A Convolution Neural Network Design for Knee Osteoarthritis Diagnosis Using X-ray Images

https://doi.org/10.3991/ijoe.v19i07.40161

Saleh Hamad Sajaan Almansour[1], Rahul Singh[2], Saleh Mabruk Hadrami Alyami[1], Neha Sharma[2], Mana Saleh Al Reshan[1(✉)], Sheifali Gupta[2], Mahdi Falah Mahdi Alyami[1], Asadullah Shaikh[1],
[1]Department of Information Systems, College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia
[2]Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
msalreshan@nu.edu.sa

Abstract—Knee osteoarthritis (OA) is a chronic degenerative joint disease affecting millions worldwide, particularly those over 60. It is a significant cause of disability and can impact an individual's quality of life. The condition occurs when the cartilage in the knee joint wears away over time, leading to bone-on-bone contact, which can result in pain, stiffness, swelling, and decreased range of motion. Deep neural networks, especially convolutional neural networks (CNN), are powerful tools in medical applications such as diagnosis and detection. This research proposes a CNN model to classify knee osteoarthritis into five categories using x-ray images. These classes are labeled: Minimal, Healthy, Moderate, Doubtful, and Severe. Furthermore, the proposed CNN model has been compared with two pre-trained transfer learning models: Xception and InceptionResNet V2. These models were evaluated based on precision, recall, F1 score, and accuracy. The results showed that although all three models performed very well, the proposed model outperformed both transfer learning models with 98% accuracy. It also achieved the highest values for other parameters such as precision, recall, and F1 score. The proposed model has several potential applications in clinical practice, such as assisting doctors in accurately classifying knee osteoarthritis severity levels by analyzing single X-ray images.

Keywords—Osteoarthritis diagnosis, deep learning, X-ray images, classification, convolution neural network.

1 Introduction

Osteoarthritis is the most prevalent kind of arthritis and affects millions of individuals all over the globe [1]. This condition arises due to the gradual wearing away of the cartilage that protects and cushions the bone ends over time. The joints in your hands, knees, hips, and spine are most likely affected by osteoarthritis, even though the disease may affect any joint [2]. Osteoarthritis of the knee is caused when the cartilage in your knee joint deteriorates [3], which allows the bones in the joint to rub against one another. Your knees will pain, become stiff, and even swell up due to the contact. There is currently no therapy that will reverse the effects of osteoarthritis in the knee; however, several medications are available that may delay the progression of the disease and improve the symptoms [4]. In cases with osteoarthritis that are more severe, surgical treatment may be a possibility. The most prevalent cause of disability in adults is osteoarthritis of the knee. Over eleven percent of those younger than sixty-four have been diagnosed with knee osteoarthritis (O.A.). Knee osteoarthritis is quite prevalent in the general population [5]. It is estimated that around 46 percent of individuals will get it at some point. Knee osteoarthritis affects women at a rate of two to three times higher than males [6]. After age 40, most individuals begin to show symptoms of this illness. However, several additional variables, such as trauma or heredity, might cause it to manifest at an earlier age [7]. It may be excruciatingly uncomfortable to engage in activities such as Jogging, running, climbing stairs, or kneeling if you suffer from osteoarthritis in the knee, the knee's most prevalent complaint. It is also possible for your knees to feel tight and swollen due to this condition [8]. The form of your knee joint might alter over time if you have knee osteoarthritis. This can give you the sensation that your joint is unstable or shaky. Joint pain from osteoarthritis is due to the slow breakdown of cartilage, which cushions the ends of bones in your joints [9]. Cartilage makes joint movement practically frictionless, a potent and slippery substance. If the cartilage wears away entirely, the bones will begin to rub against one another. "Wear and tear disease" is often used to describe osteoarthritis [9, 10]. However, osteoarthritis affects not only the cartilage, affects the joint as a whole. Damage to the connective tissues that stabilize joints and link muscles to bone directly results from this condition. Joint lining irritation is another effect. Symptoms of osteoarthritis often start mild and gradually progress to more severe forms throughout the disease. Osteoarthritis may manifest in several ways, presenting symptoms such as pain, tenderness, loss of flexibility, a grating feeling, bone spurs, and swelling [11]. The chance of developing osteoarthritis may be increased by several factors, including being older and overweight [12]. Injuries to the joint, repeated stress on the joint, genetics, and certain metabolic illnesses may all contribute to joint pain. Osteoarthritis is a progressive degenerative condition that may cause severe discomfort over time [13]. The hurt and stiffness in your joints may become so debilitating that you have trouble doing even the most fundamental duties. The pain and incapacity brought on by osteoarthritis may lead to feelings of depression and disrupted sleep. Although osteoarthritis can't be cured, some therapies may ease the discomfort and make it easier to move about. Acetaminophen, nonsteroidal anti-inflammatory medicines (NSAIDs), and duloxetine may alleviate some of the symptoms of osteoarthritis, most notably pain (Cymbalta). There is no complete method to stop knee osteoarthritis from developing, but there are things one can do to lessen the likelihood of it happening to them. These include the following: Striving to maintain a healthy weight, making sure one gets enough sleep, and Jogging or running only on grass or other surfaces that are somewhat soft. Low-impact ones like swimming or cycling should accompany alternate high-impact activities like running or jumping rope to keep workouts interesting. One should include some modest strength training into their regular workout plan. A CNN model has been proposed in this study [14].

The significant contributions of this research work are as follows:

- This work implements a deep learning-based CNN model using x-ray images to categorize knee osteoarthritis. The disease is divided into five classes: Minimal, Healthy, Moderate, Doubtful, and Severe.
- The suggested model has been equated with two transfer learning models: Xception and InceptionResNet V2.
- The model has been evaluated using performance parameters such as precision, recall, f1score, and accuracy.

