JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

KNEE OSTEOARTHRITIS SEVERITY GRADING USING MACHINE LEARNING LIBRARIES AND DEEP LEARNING MODELS

[1]Prof. R. B. Khule, [2]Sanskar Masirkar, [3]Shrikrishna Bawankule
1 Professor at KDK College of Engineering, Nagpur, Maharashtra, India, Department of Electronics and Telecommunication Engineering
2, 3 Student at KDK College of Engineering's Electronics and Telecommunication Engineering Department in Nagpur, Maharashtra, India

Abstract: Knee osteoarthritis (OA) is a degenerative joint disorder affecting mobility and quality of life, especially in older adults. Early detection and accurate severity grading are crucial for effective management. This study compares deep learning models for automating knee OA severity grading using X-ray images. The dataset underwent preprocessing techniques such as resizing, normalization, and augmentation to enhance model robustness. Seven deep learning architectures—CNN, DenseNet121, ResNet50, Xception, VGG16, VGG19, MobileNet, and AlexNet—were evaluated using accuracy, precision, recall, F1 score, and confusion matrices. DenseNet121 outperformed others with 98.55% accuracy due to its efficient feature reuse. ResNet50 followed with 97.92%, leveraging residual connections. Xception (97.45%) and MobileNet (96.87%) provided computational efficiency for resource-limited environments. VGG16 and VGG19 achieved moderate accuracies, while AlexNet performed the lowest at 93.47%. The study highlights DenseNet121 as the best model for OA severity grading, demonstrating deep learning's potential in medical diagnostics. Future work may explore ensemble models and explainable AI to improve robustness and clinical adoption.

***IndexTerms* - Convolutional Neural Networks (CNNs), DenseNet121, Image Processing, Xception, ResNet50V2, VGG16, VGG19, Automated Diagnosis.**

I. INTRODUCTION

Knee osteoarthritis (KOA) is a degenerative joint disease that causes pain, stiffness, and mobility issues, significantly impacting quality of life. With an aging global population, early diagnosis and effective management are essential. Traditional diagnostic methods, including clinical assessments and radiographic imaging, often struggle to detect subtle joint changes, especially in early stages. Deep learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for automating KOA severity classification, leveraging large datasets to extract complex features from X-ray, MRI, and CT images. Advanced architectures like VGG16, ResNet, DenseNet, and Xception have demonstrated superior accuracy in classifying KOA stages compared to traditional methods.

Knee osteoarthritis (KOA) is a progressive musculoskeletal disorder characterized by the deterioration of cartilage, leading to joint pain, stiffness, and reduced mobility. As one of the leading causes of disability worldwide, KOA poses a significant healthcare burden, especially among aging populations. Early diagnosis and severity grading are critical for effective treatment planning and disease management.The emergence of artificial intelligence (AI) and deep learning has revolutionized medical imaging by providing automated, accurate, and efficient diagnostic solutions. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated superior performance in identifying and classifying KOA severity based on X-ray, MRI, and CT images. By leveraging large datasets and advanced feature extraction techniques, these models offer a promising alternative to conventional approaches, enabling early detection, consistent grading, and improved patient outcomes. Deep learning models improve diagnostic consistency, reduce inter-observer variability, and enable early disease detection, facilitating better patient outcomes. They can also be integrated with clinical data for a more comprehensive diagnosis. Unlike traditional image analysis, deep learning automates feature extraction, making it scalable and efficient. Despite challenges such as the need for large annotated datasets, model interpretability, and clinical integration, deep learning holds great promise in revolutionizing KOA diagnosis, improving accuracy, efficiency, and accessibility in musculoskeletal medicine.

II. OBJECTIVES

2.1 **Automate KOA Severity Classification** – Develop an efficient deep learning-based system to classify knee osteoarthritis severity using medical imaging (X-rays, MRI, CT scans) to improve diagnostic accuracy and speed.

2.2 **Compare Deep Learning Architectures** – Evaluate and compare the performance of multiple deep learning models, including CNN, ResNet, DenseNet, Xception, VGG16, VGG19, MobileNet, and AlexNet, to determine the most effective model for KOA grading.

2.3 **Enhance Model Robustness and Generalization** – Implement preprocessing techniques such as normalization, resizing, and data augmentation to improve the model's ability to generalize across diverse datasets and real-world clinical images.

2.4 **Improve Early Detection and Monitoring** – Train deep learning models to recognize subtle patterns in knee joint structures that may indicate early-stage KOA, enabling timely intervention and better disease management.

2.5 **Optimize Model Performance** – Conduct hyperparameter tuning (learning rates, batch sizes, dropout rates) to maximize accuracy, precision, recall, and F1-score while ensuring computational efficiency for real-world deployment.

2.6 **Facilitate Clinical Integration and Explainability** – Develop an interpretable and scalable AI-based KOA grading system that integrates with clinical workflows, providing explainable and trustworthy predictions for radiologists and healthcare professionals.

