International Journal of Advanced Technology and Engineering Exploration, Vol 10(108)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2023.10101973

Research Article

Radiographic imaging-based joint degradation detection using deep learning

Aseel Ghazwan[1], Salma Al-Qazzaz[2] and Ashwan A. Abdulmunem[3*]
Department of Biomedical Engineering, Al-Nahrain University, Baghdad, Iraq[1]
Department of Physics, College of Sciences for Women, University of Baghdad, Baghdad, Iraq[2]
Department of Computer Science, Faculty of Computer Science and Information Technology, University of Kerbala, Kerbala, Iraq[3*]

Received: 17-July-2023; Revised: 12-November-2023; Accepted: 15-November-2023

Abstract
Osteoarthritis (OA) is a degenerative joint disease that primarily affects the knee joint. Currently, OA diagnosis relies on the examination of plain radiographs, a method susceptible to subjectivity and time consumption. This study aims to automatically assess the severity of knee OA using the "Kellgren and Lawrence (KL) grading system" based on plain X-rays. Leveraging 1650 digitized knee X-ray images, we implemented a custom MobileNetV2 architecture for a convolutional neural network with four distinct orientations. The methodology comprises two models: a fixed base and a trainable head. The MobileNetV2 network serves as the base model, while the proposed head model architecture includes an average pooling layer followed by a fully connected layer to enhance network efficiency. Results indicate that, except for grades 1 and 2, the methodology correctly identified KL grades with an accuracy of over 90%. Overall, the proposed approach demonstrates promising potential for classifying knee OA based on plain X-rays, achieving a 95% accuracy in detecting severe knee OA (stage 4). Researchers acknowledge the superior performance of their methodology compared to previous models in similar investigations, suggesting its effectiveness in forecasting OA severity based on radiographic imaging. Furthermore, the study's results support the effectiveness of deep learning-based approaches in diagnosing OA severity, with significant implications for improving patient outcomes. Although the suggested methodology shows a maximum accuracy rate of 95% in identifying severe knee OA cases (specifically, stage 4), there is a need for enhancements to address the misdiagnosis issue in stage 1 and 2 knee OA, where the accuracy rate is at 87%. This misdiagnosis arises from the similarity in characteristics between these stages.

Keywords
Knee OA, Detection disease, Deep learning, Image classification, Pre-trained networks.

1.Introduction
Osteoarthritis (OA) is a chronic joint disease that most commonly affects the knee [1], which is often accompanied by pain, altered gait mechanics, low quality of life and functioning disability [2]. The radiographic appearance is considered as a current gold standard for OA screening, where "Kellgren and Lawrence" (KL) [3] established radiographic criteria to categorise OA into five stages. Stage 0 refers to a normal, healthy knee. The highest score, 4, is given to severe OA. Visual evaluation of x-ray images, on the other hand, is time-consuming and highly relies on physicians' experience and carefulness.

*Author for correspondence

Technological advances have shown promising results in forecasting OA incidence, deterioration, progressive pain, advanced structural change, and the duration it takes to the end stage of OA [4, 5]. Researchers have attempted to utilise traditional machine learning (ML) to extract features prior to training the algorithms [6].

Several methods that use ML models have been documented in published works, to extract features prior to training the algorithms, to identify early knee OA, predict future disease scenarios, and develop novel treatments [7]. Deep learning (DL), on the other hand, feeds data directly to the algorithm and allows it to learn new features on its own. This has proven to be a novel, quick, fully-automated, and tremendously effective method for extracting meaningful diagnostic information from imaging data and opening up new opportunities for non-ML professionals to develop their own research and applications [8]. DL architecture includes recursive, recurrent, convolutional and unsupervised pre-trained networks [9]. Overfitting is a key issue in DL due to the presence of a large number of parameters and the utilisation of advanced regularisation algorithms in its design. The importance of dividing data into three sets for ensuring the effectiveness and adaptability of a model, the three sets are a test set for assessing the model's performance on unobserved data, a validation set for preventing overfitting, and a training set for optimising hyper-parameters. By using this approach, the model can be optimised and tested to ensure that it can perform accurately and effectively in real-world scenarios [10]. The significant prevalence of knee OA necessitates immediate progress towards improved techniques for detecting its existence and assessing its severity. The tiredness problem that arises from prolonged diagnosis may be avoided with fully automated knee severity grading, which can give an objective, repeatable prognosis. Knee OA severity prediction from raw screened knee x-ray images primarily entails two steps: locating the knee joint and assigning it to one of five KL classes based on the degree of damage shown. DL models can be easily scaled up to interpret huge amounts of data, and their accuracy and resilience can be improved by training on enormous datasets. These models may be integrated into healthcare systems, making diagnosis easier even in remote and low-resource settings. Accordingly, the motivation for diagnosing knee OA using plain x-ray using DL techniques is to enhance the accuracy, efficiency, and reliability of OA diagnosis, thereby contributing to the decrease of the worldwide knee OA burden.

Multiple researchers have used DL-based algorithms to analyse knee OA [10‒12]. However, knee analysis's effectiveness still has prospects for further development. Considering the ordinal nature of the KL grading assignment, a more effective loss function may lead to higher quality KL grades. To address this issue, this study investigated whether plain radiographs might simultaneously identify knee OA features and severity. In specific, the goal of this study is to develop and implement a DL-based approach for OA diagnosis that is reliable and cost-effective. Demonstrate the efficacy of the proposed approach in clinical settings, evaluate its performance using a number of performance metrics, and compare its performance to those of existing state-of-the-art procedures. This also helps doctors quickly identify and categorise knee OA. Specialists in the medical field must put out much effort in order to predict the outcome of such situations.

The article is structured to explore the complete work in different sections. Section 1 encompasses the background, challenges, inspiration, goals, and contributions. Section 2 presents relevant work, followed by an explanation of the suggested technique in Section 3. Results are detailed in Section 4, while Section 5 elaborates on the discussion. Finally, conclusions are drawn in Section 6.

2.Literature review

The ability of DL architectures to detect radiographic OA progression has revolutionised the area of medical imaging, exceeding previous computer vision technologies that needed data representation techniques to be manually coded. Yeoh et al. [13] assessed the current status of DL for predicting knee OA severity. The model's performance is equivalent to that of an attending physician with ten years of expertise [14]. This approach has the potential to minimise variability in knee OA diagnosis and treatment [15]; therefore, several approaches for detecting knee OA have been developed. For example, a network trained for multiclass classification and regression was developed by Antony et al. [12] for autonomously localising knee joints and measuring OA severity. They demonstrated that autonomously localised knee joints classify as well as manually segmented knee joints. Its multiclass classification accuracy, precision, recall, and F1 score are better than the previous approach [11]. The confusion matrix and other measures demonstrate that categorising knee OA images based on KL grade 1 is challenging due to small structural differences in the initial stage of OA. Chen et al. [16] used two deep “convolutional neural networks” (CNNs). A one-stage you only look once (YOLO) v2 network was used to locate the knee joints. The identified knee joint images were subsequently classified using a novel configurable ordinal loss in “ResNet”, “VGG2”, and “DenseNet” versions, as well as “InceptionV3”. The classification performance of “CNN” models is largely reliant on the recognition task, and the fine-tuned “VGG-19” model performed the best. In order to predict KL grades in OA, Gornale et al. (2020) explored the use of an ordinal regression module with a cumulative link loss function in six distinct neural network architectures: “VGG”, “GooLeNet”, “ResNet”, “DenseNet”, “ResNeXt”, and “MobileNetV2”. KL grades 0, 2, 3, and 4 were accurately recognised at 70% or higher; however, grade 1 performed poorly at

38.5%. This strategy improved KL grade 1 categorisation over baseline and Chen et al. [16]. Both investigations found that the ordinal regression module decreased misclassification and improved categorisation.