The remaining part of this manuscript has been arranged as section 2 considers the literature review of knee osteoarthritis, section 3 describes the proposed methodology, section 4 depicts the results and discusses them, and section 5 concludes the complete manuscript paper.

2 Literature review

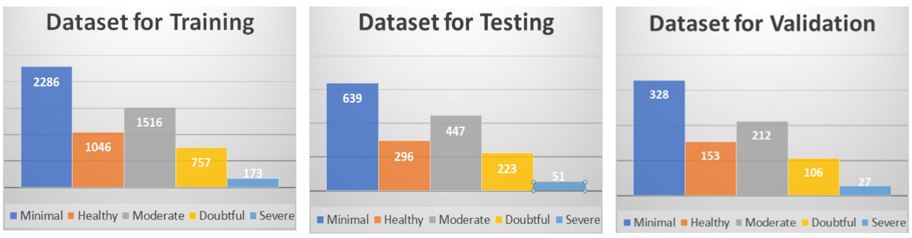
Automatic classification and diagnosis of knee OA, Kotti et al. [15] proposed a system that is based on generic subject attributes (such as age, sex, assessment of the Knee Injury, Osteoarthritis Outcome Score (KOOS)), as well as kinematic data acquired throughout a gait cycle. This system was developed with the assistance of several CNN designs. The findings demonstrated a clear distinction between the classes, despite the sample size being insufficient to provide conclusions of any practical use [16] presented a method for automatically identifying knee O.A. The authors established a methodology that is predicated on body kinetics. They validated their proposed method by applying it to a sample of 94 individuals, which included 47 patients diagnosed with OA and 47 healthy controls. The overall accuracy of the process was 72.61% plus 4.24%. A technique based on decision trees and using MLP at the tree leaves was suggested in [17]. The Kellgren-Lawrence scale classified the participants as “normal,” “mild,” “moderate,” or “severe,” depending on the severity of their symptoms. It was possible to achieve an accuracy of 80 percent on average. Uwisengeyimana and Ibrikci [18] both succeeded in their mission. In this article, we look at how artificial neural network techniques that go beyond deep learning may be used to improve the diagnosis of knee diseases. Brahim et al. [19] suggested a computer-aided diagnosis approach for significantly earlier knee identification using knee X-ray imaging and several machine-learning techniques. The system makes use of knee imaging. The investigation revealed that the proposed method had an excellent predictive classification rate for detecting O.A. The author got an accuracy score of 87.15 percent. Antony et al. [20] presented research using the pooled information from the Osteoarthritis Initiative (OAI) cohort baselines. Two thousand three hundred photographs were used for the testing, with around 1,200 individuals serving as the subjects. According to the authors, the accuracy was improved when using the combined dataset instead of using the OAI dataset only for testing or training. This was the case when comparing testing and training with the combined dataset. The methodology used in these studies utilized deep learning as its base. The author [21] described a data-driven MRI-based platform that uses T2 measurements to characterize radiographic O.A. The results showed that component learning from T2 maps could potentially disclose data that may be more helpful in O.A. diagnosis than simple averaging or linear pattern decomposition. This potential was evidenced by the fact that the findings.

3 Proposed methodology

In this research, Knee osteoarthritis is classified into five categories: Minimal, Healthy, Moderate, Doubtful, and Severe. The classification task is performed using the proposed CNN model and two transfer learning architectures: Xception and InceptionResNet V2. The dataset used in this study contains 8381 pictures of Knee Osteoarthritis, which is divided into the train, test, and validation datasets. Five thousand ninety-seven pictures were used for training, 2458 for testing, and 826 for validation. Before training the model, the images were preprocessed by resizing them to 224 x 224 image size, and the dataset was also rescaled by normalization. The Adam optimizer was used to train the model across 50 epochs with a batch size of 32 [22]. During training, the model was evaluated using accuracy, precision, recall, and F1-score metrics to measure the performance of the model. Finally, the trained model was tested on the testing set, and the performance metrics were evaluated to determine the model's effectiveness in detecting the five types of knee osteoarthritis.

3.1 Dataset description

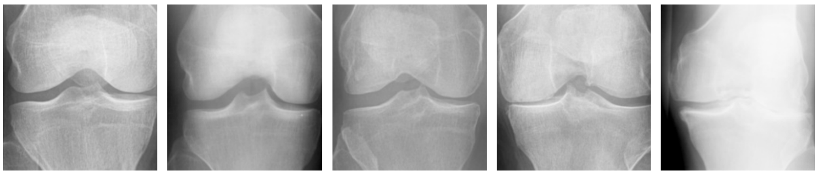
Figure 1 depicts the training, test, and validation splitting of the x-ray dataset of knee osteoarthritis. The dataset includes 8381 images labeled based on the severity of the knee condition shown in each image. The training set has the most pictures, totaling 5097 in total. This set's label distribution is as follows: 2286 are minimal, 1046 are healthy, 1516 are moderate, 757 are sceptical, and 173 are severe. Typically, this set trains machine learning models to predict the severity of knee conditions in new X-ray images. There are 2458 images in the testing set, with 639 labeled as Minimal, 296 as Healthy, 447 as Moderate, 223 as Doubtful, and 51 as Severe. Typically, this set calculates the performance of machine learning models trained on the training set.



(a) **(b)** **(c)**

Fig. 1. Dataset splitting (a) Training, (b) Testing, and (c) Validation

There are 826 images in the validation set, with 328 labeled as Minimal, 153 as Healthy, 212 as Moderate, 106 as Doubtful, and 27 as Severe. This set is used to assess the performance of machine learning models during training and, if necessary, make adjustments to the model's hyperparameters. Figure 2 depicts different input pictures. Figure 2 displays the five distinct phases of knee osteoarthritis.



