e-Prime - Advances in Electrical Engineering, Electronics and Energy 5 (2023) 100266

Contents lists available at ScienceDirect

e-Prime - Advances in Electrical Engineering, Electronics and Energy

journal homepage: www.elsevier.com/locate/prime

Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on Kellgren-Lawrence grading system

Vijaya Kishore V [a,*], V. Kalpana [a], G Hemanth Kumar [b]

[a] Department of ECE, School of Engineering, Mohan Babu University, Tirupati, India
[b] School of Electronics and Communication Engineering, Reva University, Bangalore, India

A R T I C L E I N F O

Keywords:
Knee osteoarthritis
Artificial intelligence
Deep learning
Prediction
Kellgrenlawrence grading system
Orthopaedic

A B S T R A C T

Osteoarthritis of the knee, also known as OA has been determined that osteoarthritis of the knee is the leading cause of activity limitations and the development of disability, particularly in people who are older. The utilisation of artificial intelligence (AI) methodologies grounded in deep learning (DL) has yielded promising outcomes in the realm of radiographic interpretation. The utilisation of deep learning in the healthcare industry has yielded remarkable outcomes and elevated the benchmark for the quality of medical treatment. This study used knee OA as a clinical scenario to compare twelve transfer learning DL models for detecting the grade of KOA from a radiograph, compared their accuracy, and determined the best model for detecting KOA. The models exhibited a range of 30% to 98% in detecting the KOA. It was determined that MobileNet was responsible for the highest level of accuracy, which came in at 98.36%. It has high training and validation accuracy. The maximum loss was observed for EfficientNetB7. DL approaches created by skilled radiologists and orthopaedic specialists could help smaller hospitals learn and make more emergency room. This would be especially helpful in situations when medical personnel may not be available.

1. Introduction

Osteoarthritis (OA) is a degenerative illness that affects all three knee compartments and progresses over ten to fifteen years. Knee infections can cause mobility issues, apprehension and swelling. Knee osteoarthritis is a degenerative joint condition that primarily affects the articular cartilage of the knee joint. This condition is characterised by symptoms such as pain, stiffness, and joint deformity. Knee osteoarthritis causes cartilage wear and osteophytes. KOA is more common amongst elderly, overweight, and inactive individuals. Clinical procedures and pathology are contingent upon early diagnosis [1,2]. The physical well-being of people as well as communities is therefore more dependant than ever on early examination and diagnosis. The three-dimensional structures of knee joints can be reflected using MRI technology [3]. Due to the high cost of the examination and the fact that MRI is only offered in large medical facilities, it is not appropriate for use in the routine diagnosis of knee osteoarthritis. Therefore, the X-ray has become the best model for the screening of knee osteoarthritis for the reason that it is risk-free, has a low cost, and is accessible to a vast number of individuals. Fig. 1 illustrates both the samples and the standards for each grade level.

The KL classification only applied within the scenario of osteoarthritis in the knee (OA), as stated in the information that was provided in their initial study [4] as in Fig. 1. In the beginning stages of the KL classification, anteroposterior radiographs of the knee were taken to describe the condition. A number between 0 and 4 was assigned as the grade for each radiograph. Radiographs are the method of choice for knee osteoarthritis (OA) classification. These radiographs make use of the Kellgren-Lawrence (KL) grading, which uses a scale from 0 to 4, where 0 represents normal, 1 represents doubtful symptoms of OA, 2 represents mild OA, 3 represents moderate OA, and 4 represents severe OA [5]. KL grades are normally issued to each knee joints within a relatively short amount of time following a doctor's examination of a knee X-ray. The expertise and attentiveness of the attending physicians have a significant impact on the reliability of the diagnosis. As an illustration, the presence of questionable joint space nodules and the likelihood of osteophytic lipping are two of the requirements for the KL grade 1 designation. When evaluating that knee joint at different points in time, even the same doctor may come to a different conclusion regarding the patient's KL grade. According to research [4], the KL intra-rater reliability falls somewhere in the range of 0.67–0.73. We have a theory that the low assurance of the grading done by physicians is

* Corresponding author.
E-mail address: kishiee@rediffmail.com (V.K. V).

https://doi.org/10.1016/j.prime.2023.100266
Received 31 May 2023; Received in revised form 31 July 2023; Accepted 20 August 2023
Available online 31 August 2023

because of the ambiguity of the test, which causes the knee joint's KL grade to be misidentified as one of the grades that are closer to it. Participants who did not have knee OA at start but developed knee OA throughout the follow-up period were considered to have had incident knee OA. In addition, participants with acute knee OA and associated knee pain were considered to have symptomatic knee OA for the purposes of this study. Those individuals were considered to have knee OA progression if they had knee OA at the start of the study and had an increase in KL score or equivalent at follow-up of one point or more in either of their knee compartments. There have been other studies that claim to have developed risk prediction models for OA [6–8]. Possibilities for primary as well as secondary disease prevention are being identified as a result of research conducted over the last twenty years into the epidemiology of knee osteoarthritis [9,10]. Multiple factors have been identified as contributors to the onset of knee osteoarthritis [11].

The term "artificial intelligence" (AI) refers to the rise of computer systems that mimic human intelligence. By utilising processing and analytical capacity that is unparalleled, we are expanding the limits of what is possible for humans. Deep learning (DL), a subset of artificial intelligence, has recently been applied to the field of medical image processing, and these developments have demonstrated some potential new techniques to enhance the interpretation of orthopaedic radiographs. Before beginning the training process, traditional machine learning (ML) often devotes a significant amount of time and effort to the extraction of features. In contrast, deep learning algorithms possess the capability to acquire knowledge of the characteristics inherent in the data, thereby resulting in enhancements throughout subsequent phases. The efficacy of this approach has been demonstrated, and it enables machine learning experts to devise novel techniques for implementing applied research and practical applications, medical image analysis via automating feature engineering. Poorly designed deep learning models can create side-channel vulnerabilities, making them open to timing attacks. PQC's immediate threats and countermeasures [12]. Lightweight cryptography [13] protects resource-constrained embedded systems [14]. Deep learning allows researchers to create lightweight cryptographic accelerators [15] solutions customised to reliable hardware architecture [16] has capabilities, improving efficiency and energy usage. Its completeness makes the binary Edwards curve immune to the exceptional points attack [17]. Deep learning may optimise SIDH computations [18]. Post-quantum cryptography protects private data, while deep learning teaches machines to spot patterns. Deep learning could boost SIKE post-quantum cryptography [19,20] and defend machine learning from quantum threats [21].

DL can assist medical professionals such as physicians and orthopaedic surgeons in the automatic interpretation of medical pictures, which has the potential to improve both the accuracy and speed of diagnostics. By minimising the amount of labour that medical professionals must do and, more crucially, by bringing objectivity to clinical assessment and choices about the necessity of surgical intervention, DL is able to eliminate human mistake caused by exhaustion and a lack of expertise, which is a major source of stress for these experts. Also, deep learning techniques that are trained on the knowledge of experienced radiologists and orthopaedic surgeons in large tertiary care facilities can spread knowledge to smaller institutions and create more capacity in emergency care, where skilled medical professionals may not be easy to find. This has the potential to significantly improve access to care. DL has been utilised effectively in a variety of orthopaedic applications, including the detection of fractures [22–25], the diagnosis of bone tumours [26], the identification of hip implant mechanical loosening [27], and the grading of osteoarthritis [27,28]. Deep learning has been the most important contributor to the overall progress that has been made in the health informatics [29]. In that, the primary focus is on the analysis of huge amounts of dataset in order to enhance medical decision support technologies and keep track of patient data for both assurance and patient access to healthcare services. To create reliable prediction models, sample size must be proportional to the number of variables involved in the model [30]. The quantity, precision, and quality of the dataset and image data affect the robustness of machine learning systems. Historically, models for the prediction of knee osteoarthritis were constructed through the use of conventional logistic regression techniques [30]. In more recent times, prediction models that include machine learning (ML) strategies were developed [30]. It is generally assumed that machine learning-based prediction models can manage large amounts of data as well as unpredictability in clinical and biological frameworks [31,32], despite the fact that there is no apparent contrast between ML methods and classic statistical methods.