Extraction of significant regions from distorted images may become challenging due to issues related to filming, handling, and digitalisation during capturing. Therefore, Gornale et al. [17] extracted acceptable invariant characteristics from such distorted images using Hu's invariant moments. The experimental findings for rotated and scaled pictures are generally compatible with the original image's invariants. Consequently, the recommended algorithm results, as reviewed by orthopaedics and rheumatologists, are more competitive and encouraging. The diagnostic performance of the DL models was significantly better than that of a traditional approach, which included radiographic and demographic risk markers [18]. Leung et al. [19] suggested a model that used a transfer learning (TL) strategy with sevenfold layered cross-validation based on the ResNet34 architecture. With the use of clinical data such as body mass index (BMI), Schiratti [20] developed a "multimodal" DL approach that relied on information not clearly revealed in images in conjunction with BMI. This data-driven technology provides information that cannot be directly analysed in a radiologist's clinical practice. Kondal et al. [21] introduced a methodology that utilises CNN for the automated grading of knee radiographs based on the KL scale. The suggested approach consisted of two interrelated stages. In the first stage, an object identification model was used to isolate individual knees from the surrounding picture. Subsequently, in the second stage, a regression model was utilised to automatically assign a KL scale to each identified knee. The researchers provided evidence that fine-tuning the model prior to evaluating it on a private hospital dataset resulted in a decrease in the mean absolute error from 1.09 to 0.28.

To improve knee OA severity prediction findings from plain radiographs and capture the multi-scale aspects of knee X-rays, Jain et al. [22] presented an "OsteoHRNet" DL approach using the “high-resolution network” (HRNet). The “HRNet” worked very well and produced significant advantages over the previously reported techniques due to its ability to retain high-resolution features throughout the network while collecting reliable spatial information. Additionally, radiographs with similarities and differences between classes were easier to spot according to the attention mechanism. The model's learning of the spatial properties of the radiographs was validated using “gradient-weighted class activation mapping” (Grad-CAM).

Guan et al. [23] used non-image data to examine two CNN models' capacity to predict knee OA. The support vector machine (SVM) clinical model predicts OA by combining demographic and risk factor data. There was a substantial difference between the clinical and SVM/DL models in terms of area under the curve (AUC). The highest AUC, 0.832, was achieved by the combined SVM and DL model, which was significantly higher than the clinical model. On the other hand, the effectiveness of DL over the logistic regression (LR) model was highlighted, where the latter used anthropometric, demographic data and KL grade as input.

Guan et al. [18] suggested a method that combines the YOLO model with “DenseNet” DL networks for cropping and categorisation of the joint to predict OA progression; the follow-up time in this study was only 48 months. In line with this, Wang et al. [24] combined the visual transformer with the YOLO object identification model to create a fully automated system for diagnosing OA. Compared to the standard CNN architectures, the classification outcome boosts accuracy by 2.5%. Our categorisation result is also 2.5% more accurate than that of standard CNN architectures. Using DL and ML algorithms, Ahmed and Mstafa [25] introduced new methods to classify a Knee X-ray picture according to KL grading standards. The suggested methodologies use deep hybrid learning-I (DHL-I) and DHL-II learning architectures. DHL-I, the first, uses CNN structure to train a new structure containing five prediction classes on knee X-ray images and then harvest features. SVMs may identify knee OA by pattern discrimination after a principal component analysis (PCA) reduces these learned characteristics. The second, DHL-II, is identical save for the following. DHL-I's pre-trained CNN was fine-tuned using TL to categorise knee OA into four classes, three classes, and two class labels. The suggested technique improves multiclass and binary class-based classification accuracy in the OA case study, as shown by experimental results. The empirical data showed that binary class labels outperformed all others, reaching 90.8% accuracy. Also, the developed models helped classify the condition early on, reducing its development and improving quality of life.

Khalid et al. [26] established three x-ray-based methods for diagnosing knee OA and distinguishing KL grades using the Osteoarthritis initiative (OAI) and rani channamma university (RCU) datasets. After CNN models, all approaches used PCA to remove unnecessary features and maintain significant ones. The first technique relies on VGG-19 and ResNet-101 technologies to analyse x-rays and determine knee inflammation levels. The feed forward neural network (FFNN) technique for X-ray analysis and knee OA grade diagnosis combines VGG-19 and ResNet-101 features before and after PCA. The third FFNN approach for X-ray analysis and knee OA grade diagnosis uses VGG-19 and ResNet-101 fusion features and handmade features. By combining VGG-19 and handmade features in an OAI dataset, FFNN achieved 99.25% AUC, 99.1% accuracy, 98.81% sensitivity, 100% specificity, and 98.24% precision. FFNN achieved 99.07% AUC, 98.20% accuracy, 98.16% sensitivity, 99.73% specificity, and 98.08% precision on the RCU dataset using VGG-19 fusion features and handmade features.

Abd et al. [27] employed DenseNet169 DL to fine-tune knee OA diagnosis to boost efficiency. Knee OA severity will be determined by multi-classification and binary classifications in the suggested model. Localising peripheral, diffuse, and vascular thickening opacities will be effective. The DenseNet169 model has accuracy, sensitivity, specificity, precision, and F1-score of 95.93%, 88.77%, 95.41%, 85.8%, and 87.08% in multi-classification. The “DenseNet169” model had 93.78% accuracy, 91.29% sensitivity, 91.29% specificity, 87.57% precision, and 89.27% F1-score in binary classification. Thus, the suggested paradigm has unmatched perceptual tuning compared to previous frameworks.

The suggested technique by Kokkotis et al. [28] included fuzzy logic-based feature selection, learning algorithms, and nan-explainability analysis. Fuzzy logic combined several feature relevance scores to gather more informative features, and the suggested technique aggregated filter, wrapper, and embedding feature selection methods. The suggested technique selected a subset of risk variables that improved ML model accuracy relative to conventional feature selection methods. On 21 risk indicators, the best random forest (RF) classifier model had 73.55% classification accuracy. Two verified radiographic knee OA models were suggested by McCabe et al. [29]. The diagnostic and prognostic model of KOA onset time was added. OAI supplied model development and optimisation data, while multicenter osteoarthritis study (MOST) provided external validation for both models. Diagnostic model AUC was 67% for validation and 75% for test data.

