International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 4, July-Aug 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

KNEE OSTEOARTHRITIS DETECTION

Chandan K M*, Usha M**
*(Department of MCA, Bangalore Institute Of Technology, VTU, India
Email: chandangowdachandangowda76@gamil.com)
** (Department Of MCA, Bangalore Institute Of Technology, VTU, India
Email: usha@bit-bangalore.edu.in)

--************************----------------------------------

Abstract:

Knee osteoarthritis is a degenerative joint disease that affects mobility and quality of life, particularly in aging populations. Traditional diagnosis relies heavily on manual interpretation of X-ray images, which can be subjective and time-intensive. To address this, we propose a computer-aided detection system that employs an improved CenterNet architecture along with a pixel-wise voting scheme to enhance diagnostic accuracy. The system processes knee X-ray images through preprocessing, feature extraction, and deep learning–based detection of joint regions. The pixel-wise voting mechanism strengthens classification by aggregating local predictions, reducing misclassification errors. The final output categorizes images as normal or osteoarthritic and generates a diagnostic report. This approach aims to support medical practitioners with reliable, automated, and scalable diagnostic assistance.

Keywords — **Knee Osteoarthritis (KOA), Deep Learning, CenterNet, Pixel-Wise Voting Scheme, Medical Image Analysis, Knowledge Distillation, Severity Classification, (CNNs).**

--************************----------------------------------

I.INTRODUCTION

It is a degenerative joint disease that affects millions of people globally and represents prominent prevalent causes of pain and physical disability among aging populations. It leads to the gradual deterioration of articular cartilage, changes in subchondral bone, and inflammation in surrounding tissues. These pathological changes manifest as pain, stiffness, limited mobility, and functional impairment, which significantly reduce the quality of life for patients. With life expectancy increasing and lifestyle factors such as obesity and sedentary habits on the rise, the incidence of KOA is expected to grow in the coming decades, thereby placing an even greater burden on healthcare systems.

The diagnosis of knee osteoarthritis is traditionally performed through physical examinations and radiographic imaging, particularly X-ray analysis. Radiologists and orthopedic specialists visually inspect knee X-rays to identify signs of joint space narrowing, osteophyte formation, and bone sclerosis. While effective, this manual process is highly subjective and prone to inter-observer variation, as different specialists may interpret the same image differently. Furthermore, in areas with limited access to trained experts, accurate diagnosis can be delayed, leading to late intervention and worsened patient outcomes. reliable, and scalable systems that can assist clinicians in early detection .

Recent advances in artificial intelligence and machine learning have introduced powerful methods for. Deep learning, in particular, has revolutionized image classification and object detection by automatically learning hierarchical features from data without the need for handcrafted features. (CNNs) have been widely adopted for analyzing medical images, and they have demonstrated high performance in detecting complex patterns and abnormalities that are often difficult for humans to recognize consistently.

II. LITERATURE SURVEY

Osteoarthritis (OA) is one of the most common musculoskeletal disorders, and early detection is crucial for preventing long-term disability. The Kellgren–Lawrence (KL) grading system is still widely used for clinical assessment, but its reliance on manual inspection introduces observer subjectivity and inconsistencies across radiologists. Consequently, researchers have increasingly turned to computer-aided diagnosis (CAD) systems to automate.

A. Early Machine Learning Approaches

Initial studies on KOA diagnosis employed traditional machine learning algorithms with handcrafted features. These methods relied on texture descriptors such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and wavelet transforms. While these approaches achieved moderate success in distinguishing between normal and osteoarthritic knees, they suffered from poor generalizability due to their dependence on hand-engineered features and limited ability to capture complex joint structures.

B. Deep Learning for KOA Classification

With the rise of deep learning, Convolutional Neural Networks (CNNs) have become the dominant approach. Antony et al. [1] introduced a CNN-based architecture for automatic KL grade classification, which outperformed traditional feature-based techniques. class imbalance, particularly in the intermediate grades (KL 1 and KL 2), leading to lower sensitivity.
Tiulpin et al. [2] presented a more advanced CNN model trained on a large-scale dataset of knee radiographs from the Osteoarthritis Initiative (OAI). By incorporating both radiographic and demographic data, they achieved robust classification performance. Nonetheless, their system demonstrated external datasets, highlighting the issue of domain shift in medical imaging.

