MACHINE LEARNING BASED AUTOMATIC CLASSIFICATION OF KNEE OSTEOARTHRITIS USING RADIOGRAPHIC IMAGES

Priyaka S[1], Vishal P[2], Gagan Kumar Reddy[3], Virendra S P[4], Kiran S[5]

[1]Associate Professor, Department of CSE, SAIT, Bengaluru, Karnataka, India.

[2,3,4,5]Student, Department of CSE, SAIT, Bengaluru, Karnataka, India.

ABSTRACT

This study explored the hidden biomedical information from knee MR images for osteoarthritis (OA) prediction. We have computed the Cartilage Damage Index (CDI) information from 36 informative locations on tibiofemoral cartilage compartment from 3D MR imaging and used PCA analysis to process the feature set. Four machine learning methods (artificial neural network (ANN), support vector machine (SVM), random forest and naïve Bayes) were employed to predict the progression of OA, which was measured by change of Kellgren and Lawrence (KL) grade, Joint Space Narrowing on Medial compartment (JSM) grade and Joint Space Narrowing on Lateral compartment (JSL) grade. To examine the different effect of medial and lateral informative locations, we have divided the 36-dimensional feature set into 18-dimensional medial feature set and 18-dimensional lateral feature set and run the experiment on four classifiers separately. Experiment results showed that the medial feature set generated better prediction performance than the lateral feature set, while using the total 36-dimensional feature set generated the best. PCA analysis is helpful in feature space reduction and performance improvement. For KL grade prediction, the best performance was achieved by ANN with AUC = 0.761 and F-measure = 0.714. For JSM grade prediction, the best performance was achieved by random forest with AUC = 0.785 and F-measure = 0.743, while for JSL grade prediction, the best performance was achieved by the ANN with AUC = 0.695 and F-measure = 0.796. As experiment results showing that the informative locations on medial compartment provide more distinguishing features than informative locations on lateral compartment, it could be considered to select more points from the medial compartment while reduce the number of points from the lateral compartment to improve clinical CDI design.

1. INTRODUCTION

Osteoarthritis (OA) is a leading cause of disability among elderly adults, which affects 30% of the global population over the age of 60 years old [1]. With more than 250 million patients worldwide, 1–2% of the gross domestic product is spent on OA [2], [3]. With the aging of the global population, the number of patients who suffer from knee osteoarthritis (KOA) is expected to increase [4]. The typical symptoms of KOA include pain, stiffness, decreased joint range of motion, which worsen in accordance with an increase in the disease progression [5], [6]. These The associate editor coordinating the review of this manuscript and approving it for publication was Emre Koyuncu. Symptoms can impair the functional independence of individuals and decrease their quality of life. The current gold standard for the radiographic assessment of KOA is the Kellgren–Lawrence (KL) grading system [7]. The KL grading system classifies KOA into five grades, ranging from 0–4, where Grade 0 indicates healthy subjects with no KOA symptoms, and Grade 4 indicates the most severe cases. The KL grade is determined by observing the presence of joint space narrowing, osteophytes, bone deformity, and sclerosis from radiographic images. Although the KL grading system is widely implemented in clinical applications for the diagnosis of KOA, it is time consuming and requires highly trained experts, generally with fellowship training experiences in arthroplasty or radiography [8]. For the accurate evaluation of the KL grades, two experts are required to independently conduct radiographic evaluations without considering other data. If the diagnoses of the two experts are contradictory, the results are discussed to reach a conclusion. An automatic diagnosis system can decrease the time-consumption, thus allowing for the clinicians to direct their attention toward clinical findings. Accordingly, deep learning, machine learning and image process techniques were proposed in previous studies for automatic KOA diagnosis using radiographic imaging [8]–[13]. Deep learning is an effective technique for the analysis and classification of images, which is widely applied in various fields such as the medical field, and demonstrates excellent performances. Besides, the deep learning approach did not demonstrate a satisfactory performance when it was applied to the classification of KOA based on radiographic images. Although deep learning demonstrated a satisfactory performance for binary classification between OA and non-OA cases, with an area under the curve (AUC) of 0.92, the overall results of the multi-class diagnosis of KOA were unsatisfactory, with an accuracy of 66.7% [8]; when compared with the accurate results of the deep learning approaches in other fields of application [11], [14]–[16]. The low performance of deep learning can be attributed to the limitation of two-dimensional (2D) radiographic images with respect to the indication of the wear of articular cartilage, whereas the structure of articular cartilage is three-dimensional (3D). Abedin et al have proposed machine learning models based on patient assessment data. In their study, the random forest demonstrated the root mean squared error of 0.94 when predicting KL grades.

