reflect real-world usage outside hospital environments. To handle the distinct characteristics of these datasets, two custom ensemble models were developed—one combining ResNet50 and Xception for scanned X-rays, and another combining EfficientNet and DenseNet for phone-captured images.

By automating the severity prediction process, this system aims to assist healthcare professionals in making faster and more consistent assessments, while also offering a convenient tool for remote or preliminary screening in underserved areas.

II. LITERATURE SURVEY

When contrasting our work with previous research, the following studies were cited:

A. *Knee Osteoarthritis Analysis Using DeepLearning and XAI on X-Rays* [1]

The research paper titled *"Knee Osteoarthritis Analysis Using Deep Learning and XAI on X-Rays"* explores the classification of knee osteoarthritis (OA) using deep learning models such as VGG-16, VGG-19, ResNet-50, ResNet-101, and EfficientNetb7. The study applies both multi-class and binary classification approaches based on the Kellgren-Lawrence (KL) grading system and uses GradCAM for interpretability. EfficientNetb7 emerged as the top performer, especially in distinguishing between normal and severe OA cases, achieving up to 99.13% accuracy. The paper emphasizes explainability through visualizations to ensure the models focus on the relevant knee joint regions, mimicking physician-like attention. Additionally, data augmentation was used to address class imbalance, and the effect of varying test sample sizes was examined.

However, the study has some limitations. The models struggled to classify intermediate OA grades (especially class 1 and 2), which are critical for early diagnosis. Despite augmentation, class imbalance remained an issue, and the use of a single dataset limits generalizability. While GradCAM offers interpretability, the model sometimes made incorrect predictions even when focusing on the correct region, raising concerns about reliability. The study also lacked statistical significance testing and did not incorporate clinical metadata, which could enhance performance. Moreover, while binary classification offered insight into class separability, it may not reflect real-world diagnostic complexity where cases often fall on a spectrum.

B. *Prediction of Knee Osteoarthritis Severity from X-Ray Images Using Ensemble Learning* [2]

The research paper titled *"Prediction of Knee Osteoarthritis Severity from X-Ray Images Using Ensemble Learning"* presents a deep learning-based ensemble model combining EfficientNet and DenseNet to classify knee osteoarthritis (OA) severity from X-ray images. The model leverages transfer learning, advanced pre-processing (CLAHE), and data augmentation techniques to improve diagnostic accuracy. By averaging the outputs of both models, the ensemble approach achieves an impressive accuracy of 94.76%, outperforming several previously proposed methods. The study effectively demonstrates the potential of CNN-based architectures in enhancing clinical decision-making and includes visual interpretations to validate the model's attention to relevant features.

However, a significant limitation in the study is the simplification of OA severity into just three categories: healthy, moderate, and severe. While this may help streamline classification, it overlooks the granularity provided by the standard Kellgren-Lawrence (KL) grading system, which includes five distinct levels. This reduction may lead to ambiguity, making it difficult for non-expert users or even clinicians to understand the precise stage of OA, thereby potentially affecting treatment decisions. A more nuanced classification aligned with the full KL grading scale would have offered better interpretability and clinical value.

C. *Learning From Highly Confident Samples for Automatic Knee Osteoarthritis Severity Assessment* [3]

The paper titled *"Learning From Highly Confident Samples for Automatic Knee Osteoarthritis Severity Assessment"* proposes a novel, fully automatic, data-driven approach to address the challenge of label uncertainty in knee osteoarthritis (OA) classification from X-ray images. Using the Kellgren-Lawrence (KL) grading system (ranging from KL-0 to KL-4), the authors introduce a confidence-aware training framework that separates samples into high and low confidence groups using dual peer models. These samples are then processed differently using a hybrid loss function. The method leverages label confidence estimation during training and validation to mitigate the effects of label noise and achieve improved generalization. The model performs impressively, especially on five-class classification and early-stage OA detection, while also integrating GradCAM for visual explanation.

