Automated Detection and Severity Grading of Knee Osteoarthritis from X-ray Images Using Machine Learning with detailed literature Review

Submitted: 14th August 2024; accepted: 23rd April 2025

R.Gokulapriya, M.Balamurugan, Rakoth Kandan Sambandam, Divya Vetriveeran

DOI: 10.14313/jamris-2026-048

Abstract:
This paper presents a groundbreaking approach to the detection and severity grading of knee osteoarthritis (OA) utilizing X-ray imaging and a Convolutional Neural Network (CNN) architecture, specifically EfficientNet B5. Osteoarthritis, a prevalent degenerative joint disease, requires accurate and timely diagnosis for effective management. The proposed methodology focuses on enhancing diagnostic accuracy through advanced deep learning techniques. The study leverages a diverse dataset of knee X-ray images sourced from various clinical settings, encompassing a spectrum of OA severity levels. The EfficientNet B5 architecture, known for its superior performance and efficiency, is employed for feature extraction and classification. The CNN model is trained and validated on a defined dataset, achieving an impressive accuracy of 95.84 percent in differentiating knee OA from normal conditions across multiple grades like Healthy, Minimal, Doubtful, Moderate, and Severe. The model further enhances clinical utility and incorporates a severity grading mechanism, providing clinicians with a detailed assessment of the disease progression. The severity grading is accomplished through fine-tuning the CNN model on a sub-dataset with annotated severity levels. The integration of severity grading demonstrates the model's ability to detect knee OA accurately and quantify its severity with high precision. The findings of this research highlight the potential of EfficientNet B5-based CNNs as a robust and efficient tool for knee OA diagnosis and severity grading. The reported accuracy of 95.84 percent positions the proposed model as a promising asset in clinical settings, offering a reliable and automated solution for early detection and precise evaluation of knee OA. This research, with its potential to significantly improve the diagnostic capabilities of musculoskeletal disorders, ultimately enhances patient care and outcomes by providing a more accurate and timely diagnosis.

Keywords: *Knee osteoarthritis, EfficientNet B5, severity grading, patient care*

1. Introduction

Knee osteoarthritis (OA) is a prevalent degenerative joint disorder characterized by the gradual breakdown of the protective cartilage within the knee joint. This condition affects millions of people worldwide and is a leading cause of pain and disability, particularly in the elderly population. The knee joint, a complex hinge joint crucial for mobility, undergoes changes with knee OA that involve the deterioration of cartilage, the formation of bony outgrowths (osteophytes), and inflammation of the synovial membrane. These structural alterations lead to pain, stiffness, and reduced joint function, significantly impacting the quality of life for affected individuals. The application of artificial intelligence (AI) in knee OA diagnosis has emerged as a transformative avenue with significant potential to revolutionize the detection, management, and outcomes of this prevalent musculoskeletal condition. This work stems from the historical challenges of diagnosing knee osteoarthritis promptly and accurately. Traditional methods often fail to comprehensively understand the disease's progression and deliver precise grading for effective management. Leveraging the power of AI, particularly deep learning techniques, this study aims to revolutionize the diagnosis and management of knee OA by offering a more nuanced and reliable approach.

This work, as given in figure 1, presents a new approach to the detection and severity grading of knee OA utilizing X-ray imaging and a Convolutional Neural Network (CNN) architecture, specifically EfficientNet B5. Knee OA requires accurate and timely diagnosis for effective management. The EfficientNet B5 architecture, known for its superior performance and efficiency, is employed for feature extraction and classification. The CNN model is trained and validated on a defined dataset, achieving an impressive accuracy in differentiating knee OA from normal conditions across multiple grades like Healthy, Moderate, and Severe.

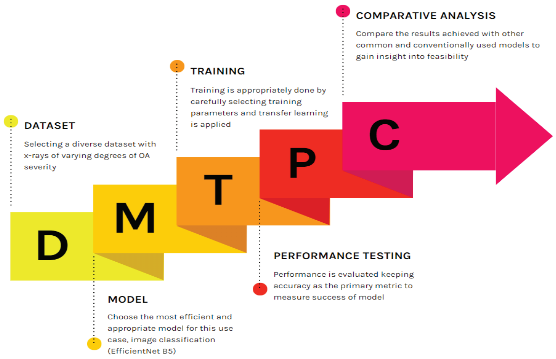


Figure 1. Flow diagram describing the model development and testing process

The study not only introduces a cutting-edge methodology for detection and grading but also compares the performance of the EfficientNet B5 model against conventional models like Random Forest and K-Nearest Neighbors (KNN), as well as advanced neural network architectures such as EfficientNetV2 B3 and Conv2D. Our findings reveal that the EfficientNet B5 model consistently outperforms conventional models Random Forest and KNN. The results showcase a remarkable accuracy for EfficientNetV2 B3 and Conv2D, highlighting the superior accuracy of the EfficientNet B5 model. The motivation behind this technical report lies in the critical need for advanced tools to improve the detection and grading of knee arthritis through x-ray imaging. Traditional methods often face challenges in providing nuanced assessments of arthritis severity, and there is a growing demand for innovative approaches to enhance diagnostic accuracy.

Many patients face the challenge of navigating busy schedules and the inconvenience of healthcare appointments. Unnecessary visits to rheumatologists contribute to patient inconvenience and strain healthcare systems grappling with high demand. By harnessing the power of deep-learning models, which can analyze medical data and images with exceptional accuracy, we can offer a transformative solution to this problem. There is potential to provide individuals with an accessible and user-friendly platform enabling them to receive an early diagnosis from their homes. This not only empowers patients to take a proactive role in their healthcare but also has the potential to significantly reduce the traffic to rheumatologists, allowing these specialists to focus on more critical cases and complex scenarios. This report also seeks to contribute to the scientific understanding of applying deep learning, specifically the EfficientNet B5 model, to the detection and severity grading of knee OA in X-ray images. By doing so, we aspire to advance the field of skeletal imaging, providing healthcare practitioners with a valuable resource for more accurate and timely diagnosis, ultimately improving patient outcomes and the overall management of knee OA.

2. Literature Review

The literature on artificial intelligence (AI), and machine learning (ML) applications in knee osteoarthritis (OA) diagnosis is diverse and impactful. Studies focus on various aspects, from diagnostic algorithms to advanced imaging techniques and prediction models.

Jang et al. [1] designed a prospective case series study to assess the effectiveness of an automated AI algorithm-based texture analysis in identifying the restoration of subchondral remodeling following therapeutic intervention. Polynucleotide (PN) filler injections were chosen as the therapeutic modality, and treatment outcomes were evaluated through symptom improvement and induced subchondral microstructural changes. The study enrolled 51 participants with knee OA, administering intra-articular PN filler injections weekly for five sessions. Knee X-rays and texture analyses with bone structure values (BSVs) were conducted at the initial screening and three months post-treatment. Visual Analogue Scale and Korean-Western Ontario MacMaster measurements demonstrated decreased scores post-PN treatment, lasting for three months after the final injection. AI algorithm-based texture analysis revealed BSV changes in the tibial bone's middle and deep layers post-PN injection, suggesting microstructural changes in the subchondral bone were detectable using this method. In conclusion, AI algorithm-based texture analysis shows promise as a tool for detecting and assessing therapeutic outcomes in knee OA.

Mohammadi et al. [2] performed a systematic review and meta-analysis to pool the data on diagnostic performance metrics of AI in OA detection and compare them with clinicians' performances. The researchers conducted a meta-analysis to aggregate data concerning diagnostic performance metrics. In order to identify potential sources of heterogeneity, the researchers undertook a subgroup analysis of the involved joint and a meta-regression based on multiple parameters.

The assessment of bias risk was done with the Prediction Model Study Risk of Bias Assessment Tool reporting guidelines. The pooled sensitivities for AI algorithms and clinicians on internal validation test sets were 88 percent and 80 percent. The pooled specificities for AI algorithms and clinicians were 81 percent and 79 percent, respectively. The pooled sensitivity and specificity for AI algorithms at external validation were 94 percent and 91 percent, respectively.

