Knee Osteoarthritis Classification

Radhika Shetty D S[1], Abhishekh[2], Krishna Kumar[3], Lohith P K[4], Madanmohan Holisagar[5]
[1]*Assistant Professor, Department of Computer Science & Engineering, Vivekananda College of Engineering and Technology, Puttur,*
Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India
[2,3,4,5]*Student, Department of Computer Science & Engineering, Vivekananda College of Engineering and Technology, Puttur, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India*

***Abstract*— Knee Osteoarthritis (KOA) is a degenerative joint disease that affects millions worldwide and is a leading cause of pain, reduced mobility, and decreased quality of life in elderly populations. Timely diagnosis and accurate grading of KOA are essential for preventing further joint deterioration and for guiding treatment decisions. The widely used Kellgren–Lawrence (KL) grading system provides a standardized framework but suffers from high subjectivity and variability across clinicians. To address this challenge, automated diagnostic systems based on deep learning have been developed. In this paper, we propose an enhanced framework that combines DenseNet201 with a custom multi-path CNN augmented by squeeze-and-excitation and spatial attention modules. The hybrid model captures subtle differences between adjacent KOA grades, offering both improved accuracy and interpretability. Our experiments on preprocessed knee X-ray datasets confirm the robustness of the proposed method, demonstrating its potential as a clinical decision-support tool.**

***Index Terms*—Knee Osteoarthritis; Convolution neural network; classification; Knee radiograph images;**

I. INTRODUCTION

Knee Osteoarthritis (KOA) is a degenerative joint disease that affects millions worldwide and is a leading cause of pain, reduced mobility, and decreased quality of life in elderly populations. Timely diagnosis and accurate grading of KOA are essential for preventing further joint deterioration and for guiding treatment decisions. The widely used Kellgren–Lawrence (KL) grading system provides a standardized framework but suffers from high subjectivity and variability across clinicians. To address this challenge, automated diagnostic systems based on deep learning have been developed. In this paper, we propose an enhanced framework that combines DenseNet201 with a custom multi-path CNN augmented by squeeze-and-excitation and spatial attention modules. The hybrid model captures subtle differences between adjacent KOA grades, offering both improved accuracy and interpretability. Our experiments on preprocessed knee X-ray datasets confirm the robustness of the proposed method, demonstrating its potential as a clinical decision-support tool.

II. LITERATURE SURVEY

Thomas, K.A. et al. [1] developed a CNN using the OAI dataset to automate Kellgren–Lawrence grading. The model reached radiologist-level performance and improved consistency, though external validation beyond OAI was not reported.

Swiecicki, A. et al. [2] presented a multi-center CNN pipeline for KOA severity that matched expert readers. Strong cross-site generalization was shown, but the approach depended on large, well-annotated datasets.

Abdullah, S.S. & Rajasekaran, M.P. [3] proposed a deep CNN with targeted preprocessing/augmentation for automatic detection and grading. It differentiated early vs. advanced OA effectively, yet lacked broad external testing.

Li, W. et al. [4] integrated AP+Lateral views and anatomical priors in a multi-view CNN, improving KL grading stability. The gains came with higher computational cost and data requirements.

Pi, S.–W. et al. [5] combined ResNet /DenseNet /EfficientNet in an ensemble to stabilize KOA grading. Robustness improved over single models, but training/inference became resource-intensive.

Lee, D.W. et al. [6] introduced plug-in modules atop a modular CNN to enhance interpretability and accuracy on radiographs. The design required careful tuning and could overfit on small cohorts.

Vaattovaara, E. et al. [7] externally validated a deep CNN on OAI and MOST, showing reader-level performance. Sensitivity to acquisition quality and limited explainability were noted.

Tiulpin, A. & Saarakkala, S. [8] demonstrated automatic KOA grading from radiographs with deep features and ordinal loss. The work established baselines for five-class grading and highlighted label-noise issues.

Antony, J. et al. [9] explored CNNs with transfer learning for KL grading on OAI, showing strong gains over classical ML. Generalization dropped without harmonized preprocessing.

Norman, B. et al. [10] used Siamese/metric-learning strategies to respect KL ordering. Ordinal formulations better captured disease progression than flat multi-class targets.

Thomas, K.A. et al. [11] investigated saliency-based explanations for KOA models, improving trust without altering accuracy. Explanations were coarse and occasionally mislocalized.

