Medical Image Analysis of Knee Osteoarthritis using Modified Deep CNN

Mohammed Zakir Bellary[1]
Assistant Professor, Dept. of Electronics & Communication, P.A. College of Engineering, Mangalore, zakir_ece@pace.edu.in
Deepthi[2]
Dept. of Electronics and Communication, P.A. College of Engineering, Mangalore, deepthikotian87@gmail.com
Tanvir Habib Sardar[3]
Assistant Professor, Dept. of CSE, GITAM School of Technology, GITAM University, Bangalore, tsardar@gitam.edu
Dr. B. Aziz Musthafa[4]
Professor, Dept of Computer science, Bearys Institute of technology Mangalore, azizmusthafa@gmail.com
Sheik Jamil Ahmed[5]
Assistant Professor, SoCSE & IS, Presidency University, Bangalore, sheik.jamilahmed@presidencyuniversity.in
Rashel Sarkar[6]
Associate Professor, Department of CSE, University of Science and Technology, Meghalaya, sarkarrashel@gmail.com

Abstract

Kellgren and Lawrence (KL) grading method is mainly used by the clinicians for grading the X-ray images. However, grading individual images are prone to errors. This study proposes an approach for automated knee osteoarthritis classification based on deep neural networks. Diagnosis of Osteoarthritis involves splitting the knee X-ray into the healthy knee or unhealthy knee with KL grading. Here we use a 20-layer deep residual networks ie ResNet 20 for automated knee osteoarthritis classification. ResNet 20 has 19 convolutional layers and 1 fully-connected layer.We analysed the dataset with different CNN models. And got the different performance.

Keywords: Knee Osteoarthritis (OA), Kellgren and Lawrence (KL) grade, Magnetic Resonance Imaging (MRI), Deep Learning (DL), Convolutional Neural Network (CNN), Neural Network (NN), Residual Network(ResNet).

I. Introduction

The largest joint in human body is the knee joint. It is composed of bones, ligaments, tendons (Patellar tendon is largest), and meniscus (medial and lateral meniscus act as cushion [1].

Osteoarthritis(OA) is a condition that develops over period when the protecting tissue ie the cartilage between the joints deteriorates. As it deteriorates, bones rub against one another, resulting in discomfort, restricted movement, stiffness and pain. Age [[2],[3]], heredity [2] or secondary conditions like obesity, injury [3], hormonal imbalances, recurrent joint damage, uric acid [3], or diabetes can all contribute to OA.

OA falls under two categories [3],

(a) Primary: As people get older, the water particles in the cartilage start reducing, due to which it weakens making it more prone to degeneration.

(b) Secondary: It may appear at a young age as a result of specific conditions like excess sugar levels, excess body fat, injury in athletes, rheumatoid arthritis, or as a result of prolonged squatting or kneeling.

Typically, doctors examine X-ray images of knees with OA infection or damage, and later grade the knee OA severity using KL grade. KL grade is considered as the best standard for the severity of knee OA grading [4].

Diagnosis of Osteoarthritis involves splitting the knee X-ray into the healthy knee or unhealthy knee [2].

