Received: July 5, 2023. Revised: August 15, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Classification of Knee X-Ray Images by Severity of Osteoarthritis Using Skip Connection Based ResNet101

Malathi Salejangal Yallappa[1]* **Geeta Rudragouda Bharamagoudar[1]**

[1]*Department of Computer Science, KLE Institute of Technology, Hubballi, Visvesvaraya Technological University, Belagavi-590018, India*
* Corresponding author's Email: malathisy@kleit.ac.in

Abstract: Knee osteoarthritis (OA) is a prevalent condition that typically affects elderly individuals. It occurs due to the gradual deterioration of hyaline cartilage located between knee joints. If it not treated in earlier stage, it causes knee replacement so the knee OA early diagnosis is essential for better treatment. The knee OA diagnosis includes knee X-ray images and classifying them using kellgren-lawrence (KL) grading system. This paper proposed the application of pre-trained models named as skip connection based ResNet101 for knee OA images from osteoarthritis initiative (OAI) dataset. The skip connection based ResNet101 is utilized to overcome the vanishing gradient problems. Moreover, this skip connection allows learning of the recognized functions. This enables the ResNet101 model have performance high-level layers compared to the low-level layers. Due to these features the skip connection based ResNet101 model is utilized in this paper. The two various classifications are performed in this paper such as binary and severity. The binary classification determines whether knee osteoarthritis (OA) is absent or present, which also categorizing the severity of KOA into three grades. The experiments are conducted on three various datasets named dataset I, II & III with five, two and three grades of knee OA images. The obtained result shows that the skip connection based ResNet101 model achieves better accuracy of 79.94%, 84.96% and 95.86% in the above-mentioned three datasets when compared to ResNet101 and InceptionNetV2 models.

Keywords: Diagnosis, Knee osteoarthritis, ResNet101, Severity classification, X-ray images.

1. Introduction

Knee osteoarthritis (OA) is a joint disorder caused by hyaline cartilage among the joints, which affects the bones of the knee touching and rubbing along with each other. In common, it happens in the diarthrosis and output from the group of injury, overuse, and genetic factors [1]. It happens with age and it affects mostly elderly people around the world. OA is the second most rheumatological disease in India and it has an occurrence of 22-39%. Heavyweight is among the several causes of the occurrence of OA [2]. Among the two bones, a heavy material is presented named cartilage which is helpful for the smoothness and flexible movement of the knee [3]. This knee joint contains two bones named as tibia and femur. This cartilage volume may be reduced because of aging loss [4]. The tibiofemoral bones generate erosion in the knee movement, which leads to knee OA [5]. The hyaline cartilage is comprised of cartilage cells that are useful for bone by load allocation and it works for lifelong [6].

The kellgren-lawrence (KL) grading system is established on the classification of radiographic of knee OA and this system determines the different stages of OA. It contains grade ranges from 0 to 4 to compute the knee OA [7]. Grade 0 indicates the healthy knee, grade 1 specifies the doubtful knee, grade 2 specifies the mild knee, grade 3 specifies the moderate knee and finally, grade 4 specifies the severe OA [8]. The initial symptoms in knee OA patients are knee pain, swelling, stiffness, tenderness grating sensation, etc. For this kind of reason, doctors reduce the occurrence of the disease [9]. In deciding this initial diagnosis technique is essential to identify the severe knee OA [10]. Some models are utilized proficient techniques like nuclear, MRI, X-ray, CT, and UV [11]. Among these techniques, the X-ray is still the gold standard technique and is widely used due to its simplicity, convenience, and cheapness [12]. Besides that, the X-ray can replicate the initial changes in the architecture of bone. The knee joint is evaluated through the extensive knee X-ray which is used to estimate Joint Space Narrowing (JSN), subchondral, and bone spurs [13]. The major contribution of this research is specified as follows,

- The preprocessing contains two stages, first stage is segmentation which contains cropping the image into the chosen area, so at the end the undesirable areas are omitted from the image. The second stage is equalization which improves the contrast by altering the distribution intensity for clear visibility of the image chosen area.
- This paper proposed the application of pre-trained CNN model named as skip connection based ResNet101 model for knee OA from the osteoarthritis initiative (OAI) dataset.
- The two types of classification are performed in this paper as binary and severity. The binary classifies the absence or presence of knee OA and the severity classifies KOA in three grades.

The rest of the manuscript is organized as follows: Section 2 illustrates the literature review. The block diagram of the proposed model is presented in section 3. The experimental result of this proposed model is illustrated in section 4. The conclusion of this paper is given in section 5.