In Alshamrani et al. [3], the researchers proposed the utilization of transfer learning models, specifically sequential CNNs, such as Visual Geometry Group 16 (VGG-16) and Residual Neural Network 50, for early detection of OA in knee X-ray images. The analysis reveals that all recommended models exhibit a predictive accuracy exceeding 90 percent in OA detection. Notably, the pre-trained VGG-16 model emerge as the top-performing model, attaining a training accuracy of 99 percent and a testing accuracy of 92 percent.

A separate study by Chen et al. [4] systematically employed two deep CNNs to autonomously gauge the severity of knee OA based on the Kellgren-Lawrence (KL) grading system. Initially, utilizing a customized one-stage YOLOv2 network, knee joints were detected in X-ray images, considering the relatively consistent size and distribution of knee joints. Subsequently, popular CNN models, including ResNet variants, VGG, DenseNet, and InceptionV3, were fine-tuned to classify the detected knee joint images. The classification incorporated a novel adjustable ordinal loss, wherein higher penalties were assigned for misclassifications with larger disparities between the predicted and actual KL grades, acknowledging the ordinal nature of the KL grading task.

Evaluation of baseline X-ray images from the Osteoarthritis Initiative (OAI) dataset yielded a mean Jaccard index of 0.858 and a recall of 92.2 percent for knee joint detection under a Jaccard index threshold of 0.75. For the knee KL grading task, the fine-tuned VGG-19 model, using the proposed ordinal loss, attained the highest classification accuracy of 69.7 percent and a mean absolute error of 0.344. The study demonstrated state-of-the-art performance in knee OA detection and KL grading.

Furthering the research around computer-aided diagnosis (CAD), Brahim et al. [5] presented a fully developed CAD system for early knee OA detection, employing knee X-ray imaging and ML algorithms. The Fourier domain was utilized for initial image preprocessing through a circular Fourier filter. Then, to reduce variability between OA and healthy participants, the researchers used a novel normalizing technique based on predictive modeling with multivariate linear regression. To reduce dimensionality, an independent component analysis technique was used during the feature selection/extraction stage. The classification challenge utilized Random Forest and Gaussian Naïve Bayes classifiers. The researchers evaluated the suggested image-based method using 1,024 knee X-ray images from the OAI database. The findings showed a strong predictive classification rate of 82.98 percent accuracy, 87.15 percent sensitivity, and up to 80.65 percent specificity for OA detection.

Further research into the use of CAD systems showed promising results in knee OA detection at various stages. Song and Zhang [6] investigated a CAD method for knee osteoarthritis (KOA-CAD) that utilized multivariate information, specifically vibroarthrographic and fundamental physiological signals. The research introduced an improved deep learning model incorporating a novel Laplace distribution-based strategy (LD-S) for classification. Additionally, an aggregated multiscale dilated convolution network (AMD-CNN) was devised to extract features from the multivariate information of knee OA patients. The proposed KOA-CAD method seamlessly integrated AMD-CNN with LD-S, achieving three computer-assisted diagnosis objectives: automatic knee OA detection, early knee OA detection, and knee OA grading detection. Clinical validation with collected multivariate information demonstrated the method's effectiveness, yielding automatic knee OA detection accuracy of 93.6 percent, early knee OA detection accuracy of 92.1 percent, and knee OA grading detection accuracy of 84.2 percent.

Focusing specifically on knee OA in overweight middle-aged women who initially presented without OA, Lazzarini et al. [7] developed an analytics pipeline to identify small models that could effectively predict the 30-month incidence of knee OA. The comprehensive dataset encompassed clinical variables, responses to food and pain questionnaires, biochemical markers, and imaging-based information. Demonstrating robust performance, all models exhibited high accuracy (area under the curve (AUC) > 0.7), while utilizing only a minimal set of variables. The study delved into the significance and directional impact of each variable within the models. Ultimately, a comparative analysis was conducted, pitting the performance of two models against state-of-the-art approaches available in the existing literature.

In S.J.O. Rytky et al. [8], the authors aimed to develop and validate a ML approach for the automatic three-dimensional histopathological grading of osteochondral samples imaged with contrast-enhanced micro-computed tomography. A dataset comprising 79 osteochondral cores from 24 total knee arthroplasty patients and two asymptomatic donors was utilized, with imaging conducted using contrast-enhanced micro-computed tomography. Volumes of interest (VOI) within surface, deep, and calcified zones were extracted depth-wise, and a dimensionally reduced Local Binary Pattern-textural feature analysis was applied. Researchers employed regularized linear and logistic regression models and validated the models using nested leave-one-out cross-validation. The models were assessed using mean squared error (MSE) and average precision metrics. An independent test set consisting of 4 mm samples was used for additional validation.

V. Pedroia et al.'s paper [9] aims to investigate the capacity of both conventional and deep learning-based T2 relaxometry patterns in distinguishing between knees afflicted with and without radiographic OA. The analysis involved T2 relaxation time maps from 4,384 subjects in the baseline Osteoarthritis Initiative Dataset. Voxel-Based Relaxometry (VBR) facilitated automatic quantification and voxel-based examination of T2 differences between subjects with and without radiographic OA.

A Densely Connected CNN (DenseNet) was trained for OA diagnosis using T2 data. To benchmark algorithmic performance, classical feature extraction techniques and shallow classifiers were employed. Evaluation of deep and shallow models, with and without the inclusion of risk factors, was conducted, utilizing Sensitivity and Specificity values along with the McNemar test to compare classifier performance.

Innovative algorithmic and deep learning models have shown promising potential in improving knee OA classification efficiency. A. Haseeb et al. [10] introduced a novel approach to classifying knee OA through the integration of deep learning and a whale optimization algorithm, aimed at addressing the time-consuming and costly manual categorization of knee joint disorders in X-ray imaging. Two pre-trained deep learning models, EfficientNet-b0 and DenseNet201, were utilized for training and feature extraction through deep transfer learning with fixed hyperparameter values. Fusion employing a canonical correlation approach resulted in a feature vector with enhanced information. An enhanced whale optimization algorithm was then employed for dimensionality reduction. The selected features underwent classification using ML algorithms, including fine-tuned support vector machines (SVMs) and neural networks. Experimental validation on a publicly available dataset yielded a maximum accuracy of 90.1 percent. The system's interpretability was enhanced through Explainable Artificial Intelligence techniques, specifically occlusion. Comparative analysis with recent research demonstrated a significant improvement in accuracy achieved by the proposed method.

Similarly, A. Rehman et al. [11] developed a model capable of effectively diagnosing early-stage OA in knee X-ray images. Utilizing advanced deep learning-based CNNs alongside various ML techniques, the research introduced a novel transfer learning-based feature engineering method known as CRK (CNN, Random Forest, KNN). This technique demonstrated high-performance OA detection. The CRK utilized a 2D-CNN to intelligently extract spatial features from X-ray images, which are then input to Random Forest and KNN techniques, generating a probabilistic feature set. This set was employed in building machine learning-based models. Experimental results indicated that the proposed model achieved a remarkable accuracy score of 99 percent in predicting OA. The performance of each model was rigorously validated through hyperparameter optimization and k-fold-based cross-validation. The study holds the potential to revolutionize OA prediction from X-ray images, showcasing high-performance scores.

In a comparative analysis, Teoh et al. [12] developed a multitask model utilizing CNN feature extractors and ML classifiers to identify nine crucial OA features. These features included the KL grade, knee osteophytes and joint-space narrowing, as well as patient-reported pain intensity from plain radiography. A novel feature extraction method was proposed, replacing the fully connected layer with a global average pooling (GAP) layer. The researchers conducted a comparative analysis to assess the effectiveness of 16 different CNN feature extractors and three ML classifiers. The most effective model incorporated the VGG16-GAP feature extractor and KNN classifier. This particular model not only surpassed the performance of the other models under examination, but also outperformed state-of-the-art methods by demonstrating superior balanced accuracy, higher Cohen's kappa, greater F1 scores, and lower Mean Squared Error when predicting seven OA features.