(a) **(b)** **(c)** **(d)** **(e)**

Fig. 2. Sample image of dataset (a) Minimal, (b) Healthy, (c) Moderate, (d) Doubtful and (e) Severe

3.2 OTSU thresholding

Binary or grayscale input pictures are necessary for deep-learning models. It is possible to binarize color pictures using a threshold, which makes it simpler to preprocess the data before training the model. Otsu's thresholding is a technique for automatically figuring out the best threshold value to employ when dividing a picture into two groups of pixels. Nobuyuki Otsu first suggested the concept in 1979, and it has subsequently gained popularity as a tool for image processing and computer vision [23]. The primary tenet of Otsu's thresholding is finding a threshold value that maximizes the inter-class variance while minimizing the intra-class variation. As a result, the pixels on each side of the threshold should be as unlike from one another as feasible, while the pixels on the same side should be as similar as possible. We first generate a histogram of the pixel intensities in the image before using Otsu's thresholding [24]. Next, for each threshold, we compute the intra-class variance and inter-class variance using a loop over all potential threshold values. The best threshold value is then determined as maximizing the interclass variance while minimizing the intraclass variance. By setting all pixels with intensities below the threshold to 0 (black) and all pixels with intensities above the threshold to 255 [25], we can utilize the ideal threshold value to binarize the picture (white). Otsu's thresholding is a powerful image segmentation method with several uses in industries, including robotics, computer vision, and medical imaging [26]. Figure 3 shows some sample images from each class for Otsu thresholding.

(a) **(b)** **(c)** **(d)** **(e)**

Fig. 3. Otsu thresholding images (a) Minimal, (b) Healthy, (c) Moderate, (d) Doubtful, and (e) Severe

3.3 Proposed CNN model

Figure 4 shows a convolutional neural network (CNN) architecture for image classification. The first block represents the network's input layer, where raw image data is fed into the network. image classification. The first block represents the network's input layer, where raw image data is fed into the network.

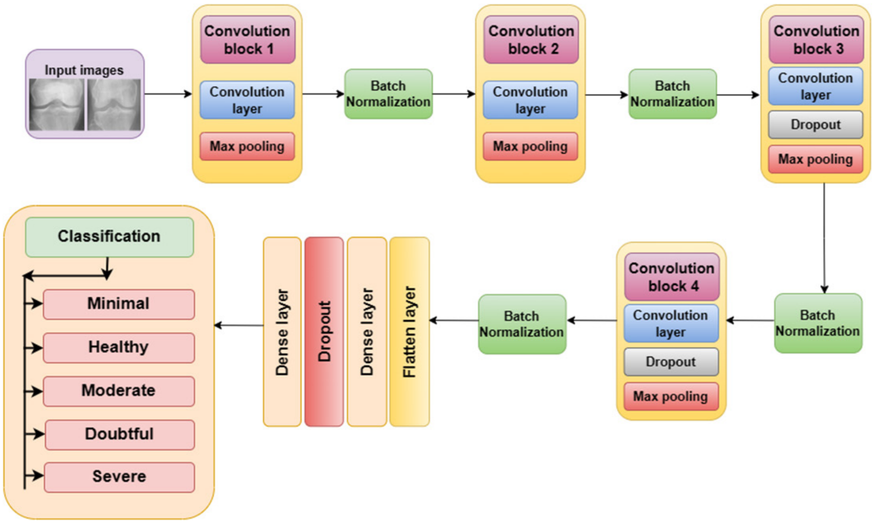


Fig. 4. Proposed CNN model

The convolutional blocks include a convolutional layer that applies a series of filters on the input picture to extract features, and a max pooling layer, which decreases the feature maps' longitudinal dimensionality without losing any relevant information. After each convolutional layer, batch normalization is applied to normalize the activations and improve the convergence of the network. The architecture includes four convolutional blocks, each with the same properties as the first two. The third convolutional block also consists of a dropout layer after the convolutional and max pooling layers to randomly drop out some activations and prevent overfitting. The fourth convolutional block has similar properties to the third convolutional block. After the convolutional blocks, the Flatten layer converts the 3D feature maps into a 1D vector. The network then includes two Dense, fully connected layers that perform the final classification based on the extracted features. A Dropout layer is formed before the final Dense layer to prevent overfitting. Finally, the output layer contains five neurons, each representing one of the five classes in the classification task.

Table 1. Detailed description of proposed CNN model layers

Layers	Input Shape	Output Shape	Number of Filters	Parameters
Conv2d	98x98x64	98x98x64	64	1792
Maxpooling2d	98x98x64	49x49x64	64	0
Batch Normalization	49x49x64	49x49x64	64	256
Conv2d_1	49x49x64	47x47x32	32	18464
Maxpooling2d_1	47x47x32	23x23x32	32	0
Batch Normalization_1	23x23x32	23x23x32	32	128
Conv2d_2	23x23x32	21x21x64	64	18496
Dropout	21x21x64	21x21x64	64	0
Maxpooling2d_2	21x21x64	10x10x64	64	0
Batch Normalization_2	10x10x64	10x10x64	64	256
Conv2d_3	10x10x64	8x8x32	32	18464
Dropout_1	8x8x32	8x8x32	32	0
Maxpooling2d_3	8x8x32	4x4x32	32	0
Batch Normalization_3	4x4x32	4x4x32	32	128
Flatten	-	-	-	0
Dense	-	-	-	32832
Dropout_2	-	-	-	0
Dense_1	-	-	-	195
Total Parameters: 91,011 Trainable Parameters: 90,627 Non-Trainable Parameters: 384				

CNN is a deep learning model often employed for image classification applications. Table 1 illustrates CNN's architecture. The CNN consists of multiple layers: convolutional layers, max-pooling layers, batch normalization layers, dropout, flattened, and dense layers. The input shape of the CNN is 98x98x64, which means the network is designed to process images with a height and width of 98 pixels and 64 color channels. The first layer in the network is a convolutional layer that uses 64 filters with a kernel size of 3x3. The output of this layer is fed into a max pooling layer with a pool size of 2x2, which reduces the spatial dimensions of the data by half. Afterward, a batch normalization layer is applied to the data, which helps to improve the stability and speed of the training process by normalizing the inputs to each neuron. The next layer in the network is another convolutional layer, which uses 32 filters with a kernel size of 3x3. The output of this layer is again fed into a max pooling layer with a pool size of 2x2. Another batch normalization layer is applied to the data to normalize the inputs to each neuron. The next layer is another convolutional layer that uses 64 filters with a kernel size of 3x3. The next layer is a dropout layer, which uses randomization to turn some input units to zero during training to prevent overfitting. The output of the dropout layer is then fed into another max pooling layer with a pool size of 2x2, and another batch normalization layer is applied. The final convolutional layer in the network uses 32 filters with a kernel size of 3x3 and is followed by another dropout layer.