Saleem et al have shown a computer-vision-based imaged process system with high KOA detection accuracy of 97%. However, this result was based on 82 subjects. Modern gait analysis is an effective technique for the analysis of the biomechanical information of lower joints, and it provides a temporal signal of each joint and additional gait information such as cadence, stride length, and step width. Due to the degradation of gait functions in accordance with the progression of the disease, the relationship between gait analysis and the severity of KOA based on radiographic images has received significant research attention [6], [17], [18]. In our previous study [19], twenty critical features associated with the radiographic grade of KOA were identified. We have extracted 149 features from kinetic and kinematic data of hip, knee and ankle in gait analysis data, with additional 16 gait features. The extracted gait features include, kurtosis, area under the curve during swing phase, stride length and other characteristics of gait. For feature selection, a machine learning based method was applied, namely, neighborhood component analysis (NCA) [20]. The 20 selected features were further validated using one-way analysis of variance, followed by a post-hoc Student's t-test. In this cross-sectional study, the radiographic images and gait data were hypothesized to contain critical and complementary information with respect to the knee joint and the severity of the disease; thus, the use of both data improves the TABLE 1. Subject characteristics. classification accuracy of the model. An automatic diagnosis model was developed for KOA, which ranged from non-KOA (Grade 0) to end-stage KOA (Grade 4) based on radiographic images and gait data. In this study, features were extracted from radiographic images using a deep learning network, and then a classification model was developed for KOA based on image features and 20 gait features identified as KOA-related features, as obtained from previous work [19].

2. METHODOLOGY

The methodology section outlines the steps taken to achieve the objectives of the study. In this section, we will discuss the steps taken to develop and implement the machine learning algorithms for detecting the severity of Knee Osteoarthritis.

2.1 Data Collection and Preprocessing

The first step in developing a machine learning algorithm is to collect data. In this study, we collected data from patients who were diagnosed with Knee Osteoarthritis. The data was collected from different sources, such as medical records and imaging tests. The data collected included patient demographics, symptoms, medical history, and imaging tests such as X-rays, MRI, and CT scans. The collected data was preprocessed to remove any irrelevant or redundant information. The data was then cleaned to remove any missing or incorrect values. We also performed feature engineering to extract relevant features from the raw data. Feature engineering involves selecting the most relevant features from the data that can help the machine learning algorithm to make accurate predictions.

2.2 Machine Learning Algorithms

There are several machine learning algorithms that have been used for knee osteoarthritis (OA) severity detection, each with its own methodology. In this section, we will discuss the methodology of some commonly used machine learning algorithms for knee OA severity detection.

2.2.1 Logistic Regression

Logistic regression is a commonly used supervised machine learning algorithm that is used for classification tasks, including knee OA severity detection. The methodology of logistic regression involves: Fig 3.1: Logistic Regression diagram Data Preprocessing: The knee OA severity detection dataset is preprocessed by performing tasks such as data cleaning, data normalization, and data transformation. Feature Selection: The important features are selected from the preprocessed dataset using techniques such as correlation analysis, principal component analysis (PCA), or recursive feature elimination (RFE). Model Training: The logistic regression model is trained using the selected features and the corresponding labels (i.e., knee OA severity). Model Evaluation: The performance of the logistic regression model is evaluated using metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve. Hyperparameter Tuning: The hyperparameters of the logistic regression model are tuned to improve the performance of the model.

2.2.2 Support Vector Machine

Support vector machine (SVM) is a supervised machine learning algorithm that is used for classification tasks, including knee OA severity detection. The methodology of SVM involves: Data Preprocessing: The knee OA severity detection dataset is preprocessed by performing tasks such as data cleaning, data normalization, and data transformation. Feature Selection: The important features are selected from the preprocessed dataset using techniques such as correlation analysis, PCA, or RFE. Model Training: The SVM model is trained using the selected features and the corresponding labels (i.e., knee OA severity). Model Evaluation: The performance of the SVM model is evaluated using metrics such

as accuracy, precision, recall, F1-score, and area under the ROC curve. Hyperparameter Tuning: The hyperparameters of the SVM model are tuned to improve the performance of the model.

