Knee Osteoarthritis Detection and Classification Using X-Rays

Mrs. Nischitha K[1], Prakruthi V S[2], Aishwarya S[3], Harshitha MC[4], Nithya shree KS[5]

[1]Assistant Professor, Department of ECE, PES College Of Engineering Mandya, Karnataka, India
[2345]Students of ECE, PES College Of Engineering Mandya, Karnataka, India

---***---

Abstract – *An AI-based system designed to help in the early detection of knee osteoarthritis using knee X-ray images. It classifies images into three categories: normal, mild osteoarthritis, and severe osteoarthritis using a Kaggle dataset. A lightweight MobileNetV2 model is used to ensure good accuracy with fast and efficient performance, and it is trained in Google Colab using GPU support. After training, the model is connected to a simple Gradio web interface where users can upload X-ray images and instantly get predictions along with confidence scores. This makes the results easy to understand and more reliable for users. The main goal is to support doctors in early screening and reduce the time needed for manual analysis. It also helps in making the diagnosis process faster, simpler, and more accessible. Overall, this system provides a practical and user-friendly tool for early knee osteoarthritis detection in real world use.*

***Key Words*: Knee Osteoarthritis, Artificial Intelligence, MobileNetV2, Deep Learning, Image Classification, Google Colab, Gradio, Medical Imaging**

1. INTRODUCTION

Knee osteoarthritis (KOA) is one of the most common joint problems seen in adults, especially in older age groups. It slowly damages the knee joint over time and leads to pain, swelling, and difficulty in everyday activities like walking or climbing stairs. The main challenge with this condition is that it develops gradually; so many people do not notice it in the early stages. That is why early detection plays a very important role, as it can help in managing the disease better and preventing it from becoming severe. In hospitals, doctors usually depend on knee X-ray images to identify and grade the severity of osteoarthritis. While this method is widely used, it still relies on visual in spection, which can sometimes take time and may vary from one expert to another.

Because of this, there has been a growing interest in using artificial intelligence to support medical diagnosis and make the process faster and more consistent. In recent years, many researchers have shown that deep learning can be very effective in analysing medical images. Several studies have successfully used convolutional neural networks to automatically detect and classify knee osteoarthritis from X-ray images with good accuracy [1][2][3]. Some improved models have also been designed to increase performance by focusing on better feature extraction and learning techniques [4][5]. Apart from this, research reviews highlight that machine learning methods are becoming an important support tool in healthcare, especially for diseases like osteoarthritis [6]. Other studies have also explored more advanced ideas such as predicting disease progression and identifying specific features from medical images [7][8]. Continuous improvements in model architecture and training strategies have further helped in making these systems more reliable and accurate [9][10]. Inspired by all these works, this project focuses on developing a simple, lightweight, and efficient deep learning-based system for classifying knee osteoarthritis using X-ray images. The main goal is to create a tool that is easy to use, fast in response, and helpful for early screening. This can assist doctors in making quicker decisions and support better patient care in real-world situations.

Furthermore, this system also plays an important role in supporting preventive healthcare. By identifying early signs of knee osteoarthritis, especially in the mild stage, patients can take timely steps such as lifestyle modifications, physiotherapy, and medical treatment to slow down disease progression. The use of deep learning ensures that subtle patterns in X-ray images, which might be difficult to detect with the human eye, are effectively captured and analyzed. This not only enhances the reliability of the diagnosis but also reduces dependency on highly specialized experts. As a result, the proposed system can be particularly useful in remote or underserved areas.

The rapid advancement of artificial intelligence and deep learning techniques has created new opportunities in the field of medical image analysis. In particular, convolutional neural networks (CNNs) have demonstrated superior performance in extracting complex patterns and hierarchical features from radiographic images. Unlike traditional machine learning methods that require handcrafted feature extraction, deep learning models automatically learn discriminative features directly from raw image data, thereby improving classification accuracy and reducing preprocessing complexity. This capability makes deep learning a highly effective approach for the automated detection and grading of knee osteoarthritis from X-ray images.