According to the projections of researchers, there will be 130 million people affected by knee OA worldwide by the year 2050. However, OA of the knee can be stopped from progressing, and people's quality of life can be improved if it is detected and treated early [33]. Many of the earlier studies had shortcomings, such as using a limited number of osteoarthritis-related variables or not having access to a sufficient amount of data. As a result, research must be conducted to develop a tool on large amounts of data in order to predict the osteoarthritis, which contributes to the difficulty a disease presents for society. To the best of our understanding, only a small number of published research have utilised DL for the purpose of KOA categorization; nevertheless, a significant number of these studies comprised applications to pre-processed and highly optimised pictures [34]. Using digital X-ray images, it is possible to detect and categorise the severity of knee osteoarthritis (OA) according to the Kellgren-Lawrence (KL) grading system [35]. Eight transfer learning deep learning models' efficacy and accuracy in determining KOA severity from a radiograph [36].

Using KOA as a clinical scenario, this study compares twelve distinct DL models to identify the grade of KOA from a radiograph, assesses the accuracy of the models' and ultimately concludes that the model best suited to detecting the grade of KOA based on KL grading. The objective of this research was to assess existing prediction models for knee osteoarthritis (OA) that have been published, and to identify potential areas for further advancement in model development. The main contributions are summarized below:

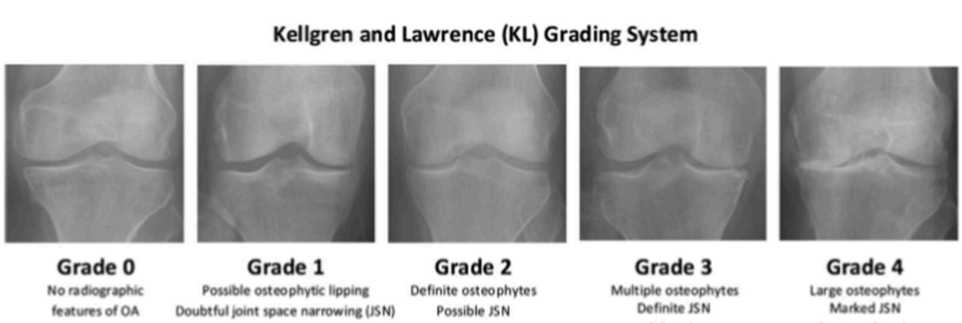

Fig. 1. KL grading for samples of knee joints.

- Model outputs were compared to independently categorized KL scale models to benchmark efficacy.
- The study allows ML extrapolation to create an automated KOA classification tool to help healthcare practitioners make better decisions.
- Compare twelve transfer learning DL models for KOA grade detection from radiographs and select the best model.
- Individual models can anticipate risk and encourage risk reduction.

Twelve different models were utilised in order to evaluate the effectiveness of deep learning in correctly identifying radiography for KOA. These models include DenseNet169, DenseNet201, EfficientNetB7, InceptionResNetV2, InceptionV3, MobileNet, MobileNetV2, NASNet-Mobile, Re sNet152V2, VGG19, and Xception. Of the total 2899 images used, 70% were put towards the training phase of the model, 10% were put towards the testing phase, and 20% were put towards the accuracy testing and validation phases.

2. Literature review

Due to the speedier diagnosis and higher rate of early identification, there is a great demand for an automatic grading of knee severity. This is due to a shortage of radiologists, especially in remote areas and the lengthy process involved in analysing X-ray pictures of the knee. Because of this, a great number of computer-aided diagnosis (CAD)-based radiology methods are suggested to identify and evaluate knee osteoarthritis [37–39]. The utilisation of machine learning and deep learning techniques in the field of medical imaging has experienced an increase in frequency in recent years. These methodologies allow for the problem-solving of classification [40,41], detecting, and other related problems with no need for the expertise of a radiologist [42]. Classification of the changes that occur in the gaps between the knee joints and the growth of osteophytes can be used to determine the level of severity of osteoarthritis of the knee. This can be performed by observing the variations in the knee joints [43,44]. A developed scoring system for predicting radiographic and symptomatic knee OA [45] risks by making use of artificial neural networks (ANN) from a survey. In a multipurpose bio-medical image classifier it was referred [46] to categorise knee OA images and for CAD-based early identification of knee OA [47–49].

To be more specific the design and implementation of DL-based detection algorithms were effective in determining the severity of knee osteoarthritis [50]. In addition to this, even though it doesn't involve human feature engineering, they demonstrate performance in the biological sector of X-ray analysis. This is due to the fact that feature engineering occurs inadvertently in training phase, when the model is adjusting its system parameters to suit the relevant data. On the other hand, in order to produce the desired results, the data must be modified beforehand using a specialised feature engineering or training technique in order to employ any of the basic machine learning algorithms. When compared the traditional ML algorithms, amount of computational power and resources that DL algorithms typically demand can be considered excessive. In addition, it leads to overfitting if it is given too little data to work with. Additionally, there are several varieties of DL, such as TL based CNN, Resnet, Inception, that produce impressive results in computer vision. In addition, an innovative method for evaluating the severity of knee OA based on X-ray images was also proposed [51]. When the statistical findings of this approach were compared to those of other ways that were currently in use, it was found that this method had advancements in accuracy rate, recall, F1 score, and precision.

Based on the KL grading method, a novel strategy has been developed to evaluate osteoarthritis (OA) in knee X-rays. This technique employs neural networks to adopt supervised procedures for exact identification from the original X-rays [52]. They said the method could be used as assistance to physicians in the process of diagnosis that is very reliable. The KL grading method was followed alongside two CNNs to determine the level of knee osteoarthritis [53]. The images of the knee joint that were detected were then classified by making use of modified ordinal loss evaluation and then CNN is performed. These CNNs included different versions of YOLO, DenseNet, VGG, ResNet, and InceptionV3. [54] A CAD method for the faster identification of knee osteoarthritis that utilises digital knee X-ray pictures and few ML algorithms are described.

Convolutional neural networks (CNNs) lately surpassed numerous methods in image processing. CNNs are taught efficient for feature sets that are ideally suit for fine-grained classification [55]. We previously showed that utilising transfer learning, pre-trained CNNs like the VGG 16-Layers network [44], the VGG-M-128 network [56], and the BVLC reference CaffeNet [5] trained on the ImageNet LSVRC dataset [53] may be improved for knee OA image classification [57]. To enhance the effectiveness of KOA diagnosis [58], a model was developed utilising the DenseNet169 deep learning (DL) technique for fine-tuning. In the past, [59] suggested employing a template match as a way to automatically recognise and extract the knee joints from an image. This was done by comparing the image to an existing template. It takes a long time for this method to process large datasets like OAI, and it has poor accuracy and precision when it comes to detecting knee joints.

In general, DL and ML approaches multiple bone illnesses have been diagnosed using in the methods that have been referred by previous works that have been published in [60,61]. The development of research into face masque identification places an emphasis on deep learning and deep transfer learning approaches [62]. After classifying VSDM methods into two broad groups, that rely on hand-crafted features and those that use deep learning, a thorough study is conducted [63,64] presents a comprehensive evaluation of several AI-based ML and DL algorithms currently being applied in PD diagnosis and their impact on future areas of study. A well-structured taxonomy is used in a comprehensive systematic assessment of AI-big data analytics to provide an overview of current frameworks [65]. When it comes to classification into binary categories, these methods have proven to be effective [66,67]. On the other hand, owing to the fact that they are not applicable to the classification tasks of knee OA, this method achieved an accuracy of 68% [42].