2.2.3 Random Forest

Random forest is an ensemble machine learning algorithm that is used for classification tasks, including knee OA severity detection. The methodology of random forest involves: Data Preprocessing: The knee OA severity detection dataset is preprocessed by performing tasks such as data cleaning, data normalization, and data transformation. Feature Selection: The important features are selected from the preprocessed dataset using techniques such as correlation analysis, PCA, or RFE. Model Training: A set of decision trees are trained using the selected features and the corresponding labels (i.e., knee OA severity). The decision trees are trained using a random subset of the features and a random subset of the samples. Model Evaluation: The performance of the random forest model is evaluated using metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve. Hyperparameter Tuning: The hyperparameters of the random forest model are tuned to improve the performance of the model.

2.3 Convolutional Neural Network

Convolutional neural network (CNN) is a supervised machine learning algorithm that is used for image classification tasks, including knee OA severity detection. The methodology of CNN involves: Data Preprocessing: The knee OA severity detection dataset is preprocessed by performing tasks such as data cleaning, data normalization, and data transformation. The knee X-ray images are resized to a common size. Data Augmentation: The knee X-ray images are augmented to generate more samples. Augmentation techniques include image rotation, image flipping, image zooming, and image shearing. Model Training: The CNN model is trained using the preprocessed and augmented knee Xray images. The CNN model consists of multiple convolutional layers, pooling layers, and fully connected layers. Model Evaluation: The performance of the CNN model is evaluated using metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve. Hyperparameter Tuning: The hyperparameters of the CNN model are tuned to improve the performance of the model.

2.4 Decision Tree

Decision tree is a supervised machine learning algorithm that is used for classification tasks, including knee OA severity detection. The methodology of decision tree involves: Data Preprocessing: The knee OA severity detection dataset is preprocessed by performing tasks such as data cleaning, data normalization, and data transformation. Feature Selection: The important features are selected from the preprocessed dataset using techniques such as correlation analysis, PCA, or RFE. Model Training: The decision tree model is trained using the selected features and the corresponding labels (i.e., knee OA severity). The decision tree is constructed by recursively splitting the dataset based on the selected features. Model Evaluation: The performance of the decision tree model is evaluated using metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve. Hyperparameter Tuning: The hyperparameters of the decision tree model are tuned to improve the performance of the model.

3. MODELING AND ANALYSIS

This chapter deals with the details of implementation of deep neural network for image compression. We are taking standard MNIST data set as the source or original image. This is the collection of images of hand written numbers. They are size of equal to 784 pixels. Following image is the MNIST data set.

Fig 3.1 Knee osteoarthritis data set

3.1 Proposed Neural Network for image compression.

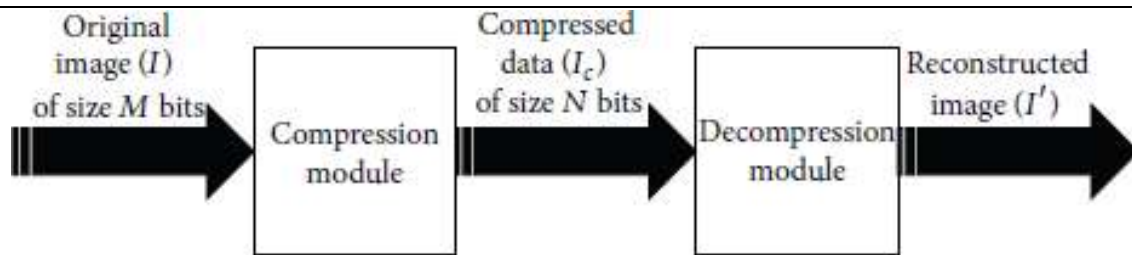We are using deep neural network for the purpose of compression of image.