The output of the dropout layer is then fed into a max pooling layer with a pool size of 2x2, and another batch normalization layer is applied. The production of the final pooling layer is then flattened into a 1D array and fed into a dense layer with 32,832 parameters. Another dropout layer is then applied before the final output dense layer with 195 parameters equivalent to the number of classes in the classification task. The CNN architecture has 91,011 parameters, with 90,627 being trainable and 384 being non-trainable, corresponding to the batch normalization layers. The architecture is suitable for image classification tasks where the input data has a similar size to the one used during training. Overall, the CNN architecture described in the table is a standard design that can be used as a starting point for many image classification tasks. It can be further optimized by tweaking the number of layers, filter sizes, and other hyperparameters based on the specific requirements of the task

3.4 Transfer learning model

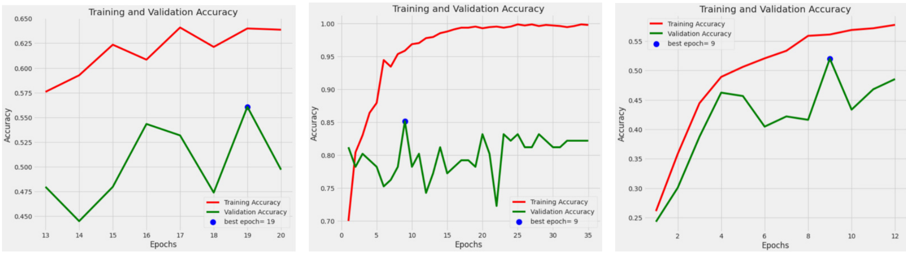
Transfer learning is used in deep understanding to retrain a model for a comparable task. It fine-tunes a previously learned model on a smaller, similar dataset [27]. The pre-trained model already understands the input features; it needs less data to train a model [28]. Transfer learning enhances model generalization by training applicable qualities across tasks. Transfer learning entails fine-tuning the pre-trained model or only the last few layers [29]. The pre-trained model might be from the target task area or a related domain. Deep learning models benefit from transfer learning, especially when data is sparse [30]. The prosed CNN model has been compared with two transfer leaning models named as Xception [31] and InceptionResNet V2 [32]. In the Xception model, depth-wise separable convolutions augment the Inception architecture. Convolutions of this type divide spatial and channel filtering into two layers. This significantly reduces the model's parameters while maintaining accuracy. The Inception-ResNetV2 approach trains deep networks using ResNet's residual connections and Inception's ability to extract features of varied sizes. With fewer parameters, this deep economic model delivers outstanding accuracy in the ImageNet dataset. Inception-ResNetV2 uses convolutional layers, pooling layers, and residual connections to extract and analyze image information. Several intermediate-layer auxiliary classifiers boost gradient flow and regularize the network during training.

4 Results and discussion

The results section includes in-depth analyses of all the parameters investigated during the study. This section also discusses the graphs depicting model accuracy and loss. The implementation has been done using the proposed CNN model, which is designed from scratch, and the CNN model is also compared with two transfer learning models: InceptionResNetV2 and Xception. The final component of the outcome is a comparison of all knee osteoarthritis classifications based on precision, recall, F1 score, and accuracy [33]. This component also includes the confusion matrix's characteristics. The batch size is 32, and every model is executed for 50 epochs.

4.1 Analysis based on accuracy plot

In machine learning, accuracy plot and loss plot are two commonly used visualization tools to monitor the performance of a model during the training process. An accuracy plot shows the model's ability to predict the output class label for each input sample correctly. It plots the model's accuracy on the y-axis, usually as a percentage, against the number of training epochs on the x-axis. Typically, the accuracy plot shows an increasing trend as the model learns from the data during training.



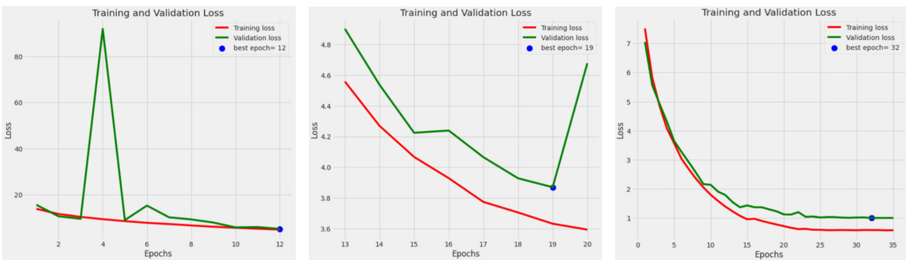
(a) (b) (c)

Fig. 5. Accuracy plots (a) Xception model, (b) InceptionResNet V2, and (c) proposed CNN

Figure 5 shows the accuracy plots for three models. Figure 5(a) shows the accuracy plot of the Xception model, Figure 5(b) shows the accuracy plot for the InceptionResNet V2 model, and Figure 5(c) shows the accuracy plot of the proposed model. The blue line demonstrates the training accuracy, while the orange line represents the validation accuracy. Amongst all three graphs, the suggested CNN model's curves attain higher values and are smoother.