II. LITERATURE REVIEW

Total of 29 papers was reviewed out of which 8 papers are review articles on the recent developments in the evaluation of knee OA severity from X-ray or Magnetic Resonance Imaging (MRI). The best method for identifying and assessing the state of the knee joint after injury is magnetic resonance imaging (MRI), which is at the forefront of medical imaging [3]. Literature review was done on the problem of KOA diagnosis and prediction. Studies and articles on ML are becoming more prevalent in the KOA sector, emphasizing the necessity for (i) improving our knowledge of the disease's start and development and (ii) innovative data-driven technologies that can help with the KOA prediction difficulty and early diagnosis [5]. Using Machine learning (ML) algorithms, the characteristics of JSW, cartilage thickness and sub-regions, and JSN are identified which are the most crucial parameters. These results might be used to develop a knee OA prediction tool [6]. Different CNN techniques are being applied to segmentation and classification models. The majority of studies categorize OA using plain radiography as their architectural input. But because knee joint alterations are multidimensional, a 3D scan will give a more accurate depiction of the joint changes than a 2D image [7]. KL system is still often utilized in both therapeutic settings and academic settings. Like any other radiographic categorization tool, the KL system is best applied in tandem with a comprehensive clinical evaluation, The creation of therapy algorithms based on classification grade should likely be the main goal of future research applications. Through the use of such algorithms and an evidence-based strategy, clinical decision-making may be better guided [8]. Deep learning techniques are not just used on HPC platforms, many mobile applications have already been created with them. Research and development are going on for hardware-friendly algorithms for DL [9]. By design, only enforce regularization using deep and thin structures, keeping the attention on the challenges of optimization. But as we shall investigate in the future, combining with better regularization may improve results [10]. Modern Neural Networks (NN) can be successfully implemented on devices with limited resources. We demonstrate that using general and hardware-specific improvements can drastically reduce the memory footprint of NNs and shorten their inference delay [11].

The other 14 papers are on the various deep learning models using either X-ray or/an MRI. Using MRI images bones and cartilage are segmented automated using 2D and 3D CNN and SSM post processing. Time is reduced for automated methods compared to manual. Database should be huge for studying the OA. Using the implemented method it takes 43 weeks for segmentation [12]. Using Efficiently-Layered Network (ELNet) architecture, the Knee MRI dataset cross-validation has consistently shown superior performance with ELNet models, demonstrating the architecture's robustness in the face of a highly imbalanced distribution. If available, future work may involve improving performance by including all three MRI volumes ie axial, coronal, and sagittal [13]. With a specific knee coil for imaging knee joints was used for magnetic resonance imaging on all patients.The best method for identifying and assessing the state of the knee joint after injury is magnetic resonance imaging (MRI), which is at the forefront of medical imaging [1]. Knee bone segmentation and Knee cartilage segmentation is done from MRI. Diagnose and stop OA development while cartilage degeneration is still reversible.Additional research is required to confirm the therapeutic viability of deep learning models [14]. KL Grading is done on X-ray and Whole Organ MRI Score (WORMS) scoring is done on MRI. The value of single biomarker in finding OA greatly increased. Future OA therapeutic trial designs and personalized medicine approaches will require the predictive biomarkers [15]. KL grading is done on MRI, more lightweight, faster in training and inference, and producing masks for all the tissues under consideration at once. The findings will assist additional work on developing reliable segmentation algorithms for knee MRI and encourage greater adoption of DL based approaches in the OA research community. We simply took 2D segmentation into account [16]. Fully Convolutional Networks (FCNs), 2D FCNs, 3D FCNs models were used on MRI. Even on high KL-grade images, models were determined to be accurate, and it could simultaneously learn across numerous cohorts. It is important to conduct more research to determine whether the findings provided here apply to other clinically significant parameters, such as surface distances, volumes, and so on [17]. On MRI CNN, Hybrid ResNet, CapsNet network models were used. The results are compared with 4 CNN models. ResNet CapsNet convolution network with auto encoder network had a better result than other models [18]. MR net, CNN are applied on MRI. Clinical specialists performance during the interpretation of medical imaging can be enhanced by deep learning models. To determine the model's applicability in the clinical situation and to validate it prospectively, more study is required [19]. X-ray is subjected to KL Grading, Knee Area Segmentation using YOLO model. (1) Preprocessing of raw radiography pictures by humans is now barely noticeable. (2) The segmentation process is almost completely automated. (3) The visual transformer blocks constructed on top of conventional CNN architecture enhance classification performance by utilizing the self-attention process. (4) Once constructed, the suggested method provides a complete answer to the OA diagnosis. Only one vendor's dataset is used for trial. Futher plan is to investigate how this technology might be used with data from various sources in the future [20]. On MRI 3D CNN is applied. MRI is regarded as an alternate imaging technique, particularly for spotting early stage osteoarthritis with a small structural alteration. Used limited number of samples [21]. On MRI ResNet14 and CNN is applied. The hybrid class balancing method, which adds 92% more data to the ResNet 14 CNN model, has the highest accuracy. Increased the overall training cost [22]. DL model is applied on X-ray. Image analysis represents a glaring milestone and is expanding rapidly. Whether a deep learning model can be developed that outperforms experts abilities is impossible to foresee. Orthopedic surgeons who want to use a deep learning algorithm for image analysis must be objective while handling data, present work that adheres to recently recommended standards, and concentrate on creating multitasking models [23]. DNN and KL Grading is done on X-ray images. It turns out that for numerous system testing yield the best accuracy. Model's dependency on training data to be reduced [24].