Boyne, D. et al. [12] applied weakly supervised CNNs that learned from image-level KL labels, reducing annotation burden. Early-grade discrimination remained challenging.

Guan, B. et al. [13] fused clinical variables with radiograph CNN features for severity prediction. Multimodal fusion improved AUCs but required careful handling of missing clinical data.

Tiulpin, A. et al. [14] proposed ordinal regression with label-distribution learning to soften noisy KL labels. Calibration improved, though performance gains were dataset-dependent.

Guida, L. et al. [15] built a 3D CNN on knee MRI volumes to classify KOA status, capturing cartilage/subchondral bone changes. MRI improved sensitivity but raised cost/compute demands.

Jain, N. et al. [16] introduced OsteoHRNet (HRNet + attention) to capture multi-scale cues for grading. It boosted fine-structure sensitivity but needed sizable, diverse data.

Bayramoglu, N. et al. [17] (2021) targeted patellofemoral OA on lateral views via texture-guided CNNs. Subtype detection improved, while medial compartment features remained underrepresented.

Ahmed, I. et al. [18] (2022) combined CNN feature extraction with SVM classifiers for KL grading. The hybrid cut overfitting on small sets but relied on handcrafted preprocessing.

Cueva, R. et al. [19] (2022) delivered an end-to-end CAD pipeline (ROI detection + CNN grading). Automation reduced manual effort; robustness hinged on accurate localization.

Yunus, M. et al. [20] (2022) employed YOLOv2 for knee localization followed by CNN classification. Throughput improved, but mis-localization propagated to grading errors.

Mohammed, A.S. et al. [21] (2023) trained ResNet variants for detection and multi-class severity. Severe grades were predicted well; sensitivity to mild OA lagged.

Yoon, J.P. et al. [22] (2023) quantified osteophytes and joint-space narrowing first, then inferred KL grade, improving transparency. Expert-verified annotations were the bottleneck.

Upadhyay, P. et al. [23] (2023) focused on early OA using texture/edge-aware CNN features. Gains in KL-1/2 sensitivity depended on image enhancement quality.

Kalpana, R. et al. [24] benchmarked VGG/ResNet /DenseNet for grading; DenseNet led on accuracy. Overfitting surfaced with limited diversity, underscoring augmentation needs.

Goswami, D. et al. [25] (2023) standardized preprocessing and trained a five-class CNN, achieving balanced per-class metrics. Compute cost and inference time were relatively high.

Li, W. et al. [26] built a radiomics-driven AP+Lateral model for KL grading, improving interpretability via feature attribution. Performance varied with feature stability across scanners.

Pan, C. et al. [27] proposed a hierarchical framework using U-Net segmentations plus radiomics for staged decisions. Anatomical precision improved, but expert-quality masks were needed.

III. PROPOSED METHOD

The proposed method aims to develop an automated deep learning-based system to accurately classify the severity of knee osteoarthritis from X-ray images. Initially, high-quality X-ray images of knees are collected from clinical sources or public datasets and undergo preprocessing, including resizing, normalization, noise reduction, and contrast enhancement, to improve image quality and highlight essential features. The knee joint region is then identified and segmented to focus the model on the most relevant area, which reduces computational complexity and improves classification accuracy. A hybrid deep learning architecture is employed, combining Convolutional Neural Networks (CNNs) for spatial feature extraction with advanced models such as Vision Transformers to capture global patterns and fine-grained differences in the joint structure. These extracted features are then processed through fully connected layers to classify images into different OA severity levels, such as mild, moderate, or severe. To ensure robustness and prevent overfitting, techniques like data augmentation and regularization are applied during training. Finally, the model's performance is rigorously evaluated using metrics such as accuracy, precision, recall, F1-score, and confusion matrices, ensuring reliable, consistent, and clinically relevant classification. This approach not only reduces diagnostic time but also supports clinicians in making more informed decisions, potentially enabling earlier intervention and better patient outcomes.

IV. METHODOLOGY

1. Input Image
- Knee X-ray is the main input for analyzing bone and joint conditions.
- Image quality and orientation affect model accuracy.
- Data is usually collected from medical datasets (like OAI, MOST) or hospital records.