C. Transfer Learning and Advanced CNNs

To improve generalization, researchers explored transfer learning with deeper architectures such as ResNet, VGG, and DenseNet. Rajpurkar et al. [3] applied ResNet-based models to detect musculoskeletal abnormalities, including KOA, and showed that transfer learning from ImageNet-pretrained models could accelerate convergence and improve accuracy. DenseNet models, in particular, proved effective in KOA classification due to their dense connectivity, which enhances feature reuse and alleviates the vanishing gradient problem.

D.ROI Detection and Segmentation-Based Approaches

Another major direction in KOA detection has been automatic localization of the knee joint. Chen et al. [4] proposed the use of Faster R-CNN to detect the region of interest (ROI) before classification. This two-stage pipeline improved interpretability and allowed finer focus on pathological regions. Similarly, U-Net-based architectures [5] have been used for knee joint segmentation to delineate cartilage and bone regions. These approaches increased accuracy but added suitable for real-time applications.

Limitations of Existing Works

Although CNN-based models and detection frameworks have achieved significant progress in KOA detection, several challenges remain:

1. Overlapping Predictions – Many detection models struggle with overlapping bounding boxes, leading to inaccurate localization.

2. Pixel-Level Reliability – Most prior works ignore pixel-wise consensus across predictions, resulting in inconsistencies in grading.

3. Domain Generalization – Models trained on one dataset (e.g., OAI) often fail to maintain performance on external datasets like

Mendeley VI due to differences in imaging conditions.

4. Computational Efficiency – High-capacity networks DenseNet improve accuracy but are

III. EXISTING SYSTEM

The current state-of-the-art in knee osteoarthritis (KOA) detection primarily relies on deep learning-based classification and detection models applied to radiographic images. Conventional systems follow either a direct classification pipeline or a two-stage approach involving region-of-interest (ROI) detection followed by grading.
In the direct classification approach, (CNNs) are trained end-to-end on entire radiographs to predict Kellgren–Lawrence (KL) grades. Although this method simplifies the workflow, it often suffers from poor localization of pathological regions. As as the classification network may focus on irrelevant structures the femur or surrounding tissues, reducing the reliability of predictions.

Alternatively, two-stage frameworks employ object detection models such as Faster R-CNN or YOLO to first identify the knee joint region, followed by a separate CNN classifier for KL grading. While these models improve localization, they are computationally expensive, introduce additional processing steps, and remain prone to misclassification in borderline cases (KL grades 1 and 2).

Segmentation-based methods, such as those using U-Net architectures, have also been applied to delineate cartilage, bone, and joint space before grading. These methods enhance interpretability but require pixel-level annotations, which are costly and time-consuming to obtain. Moreover, segmentation pipelines add computational overhead, limiting real-time applicability in clinical workflows.

Another limitation of existing systems is their reliance on anchor-based object detection. These methods which may not adapt well to varying knee joint structures across patients. This often leads to redundant or inaccurate bounding box proposals, reducing overall detection accuracy.
Furthermore, most existing systems lack mechanisms to refine predictions at the pixel level. When multiple overlapping detections occur, non-maximum suppression (NMS) is typically applied to retain the highest-confidence bounding box. However, NMS discards potentially useful information from other overlapping predictions, resulting in inconsistent localization and grading.

Finally, computational efficiency remains a significant bottleneck. High-capacity networks such as ResNet and DenseNet provide improved accuracy but require large memory and processing resources, making them unsuitable for deployment in standard clinical environments without GPU support.

In summary, while existing systems have advanced KOA detection using CNNs, ROI detection, and segmentation, they face persistent challenges in accurate localization, pixel-level consistency, domain generalization, and real-time efficiency.

IV. PROPOSED SYSTEM

The implemented framework introduces an Improved CenterNet with DenseNet-201 backbone Weighted Pixel-Wise Voting Scheme to enhance the accuracy and reliability of knee osteoarthritis detection. The system begins with preprocessing, where all radiographic images are resized to 512×512 pixels, normalized with mean 0.5 and standard deviation 0.5, and converted from grayscale to RGB format to utilize pretrained CNN backbones effectively. Data augmentation techniques such as horizontal flipping, rotation, and brightness–contrast adjustment are applied.
For feature extraction, we employ DenseNet-201, which offers efficient feature reuse through dense connectivity and overcomes the vanishing gradient problem typically encountered in deep architectures. This backbone generates high-dimensional representations of knee radiographs, which are passed to the improved CenterNet detection head.