4.2 Analysis based on loss plots

The loss plot shows the model's performance in minimizing the error between its predicted output and the actual output for each input sample. It plots the loss of the model on the y-axis against the number of training epochs on the x-axis.



(a) (b) (c)

Fig. 6. Loss plot (a) Xception model, (b) InceptionResNet V2, and (c) proposed CNN

Usually, the loss plot shows a decreasing trend as the model learns from the data during the training process. A lower loss value indicates that the model produces predictions closer to the actual values, thus improving its overall performance. These plots display the changes in training accuracy and loss over time. Figure 6 shows the loss plots for three models. Figure 6(a) shows the loss plot of the Xception model, Figure 6(b) shows the loss plot for the InceptionResNet V2 model, and Figure 6(c) shows the loss plot of the proposed model. The blue line shows the training loss, while the orange line represents the validation loss. These plots provide insights into the learning behavior of the model during the training process, with the validation data providing an additional benchmark to evaluate the model's generalization. Figure 6 shows that the suggested CNN model has smoother curves than previous models.

4.3 Analysis based on confusion matrix

A confusion matrix compares predicted and actual values to assess the effectiveness of a classification model. Instead of relying on accuracy or other summary data, it offers a more in-depth and complete perspective of the model's performance.

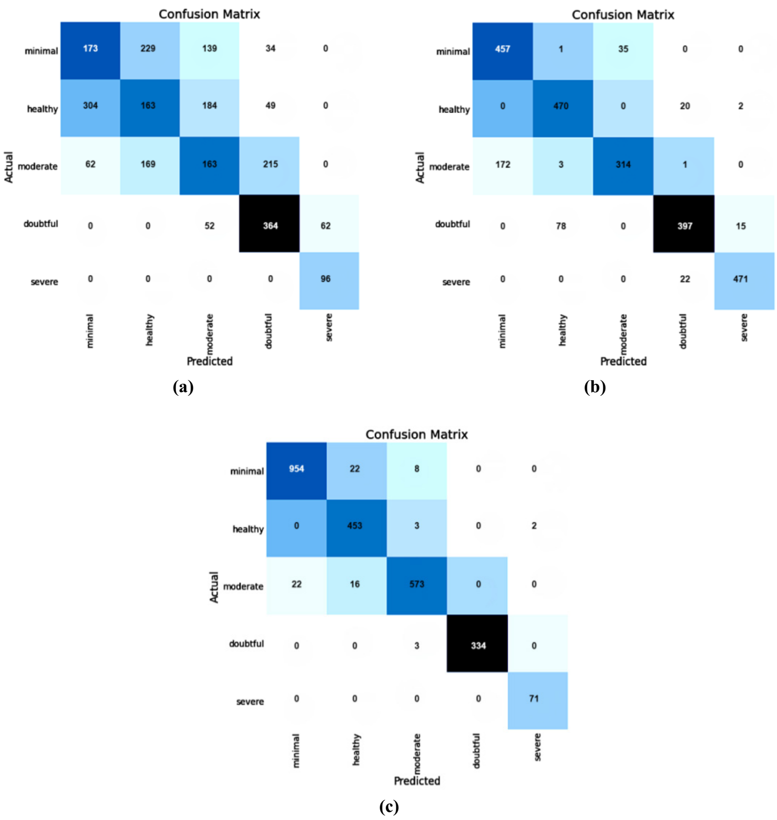


Fig. 7. Confusion matrix (a) Xception model, (b) InceptionResNet V2, and (c) proposed CNN

Figure 7 shows the confusion matrix received from three models. Figure 7(a) displays the confusion matrix of the Xception model, Figure 7(b) shows the confusion matrix of the InceptionResNet model, and Figure 7(c) depicts the confusion matrix of the proposed CNN model. This matrix gives the values of true positive, true negative, false positive, and false negative. The performance parameters such as accuracy, precision, recall, and f1 score can be calculated based on values received from the confusion matrix. Confusion matrices, in general, offer a more thorough and instructive approach to assessing the effectiveness of a classification model, and they should be used in conjunction with other assessment measures. From Figure 7, it can be concluded that the confusion matrix of the proposed CNN model has the best results compared to other transfer learning models.

4.4 Analysis based on performance parameters

Performance parameters are used to assess the performance of machine learning models. Table 2 summarizes the execution of three different machine learning models, namely Xception, InceptionResNetV2, and a proposed CNN model, in classifying different levels of disease severity in a medical dataset. The assessment metrics used to assess the models' performance include precision, recall, F1 score, and accuracy. The models are tested on five different classes of disease severity: Minimal, Healthy, Moderate, Doubtful, and Severe. According to the table, the proposed CNN model outperforms the other two models in all evaluation metrics, achieving the highest precision, recall, F1 score, and accuracy for all disease severity classes. Specifically, the expected model achieves perfect precision, recall, F1 score, and accuracy for the Severe type and very high scores for the other courses, demonstrating its superior performance in accurately classifying disease severity levels. In contrast, InceptionResNetV2 and Xception models show lower performance in some classes. For example, InceptionResNetV2 shows lower recall and F1 scores for Minimal, Healthy, and Moderate, while Xception offers lower precision and F1 scores for Moderate classes. The proposed CNN model's processing time is also less than other transfer learning models. Overall, the expected CNN model achieves better than the additional two models and can be a better choice for accurately classifying disease severity levels.