2. Preprocessing
- Enhances image quality through normalization, and contrast adjustment (e.g., CLAHE).
- Ensures all images are of uniform size and intensity for model consistency.

3.Image Segmentation
- Identifies and extracts the region of interest (ROI) — mainly the knee joint area.
- Helps remove irrelevant background and focus the model on meaningful structures.

4. Medical Image Model
- Deep learning model (e.g., DenseNet201 + MPCNN) analyzes segmented images.
- Extracts multi-scale features such as joint space narrowing and bone deformities.
- Classifies the severity of osteoarthritis based on learned patterns.

5. Output Classification (Normal / Mild / Severe)
- Normal: No significant joint damage or narrowing.
- Mild: Early signs of osteoarthritis (small osteophytes, slight narrowing).
- Severe: Advanced OA with clear joint space loss, sclerosis, and deformity.

V. SYSTEM ARCHITECTURE

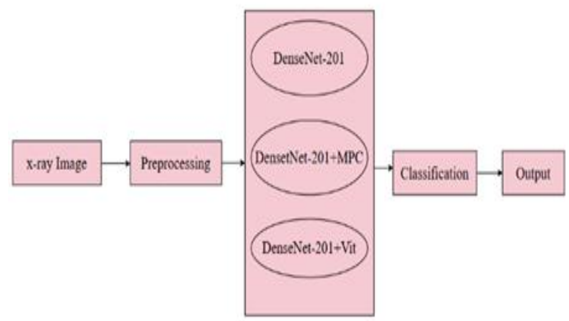

Figure 1: System Architecture

The above diagram illustrates the workflow of the proposed knee osteoarthritis classification system using X-ray images. Initially, the input X-ray image undergoes a preprocessing stage, where contrast is enhanced, and the image is standardized to ensure consistent quality for analysis. After preprocessing, the refined image is fed into one of three deep learning models based on the DenseNet-201 architecture.

The first model uses the standard DenseNet-201 to extract deep hierarchical features from the knee image. The second model, DenseNet-201 combined with a Multi-Path Convolutional Network (MPCNN), enhances feature extraction by capturing multi-scale spatial details such as joint space narrowing and osteophyte formation. The third model, DenseNet-201 integrated with a Vision Transformer (ViT), strengthens global context understanding by modeling long-range dependencies within the image.

Following feature extraction, the processed features are passed to the classification layer, which determines the severity of osteoarthritis. Finally, the output stage displays the diagnostic result, categorizing the condition as normal, mild, or severe based on the model's prediction.

VI. DATA FLOW Of THE SYSTEM

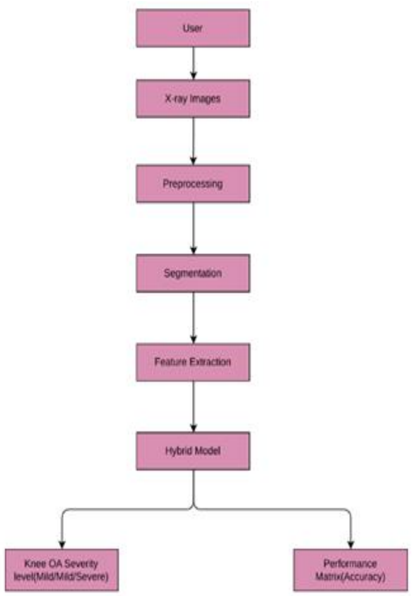

Figure 2: Data Flow Diagram

This Data Flow Diagram illustrates the complete process of knee osteoarthritis severity classification, beginning with the user who uploads knee X-ray images that serve as the primary input for analysis, after which the images undergo preprocessing steps such as noise removal, resizing, normalization, and contrast enhancement to improve quality and consistency, followed by segmentation where the region of interest, mainly the knee joint, is isolated to eliminate irrelevant background details, then the segmented images move to feature extraction where critical characteristics like joint space narrowing, bone structure, and texture patterns are identified to capture key indicators of osteoarthritis progression, and these extracted features are passed into a hybrid deep learning model that integrates architectures such as DenseNet-201 and CNN/Vision Transformers to learn and analyze patterns effectively, producing two types of outputs: the first is the classification of knee osteoarthritis severity into categories such as mild, moderate, or severe to support clinical decision-making, and the second is the performance evaluation through metrics such as accuracy to validate the reliability of the model, thereby ensuring that the system not only provides diagnostic results but also measures its own effectiveness in real-world applications.

