Knee Osteoarthritis Detection Using Deep Learning Algorithms

*Retaj Muftah Taboun re.taboun@uot.edu.ly **Dr. Youssef Gdura ***Dr. Sarra Elrabiei

Abstract. Osteoarthritis is one of the most common form of arthritis that affects middle-aged and elderly people, and osteoarthritis (OA) usually affects knees and small finger joints, as well as thumb. The conventional prevalent practice to diagnose such as osteoarthritis diseases relies on human interpretation of medical images. This manual diagnosis approach is prone to the medical skills of individuals, time consuming, and error-prone task. However, the Artificial Intelligence (AI) has emerged in the recent years as a new diagnosis approach with the potential to overcome these limitations, and it shown outstanding success in processing and analysis of medical images. This paper presents a convolutional neural network based model and a mobile application that can help healthcare specialists and nonprofessional individuals in diagnosing knee OA from 2D (X-ray) images. The proposed model was trained and was tested using images from three different sources: an online reputed medical source, local diagnostic centers, and from local hospitals. The user-friendly mobile application was designed to take an X-Ray image as an input to this model and then displays the level of severity of the knee OA.

The obtained results were very satisfactory the model yields up to 92% accuracy to predict the presence of osteoarthritis or not. The model also achieved about 86% accuracy to predict the Knee OA severity level (five grades) based on the KL system.

Keywords: Knee, Osteoarthritis, Arthritis, Convolutional Neural Network, Deep Learning Algorithms, Radiograph, X-ray.

*Staff member faculty of engineering Tripoli University Libya

**Staff member faculty of engineering Tripoli University Libya

***Staff member faculty of engineering Tripoli University Libya

1. Introduction

Osteoarthritis (OA) is the damage of joint cartilage and leads to damage of functionality in the knee and hips, and the early signs can be observed with tears in cartilage (S. Castañeda, et al. 2012 & D. Lubar, et al 2010). Nowadays, there are over 250 million patients suffer from arthritis globally, and around 70% of them are 60 years of age or over(Arthritis and its Public Health Burden, 2017 & A. O. Akinpelu, et al. 2009). The primary knee OA symptoms are pain, stiffness, decreased range of joint motion, and malfunctioning gait that ultimately increases the progression rate of the disease (Global Burden of Disease Study, 2013). These indications affect individuals' functionality, degrade their life quality, and lead to joint replacement surgery in many cases.

The traditional or manual procedures of diagnosing and detecting knee OA has some drawbacks because it relies on manual interpretation of medical images, which is an error-prone task. However, in the last two decades, deep learning algorithms have been used extensively in medical image analysis and classification problems and accomplished substantial achievements across variety of domains (Yamashita, R., *et al.* 2018 pp611-629). In fact, some researchers anticipate that some deep learning algorithms, such as Convolutional Neural Network (CNN), could achieve performances that resemble human experts (http://youtu.be/FxlSzS76VTY). Hence, it becomes imperative to find innovative solutions or to improve existing deep learning model's overall accuracy in predicting the levels of severity of Knee OA.

1. One of the main features to determine the severity of knee OA is the joint space narrowing, which could be easily visualized using existing diagnostic imaging techniques such as Magnetic Resonance Imaging (MRI) and X-Rays. The MRI technique is a valuable tool and provide multiple images,

which can be used to produce 3D images to extract essential information especially on knee kinematic data (Mezghani, N.; et al.2018). However, processing and analysing MRI multiple images need considerably a powerful computing device. On the other hand, the most popular procedures nowadays for knee OA diagnosis are based on plain radiographs; X-Rays, because 2D radiographs would be sufficient to diagnose patients with knee OA (Newman, Samuel et al. 2022).

This work includes building a model based on artificial neural network (CNN) for classifying Knee OA severity levels and developing a mobile application that can use this model to scan X-ray images and detect the level of severity. The predication of our model will be based on the Kellgren–Lawrence (KL) grading system, which classifies knee OA severity into five grades, where grade 0 represents healthy with no symptoms of knee OA (normal case), while grade 4 presents a severest stage (Murphy, L.; Helmick, C.G., 2012).

The 2D radiographs images were chosen here because first X-rays images are available, more accessible than MRI images, and cost-saving procedure. Secondly, radiographs images need less computation power and less space to load the image than MRIs. Thus, these features help us to launch the proposed model on mobile devices.