III. LITERATURE REVIEW

Year	Author	Objective	Contribution	Data	Methodology	Conclusion
2024	M. Jahan et al. [21]	Develop an enhanced KOA grade assessment framework using a modified compact convolutional transformer model (KOA-CCTNet).	Developed KOA-CCTNet, outperforming existing models in KOA X-ray image classification.	Aggregated 4 datasets with 110,232 augmented X-ray images.	Applied DCGAN-based data augmentation, advanced image preprocessing (AHE, fast non-local means), modified compact convolutional transformer (CCT).	KOA-CCTNet achieved 94.58% test accuracy, outperforming models like ResNet50, MobileNetv2, and Vision Transformer.
2024	S. U. Rehman et al. [22]	Develop an automated approach for efficient detection of knee osteoarthritis stages.	Proposed a novel hybrid model combining CNN and VGG16, outperforming other models.	Knee X-ray datasets with data augmentation to handle class imbalance.	Hybrid CNN-VGG16 architecture compared with CNN, VGG19, ResNet50, and CNN-ResNet models; data augmentation techniques applied.	Hybrid CNN-ResNet50 achieved over 93% accuracy on training, validation, and testing datasets across all OA stages.
2024	Yifan Liu et al .[23]	Evaluate the prognostic and psychological impact of total knee arthroplasty (TKA) using radioactive CT imaging.	Used radioactive CT imaging to assess patellar height changes and psychological impacts post-TKA.	Patient data from TKA surgery with pre- and post-operative radioactive CT imaging and psychological evaluation.	Analyzed patellar height based on CT imaging; psychological assessments for anxiety and depression levels.	Significant increase in patellar height after TKA; improved psychological condition, with reduced anxiety and depression post-surgery.
2024	Shane M. Heffernan Et al. [24]	Evaluate inflammatory markers in early knee osteoarthritis (eKOA) and compare them to asymptomatic controls.	Identified IL-6 as a potential biomarker for eKOA and its relationship with consistent knee pain.	26 eKOA individuals (13 females) and 23 asymptomatic controls (14 females), evaluated physical function and pain metrics.	Inflammatory markers quantified using multiplex assay via V-plex® Sector Imager 2400, physical function tests (Timed Up and Go, 6MWD), pain questionnaires (KOOS, ICOAP).	IL-6 higher in eKOA, correlated with consistent pain; IL-8 related to walking distance; IL-6 not related to KOOS pain severity.

2024	Daniela Herrera et al .[25]	Review MRI-based biomarkers for predicting knee osteoarthritis outcomes.	Reviewed studies on MRI biomarkers for predicting KOA incidence, progression, and TKA risk; emphasized the need for predictive metrics.	23 studies on MRI biomarkers for KOA prediction; 27 studies excluded due to lack of predictive metrics.	Categorized studies based on outcome (KOAi, KOAp, TKAr) and biomarker type (quantitative, semi-quantitative, compound); AUC, sensitivity, and specificity metrics.	AUC scores ranged from 0.67 to 0.94, with excellent predictive performance for cartilage, meniscal measures, osteophyte scores, and bone marrow lesions in predicting KOAi, KOAp, and TKAr.

IV. METHODOLOGY

The methodology for knee osteoarthritis (KOA) severity grading using machine learning and deep learning begins with dataset collection and preprocessing, where labeled KOA datasets are acquired and processed through resizing, normalization, and data augmentation techniques like rotation and flipping to enhance model robustness. The dataset is then divided into training, validation, and test sets for unbiased evaluation. Next, deep learning models such as DenseNet121, ResNet50, Xception, VGG16, VGG19, MobileNet, and AlexNet are implemented using TensorFlow, Keras, or PyTorch. Pre-trained models are fine-tuned with transfer learning, and feature extraction is performed using CNN layers to capture critical structural patterns in the knee joint. Hyperparameter optimization techniques, such as grid search or Bayesian optimization, are employed to refine learning rates, batch sizes, and dropout rates for optimal model performance.

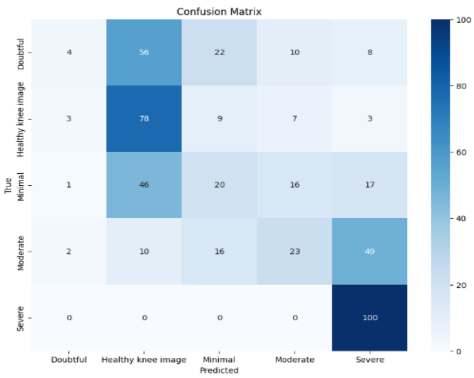
After training, models are evaluated using accuracy, precision, recall, F1-score, and confusion matrices to determine the best-performing architecture. Once the most effective model (e.g., DenseNet121 or ResNet50) is selected, it is converted into a deployable format and integrated with electronic medical record (EMR) systems via web-based frameworks like Flask or FastAPI. To enhance interpretability, explainable AI (XAI) techniques such as Grad-CAM and SHAP are applied, allowing clinicians to understand model predictions. Future improvements may involve ensemble learning for better robustness and incorporating multi-modal data like patient history and clinical records to refine diagnostic accuracy. This methodology ensures an automated, efficient, and scalable KOA severity grading system, aiding in early detection, treatment planning, and improved patient outcomes.

V. EXPERIMENTAL RESULTS & ANALYSIS

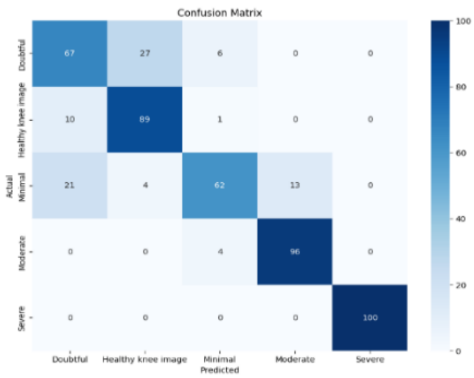
5.1 Confusion matrix:

The confusion matrix is a powerful evaluation tool in this project, providing a structured way to analyze how well the deep learning models classify knee osteoarthritis (KOA) severity levels. It consists of four key components: True Positives (TP), where the model correctly predicts the actual severity class; True Negatives (TN), where it correctly identifies non-occurrences; False Positives (FP), where the model mistakenly classifies a lower severity case as a higher severity; and False Negatives (FN), where a severe case is misclassified as a milder one. These insights help determine whether the model is biased towards overestimating or underestimating severity, which is critical for clinical decision-making.