Additionally, the integration of automated diagnostic systems into healthcare environments has the potential to improve clinical efficiency and diagnostic reliability. A computer-aided diagnosis framework for knee osteoarthritis can assist orthopedic specialists by providing fast and objective assessments, minimizing inter-observer variability, and supporting early-stage disease detection. Such systems are particularly valuable in large-scale screening applications and healthcare centers with limited access to experienced radiologists. Therefore, the proposed work aims to develop an efficient and lightweight deep learning model capable of accurately classifying knee osteoarthritis severity, contributing toward intelligent and accessible healthcare solutions.

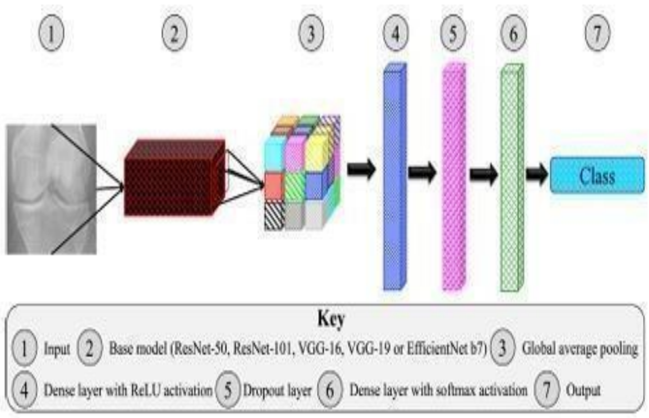

Fig-1: General architecture of models.

2. LITERATURE REVIEW

The application of deep learning for knee osteoarthritis (KOA) classification has gained significant attention in recent years, particularly with the advancement of convolutional neural networks (CNNs) for medical image analysis. Early work in this domain primarily focused on traditional radiographic grading methods such as the Kellgren–Lawrence (KL) system [11], which relies heavily on manual interpretation and is often subject to inter-observer variability.

With the emergence of deep learning, several researchers have explored automated classification approaches using X-ray images. Thomas et al. [1] proposed a deep neural network-based framework for automated grading of knee osteoarthritis severity, demonstrating that deep learning models can effectively learn radiographic patterns 3 without manual feature extraction. Similarly, Antony et al. [2] utilized CNN-based architectures to quantify KOA severity, achieving improved classification performance compared to conventional machine learning methods. Saarakkala [3] further highlighted the potential of deep learning in radiographic analysis, showing its capability to identify subtle structural changes in knee joints.

Subsequent studies have investigated the use of pre-trained architectures such as VGG16, ResNet, and Inception networks for KOA classification tasks [4], [5]. These transfer learning approaches improved accuracy, particularly in scenarios with limited medical datasets. However, most of these models remain computationally heavy and are not optimized for real-time or low-resource deployment environments.

Recent research has also explored lightweight architectures such as MobileNetV2 for efficient medical image classification [9]. These models provide a balance between ac curacy and computational efficiency, making them suitable for deployment in web based or mobile-assisted diagnostic systems. Additionally, techniques such as data aug mentation, class balancing, and optimization strategies have been widely adopted to improve generalization performance [6].

Despite these advancements, existing approaches remain fragmented in terms of deployment readiness and usability. Most studies focus primarily on classification accuracy without integrating user-friendly diagnostic interfaces or real-time prediction systems. Furthermore, limited attention has been given to end-to-end systems that combine model training, inference, and deployment in a unified framework.

Unlike existing works, the proposed system integrates a lightweight MobileNetV2 based classifier with a Gradio based interactive interface, enabling real-time prediction of KOA severity levels. This addresses the gap between high-performing research models and practical clinical usability, making the system more accessible for preliminary screening applications.

3. METHODOLOGY

The proposed system follows a structured pipeline consisting of four main stag-es: data collection, preprocessing, model training, and deployment.