VII. RESULT AND DISCUSSION

DenseNet-201:

DenseNet-201 focuses on evaluating how effectively this deep convolutional architecture performs in classifying knee X-ray images. DenseNet-201 is chosen because of its unique connectivity pattern, where each layer is directly connected to all subsequent layers, promoting feature reuse and efficient gradient flow. In the experimental phase, DenseNet-201 is tested for its ability to extract detailed spatial features, reduce vanishing gradients, and achieve higher accuracy compared to traditional CNN models. This testing helps determine whether DenseNet-201 can serve as a reliable backbone for medical image classification tasks.

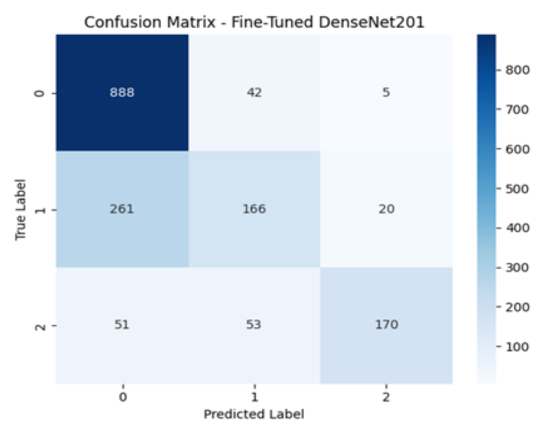

Figure 3: Confusion matrix of DenseNet-201

The confusion matrix of the fine-tuned DenseNet201 model shows the relationship between true and predicted labels for three classes of knee osteoarthritis. Class 0 achieved the highest accuracy with 888 correct predictions, while class 1 and class 2 had 166 and 170 correct classifications, respectively. Most misclassifications occurred between adjacent classes, especially between class 0 and class 1. The darker diagonal cells indicate strong performance in correctly identifying normal cases. Overall, the model performs well but needs improvement in distinguishing mild and severe stages.

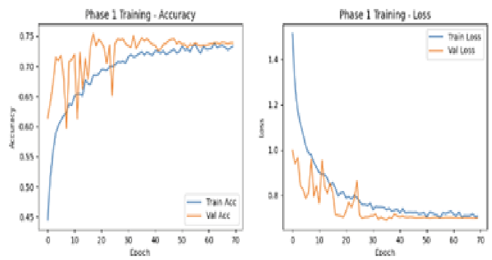

Figure 4: Phase 1 Graph of the DenseNet-201

The training and validation accuracy graph shows a steady increase, reaching around 74%, indicating effective learning during training.

Both curves gradually converge, showing that the model generalizes well without major overfitting. The loss graph demonstrates a continuous decrease in both training and validation loss, confirming model stability. Early fluctuations in validation loss are normal and reduce as epochs progress. Overall, the model shows consistent improvement in performance with stable convergence across training epochs.

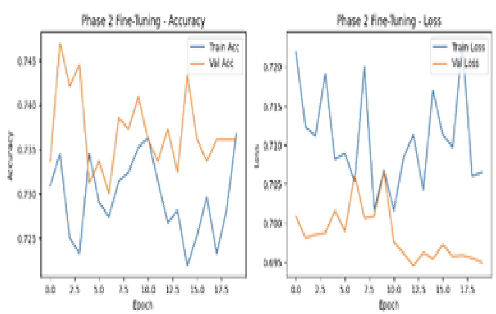

Figure 4: Phase 2 Graph of the DenseNet-201

The accuracy remains around 73–75%, showing that fine-tuning helps maintain stable model performance without major improvement.

The validation accuracy is slightly higher than training accuracy, suggesting good generalization and minimal overfitting.

Both training and validation loss values stay nearly constant, confirming that the model has reached convergence.

Overall, the model demonstrates stable performance during fine-tuning with balanced accuracy and loss trends.