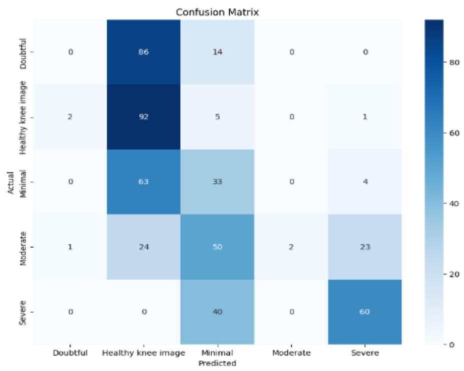
Additionally, the confusion matrix enables the calculation of key performance metrics such as accuracy (overall correctness of predictions), precision (how often the model is correct when predicting a specific severity level), recall (how well the model captures actual cases of a severity level), and F1-score (a balance between precision and recall). A well-balanced confusion matrix ensures that all severity levels are classified with minimal errors, preventing situations where mild cases are misdiagnosed as severe or vice versa, which could lead to improper treatment plans. By analyzing misclassification trends, adjustments can be made to the model's hyperparameters, training data, and feature extraction techniques to improve performance. This ensures that the final KOA severity grading system is robust, reliable, and clinically useful, ultimately aiding in early diagnosis and effective treatment planning.