In their paper, researchers Du, Shan, and Zhang [13] discovered hidden biomedical insights from knee magnetic resonance imaging (MRIs)s to predict OA. The Cartilage Damage Index information was computed from 36 strategic locations in the tibiofemoral compartment through 3D MRI reconstruction. Principal component analysis was applied to process the feature set, which was then utilized along with the original raw feature set as inputs for four distinct ML methods (artificial neural network (ANN), SVM, Random Forest, and Gaussian Naïve Bayes). The most optimal performance was observed with ANN, achieving an area under the receiver operating characteristic (ROC) curve of 0.761 and an F-measure of 0.714. Experimental results highlighted that the informative locations on the medial tibiofemoral compartment held more valuable information than those on the lateral tibiofemoral compartment for predicting the severity of osteoarthritis.

G.B. Joseph et al. [14] aimed to formulate a machine learning-based predictive model for incident radiographic knee OA over an eight-year period, utilizing MRI-based cartilage biochemical composition, knee joint structure, demographics, and clinical predictors, including muscle strength and symptoms. Three distinct models were developed and compared: Model 1, incorporating 112 predictors based on OA risk factors; Model 2, featuring the top ten predictors determined by feature importance score from Model 1, alongside clinical relevance; and Model 3, representing Model 2 without the inclusion of imaging predictors. Model performance was evaluated using the area under the ROC curve derived from hold-out data. The ten-predictor model (Model 2), inclusive of cartilage and meniscus WORMS scores and cartilage T2, exhibited a slightly lower AUC compared to the model with 112 predictors but significantly outperformed the model without MRI predictors.

Further studies address the challenge of automatically detecting knee OA through the development of a computer system. The system developed in M. Kotti et al.'s study [15] takes body kinetics as input and provides an output that not only estimates the presence of knee OA as seen in previous literature, but also identifies the most discriminative parameters and outlines a set of rules guiding the decision-making process. This approach bridges the interpretability gap between medical and engineering perspectives. Participants walked on a walkway equipped with two force plates containing piezoelectric three-component force sensors. Parameters related to vertical, anterior-posterior, and mediolateral ground reaction force—such as mean value, push-off time, and slope—were extracted. Random Forest regressors then mapped these parameters via rule induction to determine the degree of knee OA. To enhance generalization ability, a subject-independent protocol was implemented.

A different deep learning model, Deep-KOA, developed by D.N.A. Ningrum et al. [16] aimed to predict the risk of knee OA over the next year using non-image-based electronic medical record data from the previous three years. A feature matrix was constructed based on the three-year sequential history of diagnoses, drug prescriptions, age, and sex. The risk prediction model was developed using a combination of CNN and ANN deep learning methods. Performance evaluation metrics, such as the area under the ROC, sensitivity, specificity, and precision, were employed to assess the efficacy of Deep-KOA. Furthermore, the study explored important features through stepwise feature selection.

L.S. Lee et al. [17] provide a comprehensive overview of the current evidence and recent advancements in the application of AI for diagnosing knee OA and predicting outcomes of total knee arthroplasty. The researchers conducted searches in the PubMed and EMBASE databases for articles published in peer-reviewed journals from January 1, 2010, to May 31, 2021. Exclusions comprised non-English language articles and those lacking English translation. A designated reviewer assessed the articles for relevance to the research questions and the strength of evidence.

In their paper, A.J. Prabhakar et al. [18] aimed to introduce a novel algorithm designed to classify the presence and severity of knee OA using parameters derived from a force plate. Forty-four sway movement graphs were measured for analysis. Various ML algorithms, including K-Nearest Neighbours (KNN), Logistic Regression, Gaussian Naïve Bayes, Support Vector Machine (SVM), Decision Tree Classifier, and the Random Forest (RF) Classifier, were applied to the dataset. The proposed method demonstrated a commendable 91 percent accuracy in detecting sway variation, providing valuable support for rehabilitation specialists in objectively identifying a patient's condition at an early stage and facilitating patient education regarding disease progression.

Returning to predictive OA research, A. Tiulpin et al. introduced [19] a multi-modal machine learning-based model for predicting knee OA progression. This model incorporated raw radiographic data, clinical examination results, and the patient's previous medical history. The validation of this approach was carried out on an independent test set comprising 3,918 knee images from 2,129 subjects. The method demonstrated notable performance with an area under the ROC curve (AUC) of 0.79 and an average precision (AP) of 0.68. In comparison, a reference approach relying on logistic regression produced an AUC of 0.75 and an AP of 0.62. This proposed method holds the potential to significantly enhance the subject selection process for OA drug development trials and contribute to the formulation of personalized therapeutic plans.

In a review of the field of knee OA, P.S.Q. Yeoh et al. [20] offer a broad overview of current two-dimensional and three-dimensional CNN approaches. A total of 74 relevant studies about the classification and segmentation of knee OA were reviewed and extracted from the Web of Science database. The review delves into discussions on various state-of-the-art deep learning approaches proposed in this context, with a particular emphasis on highlighting the potential feasibility of employing three-dimensional CNNs in the field of knee OA. The authors conclude by addressing potential challenges and envisioning advancements associated with the adoption of three-dimensional CNNs in this particular research domain.

In their paper, C. Ntakolia et al. [21] use an ML pipeline for predicting knee joint space narrowing (JSN) in patients with knee OA utilizing multi-disciplinary data from the OAI database, the proposed methodology encompasses several key components. Firstly, a clustering process is employed to identify distinct groups of individuals with progressing and non-progressing OA. Subsequently, a robust feature selection process is implemented, comprising filter, wrapper, and embedded techniques, aimed at identifying the most informative risk factors contributing to JSN prediction. Lastly, a decision-making process evaluates and compares various classification algorithms to select and develop the final prediction model for JSN. Evaluation criteria included the model's overall performance, robustness, and highest achieved accuracy. Notably, Logistic Regression achieved accuracies of 78.3 percent and 77.7 percent for the left and right leg, respectively, based on a group of 164 risk factors. Additionally, SVM attained accuracies of 78.3 percent and 77.7 percent for the left and right leg, respectively, utilizing a reduced set of 88 and 90 risk factors.

The KL grading scale is a highly useful measure of analysis in knee OA severity research. In one study, utilizing deep CNNs alongside the KL grading system, researchers Ganesh Kumar and Goswami [22] evaluated the severity of knee OA using a form of image filtering. Image sharpening enhances image clarity and addresses noise in knee X-ray images. The evaluation used baseline X-ray images from the OAI. Upon analyzing enhanced images generated through the image sharpening process, a mean accuracy of 91.03 percent was achieved. This marks a significant improvement of 19.03 percent over the previous accuracy rate of 72 percent obtained using original knee X-ray images for OA detection with five gradings. The utilization of image sharpening techniques serves to enhance knee joint recognition and facilitate knee KL grading.

V.V. Kishore et al. [23] compared the accuracy of models to identify the most effective model for detecting knee OA. In their paper, knee OA served as a clinical scenario to assess twelve transfer learning deep learning models for their ability to detect the grade of knee OA from radiographs. The evaluated models demonstrated a wide range of accuracies, varying from 30 percent to 98 percent in detecting knee osteoarthritis. Notably, MobileNet exhibited the highest accuracy, achieving 98.36 percent. This model showed high levels of training and validation accuracy. Conversely, EfficientNetB7 displayed the maximum loss among the evaluated models. The study suggests that deep learning approaches developed by experienced radiologists and orthopedic specialists could benefit smaller hospitals, enabling them to enhance their learning capabilities and provide improved emergency room services, particularly in scenarios where medical personnel may be scarce.