2. Literature Review

In 2018, Aleksei Tiulpin, et al. introduced a new transparent computer-aided diagnosis method based on the Deep Siamese Convolutional Neural Network [9]. This method can automatically score knee OA severity according to the KL grading scale. It was trained using data from the Multicenter Osteoarthritis Study and validated it on randomly selected 3,000 subjects (5,960 knees) from Osteoarthritis Initiative dataset. The method yielded an average multi-class accuracy of 66.71% compared to the annotations given by a committee of clinical experts.

Another method proposed a new learning model for knee OA detection using X-ray images (Yassine Nasser, et al. 2020). The proposed framework; called Discriminative Regularized Auto-Encoder (DRAE), is based on Auto-Encoders that allows to learn both relevant and discriminative properties to improve the classification performance. Their experimental results on data from the public multicenter Osteo Arthritis Initiative (OAI) show that the developed method presents potential results for early knee OA detection.
Recently, Md. Rezaul Karim, et al proposed Deep Knee Explainer to leverage explainable knee OA diagnosis based on radiographs and MRIs (Yassine Nasser, et al. 2020). First, they comprehensively preprocess MRIs and radiographs through the deep-stacked transformation technique against possible noises and artifacts that could contain unseen images for domain generalization. Then, they extract the region of interests (ROIs) by employing U-Net architecture with ResNet backbone. To classify the cohorts, they train DenseNet and VGG architectures on the extracted ROIs. Finally, they highlight class-discriminating regions using gradient-guided class activation maps (Grad-CAM++) and layer-wise relevance propagation (LRP), followed by providing human-interpretable explanations of the predictions. The training and testing was based on the Multicenter Osteoarthritis Study (**MOST**), which is a prospective and observational study of knee osteoarthritis (OA) in older Americans with OA disease. This approach yields up to 91% classification accuracy for four classes.

In 2022, two models were introduced to assist orthopedists and radiologists in the detection and classification of knee osteoarthritis in accordance with the KL classification system using X-rays images. One model separated the first two grades of (KL 0-1) system from the last three severest grades (II-IV) (Md. RezaulKarim, et al. 2021). The other model grouped the five levels of the KL system into three groups: The first two grades (KL 0-I) as normal, the third level (KL II) as non-severe and the last two levels (KL III-IV) as severe (N. Pongsakonpruttikul et al, 2022 pp1549-1558). The average accuracy of the first model in detecting and classifying the two groups was about 85%, while the second mode's average accuracy based on the three grouped was around 86.7%.

Our approach differs from the above-mentioned approaches in that our model; in addition to the online dataset, has been tested using unseen dataset from local health institutions. It was first tested on detecting the five severity levels and then tested on detecting two classes: normal (KL 0) or abnormal (K I-IV). Our model achieved, as we shall see in the results, better performance than the two recent models in (N. Pongsakonpruttikul et al, 2022 & Md. RezaulKarim, et al. 2021).

3. Methodology

Figure 1 shows an overview of the development process. The first step was collecting the dataset and then preprocessing the selected images. After that, several machine leaning and deep learning algorithms to select the proper algorithm. The CNN algorithm was chosen and tuned to the best performance. In the last step, we developed a user interface, which can be used for examining an X-ray image and shows the model's predication.