2. Literature review

Mohammed [14] introduced a detection and severity classification of knee osteoarthritis by utilizing residual neural networks on X-ray images. In this paper, the six various models are trained such as ResNet101, InceptionResNetV2, VGG16, VGG18, DenseNet121 and MobileNetV2 on osteoarthritis initiative (OAI) dataset, which includes 9786 images. Then, the information is given to the model, that predicts KOA severity based on KL grading systems. The high KOA occurrence involves the requirement for reliable, precise and automatic severity classification. The advantage of using this model is to provide focus on identifying features in knee images that affects the decision of the network. The highly unbalanced and small dataset makes this model performs difficult with better accuracy.

Kumar and Goswami [15] developed an automatic severity classification of knee osteoarthritis by using convolutional neural network and enhanced image sharpening methods. The kellgren-lawrence system was utilized to train this model for producing exact predictions of the severity of knee disease. The image sharpening technique enhance the performance of five-grade KOA severity classification. Through, combining patient data allows the element identification that impacts the disease development which is necessary for medical specialist to generate accurate diagnosis. The advantage of this model is that automatically learns the contextual information without the need for any computing spatial structure method. The variety of samples was limited and it required maximum training time.

Sarvamangala and Kulkarni [16] implemented an automatic grading and classification of knee osteoarthritis for multiscale convolutional blocks in convolutional neural networks (MCBCNN). This model was implemented using pre-trained CNN models like ResNet50, MobileNet2, and InseptionNetV3. The pretrained models are given into MCBCNN and fine-tuned on knee OA dataset. The multiscale convolutional block contains varying size of ReLU activations, conv filters and highest pooling layers. The pre-trained model minimizes the training time and the multiscale convolutional block was used for fine feature extraction at various scales and it improve the effectiveness of the classification. Information about the patient like age, demographic location, and history of injury is not available.

Liu [17] introduced an automatic quantification of Knee Osteoarthritis based on an improved faster R-CNN model. This model was used for processing the input images with the classification and location and it contains region proposal network (RPN). The RPN is trained for producing region proposals that consist of knee joints then it is utilized by classification for the faster R-CNN. Because of the localized classification using CNN, the unusable informations in X-ray images are filtered and extracted relevant clinical features. This model is efficiently achieving automatic diagnosis of the knee and used as a CAD kit in medical applications. This method only investigates the image data and does not produce other informations about the patient.

Tiulpin and Saarakkala [18] developed an automatic grading of knee osteoarthritis in plain radiographs using deep CNN. This method utilized transfer learning from ImageNet with fine-tuning on the OAI and MOST datasets with 50 layers. This method was used to predict the osteoarthritis research society international (OARSI) and kellgren-lawrence systems from the knee radiograph. This method determines the probability for predict the individual knee OA accurately and all knee OA severity from radiographs concurrently. This method is helpful for

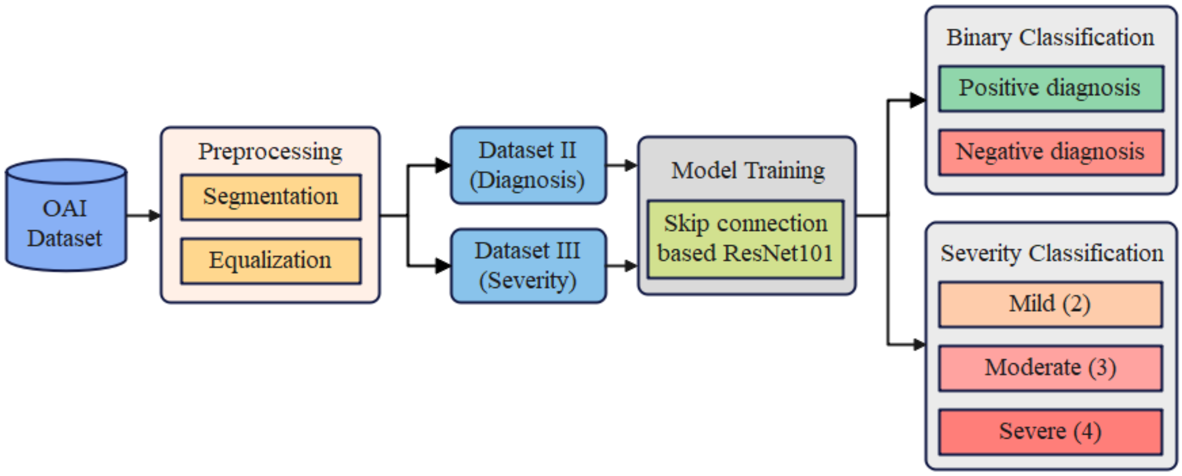

Figure. 1 Block diagram of the proposed method

medical OA and produce better information about the knee of the patient. The size of the sample set was quite small while compared to various other OA methods.