Recent studies demonstrate the potential of machine learning (ML) in automating knee OA diagnosis and prognosis through knee joint localization, classification of OA severity, and prediction of disease progression. Y.X. Teoh et al. [24] provide an overview of knee OA imaging features across various practices for traditional OA diagnosis, alongside recent advancements in image-based ML approaches for knee OA diagnosis and prognosis. While X-ray remains the standard imaging modality for knee OA diagnosis, it is limited in its ability to capture short-term OA changes, prompting recommendations for MRI exploration to uncover hidden OA-related radiomic features. Additionally, ultrasound imaging features should be further explored for improved point-of-care diagnosis. Traditional knee OA diagnosis, reliant on manual interpretation of medical images using the

Kellgren-Lawrence (KL) grading scheme, faces challenges of human resources and time constraints, hindering effective OA prevention. AI-aided diagnostic models significantly enhance the efficiency, reproducibility, and accuracy of knee OA diagnosis. Prognostic capabilities have been exhibited by several prediction models in estimating OA onset, deterioration, pain progression, structural changes, and time to total knee replacement incidence.

Despite existing research gaps, ML techniques hold promise for addressing demanding tasks, such as early knee OA detection and forecasting future disease events, as well as fundamental objectives like discovering new imaging features and establishing novel OA status measures. Continuous enhancement of ML models may lead to the discovery of innovative OA treatments in the future. Cigdem and Deniz's literature review [25] provides a comprehensive overview of the current evidence and recent applications of AIin knee OA detection. The review focuses on articles that examine AI's role in diagnosing and predicting the prognosis of knee OA and in accelerating image acquisition. Each selected study was analyzed for various factors, including code availability, patient and knee count, imaging modalities, covariates, OA grading methods, utilized models, validation techniques, objectives, and outcomes.

A.E. Nelson et al. [26] used innovative ML approaches for knee OA phenotyping in order to identify progression phenotypes potentially more responsive to interventions. Publicly available data from the Foundation for the National Institutes of Health's Osteoarthritis Biomarkers Consortium were utilized, wherein radiographic and pain progression over 48 months were categorized into four mutually exclusive outcome groups (none, both, pain only, radiographic only), alongside an extensive set of covariates. Distance weighted discrimination, direction-projection-permutation testing, and clustering methods were employed to focus on the contrast (z-scores) between individuals progressing by both criteria ("progressors") and those not progressing by either ("non-progressors").

A separate investigation aimed to determine if a significant loss of surface stiffness could be detected in early OA superficial zone chondrocyte spatial organizations (SCSO) stages. This study [27] examined the relationship between two sensitive early osteoarthritis (OA) markers: atomic force microscopy -measured human articular cartilage surface stiffness and location-matched SCSOs. Subsequently, the study assessed whether current clinical technology could visualize and accurately diagnose the SCSOs using an approved probe-based confocal laser-endomicroscopic system and a RF model.

Another study found that AI-aided diagnostic ratings exhibited a stronger association with the overall KL score and the Knee Injury and Osteoarthritis Outcome Score (KOOS). In this study [28], M. Neubauer et al. used seventy-one DICOMs for AI analysis using KOALA software in a study involving subjects recruited from a physiotherapy trial (MLKOA). At baseline, each subject received a knee X-ray along with an assessment using five main scores: Tegner Scale, KOOS, International Physical Activity Questionnaire, Star Excursion Balance Test, and a Six-Minute Walk Test. Clinical assessments were repeated at weeks 6, 12, and 24. Three physicians analyzed the presented X-rays both with and without AI assistance using KL grading. Interrater reliability (IRR) analyses and Spearman's Correlation Test were conducted to assess the overall KL score for each individual rater with clinical scores. The extent of improvement due to AI varied among individual raters.

In their study, J.J. Lee et al. [29] aimed to delineate distinctive pain trajectories among knee OA patients and also to explore the correlation between MRI biomarkers acquired through a three-dimensional CNN and the identified pain trajectories. Data encompassing repeated measures of KOOS pain scores for both knees over a decade-long period were sourced from the OAI, totaling 4,796 subjects. Three-dimensional Double Echo Steady State images of the knee at baseline were employed for image biomarker discovery.

To mitigate inherent noise and address missing data, individual pain curves were temporally smoothed using a regression model fitted with orthogonal polynomials of degrees one and two. Standardized estimated parameters were utilized as input for clustering analysis via a Gaussian Naïve Bayes mixture model, with the optimal model selected using the Silhouette approach to effectively capture distinct pain patterns. The study's CNN architecture is a three-dimensional extension of the DenseNet121, with weights initialized using the He initialization. The network was trained to ascertain posterior probabilities of pain trajectory membership, employing the mean squared error function as the loss function for regression.

N. Pongsakonpruttikul et al. [30] employed and evaluated the performance of a CNN aimed at aiding orthopedists and radiologists in detecting and categorizing knee OA across various degrees, following the KL classification system. A dataset comprising 1,650 knee joint radiographs (anteroposterior view) was amassed from the Osteoarthritis Initiative's (OAI) public resources. Two models were constructed: one for distinguishing normal (KL 0-I) from osteoarthritic knees (KL II-IV); The other for classifying severity as normal (KL 0-I), non-severe (KL II), or severe (KL III-IV). Expert supervision was utilized for labeling regions of interest.

The AI models were trained utilizing the You Only Look Once version (YOLOv3) detection algorithm. The initial AI model employing YOLOv3 demonstrated an 85 percent accuracy and 81 percent mean average precision in detecting and classifying normal and osteoarthritic knees on plain knee joint radiographs. The subsequent AI model, focusing on severity classification, achieved an overall accuracy of 86.7 percent and mean average precision of 70.6 percent.

It is also useful to look at the use of AI in OA diagnosis in other joints. One paper [31] aimed to develop and evaluate an AI model for diagnosing Temporomandibular joint (TMJ) osteoarthritis from cone-beam computed tomography (CBCT). W.M. Talaat et al. used a dataset comprising 2,737 CBCT images from 943 patients for training and validation. The model, employing a single CNN and a single regression model for object detection, underwent assessment against a model-testing set of 350 images, established by two experienced evaluators using the Diagnostic Criteria for Temporomandibular Disorders. The diagnosis concluded by the evaluators served as the golden reference for comparison. The diagnostic performance of the AI model was then juxtaposed with that of an experienced oral radiologist, revealing statistically higher agreement between the AI diagnosis and the golden reference compared to the radiologist.

Another paper centered on TMJ [34] aimed to develop a diagnostic support tool utilizing pretrained models to classify panoramic images of the TMJ into normal and OA cases. The dataset was randomly divided into training, validation, and testing subsets. Pretrained ResNet152 and EfficientNet-B7 models were employed as transfer learning models. The accuracy, specificity, sensitivity, area under the curve, and gradient-weighted class activation mapping of both trained models were assessed. The performance of the trained models was compared to that of dentists. The classification accuracies of ResNet-152 and EfficientNet-B7 were 0.87 and 0.88, respectively, with the trained models demonstrating the highest accuracy in OA classification.

Returning to knee OA, M.W. Brejnebøl et al. [32] sought to validate an AI tool for classifying radiographic knee OA severity using weight-bearing, non-fixed-flexion posterior/anterior knee radiographs from a clinical production PACS system. The index test involved ordinal KL grading by an AI tool, two musculoskeletal radiology consultants, two reporting technologists, and two resident radiologists. All readers repeated grading after at least four weeks. The reference test was the consensus of the two consultants. The primary outcome measure was quadratic weighted kappa, with secondary outcomes including ordinal weighted accuracy, multiclass accuracy, and F1-score.

In another study, S. Olsson et al. [33] aimed to assess the effectiveness of an AI system in classifying the severity of knee OA using entire image series, without excluding common visual disturbances such as implants, casts, and non-degenerative pathologies. They gathered 6,103 radiographic knee exams conducted at Danderyd University Hospital between 2002 and 2016 and manually categorized them based on the KL grading scale. Subsequently, a CNN of ResNet architecture was trained using PyTorch. The study's outcomes were evaluated against a test set comprising 300 exams, independently reviewed by two senior orthopedic surgeons who resolved any inter-observer disagreements through consensus sessions.