For our study we used mainly 7 papers out of which 3 papers are on the ResNet modules using KL grading methods on X-ray images and the other 4 papers are on CNN using KL grading methods on X-ray images.

Deep Learning (DL) model is developed to predict 1) probability of total knee replacement 2)KL grade by using baseline X-ray. Analysis to be done to find the actual time to TKR, which would give doctors helpful additional information[5]. Image analysis represents a glaring milestone and is expanding rapidly. Whether a deep learning model can be developed that outperforms experts abilities is impossible to foresee. Orthopedic surgeons who want to use a deep learning algorithm for image analysis must be objective while handling data, present work that adheres to recently recommended standards, and concentrate on creating multitasking models [25]. Discovered that utilizing the KL grading system, we could train a CNN to accurately diagnose and categorize the severity of knee OA without first removing important visual abnormalities like implants from the input data. The absence of various DL architectures investigated is one restriction. There are many different DL designs, and perhaps some of them could have produced superior outcome [26]. Here, a brand new end-to-end architecture was proposed that includes trainable attention modules that serve as unsupervised, fine-grained Region Of Interest (ROI) detectors. The results for the publicly available knee OA datasets OAI and MOST were satisfactory, although there is still a lot of room for improvement [27].

JCRegNet method can effectively improve the accuracy of joint center recognition on knee X - ray images and also with better generalization capability. CT or MRI dataset can be used to help improve the model accuracy during the training to get more depth information. And other image features such as the trabecular thickness, trabecular bone density and trabecular bone texture [4]. The model performs significantly better in simultaneous OARSI and KL grading as well as in the detection of radiographic OA presence. Analysis not done on the attention maps produced by the method. The presented ensemble approach is computationally heavy due to ensembling and hypothetically, could affect the real-life use of the developed method unless the model is deployed on GPU [28]. Demonstration of the model’s ability to learn relevant OA features that are transferable across different datasets Attention maps show that our model reacts to the relevant radiological findings –osteophytes – while the baseline reacts to the joint center. Investigation is required for finding the generalisability of the method across multiple datasets using larger amounts of data. Another limitation is that the image resolution is reduced to 8-bit, which could have led to the loss of fine-grained information stored in the images [29].

III. ARCHITECTURE

A. Convolutional Neural Network (CNN)

CNN was proposed first by Fukushima in the year 1988. However, due to the limitations for training the network of computation hardware, it was not used widely. In the year 1990s, LeCun et al solved the classification problem with CNNs and a gradient-based learning algorithm. Following that, researchers developed improved CNNs.

CNNs have numerous advantages over DNNs, including a processing system that is more similar to that of the human eye, excellent structural optimization for image processing, and the ability to learn and extract 2D feature abstractions.

There are three types of layers (Figure 1) in CNN they are:

1. Input Layers

This layer is the first block which receives the input data. The number of features in the data and the number of neurons in the layer are same.

2. Hidden Layers

This layer gets data from Input Layer. Depending on our model and the data, many hidden levels are present. The hidden layer neurons varies and always exceed the feature numbers. The output of each layer is determined by multiplying its learnable weights by the output from the layer preceding it in a matrix. The network becomes nonlinear by combining the learnable biases from that layer and then computing the activation function.

Convolutional, ReLU, max pooling are among some of the hidden layers, and they all have a significant impact.

3. Output Layers

The data from hidden layer is then used as the input for a logistic function, such as sigmoid or softmax, to turn the output of each class into the likelihood score for each class.