Sozan Mohammed Ahmed and Ramadhan J. Mstafa [19] suggested a severity grading of knee osteoarthritis from X-ray images based on effective deep learning and machine learning models. The developed model split into two techniques by applying pretrained CNN for feature extraction and fine-tuned the pretrained CNN using transfer learning method. This developed model integrates deep and hand-crafted features attained by pretrained CNN and PCA algorithm to produce significant feature set previously sent to the SVM for classification. The transfer learning is utilized for fine tuning the pretrained CNN which provides five class labels. The method minimizes the computational load and low cost thus the CNN model is faster and easier. The limitation of the developed model is class imbalance which detect the minority class incorrectly.

3. Proposed method

This section represents four different stages such as dataset, preprocessing, model training, and classification. The knee OA x-ray image is attained from osteoarthritis initiative (OAI) dataset which is accessible on Kaggle. This dataset has five various grade images such as grade 0 indicates the healthy knee, grade 1 specifies the doubtful knee, grade 2 specifies the mild knee, grade 3 specifies the moderate knee and finally, grade 4 specifies the severe OA. In the preprocessing stage, segmentation, and equalization are performed in the image. These preprocessed images were considered in Dataset 1. In this paper, two types of classification are performed in knee OA so preprocessed images are arranged in new two datasets named Dataset II & III. Dataset II contains two grades named positive diagnosis (grades 2 to 4) and negative diagnosis (grades 0 and 1) which is used for binary classification. Dataset III contains three grades (2, 3, and 4) which are used for the classification of severity. For experimental, all datasets are split into train, test, and validation sets of 7:2:1 ratio respectively. The overall process involved in the proposed method is presented in Fig. 1.

3.1 Dataset

The knee OA x-ray image is attained from osteoarthritis initiative (OAI) dataset [20] which is available on Kaggle. There are 9786 images, which are split into five various stages based on kellgren-lawrence (KL) grading systems as grade 0 indicates the healthy knee, grade 1 specifies the doubtful knee, grade 2 specifies the mild knee, grade 3 specifies the moderate knee and finally, grade 4 specifies the severe OA. Each image resolution is 224×224 pixels. 40% of the dataset images belong to the healthy grade compared to 18% of doubtful, 26% of mild, 13% of moderate and 3% of severe images approximately. From the original dataset, two more datasets are derived, named Dataset II & III. Dataset II is the binary dataset that contains two grades named positive diagnosis (grades 2 to 4) and negative diagnosis (grades 0 and 1) which is used for binary classification. Dataset III is the severity dataset which contains three grades (2, 3 and 4) which is used for severity classification.

3.2 Preprocessing

The three dataset images are gone through two various preprocessing stages, first stage is segmentation which contains cropping the image into the chosen area, so at the end the undesirable areas are omitted from the image. This stage is obtained by cropping the image from both upper and lower by 60 pixels. Then the resolution was taken to 224×104. The second stage is equalization which improves the contrast by altering the distribution intensity for clear visibility of the image chosen area. The equalization is performed on images by using Eq. (1),

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) \quad (1)$$

Where, s and L represent the input, output, and highest pixel score in the image. The occurrence of r_j the intensity level is presented using below Eq. (2),

$$p_r(r_j) = \frac{n_j}{MN} \quad (2)$$

Where, n_j is the number of pixels, MN is the entire number of pixels in the image that have an r_j intensity.

3.3 Convolutional neural network (CNN)

The AI domain has quick development in recent years and has been employed in different areas such as computer vision. The purpose of computer vision is to allow computers which able to view the world humanely. The desired output is automatically removed and process the applicable data from the atmosphere. Different types of algorithms have been developed to accomplish the goal and the CNN has specifically applied and achieved better results. The CNN takes an input and allocates weights to several image features such as that the image is different from various images that are managed by similar algorithms. CNN has a unique capability to capture the spatial dependencies of the images [21]. This CNN converts the image into one form that is geometrically simple while managing the crucial features that are obtainable in the image. The essential block of CNN is convolution operation which is a distinctive feature among regular neural networks and CNN. The main role of the convolution operation is to extract the highest feature from images. Another important operation is the pooling operation which mainly removes image dimensionality in energy to minimize computational complications that need to develop the input information. Both convolutional and pooling operations are presented as layers in CNN. Both are combined and form CNN i th layer and it is dependable for CNN feature extraction. These features are stuffed into a neural network for classification. Transfer learning is an efficient method to utilize problems in DL. Transfer learning is utilized to learn information and features of CNN and is trained on huge datasets. Over the past few years, certain architectures have been accomplished specifically on the benchmark and verified to be extremely effective. In the domain of DL, some popular names are MobileNetV2, ResNet, VGGNet, etc. The ResNet101 architecture is considered as best CNN model for detecting and classifying knee OA. All layer that minimizes the model performance is skipped with the help of residual block regulation represented in this architecture. The next section presented the architecture of ResNet101.