Among studies investigating the relationship between pain progression and quantitative imaging outcomes, Guan et al. [30] predicted the progression of pain from baseline X-ray images based on DL. The AUC for this approach is 0.770. With the addition of demographic and clinical data, the AUC improved to 0.807. In order to extract features from knee radiographs, a DL model was developed and trained. The extracted features were then combined with demographic and radiographic risk factor data to predict outcomes for OA. In this study, the authors combined the DL model with a DL/LR model using non-image data. Compared to conventional models, this hybrid approach was able to achieve a much higher AUC, demonstrating the effectiveness of combining non-image and image data with ML models to improve predictions for OA outcomes. The method used in this study was similar to that of Guan et al. [23], who also achieved success in predicting OA outcomes using a similar approach. Pedoia et al. study [31] tested DL models' ability to identify and stage meniscus and femoropatellar cartilage abnormalities in OA patients with anterior cruciate ligament (ACL) rupture. The sensitivity and specificity for identifying meniscus lesions were 90 percent and 82 percent, respectively; for detecting femoropatellar cartilage lesions, the sensitivity and specificity were both 80 percent, with better performance when demographic data were incorporated.

Two previous studies [32, 33] reported on traditional OA risk assessment models for predicting pain progression in individuals at risk for knee OA based on demographic, clinical, and radiographic risk characteristics. In the prevention of knee OA in overweight females experiment, Landsmeer et al. [32] applied a conventional model to predict the onset of frequent knee pain over a six-year follow-up period in 472 knees of overweight and obese women without knee OA. The AUC for a multivariate LR model, including BMI, knee pain at baseline, knee pain going up stairs, morning stiffness, being postmenopausal, and heavy work, was 0.71. To predict pain progression in 1243 knees in the OAI database, Halilaj et al. [33] employed number of potential risk factors, including demographics, knee symptoms and grades, medication use, history of the family, general health status, gound level walking ability, and knee alignment measurements on radiographs. The AUC for predicting pain worsening during an 8-year follow-up period using a LASSO regression model was 0.79. Halilaj et al.'s [33] model had excellent diagnostic performance, but it would be challenging to implement in routine clinical use because it required analysis of a large number of risk factors gleaned from exhaustive and time-consuming clinical history, physical examination, and radiographic evaluations.

Widera et al. [34] evaluated six ML approaches for predicting knee OA. There were four predicted categories: progressive structural change, non-progressive pain, progressive pain, and progressive structural change with pain. On a down sampled training set, RF outperformed balanced learning. The dual classifier enhanced the findings. Due to the nature of clinical studies, this research focused on a limited progression time frame.

Huang et al. [35] aimed to track OA development over time and among patients. The dynamic functional mixed-effects model was suggested to identify specific aberrant areas at the baseline, 1, 2, and 4-year MRI scans. The model accommodated spatial-temporal heterogeneity. By predicting how cartilage changes over time, the model can provide insight into how the disease progresses and how it affects the body. This information can be useful in developing treatments and interventions that target specific stages of the disease, potentially improving outcomes for patients with OA. Overall, the model's ability to accurately predict changes in cartilage over time is a significant contribution to our understanding of OA and its pathophysiology. Tolpadi et al. [36] employed a DL pipeline comprised of DenseNet-121 and LR to predict total knee replacement (TKR). Both non-imaging data and imaging data (x-ray and MR images) were used to evaluate the model's performance. Throughout all OA phases, this model was more sensitive than the integrated x-ray model (88.4). An integrated MRI model had an AUC of 0.834, exceeding x-ray models in the non-OA group. This model accurately identified TKR events in individuals who did not have OA at baseline (AUC 0.943). Advances in computer power and data accessibility provided by these technologies, potentially make early identification of OA easier using three-dimensional (3D) DL [13].

While most studies on multiclass OA categorisation ignore the actual continuous range of OA progression, Li et al. [37] created a Siamese neural network to categorise OA at specified time intervals. The AUC of the proposed model is 0.9. In the proposed technique, the authors claim that image classification may remove the necessity for manual ROI localisation.

Swiecicki [38] created an automated DL based approach that employs Posterior-Anterior (PA) and lateral (LAT) views of knee radiographs to score OA severity. The results demonstrate the limitations of the LAT views' data set.

Saini et al. [39] suggested a three-stage pre-processing approach that uses gaussian-filter noise reduction, pixel-centering normalisation, and balanced contrast enhancement successively. The severity classification architecture was transfer-learning-based VGG16. The developed categorisation system surpasses current approaches with an accuracy of 89.95%. Norman et al. [40] proposed an automatic OA detection technique. A U-net model was used to identify the joint in knee radiographs at six distinct time intervals. These images were used to train a “DenseNet” neural network architectural ensemble for OA severity prediction. “DenseNets” ensemble has 83.7, 70.2, 68.9 and 86.0 percent testing sensitivity for the non-OA group, mild, moderate and severe OA, respectively. The model was validated using saliency maps to ensure the neural networks selected the right osteoarthritic properties for identification. The findings of this study suggest the use of automated classifiers to help radiologists make more precise diagnoses.

Liu et al. [41] employed FasterR-CNN, which combines region proposal network (RPN) and Fast R-CNN, to analyse the input images for location and classification. The RPN is trained to propose knee joint regions that are subsequently classified using Fast R-CNN. Using CNNs, clinically valuable properties may be retrieved from X-ray pictures. To correct the class imbalance and improve the model's performance, a weighted loss function was used. Larger anchors are utilised with larger X-ray image input sizes to address anchor mismatch issues. The model's performance is widely evaluated. In terms of accuracy, sensitivity, and specificity, the improved model performs better than the Faster R-CNN. Each image is evaluated in 0.33 seconds.

Many researchers in OA have shown success in analysing MRI data, predicting OA development, and automating KL-grading of knee radiographs using CNN. However, there have been few attempts in utilising plain radiography to determine individual knee OA characteristics. Thus, this research was undertaken to determine the possibility of concurrently recognising particular knee OA characteristics and total knee OA severity using plain x-rays.

3.Materials and methods

3.1Dataset

The dataset of 1650 digitised knee X-ray images supplied by Gornale and Patravali [42] was used in this study. The original images are 8-bit grayscale. Each radiographic knee X-ray image is individually marked/labelled in accordance with KL grades. It contains anteroposterior knee radiograph assessments of 400 Chang Gung Memorial Hospital participants, and the diagnosis was established by a doctor on each X-ray imaging with the original resolution. These data cover patients with KL grading scale grades 0 to 4, and the amount of data is the same for all grades. A score of 0 indicates that there is no evidence of OA, while a score of 4 indicates severe OA. Grades 1, 2, and 3 indicate increasingly severe stages of OA. This categorisation allows the degree of OA to be defined into a numerical number, allowing clinicians to differentiate between different degrees of OA and choose the best treatment option. *Figure 1* shows samples of the dataset. The representative knee radiographs with the corresponding KL grading are illustrated in *Figure 2*.