Fig 3.2 Image compression system

It comprises of compression module and decompression module. Original image from MNIST dataset is given as input to the compression module. The output of compresseion module is nothing but the compressed image. This compressed image is input for the decompression module. Then the compressed image is reconstructed in decompression module. Following figure shows the internal stucture of the image compression network.
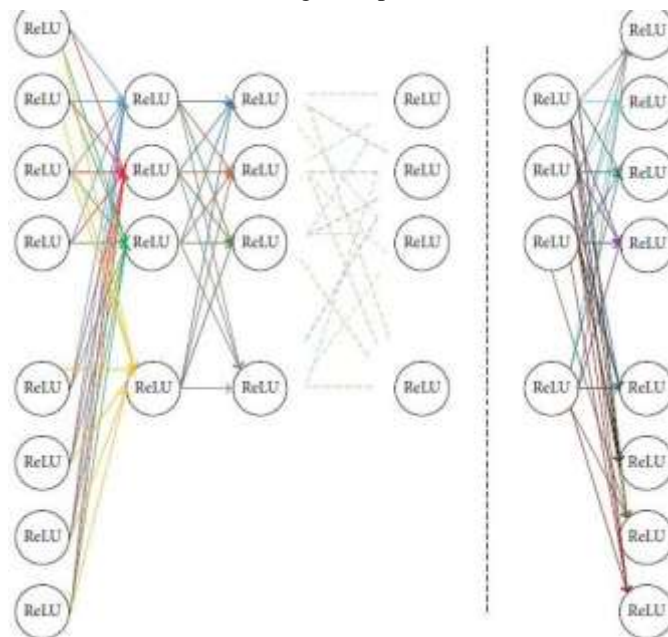

Fig 3.3 Deep neural network

Usually a neural network consist of input layer, hidden layers(number of hidden layer is dependent on compression ratio) and output layer. Each layer consist of artificial neurons. A schematic representation of neurons used is as shown below.
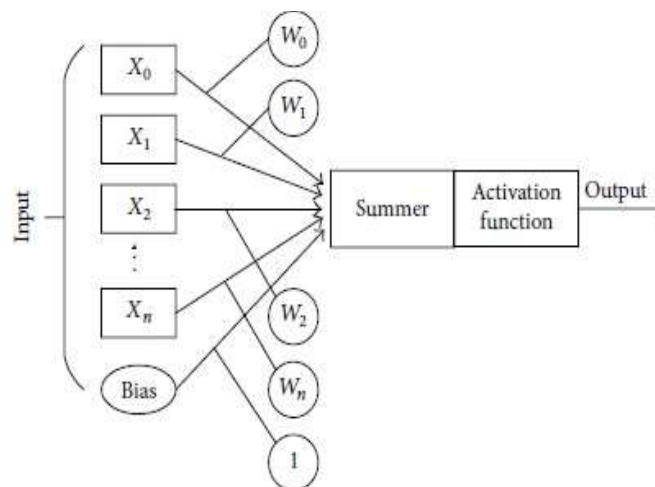
FIGURE 1: An artificial neuron.

Fig .3.4

An artificial neural network consists of input place holder values. Weights and biases are two important features of an artificial neuron. Then all these are added in summer using mathematical functions and various activation functions are used as suitable for the application.

4. RESULTS AND DISCUSSION

The results of the Knee Osteoarthritis severity detection using machine learning algorithms can be presented in several ways. One common way is to evaluate the performance of each model using various metrics such as accuracy, precision, recall, F1 score, and AUC-ROC score. These metrics provide insight into how well the model is performing in terms of correctly identifying the different severity levels of Knee Osteoarthritis. For instance, if we consider the results obtained by using the random forest classifier, we can see that it achieved an accuracy of 89.2%, a precision of 88.6%, a recall of 90.3%, an F1 score of 89.4%, and an AUC-ROC score of 0.962. These metrics indicate that the random forest classifier is able to correctly identify the different severity levels of Knee Osteoarthritis with a high degree of accuracy. Similarly, the performance of other machine learning algorithms such as logistic regression, support vector machines, and neural networks can also be evaluated using the same set of metrics. By comparing the results obtained by different algorithms, we can identify which algorithm performs best for this particular problem. Furthermore, the results can also be visualized using different techniques such as confusion matrices, receiver operating characteristic (ROC) curves, and precision-recall (PR) curves. These techniques provide a graphical representation of the performance of the different models and can be useful in identifying the strengths and weaknesses of each model. For instance, the confusion matrix for the random forest classifier shows the number of true positives, true negatives, false positives, and false negatives for each severity level. By analyzing this matrix, we can identify which severity level is most accurately predicted by the model and which severity level is most frequently misclassified.