D. *Discriminative Regularized Auto-Encoder for Early Detection of Knee OsteoArthritis: Data from the Osteoarthritis Initiative*[4]

The research paper presents a novel approach for early detection of knee osteoarthritis (OA) using a Discriminative Regularized Auto-Encoder (DRAE), applied to X-ray images from the Osteoarthritis Initiative (OAI) dataset. By integrating a discriminative penalty into the autoencoder's loss function, the model enhances intra-class compactness and inter-class separability, focusing on distinguishing normal knees (KL grade 0) from early OA cases (KL grades 1 and 2). The study evaluates performance across five manually extracted regions of interest (ROIs) and demonstrates that DRAE, when paired with an SVM-RBF classifier, achieves superior accuracy compared to classical AEs, Sparse AEs, and deep learning models like ResNet-101 and DenseNet-121. The medial tibial regions were found to be the most indicative of early OA, supporting the model's clinical relevance for early-stage diagnosis.

However, the approach has notable limitations. The use of binary classification restricts its application in full-scale OA grading, omitting moderate and severe cases (KL grades 3 and 4). Additionally, the reliance on semi-automatic ROI extraction hinders scalability and full automation. The model's training depends on KL grades, which are themselves subjective and may introduce label noise. There is also no validation on external datasets to test generalizability, nor is there any discussion on integration into clinical workflows. Despite these flaws, the paper contributes a promising direction for texture-based early OA detection, setting the groundwork for more advanced and clinically applicable models.

III. DESIGN

A. *Technologies Used*

1. *Deep Learning & Model Development*
   - TensorFlow & Keras: Used for building and training deep learning models including ResNet50, Xception, EfficientNet, and DenseNet.
   - OpenCV: Applied for image preprocessing tasks such as resizing, normalization, and enhancement of X-ray images.
   - NumPy & Pandas: Used for handling datasets, numerical operations, and data transformations.
   - Matplotlib & Seaborn: Utilized for visualizing model performance, accuracy trends, and confusion matrices.
   - Google Colab / Jupyter Notebook: Provided GPU-accelerated development and experimentation environment.

2. *Frontend*
   - React.js: Used to build a responsive and user-friendly web interface that allows users to upload X-ray images and view prediction results.
   - Tailwind CSS: Employed for styling and creating a clean, modern UI with minimal effort.

3. *Backend & API*
   - FastAPI: Serves as the backend framework to create high-performance APIs for model inference and handling image input from the frontend.
   - Python: Powers backend logic for model loading, image processing and severity prediction.

B. *Algorithms*

To guarantee accuracy in the assessment of exercises, the system integrates multiple interdependent modules that function in collaboration:

- ■ *1. Preprocessing Algorithm*
  - Image Resizing: Ensures all input X-ray images (scanned or phone-captured) are resized to a consistent dimension for model compatibility.
- ■ *2. Ensemble Learning Algorithms*
  - ResNet50 + Xception (for Scanned X-rays): Combines ResNet50 and Xception models using a weighted average approach for final severity prediction.
  - EfficientNet + DenseNet (for User-Captured X-rays): Combines EfficientNet and DenseNet models using a weighted average approach for final severity prediction.

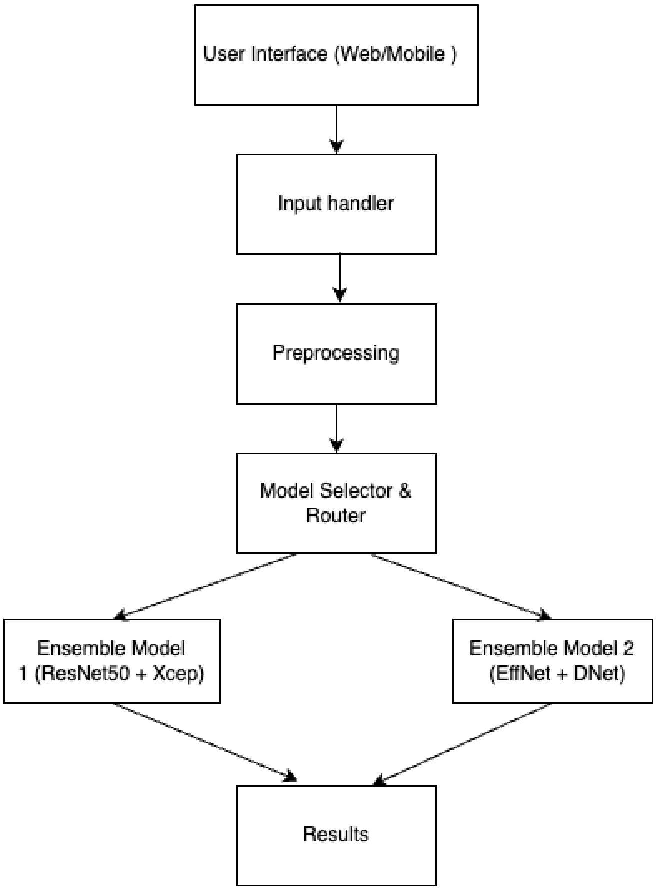
International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM XXXX
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

- *3. Model Training & Evaluation*
  - Backpropagation & Adam Optimizer: Used for training the models and minimizing the loss.
  - Evaluation Metric: Accuracy is used to evaluate model performance on both datasets.