Figure 1: CNN Layers

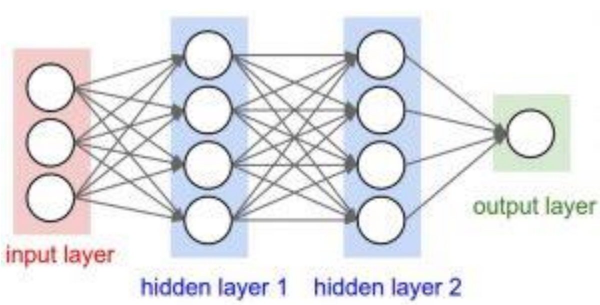
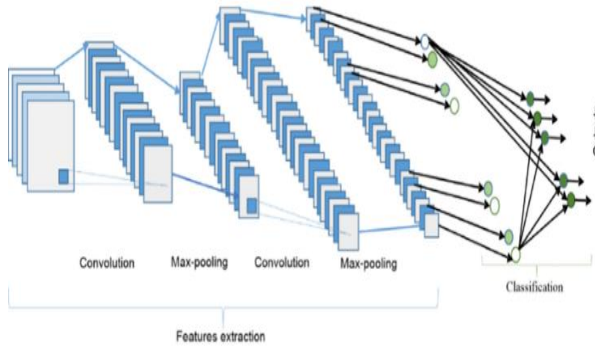


Figure 2 shows CNN architecture [9].



B. Residual Network (ResNet)

In image categorization, Deep convolutional neural network (DCNN) has made number of significant advances. The accuracy of a deepening neural network will first increase before reaching saturation, however as deep learning advances and there is problem with degradation.

Accuracy will decrease, if the depth is increased. As the accuracy decreases in training set as well as testing set, its clear that its not the result of over fitting. With fusion of cross layer feature, ResNet improves feature extraction abilities, and performance of the network gradually rises as the depth is increased.

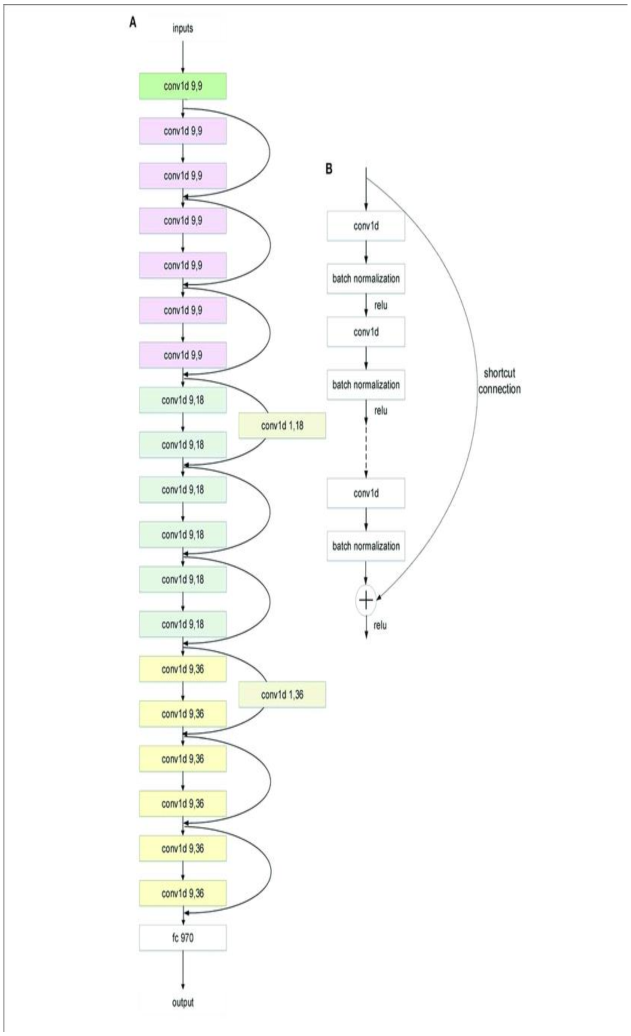
ResNet design is primarily motivated by the need to address the degradation problem with neural networks, which is that as neural networks get more complex, their training error rates increase. To solve this issue residual structure is implemented.