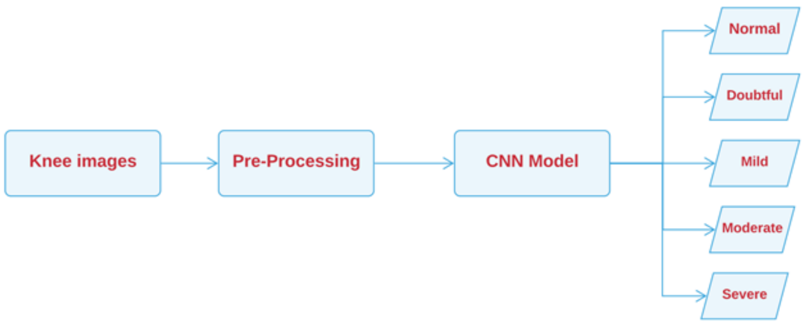

Fig. 1.Overview of the proposed model's development process

- **Dataset Description**

The dataset; X- Rays images, which was used to train and test our model, were collected from this online source; Mendeley Data (N. Pongsakonpruttikul et al, 2022 pp1549-1558). This dataset was used in many studies for training and testing several developed models. The online dataset consists of 1650 digital X-ray images of knee OA with different features such as osteophytes, narrowing of the joint space, subchondral bone sclerosis and subchondral bone cysts (https://data.mendeley.com/datasets/t9ndx37v5h/1). This dataset was split into 80% for training set, 10% for validation, and 10% for testing. The images are 8-bit grayscale image, and each image was manually labeled by two medical specialists as per KL grading system, which comprises, as shown in Fig.2, the following five grades:

1. Grade 0 (Normal): No radiological findings of OA.
2. Grade I (Doubtful): Suspicious narrowing of the joint space and possible osteophytic lipping.

3. Grade II (Mild): Definite osteophytes and possible narrowing of the joint space.
4. Grade III (Moderate): Moderate multiple osteophytes, definite narrowing of the joint space, small pseudocystic areas with sclerotic walls, and possible bone contour deformity.
5. Grade IV (Severe): Large osteophytes marked narrowing of the joint space, severe sclerosis, and definite deformity of the bone contour.

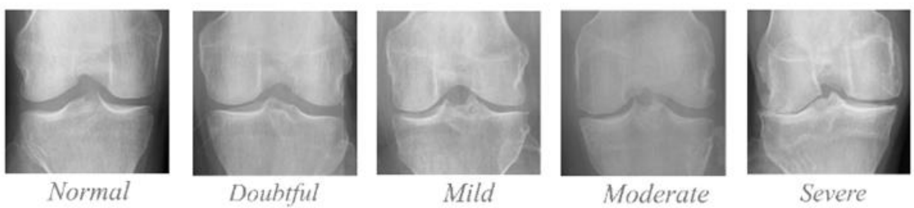

Fig. 2. Knee X-ray images (KL classification)

In addition to the online dataset, we collected other knee images of resident patients from the local diagnostic centers to test the model. The local dataset consists of 20 X-ray images that were collected from domestic diagnostic centers and local hospitals for resident patients. This local dataset was used to check how well the model is performing on new, never-before-seen-data.

○ Preprocessing Stage

The preprocessing stage is a very important stage in order to enhance the images and increase the model's accuracy. The main steps in this are shown in Fig. 3.

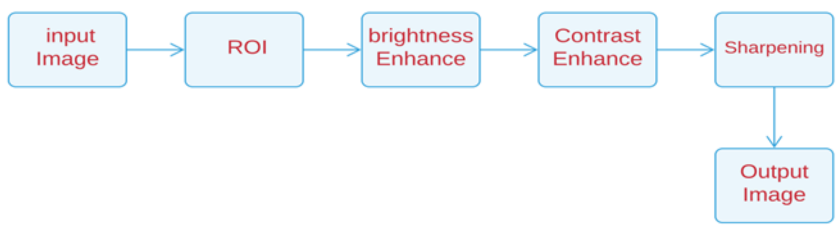

Fig. 3.Overview of the dataset preprocessing

○ Region of Interests (ROI) Segmentation

A region of interest (ROI) means the important region in an image. In this step, we manually specified the ROI in these images and provided its dimensions as a rectangle in order to focus on the knee joint space. The ROI segmentation has been applied on each image in the dataset. Fig. 4 shows two images with their original size; 224×224 pixel, and Fig. 5 shows the same images after applying ROI. As a result, all images in the dataset will come to be with dimensions (224×160). The procedure was done an attempt to focus on the important part while processing the image and to reduce the computation process shall

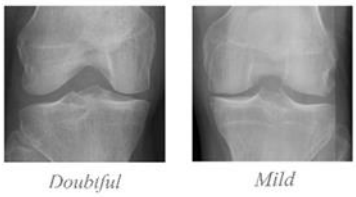

Fig. 4 Original Images

Fig. 5 Images after ROI

- **Image Enhancement**

This step is applied mostly on severe class images because most images in this class are not clear, as shown in Fig.6. For this reason, fixing this problem requires using two method of image enhancement; brightness and contrast enhancement.