3.1 Dataset Collection

The knee X-ray dataset is obtained from Kaggle and contains images categorized into three classes:

- Normal
- Osteopenia
- Osteoporosis

3.2 Data Preprocessing

The proposed system consists of three main phases: data preprocessing, model training, and deployment. In preprocessing, bone density images are resized to 224×224 pixels, normalized, and augmented using rotation and flipping to im-prove generalization. The dataset is labeled into three classes Normal, Osteopenia,

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 05 | May 2026 www.irjet.net p-ISSN: 2395-0072

and Osteoporosis and split into training, validation, and test sets. In the training phase, MobileNetV2 is used with transfer learning from ImageNet weights, and a custom classification layer is added. The model is trained using suitable loss functions while monitoring validation performance to select the best model. In deployment, the trained model is integrated into a web-based interface that preprocesses input images consistently and outputs class predictions with probability scores

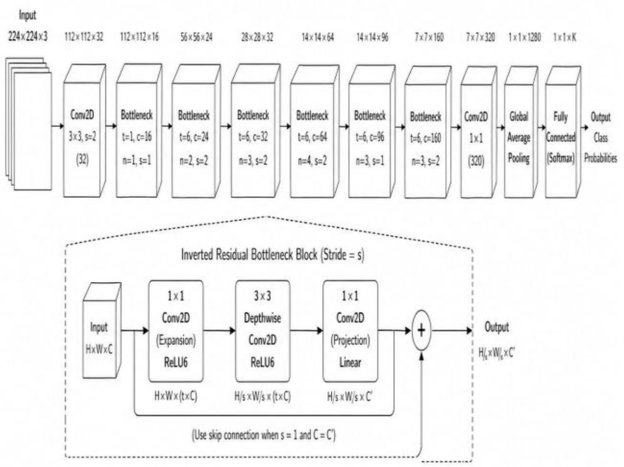

Fig:2- MobileNetV2 Architecture Diagram

The proposed model is built using the MobileNetV2 architecture, which is known for being lightweight and efficient while still providing strong classification performance. The input X-ray image of size 224 × 224 × 3 is first passed through an initial convolution layer to extract basic features such as edges and textures. After this, the network uses inverted residual bottleneck blocks, which are the main building units of MobileNetV2 and help in learning features efficiently using depth wise separable convolutions. In these blocks, features are first expanded, then processed using depth wise convolution, and finally compressed back, with skip connections added at certain stages to preserve important information and improve learning. As the network goes deeper, the spatial size of the image reduces while the number of extracted features increases, allowing the model to capture more detailed and complex patterns. At the end, global average pooling is applied to convert all extracted features into a single compact representation, which is then passed to a SoftMax classification layer for final prediction. The model also uses transfer learning with pre-trained Image Net weights, which are fine-tuned for the specific task of knee osteoarthritis detection. The complete workflow of the proposed system is clearly represented in Fig. 2.