As residual function of input from each layer, the network layer's functionality is reprogrammed. The concept of residual in mathematics refers to the difference between the value of actual observation and estimated value ie fitting value.

Here we use a 20-layer deep residual networks ie ResNet 20. ResNet 20 has 19 convolutional layers and 1 fully-connected layer. Shortcuts are used to link feature maps of the same dimension, and 1x1 convolutional layers are selected as shortcuts between feature maps of different dimensions. After convolution and before activation batch normalization is adopted.

Figure 3 shows the architecture of 20-layer ResNet.

Figure 3: ResNet20 Architecture [30]



IV. PROPOSED METHOD

A. Data Collection

Dataset used in this paper is obtained from the below link https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity

Total of 1464 images were taken under 5 grades ie Grade0, Grade1, Grade 2, Grade 3, Grade 4 and were split into 1171 under training dataset and 298 images under testing dataset.

Here the ratio for Training Data and Testing Data is taken as 80:20.The following are the Grade descriptions:

Grade 0: A normal knee joint X-ray image with no joint space reduction and no osteophytes seen.

Grade 1: A doubtful knee joint X-ray image where joint narrowing is doubtful.

Grade 2: A mild knee joint X-ray image where osteophytes are definitely present and mild reduction in joint space are seen.

Grade 3: A moderate knee joint X-ray where joint space is moderately reduced but bones are not touching each other and osteophytes are seen.

Grade 4: A severe knee joint X-ray where joint space is significantly reduced and significant osteophytes, scleosis seen.

B. Data Processing

1. Convolutional Neural Network (CNN)

The image is fed to the model and the output from each of the layer is got through the process called feedforward. Then we calculate the error using error function, and later backpropagate the model to minimize the loss.

Convolution is used for extracting the features of the image. Convolution layers has filters which does the convolution operation. Here image is considered as matrix of pixels. The dot product of all the filters and image patches are done in convolution layer. Activation function Relu is applied to the output of convolutional layer. Element wise operation is done by Relu which sets negative pixel to 0. The output of this operation generates rectified feature map. This helps in keeping the image dimension unchanged. Down sampling operation is done through pooling layer which reduces the dimension of rectified feature map. Pooled feature map is generated from the pooling layer. Max pooling is mainly used so that the computation is faster and reduces the memory and helps in preventing overfitting. The pooled feature map undergoes a process called flattening, ie the pooled feature map is converted to continuous single long linear vector. The flattened pooled feature map is fed to fully connected layer to get the final output.

2. Residual Networks (ResNet)

To solve the problem associated with varnishing/exploding gradient, Residual Blocks were introduced by ResNet architecture. A residual block or identity block is a fundamental component of a ResNet. Simply put, a residual block occurs when a layer's activation in the neural network is accelerated to a deeper layer.

The vanishing/exploding gradient problem was addressed in this design by the introduction of the Residual Blocks concept. In this network, we use a technique called as skip connections. Skip connection bypasses some levels in between to link layer activation's to subsequent layers, Residual block is formed this way. By stacking these residual blocks together ResNets are made.

ResNet with 20 layers for image classification using Keras is built. The identity block is defined so that it makes the neural network to residual network which represents the skip connection. Then convolution block is built which combines main path and shortcut. Now the blocks are combined building 20 layer residual network.

C. Equations used

1. Accuracy

To evaluate how well the suggested method performed overall on the data, accuracy is computed. The equation for accuracy is as below [18] ,

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

Where,

True Positive(TP) defines the total count of images that were accurately classified as non healthy class.

False Positive(FP) defines the total count of images that were falsely classified as non healthy.

False Negative(FN) defines the total count of images that our model failed to detect as non healthy but were non healthy.

False Positive(FP) defines the total count of images that were falsely predicted as non healthy, but were healthy.

2. Recall/ Sensitivity

The percentage of images that the system recalled that were not healthy is calculated as Recall R.