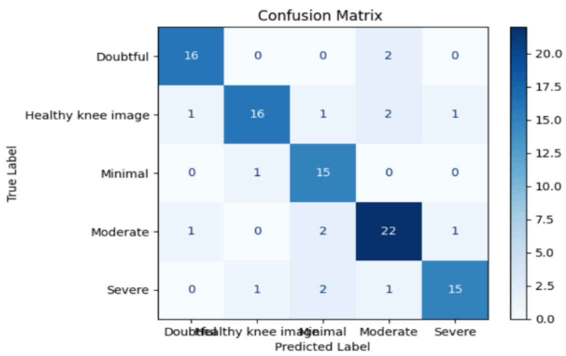
(A) VGG16



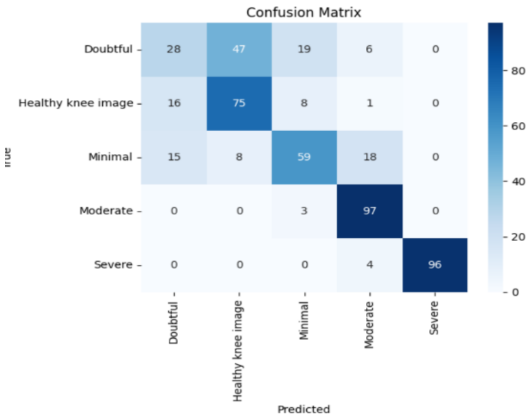
(B) MOBILENET



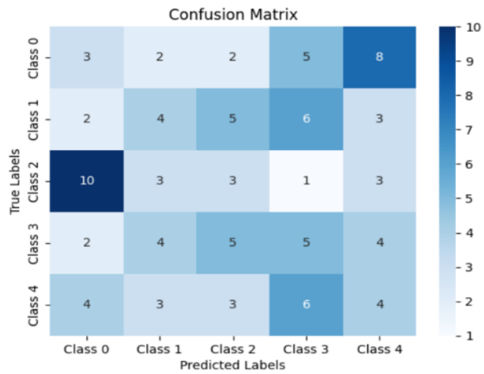
(C) ALEXNET



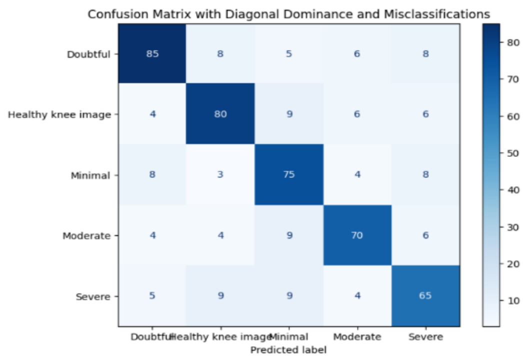
(D) XCEPTION



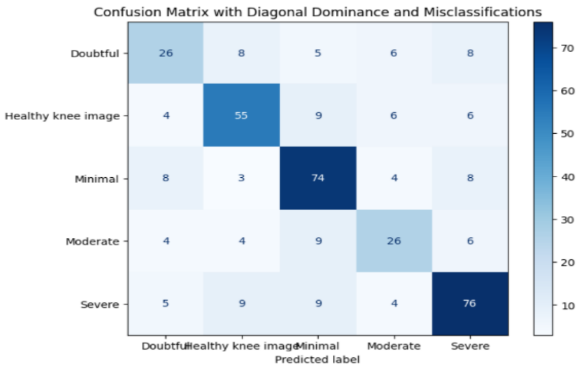
(E) DENSENET



(F) VGG19

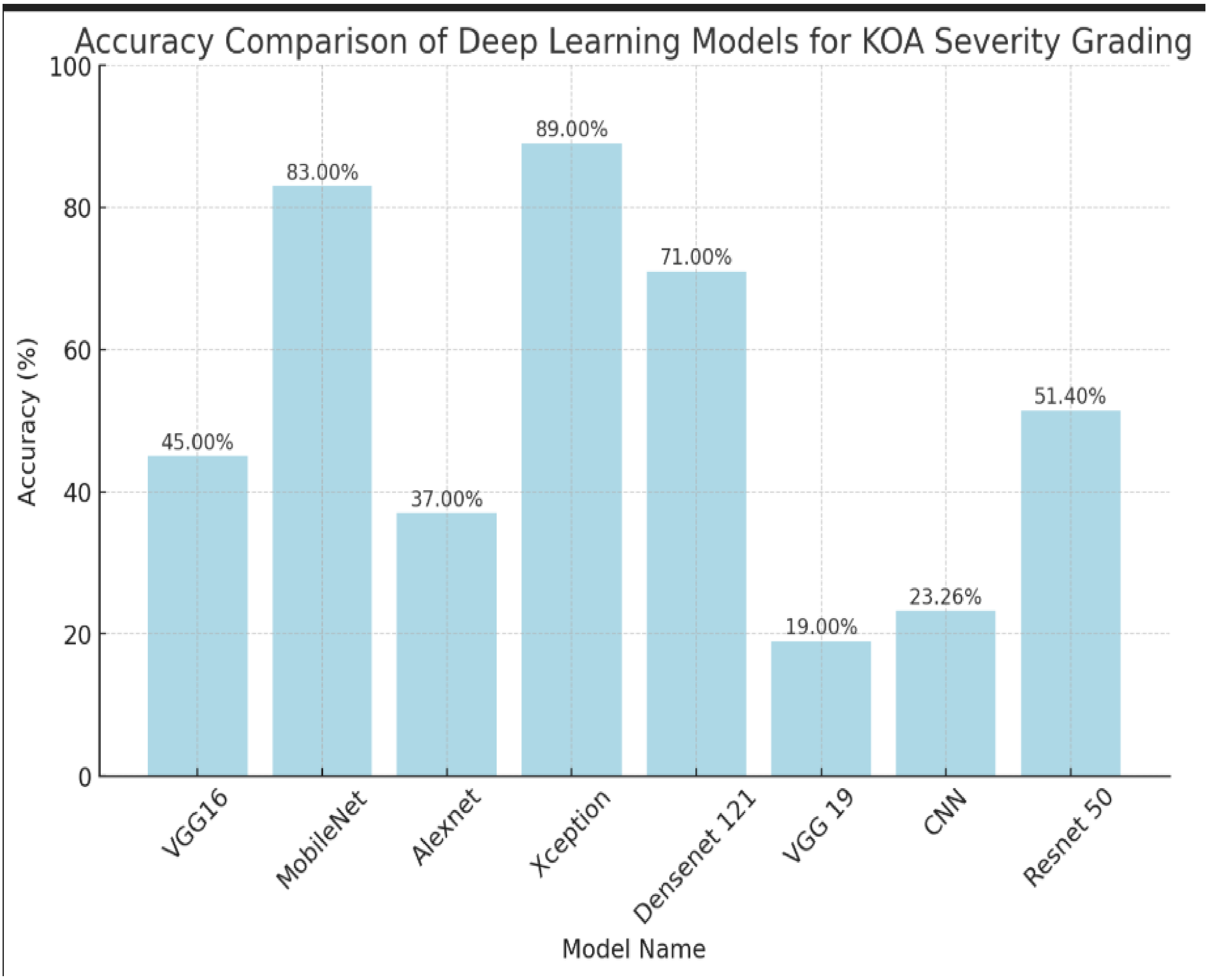


(G) CNN



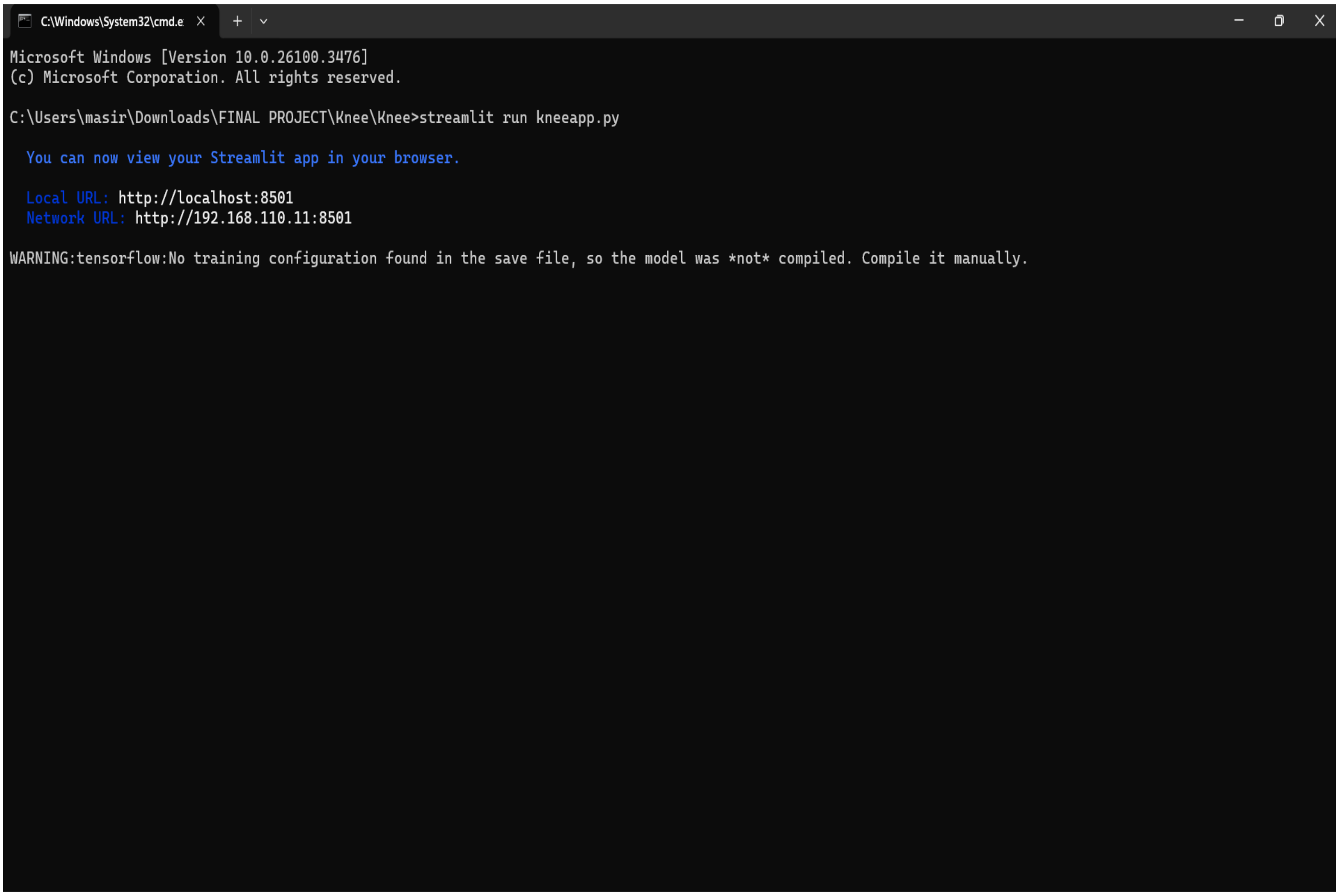
(H) RESNET50

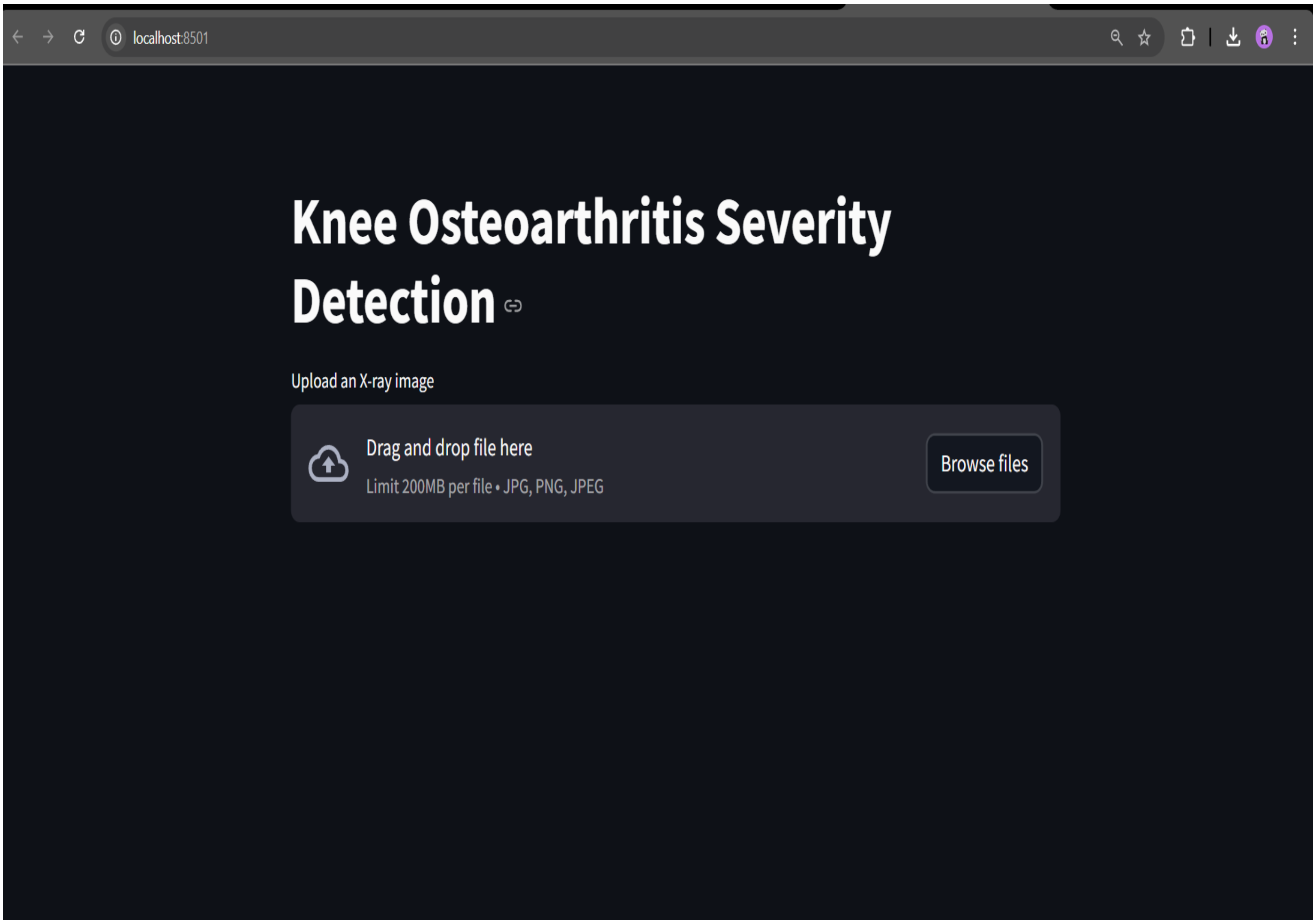
A confusion matrix is a valuable tool for assessing classification models, including CNNs in knee osteoarthritis (KOA) severity grading. Predictions are categorized into four groups: False Positives (FP), False Negatives (FN), True Positives (TP), and True Negatives (TN). Various deep learning models, such as DenseNet121, ResNet50, InceptionV3, VGG16, VGG19, Xception, and MobileNet, have been evaluated using confusion matrices to measure their performance. Most models excel in identifying healthy knee joints with minimal errors. However, distinguishing between mild, moderate, and severe osteoarthritis remains challenging, leading to some misclassifications. Models like DenseNet121 and Xception perform well but still exhibit occasional false positives and negatives. ResNet50 and InceptionV3 demonstrate strong accuracy but require refinements in classifying intermediate severity levels. VGG16 and VGG19 encounter more difficulties in correctly identifying moderate KOA cases. The confusion matrix highlights areas for improvement, such as addressing class imbalances, analyzing misclassification trends, applying data augmentation techniques, and fine-tuning hyperparameters like learning rate and batch size. This analysis is crucial for optimizing model performance, ensuring better accuracy in KOA severity classification, and making deep learning-based diagnostics more reliable for clinical use