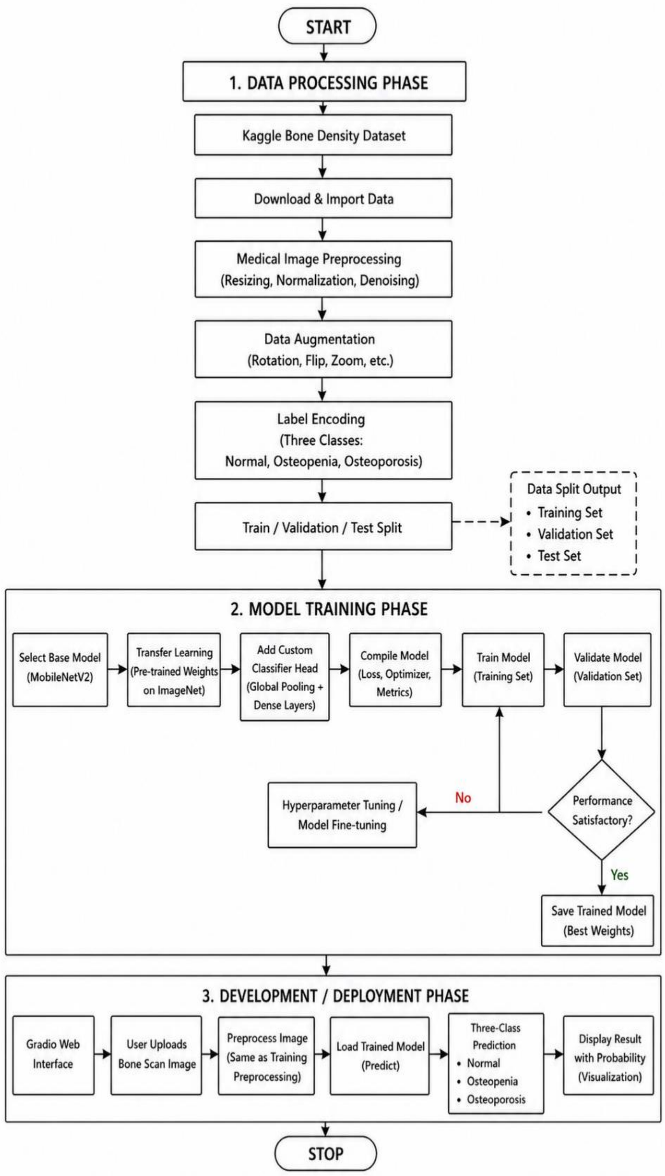

Fig-3: Flowchart

3.3 Training Process

During the training phase, the deep learning model learns meaningful patterns from knee X-ray images using the prepared training dataset. The dataset is divided into training and validation sets, where the training data is used for learning and the validation data is used to evaluate performance during training.

The model's performance is continuously monitored using key evaluation metrics such as accuracy, loss, precision, recall, and F1-score. If the validation performance is not satisfactory (for example, low accuracy or high loss), hyperparameter tuning and model fine-tuning are performed to improve learning efficiency. This process is repeated until the model achieves stable and optimal performance. Once satisfactory results are obtained, the best-performing model is saved for deployment.

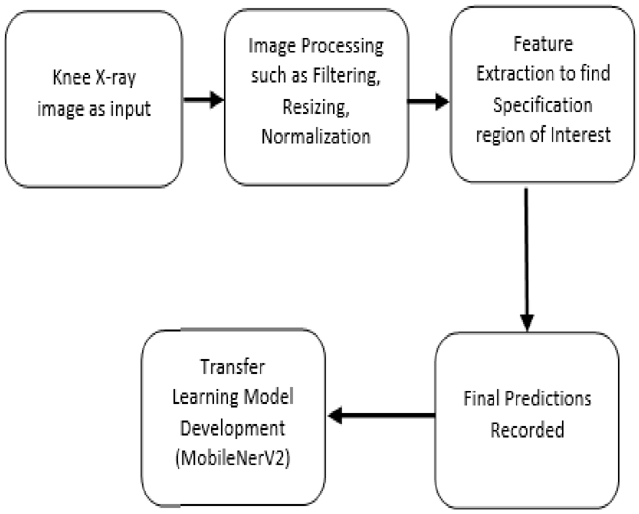

Fig:4- Block Diagram

To understand model performance in a more detailed way, evaluation is based on a confusion matrix, which consists of four outcomes: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

- **True Positive (TP):** The model correctly predicts a positive case (e.g., correctly identifying osteoarthritis).
- **True Negative (TN):** The model correctly predicts a negative case (e.g., correctly identifying a healthy knee).
- **False Positive (FP):** The model incorrectly predicts a positive case when it is actually negative.
- **False Negative (FN):** The model incorrectly predicts a negative case when it is actually positive.

Using these values, standard performance metrics are calculated as follows:

Accuracy=TP+TN/ TP+TN+FP+FN (1)

Accuracy represents the overall correctness of the model's predictions.

Precision=TP/TP+FP (2)
Precision indicates how many of the predicted positive cases are actually correct.

Recall=TP/TP+FN (3)
Recall measures how many actual positive cases are correctly identified.

Specificity=TN/TN+FP (4)
Specificity shows how well the model identifies negative cases correctly.