Seeking to address a topic not previously explored in knee OA research, Sozan Ahmed and Ramadhan Mstafa [35] developed a novel method to efficiently diagnose and classify knee OA severity based on X-ray images, addressing the issue of classifying knee OA in binary and multiclass categories. The proposed models were categorized into two frameworks: one utilizing pre-trained convolutional neural networks (CNNs) for feature extraction and the other involving fine-tuning of pre-trained CNNs using transfer learning (TL) techniques. Additionally, traditional ML) classifiers were employed to leverage the enriched feature space for improved knee OA classification. The first framework comprised five class-based models employing a proposed pre-trained CNN for feature extraction, principal component analysis for dimensionality reduction, and SVM for classification. In the second framework, adjustments were made to the steps of the first framework, incorporating TL to calibrate the pre-trained CNNs for two-, three-, and four-class-based models.

In their paper, H. Bonakdari et al. [36] constructed a comprehensive ML model aimed at linking major OA risk factors with serum levels of adipokines and related inflammatory factors at baseline to predict early progression of knee OA among at-risk patients. A total of 677 subjects were selected from the OAI progression sub-cohort for the study. Probability values indicating the likelihood of being structural progressors were generated using a previously developed prediction model, which incorporated five baseline structural knee features, including two X-rays and three magnetic resonance imaging variables. To identify the most influential variables among the 47 studied in relation to progressive knee OA, the researchers utilized ML feature classification methodology. Among the five supervised ML algorithms tested, SVM exhibited the highest accuracy and was chosen for gender-based classifier development. Feature selection analysis revealed that age, BMI, and the ratios of CRP/MCP-1 and leptin/CRP were the most crucial variables for predicting OA structural progressors in both genders.

A paper by S.B. Kwon et al. [37] used patient-reported outcome measures to identify gait features related to knee OA and develop regression models using ML algorithms to estimate knee OA severity. Knee OA severity was assessed using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Linear regression models and a Random Forest regression were employed to estimate WOMAC scores for patients. A range of features were selected from each joint, including 12 from the hip, 1 from the pelvis, 17 from the knee, 9 from the ankle, 1 from the foot, and 3 from spatiotemporal parameters. Gait analysis features were utilized to predict knee OA severity, contributing to the development of an objective estimation model for knee OA severity.

A.S. Mohammed et al. [38] conducted binary classification to detect the presence or absence of knee OA and classify the severity of knee OA into three classes. This paper proposed the application of six pre-trained deep neural network (DNN) models, including VGG16, VGG19, ResNet101, MobileNetV2, InceptionResNetV2, and DenseNet121, for knee OA diagnosis using images sourced from the OAI dataset. For comparative analysis, experiments were performed on three datasets—Dataset I, Dataset II, and Dataset III—with five, two, and three classes of knee OA images. The ResNet101 DNN model achieves maximum classification accuracies of 69 percent, 83 percent, and 89 percent for Dataset I, Dataset II, and Dataset III, respectively. The results demonstrate improved performance compared to existing literature.

Continuing research into computer-aided diagnosis (CAD), A. Tiulpin et al. [39] present a novel transparent CAD method based on the Deep Siamese CNN to automatically score knee OA severity according to the KL grading scale. The method is trained using data exclusively from the Multicenter Osteoarthritis Study and validated on a randomly selected subset of 3,000 subjects (5,960 knees) from the OAI dataset. The proposed method achieves a quadratic Kappa coefficient of 0.83 and an average multiclass accuracy of 66.71 percent compared to annotations provided by a committee of clinical experts. Additionally, a radiological OA diagnosis area under the ROC curve of 0.93 is reported.

S. Abd El-Ghany, M. Elmogy, and A.A. Abd El-Aziz's [40] model aims to classify the severity of knee OA diseases through multi-classification and binary classification. Their paper proposed a fine-tuned knee OA diagnosis model utilizing the DenseNet169 deep learning) technique to enhance the efficiency of knee OA diagnosis. The model was designed to accurately localize peripheral opacities, diffuse distribution, and vascular thickening, thereby assisting clinicians in understanding the primary causes of knee OA. Pre-processing of the OAI dataset involved artifact removal, resizing, contrast handling, and normalization techniques. The proposed model underwent evaluation and comparison with recent classifiers.

In multi-classification, the DenseNet169 model demonstrated performance metrics of 95.93 percent accuracy, 88.77 percent sensitivity, 95.41 percent specificity, 85.8 percent precision, and 87.08 percent F1-score.

3. Proposed Methodology

In the initial phase of data collection and preprocessing, a diverse and well-annotated dataset of knee X-ray images, labeled with OA severity grades, was carefully curated. To mitigate potential biases, particular attention was given to ensuring a balanced distribution across all severity levels. The dataset was then partitioned into training, validation, and test sets while maintaining representative distributions in each subset. During preprocessing, the images were standardized and resized to ensure a uniform format suitable for model input. To improve the model's generalization capabilities, augmentation techniques, such as rotation, flipping, and scaling, were applied. Additionally, pixel values were normalized to a specific range to support efficient model convergence during training. This comprehensive and systematic approach to data collection and preprocessing provides a solid foundation for the subsequent phases of model development and evaluation.

The primary deep learning architecture chosen for model selection is the pre-trained EfficientNet B5 model, tailored explicitly for X-ray detection. The model parameters were adapted to the distinctive features of X-ray images. Leveraging transfer learning, the pre-trained EfficientNet B5 will capitalize on the knowledge gained from general image datasets, enhancing its ability to discern nuanced patterns in knee X-rays. Additionally, alternative ML models, including Random Forest and KNN, chosen for their simplicity and interpretability, are utilized for benchmarking purposes. EfficientNet V2B3, a variant of the primary model, will also be deployed to compare its performance. A primary CNN (Conv2D) with MaxPooling2D layers is used to diversify the comparison, offering insights into the trade-offs between model complexity and performance. This comprehensive model selection strategy aims to evaluate the efficacy of the chosen deep learning architecture against various benchmarks, providing a thorough understanding of the model's strengths and potential improvements.

In the model training phase, the pre-trained EfficientNet B5 underwent fine-tuning on the designated knee OA training dataset using categorical cross-entropy as the loss function for multi-class classification. Early stopping was implemented to mitigate overfitting risks. An overall view of the proposed work is given in Figure 2.

Random Forest and KNN were trained on flattened X-ray images, utilizing cross-validation for robust model evaluation. EfficientNet V2B3 and Conv2D models were trained with configurations similar to EfficientNet B5. In the model evaluation phase, a comprehensive set of evaluation metrics, was employed to assess the performance of the trained models on the

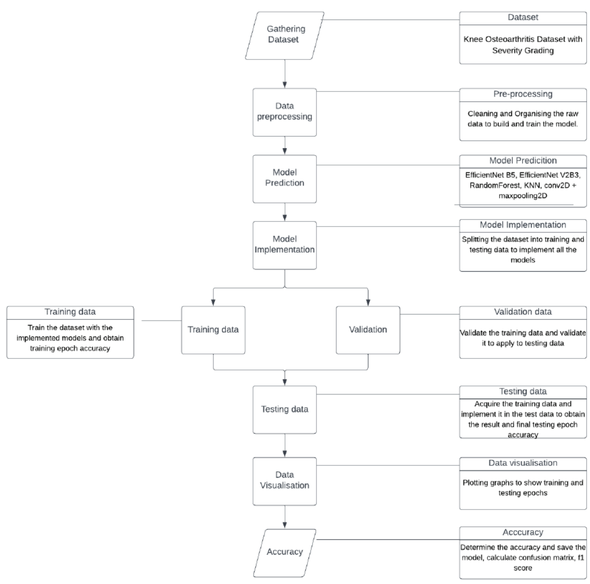
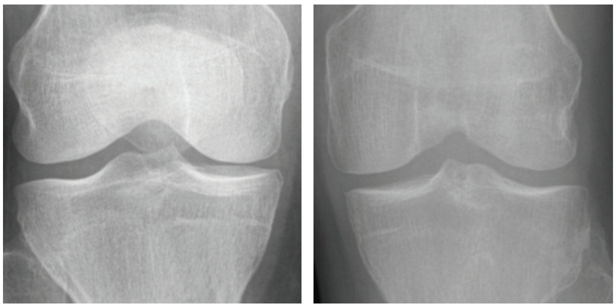
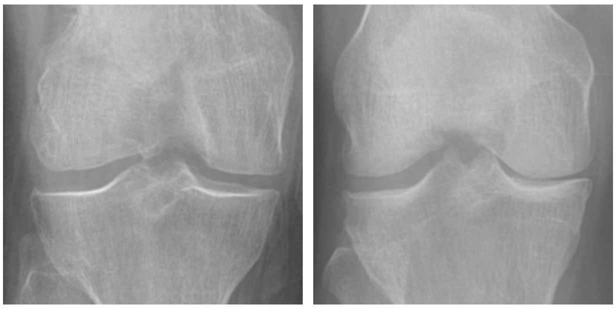