2.1. Evaluation of osteoarthritis of the knee and its progression

Contact forces in the tibiofemoral compartment can be predicted using machine learning. However, due to individual variances in anthropometry, making reliable projections is difficult. This research aimed to create transfer learning models that may predict medial KCF in rehabilitating patients with knee valgus [68]. This study explores the application of transfer learning models, specifically sequential convolutional neural networks (CNNs), Visual Geometry Group 16 (VGG-16), and Residual Neural Network 50 (ResNet-50), in the early detection of osteoarthritis using knee X-ray images. Based on our analysis, it was determined that all of the proposed models demonstrated a predictive accuracy level exceeding 90% in the detection of osteoarthritis [69]. Recently, potential ways to ease these input data to generalisation difficulties have been presented, including deep transfer learning (DTL) and deep domain adaption (DDA). By transferring information from one domain to another or from one task to another, they can (i) make training easier, (ii) boost the generalizability of ML and DL models, and (iii) get around data scarcity issues [70]. Using heterogeneous clinical MR images and a short training data set, a transfer learning model segmented knee cartilage like a human. The model also passed cross validation and vendor image tests. Fully automated cartilage segmentation of clinical knee MR images was possible [71].

Therefore, it is challenging to recommend a useful instrument for the early categorization of osteoarthritis of the knee. Therefore, for the purpose of enhancing performance in classification, an approach that makes use of both types of learning is essential.

3. Materials and methods

The diagnostic strategy adopted by the KOA was founded on the utilisation of radiographic examinations that had been obtained in the past. The analyses were carried out by a neural network, using the KL system being used to evaluate both the existence of KOA and the severity of the condition [72]. The KL system was utilised for assessing both the occurrence of KOA and also the severity of the condition. Artificial intelligence can recognise patterns in pictures by analysing incoming data and using associated ongoing learning. It receives its input, which consists of the radiographic images, and the information that is supposed to be labelled on its output is the classifications of OA grade [73]. There is a wide range of variation in the levels of precision and accuracy achieved through learning when using real-time datasets and simulation datasets [74,75].

3.1. Data group

All of the medical images were anonymized, so no personal information about the patients was included. Following the completion of the de-identification process, the inclusion and exclusion criteria were applied in a manner that was in accordance with international standards.

3.2. Pre-processing

Twenty-six hundred and eighteen X-ray images were used to train the models. The dataset was divided into three parts: a training set, a test set, and a validation set, using a split ratio of 70:10:20. Initially, the model was trained using 70% of the available images. After that, 10% of the data was utilised to validate the models and 20% was used for testing. Complete knee radiographs were inputted into a fully interpretable model, which then ascribed severity grades to KOA based on the data (Fig. 1).

3.3. Input images

The images that were collected for each category were processed by the algorithm. Each radiograph in DICOM format was automatically cropped to the active imaging area, removing any black borders and reducing the image size to a maximum of 224 pixels. The dimensions of the image were preserved by padding the rectangular image to a square of 224 by 224 pixels. Due to their higher resolution and size, larger medical images, such as those from DICOM, are typically computationally more expensive to process. By resizing images to 224 $\times$ 224 pixels. Especially when operating on GPUs, deep learning models have memory limitations. Ensuring that the images and model parameters fit within the GPU memory by resizing them to 224 $\times$ 224 enables efficient training and inference. It assists in standardising the AI model's input data and assures compatibility with pre-trained models and frameworks. It is common and practicable for AI models to convert DICOM images to 224 $\times$ 224, especially when using pre-trained CNNs and addressing computational and memory constraints. As shown in Fig. 1, the images are categorised into five classes. The distribution of the picture data amongst the training, testing, and validation subsets is outlined in Table 1, which contains the KL grades.

3.4. Necessity of deep learning

AI models employ a variety of algorithms to first analyse the information as it exists inside a dataset and then acquire this knowledge with the specific intention of resolving a variety of business difficulties and overarching concerns. A DL model refers to a piece of software that has been taught to recognise specific kinds of patterns by being exposed to a specific set of data, often known as the training set. Transfer learning, on the other hand, is a method that takes knowledge gained from completing a variety of tasks and applies it to new endeavours for which it has not been prepared. The transmission of knowledge can help to accelerate training, which helps to minimise overtraining and improves the final performance that can be achieved.

3.5. Proposed model development

Python was chosen as the programming language for the development of DL models in this work, and the Google Colab platform were used to compile the code on a GPU configuration. This enabled the models to be processed more quickly. The proposed algorithm is composed of the following steps:

1. Getting the Knee OA data set, bringing in library files and making a personalized dataset.
2. Pre-processing the images – Resizing and normalizing the images with dimensions 224 $\times$ 224.
3. Prepare Data Loader.
4. Creating augment and data generator routine functions for use with testing, training, and validation datasets.
5. Build and compile the models.
6. Training and testing using 12 different DL models.
7. KL grading and computing the values of accuracy and loss.

4. Results

In the present study, twelve DL models are applied on the data sets as mentioned in Table 1 to compare the following parameters, (1) Accuracy is a measure that tells us which fractions have been correctly classified. For the sake of validation and testing, the accuracy was computed. (2) Loss, which reveals an inaccuracy in the model's ability to predict the future. For the sake of validation and testing, the loss was computed. The purpose of this comparison was to investigate which model performed superiorly with relation to the picture categorization of KOA grades.

4.1. Comparisons of the learning models

Table 2 presents a comparison of the several models that were utilised for transfer learning models for KOA. This comparison takes into account the validation and training of the two metrics accuracy and loss. The models exhibited a range of 30% to 98% in detecting the KOA. It was determined that MobileNet was responsible for the highest level of

Table 1
The partitioning of image data in to training subset, testing subset, and validation subset, corresponding to KL grades.

KL grades	Training		Testing		Validation	
	Samples n	Proportion,%	Samples n	Proportion,%	Samples n	Proportion,%
0	515	31	73	30.9	219	30.9
1	478	29	68	28.8	204	28.8
2	233	14	33	13.9	99	13.9
3	222	13.5	32	13.6	96	13.7
4	207	12.5	30	12.8	90	12.7
Total	1655	100	236	100	708	100

Table 2
Comparison of validation and training accuracy and loss of twelve different learning models.

Model Name	Val_Accuracy	Train_Accuracy	Val_Loss	Train_Loss
DenseNet169	0.937	0.9582	0.567	0.3149
DenseNet201	0.9455	0.9333	0.4538	0.7055
EfficientNetB7	0.3048	0.2521	6.0686	17.939
InceptionResNetV2	0.9097	0.9103	0.7195	0.7296
InceptionV3	0.8994	0.9503	0.8727	0.4431
MobileNet	0.9836	0.9624	0.5168	0.5164
MobileNetV2	0.8345	0.9285	2.8252	1.0957
NASNetLarge	0.9721	0.9582	2.1047	2.0527
NASNetMobile	0.9388	0.9279	0.4111	0.4522
ResNet152V2	0.9588	0.9364	0.6546	1.0154
VGG19	0.883	0.8073	0.3389	0.489
Xception	0.8497	0.9042	1.7289	0.6782

accuracy, which came in at 98.36%. It has high training and validation accuracy. The maximum loss was observed for EfficientNetB7. According to the findings, MobileNet was the most effective model to use when attempting to create a KOA image categorization system that was based on KL-grade. After selecting the best medical image categorization models for OA, these models can be implemented to automate preliminary X-ray processing and step selection. An experienced person should manually categorise a large number of images for each image classification project. This input will help model training. This study prepares the orthopaedic department for an end-to-end machine recommendation engine. This engine will provide enough inputs for OA surgeons to evaluate OA surgery, patient therapy, and patient prioritising based on model output.