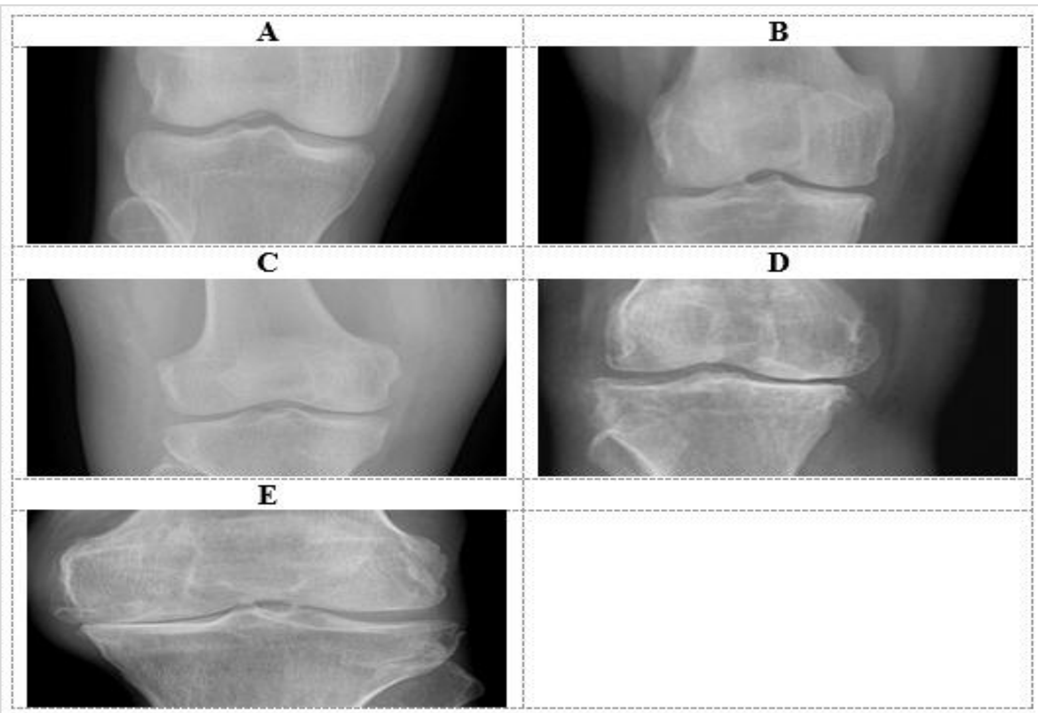

Figure 1 Representative antero-posterior (AP) knee radiographs of KL classification Grade 0 – A, Grade 1 – B, Grade 2 – C, Grade 3 – D, Grade 4 – E [42].

3.2Preprocessing

This section tackles the issue of a dataset with a limited number of images. We acknowledge that a ML model requires a sizable number of statistically significant samples in the training dataset in order to generalise beyond the training data. The proposed data augmentation technique (rescale= 1./255, shear_range = 0.5, zoom_range = 0.25, rotation_range = 45) is used to increase the size of the dataset because the chosen dataset only contains a small number of images with suitable features. All of the images were scaled to 224 pixels in width and height. Following that, a feature map was constructed and targeted in order to train the model. This step is critical in enhancing the results of the proposed model.

3.3OA classification by Mobilenetv2 networks

The base model is the pre-trained model, which is often used to accurately categorise limited datasets. Because a DL model trained from scratch on limited data is unlikely to achieve high accuracy. So it can't learn data characteristics. A big dataset is used to train the model, which is represented in the network by weights. Later on, the weights are employed in a second network for a new task and dataset. So, rather than starting from scratch with the second network, we "transfer" the first network's learnt properties to the second. Deep convolutional networks' first layers are typically used to learn visual features like lines and shapes. The network's last layers learn task-specific information, such as image classification. Thus, in TL, the core model weights stay unchanged. The network's last layer, which is often a fully connected layer, then learns the new dataset's specific features.

The approach relies on Mobilenetv2 [43], which utilised depthwise separable convolution. In two-dimensional (2D) convolution, the depth dimension (channel) is included; thus, all input channels are processed into one. Each input channel is then depth wise convolved with its filter channel.

The stacked filtered output channels, known as a pointwise convolution, are used to merge the stacked output channels. Mobilenetv2 accepts images up to 224×224×3. These are scaled and trimmed to 224×224 pixels. We set to include top = False to remove the top layers of the pre-trained model, which is appropriate for feature extraction. Following a convolution layer with 32 filters, 19 inverted residual bottleneck layers were added [43]. Residual blocks are used to connect the start and end of convolutional blocks. The flowchart of the proposed pipeline is illustrated in *Figure 2*. As shown in the figure, the model is based on a pretrained Mobilenetv2 with modification on the last layers by adding the average pooling layer and dropout to avoid overfitting.

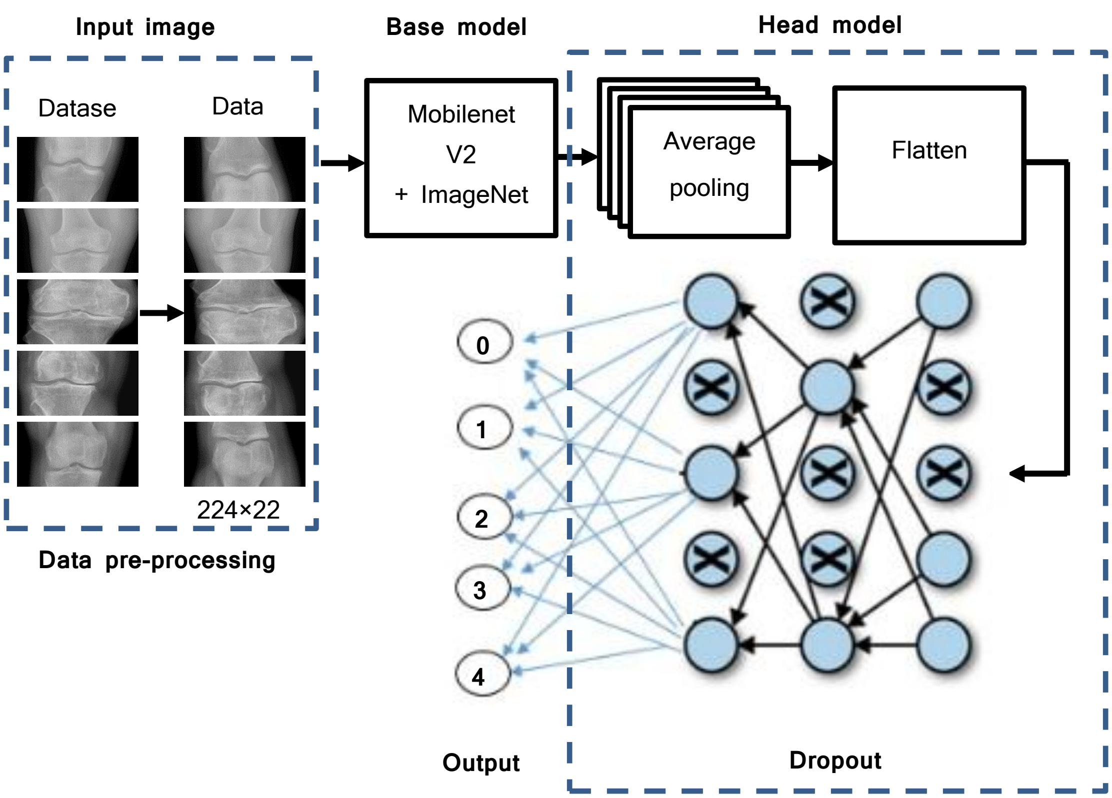

Figure 2 The flowchart of the proposed DL model

3.4Training and implementation details

Deep convolutional networks' first layers are typically used to learn visual features like lines and shapes. The suggested approach was written in Python (3.6) and tested using Windows 7 on an Intel Core i5 CPU with 12 GB of RAM. DL toolkits for Tensorflow and Keras toolbox were used to train the model, which took about seven hours. In stochastic gradient descent (SGD), the following parameters were adjusted in order to update the loss function (categorical_crossentropy) on the training data set: The learning rate is 0.0001, and the maximum number of epochs is 300 and other experiment with 100.