3.4 ResNet101 architecture

The current research performed in the domain of DL appeared to confirm that comes to CNN, a deep model is best in every time. The neural network is coming in to deep, the gradient loss function is zero after various chain rule applications. When the gradient is zero, then weight of the method is stopped and learning not conducted. This architecture is used to solve the gradient problems by utilizing the residual blocks [22]. This residual block consists of skip connections, that connect the activation from one layer to another layer by skipping in between layers. The ResNet model is created by stacking the numerous residual blocks. The benefit of utilizing skip connection is regularization. By using regularization, all layer minimizes the model performance is need to be skipped. This enables the DNN model without considering the gradient problems. This ResNet101 model utilized the 101 layers that are presented in Fig. 2.

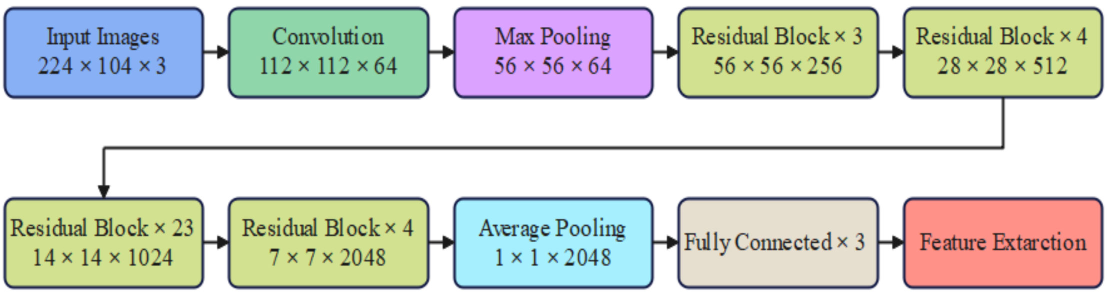

Figure. 2 Architecture of ResNet101

Received: July 5, 2023. Revised: August 15, 2023.

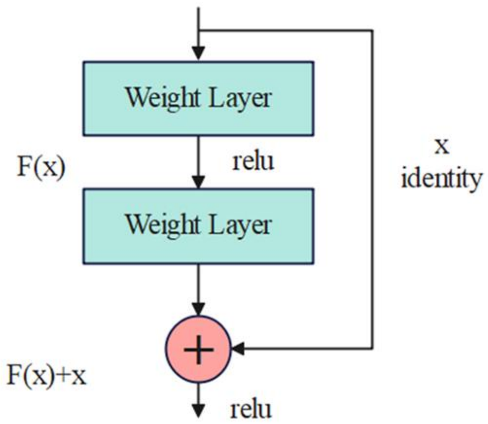

Figure. 3 Skip connection based ResNet101

3.4.1. Skip conection based ResNet101

The popular neural network utilizes 16-30 layers in deep models for classifying images. For these deep networks, the convergence is restricted by a vanishing gradient problem. To resolve this problem, the intermediate and normalized initialization is employed which allows numerous layers to connect backpropagation with stochastic gradient descent. Through, this cause a degradation problem where accuracy become saturated and starts degrading with the improvement in network. The skip connection is utilized to overcome the degradation problems which is shown in Fig. 3.

This skip connection is an essential part of residual networks and act in skipping connection which producing various outputs on the certain layer. Without skip connection in network, an input x is multiplied by the bias and weight addition layers. Then it goes over an activation function $f(x)$ which generates an output $H(x)$ is formulated in Eq. (3) and (4),

$$H(x) = f(wx + b) \tag{3}$$

$$H(x) = f(x) \tag{4}$$

Once the skip connection is introduced then the output converts to Eq. (5). But, the convolutional and pooling layers are impediment exterior where the input and output dimensions are different.

$$H(x) = f(x) + x \tag{5}$$

To resolve this issue, the dimensions are equaled by adding 1×1 convolution layers to input, this is known as projection technique which is formulated in Eq. (6), The next approach is occupied by filling the skip connection with additional zero for improving the dimensions where calculation of $w1$ is evaded.

$$H(x) = f(x) + w1.x \tag{6}$$

In the vanishing gradient problems, the activation function with numerous layers can cause the gradient loss function to zero which makes difficult to train the neural network. The skip connection based ResNet101 is useful for avoiding this kind of problems and these skip connections are alternative way for gradient. Moreover, this skip connection allows learning of the recognized functions. This enables the ResNet101 model have performance high-level layers compared to the lower-level layers. Due to these features the skip connection based ResNet101 is utilized in this paper.