A color screens use three colors i.e., RGB scheme (red, green and blue) the brightness of the screen depends upon the sum of the amplitude of red green and blue pixels, and it is divided by 3.

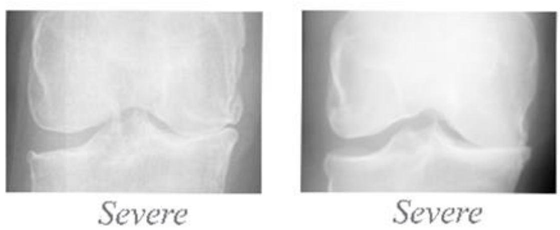

Fig. 6 Before brightness enhancement

2. The perception of brightness depends upon the optical illusions to appear brighter or darker. When the brightness is decreased, the color appears dull, and when brightness increases, the color is clearer (Kellgren and Lawrence system for classification of osteoarthritis 2021).

Fig. 7 shows images after applying brightness enhancement, and then applying contrast enhancement as shown in Fig. 8.

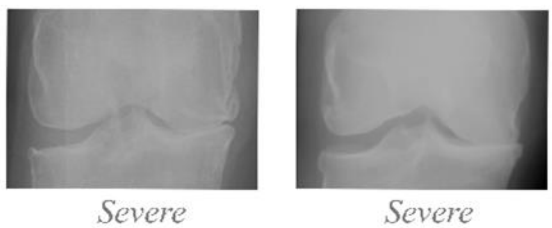

Fig. 7. After Brightness Enhancement

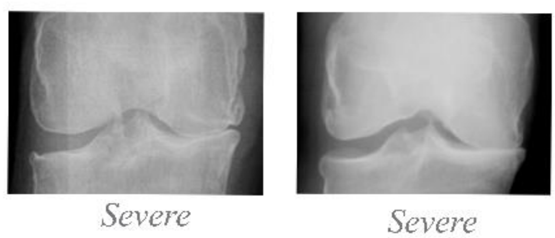

Fig. 8. After Contrast Enhancement

- **Image Filters**

The authors applied sharpening filters to the images using convolution operation. Fig. 9 and Fig. 10 shows an image before and after applying sharpening filter.

Fig. 9 Images before sharpening

Fig. 10 Images after sharpening

However, our model was tested using the original dataset (images) before and after the preprocessing stage or image enhancement steps.

- **Deep Learning Model Configuration**

In this project, authors chose the Convolution Neural Network (CNN) as the common deep learning type for image processing and gave us satisfactory results. The CNN architecture, as shown in Fig. 11, contains of three main layers: Convolution layers, pooling layers, and fully connected layer.

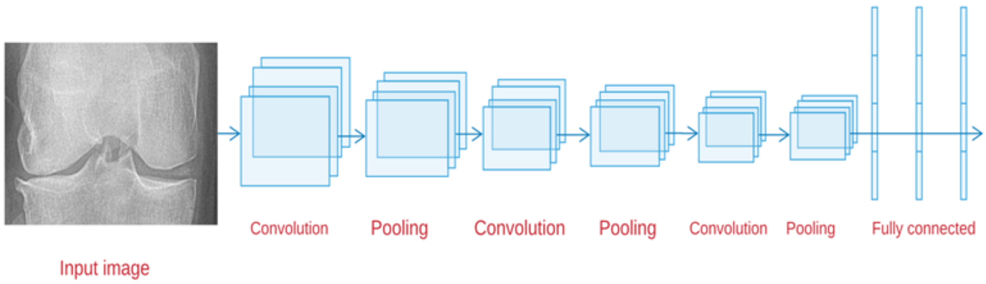

Fig. 11. CNN Architecture

- **Convolution layer**

The main task of the convolutional layer is to detect local conjunctions of features from the previous layer and mapping their appearance to a feature map. As a result of convolution in neural networks, the image is split into perceptrons, creating local receptive fields and finally compressing the perceptrons in feature maps of size $m_2 \times m_3$. Thus, this map stores the information where the feature occurs in the image and how well it corresponds to the filter. Hence, and each filter is trained spatial about the position in the volume it is applied to.