Formula is given below [18],

$$R = \frac{TP}{TP + FP} X100$$

3. Precision

Precision P calculates and graphically illustrates the proportion of people who were correctly categorised by the suggested system.

The equation for precision is given below ,

$$P = \frac{TP}{TP + FN} X100$$

4. F1 score

F1 score gives the harmonic mean of precision and recall. It mainly tells the overall performance of the model.

$$F1Score = \frac{Precision * Recall}{Precision + Recall} X2$$

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

Using the trained CNN and ResNet model, the classification results of the test dataset are shown in Table 1, Table 2, Table 3, Table 4, Table-5 and Table 6.

Figure 4 shows grade 0 images being misclassified as grade 2, Figure 5 shows grade 1 misclassified as grade 0,Figure 6 shows grade 2 misclassified as grade 4, Figure 7 shows grade 3 misclassified as grade 4 and Figure 8 shows grade 4 misclassified as grade 3.

Table 1: Accuracy, recall rate, and F1 score for CNN

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	60.00	83.33	62.50	71.43
1	70.00	100	57.14	72.73
2	70.00	100	57.14	72.73
3	70.00	100	57.14	72.73
4	60.00	83.33	62.50	71.43

Table 2: Accuracy, recall rate, and F1 score for ResNet

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	70.00	85.71	75.00	80.00
1	80.00	100	71.43	83.33
2	80.00	100	71.43	83.33
3	80.00	100	71.43	83.33
4	70.00	85.71	75.00	80.00

Table 3: Accuracy, recall rate, and F1 score for AlexNet

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	74.00	85.71	76.00	82.00
1	81.00	100	77.43	83.33
2	81.00	100	76.43	85.33
3	80.00	100	76.43	85.33
4	73.00	85.71	75.00	83.00

Table 4: Accuracy, recall rate, and F1 score for SegNet

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	84.00	86.51	77.01	86.00
1	83.00	100	76.43	86.33
2	83.00	100	76.43	86.33
3	82.00	100	76.43	85.33
4	82.00	85.71	75.00	84.01

Table 5: Accuracy, recall rate, and F1 score for googleNet

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	83.00	84.71	75.00	80.00
1	82.00	100	71.43	83.33
2	83.00	100	71.43	83.33
3	82.00	100	71.43	83.33
4	82.56	84.71	75.00	82.00

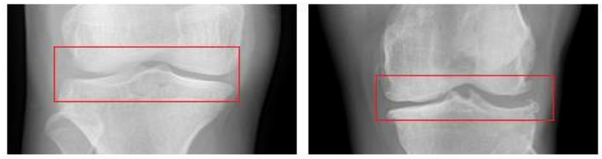
Table 6: Accuracy, recall rate, and F1 score for VGGNet

Grade	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0	70.00	85.71	76.00	82.00
1	80.00	100	74.43	81.34
2	80.00	100	75.49	85.34
3	80.00	100	75.43	84.34
4	80.00	85.71	77.00	83.01

Figure 4: Normal Knee X-ray Images

No joint space reduction. No osteophytes seen.