In the training phase, the feature extraction will be achieved. The first stage of feature extraction involves a series of 2D convolution layers (by adding 6 convolution layers with 64 size of fully connected layer and 25% dropout rate), with each layer being followed by a non-linear activation function. This helps to identify valuable features in each hidden layer. However, overfitting can be a problem with this model, especially with big models and huge datasets. Using "Rectified Linear Units (ReLU)" for all layers in the feature extraction step can help prevent overfitting while ensuring quicker learning and better performance. *Table 1* explains the hyper parameter which affect the performance of the model. To put it simply, the convolution layer acts as a filter on the input data using filter kernel coefficients that are determined during the training process. The first convolution layer extracts basic patterns in the incoming data, resulting in low-level features.

These key properties may be combined in the subsequent convolution layer to generate patterns of patterns and so on. Higher-level feature patterns are created by the combination of these secondary characteristics.

Table 1 The hyper-parameter which affects the performance of the model

Layers	Parameter	Hyper parameter
Convolutions	Kernels	Adjustable kernel number and size, activation functions, stride, and padding
Pooling	-	Kernel number and size, activation functions, stride, and padding can all be altered.
Fully Connected	Weights	Number of total weights and activation function
Other		The effectiveness of the model is influenced by several factors. The model’s loos function, framework, epochs, optimiser, learning rate, mini batch size, dataset partitioning, weight initialization, and regularisation are some of these components.

4.Results

The dataset has been splatted into a train, test, and validation sets (80%, 10%, and 10%, respectively) before the model has been trained. The train set has been used to train the knee classification model, and the validation set has been used to assess the model while it is being trained, and the test set is used to evaluate the recognised knee condition on the trained model. The architecture was modified by adding convolutional layers; this step increased the recognition rate from 87% to 93%. The following information pertains to the proposed model. As shown below.

Total parameters of the model: 2,456,261

Trainable parameters: 198,277

Non-trainable parameters: 2,257,984.

An automated OA categorisation will help physicians distinguish between OA degrees and choose the appropriate treatment approach. The performance of the recommended model was evaluated using three common metrics: precision, recall, and F1-measure with epoch 300 and epoch 100, as presented in *Table 2*. The F1-measure gives an intersection measurement between the manually defined OA and the prediction outcomes of the fully automated technique. Where 0-4 represent the classes that will be classified. Precision, recall and F1- measures are also calculated for each class. *Table 2* depicts the performance evaluation of the proposed approach based on the OA grades. According to the results, the sever OA (stage 4) earned the greatest precision score of 95%. The recall and f1-performance metrics ratings are 92% and 94%, respectively. The present research yields the finding that the proposed approach used in this study exhibited suboptimal performance metrics in the task of diagnosing knee OA via the utilisation of knee x-ray pictures. *Table 3* explains these metrics and compares the performance of the proposed model with other models, [44, 11].

Table 2 The proposed DL model's performance with epoch 300 and 100, with respect to OA grades

	Epoch = 300			**Epoch =100**		
Parameter	Precision	Recall	F1-measure	Precision	Recall	F1-measure
KL score						
0	0.94	0.92	0.93	0.62	0.81	0.70
1	0.87	0.95	0.91	0.81	0.72	0.76
2	0.87	0.84	0.86	0.65	0.56	0.61
3	0.90	0.89	0.90	0.79	0.78	0.79
4	0.95	0.92	0.94	0.74	0.70	0.72

Table 3 The proposed DL model's performance, with respect to OA grades, compared with other models

	The proposed model			**Thomas et al. 2020 [44]**			**Antony et al. 2016 [11]**		
Parameter	Precision	Recall	F1-measure	Precision	Recall	F1-measure	Precision	Recall	F1-measure
KL score									
0	0.94	0.92	0.93	0.73	0.87	0.79	0.57	0.92	0.71
1	0.87	0.95	0.91	0.38	0.27	0.31	0.32	0.14	0.20
2	0.87	0.84	0.86	0.71	0.67	0.69	0.71	0.46	0.56
3	0.90	0.89	0.90	0.82	0.81	0.81	0.78	0.73	0.76
4	0.95	0.92	0.94	0.87	0.86	0.87	0.89	0.73	0.80

Figure 3 compares the outcomes of the training and validation sets and displays the model's accuracy and loss with time. It can be observed that the curves in the validation and train sets are comparable and that the generalisation gap between the train and test loss is minimal when using epoch 300 and epoch 100, as shown in *Figure 3*. This suggests that the categorisation outcomes in the training and validation sets were comparable. Moreover epoch 300 improve the performance of the suggested model. In ML, precision and recall are both significant evaluation metrics, but they have different functions and are applied in various situations. Saying that one is always superior to the other is untrue. Depending on the particular requirements of the current task, one must choose between precision and recall.