In each layer, there is a bank of m_1 filters. The number of filters, which are applied in one stage, is equivalent to the depth of the volume of output feature maps. Each filter detects a particular feature at every location on the input. The output $Y_i^{(L)}$of layer Lconsists of $m_1^{(L)}$feature maps of size $m_2^{(L)} \times m_3^{(L)}$. The i^{th}feature map, denoted$Y_i^{(L)}$, is computed as

$$Y_i^{(L)} = B_i^{(L)} + \sum_{j=1}^{m_1^{(L-1)}} K_{i,j}^{(L)} * Y_j^{(L-1)} \quad (1)$$

Where $B_i^{(L)}$ is a bias matrix and $K_{i,j}^{(L)}$ is the filter of size $2h_1^{(L)} + 1 \times 2h_2^{(L)} + 1$ connecting the j^{th} feature map in layer $(L\text{-}1)$with i^{th} feature map in layer (YannLeCun, KorayKavukcuoglu, 2010).

The result of staging these convolutional layers in conjunction with the following layers is that the information of the image is classified like in vision.

- **Pooling layer**

The pooling or down sampling layer is responsible for reducing the special size of the activation maps. In general, they are used after multiple stages of other layers (i.e. convolutional and non-linearity layers) in order to reduce the computational requirements progressively through the network as well as minimizing the likelihood of overfitting. The pooling layer L has two hyper parameters, the spatial extent of the filter $F^{(L)}$ and the stride$S^{(L)}$.It takes an input volume of size $m_1^{(L-1)} \times m_2^{(L-1)} \times m_3^{(L-1)}$ and provides an output volume of size $m_1^{(L)} \times m_2^{(L)} \times m_3^{(L)}$ where;

$$m_1^{(L)} = m_1^{(L-1)} \quad (2)$$

$$m_2^{(L)} = (m_2^{(L-1)} - F^{(L)})/S^{(L)} + 1 \quad (3)$$

$$m_3^{(L)} = (m_3^{(L-1)} - F^{(L)})/S^{(L)} + 1 \quad (4)$$

The key concept of the pooling layer is to provide translational invariance since particularly in image recognition tasks, the feature detection is more important compared to the feature's exact location. Therefore, the pooling operation aims to preserve the detected features in a smaller representation and does so, by discarding less significant data at the cost of spatial resolution.

The pooling layer operates by defining a window of size $F^{(L)} \times F^{(L)}$and reducing the data within this window to a single value. The window is moved by $S^{(L)}$positions after each operation similarly to the convolutional layer and the reduction is repeated at each position of the window until the entire activation volume is spatially reduced.

It is noteworthy that the window for pooling layers does not have to be a square and can be parameterized with $F_1^{(L)}$and $F_2^{(L)}$resulting in a rectangle of size $F_1^{(L)} \times F_2^{(L)}$. This is extremely uncommon and is therefore left out of the notation.

- **Fully connected layer**

The fully connected layers in a convolutional network are practically a multilayer perceptron (MLP), generally a two or three layer MLP that aims to map the $m_1^{(L-1)} \times m_2^{(L-1)} \times m_3^{(L-1)}$ activation volume from the combination of previous different layers into a class probability distribution. Thus, the output layer of the multilayer perceptron will have $m_1^{(L-i)}$outputs, i.e. output neurons where *i* denotes the number of layers in the multilayer perceptron.

The key difference from a standard multilayer perceptron is the input layer where instead of a vector, activation volume is taken as the input. As a result the fully connected layer is defined as:

IfL$-1$is a fully connected layer;

$$y_i^{(L)} = f\left(z_i^{(L)}\right) \text{ with } z_i^{(L)} = \sum_{j=1}^{m_1^{(L-1)}} \omega_{i,j}^{(L)} y_i^{(L-1)} \quad (5)$$

Otherwise;

$$y_i^{(L)} = f\left(z_i^{(L)}\right) \text{ with } z_i^{(L)} = \sum_{j=1}^{m_1^{(L-1)}} \sum_{r=1}^{m_2^{(L-1)}} \sum_{s=1}^{m_3^{(L-1)}} \omega_{i,j,r,s}^{(L)} (Y_i^{(L-1)})_{r,s} \quad (6)$$

The goal of the complete fully connected structure is to tune the weight parameters $\omega_{i,j}^{(L)}$ or $\omega_{i,j,r,s}^{(L)}$ to create a stochastic likelihood representation of each class based on the activation maps generated by the concatenation of convolutional, non−linearity, rectification and pooling layers. Individual fully connected layers operate identically to the layers of the multilayer perceptron with the only exception being the input layer.