Figure 5: Doubtful Knee OA X-ray Images



Joint narrowing is doubtful

Figure 6: Mild Knee OA X-ray Images

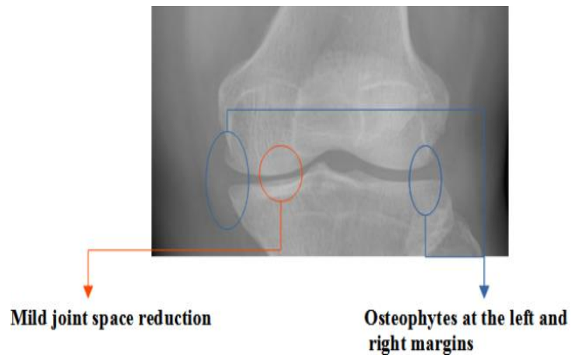


Figure 7: Moderate Knee OA X-ray Images

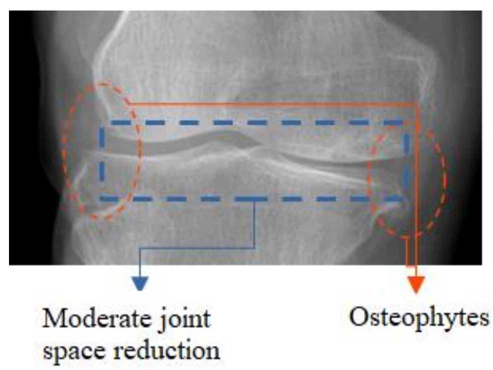
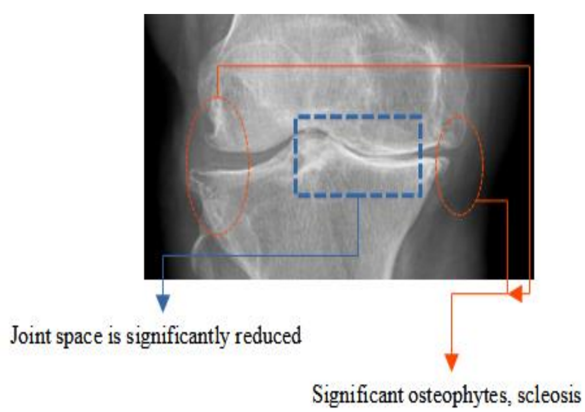


Figure 8: Severe Knee OA X-ray Image



VI. CLINICAL VALIDATION

Clinical validation of 50 X-ray images, ie 10 X-ray images of each grade is done by Dr Dileep K.S, Associate Professor, Department of Orthopaedics, K S Hedge Medical Academy, Mangalore.

VII. CONCLUSION

Deep learning has recently gained popularity as a field of study in academia, particularly in the area of medical imaging. For automated classification of knee osteoarthritis on X-ray image, We used the ResNet 20 architecture to classify the grades of Knee osteoarthritis.

Finally, accuracy of 66% is achieved for CNN and 76% is achieved for ResNet20. Comparing both we found ResNet architecture is better than CNN.

VIII. FUTURE WORK

For our trial, we only used data from one provider. In the future, we intend to look at how our technique might be applied to data from various sources and also need to implement latest deep learning algorithms for Knee OA detection.

References

[1]Saeed, I.O., 2018. MRI evaluation for post-traumatic knee joint injuries. IOSR Journal of Nursing and Health Science (IOSR-JNHS), 7(2), pp.48-51.

[2]Mahum, R., Rehman, S.U., Meraj, T., Rauf, H.T., Irtaza, A., El-Sherbeeny, A.M. and El-Meligy, M.A., 2021. A novel hybrid approach based on deep cnn features to detect knee osteoarthritis. Sensors, 21(18), p.6189.

[3]Saini, D., Chand, T., Chouhan, D.K. and Prakash, M., 2021. A comparative analysis of automatic classification and grading methods for knee osteoarthritis focussing on X-ray images. Biocybernetics and Biomedical Engineering.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4]Wang, Y., Li, S., Zhao, B., Zhang, J., Yang, Y. and Li, B., 2022. A ResNet‐based approach for accurate radiographic diagnosis of knee osteoarthritis. CAAI Transactions on Intelligence Technology.

[5]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.

[6]Jamshidi, A., Leclercq, M., Labbe, A., Pelletier, J.P., Abram, F., Droit, A. and Martel-Pelletier, J., 2020. Identification of the most important features of knee osteoarthritis structural progressors using machine learning methods. Therapeutic advances in musculoskeletal disease, 12, p.1759720X20933468.

[7]Yeoh, P.S.Q., Lai, K.W., Goh, S.L., Hasikin, K., Hum, Y.C., Tee, Y.K. and Dhanalakshmi, S., 2021. Emergence of deep learning in knee osteoarthritis diagnosis. Computational intelligence and neuroscience, 2021.

[8]Kohn, M.D., Sassoon, A.A. and Fernando, N.D., 2016. Classifications in brief: Kellgren-Lawrence classification of osteoarthritis. Clinical Orthopaedics and Related Research®, 474(8), pp.1886-1893.