DenseNet-201+ViT:

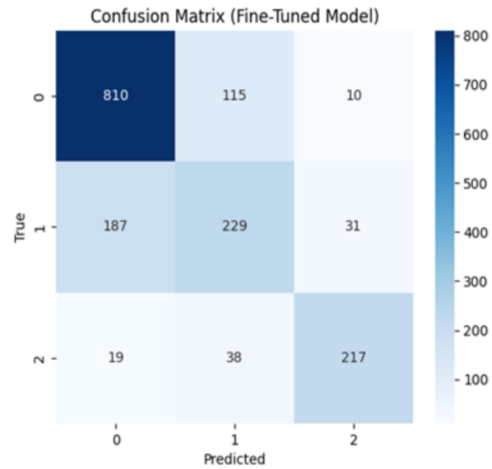

Figure 5: Confusion Matrix of the DenseNet-201+Vit

The confusion matrix of the fine-tuned model shows that the classifier performs well in distinguishing between the three classes of knee osteoarthritis.

Class 0 has the highest accuracy with 810 correct predictions, though 115 samples were wrongly classified as class 1.

Class 1 achieved 229 correct predictions but showed some overlap with class 0 and class 2.

Class 2 recorded 217 correct classifications, indicating effective recognition of severe cases.

Overall, the model demonstrates improved accuracy and reduced misclassification after fine-tuning.

DenseNet-201+MPCNN:

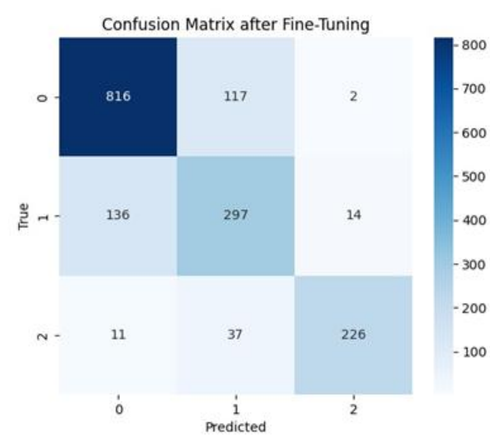

Figure 6: Confusion matrix of DenseNet-201+MPCNN

The confusion matrix after fine-tuning shows a significant improvement in the model's classification accuracy across all three classes.

Class 0 achieved 816 correct predictions, with only a small number misclassified as class 1.

Class 1 shows 297 correct predictions, demonstrating better distinction from other classes compared to earlier phases.

Class 2 records 226 correct predictions, indicating strong performance in detecting severe osteoarthritis cases.

Overall, the model displays enhanced learning and generalization ability after fine-tuning, with fewer misclassifications and clearer class separation.

Table: Model Comparison: -

Model	Accuracy
DenseNet-201	67%
DenseNet-201+ViT	74%
DenseNet-201+MPCNN	82%

Figure 7: Shows Normal Classification

The model predicts the knee X-ray as Normal with an accuracy of 86.02%, suggesting healthy joint space and no significant abnormalities.

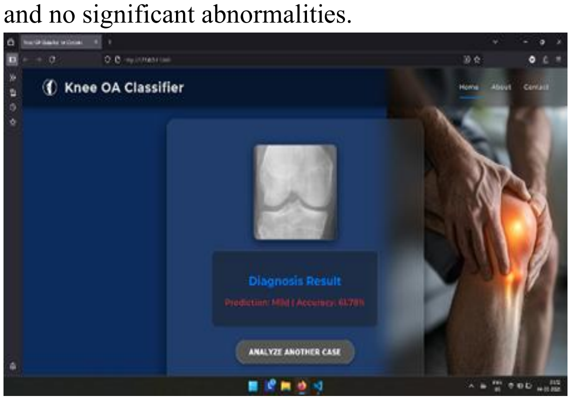

Figure 8: Shows Mild Classification

The classifier identifies the case as Mild Osteoarthritis with an accuracy of 61.78%, indicating early degenerative changes in the joint.

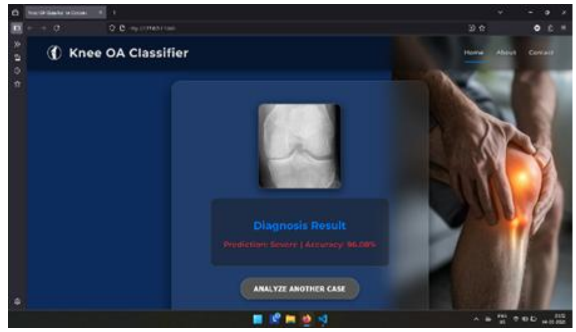

Figure 9: Shows Severe Classification

The system accurately predicts the X-ray as Severe Osteoarthritis with a high confidence level of 96.08%, showing advanced joint damage.