It is noteworthy that the function f once again represents the non−linearity, in a fully connected structure the non−linearity is built within the neurons and is not a separate layer (Diego Unzueta, 2021).

- **Model Configuration**

1. The first layer uses a large kernel size, but no stride because the input images are not very large, followed by the activation ReLu layer; see Fig. 11.
2. Next, the max pooling layer divides each spatial dimension by a factor of two (since pool_size=(2,2)).
3. Then by repeating the same steps twice: one convolutional layer followed by the activation ReLu layer and max pooling layer.
4. The number of filters that has been used in the CNN towards the output layer (it is initially 128, then 64, then 32).
5. The next layer is the fully connected network. It composed of two hidden dense layers and a dense output layer. Fully connected network must flatten its inputs, since a dense network expects a 1D array of features for each instance. By adding two dropout layers, with a dropout rate of 20% at the first one, and dropout rate of 10% at the second one, to reduce the overfitting.

- **Optimization Process**

In attempt to improve the model's accuracy, we applied an optimization tactic by manually adjusting two common hyper-parameters; epochs and batches,. These two parameters were set to different values across experiments, and each time the model retrained and revalidated. The best performance was achieved when the batch size was 64 and number of epochs was 80.

After the optimization process, we built three different versions of the model. The first two versions of the model were designed to predicate the severity; five grades, of knee OA based on the KL grading system. The first version was trained using the original dataset without any image enhancement, and the second version of the model was trained and was validated after we applied image enhancement on the online dataset. The third version was a binary classification model to predict if the knee X-ray images as normal (no OA) or not normal (with one of the knee OA stages).

- **User Interface**

To assess the model by healthcare professionals or any individuals, we develop an Android mobile application that allows easy access to our CNN model. This mobile App provides a user interface that can interact with the model offline. The user interface consists of one button and two viewing windows. The button can be used to select the image to be diagnosed, and one window to view the selected image, and a text window to show the predication of the model. Once the user click on the button, the App will allows the user to select any X-ray image, which is previously stored on the mobile device, for diagnostic purpose. The selected image will then appear at the Image View window, and the model will immediately classify it and show the predication at the Text View as shown in Figure 12.

Fig. 12. The application's final results

The Tensor Flow Lite Python library was used to be able to launch the model on mobile devices because they have limited computation power and storage capacity.

- **Training and Validation**

The model was first trained and was tested using the original online dataset without any image enhancements. This dataset was split into 80%, 10%, and 10% for training, validation, and testing respectively. The model was also tested using X-ray images that were collected from domestic diagnostic centers and local hospitals.

- **Results and Discussions**

The subsequent results are obtained using three different testing sets. The accuracy of the model on the first testing set, which presents 10% of original online dataset before we applied an image enhancement, is discussed in the first approach. In the second approach, we present the results obtained using 10% of original dataset after image enhancement, and then present the results obtained using the images, which was collected locally from diagnostic centres and local hospitals. The results of the binary classification model are stated in the third approach.

- **First Approach**

The model was first built to predicate the severity of knee OA based on the KL system without image enhancements. Through the conducted experiments, the model showed slight improvement and the accuracy stays roughly within a range of 2%. The accuracy on the training set was around 80%, and the accuracy on the validation set was less than 73%. However, the accuracy declined even more in the testing phase and dropped to 65%. Fig. 13 shows the accuracy and loss of the model on the training and validation data sets.

The authors noticed that the validation accuracy in this approach improved quite slowly. This is mainly due to multiple connections within the dense blocks. However, varying the batch size does not have a decisive impact on the performance, and increasing the epochs showed an acceptable improvement, yet it took substantial time for the training.

Fig. 13 Training Loss and Accuracy.

- **Second Approach**

In this approach, we tested the model using the 10% of the enhanced images. Though the model's accuracy to predict the five knee OA grades was above 96% in the training phase and around 90% in the validation phase, yet our model achieved better performance as compared to the very recent model, which were mentioned in [11 , 12]. Fig. 14 and Fig. 15 show the accuracy and loss in the training and validation stages against the number of epochs. However, the model achieved above 86% on the testing stage.