Figs. 2–5 depicts the twelve unique models and their corresponding confusion matrices, as well as the loss and accuracy metrics plotted

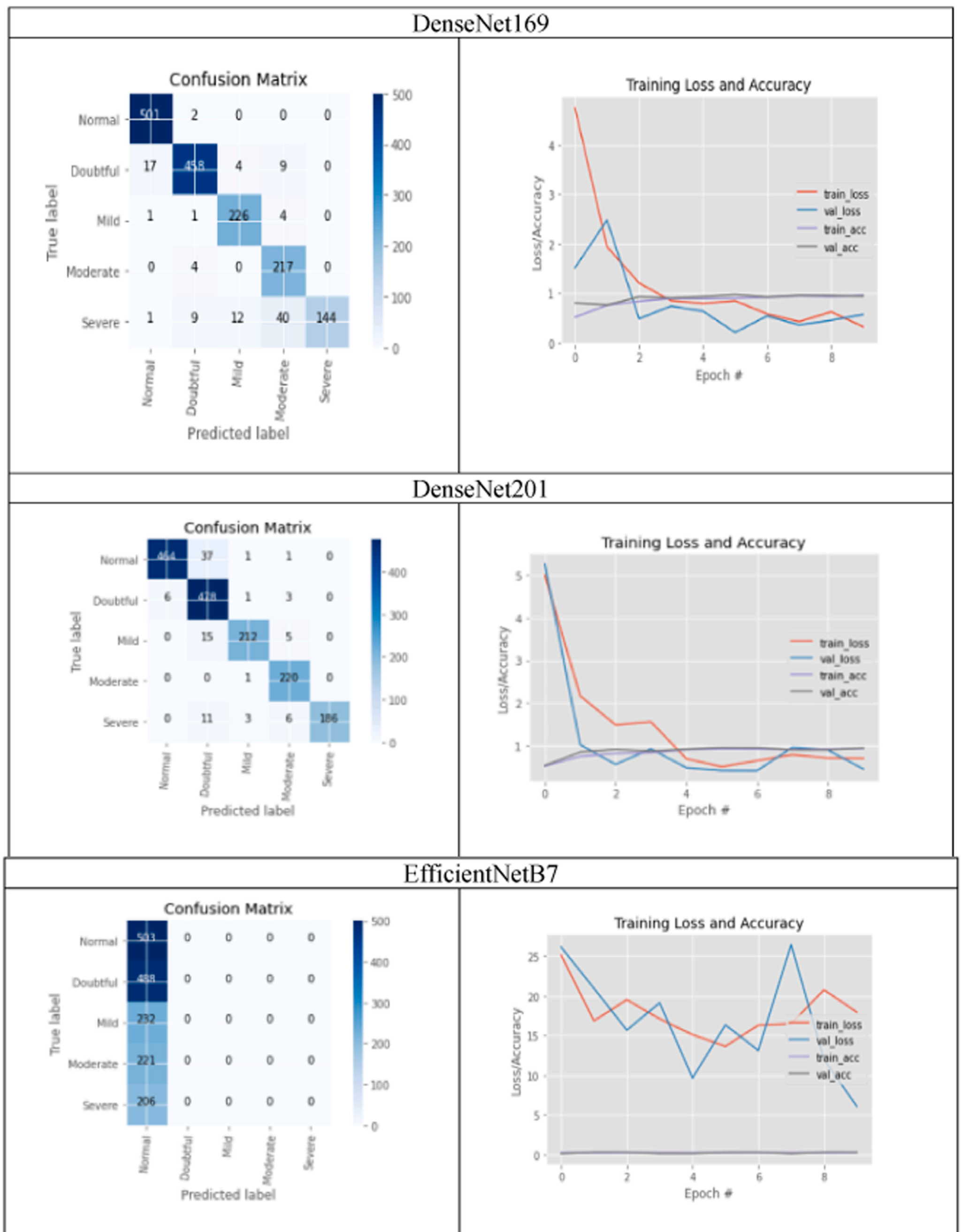

Fig. 2. Confusion matrices, as well as the loss and accuracy metrics plotted against the number of epochs for DenseNet169, DensNet201 and EfficientNetB7 models.

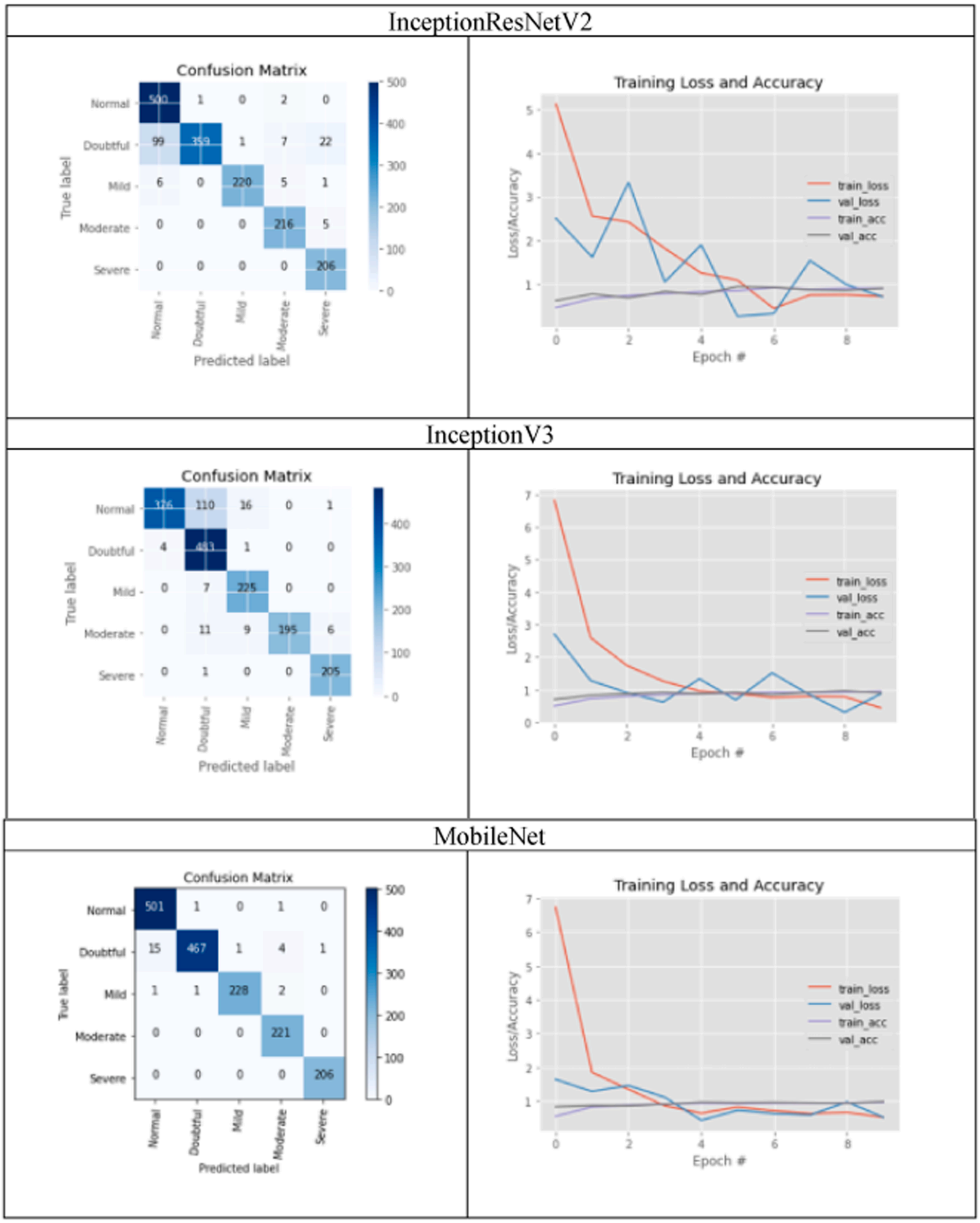

Fig. 3. Confusion matrices, as well as the loss and accuracy metrics plotted against the number of epochs for InceptionResNetV2, InceptionV3 and MobileNet models.

against the number of epochs for each model. Model selection, evaluation, and fine-tuning require plotting loss and accuracy vs. epochs. It helps to comprehend how each model learns and performs, so one can choose the right model for a task or optimise it. Loss and accuracy plots reveal overfitting or underfitting. Underfitting occurs when the model fails to capture data patterns, while overfitting occurs when the model performs well on training data but badly on unknown data. Seeing the model's performance increase over time might be reassuring because it indicates the amount of training that is still required to reach the desired level of accuracy. By monitoring the loss's response to different learning rates, we can also fine-tune the learning process.