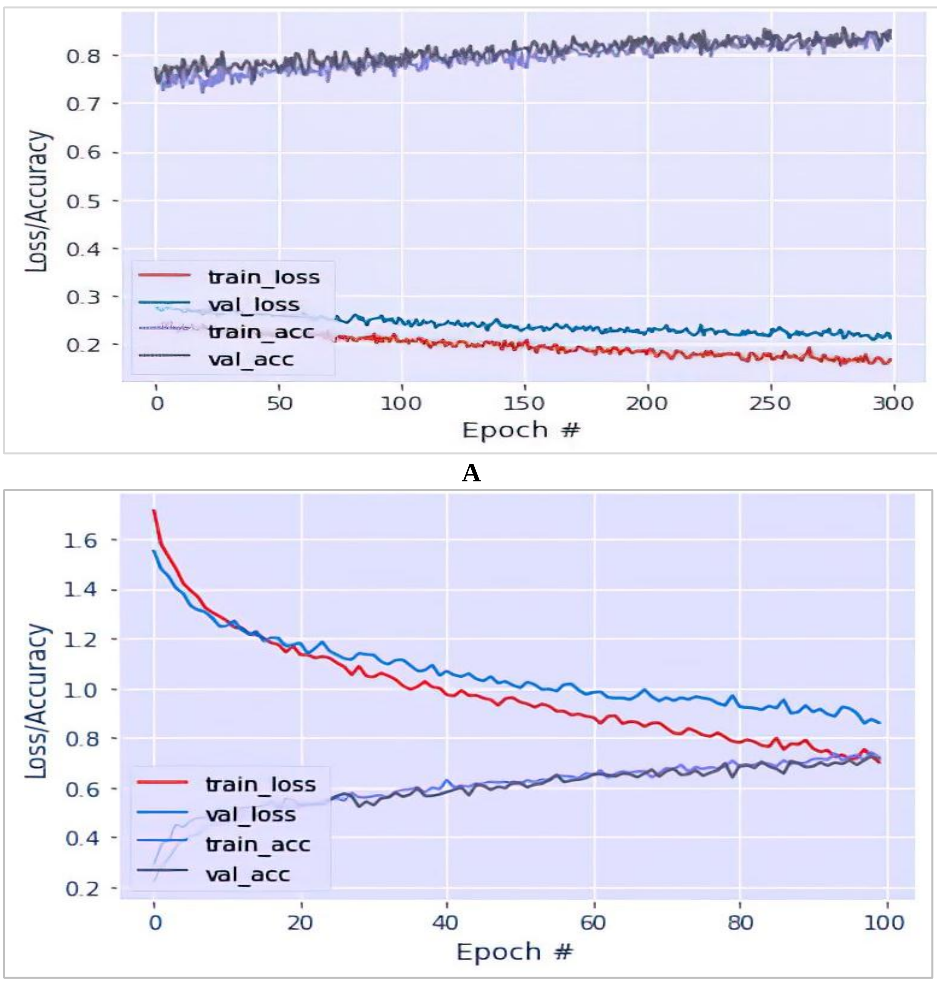

Figure 3 Loss and accuracy for the proposed DL model for a different number of epochs; A) Number of epochs=300; B) Number of epochs=100

5.Discussion

The current study illustrated the applicability of utilising DL to predict OA progression using plane knee radiographs. The model demonstrated exceptional diagnostic efficacy, properly predicting knee OA with a precision of 90%. Nevertheless, it is important to acknowledge that the aforementioned level of precision was not found in instances categorised as grades 1 and 2, which achieved just 87%. The results of this study indicate that the diagnostic accuracy of the model is superior in predicting the advancement of pain in knees afflicted by OA in comparison to knees with risk indicators for OA that have not yet shown radiographic evidence of the illness. Similarly, the same finding was obtained for the baseline technique in [11, 16, 40, 44, 45]'s investigation, where the rate of accurate prediction on grade 1 pictures was lower than on subsequent grades. However, our method successfully predicted the majority of knee joint radiographs and beat Antony and Thomas's model's performance, as shown in *Table 1*. The KL grades of 0 and 1 were combined by [40], since they both indicate the absence of OA. One potential explanation for the misclassification of grades 1 and 2 may be attributed to the unclear criteria used. This ambiguity arises from the fact that knees with risk factors for OA had not yet shown radiographic symptoms that could accurately identify the relevant regions within the picture for categorisation. Our modules limit the possibility of remote misclassification instances, specifically for moderate and severe OA, KL grade 3 and grade 4, respectively. In terms of clinical performance, the suggested method has great potential for learning local radiography characteristics and categorising knee OA based on plain x-ray, with 95 percent accuracy in detecting severe knee OA (stage 4). The observed accuracy was outerperformed that used TL-based EfficientNetB1, with an accuracy of 89% on Indian data [46]. In a separate study, several ML algorithms including SVM, LR, Adaboost, gradient boosting, and multi-layer perceptron were used for comparative analysis... The research findings indicate that the used LR approach yielded a performance score of 84%, which is relatively poor [47].

The suggested approach did not overfit. Overfitting occurs when the final model fits the training data too well and lacks generalisation capabilities, resulting in poor accuracy when employed on new data. Each DL model training iteration is validated to check for overfitting. During validation, the network with the updated weights is tested on non-training data. *Figure 3* shows the training and validation data loss and accuracy results for each training iteration. The validation data used in this research are from the testing set. An overfitted model performs well on training data but badly on new data. As a consequence, each repetition reduces the training loss but raises the validation loss. As seen in *Figure 3*, both training and validation losses decrease with each training cycle. Thus, the suggested model is not overfitting and has a high generalisation ability. This is because the suggested model employs a modest epoch value and augments the training data to enhance its variability.

Table 2 demonstrated that the proposed model outperforms Thomas et al. [44], Antony et al. [11], and Rehman et al. technique [48], where the former used the "saliency map algorithm" to identify and highlight the most visually significant regions within an image. The derivation of the contribution of each pixel is obtained by the backpropagation technique. The network propagates the variability in the output layer, revealing the pixels that have made the greatest contribution to this fluctuation. The pixels that would have the most significant impact on the projected score are those that undergo alterations in the input picture. The pixels that we perceive as the most predictive ones. In this representation, the intensity mapping is shown using an image that exhibits transparency for extremely low values, green for low values, and red for high values. The matrix of intensity is often known as a saliency map. Whereas Antony et al.'s technique [11] used mean squared error as a metric for evaluating the effectiveness of an automated knee OA severity assessment instead of relying on binary and multi-class classification accuracy. In order to get an accurate model, several researchers have used a multimodel method rather than augmentation. They employ an ensemble of models, each contributing to the classification of a picture, with one model cropping the joint space. When compared to our method, which employs a single model for the whole of the study, methods that use many models demand more time and resources to compute. Rehman et al. technique [48] carried out a comparative performance analysis of using different ML methods. The accuracy score for SVM and SGD approaches is 0.29, the lowest among comparisons. Metric performance data indicate poor scores for SVM and SGD algorithms in each class. The K-nearest neighbours (KNN) earned a performance score of 0.63, outperforming SVM and SGD approaches. The KN approach earned a 0.69 precision score. The tree-based RF approach was the only one that scored well in comparison. The RF method scored 0.77 precision, which is lower than that earned in the proposed method. Clinical decisions about joint replacement surgery often collaborate both pain assessment and KL score; however, evidence indicates that the preoperative KL grade might be a prognostic factor for surgical outcomes. This finding provides evidence for using an automated tool, such as the one we have developed, to enhance the decision-making process with more informed choices. Therefore, the proposed model has the potential to be used in healthcare facilities for the prompt and precise prediction of knee OA or those who possess risk factors for the condition.

A limitation of the proposed method is the time-consuming nature of the training stage, a common challenge in DL and ML, especially when dealing with large datasets. Nevertheless, the authors highlight that the prediction phase of the model is considerably faster, swiftly processing the testing dataset for accurate classification. This suggests that the approach can efficiently analyze test data, even if the training step is time-intensive. The scalability of the approach will be assessed in future studies using more extensive datasets.

In this study, picture resolution was reduced to 8-bit, potentially resulting in the loss of fine-grained image data. Utilizing the original image's resolution and filtering data could enhance the findings. It is important to note that the proposed method may benefit from additional validation studies to assess its effectiveness in real-world settings and to evaluate its generalizability to other populations and imaging modalities.