Table 2. Performance parameters comparison

Model Name	Class Name	Precision	Recall	F1 Score	Processing Time	Accuracy
InceptionResNetV2	Minimal	0.30	0.32	0.30	2h 37 m	0.68
	Healthy	0.23	0.29	0.25		
	Moderate	0.26	0.30	0.28		
	Doubtful	0.76	0.55	0.64		
	Severe	1.00	0.60	0.75		
Xception	Minimal	0.93	0.73	0.82	2 h 18m	0.91
	Healthy	0.96	0.85	0.90		
	Moderate	0.64	0.90	0.75		
	Doubtful	0.81	0.90	0.85		
	Severe	0.95	0.96	0.95		
Proposed CNN Model	**Minimal**	**0.97**	**0.98**	**0.99**	**1h 47m**	**0.98**
	Healthy	**0.99**	**0.92**	**0.95**		
	Moderate	**0.94**	**0.98**	**0.96**		
	Doubtful	**1.00**	**0.99**	**0.99**		
	Severe	**1.00**	**1.00**	**1.00**		

From the results, it can be concluded that the proposed CNN model, which is designed from scratch, outperformed the transfer learning model with 98% accuracy. The model also takes the least time to process, which is 1 hour and 47 minutes. The model achieved the highest precision, recall, and f1 scores for all five classes.

4.5 State art comparison

The State-of-the-Art comparison Table 3 presents various techniques and their results for classifying medical images using MRI and X-ray datasets. The methods used in the comparison include SDM-NN and SDM-SVM, Random Forest Classifier, Proposed Deep neural networks, DenseNet, R-CNN, and the Proposed Model, which utilizes a CNN. The results indicate that the proposed CNN model achieves the highest accuracy rate of 98% using X-ray images, significantly improving over the other techniques. The Random Forest Classifier has an accuracy rate of about 87.92%, while the SDM-SVM technique achieves the highest accuracy of 85.4% using X-ray images. Overall, the State of the art comparison highlights the effectiveness of deep learning techniques, particularly CNNs, in achieving high accuracy rates in medical image classification tasks.

Table 3. State of art comparison

Reference No./ Year of Publishing	Dataset Used	Technique	Results
[34]/ 2010	X-ray images	SDM-NN and SDM-SVM	78.8% accuracy with SDM-NN and 85.4% accuracy with SDM-SVM
[35]/ 2016	X-ray images	Random Forest Classifier	An accuracy of about 87.92 % is achieved.
[36]/ 2019	X-ray images	Proposed Deep neural networks	79.39% accuracy
[21]/ 2019	MRI images	Deep learning-based convolutional neural network (DenseNet)	75% accuracy
[37]/ 2020	X-ray images	R-CNN	74.3% accuracy 93.6% sensitivity 74.2% specificity
[38]/ 2021	MRI images	CNN	83% accuracy
Proposed Model	**X-ray images**	**CNN**	**98% accuracy**

5 Conclusions

Knee osteoarthritis (O.A.) is a chronic degenerative joint disease that affects the knee joint and causes pain, stiffness, swelling, and a limited range of motion. It happens when the cartilage in the knee joint wears away, resulting in bone-on-bone contact. Deep neural networks effectively tackle various machine learning challenges in medicine, including diagnosis, prediction, and picture categorization. Knee osteoarthritis is classified into five categories in this study: minimal, healthy, moderate, doubtful, and severe. A deep learning-based CNN model has been designed from scratch for the classification task. The proposed CNN model is compared with two transfer learning models named Xception and InceptionResNet V2 model. The implementation has been performed using x-ray images of knee osteoarthritis. The models are evaluated based on precision, recall, F1 score, and accuracy during the model performance. The proposed CNN model, designed from scratch, outperformed the transfer learning model with 98% accuracy. Xception model has 91% accuracy, and InceptionResNet V2 has 68% accuracy. The proposed model assists doctors in classifying the severity of patients based on single images quickly and accurately. The proposed model also achieved a state of the art results compared to existing models and can be a valuable tool for accurate and efficient diagnosis and treatment planning.

6 Acknowledgement

The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Future Funding program grant code (NU/SRP/SERC/12/29).

7 References

[1] Wang, Y., Wang, X., Gao, T., Du, L. and Liu, W., 2021. An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative. *Journal of Healthcare Engineering*, *2021*, pp. 1–10. https://doi.org/10.1155/2021/5586529

[2] Tolpadi, A.A., Lee, J.J., Pedoia, V. and Majumdar, S., 2020. Deep learning predicts total knee replacement from magnetic resonance images—*scientific reports*, *10*(1), p. 6371. https://doi.org/10.1038/s41598-020-63395-9

[3] Kokkotis, C., Moustakidis, S., Papageorgiou, E., Giakas, G. and Tsaopoulos, D.E., 2020. Machine learning in knee osteoarthritis: a review. *Osteoarthritis and Cartilage Open*, *2*(3), p. 100069. https://doi.org/10.1016/j.ocarto.2020.100069

[4] Guan, B., Liu, F., Mizaian, A.H., Demehri, S., Samsonov, A., Guermazi, A. and Kijowski, R., 2022. Deep learning approach to predict pain progression in knee osteoarthritis. *Skeletal radiology*, pp. 1–11.

[5] Yeoh, P.S.Q., Lai, K.W., Goh, S.L., Hasikin, K., Hum, Y.C., Tee, Y.K. and Dhanalakshmi, S., 2021. The emergence of deep learning in knee osteoarthritis diagnosis. *Computational intelligence and neuroscience*, *2021*, pp. 1–20. https://doi.org/10.1155/2021/4931437

[6] Kim, D.H., Lee, K.J., Choi, D., Lee, J.I., Choi, H.G. and Lee, Y.S., 2020. Can additional patient information improve the diagnostic performance of deep learning for the interpretation of knee osteoarthritis severity. *Journal of Clinical Medicine*, *9*(10), p. 3341. https://doi.org/10.3390/jcm9103341

[7] Ahmed, S.M. and Mstafa, R.J., 2022. A comprehensive survey on bone segmentation techniques in knee osteoarthritis research: from conventional methods to deep learning. *Diagnostics*, *12*(3), p. 611. https://doi.org/10.3390/diagnostics12030611