5. Discussions

Previous systematic reviews have documented occupational activities as potential risk factors for knee osteoarthritis (OA). However, a limited number of studies have provided quantitative data, potentially attributed to the diverse range of risk factors and measurements found in the existing literature. The utilisation of DL algorithms enables the differentiation between symptomatic osteoarthritis (OA) requiring treatment and radiological OA without symptoms that only necessitate observation. The implementation of deep learning in the healthcare industry has demonstrated remarkable advancements and significantly enhanced the standard of patient care. Numerous organisations have dedicated substantial resources to the development of deep learning models that provide predictions regarding hospitalised patients and facilitate the management of patient data and outcomes. A crucial factor to ensure reliable predictive performance of prediction models is the appropriate consideration of sample size relative to the number of predictors incorporated in the model. The size and quality of the dataset, as well as the quality of image data, have a noteworthy influence, especially on the resilience of machine learning-based methodologies. The future of healthcare presents an intriguing landscape, with the emergence of artificial intelligence (AI) and machine learning (ML) offering

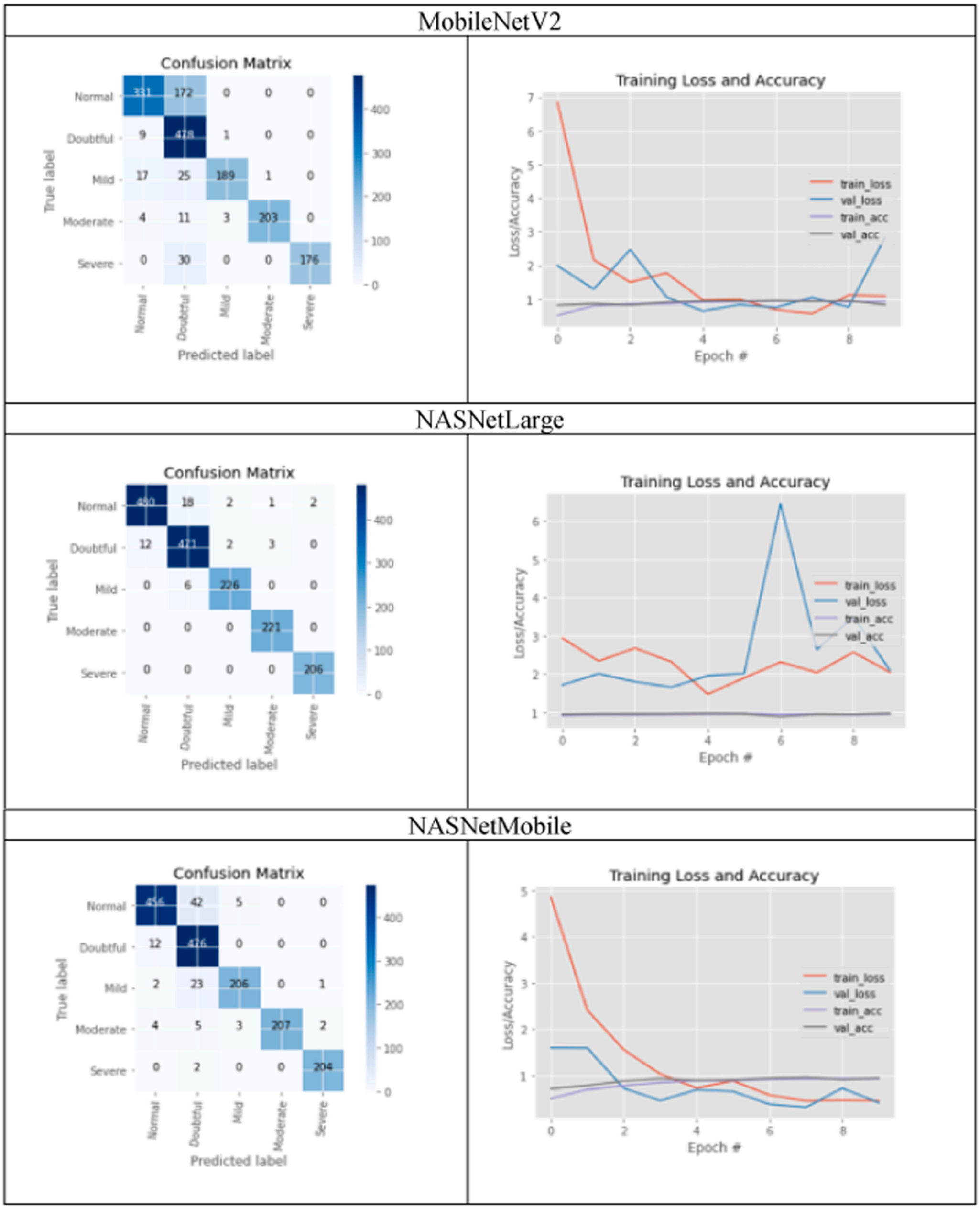

Fig. 4. Confusion matrices, as well as the loss and accuracy metrics plotted against the number of epochs for MobileNetV2, NASNetLarge and NASNetMobile models.

promising avenues for developing tailored solutions to address specific healthcare needs. Having a continuous, reliable, fast, effective, and accurate machine would be advantageous in situations where there is no access to a radiology expert on-site for preliminary diagnosis. Artificial intelligence (AI) solutions have the capability to integrate image data and radiology text reports to enhance image analysis and facilitate ongoing learning of the network. The utilisation of artificial intelligence (AI) in the field of orthopaedics has predominantly centred around the integration of deep learning (DL) techniques with clinical images. Transfer learning is a technique that utilises data obtained from one set of distributions to enhance its learning and make predictions or classifications on other sets.

6. Conclusions

According to the findings of the current research, a neural network may be educated to categorise and diagnose KOA severity and direction in accordance with the KL grading scale. According to the findings of the study, the accuracy of the MobileNet model for determining the degree of KOA based on KL grading was good, coming in at 98.3%. Because there were no tools and methods of this kind available in the past, it was difficult for experts working in the healthcare industry to collect and analyse vast amounts of data for the purpose of classification, analysis, and the prediction of therapies. However, ever with the development of ML, it has been significantly more accessible. The findings of this study have a variety of potential applications, including the development of new orthopaedic technologies in the future. The study's goals include enhancing the accuracy, which can be accomplished with more image and the characteristics of the algorithm. The implications of this study extend more than just the straightforward categorization of KOA. The research paves the path for extending what is learned from KOA classification using DL to construct a computerised KOA classification tool and provide healthcare practitioners with the ability to make more informed decisions. This study contributes to the exploration of trends