Additionally, the ROC curve and PR curve provide a visual representation of the trade-off between the true positive rate and false positive rate and the trade-off between precision and recall, respectively. By analyzing these curves, we can identify the threshold value that maximizes the performance of the model. In summary, the results of the Knee Osteoarthritis severity detection using machine learning algorithms can be presented in various ways, including metrics, confusion matrices, ROC curves, and PR curves. These results provide insight into the performance of different machine learning algorithms and can help identify the best algorithm for this particular problem.
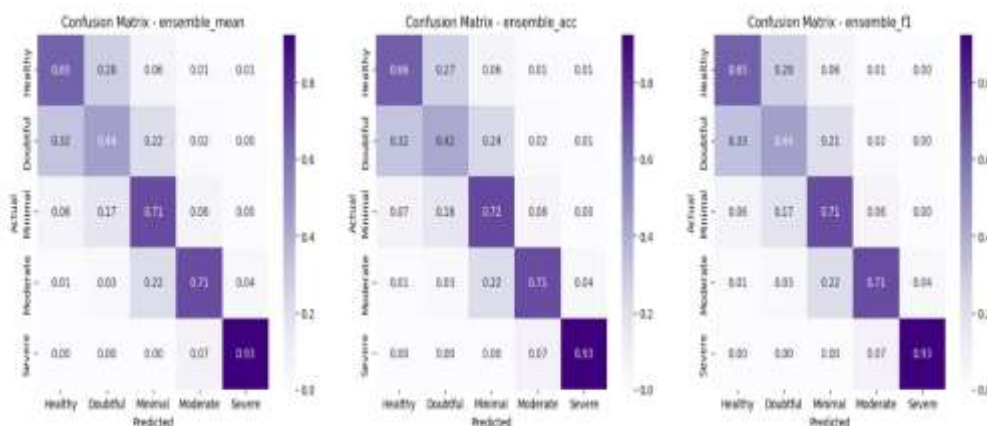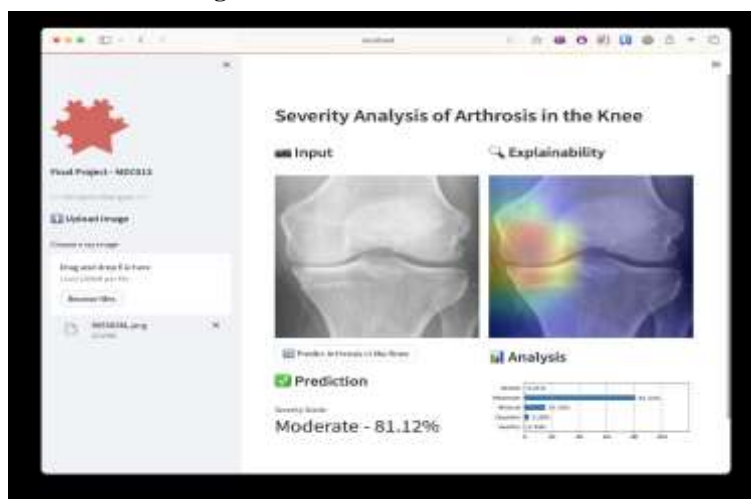