[9]Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A. and Asari, V.K., 2019. A state-of-the-art

survey on deep learning theory and architectures. Electronics, 8(3), p.292.

[10] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[11] Heim, L., Biri, A., Qu, Z. and Thiele, L., 2021. Measuring what really matters: Optimizing neural networks for tinyml. arXiv preprint arXiv:2104.10645.

[12] Ambellan, F., Tack, A., Ehlke, M. and Zachow, S., 2019. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the Osteoarthritis Initiative. Medical image analysis, 52, pp.109-118.

[13] Tsai, C.H., Kiryati, N., Konen, E., Eshed, I. and Mayer, A., 2020, September. Knee injury detection using MRI with efficiently-layered network (ELNet). In Medical Imaging with Deep Learning (pp. 784-794). PMLR.

[14] 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.

[15] Attur, M., Krasnokutsky, S., Zhou, H., Samuels, J., Chang, G., Bencardino, J., Rosenthal, P., Rybak, L., Huebner, J.L., Kraus, V.B. and Abramson, S.B., 2020. The combination of an inflammatory peripheral blood gene expression and imaging biomarkers enhance prediction of radiographic progression in knee osteoarthritis. Arthritis research & therapy ,22(1), pp.1-17.

[16] Panfilov, E., Tiulpin, A., Klein, S., Nieminen, M.T. and Saarakkala, S., 2019. Improving robustness of deep learning based knee mri segmentation: Mixup and adversarial domain adaptation. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (pp. 0-0).

[17] Perslev, M., Pai, A., Runhaar, J., Igel, C. and Dam, E.B., 2021. Cross ‐ Cohort Automatic Knee MRI Segmentation With Multi ‐ Planar U ‐ Nets. Journal of Magnetic Resonance Imaging.

[18] Akbari, M. and Momeni, M., A Hybrid Deep Learning Methodology for Knee Osteoarthritis Diagnosis Using Magnetic Resonance Images.

[19] Bien, N., Rajpurkar, P., Ball, R.L., Irvin, J., Park, A., Jones, E., Bereket, M., Patel, B.N., Yeom, K.W., Shpanskaya, K. and Halabi, S., 2018. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. PLoS medicine, 15(11), p.e1002699.

[20] 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.

[21] Guida, C., Zhang, M. and Shan, J., 2021. Knee osteoarthritis classification using 3d cnn and mri. Applied Sciences, 11(11), p.5196.

[22] Javed Awan, M., Mohd Rahim, M.S., Salim, N., Mohammed, M.A., Garcia-Zapirain, B. and Abdulkareem, K.H., 2021. Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach.

Diagnostics, 11(1), p.105.

[23] Lee, J. and Chung, S.W., 2022. Deep learning for orthopedic disease based on medical image analysis: Present and future. Applied Sciences, 12(2), p.681.

[24] Tri Wahyuningrum, R., Yasid, A. and Jacob Verkerke, G., 2020, December. Deep Neural Networks for Automatic Classification of Knee Osteoarthritis Severity Based on X-ray Images. In 2020 The 8th International Conference on Information Technology: IoT and Smart City (pp. 110-114).

[25] 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), p.584.

[26] Olsson, S., Akbarian, E., Lind, A., Razavian, A.S. and Gordon, M., 2021. Automating Lawrence in the knee using deep learning in an unfiltered adult population. BMC Musculoskeletal Disorders, 22(1), pp.1-8.classification of osteoarthritis according to Kellgren-

[27] Górriz, M., Antony, J., McGuinness, K., Giró-i-Nieto, X. and O'Connor, N.E., 2019, May. Assessing knee OA severity with CNN attention-based end-to-end architectures. In International conference on medical imaging with deep learning (pp. 197-214). PMLR.

[28] Tiulpin, A. and Saarakkala, S., 2020. Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks. Diagnostics, 10(11), p.932.

[29] 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.

[30] Cui, Q., Lu, S., Ni, B., Zeng, X., Tan, Y., Chen, Y.D. and Zhao, H., 2020. Improved prediction of aqueous solubility of novel compounds by going deeper with deep learning. Frontiers in oncology, 10, p.121.