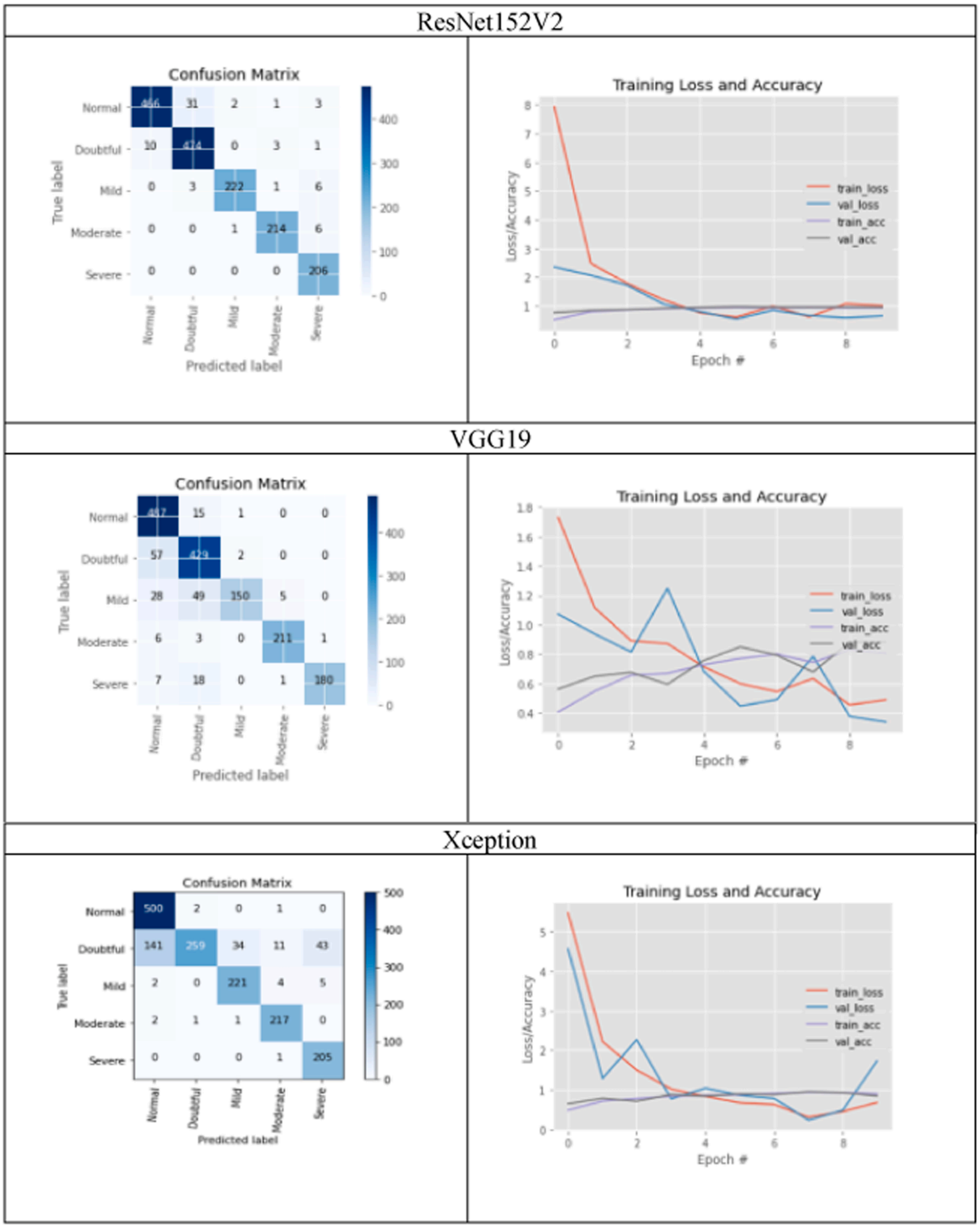

Fig. 5. Confusion matrices, as well as the loss and accuracy metrics plotted against the number of epochs for ResNet152V2, VGG19 and Xception models.

and developments in the field of AI-based medical picture categorization, as well as to the improvement of the accuracy of current studies and trials in this area. When it comes to patients' diagnosis and treatment options, medical professionals utilising AI technologies are able to make more informed judgements, which ultimately leads to improved and more efficient healthcare services.

Future prospects

Knee osteoarthritis (OA) risk variables and radiographic markers from knee x-rays are complex, hence more study is needed to build and validate knee OA prediction models in different populations. Clinical decision-making would be greatly aided by the incorporation of such models into everyday practise. As a result, it is necessary to modify the models and the parameters associated with them in order to investigate the effect that tuning and epochs have on the level of the model's performance or accuracy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

Data availability

Data will be made available on request.

References

[1] H. Hsu, R.M. Siwiec, Knee arthroplasty. StatPearls [Internet], StatPearls Publishing, Treasure Island, 2022 [PMID: 29939691].
[2] O. Aweid, Z. Haider, A. Saed, Y. Kalairajah, Treatment modalities for hip and knee osteoarthritis: a systematic review of safety, J. Orthop. Surg. (26) (2018).