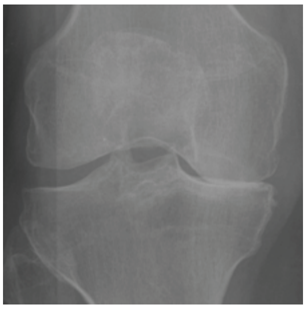
Figure 2. Flow chart detailing the data training and validation stages of the methodology



Grade 0 (Healthy) and Grade 1 (Doubtful):



Grade 2 (Minimal) and Grade 3 (Moderate



Grade 4 (Severe)

Figure 3. Sample images of each class of knee OA severity

validation set, including accuracy, precision, recall, F1-score, and confusion matrices. The final evaluation was based on the independent test set, providing a robust assessment of the model's effectiveness in accurately classifying and grading knee OA severity. The evaluation initially took place on the validation set, allowing for the fine-tuning of hyperparameters and guarding against overfitting.

A diverse dataset comprising knee X-ray images with varying degrees of OA severity was meticulously selected to achieve comprehensive evaluations. The dataset was annotated and graded by medical professionals to ensure accuracy. The dataset is already a benchmark dataset, which can be found in: *Chen, Pingjun (2018), "Knee Osteoarthritis Severity Grading Dataset", Mendeley Data, V1, doi: 10.17632/56rmx-5bjcr.1.* Sample data from each class of the dataset is given in Figure 3.

Standard preprocessing techniques, including resolution standardization and image augmentation, were applied to enhance the robustness of the models. The choice of machine learning models was driven by a desire to explore diverse architectures. EfficientNet B5, renowned for its success in image classification tasks, was selected for its potential in medical image analysis. EfficientNet B5 is a member of the EfficientNet family that uses a compound scaling method to grow the model's depth, width, and resolution. It has been described as a new breed of CNNs that, with fewer parameters, are able to achieve very high accuracy.

The Random Forest model, known for its interpretability, was chosen to assess its capability to distinguish severity levels. The KNN algorithm, valued for its simplicity, was included as a viable option for knee OA detection. Additionally, the EfficientNet V2B3 model and traditional Conv2D and MaxPooling2D architectures were employed to understand model versatility comprehensively.

Highlights of Model Architecture:

- MBConv Blocks: Mobile Inverted Bottleneck Convolution (MBConv) blocks form the core part of the architecture;
- Activation Function: Swish (smooth; non-monotonic);
- Squeeze-and-Excitation (SE): Applied in every block to enhance the recalibration of features;
- Input resolution: 456×456 pixels;
- Total layers: ~570 layers (not all of them are convolution; includes batch norm, pooling,etc.);
- Number of Params: ~30 million (29.34M); and
- Number of MBConv Blocks: 39 MBConv blocks with differing kernel sizes (3×3 or 5×5) and expansion ratios.

Fine-tuning Procedure:

- Pretrained on: ImageNet;
- Layers Frozen: During early phases of training, the initial layers were frozen;
- Unfrozen Layers: The last 3 blocks were unfrozen to be fine-tuned for the purpose of specific tasks;
- Optimizer: Adam, LR=1e-4, ReduceLROnPlateau;
- Batch size: 16;
- Epochs: 50 with early stopping.

A comparative analysis with well-known architecture is given in Table 1 for better justification.

4. Results and Analysis

The application of ML models for the detection of knee OA using X-ray images yielded compelling outcomes. Notably, the EfficientNet B5 model demonstrated exceptional performance, achieving an impressive accuracy of 95.84 percent. This result highlights the efficacy of leveraging state-of-the-art neural network architectures for medical image analysis. The EfficientNet B5 model's ability to discern intricate features within X-ray images contributed significantly to its superior performance in accurately grading the severity of knee OA. Sample images and a predicted image sample are given in Figure 4 and Figure 5, respectively.

Table 1. Comparative analysis of models

Model	Parameters	Input Size	Conv Layers	Key Features
EfficientNet-B5	~29.3M	456×456	MBConv (39)	Compound scaling, SE blocks, Swish
ResNet50	~25.6M	224×224	50	Residual blocks, skip connections
VGG16	~138M	224×224	13	Deep stack, large fully connected head
MobileNetV2	~3.4M	224×224	~53	Lightweight, depthwise separable conv