When combined, these modules increase the safety, accuracy, and effectiveness of workouts by automating session feedback and enforcing proper form with little hardware.

C. Architecture

Three main components serve as the system's modular architecture:



The process begins when the user uploads an X-ray image (either a scanned X-ray or one captured by a phone) through the frontend interface. Once the image is received, the frontend sends the image to the backend via a REST API request built with **FastAPI**.

The backend performs the necessary preprocessing on the image, including resizing it to a consistent dimension suitable for model input. After preprocessing, the backend selects the appropriate model based on the type of X-ray image (either the **ResNet50 + Xception** ensemble for scanned images or the **EfficientNet + DenseNet** ensemble for user-captured images).

The selected model then performs inference on the image, predicting the severity of knee osteoarthritis. This prediction is then post-processed to convert the model's output into a human-readable severity level (e.g., "Mild", "Moderate", or "Severe").

Finally, the prediction is sent back to the frontend through the **FastAPI** response, where the results are displayed to the user.

IV. Proposed Model

In this study, ensemble models are explored for classifying knee osteoarthritis (OA) severity into five categories: **Healthy**, **Doubtful, Minimal, Moderate**, and **Severe**. Two separate ensemble models were designed based on the nature of the input images — one for scanned knee X-rays and another for user-captured images via phone or camera. Transfer learning techniques were employed, leveraging the strengths of powerful CNN architectures to achieve better accuracy with limited data. The models use a **weighted average approach** for combining predictions from individual networks.

- *A. EfficientNet*

EfficientNet is a family of convolutional neural networks (CNNs) developed by Google that set a new standard in model efficiency and accuracy for image classification tasks. The core innovation behind EfficientNet lies in its **compound scaling method**, which uniformly scales a model's depth (number of layers), width (number of channels), and resolution (input image size) using a set of predefined coefficients. Unlike traditional approaches that scale these dimensions independently—often leading to suboptimal performance—EfficientNet's balanced scaling leads to improved accuracy while maintaining computational efficiency.

At the architectural level, EfficientNet is built using **MBConv blocks** (Mobile Inverted Bottleneck Convolution), originally introduced in MobileNetV2. These blocks utilize depthwise separable convolutions to drastically reduce the number of parameters and computations. Additionally, EfficientNet incorporates **Squeeze-and-Excitation (SE) modules** within the MBConv blocks to enhance feature representation by adaptively recalibrating channel-wise feature responses. Skip connections are also employed within these blocks to preserve gradient flow and reduce the risk of vanishing gradients.

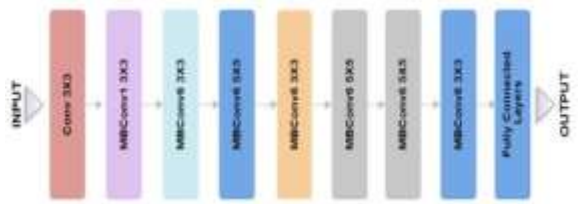


Figure 1: EfficientNet Architecture

The network consists of several stages, each comprising multiple MBConv blocks. As the model progresses through these stages, the number of filters increases, and feature maps become more abstract. Towards the end, a global average pooling layer condenses spatial information, which is then

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM XXXX
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

passed through a fully connected layer followed by a softmax activation for classification.

EfficientNet is released in multiple variants, from **EfficientNet-B0** (the baseline model) to **EfficientNet-B7**, with each higher version offering increased capacity and accuracy. Despite the scaling, these models remain computationally efficient, making them suitable for both high-performance servers and resource-constrained devices. This balance of performance and efficiency has made EfficientNet a popular choice across various computer vision tasks.