[3] N.C. Nacey, M.G. Geeslin, G.W. Miller, J.L. Pierce, Magnetic resonance imaging of the knee: an overview and update of conventional and state of the art imaging, J. Magn. Reson. Imaging (2017).
[4] B. Guan, F. Liu, A. Haj-Mirzaian, et al., Deep learning risk assessment models for predicting progression of radiographic medial joint space loss over a 48-month follow-up period, Osteoarthr. Cartil. 29 (4) (2020) 428–437.
[5] J. Kellgren, J. Lawrence, Radiological assessment of osteoarthrosis, annals of the rheumatic diseases (1957) (16) (4),494.
[6] G. Schett, S. Kiechl, E. Bonora, et al., Vascular cell adhesion molecule 1 as a predictor of severe osteoarthritis of the hip and knee joints, Arthritis Rheum 60 (2009) 2381–2389.
[7] H. Takahashi, M. Nakajima, K. Ozaki, et al., Prediction model for knee osteoarthritis based on genetic and clinical information, Arthritis Res. Ther. 12 (R187) (2010).
[8] E.W. Watt, E. Watt, A.A. Bui, et al., Evaluation of a dynamic Bayesian belief network to predict osteoarthritic knee pain using data from the osteoarthritis initiative, in: Proceedings of the AMIA Annual Symposium Proceeding, 2008, pp. 788–792.
[9] W. Zhang, D.F. McWilliams, S.L. Ingham, S.A. Doherty, S. Muthuri, et al., Nottingham knee osteoarthritis risk prediction models, Ann. Rheum Dis. 70 (9) (2011) 1599–1604.
[10] W. Zhang, Risk factors of knee osteoarthritis–excellent evidence but little has been done, Osteoarthr. Cartil. 18 (2010) 1–2.
[11] M. Blagojevic, C. Jinks, A. Jeffery, et al., Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis, Osteoarthr. Cartil. 18 (2010) 24–33.
[12] C.C., Alvaro & J. Kaur & M.K., Mehran & R. Azarderakhsh, Algorithmic security is insufficient: a comprehensive survey on implementation attacks haunting post-quantum security (2023). doi:10.36227/techrxiv.23071079.v1.
[13] B.S. Siavash, M.K. Mehran, R. Azarderakhsh, C.Y. Lee, Dual-basis superserial multipliers for secure applications and lightweight cryptographic architectures. Circuits and systems II: express briefs, IEEE Trans. 61 (2014) 125–129, https://doi.org/10.1109/TCSII.2013.2291075, on.
[14] J. Kaur & C.C., Alvaro & M. Kermani & R. Azarderakhsh, A comprehensive survey on the implementations, attacks, and countermeasures of the current NIST lightweight cryptography standard (2023). doi:10.36227/techrxiv.22970855.
[15] M. Bisheh-Niasar, R. Azarderakhsh, M. Mozaffari-Kermani, Cryptographic accelerators for digital signature based on Ed25519, IEEE Trans. Very Large Scale Integr. VLSI Syst. 29 (7) (2021) 1297–1305, https://doi.org/10.1109/TVLSI.2021.3077885.
[16] M.K. Mehran, R. Azarderakhsh, Reliable architecture-oblivious error detection schemes for secure cryptographic GCM structures, IEEE Trans. Reliab. (2019) 1–9, https://doi.org/10.1109/TR.2018.2882484.
[17] B. Koziel, R. Azarderakhsh, M. Mozaffari-Kermani, Low-resource and fast binary Edwards curves cryptography. In: A. Biryukov, V. Goyal (eds) Progress in Cryptology – INDOCRYPT 2015. INDOCRYPT 2015. Lecture Notes in Computer Science (2015) (9462). doi:10.1007/978-3-319-26617-6_19.
[18] A. Jalali, R. Azarderakhsh, M.M. Kermani, D. Jao, Supersingular isogeny Diffie-Hellman key exchange on 64-Bit ARM, IEEE Trans. Dependable Secure Comput. 16 (5) (2023) 902–912, https://doi.org/10.1109/TDSC.2017.2723891, 1029.
[19] M. Anastasova, R. Azarderakhsh, M. Mozaffari-Kermani, Fast strategies for the implementation of SIKE Round 3 on ARM Cortex-M4, IEEE Trans. Circuits Syst. I Reg. Pap. 68 (10) (2021) 4129–4141.
[20] S. Subramanian, M. Mozaffari-Kermani, R. Azarderakhsh, M. Nojoumian, Reliable hardware architectures for cryptographic block ciphers LED and HIGHT, IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 36 (10) (2017) 1750–1758, https://doi.org/10.1109/TCAD.2017.2661811.
[21] A. Sarker, M.M. Kermani, R. Azarderakhsh, Error detection architectures for ring polynomial multiplication and modular reduction of ring-LWE in benchmarked on ASIC, IEEE Trans. Reliab. 70 (1) (2021) 362–370.
[22] T. Urakawa, Y. Tanaka, S. Goto, H. Matsuzawa, K. Watanabe, N. Endo, Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network, Skeletal Radiol. 48 (2019) 239–244, https://doi.org/10.1007/s00256-018-3016-3. PMID: 29955910.
[23] D.H. Kim, T. MacKinnon, Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks, Clin. Radiol. 73 (2018) 439–445, https://doi.org/10.1016/j.crad.2017.11.015.
[24] J. Olczak, N. Fahlberg, A. Maki, A.S. Razavian, A. Jilert, A. Stark, O. Sköldenberg, M. Gordon, Artificial intelligence for analyzing orthopedic trauma radiographs, Acta Orthop. 88 (2017) 581–586, https://doi.org/10.1080/17453674.2017.1344459. PMID: 28681679.
[25] R. Lindsey, A. Daluiski, S. Chopra, A. Lachapelle, M. Mozer, S. Sicular, D. Hanel, M. Gardner, A. Gupta, R. Hotchkiss, H. Potter, Deep neural network improves fracture detection by clinicians, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 11591–11596, https://doi.org/10.1073/pnas.1806905115. PMID: 30348771.
[26] B.H. Do, C. Langlotz, C.F. Beaulieu, Bone tumor diagnosis using a naïve Bayesian model of demographic and radiographic features, J. Digit Imaging 30 (2017) 640–647, https://doi.org/10.1007/s10278-017-0001-7. PMID: 28752323.
[27] J. Langhorn, A. Borjali, E. Hippensteel, W. Nelson, B. Raeymaekers, Microtextured CoCrMo alloy for use in metal-on-polyethylene prosthetic joint bearings: multi-directional wear and corrosion measurements, Tribol. Int. 124 (2018) 178–183, https://doi.org/10.1016/j.triboint.2018.04.007.
[28] Y. Xue, R. Zhang, Y. Deng, K. Chen, T. Jiang, A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis, PLoS One 12 (2017), https://doi.org/10.1371/journal.pone.0178992.
[29] J. Lim, J. Kim, S. Cheon, A deep neural network-based method for early detection of osteoarthritis using statistical data, Int. J. Environ. Res. Public Health 16 (7) (2019) 1281.
[30] T. Ramazanian, S. Fu, S. Sohn, M.J. Taunton, H.M. Kremers, Prediction models for knee osteoarthritis: review of current models and future directions, Arch. Bone Jt. Surg. 11 (1) (2023) 1–11.
[31] A. Jamshidi, J.P. Pelletier, J. Martel-Pelletier, Machine-learning-based patient-specific prediction models for knee osteoarthritis, Nat. Rev. Rheumatol. 15 (1) (2019) 49–60.
[32] L. Chen, Overview of clinical prediction models, Ann. Transl. Med. 8 (4) (2020) 71.
[33] Y. Wang, X. Wang, T. Gao, L. Du, W. Liu, An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative, J. Healthc. Eng (2021).
[34] A. Tiulpin, S. Klein, S.M.A. Bierma-Zeinstra, et al., Multimodal machine learning-based knee osteoarthritis progression prediction from plain radiographs and clinical data, Sci. Rep. 9 (2019), https://doi.org/10.1038/s41598-019-56527-3.
[35] S.S. Abdullah, M.P. Rajasekaran, Automatic detection and classification of knee osteoarthritis using deep learning approach, Radiol. Med. 127 (4) (2022) 398–406, https://doi.org/10.1007/s11547-022-01476-7.
[36] A. Tiwari, M. Poduval, V. Bagaria, Evaluation of artificial intelligence models for osteoarthritis of the knee using deep learning algorithms for orthopedic radiographs, World J. Orthop. 13 (6) (2022) 603–614, https://doi.org/10.5312/wjo.v13.i6.603.
[37] L. Anifah, I.K.E. Purnama, M. Hariadi, M.H. Purnomo, osteoarthritis classification using self organizing map based on gabor kernel and contrast-limited adaptive histogram equalization, Open Biomed. Eng. J 7 (2013) 18–28.
[38] M. Kotti, L.D. Duffell, A.A. Faisal, A.H. McGregor, Detecting knee osteoarthritis, and its discriminating parameters using random forests, Med. Eng. Phys. 43 (2017) 19–29.
[39] R.T. Wahyuningrum, L. Anifah, I.K.E. Purnama, M.H. Purnomo, A novel hybrid of S2DPCA and SVM for knee osteoarthritis classification, in: Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, CIVEMSA, Budapest, Hungary, 2016, pp. 27–29.
[40] M. Varma, M. Lu, R. Gardner, J. Dunnmon, N. Khandwala, P. Rajpurkar, J. Long, C. Beaulieu, K. Shpanskaya, L. Fei-Fei, et al., Automated abnormality detection in lower extremity radiographs using deep learning, Nat. Mach. Intell. 1 (2019) 578–583.
[41] M.K. Abd Ghani, M.A. Mohammed, N. Arunkumar, S.A. Mostafa, D.A. Ibrahim, M. K. Abdullah, M.M. Jaber, E. Abdulhay, G. Ramirez-Gonzalez, M.A. Burhanuddin, Decision-level fusion scheme for nasopharyngeal carcinoma identification using machine learning techniques, Neural Comput. Appl. 32 (2020) 625–638.
[42] D. Wang, J. Mo, G. Zhou, L. Xu, Y. Liu, An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest x-ray images, PLoS One 15 (2020), e0242535.
[43] H.S. Gan, M.H. Ramlee, A.A. Wahab, Y.S. Lee, A. Shimizu, From classical to deep learning: review on cartilage and bone segmentation techniques in knee osteoarthritis research, Artif. Intell. Rev. 2020 (2021) 1–50.
[44] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[45] F. Eckstein, W. Wirth, A. Culvenor, Osteoarthritis year in review 2020: imaging, Osteoarthr. Cartil. 29 (2021).
[46] F. Razmjoo, F. Liu, A.M. Caliva, S. Martinez, Majumdar, V. Pedoia, Semi-supervised graph-based deep learning for multi-modal prediction of knee osteoarthritis incidence, Osteoarthr. Cartil. 28 (2020) S305–S306.
[47] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, S. Saarakkala, Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach, Sci. Rep. (2018) 17–27.
[48] A. Tack, A. Mukhopadhyay, S. Zachow, Knee menisci segmentation using convolutional neural networks: data from the osteoarthritis initiative, Osteoarthr. Cartil. 26 (5) (2018) 680–688.
[49] S. Ebrahimkhani, M.H. Jaward, F.M. Cicuttini, A. Dharmaratne, Y. Wang, A.G.S. de Herrera, A review on segmentation of knee articular cartilage: from conventional methods towards deep learning, Artif. Intell. Med. 106 (2020).
[50] M. Varma, M. Lu, R. Gardner, J. Dunnmon, N. Khandwala, P. Rajpurkar, J. Long, C. Beaulieu, K. Shpanskaya, L. Fei-Fei, et al., Automated abnormality detection in lower extremity radiographs using deep learning, Nat. Mach. Intell 1 (2019) 578–583.
[51] J. Antony, K. McGuinness, K. Moran, N.E. O'Connor, Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks, in: Proceedings of the International Conference on Machine Learning and Data Mining in Pattern Recognition, Berlin/Heidelberg, Germany, Springer, 2017, pp. 376–390.
[52] B. Norman, V. Pedoia, A. Noworolski, T.M. Link, S. Majumdar, Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs, J. Digit. Imaging 32 (3) (2019) 471–477.
[53] P. Chen, L. Gao, X. Shi, K. Allen, L. Yang, Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss, Comput. Med. Imaging Graph 75 (2019) 84–92.
[54] A. Brahim, R. Jennane, R. Riad, T. Janvier, L. Khedher, H. Toumi, E. Lespessailles, A decision support tool for early detection of knee osteoarthritis using x-ray imaging and machine learning: data from the osteoarthritis initiative, Comput. Med. Imaging Graph 73 (2019) 11–18.