3.5 Performance metrics

In this section, various performance estimation metrics are utilized for estimate the performance of classifiers and it is essential to estimate the model performance. The accuracy, precision, recall and F1-score are utilized for estimate model performance. When the test case is in the positive section and the classifier accurately predicts it is a positive test case, it is defined as true positive (TP). When a positive test case is incorrectly predicted to be a negative test case, it is known as false positive (FP). Likewise, when a negative test case is accurately predicted it is a negative test case, it is defined as true negative (TN). When a positive test case is incorrectly predicted to be a negative test case, it is known as false negative (FN).

Accuracy:

Accuracy is a metric that is used for estimating the model performance. It is determined by using the ratio values of accurate predictions to the total amount of test cases in the dataset. The mathematical formula of accuracy is represented in Eq. (7),

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{7}$$

Precision:

Precision is a metric that calculates the accuracy of positive predictions. It is determined using the ratio of the amount of true positive and false positive. The precision establishes the model consistency by classifying the test case as positive. The mathematical formula of precision is represented in Eq. (8),

$$Precision = \frac{TP}{TP+FP} \tag{8}$$

Recall:

The recall is a metric that calculates the

Received: July 5, 2023. Revised: August 15, 2023.

Table 1. Represents the various classifiers on dataset I and contains 5 grades

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	77.35	77.63	76.59	75.46
EfficientNetB5	74.69	75.51	77.85	74.85
ResNet152V2	76.12	76.38	75.90	73.97
skip connection based ResNet101	79.94	78.82	78.21	77.51

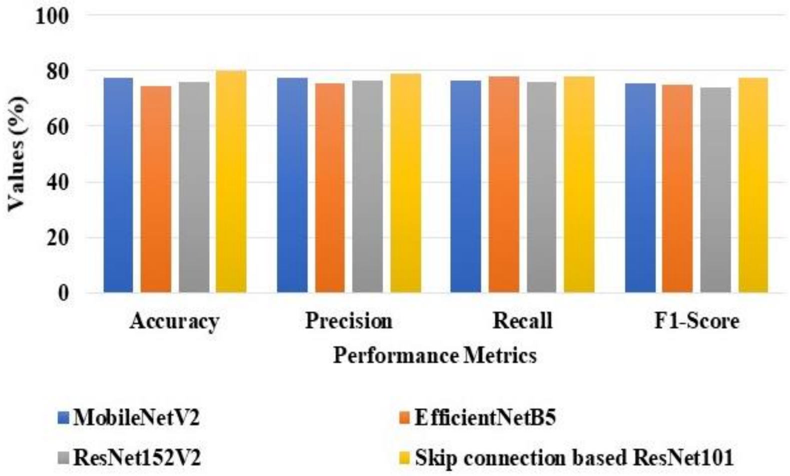

Figure. 4 Performance of the various classifier on dataset I contains 5 grades

completeness of positive predictions. It is determined using a ratio among the variety of true positives and amount of test cases in the dataset. The recall establishes model capability by accurately classifying test cases as positive. The mathematical formula of recall is represented in Eq. (9),

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

F1-score:

F1-score is a metric that calculates the model accuracy to combine the precision and recall values. The mathematical formula of the f1-score is represented in Eq. (10),

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

Where,

- True positives (TP) – classifies the positive classes.
- True negatives (TN) – classifies the negative classes.
- False positives (FP) – misclassification the predicted outcome is “yes” but the actual outcome is “no”.
- False negatives (FN) – misclassification of the predicted outcome is “no” but the actual outcome is “yes”.

4. Result

This section demonstrates the experimental results attained from this proposed methodology. The experiments are managed on three various datasets, the original dataset named, OAI and two extracted datasets, named Dataset II and III correspondingly. Dataset II is the binary dataset which contains two grades named positive diagnosis (grades 2 to 4) and negative diagnosis (grades 0 and 1) which is used for binary classification. Dataset III is the severity dataset which contains three grades (2, 3 and 4) used for severity classification. All the datasets are split in train, test and validation set 7:2:1 ratio respectively. This experiment is performed by using Python programming language. The derived two datasets are used for creating a multistage diagnosis system. The first stage is the detection of knee OA and the next step is to diagnose the severity.