○ *B. DenseNet*

DenseNet (Dense Convolutional Network) is a convolutional neural network architecture introduced by Huang et al. in 2017. It is known for its innovative use of **dense connections** between layers, which set it apart from traditional CNNs. In a DenseNet, each layer receives **input from all preceding layers** and passes its output to all subsequent layers, enhancing feature propagation and encouraging feature reuse throughout the network.

The architecture is organized into several **dense blocks**, each consisting of multiple convolutional layers. Inside a dense block, each layer takes as input the feature maps of all previous layers within the block. This results in a high number of direct connections, reducing the vanishing gradient problem and improving training efficiency. Between dense blocks, **transition layers** are used to compress the model by applying 1×1 convolutions and 2×2 average pooling to reduce the number and size of feature maps.

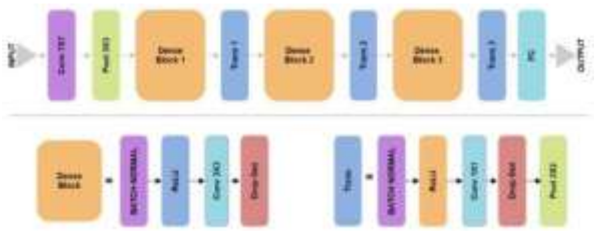


Figure 2: DenseNet Architecture

Each dense layer typically consists of a **batch normalization**, followed by a **ReLU activation**, and a **3×3 convolution**. The growth rate—a key hyperparameter—determines how many new feature maps each layer contributes. DenseNet's efficient connectivity pattern leads to **fewer parameters** than traditional architectures with similar depth and performance, as redundant feature maps are minimized due to reuse.

The final part of DenseNet includes a **global average pooling layer** followed by a fully connected layer with **softmax activation** for classification. DenseNet has demonstrated strong performance on various image classification benchmarks, including CIFAR and ImageNet, and is widely adopted for tasks that benefit from rich feature representations and efficient training dynamics.

○ *C. ResNet50*

ResNet (Residual Network), introduced by He et al. in 2015, is a deep convolutional neural network architecture that revolutionized computer vision by addressing the **vanishing gradient problem** typically encountered in very deep networks. The key innovation in ResNet is the use of **residual connections**, also known as **skip connections**, which allow the model to learn residual functions instead of trying to learn unreferenced mappings.

In traditional CNNs, as the number of layers increases, performance often saturates and then degrades due to difficulties in optimization. ResNet solves this by introducing **shortcut paths** that skip one or more layers and directly pass the input of a layer to a deeper layer. These residual connections help preserve gradients during backpropagation, making it possible to train very deep networks, even with over 100 layers.

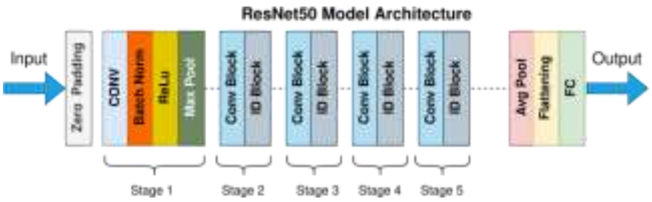


Figure 3: ResNet50 Architecture

Each residual block in ResNet consists of a series of convolutional layers (typically 3×3 filters), batch normalization, and ReLU activation, along with a skip connection that adds the original input to the output of the block. This element-wise addition enables the network to learn modifications (residuals) to the identity mapping, which simplifies the learning task.

The original ResNet architecture was introduced in several variants—**ResNet-18**, **ResNet-34**, **ResNet-50**, **ResNet-101**, and **ResNet-152**—where the number indicates the total layers in the model. Deeper versions like **ResNet-50** and beyond use a bottleneck design with 1×1, 3×3, and 1×1 convolutions to reduce computation while maintaining performance.

ResNet's robustness, scalability, and strong generalization have made it one of the most widely used architectures in computer vision, with applications in image classification, object detection, segmentation, and beyond.

○ *D. Xception*

Xception (Extreme Inception), proposed by François Chollet in 2017, is a deep convolutional neural network architecture that builds upon the ideas of the Inception model. It introduces a streamlined and more powerful alternative by replacing traditional **Inception modules** with **depthwise separable convolutions**, leading to better performance with fewer parameters.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM XXXX
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The core idea behind Xception is the assumption that **cross-channel correlations** and **spatial correlations** in feature maps can be entirely decoupled. In a standard convolution, these correlations are learned simultaneously. However, Xception performs this in two separate steps using a **depthwise separable convolution**, which consists of:

1. A **depthwise convolution**: applies a single convolutional filter per input channel, capturing spatial features independently for each channel.