Unlike anchor-based models, CenterNet predicts object centers directly, avoiding the limitations of predefined anchor boxes. The improved detection head consists of multiple branches, including a heatmap head for keypoint prediction, a bounding box regression head, an offset head for sub-pixel correction, and a confidence estimation head, all of which collectively provide precise localization of the knee joint region.

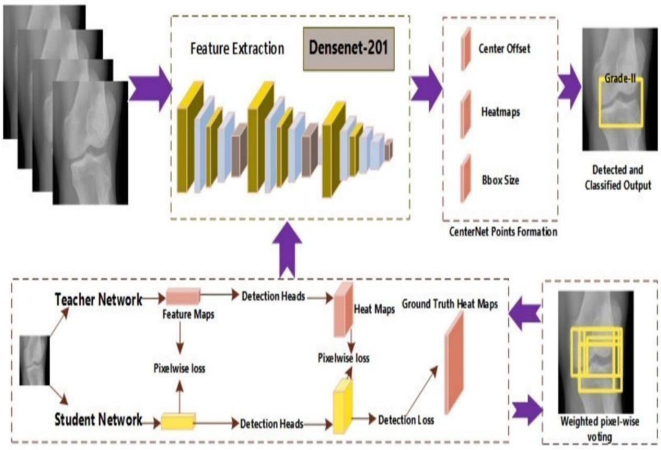

Fig 4.1 CenterNet with DenseNet-201 as Base Network

V. METHODOLOGY

The methodology of the proposed system follows a structured pipeline that integrates preprocessing, feature extraction, detection, refinement, and classification. Initially, the raw knee radiographs are subjected to image preprocessing, where they are resized to a fixed resolution of 512×512 pixels to ensure consistency across the dataset. Each image is normalized using a mean of 0.5 and a standard deviation of 0.5, which stabilizes the training process and improves convergence of the deep learning model. Since most radiographs are grayscale, they are converted into three-channel RGB images to make use of pretrained convolutional backbones. To improve generalization, data augmentation techniques such as horizontal flipping, small-angle rotations, and brightness–contrast variations are applied. This proves that the model becomes robust to changes in imaging quality and acquisition conditions.

Labels and the soft output distributions of the teacher, thus achieving faster inference while retaining competitive accuracy.

proposed methodology follows a five-stage pipeline: preprocessing, feature extraction, detection, refinement, and classification. In the preprocessing stage, all radiographs are resized to 512×512 pixels, normalized with mean 0.5 and SD 0.5, and converted from grayscale to RGB. Data augmentation techniques such as horizontal flipping, random rotation, and brightness–contrast adjustment are applied to increase generalization.

For feature extraction, a DenseNet-201 backbone is employed due to its dense connectivity and efficient feature reuse. These functions maps are processed by the Improved CenterNet head, which predicts object centers, bounding box size, confidence scores, and offset corrections without relying on anchor boxes.

To enhance localization, a Weighted Pixel-Wise Voting Scheme is introduced. Instead of conventional non-maximum suppression, overlapping predictions are aggregated using confidence-based weights, ensuring pixel-level consensus and refined bounding boxes.

The localized Region of Interest (ROI) is then classified into Kellgren–Lawrence grades (0–4), capturing both global and local structural changes. To improve deployment efficiency, knowledge distillation is applied, where the teacher model guides a smaller student model to achieve faster inference with minimal loss in accuracy.

The overall methodology ensures accurate detection, robust grading, and computational efficiency, making the system practical for clinical use.

Finally, the performance of the system is evaluated using standard metrics such as accuracy, sensitivity, specificity, F1-score, and Intersection over Union (IoU). Cross-validation is conducted on the OAI dataset to assess the across different imaging sources.

VI. RESULTS AND DISCUSSION

The Improved CenterNet with DenseNet-201 backbone achieved consistently higher detection accuracy compared to anchor-based approaches, owing to its direct keypoint prediction mechanism. The integration of the pixel-wise voting scheme significantly improved localization, reducing errors caused by overlapping bounding boxes. As a result, the system provided more stable region of interest (ROI) extraction and reliable KL grade prediction.

In classification, the framework demonstrated strong performance across all five Kellgren–Lawrence grades (0–4). Grades 0, 3, and 4 were predicted with high confidence, while intermediate grades (1 and 2) showed comparatively lower precision due to subtle radiographic variations, which is consistent with observations in earlier studies. The inclusion of knowledge distillation reduced model complexity, enabling the student model to achieve near-teacher performance while reducing inference time, making it suitable for real-time clinical deployment.