4.1 Quantitative analysis

This section shows the quantitative analysis of the skip connection based ResNet101 model in the performance evaluation of accuracy, precision, recall, and F1-score. Table 1 represents the various classifiers on the original dataset (Dataset I) which contains 5 grades. Table 2 represents the various classifiers on Dataset II which consists of 2 grades.

Table 2. Represents the various classifiers on dataset II and contains 2 grades

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	79.64	78.76	79.19	78.94
EfficientNetB5	81.96	80.49	78.91	79.48
ResNet152V2	83.21	82.12	80.25	79.98
skip connection based ResNet101	84.96	83.26	81.92	81.63

Table 3. Represents the various classifiers on Dataset III and contains 3 grades

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	90.13	90.53	89.95	89.98
EfficientNetB5	91.62	89.76	89.28	90.69
ResNet152V2	93.75	91.49	90.74	90.16
skip connection based ResNet101	95.86	93.72	92.51	91.89

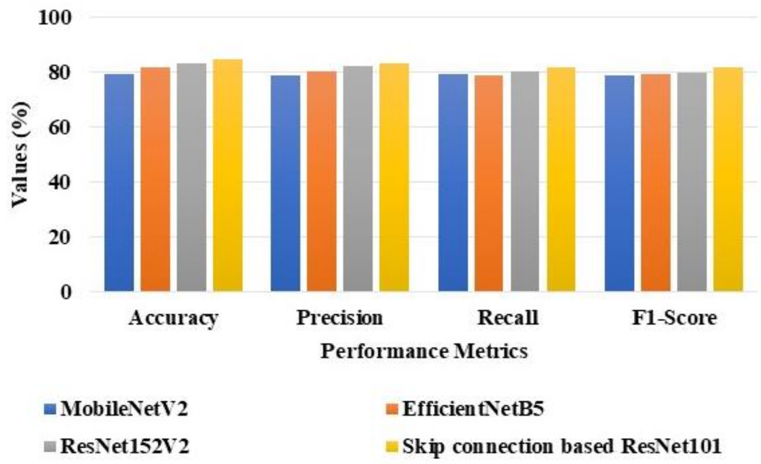

Figure. 5 Performance of the various classifiers on dataset II contains 2 grades

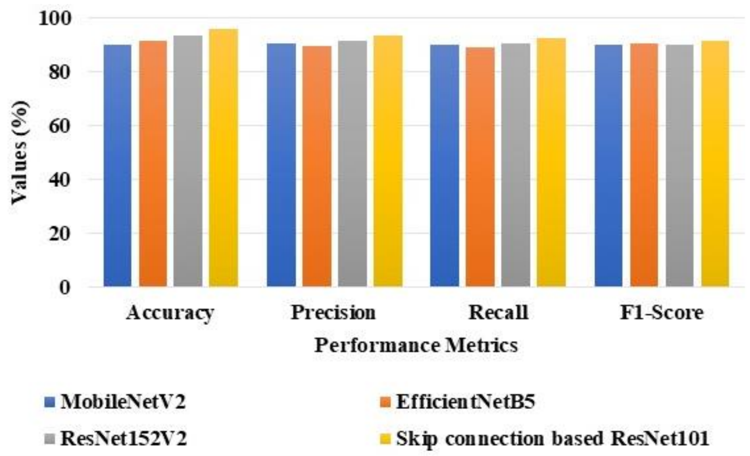

Figure. 6 Performance of the various classifiers on dataset III contains 3 grades

Table 3 represents the various classifiers on Dataset III which consists of 3 grades.

Fig. 4 represents the performance measure of various classifiers on the original (OAI) dataset named (Dataset I) which contains 5 grades (0 to 4). Accuracy, precision, recall and F1-Score of MobileNetV2, EfficientNetB5 and ResNet152V2 are measured and matched with the proposed skip connection based ResNet101 model. The skip connection based ResNet101 achieves better results by using accuracy, precision, recall, and f1-scores values of about 79.94%, 78.82%, 78.21% and 77.51% respectively when compared to various other classifiers.