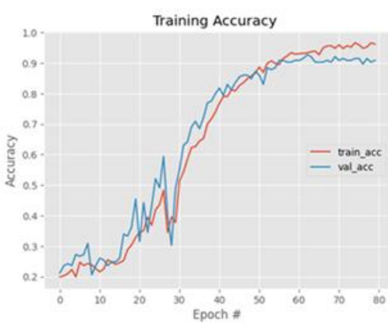

Fig. 14 Training Accuracy

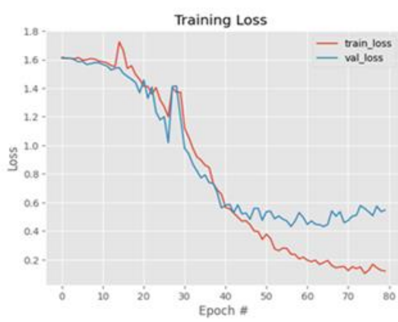

Fig. 15 Training Loss.

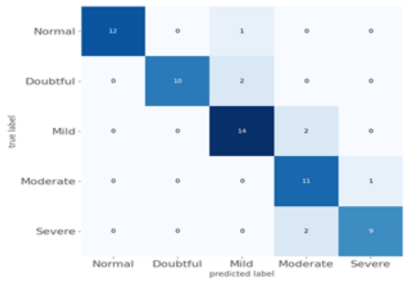

Fig. 16 Confusion Matrix

This version was also tested on a local testing set that was collected from a local radiology and medical imaging center. The confusion matrix given in Fig. 16 shows the prediction accuracy of the five knee OA grades using this testing set. The confusion matrix-based measures using the same testing set are shown in Table 1.

Table 1. Performance results

Class	Precision	Recall	F_1-Score	Overall Accuracy
Normal	1.000	0.923	0.960	0.875
Doubtful	1.000	0.833	0.909	
Mild	0.824	0.875	0.848	
Moderate	0.733	0.917	0.815	
Severe	0.900	0.818	0.857	

Figure 17 shows nine samples of this testing set, and on the top of each image, it shows the model's predication and the real knee OA grade based on KL grading system.

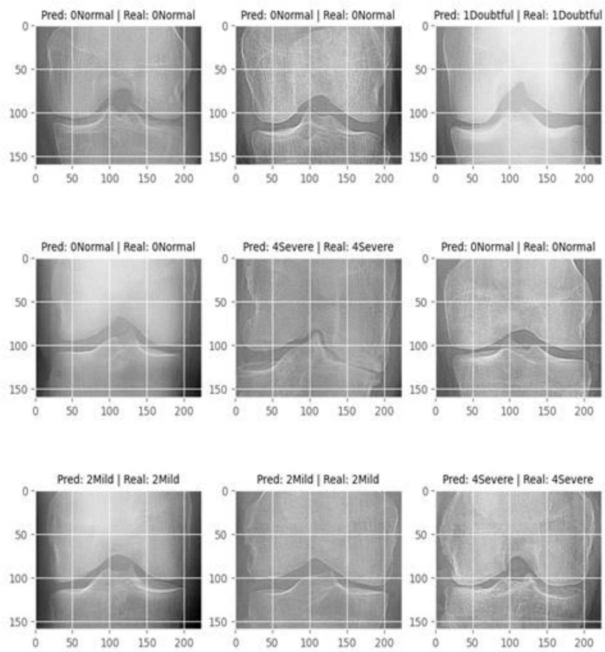

Fig. 17. Images with the Predicted and Real KL Grades

The model was also tested on a local testing set using the light version of the model, which was built specifically for mobile devices. For testing our mobile application on the spot, we visited a local hospital that provided us with several X-ray images and one radiologist, who tried the application using his own patients' images. The application won his approval, as his assessment of the model's accuracy on the local data reaches up to 95%.

The observation here is that the model's accuracy during the validation and testing declined slightly which is usually expected since there are often some differences between the data, which the model was trained on, and the testing data that is used for evaluation the model.