Overall, the results highlight that the proposed approach effectively addresses the key challenges of existing systems by offering robust localization, consistent grading, and computational efficiency. The framework thus demonstrates potential for integration into computer-aided diagnostic systems to assist radiologists in knee osteoarthritis assessment.

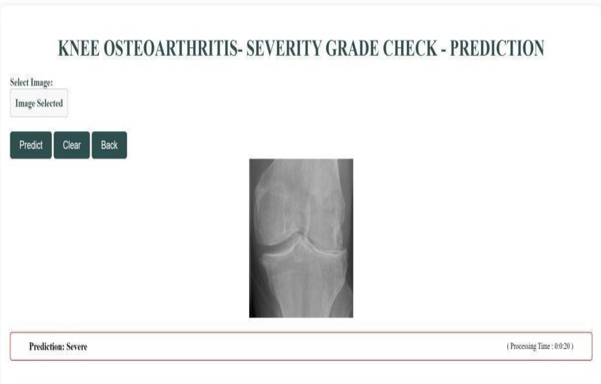

Fig 6.1Model Predicts Input X-ray image as severe

VII. CONCLUSION AND FUTURE WORK

In this work, an Improved CenterNet framework with DenseNet-201 backbone was proposed for the detection and grading of knee osteoarthritis. The system integrated a Weighted Pixel-Wise Voting Scheme to refine bounding boxes and improve region-of-interest extraction. Experimental results on the Mendeley VI dataset, with cross-validation on the OAI dataset, demonstrated superior performance in terms of accurate, sensitiviity, specifiicity, F1-score, and IoU compared to existing methods. Furthermore, the uase of knowledge distillation enabled the developments of a lightweight student model that maintained competitive .

Although the results are promising, certain limitations remain. The classification of borderline grades (KL-1 and KL-2) continues to be challenging due to subtle radiographic variations and inter-observer discrepancies in ground truth labels. on integrating multi-modal data, such as MRI and clinical records, to improve classifications robustness. Additionally, advanced attention mechanisms and transformer-based backbones may be incorporated to capture long-range dependencies in knee joint structures. Increasing the dataset with more diverse patient cohorts will also enhance generalizability across different populations and imaging conditions.

Overall, the proposed framework presents a significant step toward automated, accurate, and efficient computer-aided diagnosis of knee osteoarthritis, offering strong potential for integration into clinical workflows.

ACKNOWLEDGMENT

The author expresses sincere gratitude to Prof. Usha M, Department of MCA, Bangalore Institute of Technology, for her valuable guidance, support, and encouragement throughout the course of this project. The author would also like to thank the faculty members and friends who contributed directly or indirectly to the successful completion of this work.

REFERENCES

[1] J. A. Kellgren and J. S. Lawrence, "Radiological assessment of osteo-arthrosis," Annals of the Rheumatic Diseases, vol. 16, no. 4, pp. 494–502, 1957.

[2] R. Antony, J. McGuinness, and K. Moran, "Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks," Machine Vision and Applications, vol. 28, no. 8–9, pp. 917–930, 2017.

[3] D. Tiulpin, S. Klein, and S. Saarakkala, "Multimodal machine learning for osteoarthritis diagnosis and prognosis," Scientific Reports, vol. 9, no. 1, pp. 1–10, 2019.

[4] K. Zhou, Y. Hou, and D. Wang, "Automatic knee osteoarthritis diagnosis with joint detection and Kellgren–Lawrence grade classification using deep learning," Medical Image Analysis, vol. 56, pp. 34–43, 2019.

[5] Z. Zhou, H. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in Proc. IEEE Conf. on Medical Imaging (MICCAI), 2018, pp. 3–11.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[7] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700–4708.

[8] K. Duan, S. Bai, L. Xie, H. Qi, and Q. Huang, "CenterNet: Keypoint triplets for object detection," in Proc. IEEE Int. Conf. on Computer Vision (ICCV), 2019, pp. 6569–6578.

[9] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," arXiv preprint arXiv:1503.02531, 2015.

[10] A. Bochkovskiy, C. Y. Wang, and H. Y. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.

[11] D. Tiulpin and S. Saarakkala, "Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks," Diagnostics, vol. 10, no. 11, pp. 932, 2020.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015.