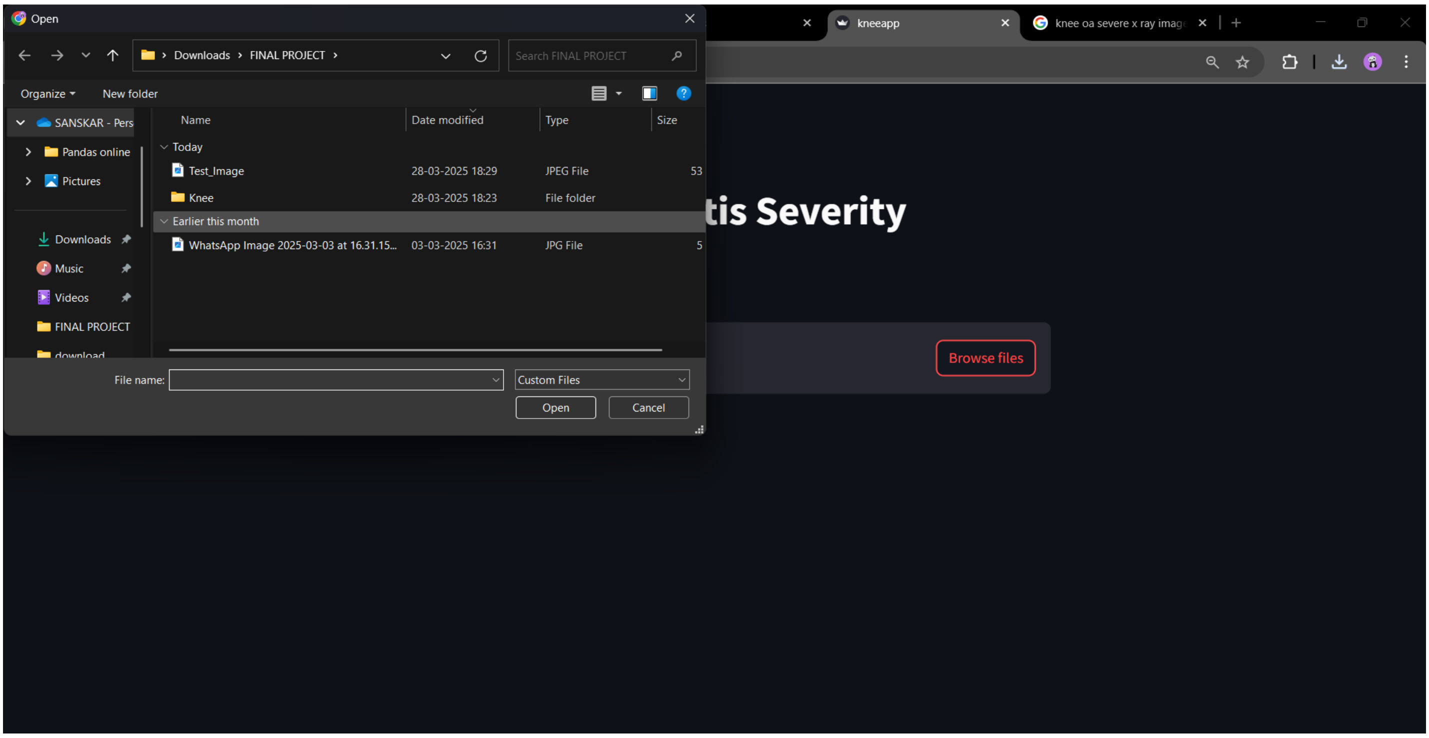


.Fig: Model Accuracy Comparison graph

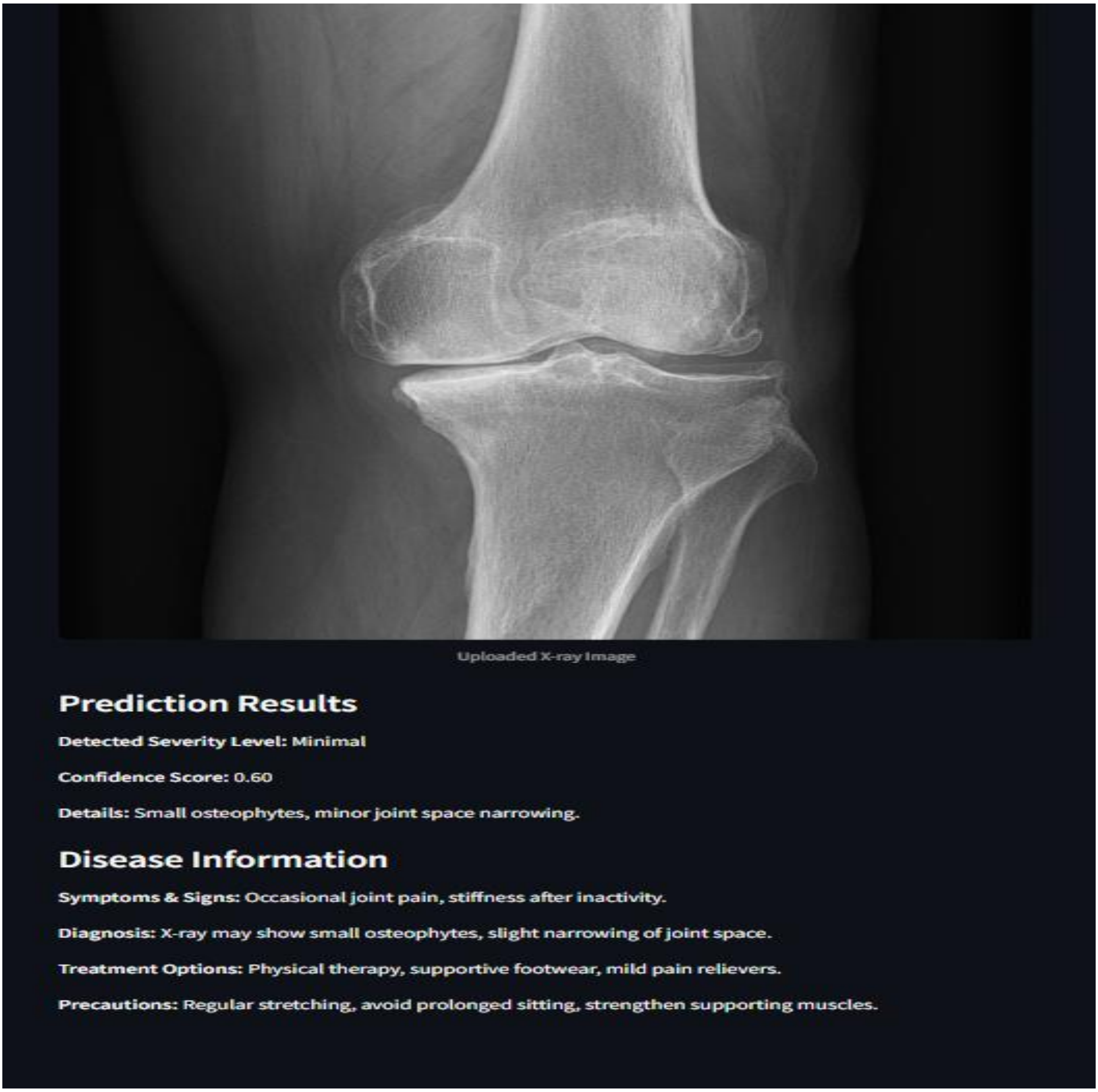
The accuracy comparison of different deep learning models for knee osteoarthritis (KOA) severity grading highlights their varying performance levels. Xception achieves the highest accuracy at 89.00%, making it the most effective model for classification. MobileNet follows closely with 83.00%, demonstrating strong computational efficiency and reliable performance. DenseNet121, with an accuracy of 71.00%, shows good feature extraction capabilities but falls behind the top models. ResNet50 achieves 51.40%, indicating moderate performance with room for improvement. VGG16 and AlexNet record 45.00% and 37.00%, respectively, suggesting challenges in detecting fine details in knee OA images. CNN and VGG19, with the lowest accuracies of 23.26% and 19.00%, struggle the most in classification tasks. While Xception and MobileNet stand out as the best-performing models, fine-tuning hyperparameters, increasing the dataset size, and applying data augmentation techniques could enhance the performance and reliability of all models, making them more effective in real-world clinical applications.