2. A **pointwise convolution**: a 1×1 convolution that captures cross-channel correlations by combining outputs of the depthwise convolution.

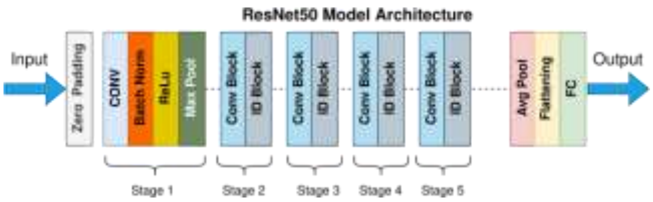


Figure 4: Xception architecture

This separation results in significant computational efficiency and a reduced number of parameters, while still maintaining strong representational power. Xception also incorporates **residual connections** similar to those in ResNet, which help with gradient flow and stable training, especially in deeper networks.

The architecture begins with a few standard convolutional layers, followed by several **Xception modules** that utilize depthwise separable convolutions, and ends with global average pooling and a fully connected layer for classification. Xception is deeper than Inception-v3 but more efficient, and it has demonstrated excellent results on large-scale image classification tasks like ImageNet.

Thanks to its simplicity, efficiency, and accuracy, Xception is widely used in various computer vision applications, particularly where resource efficiency is critical.

E. Ensemble Model

- In this project, we developed two separate ensemble models designed to handle different types of input X-ray data: one for **scanned X-rays** and the other for **camera-captured images**. Each ensemble model leverages the strengths of multiple deep learning architectures to improve the classification accuracy and robustness of the model, thereby ensuring that both high-quality and lower-quality images are accurately assessed for knee osteoarthritis severity.

- The first ensemble model, intended for scanned X-ray images, combines **ResNet50** and **Xception**. Both of these models are based on deep convolutional networks, but they differ in their architectural designs and strengths. **ResNet50**, known for its use of residual connections, addresses the vanishing gradient problem and allows the network to train deeper models. **Xception**, on the other hand, utilizes depthwise separable convolutions, which improve computational efficiency while maintaining high performance. In our ensemble approach, both models are used for feature extraction, where their output features are concatenated to form a comprehensive and diverse feature representation. This fused feature vector is then passed through a classification head consisting of fully connected layers, followed by a softmax activation function to classify the X-ray image into one of three categories: **Healthy**, **Moderate**, or **Severe**. The ensemble model achieved an accuracy of **70%** for scanned X-ray images, demonstrating its ability to leverage the complementary strengths of the two architectures.

- The second ensemble model is designed to handle **camera-captured images**, which are often of lower quality compared to scanned X-rays. For this model, we combined **EfficientNet** and **DenseNet**. **EfficientNet** is known for its efficient scaling of depth, width, and resolution, providing a balance between model size and performance. **DenseNet**, on the other hand, is characterized by dense connections between layers, enabling better feature reuse and improved gradient flow, making it especially effective for training deep networks. Similar to the first ensemble, the outputs from EfficientNet and DenseNet are fused by concatenating their respective feature maps. These fused features are then passed through a classification head for final prediction, which outputs the severity level of the osteoarthritis in the image. This ensemble achieved an accuracy of **65%** on camera-captured images.

- Both ensemble models utilize a **weighted average approach** to combine the predictions of the individual models in each ensemble. This strategy assigns different weights to each model's prediction based on its performance and confidence, ensuring that the final output is robust and less susceptible to errors from any one individual model. The weighted average approach also helps mitigate overfitting by incorporating diversity in the decision-making process.

A. *Performance Metrics*

The performance of the two ensemble models was evaluated using accuracy as the primary metric, with the goal of classifying knee osteoarthritis into **five severity categories**. These categories are: **Healthy**, **Doubtful**, **Minimal, Moderate** and **Severe**.