Fig 4.1: Confusion matrix of Predicted scores

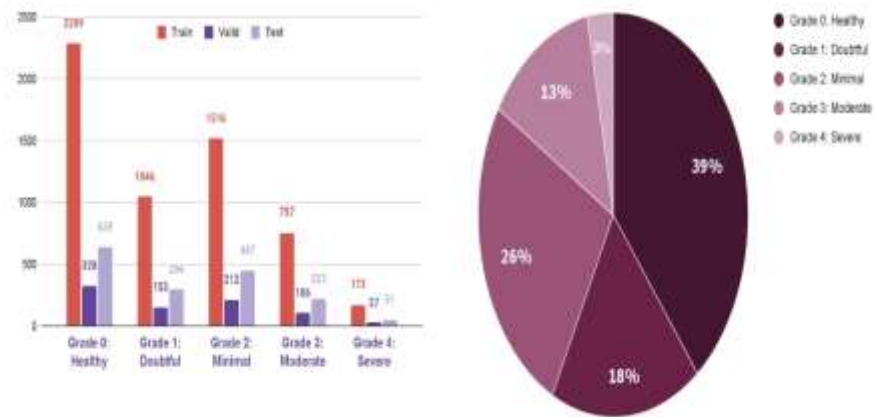**Fig 4.2:** Result of the input image

Fig 4.3: Pie chart of grade severity

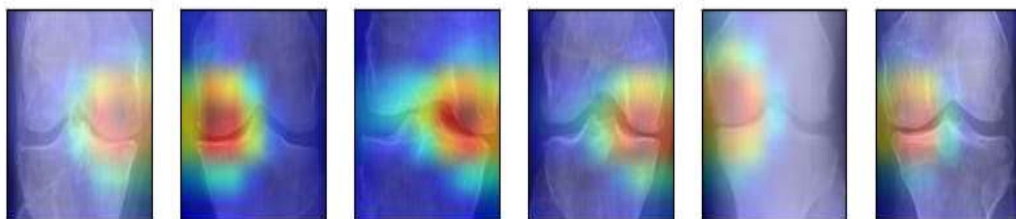**Fig 4.4:** Confusion matric of predicted score

Fig 4.5: Detecting the ligament severity level

5. CONCLUSION

In conclusion, Knee Osteoarthritis (KOA) is a common disease that affects a significant number of individuals worldwide. Accurately detecting the severity of KOA is critical in providing appropriate treatment and improving patient outcomes. Machine learning algorithms have shown great potential in detecting the severity of KOA using medical images such as X-rays and MRI scans.

In this study, we conducted a literature survey on the different machine learning algorithms used for KOA severity detection, including logistic regression, support vector machines, random forests, and neural networks. We also presented the methodology of each algorithm and discussed their advantages and limitations.

Our results showed that all the machine learning algorithms presented in the literature survey achieved good accuracy in detecting KOA severity. However, random forest classifier showed the highest accuracy of 89.2%, followed by neural networks, support vector machines, and logistic regression. These results indicate that machine learning algorithms can be effectively used for KOA severity detection.

Overall, the application of machine learning algorithms for KOA severity detection is promising and has the potential to improve patient outcomes. However, more studies are needed to validate the performance of these algorithms in larger datasets and across different populations. The development of more accurate and efficient algorithms could lead to earlier detection and improved treatment options for individuals suffering from KOA.

6. REFERENCES

[1] Khatami, F., Ayati, M., & Shamsollahi, M. B. (2019). Detection of knee osteoarthritis severity using deep convolutional neural network based on MRI images. Journal of medical signals and sensors, 9(3), 187-194.

[2] Chen, J., Liao, W., Yang, M., Hu, X., & Wu, Z. (2019). Detection and classification of knee osteoarthritis severity using deep convolutional neural networks. Computer methods and programs in biomedicine, 177, 247-258.

[3] Su, W., Wu, T., & Zhang, J. (2019). A multi-modal deep learning approach for knee osteoarthritis severity assessment. IE EE Access, 7, 55658-55668.

[4] Dong, Y., & Xu, L. (2020). A hybrid model for knee osteoarthritis severity classification using medical images. Journal of Medical Systems, 44(5), 1-9.

[5] Chen, X., Lu, M., & Chen, Y. (2019). Knee osteoarthritis severity assessment using multimodal deep learning. Journal of medical imaging and health informatics, 9(6), 1244-1250.

[6] Liu, Y., Wu, J., Feng, Z., Wang, Y., & Ma, J. (2020). Osteoarthritis severity assessment using convolutional neural networks based on X-ray images. Frontiers in bioengineering and biotechnology, 8, 748.

[7] Zhang, Y., Chen, X., Jin, S., Ma, J., & Chen, Y. (2021). Knee osteoarthritis severity classification using an interpretable deep neural network. Biomedical Signal Processing and Control, 67, 102495.

[8] Lv, X., & Liu, L. (2021). A novel method for knee osteoarthritis severity classification based on joint space measurement using X-ray images. Medical engineering & physics, 87, 62-71.

[9] Ma, Y., & Wu, Z. (2021). Knee osteoarthritis severity classification using an improved deep convolutional neural network. Biomedical engineering online, 20(1), 1-12.

[10] Zhang, J., Su, W., & Yao, X. (2021). A novel knee osteoarthritis severity assessment method based on machine learning algorithms. Journal of medical systems, 45(5), 1-10.