Fig. 5 represents the performance measure of various classifiers on dataset II. Dataset II is the

Table 4. Comparative analysis of the proposed method with other various methods

Author	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Mohammed [14]	ResNet101	89	87	86	86
Kumar and Goswami [15]	InceptionNetV2	91.02	89	N/A	88
Liu [17]	Faster R-CNN	79.3	76.5	76.5	76.2
Sozan Mohammed Ahmed and Ramadhan J. Mstafa [19]	CNN	62	N/A	59	N/A
Proposed method	Skip connection based ResNet101	95.86	93.72	92.51	91.89

binary dataset that contains two grades named positive diagnosis (grades 2 to 4) and negative diagnosis (grades 0 and 1) which is used for binary classification. Accuracy, precision, recall, and F1-Score of MobileNetV2, EfficientNetB5 and ResNet152V2 are measured and matched with the proposed skip connection based ResNet101 model. The skip connection based ResNet101 achieves better results by using accuracy, precision, recall and f1-scores values of about 84.96%, 83.26 %, 81.92%, and 81.63% respectively when compared to various other classifiers.

Fig. 6 represents the performance measure of various classifiers on dataset III. Dataset III is the severity dataset which contains three grades (2, 3, and 4) that are used for severity classification. Accuracy, precision, recall, and F1-Score of MobileNetV2, EfficientNetB5 and ResNet152V2 are measured and matched with the proposed skip connection based ResNet101 model. The skip connection based ResNet101 achieves better results using accuracy, precision, recall, and f1-scores values of about 95.86%, 93.72 %, 92.51% and 91.89% respectively when compared to various other classifiers.

4.2 Comparative analysis

This section demonstrates the comparative analysis of skip connection based ResNet101 model with performance metrics like Accuracy, Precision, Recall and F1-Score as presented in Table 4. Existing research such as [14, 15, 17, 19] are used for evaluating the ability of the proposed method. The skip connection based ResNet101 is utilized to overcome the limitations of existing methods and vanishing gradient problems. Moreover, this skip connection allows learning of the recognized functions. This enables the ResNet101 model have performance high-level layers compared to the low-level layers. Due to these features the skip connection based ResNet101 model is utilized in this paper. This proposed method is trained, tested, and verified by utilizing three various datasets, the original dataset named, OAI (dataset I) and two derivative datasets, named as dataset II & III respectively. Dataset II is the binary dataset which contains two grades named positive diagnosis (grades 2 to 4) and negative diagnosis (grades 0 and 1) which is used for binary classification. Dataset III is the severity dataset which contains three grades (2, 3, and 4) that are used for severity classification. The accuracy, precision, recall and f1-score were improved by 95.86%, 93.72%, 92.51%, and 91.89% respectively.

5. Conclusion

This manuscript focused on the recognition and classification of knee osteoarthritis (OA) which is a challenging disease in aged people. In this paper, deep neural network (DNN) is used which has the advantage of automatic feature extraction in X-ray images. The kellgren-lawrence system is utilized to train the pre-trained model to produce exact predictions about the disease severity. This paper experiments with pre-trained CNN model named as, Skip connection based ResNet101 for knee OA from the OAI dataset. The obtained result shows that the Skip connection based ResNet101 achieves better accuracy of 79.94%, 84.96% and 95.86% in three datasets named dataset I, II & III respectively. Dataset II is a binary classification of knee OA diagnosis and dataset III is the of knee OA severity classification in three-grades are established to be better than dataset I which is the original dataset that contains a five-grade classification based on the KL grading system. The future work includes hyperparameter tuning in optimization algorithms for enhancing the performance of the model.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

For this research work all authors' have equally contributed in conceptualization, methodology, validation, resources, writing—original draft preparation, writing—review and editing.

References

[1] R. T. Wahyuningrum, A. Yasid, and G. J. Verkerke, "Deep Neural Networks for Automatic Classification of Knee Osteoarthritis Severity Based on X-ray Images", In: *Proc. of 2020 The 8th International Conference on Information Technology: IoT and Smart City*, Xi'an, China, pp. 110-114, December 2020.

[2] U. Yunus, J. Amin, M. Sharif, M. Yasmin, S. Kadry, and S. Krishnamoorthy, "Recognition of knee osteoarthritis (KOA) using YOLOv2 and classification based on convolutional neural network", *Life*, Vol. 12, No. 8, p. 1126, 2022.

[3] S. S. Abdullah and M. P. Rajasekaran, "Automatic detection and classification of knee osteoarthritis using deep learning approach", *La Radiologia Medica*, Vol. 127, No. 4, pp. 398-406, 2022.

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

[5] M. Mezghani, N. Hagemeister, Y. Ouakrim, A. Cagnin, A. Fuentes, and N. Mezghani, "3D kinematics and decision trees to predict the impact of a physical exercise program on knee osteoarthritis patients", *Applied Sciences*, Vol. 11, No. 2, p. 834, 2021.