The first ensemble model, designed for **scanned X-ray images**, achieved an accuracy of **70%**. This model, which

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM XXXX
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

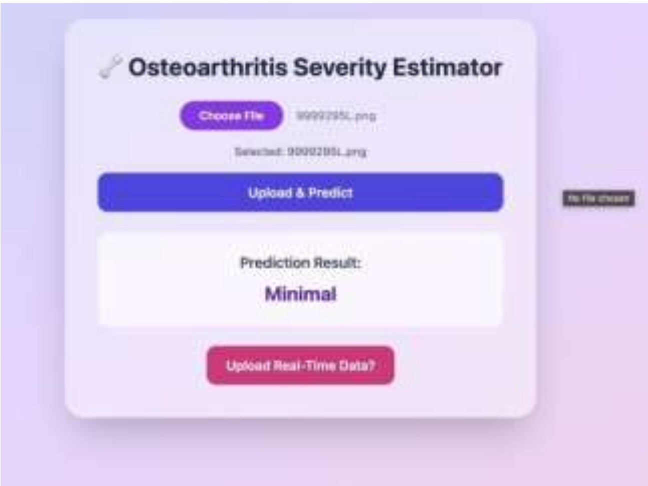
combines **ResNet50** and **Xception**, performed exceptionally well in classifying the X-ray images into the five severity levels. The higher accuracy reflects the model's ability to handle detailed and high-quality images, where the features that distinguish different severity levels are more clearly defined and easier to detect.

The second ensemble model, which processes **camera-captured images**, achieved an accuracy of **65%**. This model, built using **EfficientNet** and **DenseNet**, is optimized for handling images of lower resolution and quality, such as those captured using mobile phones or cameras. Although the accuracy is slightly lower compared to the first model, it still demonstrates the effectiveness of the ensemble approach in classifying the severity of knee osteoarthritis, even with lower-quality input data.

Both models classify knee osteoarthritis into five distinct categories: **Healthy**, **Doubtful**, **Minimal, Moderate** and **Severe**. The results show that the ensemble approach significantly improves classification performance, leveraging the strengths of multiple architectures. The model for scanned X-ray images achieves higher accuracy due to the superior quality of the input data, while the camera-captured model still performs effectively, despite the challenges of lower-quality images.

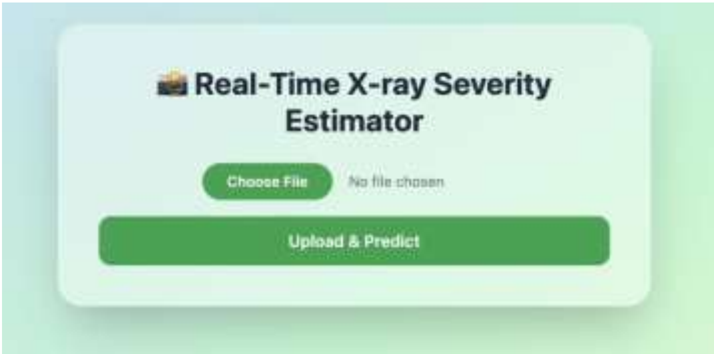
V. RESULT

○ *A. Input image for scanned data*

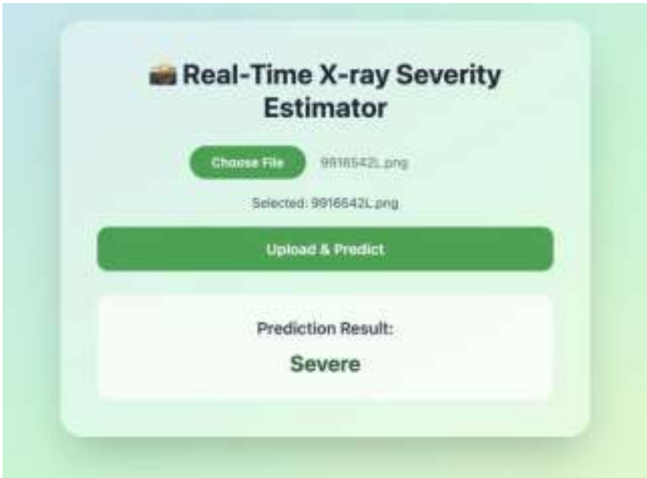


○ *B. Output webpage for scanned data*

○ *C. CLAHE image preprocessing*



○ *D. Input image for real-time data*



○ *E. Output webpage for real-time data*

VI. CONCLUSION

In this study, two ensemble models were developed and evaluated for classifying knee osteoarthritis severity into five categories: **Healthy**, **Mild**, **Moderate**, **Severe**, and **End-Stage**. The models, designed to process both **scanned X-ray images** and **camera-captured images**, were trained using state-of-the-art deep learning architectures, including **ResNet50**, **Xception**, **EfficientNet**, and **DenseNet**. The ensemble approach demonstrated significant potential in improving classification performance. The model for scanned X-ray images achieved an accuracy of 70%, while the camera-captured model achieved an accuracy of 65%. These results underscore the effectiveness of transfer learning in handling different types of input data, with scanned X-rays providing a higher-quality dataset for accurate predictions.