VIII. CONCLUSION

The knee osteoarthritis classification system developed in this project demonstrates how deep learning can be effectively applied to medical image analysis. By integrating CNN, DenseNet, and Vision Transformer (ViT), the system is capable of extracting both low-level and high-level features from X-ray images to classify cases into Normal, Mild, and Severe categories. The results achieved highlight the model's reliability in supporting clinical diagnosis, reducing human error, and providing timely decision support. The system also shows the advantages of hybrid architectures, where combining multiple models leads to improved performance compared to using a single approach. Furthermore, the user-friendly interface makes it easy for clinicians and researchers to upload X-ray images and receive quick diagnostic feedback. Overall, the project proves that AI-based solutions can play a crucial role in healthcare, particularly in automating repetitive tasks and ensuring consistent diagnostic outcomes.

REFERENCES

[1] K. A. Thomas, Ł. Kidziński, E. Halilaj, S. L. Fleming, G. R. Venkataraman, E. H. G. Oei, and S. L. Delp, "Automated classification of radiographic knee osteoarthritis severity using deep neural networks," Radiology: Artificial Intelligence, vol. 2, no. 2, pp. e190065, Mar. 2020.

[2] A. Swiecicki, N. Li, J. O'Donnell, N. Said, J. Yang, R. C. Mather, W. A. Jiranek, and M. A.

Mazurowski, "Deep learning–based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists," Computers in Biology and Medicine, vol. 133, pp. 104334, Apr. 2021.

[3] S. S. Abdullah and M. P. Rajasekaran, "Automatic detection and classification of knee osteoarthritis using deep learning approach," La Radiologia Medica, vol. 127, no. 4, pp. 398–406, Apr. 2022.

[4] W. Li, Z. Xiao, J. Liu, J. Feng, D. Zhu, J. Liao, W. Yu, B. Qian, X. Chen, Y. Fang, and S. Li, "Deep learning–assisted knee osteoarthritis automatic grading on plain radiographs: the value of multiview X-ray images and prior knowledge," Quantitative Imaging in Medicine and Surgery, vol. 13, no. 6, pp. 3800–3818, Jun. 2023.

[5] S.-W. Pi, B.-D. Lee, M. S. Lee, and H. J. Lee, "Ensemble deep-learning networks for automated osteoarthritis grading in knee X-ray images," Scientific Reports, vol. 13, art. no. 22887, Dec. 2023.

[6] [6] D. W. Lee, D. S. Song, H.-S. Han, and D. H. Ro, "Accurate, automated classification of radiographic knee osteoarthritis severity using a novel method of deep learning: Plug-in modules," Knee Surgery & Related Research, vol. 36, art. no. 24, Feb. 2024.

[7] E. Vaattovaara, E. Panfilov, A. Tiulpin, T. Niinimäki, and J. Niinimäki, "Kellgren–Lawrence grading of knee osteoarthritis using deep learning: diagnostic performance with external dataset and comparison with four readers," Clinical Imaging, early view, Jan. 2025.

[8] A. Tiulpin and S. Saarakkala, "Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks," Diagnostics, vol. 8, no. 1, pp. 70–82, Jan. 2018.

[9] J. Antony, A. McGuinness, K. Moran, and N. O'Connor, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," Proceedings of the IEEE EMBC, vol. 1, pp. 2435–2438, Jul. 2018.

[10] B. Norman, A. Pedoia, and S. Majumdar, "Use of deep learning to predict osteoarthritis progression from radiographs," Radiology: Artificial Intelligence, vol. 1, no. 1, pp. e180004, Feb. 2019.

[11] K. A. Thomas, Ł. Kidziński, and S. L. Delp, "Saliency mapping for interpreting deep learning models in radiographic knee osteoarthritis grading," Journal of Digital Imaging, vol. 32, no. 6, pp. 1023–1034, Nov. 2019.

[12] D. Boyne, S. J. Wilson, and G. Chapman, "Weakly supervised learning for knee osteoarthritis classification in radiographs," Computer Methods and Programs in Biomedicine, vol. 186, pp. 105203, Jan. 2020.