[6] P. H. Basavaraja and S. Ganesarathinam, "An Ensemble-Of-Deep Learning Model with Optimally Selected Features for Osteoporosis Detection from Bone X-Ray Images", *International Journal of Intelligent Engineering & Systems*, Vol. 15, No. 5, pp. 194-206, 2022, doi: 10.22266/ijies2022.1031.18.

[7] Y. Wang, X. Wang, T. Gao, L. Du, and W. Liu, "An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative", *Journal of Healthcare Engineering*, Vol. 2021, p. 5586529, 2021.

[8] S. Misu, S. Tanaka, K. Ishihara, T. Asai, and T. Nishigami, "Applied assessment method for varus thrust during walking in patients with knee osteoarthritis using acceleration data measured by an inertial measurement unit", *Sensors*, Vol. 22, No. 17, p. 6460, 2022.

[9] R. Yan, "The value of convolutional-neural-network-algorithm-based magnetic resonance imaging in the diagnosis of sports knee osteoarthropathy", *Scientific Programming*, Vol. 2021, p. 2803857, 2021.

[10] L. Liu, J. Chang, P. Zhang, Q. Ma, H. Zhang, T. Sun, and H. Qiao, "A joint multi-modal learning method for early-stage knee osteoarthritis disease classification", *Heliyon*, Vol. 9, No. 4, p. E15461, 2023.

[11] J. B. Schiratti, R. Dubois, P. Herent, D. Cahané, J. Dachary, T. Clozel, G. Wainrib, F. K. Guibert, A. Lalande, M. Pueyo, R. Guillier, C. Gabarroca, and P. Moingeon, "A deep learning method for predicting knee osteoarthritis radiographic progression from MRI", *Arthritis Research & Therapy*, Vol. 23, p. 262, 2021.

[12] R. Mahum, S. U. Rehman, T. Meraj, H. T. Rauf, A. Irtaza, A. M. E. Sherbeeny, and M. A. E. Meligy, "A novel hybrid approach based on deep cnn features to detect knee osteoarthritis", *Sensors*, Vol. 21, No. 18, p. 6189, 2021.

[13] M. J. Awan, M. S. M. Rahim, N. Salim, A. Rehman, H. Nobanee, and H. Shabir, "Improved deep convolutional neural network to classify osteoarthritis from anterior cruciate ligament tear using magnetic resonance imaging", *Journal of Personalized Medicine*, Vol. 11, No. 11, p. 1163, 2021.

[14] A. S. Mohammed, A. A. Hasanaath, G. Latif, and A. Bashar, "Knee Osteoarthritis Detection and Severity Classification Using Residual Neural Networks on Preprocessed X-ray Images", *Diagnostics*, Vol. 13, No. 8, p. 1380, 2023.

[15] G. M. Kumar and A. D. Goswami, "Automatic classification of the severity of knee osteoarthritis using enhanced image sharpening and CNN", *Applied Sciences*, Vol. 13, No. 3, p. 1658, 2023.

[16] D. R. Sarvamangala and R. V. Kulkarni, "Grading of Knee Osteoarthritis Using Convolutional Neural Networks", *Neural Processing Letters*, Vol. 53, No. 4, pp. 2985-3009, 2021.

[17] B. Liu, J. Luo, and H. Huang, "Toward automatic quantification of knee osteoarthritis severity using improved Faster R-CNN", *International Journal of Computer Assisted Radiology and Surgery*, Vol. 15, No. 3, pp. 457-466, 2020.

[18] A. Tiulpin and S. Saarakkala, "Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks", *Diagnostics*, Vol. 10, No. 11, p. 932, 2020.

[19] S. M. Ahmed and R. J. Mstafa, "Identifying Severity Grading of Knee Osteoarthritis from X-ray Images Using an Efficient Mixture of Deep

Learning and Machine Learning Models”, *Diagnostics*, Vol. 12, No. 12, p.2939, 2022.

[20] Dataset link: https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity?resource=download&select=auto_test.

[21] A. Khalid, E. M. Senan, K. A. Wagih, M. M. A. A. Azzam, and Z. M. Alkhraisha, “Hybrid Techniques of X-ray Analysis to Predict Knee Osteoarthritis Grades Based on Fusion Features of CNN and Handcrafted”, *Diagnostics*, Vol. 13, No. 9, p. 1609, 2023.

[22] M. Habib, M. Ramzan, and S. A. Khan, “A Deep Learning and Handcrafted Based Computationally Intelligent Technique for Effective COVID-19 Detection from X-ray/CT-scan Imaging”, *Journal of Grid Computing*, Vol. 20, No. 3, p. 23, 2022.