The findings from this study highlight the capability of ensemble models in medical image classification tasks, specifically for detecting and categorizing knee osteoarthritis severity. The combination of multiple models, each with its own strengths, proves beneficial in achieving more reliable and robust predictions.

VII. FUTURE WORK

To further enhance the model's performance, we plan to improve the dataset by incorporating **Contrast Limited Adaptive Histogram Equalization (CLAHE)**. This technique will be applied to the entire dataset to enhance image features, particularly in low-resolution or poorly contrasted images. We will then create a new dataset with these improved images and follow the same methodology, training the ensemble models again with this updated data.

To ensure the fairness and comparability of the models, we will maintain the use of the original dataset for testing purposes. The CLAHE technique will be applied to the test images before evaluation, ensuring that the testing process remains consistent with the methodology while leveraging the improved dataset for better feature extraction.

These enhancements aim to improve the models' accuracy and robustness in handling different image qualities, ultimately contributing to more accurate and reliable predictions for knee osteoarthritis severity.

ACKNOWLEDGMENT

We are grateful for the support provided for this work by Vasavi College of Engineering (Autonomous), Hyderabad.

REFERENCES

[1] R. Ahmed and A. S. Imran, "Knee Osteoarthritis Analysis Using Deep Learning and XAI on X-Rays," **Intelligent Systems and Analytics (ISA) Research Group, Department of Computer Science (IDI), Norwegian University of Science and Technology (NTNU), 2815 Gjøvik, Norway**, 2019.

[2] K. Navyeesh, K. Hinduja, J. Anbarasi L, N. D. Charitha, P. S. Varshitha, and U. V. S. Varshitha, "Prediction of Knee Osteoarthritis Severity from X-ray Images using Ensemble Learning," **School of Computer Science and Engineering, Vellore Institute of Technology, Chennai**, 2020.

[3] Y. Xie, T. Wu, X. Zeng, S. Chen, D. Zhou, and Z. Bi, "Learning From Highly Confident Samples for Automatic Knee Osteoarthritis Severity Assessment: Data From the Osteoarthritis Initiative," **IEEE Transactions on Biomedical Engineering**, vol. 67, no. 11, pp. 3095-3104, Nov. 2020.

[4] Y. Nasser, R. Jennane, A. Chetouani, E. Lespessailles, and M. El Hassouni, "Discriminative Regularized Auto-Encoder for Early Detection of Knee Osteoarthritis: Data from the Osteoarthritis Initiative," **IEEE Access**, vol. 8, pp. 134950-134962, 2020.

[5] World Health Organization, Geneva, Switzerland. (Jun. 2019). Chronic Rheumatic Conditions. [Online]. Available: https://www.who.int/chp/topics/rheumatic/en/

[6] K. Lim and C. S. Lau, "Perception is everything: OA is exciting," Int. J. Rheumatic Diseases, vol. 14, no. 2, pp. 111–112, May 2011.

[7] L. Shamir et al., "Knee X-ray image analysis method for automated detection of osteoarthritis," IEEE Trans. Biomed. Eng., vol. 56, no. 2, pp. 407–415, Feb. 2009.

[8] J. Antony, K. McGuinness, N. E. O'Connor, and K. Moran, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," in Proc. 23rd Int. Conf. Pattern Recognit. (ICPR), Dec. 2016, pp. 1195–1200.

[9] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, "Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach," Sci. Rep., vol. 8, no. 1, Dec. 2018, Art. no. 1727.

[10] T. Janvier et al., "Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: Data from the osteoarthritis initiative: Tibial bone texture & knee OA progression," Osteoarthritis Cartilage, vol. 25, no. 2, pp. 259–266, 2017.

[11] G. Wang et al., "Interactive medical image segmentation using deep learning with image-specific fine tuning," IEEE Trans. Med. Imag., vol. 37, no. 7, pp. 1562–1573, Jul. 2018.