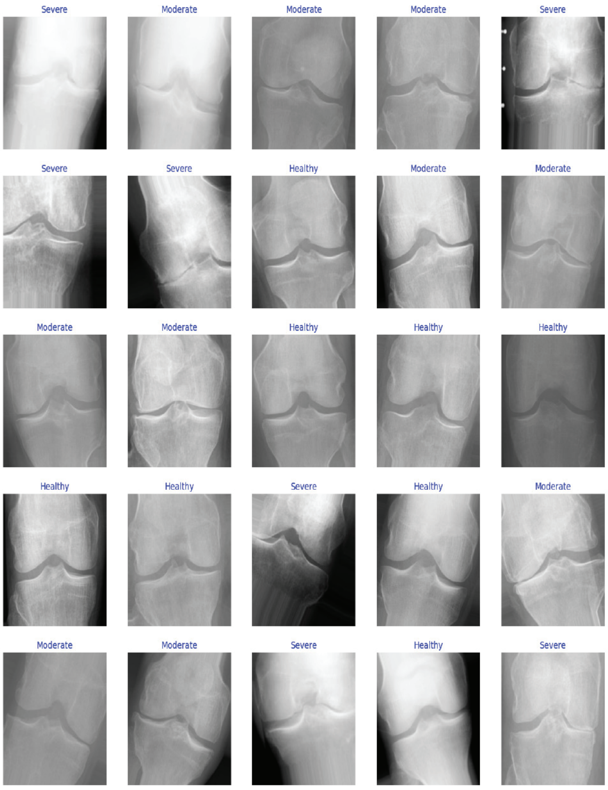


Figure 4. Sample Images

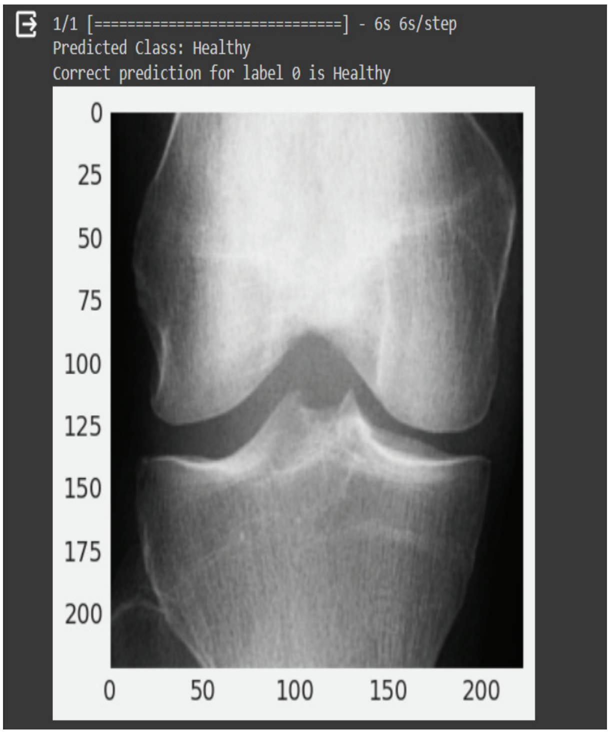


Figure 5. Predicted Image

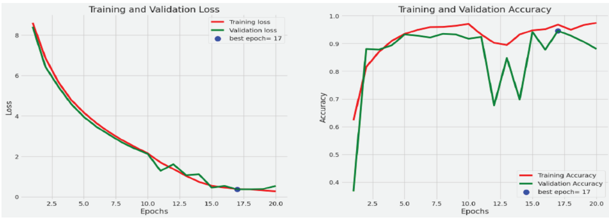


Figure 6. Output Graph

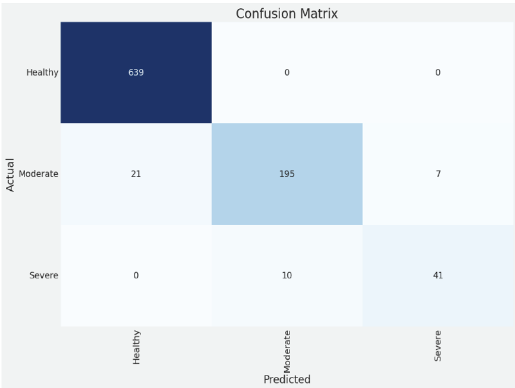


Figure 7. Confusion Matrix Efficient Net B5

The impressive outcomes from applying ML models for knee OA detection using X-ray images underscore the significance of advanced neural network architectures. The model's ability to discern intricate features within X-ray images was crucial in achieving superior performance, showcasing its potential for accurate severity grading in knee osteoarthritis cases. The generated graph is given in Figure 6. Figures 7, 8, and 9 show the confusion matrices of the algorithms under comparison.

Comparatively, the Random Forest model exhibited a moderate accuracy of 45 percent, indicating a reasonable capability to distinguish between different severity levels. While not as high as the EfficientNet B5, this accuracy suggests a pragmatic performance level for severity classification. The KNN algorithm achieved an accuracy of 41 percent, highlighting its potential as a viable option for knee OA detection, albeit with room for improvement.

The EfficientNet V2B3 model demonstrated competitive performance with an accuracy of 65.78 percent, further emphasizing the versatility of different EfficientNet variants in handling medical image classification tasks. This result positions EfficientNet as a robust choice for knee OA diagnosis. On the other hand, the Conv2D and MaxPooling2D models, repre-

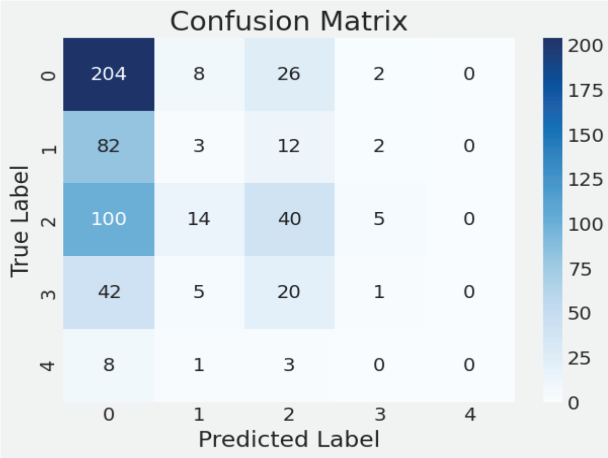


Figure 8. Confusion Matrix KNN

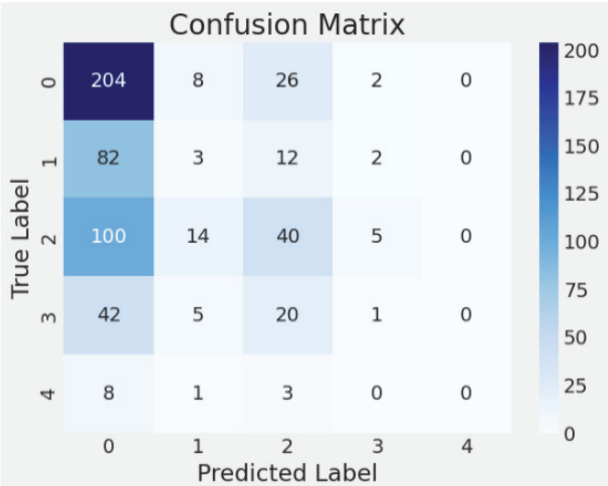


Figure 9. Confusion Matrix Random Forest

senting traditional CNN architectures, achieved an accuracy of 39.56 percent, providing insights into the effectiveness of these established approaches. The discussions center around the remarkable accuracy achieved by the EfficientNet B5 model, showcasing its potential for accurate knee OA severity grading. The moderate accuracies of Random Forest, KNN, EfficientNet V2B3, Conv2D, and MaxPooling2D models offer a nuanced understanding of their capabilities. Future considerations may involve refining model parameters, exploring ensemble approaches, and addressing potential limitations to further enhance the practical implementation of these models in clinical settings.

5. Conclusion and Future Work

Our research contributes to the scientific discourse surrounding knee OA detection and underscores the transformative potential of ML in revolutionizing healthcare practices. The journey towards a more technologically advanced and patient-centric healthcare system is underway, and the successful fusion of ML with medical imaging marks a significant step forward in this transformative process. Future research should focus on developing guidelines and frameworks that facilitate seamless integration, addressing privacy, security, and interoperability concerns within healthcare systems. Furthermore, continuous exploration and incorporation of the latest advancements in ML and medical imaging are essential to staying at the forefront of technological progress, ensuring sustained improvement in the performance and efficiency of knee OA detection models.

AUTHORS

R. Gokulapriya – Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST (Deemed to be University), Bangalore Kengeri Campus, Bangalore, India, email: r.gokulapriya@christuniversity.in.

M. Balamurugan – Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST (Deemed to be University), Bangalore Kengeri Campus, Bangalore, India, email: balamurugan.m@christuniversity.in.

Rakoth Kandan Sambandam – Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST (Deemed to be University), Bangalore Kengeri Campus, Bangalore, India, email: rakoth.kandan@christuniversity.in.

Divya Vetriveeran* – Department of Computer Science and Engineering, University College of Engineering Thirukkuvalai, Anna University, Thirukkuvalai, Tamil Nadu, India, email: vdivya.3793@gmail.com

*Corresponding author

References

[1] J.Y. Jang et al., "Study of the Efficacy of Artificial Intelligence Algorithm- Based Analysis of the Functional and Anatomical Improvement in Polynucleotide Treatment in Knee Osteoarthritis Patients: A Prospective Case Series," *Journal of Clinical Medicine,* 2022; doi:10.3390/jcm11102845.

[2] S. Mohammadi et al., "Artificial intelligence in osteoarthritis detection: A systematic review and meta-analysis," *Osteoarthritis and Cartilage*, 2023, pp. 241-253.

[3] H.A. Alshamrani, M. Rashid, S.S. Alshamrani, A.H.D. Alshehri, "Osteo-NeT: An Automated System for Predicting Knee Osteoarthritis from X-ray Images Using Transfer-Learning-Based Neural Networks Approach," *Healthcare*, 2023; doi:10.3390/healthcare11091206.

[4] P. Chen et al., "Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss," *Computerized Medical Imaging and Graphics,* vol 75, 2019, pp. 84-92.

[5] B. Abdelbasset et al., "A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the Osteoarthritis Initiative," *Computerized Medical Imaging and Graphics*, vol. 73, 2019, pp. 11-18.

[6] J. Song and R. Zhang, "A novel computer-assisted diagnosis method of knee osteoarthritis based on multivariate information and deep learning model," *Digital Signal Processing*, 2023; doi:10.1016/j.dsp.2022.103863.