[55] H.H. Nguyen, S. Saarakkala, A. Tiulpin, Deep semi-supervised learning for knee osteoarthritis severity assessment from plain radiographs, Osteoarthr. Cartil. 28 (2020) S311–S312.
[56] T. Neogi, The epidemiology and impact of pain in osteoarthritis, Osteoarthr. Cartil. 21 (9) (2013) 1145–1153.
[57] P.G. Conaghan, M. Porcheret, S.R. Kingsbury, A. Gammon, A. Soni, M. Hurley, M. P. Rayman, J. Barlow, R.G. Hull, J. Cumming, et al., Impact and therapy of osteoarthritis: the arthritis care a nation 2012 survey, Clin. Rheumatol. 34 (9) (2015) 1581–1588.
[58] S.A. El-Ghany, M. Elmogy, A.A. Abd El-Aziz, A fully automatic fine tuned deep learning model for knee osteoarthritis detection and progression analysis, Egypt. Inform. J. 24 (2) (2023) 229–240.
[59] S. Ebrahimkhani, M.H. Jaward, F.M. Cicuttini, A. Dharmaratne, Y. Wang, A.G.S. de Herrera, A review on segmentation of knee articular cartilage: from conventional methods towards deep learning, Artif. Intell. Med. (106) (2020). Article ID 101851.
[60] E. Christodoulou, S. Moustakidis, N. Papandrianos, D. Tsaopoulos, E. Papageorgiou, Exploring deep learning capabilities in knee osteoarthritis case study for classification, in: Proceedings of the 10th International Conference on Information, Intelligence, Systems and Applications (IISA), Achaia, Greece, IEEE, 2019, pp. 1–6.
[61] Y. Du, R. Almajalid, J. Shan, M. Zhang, A novel method to predict knee osteoarthritis progression on MRI using machine learning methods, IEEE Trans. Nanobiosci. 17 (2018) 228–236.
[62] Y. Himeur, S. Al-Maadeed, I. Varlamis, N. Al-Maadeed, K. Abualsaud, A. Mohamed, Face mask detection in smart cities using deep and transfer learning: lessons learned from the COVID-19 pandemic, Systems 11 (2023) 107, https://doi.org/10.3390/systems11020107.
[63] Y. Himeur, S. Al-Maadeed, N. Almaadeed, K. Abualsaud, A. Mohamed, T. Khattab, O. Elharrouss, Deep visual social distancing monitoring to combat COVID-19: a comprehensive survey, Sustain Cities Soc. 85 (2022), 104064, https://doi.org/10.1016/j.scs.2022.104064.
[64] S. Dixit, K. Bohre, Y. Singh, Y. Himeur, W. Mansoor, S. Atalla, K. Srinivasan, A comprehensive review on AI-enabled models for Parkinson's disease diagnosis, Electronics 12 (2023) 783, https://doi.org/10.3390/electronics12040783 (Basel).
[65] Y. Himeur, M. Elnour, F. Fadli, et al., AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives, Artif. Intell. Rev. 56 (2023) 4929–5021, https://doi.org/10.1007/s10462-022-10286-2.
[66] J. Hirvasniemi, J. Runhaar, R.A. Van der Heijden, M. Zokaeinikoo, M. Yang, X. Li, J. Tan, H.R. Rajamohan, Y. Zhou, C.M. Deniz, The KNee Osteoarthritis prediction (knoap2020) challenge: an image analysis challenge to predict incident symptomatic radiographic knee osteoarthritis from MRI and X-Ray images, Osteoarthr. Cartil (2022).
[67] M A. Sharma, D.P. Yadav, H. Garg, M. Kumar, B. Sharma, D. Koundal, Bone cancer detection using feature extraction based machine learning model, Comput. Math. Methods Med. (2021).
[68] J. Zou, X. Zhang, Y. Zhang, J. Li, Z. Jin, Prediction on the medial knee contact force in patients with knee valgus using transfer learning approaches: application to rehabilitation gaits [published online ahead of print, 2022 Sep 17], Comput. Biol. Med. 150 (2022), 106099, https://doi.org/10.1016/j.compbiomed.2022.106099.
[69] 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 11 (9) (2023) 1206, https://doi.org/10.3390/healthcare11091206.
[70] Y. Himeur, S. Al-Maadeed, H. Kheddar, N. Al-Maadeed, K. Abualsaud, A. Mohamed, T. Khattab, Video surveillance using deep transfer learning and deep domain adaptation: towards better generalization, Eng. Appl. Artif. Intell. 119 (C) (2023), https://doi.org/10.1016/j.engappai.2022.105698.
[71] M. Yang, C. Colak, K.K. Chundru, et al., Automated knee cartilage segmentation for heterogeneous clinical MRI using generative adversarial networks with transfer learning, Quant Imaging Med. Surg. 12 (5) (2022) 2620–2633, https://doi.org/10.21037/qims-21-459.
[72] V. Pai, H. Knipe, Kellgren and Lawrence system for classification of osteoarthritis. https://radiopaedia.org/articles/27111.
[73] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, Artif. Intell. Rev. 53 (2020) 5455–5516, https://doi.org/10.1007/s10462-020-09825-6.
[74] S.C. Sangapu, K.S.N. Prasad, R.J. Kannan, et al., Impact of class imbalance in VeReMi dataset for misbehavior detection in autonomous vehicles, Soft Comput. (2023), https://doi.org/10.1007/s00500-023-08003-4.
[75] V. Kishore Veparala, V. Kalpana, Big data and different subspace clustering approaches: from social media promotion to genome mapping Salud, Ciencia y Tecnología 3 (2023) 413. https://revista.saludcyt.ar/ojs/index.php/sct/article/view/413.

Dr. V. Vijaya Kishore is working as Professor and Head, Department of Electronics and Communication Engineering at Mohan Babu University, Tirupati, AP. He has a teaching experience of 21+ years. He holds Ph.D. degree in Electronics and Communication Engineering department from S. V. University, Tirupati in the field of Bio-medical Image Processing. He is a Senior member IEEE, Fellow IETE, Member IEU, Life member ISTE, Senior member IRED, Professional member IERP and member INSTICC. He has published more than 35 research papers in journals of both international and national repute. He published three book articles on Effective Engineering Teaching. He has published two books in the area of Biomedical Image Processing. His-research areas are Biomedical Image Processing, Computer Aided Diagnosis tools design and development. His-main focus is on developing tools that help for segmentation and extraction of region of interest in medical images and for decision making on the diagnosis. He received National award of Excellence in teaching from Global Management Council. He is acting as reviewer for five SCI journals (IEEE-TMI, IEEE-JBHI, Tand F-CMB & BE, and Imaging Science) and two Scopus (IJBCE and IJEHMC) journal related to Medical and Image Processing. He is also a member of Editorial Review Board of International Journal of Biomedical and Clinical Engineering journal, American Journal of Science, Engineering and Technology, Science Publishing Group, USA and Glacier Journal of Scientific Research, Ahmedabad, India.

Ms. V. Kalpana is working as Assistant Professor in Department of Electronics and Communication Engineering at Mohan Babu University, Tirupati, AP, India. She has more than 15 years of experience in teaching and research. She has published more than 25 papers in reputed international and national journals. Her research areas are Biomedical Image Processing, Computer Aided Diagnosis tool design and development.

Dr. Hemanth Kumar G, Ph.D., is a highly accomplished individual with a strong background in the field of the Internet of Things. As a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE), he has demonstrated exceptional expertise and dedication in his field. Of his 10 years of teaching career, He has been working as an Assistant Professor at REVA University, School of ECE. His-research interests primarily revolve around the fascinating domains of the Internet of Things (IoT) and Machine Learning. Throughout his career, made significant contributions to academia through his teaching and research endeavours. By staying abreast of the latest advancements in IoT and Machine Learning, he ensures that his teaching and research are relevant and cutting-edge.