Furthermore, exploring advanced regularization techniques or other strategies to reduce the model's complexity could optimize the training process and decrease training time. Overall, while the proposed method holds promise for detecting OA severity from radiographic imaging using DL, further research is essential to comprehensively evaluate its potential and limitations. A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion

OA classification has become a vital problem in medical imaging to diagnose the degree of the disease. DL proved an effective methodology for implementing a system to deal with this challenge. This study proposes a DL model and a TL approach for categorizing knee OA radiographs into five stages. The technique proposes two models: a fixed base and a trainable head. The Mobilenetv2 network was selected as a base model. The recommended head model design comprises an average pooling layer followed by a fully connected layer to ensure network efficiency. While the proposed method utilized in this study was effective in identifying KL grades for most classes, it had slightly lower accuracy for grades 1 and 2, achieving an accuracy of 87%. This finding is consistent with other state of the art methods, which also found that predicting the KL grade for grade 1 images was more challenging than for subsequent grades. These findings suggest that predicting the severity of OA in its early stages may be more difficult than in later stages, which could have implications for early detection and treatment. Despite these challenges, the proposed method still achieved high accuracy overall, with an accuracy of over 90% for all other grades. The authors acknowledge that their methodology shown superior performance compared to previous models used in similar investigations, suggesting that their technique might potentially provide more efficacy in predicting the severity of OA based on radiographic imaging. Moreover, this study's findings provide further evidence that DL-based methods can be effective for detecting OA severity, which has important implications for improving patient outcomes.

Overall, the experimental findings on the knee OA classification dataset suggest that the proposed approach can provide up to 95% accuracy in severe knee OA detection (stage 4). However, improvements are needed to overcome the misdiagnosis of stage 1 and 2 knee OA, with 87% accuracy, due to their comparable features. Our future work will investigate more modification methods to improve knee OA classification in stage 1 and 2 while keep the best in stage 4 detection.

Acknowledgment
None.

Conflicts of interest
The authors have no conflicts of interest to declare.

Author’s contribution statement
Aseel Ghazwan: Conceptualization and design of the work, interpretation, and analysis, writing and editing original draft. **Salma Al-Qazzaz:** Conceptualization and design of the work, interpretation, and analysis. **Ashwan A. Abdulmunem:** Conceptualization and design of the work, implementation, interpretation, and analysis.

References

[1] Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology. 1995; 38(8):1134-41.

[2] Kawano MM, Araújo IL, Castro MC, Matos MA. Assessment of quality of life in patients with knee osteoarthritis. Acta Ortopedica Brasileira. 2015; 23:307-10.

[3] Kellgren JH, Lawrence J. Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases. 1957; 16(4):494-502.

[4] Mahum R, Rehman SU, Meraj T, Rauf HT, Irtaza A, El-sherbeeny AM, et al. A novel hybrid approach based on deep CNN features to detect knee osteoarthritis. Sensors. 2021; 21(18):1-18.

[5] Jamshidi A, Pelletier JP, Labbe A, Abram F, Martel-pelletier J, Droit A. Machine learning–based individualized survival prediction model for total knee replacement in osteoarthritis: data from the osteoarthritis initiative. Arthritis Care & Research. 2021; 73(10):1518-27.

[6] Kuran EC, Kuran U, Er MB. Sub-image histogram equalization using coot optimization algorithm for segmentation and parameter selection. CS & IT conference proceedings. 2022(pp. 33-46).

[7] Teoh YX, Lai KW, Usman J, Goh SL, Mohafez H, Hasikin K, 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; 2022:1-20.

[8] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annual Review of Biomedical Engineering. 2017; 19:221-48.

[9] Gan HS, Ramlee MH, Wahab AA, Lee YS, Shimizu A. From classical to deep learning: review on cartilage and bone segmentation techniques in knee osteoarthritis research. Artificial Intelligence Review. 2021; 54(4):2445-94.

[10] Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. Scientific Reports. 2018; 8(1):1-10.

[11] Antony J, Mcguinness K, O'connor NE, Moran K. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In 23rd international conference on pattern recognition 2016 (pp. 1195-200). IEEE.

[12] Antony J, Mcguinness K, Moran K, O’connor NE. Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks. In machine learning and data mining in pattern recognition: 13th international conference, New York, USA, 2017 (pp. 376-90). Springer International Publishing.

[13] Yeoh PS, Lai KW, Goh SL, Hasikin K, Hum YC, Tee YK, et al. Emergence of deep learning in knee osteoarthritis diagnosis. Computational Intelligence and Neuroscience. 2021; 2021:1-20.

[14] Xue Y, Zhang R, Deng Y, Chen K, Jiang T. A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. PloS one. 2017; 12(6):1-9.

[15] Schwartz AJ, Clarke HD, Spangehl MJ, Bingham JS, Etzioni DA, Neville MR. Can a convolutional neural network classify knee osteoarthritis on plain radiographs as accurately as fellowship-trained knee arthroplasty surgeons?. The Journal of Arthroplasty. 2020; 35(9):2423-8.

[16] Chen P, Gao L, Shi X, Allen K, Yang L. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. Computerized Medical Imaging and Graphics. 2019; 75:84-92.

[17] Gornale SS, Patravali PU, Hiremath PS. Automatic detection and classification of knee osteoarthritis using hu's invariant moments. Frontiers in Robotics and AI. 2020; 7:591827.

[18] Guan B, Liu F, Haj-mirzaian A, Demehri S, Samsonov A, Neogi T, et al. Deep learning risk assessment models for predicting progression of radiographic medial joint space loss over a 48-MONTH follow-up period. Osteoarthritis and Cartilage. 2020; 28(4):428-37.

[19] Leung K, Zhang B, Tan J, Shen Y, Geras KJ, Babb JS, et al. Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: data from the osteoarthritis initiative. Radiology. 2020; 296(3):584-93.

[20] Schiratti JB, Dubois R, Herent P, Cahané D, Dachary J, Clozel T, et al. A deep learning method for predicting knee osteoarthritis radiographic progression from MRI. Arthritis Research & Therapy. 2021; 23:1-10.

[21] Kondal S, Kulkarni V, Gaikwad A, Kharat A, Pant A. Automatic grading of knee osteoarthritis on the Kellgren-Lawrence scale from radiographs using convolutional neural networks. In advances in deep learning, artificial intelligence and robotics: proceedings of the 2nd international conference on deep learning, artificial intelligence and robotics, 2022 (pp. 163-173). Cham: Springer International Publishing.

[22] Jain RK, Sharma PK, Gaj S, Sur A, Ghosh P. Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. Multimedia Tools and Applications. 2023:1-8.

[23] Guan B, Liu F, Mizaian AH, Demhri S, Neogi T, Guermazi A, et al. Deep learning approach to predict radiographic knee osteoarthritis progression. Osteoarthritis and Cartilage. 2019; 27:S395-6.

[24] Wang Y, Wang X, Gao T, Du L, Liu W. An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative. Journal of Healthcare Engineering. 2021; 2021:1-10.

[25] Ahmed SM, Mstafa RJ. Identifying severity grading of knee osteoarthritis from X-ray images using an

efficient mixture of deep learning and machine learning models. Diagnostics. 2022; 12(12):1-25.

[26] Khalid A, Senan EM, Al-wagih K, Ali AMM, Alkhraisha ZM. Hybrid techniques of X-ray analysis to predict knee osteoarthritis grades based on fusion features of CNN and handcrafted. Diagnostics. 2023; 13(9):1-26.

[27] Abd ES, Elmogy M, Abd EAA. A fully automatic fine tuned deep learning model for knee osteoarthritis detection and progression analysis. Egyptian Informatics Journal. 2023; 24(2):229-40.