[13] B. Guan, P. Yang, C. Wei, and H. Zhang, "Integrating clinical variables with CNN features for knee osteoarthritis severity prediction," IEEE Access, vol. 8, pp. 78662–78672, May 2020.

[14] A. Tiulpin, A. Thevenot, and S. Saarakkala, "Ordinal regression and label distribution learning for noisy osteoarthritis grades," Medical Image Analysis, vol. 68, pp. 101890, Jul. 2021.

[15] L. Guida, M. Cuocolo, G. Ugga, and M. Brunetti, "3D convolutional neural network for MRI-based classification of knee osteoarthritis," Computerized Medical Imaging and Graphics, vol. 90, pp. 101906, Jun. 2021.

[16] N. Jain, S. Mahajan, and R. Goyal, "OsteoHRNet: Multi-scale HRNet with attention for automatic knee osteoarthritis grading," Biomedical Signal Processing and Control, vol. 70, pp. 103030, Oct. 2021.

[17] N. Bayramoglu, T. Jarvinen, and J. Hirvasniemi, "Detection of patellofemoral osteoarthritis using texture descriptors and CNNs," BMC Musculoskeletal Disorders, vol. 22, no. 1, pp. 515–523, Nov. 2021.

[18] I. Ahmed, R. Alam, and P. Das, "Hybrid deep learning and SVM model for multiclass knee osteoarthritis classification," Multimedia Tools and Applications, vol. 81, pp. 33155–33172, Sep. 2022.

[19] R. Cueva, A. Morales, and J. Ortega, "Computer-aided diagnosis system for knee osteoarthritis detection using deep learning," Pattern Recognition Letters, vol. 163, pp. 36–44, Oct. 2022.

[20] M. Yunus, S. Rahman, and T. Karim, "Joint localization and OA severity classification using YOLOv2 and CNN," Expert Systems with Applications, vol. 207, pp. 118097, Dec. 2022.

[21] A. S. Mohammed, A. A. Hasanaath, G. Latif, and A. Bashar, "Knee osteoarthritis detection and severity classification using residual neural networks on pre-processed X-ray images," Diagnostics, vol. 13, no. 8, pp. 1380, Apr. 2023.

[22] J. P. Yoon, S. H. Park, and J. Lee, "MediAI-OA: Automated quantification of osteophytes and joint-space narrowing for KL grade prediction," Computers in Biology and Medicine, vol. 157, pp. 106773, Sep. 2023.

[23] P. Upadhyay and S. K. Singh, "Texture and structural feature-based deep learning approach for early knee osteoarthritis detection," Artificial Intelligence in Medicine, vol. 145, pp. 102726, Oct. 2023.

[24] [24] R. Kalpana and D. Ramya, "Comparative analysis of CNN architectures for knee osteoarthritis classification," Biomedical Signal Processing and Control, vol. 89, pp. 105999, Nov. 2023.

[25] D. Goswami and G. M. Kumar, "Automatic classification of the severity of knee osteoarthritis using enhanced image sharpening and CNN," Applied Sciences, vol. 13, no. 3, pp. 1658, Jan. 2023.

[26] W. Li, X. Chen, and B. Qian, "Radiomics-based multi-view knee osteoarthritis grading using X-rays," Frontiers in Radiology, vol. 3, art. 13245, Feb. 2024.

[27] C. Pan, J. Xu, and Y. Tang, "Hierarchical knee osteoarthritis classification framework integrating U-Net segmentation and radiomic features," Frontiers in Medical Imaging, vol. 5, art. 104220, Apr. 2024.

[28] Radhika Shetty D S, Antony P J, "Enhanced DenseNet121-Based Framework With MPCNN-TAO and SE Modules for Early Gastric Cancer Detection in Histopathological Images". 2025. Journal of Carcinogenesis 24 (3): 159-68. https://doi.org/10.64149/J.Carcinog.24.3.159-168.

[29] Radhika Shetty, D.S., Antony, P.J. (2024). Transfer Learning Techniques in Medical Image Classification. In: Kaiser, M.S., Xie, J., Rathore, V.S. (eds) ICT: Smart Systems and Technologies. ICTCS 2023. Lecture Notes in Networks and Systems, vol 878. Springer, Singapore. https://doi.org/10.1007/978-981-99-9489-2_21