[7] N. Lazzarini et al., "A machine learning approach for the identification of new biomarkers for knee osteoarthritis development in overweight and obese women," *Osteoarthritis and Cartilage* 2014-2021 (epub 2017); doi:10.1016/j.joca.2017.09.001.

[8] S.J.O. Rytky et al., "Automating three-dimensional osteoarthritis histopathological grading of human osteochondral tissue using machine learning on contrast-enhanced micro-computed tomography," *Osteoarthritis and Cartilage* vol. 28, no. 8, 2020, pp. 1133-1144.

[9] V. Pedroia et al., "Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort," *Osteoarthritis and Cartilage* vol. 27, no. 7, 2019, pp. 1002-1010.

[10] A. Haseeb, et al., "Knee osteoarthritis classification using x-ray images based on optimal deep neural network," *Computer Systems Science and Engineering*, vol. 47, no. 2, 2023, pp. 2397–2415.

[11] A. Rehman et al., "Transfer Learning- Based Smart Features Engineering for Osteoarthritis Diagnosis From Knee X-Ray Images," *IEEE Access*, vol. 11, 2023, pp. 71,326-71,338.

[12] Y.X. Teoh et al. "Stratifying knee osteoarthritis features through multi-task deep hybrid learning: Data from the osteoarthritis initiative," *Computer Methods and Programs in Biomedicine*, 2023; doi:10.1016/j.cmpb.2023.107807.

[13] Y. Du, J. Shan, and M. Zhang, "Knee osteoarthritis prediction on MR images using cartilage damage index and machine learning methods," *2017 IEEE International Conference on Bioinformatics and Biomedicine* (BIBM), 2017, pp. 671-677; doi: 10.1109/BIBM.2017.8217734.

[14] G.B. Joseph et al., "Machine learning to predict incident radiographic knee osteoarthritis over 8 Years using combined MR imaging features, demographics, and clinical factors: data from the Osteoarthritis Initiative," *Osteoarthritis and Cartilage,* vol. 30, no. 2, 2022, pp. 270-279.

[15] M. Kotti et al., "Detecting knee osteoarthritis and its discriminating parameters using random forests." *Medical Engineering & Physics*, vol. 43, 2017 pp. 19-29.

[16] D.N.A. Ningrum et al., "A Deep Learning Model to Predict Knee Osteoarthritis Based on Nonimage Longitudinal Medical Record," *Journal of Multidisciplinary Healthcare*, vol. 14, 2021, pp. 2477-2485.

[17] L.S. Lee et al., "Artificial intelligence in diagnosis of knee osteoarthritis and prediction of arthroplasty outcomes: a review," *Arthroplasty,* 5 Mar. 2022; doi:10.1186/s42836-022-00118-7.

[18] A.J. Prabhakar et al.,"Use of Machine Learning for Early Detection of Knee Osteoarthritis and Quantifying Effectiveness of Treatment Using Force Platform," *Journal of Sensor and Actuator Networks*, 2022; doi:10.3390/jsan11030048

[19] A. Tiulpin et al., "Multimodal Machine Learning-based Knee Osteoarthritis Progression Prediction from Plain Radio-graphs and Clinical Data," *Sci Rep,* 2019; doi:10.1038/s41598-019-56527-3.

[20] P.S.Q. Yeoh et al., "Emergence of Deep Learning in Knee Osteoarthritis Diagnosis", *Computational Intelligence and Neuroscience*, 2021; doi:10.1155/2021/4931437.

[21] C. Ntakolia et al., "A machine learning pipeline for predicting joint space narrowing in knee osteoarthritis patients," *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering* (BIBE), 2020, pp. 934-941.

[22] M. Ganesh Kumar and A.D. Goswami, "Automatic Classification of the Severity of Knee Osteoarthritis Using Enhanced Image Sharpening and CNN," *Applied Sciences,* , 2023; doi:10.3390/app13031658.

[23] V.K. V, V. Kalpana, G.H. Kumar, "Evaluating the efficiency of deep learning models for knee osteoarthritis prediction based on Kellgren-Lawrence grading system," *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, , 2023; doi:10.1016/j.prime.2023.100266.

[24] Y.X. Teoh et al., "Discovering Knee Osteoarthritis Imaging Features for Diagnosis and Prognosis: Review of Manual Imaging Grading and Machine Learning Approaches," *Journal of Healthcare Engineering,* 2022; doi:10.1155/2022/4138666.

[25] O. Cigdem and C.M. Deniz, "Artificial intelligence in knee osteoarthritis: A comprehensive review for 2022," *Osteoarthritis Imaging*, 2023; doi:10.1016/j.ostima.2023.100161.

[26] A.E. Nelson et al., "A machine learning approach to knee osteoarthritis phenotyping: data from the FNIH Biomarkers Consortium," *Osteoarthritis and Cartilage*, vol. 27, no. 7, 2019, pp. 994-1001.

[27] M. Tschaikowsky et al., "Proof-of-concept for the detection of early osteoarthritis pathology by clinically applicable endomicroscopy and quantitative AI-supported optical biopsy," *Osteoarthritis and Cartilage*, vol 29, no. 2, 2021, pp. 269-279.

[28] M. Neubauer M et al., "Artificial-Intelligence-Aided Radiographic Diagnostic of Knee Osteoarthritis Leads to a Higher Association of Clinical Findings with Diagnostic Ratings," *Journal of Clinical Medicine*, 2023; doi:10.3390/jcm12030744.

[29] J.J. Lee et al., "Can AI Predict Pain Progression in Knee Osteoarthritis Subjects From Structural

MRI," *Osteoarthritis and Cartilage*, April 2019;doi:10.1016/j.joca.2019.02.036.

[30] N. Pongsakonpruttikul et al., "Artificial intelligence assistance in radiographic detection and classification of knee osteoarthritis and its severity: a cross-sectional diagnostic study," *European Review for Medical and Pharmacological Sciences*, vol. 26, no. 5, 2022, pp. 1549-1558.

[31] W.M. Talaat et al., "An artificial intelligence model for the radiographic diagnosis of osteoarthritis of the temporomandibular joint," *Scientific Reports,* 2023;doi:10.1038/s41598-023-43277-6.

[32] M.W. Brejnebøl et al., "External validation of an artificial intelligence tool for radiographic knee osteoarthritis severity classification," *European Journal of Radiology*, 2022; doi:10.1016/j.ejrad.2022.110249.

[33] S. Olsson, et al., "Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population," *BMC Musculoskeletal Disorders*, 2021; doi:10.1186/s12891-021-04722-7.

[34] W. Jung et al., "Deep learning for osteoarthritis classification in temporomandibular joint," *Oral Diseases,* vol. 29, no. 3, 2023, pp. 1050-1059.

[35] S.M. Ahmed and R.J. Mstafa, "Identifying Severity Grading of Knee Osteoarthritis from X-ray Images Using an Efficient Mixture of Deep Learning and Machine Learning Models," *Diagnostics (Basel),* , 2022; doi:10.3390/diagnostics12122939.

[36] H. Bonakdari et al., "A warning machine learning algorithm for early knee osteoarthritis structural progressor patient screening," *Therapeutic Advances in Musculoskeletal Disease*, 23 Feb. 2021, doi:10.1177/1759720X21993254.

[37] S.B. Kwon et al., "A machine earning-based diagnostic model associated with knee osteoarthritis severity," *Scientific Reports*, 2020; doi:10.1038/s41598-020-72941-4.

[38] A. S. Mohammed et al., "Knee Osteoarthritis Detection and Severity Classification Using Residual Neural Networks on Preprocessed X-ray Images," *Diagnostics (Basel)*, 2023; doi:10.3390/diagnostics13081380

[39] A. Tiulpin, et al., "Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach," *Scientific Reports*, 2018; doi:10.1038/s41598-018-20132-7

[40] S. Abd El-Ghany, M. Elmogy, A. A. Abd El-Aziz, "A fully automatic fine tuned deep learning model for knee osteoarthritis detection and progression analysis," *Egyptian Informatics Journal*, vol 24, no. 2, 2023, pp. 229-240.