[28] Kokkotis C, Ntakolia C, Moustakidis S, Giakas G, Tsaopoulos D. Explainable machine learning for knee osteoarthritis diagnosis based on a novel fuzzy feature selection methodology. Physical and Engineering Sciences in Medicine. 2022; 45(1):219-29.

[29] Mccabe PG, Lisboa P, Baltzopoulos B, Olier I. Externally validated models for first diagnosis and risk of progression of knee osteoarthritis. PloS one. 2022; 17(7):1-23.

[30] Guan B, Liu F, Mizaian AH, Demehri S, Samsonov A, Guermazi A, Kijowski R. Deep learning approach to predict pain progression in knee osteoarthritis. Skeletal Radiology. 2022:1-11.

[31] Pedoia V, Norman B, Mehany SN, Bucknor MD, Link TM, Majumdar S. 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. Journal of Magnetic Resonance Imaging. 2019; 49(2):400-10.

[32] Landsmeer ML, Runhaar J, Van MM, Oei EH, Schiphof D, Bindels PJ, et al. Predicting knee pain and knee osteoarthritis among overweight women. The Journal of the American Board of Family Medicine. 2019; 32(4):575-84.

[33] Halilaj E, Le Y, Hicks JL, Hastie TJ, Delp SL. Modeling and predicting osteoarthritis progression: data from the osteoarthritis initiative. Osteoarthritis and Cartilage. 2018; 26(12):1643-50.

[34] Widera P, Welsing PM, Ladel C, Loughlin J, Lafeber FP, Petit DF, et al. Multi-classifier prediction of knee osteoarthritis progression from incomplete imbalanced longitudinal data. Scientific Reports. 2020; 10(1):1-15.

[35] Huang C, Xu Z, Shen Z, Luo T, Li T, Nissman D, et al. DADP: dynamic abnormality detection and progression for longitudinal knee magnetic resonance images from the osteoarthritis initiative. Medical Image Analysis. 2022; 77:102343.

[36] Tolpadi AA, Lee JJ, Pedoia V, Majumdar S. Deep learning predicts total knee replacement from magnetic resonance images. Scientific Reports. 2020; 10(1):1-12.

[37] Li MD, Chang K, Bearce B, Chang CY, Huang AJ, Campbell JP, et al. Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging. NPJ Digital Medicine. 2020; 3(1):1-9.

[38] Swiecicki A, Li N, O'donnell J, Said N, Yang J, Mather RC, et al. Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. Computers in Biology and Medicine. 2021; 133:104334.

[39] Saini D, Khosla A, Chand T, Chouhan DK, Prakash M. Automated knee osteoarthritis severity classification using three-stage preprocessing method and VGG16 architecture. International Journal of Imaging Systems and Technology. 2023; 33(3):1028-47.

[40] Norman B, Pedoia V, Noworolski A, Link TM, Majumdar S. Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. Journal of Digital Imaging. 2019; 32(3):471-7.

[41] Liu B, Luo J, Huang H. Toward automatic quantification of knee osteoarthritis severity using improved faster R-CNN. International Journal of Computer Assisted Radiology and Surgery. 2020; 15:457-66.

[42] https://data.mendeley.com/datasets/t9ndx37v5h/1. Accessed 30 October 2023.

[43] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: inverted residuals and linear bottlenecks. In proceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 4510-20).

[44] Thomas KA, Kidziński Ł, Halilaj E, Fleming SL, Venkataraman GR, Oei EH, et al. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. Radiology: Artificial Intelligence. 2020; 2(2):1-10.

[45] Yong CW, Teo K, Murphy BP, Hum YC, Tee YK, Xia K, et al. Knee osteoarthritis severity classification with ordinal regression module. Multimedia Tools and Applications. 2021:1-3.

[46] Dharmani BC, Khatri K. Deep learning for knee osteoarthritis severity stage detection using X-Ray images. In international conference on communication systems & networks 2023 (pp. 78-83). IEEE.

[47] Li W, Feng J, Zhu D, Xiao Z, Liu J, Fang Y, et al. Nomogram model based on radiomics signatures and age to assist in the diagnosis of knee osteoarthritis. Experimental Gerontology. 2023; 171:112031.

[48] Rehman A, Raza A, Alamri FS, Alghofaily B, Saba T. Transfer learning-based smart features engineering for osteoarthritis diagnosis from knee X-ray images. IEEE Access. 2023; 11:71326-38.

Aseel Ghazwan received the B.Sc. and M.Sc. degree in Medical Engineering from Al-Nahrain University, Iraq, and the Ph.D. degree in Medical Engineering, in 2017, from Cardiff University, U.K. She is a Lecturer in Biomedical Engineering Department /College of Engineering / Al-Nahrain University, Baghdad, Iraq, and worked in the Biomechanics and Bioengineering Research Centre Versus Arthritis, Cardiff University, Cardiff, UK. Her research areas are OA

Classification, Gait analysis, Musculoskeletal modelling, Deep learning, Muscle forces, Foot deformity, and Prosthetics.
Email: aseel_ghazwan@yahoo.com

Salma Al-Qazzaz is a lecturer in Department of Physics, College of Science for Women, University of Baghdad, Iraq. She received the B.Sc. and M.Sc. degree in Medical Engineering from the Al-Nahrain University, Iraq, and the Ph.D. degree in Medical Engineering, in 2020, from Cardiff University, U.K. Her research interests are Brain Tumor Segmentation, Artificial Intelligence, Machine and Deep Learning
Email: zhraaali18@gmail.com

Ashwan A. Abdulmunem Received her Ph.D. degree in computer vision, Artificial Intelligence from Cardiff University, UK. She is an Assistnt Professor at College of Computer Science and Information Technology, University of Kerbala. Her research interests include Computer Vision and Graphics, Computational Imaging, Pattern Recognition, Artificial Intelligence, Machine and Deep Learning.
Email: ashwan.a@uokerbala.edu.iq

Appendix I

S. No.	Abbreviation	Description
1	ACL	Anterior Cruciate Ligament
2	AUC	Area Under the Curve
3	BMI	Body Mass Index
4	CNN	Convolutional Neural Network
5	2D	Two-Dimensional
6	3D	Three-Dimensional
7	DL	Deep Learning
8	DHL-I	Deep Hybrid Learning-I
9	DHL-II	Deep Hybrid Learning-I
10	FFNN	Feed Forward Neural Network
11	Grad-CAM	Gradient-Weighted Class Activation Mapping
12	HRNet	High-Resolution Network
13	KL	Kellgren and Lawrence
14	KNN	K-Nearest Neighbours
15	LAT	Lateral
16	LR	Logistic Regression
17	ML	Machine Learning
18	MOST	Multicenter Osteoarthritis Study
19	OA	Osteoarthritis
20	OAI	Osteoarthritis Initiative
21	PA	Posterior-Anterior
22	PCA	Principal Component Analysis
23	RCU	Rani Channamma University
24	RF	Random Forest
25	ReLU	Rectified Linear Units
26	RPN	Region Proposal Network
27	SGD	Stochastic Gradient Descent
28	SVM	Support Vector Machine
29	TL	Transfer Learning
30	TKR	Total Knee Replacement
31	YOLO	You Only Look Once