F1=2×Precision×Recall/Precision + Recall (5)
F1-score provides a balanced measure between precision and recall, especially useful when class distribution is uneven.

3.4 Deployment

After training, the model is integrated into a Gradio-based web interface. Users can upload X-ray images and instantly receive predictions with confidence scores.

4. RESULTS AND DISCUSSION

The proposed system successfully classifies knee X-ray images into three severi-ty levels with good consistency. The use of MobileNetV2 helps in achieving a balance between accuracy and computational efficiency. The model shows stable performance during training and validation, indicating that it has learned meaningful radiographic patterns such as joint space narrowing and bone deformation.

The integration of the Gradio interface makes the system practical and easy to use. Users can obtain instant results, which reduces dependency on manual diagnosis and speeds up initial screening. the system demonstrates that deep learning can be effectively used as a supportive tool in medical image analysis.

5 CONCLUSIONS

The system improves the efficiency of diagnosing knee osteoarthritis by significantly reducing the time required to analyze X-ray images compared to traditional manual methods. It enhances reliability by minimizing human errors and ensuring more consistent, data-driven predictions. Early detection through this approach enables timely medical intervention, which can greatly improve patient recovery and quality of life. Additionally, the proposed method makes advanced diagnostic support more accessible, even in healthcare settings with limited re-sources. The integration of deep learning allows the model to continuously im-prove with new data, increasing its accuracy over time. In essence, this system serves as a supportive tool for medical professionals by assisting in faster, more informed, and efficient decision-making in real-world clinical scenarios.

References

1. Tariq, T. – Knee Osteoarthritis Detection and Classification Using X-Rays – 2023. https://doi.org/10.1148/ryai.2020190065

2. Kondal – Automatic Grading of Knee Osteoarthritis on the Kellgren–Lawrence Scale from Radiographs Using Convolutional Neural Networks – 2020. https://ieeexplore.ieee.org/

3. Li – Automatic grading of knee osteoarthritis with a plain radiograph radiomics model: combining anteroposterior and lateral images – 2024. https://doi.org/10.1186/s13244-024-01525-4

4. Goswami – Automatic Classification Using Enhanced Image Sharpening and CNN – 2023 https://www.mdpi.com/

5. Subramanian – A Multi-Site Study on AI-Driven Pathology Detection and Osteoarthritis Grading from Knee X-Ray – 2025. https://arxiv.org/

6. Thomas, K. A. – Automated classification of radiographic knee osteoarthritis severity using deep neural networks – 2020. https://doi.org/10.1148/ryai.2020190065

7. Sarvamangala, D. R. – Grading of knee osteoarthritis using convolutional neural networks – 2021 https://doi.org/10.1007/s11063-020-10327-9

8. Wang, Y. – Learning from highly confident samples for automatic knee osteoarthritis severity assessment – 2022. https://doi.org/10.1109/JBHI.2021.3108395

9. Antony, J. – Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks – 2017. https://doi.org/10.1109/ISBI.2017.7950660

10. Thomas, C. – Automated detection of knee osteoarthritis severity using deep learning on radiographs – 2019. https://ieeexplore.ieee.org/

11. Antony, J. – Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks – 2017. https://doi.org/10.1109/ISBI.2017.7950660

12. Imran, S. – Knee osteoarthritis detection and severity classification using deep learning – 2021 https://doi.org/10.1186/s12880-021-00600-3

13. Kokkotis, C. – Machine learning in knee osteoarthritis: Areview2022.https://doi.org/10.1016/j.ocarto.2022.100222

14. Neogi, T. – The epidemiology and impact of pain in osteoarthritis2013https://doi.org/10.1016/j.joca.2013.03.018

15. Hunter, D. J. – Osteoarthritis – 2019 https://doi.org/10.1016/S0140-6736(19)30417-9

16. Guermazi – Imaging of osteoarthritis: update from a radiological perspective – 2013 https://doi.org/10.1038/nrrheum.2013.181

17. Arden, N. K. – Osteoarthritis: epidemiology and classification – 2011 https://doi.org/10.1016/j.berh.2010.08.001