[8] Swiecicki, A., Li, N., O'Donnell, J., Said, N., Yang, J., Mather, R.C., Jiranek, W.A. and Mazurowski, M.A., 2021. Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. *Computers in biology and medicine*, *133*, p. 104334. https://doi.org/10.1016/j.compbiomed.2021.104334

[9] Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P. and Saarakkala, S., 2018. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. *Scientific reports*, *8*(1), pp. 1–10. https://doi.org/10.1038/s41598-018-20132-7

[10] Gan, H.S., Ramlee, M.H., Wahab, A.A., Lee, Y.S. and Shimizu, A., 2021. From classical to deep learning: review on cartilage and bone segmentation techniques in knee osteoarthritis research. *Artificial Intelligence Review*, *54*(4), pp. 2445–2494. https://doi.org/10.1007/s10462-020-09924-4

[11] Leung, K., Zhang, B., Tan, J., Shen, Y., Geras, K.J., Babb, J.S., Cho, K., Chang, G. and Deniz, C.M., 2020. Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: data from the osteoarthritis initiative. *Radiology*, *296*(3), pp. 584–593. https://doi.org/10.1148/radiol.2020192091

[12] Jakaite, L., Schetinin, V., Hladůvka, J., Minaev, S., Ambia, A. and Krzanowski, W., 2021. Deep learning for early detection of pathological changes in x-ray bone microstructures: case of osteoarthritis. *Scientific Reports*, *11*(1), pp. 1–9. https://doi.org/10.1038/s41598-021-81786-4

[13] Chang, G.H., Park, L.K., Le, N.A., Jhun, R.S., Surendran, T., Lai, J., Seo, H., Promchotichai, N., Yoon, G., Scalera, J. and Capellini, T.D., 2021. Subchondral bone length in knee osteoarthritis: a deep learning–derived imaging measure and its association with radiographic and clinical outcomes. *Arthritis & Rheumatology*, *73*(12), pp. 2240–2248. https://doi.org/10.1002/art.41808

[14] Panfilov, E., Tiulpin, A., Nieminen, M.T., Saarakkala, S. and Casula, V., 2022. Deep learning-based segmentation of knee MRI for fully automatic subregional morphological assessment of cartilage tissues: data from the osteoarthritis initiative. *Journal of Orthopaedic Research®, 40*(5), pp. 1113–1124. https://doi.org/10.1002/jor.25150

[15] Kotti, M., Duffell, L.D., Faisal, A.A. and McGregor, A.H., 2013. Towards automatically assessing osteoarthritis severity by regression trees & SVMs.

[16] Kotti, M., Duffell, L.D., Faisal, A.A. and McGregor, A.H., 2017. Detecting knee osteoarthritis and its discriminating parameters using random forests. *Medical engineering & physics*, *43*, pp. 19–29. https://doi.org/10.1016/j.medengphy.2017.02.004

[17] Köktaş, N.Ş., Yalabik, N., Yavuzer, G. and Duin, R.P., 2010. A multi-classifier for grading knee osteoarthritis using gait analysis. *Pattern Recognition Letters*, *31*(9), pp. 898–904. https://doi.org/10.1016/j.patrec.2010.01.003

[18] de Dieu Uwisengeyimana, J. and Ibrikci, T., 2017. Diagnosing knee osteoarthritis using artificial neural networks and deep learning. *Biomedical Statistics and Informatics*, *2*(3), p. 95.

[19] Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H. and Lespessailles, E., 2019. A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: data from the osteoarthritis initiative. *Computerized Medical Imaging and Graphics*, *73*, pp. 11–18. https://doi.org/10.1016/j.compmedimag.2019.01.007

[20] Antony, J., McGuinness, K., Moran, K. and O'Connor, N.E., 2017. 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, MLDM 2017, New York, NY, USA, July 15-20, 2017, Proceedings 13* (pp. 376–390). Springer International Publishing. https://doi.org/10.1007/978-3-319-62416-7_27

[21] Pedoia, V., Lee, J., Norman, B., Link, T.M. and Majumdar, S., 2019. Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort. *Osteoarthritis and cartilage*, *27*(7), pp. 1002–1010. https://doi.org/10.1016/j.joca.2019.02.800

[22] Anand, V., Gupta, S., Koundal, D., Nayak, S.R., Barsocchi, P. and Bhoi, A.K., 2022. Modified U-net architecture for segmentation of skin lesion. *Sensors*, *22*(3), p. 867. https://doi.org/10.3390/s22030867

[23] Kate, V., & Shukla, P. (2021). Breast cancer image multi-classification using random patch aggregation and depth-wise convolution based deep-net model. *International Journal of Online and Biomedical Engineering (iJOE)*, *17*(01), pp. 83–100. https://doi.org/10.3991/ijoe.v17i01.18513

[24] Al-Rahlawee, A.T.H. and Rahebi, J., 2021. Multilevel thresholding of images with improved Otsu thresholding by black widow optimization algorithm. *Multimedia Tools and Applications*, *80*(18), pp. 28217–28243. https://doi.org/10.1007/s11042-021-10860-w

[25] Yousefi, J., 2011. Image binarization using Otsu thresholding algorithm. *Ontario, Canada: University of Guelph*, *10*.