○ **Third Approach**

The binary classification version of the model has achieved 92.5% of testing accuracy while the previous version to predict the five Knee OA achieved around 87% accuracy. The improvement here, which was roughly 5%, was a result of combining the four classes of severity as on class; that is, not normal without determining the grade of the severity.

• 6. Conclusions

Knee OA is a degenerative joint disease that significantly affects middle-aged and elderly people, yet providing a precise and accurate diagnosis of Knee OA is a challenging task. This due to the similarity between different KL grades that makes it hard to recognize each grade, and it might lead to misclassification issue.

This paper presented a CNN based model that can detect the five severity levels of knee OA based on X-ray images. The development process went through several stages starting by optimizing the model, and then pre-processing the images. After that, we build two versions of the model using three different testing set.

The two versions showed very satisfactory results as compared to the other approaches, which presented in the literature review. Our model was capable to detect the five severity levels based on the KL system with quite remarkable accuracy that ranges between 87% up to 92%.

The overall performance of the model after the image enhancement process was remarkably improved as compared to the achieved accuracy in the first approach. The mobile application was also tested and used by several users who were very satisfied.

References

1. S. Castañeda, J. A.-Blas, R. Largo, and G. Herrero- Beaumont, "Subchondral bone as a key target for osteoar- thritis treatment," *Biochemical Pharmacology*, (2012).
2. D. Lubar, P. H. White, L. F. Callahan et al., "A national public health agenda for osteoarthritis 2010," *Seminars in Arthritis and Rheumatism*, (2010).
3. Arthritis and its Public Health Burden, Dela J Public Health, (2017).
4. A. O. Akinpelu, T. O. Alonge, B. A. Adekanla, and A. C. Odole, "Prevalence and pattern of symptomatic knee osteoarthritis in Nigeria: a community-based study," *The Internet Journal of Allied Health Sciences and Practice*, (2009).
5. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013. *Lancet,* (2015).
6. Murphy, L.; Helmick, C.G. The impact of osteoarthritis in the United States: A population-health perspective: A population-based review of the fourth most common cause of hospitalization in US adults. *Orthop. Nurs,* (2012).
7. Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9
8. Stages Of Osteoarthritis Of The Knee| Orthopedic Surgeon in Bangalore| DrHemantL Kalyan| Manipal: http://youtu.be/FxlSzS76VTY.

9. Aleksei Tiulpin, et al " Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach", (2018).
10. Yassine Nasser, et al. "Discriminative Regularized Auto-Encoder for Early Detection of Knee OsteoArthritis: Data from the Osteoarthritis Initiative", (2020).
11. Md. RezaulKarim, et al. "DeepKneeExplainer: Explainable Knee Osteoarthritis Diagnosis From Radiographs and Magnetic Resonance Imaging", (2021)
12. N. Pongsakonpruttikul et al, "Artificial intelligence assistance in radiographic detection and classification of knee osteoarthritis and its severity: a cross-sectional diagnostic study. Eur Rev Med Pharmacol Sci. 2022 Mar;26(5):1549-1558.
13. *Digital Knee X-ray Images.* From Mendeley Data, https://data.mendeley.com/datasets/t9ndx37v5h/1
14. "Kellgren and Lawrence system for classification of osteoarthritis," Radiopaedia.
15. "Brightness and Contrast," javaTpoint, (2021).
16. YannLeCun, KorayKavukcuoglu, Clement Farabet, "Convolutional networks and applications in vision," (2010).
17. Diego Unzueta,"Convolutional Layers vs Fully Connected Layers," Towards Data Science, (2021).
18. Mezghani, N.; Mechmeche, I.; Mitiche, A.; Ouakrim, Y.; De Guise, J.A. An analysis of 3D knee kinematic data complexity in knee osteoarthritis and asymptomatic controls. *PLoS ONE* **2018**.

19. Newman, Samuel et al., "Radiographic vs. MRI vs. arthroscopic assessment and grading of knee osteoarthritis - are we using appropriate imaging?." *Journal of experimental orthopaedics* vol. 9, 1 2. 3 Jan. 2022.

20. Prezja, F., Paloneva, J., Pölönen, I. *et al.*, "Deep Fake knee osteoarthritis X-rays from generative adversarial neural networks deceive medical experts and offer augmentation potential to automatic classification. *Sci Rep* **12**, 18573 (2022)