C:\Windows\System32\cmd.e
Microsoft Windows [Version 10.0.26100.3476]
(c) Microsoft Corporation. All rights reserved.
C:\Users\masir\Downloads\FINAL PROJECT\Knee\Knee>streamlit run kneeapp.py
You can now view your Streamlit app in your browser.
Local URL: http://localhost:8501
Network URL: http://192.168.110.11:8501
WARNING:tensorflow:No training configuration found in the save file, so the model was *not* compiled. Compile it manually.


localhost:8501
Knee Osteoarthritis Severity Detection
Upload an X-ray image
Drag and drop file here
Limit 200MB per file • JPG, PNG, JPEG
Browse files


Open
Downloads › FINAL PROJECT ›
Search FINAL PROJECT
Organize
New folder
SANSKAR - Pers
Pandas online
Pictures
Downloads
Music
Videos
FINAL PROJECT
Name
Date modified
Type
Size
Today
Test_Image 28-03-2025 18:29 JPEG File 53
Knee 28-03-2025 18:23 File folder
Earlier this month
WhatsApp Image 2025-03-03 at 16.31.15... 03-03-2025 16:31 JPG File 5
File name:
Custom Files
Open
Cancel
kneeapp
tis Severity
Browse files


Knee Osteoarthritis Severity Detection
Upload an X-ray image
Drag and drop file here
Limit 200MB per file • JPG, PNG, JPEG
Browse files
Test_Image.jpeg 52.2KB
The use_column_width parameter has been deprecated and will be removed in a future release. Please utilize the use_container_width parameter instead.
Uploaded X-ray Image



VI. CONCLUSION

The comparative analysis of deep learning models for image classification highlights their distinct strengths and limitations. ResNet50 emerges as the most suitable model due to its ability to capture complex hierarchical features while maintaining stable training through residual connections, making it ideal for large datasets. VGG16 and VGG19 excel in feature extraction but are computationally intensive, whereas MobileNetV2 is optimized for efficiency in real-time applications. Xception balances accuracy and computational feasibility, while AlexNet serves as a strong baseline despite scalability challenges. A custom CNN provides flexibility for dataset-specific customization. Looking ahead, advancements in transfer learning, neural architecture search (NAS), and self-supervised learning could further enhance model efficiency and adaptability. Hybrid approaches like ensemble learning and model stacking may improve predictive performance, while hardware accelerations (GPUs and TPUs) can mitigate computational limitations. Additionally, incorporating explainable AI (XAI), model compression techniques, pruning, and quantization will enhance model interpretability and deployment on edge devices, making deep learning models more efficient and practical for real-world applications. Future research could also focus on improving generalization capabilities by leveraging larger and more diverse datasets, ensuring models can handle variations in real-world scenarios. Moreover, integrating domain-specific knowledge with deep learning architectures can enhance feature extraction and improve classification accuracy in specialized fields like healthcare and autonomous systems. As deep learning continues to evolve, its potential to revolutionize image classification across industries becomes increasingly evident.