[26] Singh, S., Aggarwal, A.K., Ramesh, P., Nelson, L., Damodharan, P. and Pandian, M.T., 2022, August. COVID 19: Identification of Masked Face using CNN Architecture. In *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1045–1051). IEEE. https://doi.org/10.1109/ICESC54411.2022.9885327

[27] Sharma, N., Gupta, S., Mehta, P., Cheng, X., Shankar, A., Singh, P. and Nayak, S.R., 2022. Offline signature verification using deep neural network with application to computer vision. *Journal of Electronic Imaging*, *31*(4), pp. 041210–041210. https://doi.org/10.1117/1.JEI.31.4.041210

[28] Prema, C.E., Suresh, S., Krishnan, M.N. and Leema, N., 2022. A novel efficient video smoke detection algorithm using co-occurrence of local binary pattern variants. *Fire Technology*, *58*(5), pp. 3139–3165. https://doi.org/10.1007/s10694-022-01306-2
[29] Irawati, I. D., Andrea Larasaty, I. and Hadiyoso, S., 2022. Comparison of Convolution Neural Network Architecture for Colon Cancer Classification. *International Journal of Online and Biomedical Engineering (iJOE)*, *18*(03), pp. 164–172. https://doi.org/10.3991/ijoe.v18i03.27777
[30] Soni, A.N., 2018. Application and analysis of transfer learning-survey. *International Journal of Scientific Research and Engineering Development*, *1*(2), pp. 272–278.
[31] Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251–1258). https://doi.org/10.1109/CVPR.2017.195
[32] Baldassarre, F., Morín, D.G. and Rodés-Guirao, L., 2017. Deep koalarization: Image colorization using cnns and inception-resnet-v2. *arXiv preprint arXiv:1712.03400*.
[33] Singh, R., Sharma, A., Sharma, N. and Gupta, R., 2023, January. Impact of Adam, Adadelta, SGD on CNN for White Blood Cell Classification. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1702–1709). IEEE. https://doi.org/10.1109/ICSSIT55814.2023.10061068
[34] Woloszynski, T., Podsiadlo, P., Stachowiak, G.W. and Kurzynski, M., 2010. A signature dissimilarity measure for trabecular bone texture in knee radiographs. *Medical physics*, *37*(5), pp. 2030–2042. https://doi.org/10.1118/1.3373522
[35] Gornale, S.S., Patravali, P.U. and Manza, R.R., 2016. Detection of osteoarthritis using knee x-ray image analyses: a machine vision based approach. *Int. J. Comput. Appl*, *145*(1), pp. 20–26. https://doi.org/10.5120/ijca2016910544
[36] Christodoulou, E., Moustakidis, S., Papandrianos, N., Tsaopoulos, D. and Papageorgiou, E., 2019, July. Exploring deep learning capabilities in knee osteoarthritis case study for classification. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (pp. 1–6). IEEE. https://doi.org/10.1109/IISA.2019.8900714
[37] Samreen, A., Taha, A. M., Reddy, Y. V. and P, S. (2020). Brain tumor detection by using convolution neural network. *International Journal of Online and Biomedical Engineering (iJOE)*, *16*(13), pp. 58–69. https://doi.org/10.3991/ijoe.v16i13.18545
[38] Guida, C., Zhang, M. and Shan, J., 2021. Knee osteoarthritis classification using 3D CNN and MRI. *Applied Sciences*, *11*(11), p. 5196. https://doi.org/10.3390/app11115196

8 Authors

Saleh Hamad Sajaan Almansour did his bachelor's from Najran Technical College in computer network systems support. He is currently a master's student of cybersecurity at the Department of Information Systems, College of Computer Science and Information Systems, Najran University (email: shsajaan@nu.edu.sa).

Rahul Singh is currently pursuing ME fellowship in electronics and communication engineering from Chitkara University, Punjab. He has completed his B.Tech degree in electrical engineering from Sri Sukhmani Institute of Engg. & Tech. Derabassi, Punjab. His areas of interest are artificial intelligence, digital Image processing, machine learning and deep learning (email: rahul.2414@chitkara.edu.in).

Saleh Mabruk Hadrami Alyami did his bachelors from Incarnate Word University in the USA. He is currently working in the Ministry of Health Saudi Arabia as a software designer and doing a master's in cybersecurity at the Department of Information Systems, College of Computer Science and Information Systems, Najran University. (email: salyami4111@gmail.com).

Neha Sharma is currently pursuing her Ph.D. degree in electronics and communication engineering from Chitkara University, Punjab. She has completed her master's degree in electronics and communication engineering from Chitkara University, Punjab. She received her bachelor's degree in electronics and communication engineering from Kurukshetra University, Haryana. Her areas of interest are artificial intelligence, digital Image processing, machine learning, deep learning, and computer vision (email: sharma.neha@chitkara.edu.in).

Mana Saleh Al Reshan received a B.S. degree in information systems from King Khalid University, Abha, Saudi Arabia, in 2007, an M.S. degree (Hons.) in computer, information, and network security from DePaul University, Chicago, USA, in 2011, and the Ph.D. degree in computer science from The Catholic University of America (CUA), Washington, DC, USA, in 2019. He is currently an Assistant Professor and the Head of the Network Engineering Department. His current research interests include computer network and security, system security, wireless and mobile security, body area networks, and cloud security. (email: msalreshan@nu.edu.sa).

Sheifali Gupta is a professor in the Chitkara University Research and Innovation Network (CURIN) at Chitkara University, Punjab Campus-India. She specializes in the areas of digital image processing, pattern recognition, machine intelligence, biomedical image processing, agri- culture based image processing, and deep learning. She has published more than 100 research papers and articles in reputed national & international journals and conferences. She has received a prestigious IRDP Award 2018 for remarkable achievements in teaching, research, and publications (sheifali.gupta@chitkara.edu.in).

Mahdi Falah Mahdi Alyami did his bachelors from the Gulf University of Science and Technology (Kuwait) and he is currently a master's student of cybersecurity at the Department of Information Systems, College of Computer Science and Information Systems, Najran University (email: 444100188@nu.edu.sa).

Asadullah Shaikh received a B.Sc. degree in software development from the University of Huddersfield, U.K., an M.Sc. degree in software engineering and management from Goteborg University, Sweden, and a Ph.D. degree in software engineering from the University of Southern Denmark. He was a Researcher at UOC Barcelona, Spain. He is currently a Professor, the Head of research and graduate studies, and the Coordinator of seminars and training with the College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia (email: asshaikh@nu.edu.sa).

Article submitted 2023-03-05. Resubmitted 20232-04-12. Final acceptance 2023-04-13. Final version published as submitted by the authors.