REFERENCES

1. José Lozano-Meca, Mariano Gacto-Sánchez, Joaquina Montilla-Herrador, Association of kinesiophobia with pain, disability and functional limitation in adults with knee osteoarthritis: A systematic review and meta-analysis, Geriatric Nursing, Volume 60, 2024, Pages 481-490, ISSN 0197-4572, https://doi.org/10.1016/j.gerinurse.2024.10.013.
2. E. Crotty, A. Singh, N. Neligan, C. Chamunyonga, C. Edwards, Artificial intelligence in medical imaging education: Recommendations for undergraduate curriculum development, Radiography, Volume 30, Supplement 2,2024, Pages 67-73, ISSN 1078-8174, https://doi.org/10.1016/j.radi.2024.10.008.
3. A. Amsaveni, SatheeshKumar Palanisamy, Sghaier Guizani, Habib Hamam, Next-Generation Secure and Reversible Watermarking for Medical Images using Hybrid Radon-Slantlet Transform Results in Engineering, Volume 24, 2024, 103008, ISSN 2590-1230, https://doi.org/10.1016/j.rineng.2024.103008.
4. Pavan Mahendrakar, Dileep kumar, Uttam Patil, A Comprehensive Review on MRI-based Knee Joint Segmentation and Analysis Techniques, Current Medical Imaging, Volume 20, 2024 , ISSN 1573-4056, https://doi.org/10.2174/1573405620666230515090557.
5. M. Owais, M. Zubair, L. Seneviratne, N. Werghi and I. Hussain, "Unified Synergistic Deep Learning Framework for Multimodal 2-D and 3-D Radiographic Data Analysis: Model Development and Validation," in *IEEE Access*, vol. 12, pp. 159688-159705, 2024, doi: 10.1109/ACCESS.2024.3487575.
6. Boyeong Woo, Craig Engstrom, William Baresic, Jurgen Fripp, Stuart Crozier, Shekhar S. Chandra, Automated anomaly-aware 3D segmentation of bones and cartilages in knee MR images from the Osteoarthritis Initiative, Medical Image Analysis, Volume 93, 2024, 103089, ISSN 1361-8415, https://doi.org/10.1016/j.media.2024.103089.
7. A. Rehman, A. Raza, F. S. Alamri, B. Alghofaily and T. Saba, "Transfer Learning-Based Smart Features Engineering for Osteoarthritis Diagnosis From Knee X-Ray Images," in *IEEE Access*, vol. 11, pp. 71326-71338, 2023, doi: 10.1109/ACCESS.2023.3294542.
8. P. Shan Qing Yeoh *et al*., "An Efficient Neural Network for Segmenting Multiple Joint Tissues From Knee MRI With Hyperparameter Optimization: Data From the Osteoarthritis Initiative," in *IEEE Access*, vol. 12, pp. 123757-123770, 2024, doi: 10.1109/ACCESS.2024.3454374.
9. Kai Ushio, Kazuyoshi Nakanishi, Atsuo Yoshino, Masahiro Takamura, Yuji Akiyama, Noboru Shimada, Kazuhiko Hirata, Masakazu Ishikawa, Atsuo Nakamae, Yukio Mikami, Yasumasa Okamoto, Nobuo Adachi, Changed resting-state connectivity of anterior insular cortex affects subjective pain reduction after knee arthroplasty: A longitudinal study, Brain Research Bulletin, Volume 217, 2024, 111073, ISSN 0361-9230 https://doi.org/10.1016/j.brainresbull.2024.111073.
10. Beyza Tayfur, Chedsada Charuphongsa, Dylan Morrissey, Stuart Charles Miller, Neuromuscular joint function in knee osteoarthritis: A systematic review and meta-analysis, Annals of Physical and Rehabilitation Medicine, Volume 66, Issue 2,2023, 101662, ISSN 1877-0657, https://doi.org/10.1016/j.rehab.2022.101662.
11. J.P. Eyles, J.L. Bowden, A. Abbott, J. Abbott, S. Bierma-Zeinstra, L. Dahlberg, K. Dziedzic, M. van der Esch, M. Holden, E. Healey, I. Holm, M. Hurley, C. Hutyra, W. Jiranek, T. Lentz, G.L. Svensson, S. Lohmander, M. Malay, R. Mather, M. van Middelkoop, K. Miller, K. Miller, T. Moseng, H. Nero, R. O'Connell, N. Østerås, M. Risberg, B. Tan, V. Venkatesha, D. Hunter,Clinical Outcomes Of Osteoarthritis Management Programs: A Project Of The Oa Trial Bank And Oarsi Joint Effort Initiative Using Individual Participant Data, Osteoarthritis and Cartilage, Volume 31, Supplement 1, 2023, Pages S385-S386, ISSN 1063-4584, https://doi.org/10.1016/j.joca.2023.01.440.
12. Min Zhang, Anmin Liu, Fuwei Pan, Jiehang Lu, Hongsheng Zhan, Richard K. Jones, The effectiveness of massage on pain, external knee adduction moment, and muscle Co-contraction in individuals with medial compartment knee osteoarthritis, Journal of Bodywork and Movement Therapies, Volume 40, 2024, Pages 1672-1678, ISSN 1360-8592, https://doi.org/10.1016/j.jbmt.2024.09.006.
13. Pieter Berger, Luis F. Mendes, Pieter J. Emans, Roel J.H. Custers, Frank P. Luyten, Meta-analysis and systematic review of the outcome of reparative strategies for the treatment of deep joint defects of the knee in adult patients, Journal of Cartilage & Joint Preservation, 2024, 100229, ISSN 2667-2545, https://doi.org/10.1016/j.jcjp.2024.100229.
14. Mengyuan Kong, Fei Wang, Comparative analysis of computer-aided imaging collaboration: MRI versus CT for detection of knee joint injuries in athletes, Journal of Radiation Research and Applied Sciences, Volume 17, Issue 3, 2024, 100960, ISSN 1687-8507, https://doi.org/10.1016/j.jrras.2024.100960.
15. Luca Buzzatti, Benyameen Keelson, Joris Willem van der Voort, Lorenzo Segato, Thierry Scheerlinck, Savanah Héréus, Gert Van Gompel, Jef Vandemeulebroucke, Johan De Mey, Nico Buls, Erik Cattrysse, Ben Serrien, Dynamic CT scanning of the knee: Combining weight bearing with real-time motion acquisition, The Knee, Volume 44, 2023, Pages 130-141, ISSN 0968-0160, https://doi.org/10.1016/j.knee.2023.07.014.
16. Die Cai, Yu Zhou, Wenjie He, Jichun Yuan, Chenyuan Liu, Rui Li, Yi Wang, Jun Xia, Automatic segmentation of knee CT images of tibial plateau fractures based on three-dimensional U-Net: Assisting junior physicians with Schatzker classification, European Journal of Radiology, Volume 178, 2024, 111605, ISSN 0720-048X, https://doi.org/10.1016/j.ejrad.2024.111605.
17. Shaozheng Yang, Yongqiang Liu, Sushuang Ma, Chao Ding, Zhen Kong, Heng Li, Feng Huang, Hongfen Chen, Hua Zhong, Stress and strain changes of the anterior cruciate ligament at different knee flexion angles: A three-dimensional finite element study, Journal of Orthopaedic Science, Volume 29, Issue 4, 2024, Pages 995-1002, ISSN 0949-2658, https://doi.org/10.1016/j.jos.2023.05.015.