International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 9 Issue 2, Mar-Apr 2025 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Knee Cartilage Degeneration Detection using Machine Learning Algorithm

Dhanashree D. Kul, Dr. Prabhat Pallav, Prof. M. U. Inamdar

E&TC, SPPU, Maharashtra, India

ABSTRACT

Connective tissue that helps the joints and bones is named as Cartilage which is present between the bones. Degradation of the tissues between the bones is known as osteoarthritis (OA). It affects the lots of population worldwide which leads to Ache, Rigidity, and Immobility. Early osteoarthritis detection and categorization are challenging for precise analysis and appropriate treatment. Treatments which are followed earlier are totally depends upon the clinical investigations like radiographic x-ray images which is very time-consuming. Subsequently, these processed images are categorized as normal & osteoarthritis. Our goal is to detect knee osteoarthritis. Latest research on machine learning Artificial intelligence(AI) techniques provide much better solutions for the medical image findings and give better solutions for the analysis of osteoarthritis (OA) using different types of imaging technologies. In order to detect OA, this study focuses on Grey-Level Co-Occurrence Matrix (GLCM) feature extraction in conjunction with ML classifiers, specifically Naïve Bayes, Random Forest, and Decision Tree. The Random Forest classifier reaches an accuracy of 89%, according to experimental results. This will help to improve the quality of life.

KEYWORDS: *Osteoarthritis detection, X-ray images, segmentation, machine learning classifier*

How to cite this paper: Dhanashree D. Kul | Dr. Prabhat Pallav | Prof. M. U. Inamdar "Knee Cartilage Degeneration Detection using Machine Learning Algorithm" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-9 | Issue-2, April 2025, pp.917-925, URL: www.ijtsrd.com/papers/ijtsrd78457.pdf

IJTSRD78457

1. INTRODUCTION

Osteoarthritis is a disease of bone joins which is typically identified by decay of connective tissue present between two bones. It is characterized by discomfort while moving or any movement of particular bone areas, hardness which is typically seen in older persons. Older methods of detecting osteoarthritis include radiography-based medical imaging investigations. Osteoarthritis is typically detected by the thickness and hardening of the bone, bone spurs, and the breakdown of materials between joints. Bone spurs, thickening and hardening of the bone, and change in the shape, size, or alignment of a bone [1].

However, this manual detection of the images is cost effective and prone to variability between radiologists. Cartilage protects the knee's bones from rubbing against each other. Healthy cartilage helps bones move smoothly in the joint without friction. It minimizes friction between the bones.

It affects millions of patients globally. Advancements in Artificial Intelligence(AI), specifically present promising solutions to these challenges by automating the analysis of medical images, AI systems can offer faster, more accurate and consistent diagnostic support for OA detection. Detection of knee osteoarthritis more efficiently and accurately is possible with help of different transfer learning, deep learning techniques with the help of X-ray images [1].

This paper strives to offer a thorough understanding of current AI-based techniques, focusing on machine learning model utilized in OA detection and classification.

2. LITERATURE REVIEW

Sl. No.	Author & Year	Methodology	Advantages	Gap
1	Tiwari A[2], 2022	CNN with deep learning	A deep learning model utilizing CNN on X-rays, resulting in high accuracy.	Deep learning models need a substantial quantity of data for optimal accuracy.
2	Shivanand S. Gornale [3], 2016	Active contour segmentation technique	Uses Random Forest for classification.	With current algorithms technique used produces more encouraging and Competitive results.
3	R. S. Hegadi (2019)[4].	ANN	High accuracy in OA detection via X-rays.	ANN typically needs large amounts of data to achieve accuracy.
4	Khalid H (2020) [5]	Deep learning Machine learning	AI advancements in OA detection and grading.	Deep learning Models require high accuracy
5	R Mahum, et,al (2021) [6]	Deep CNN	Captures complex features in medical images	Deep learning Models require high accuracy
6	Phat Nguyen Hue[2021][7]	YoLOv3 &VGG-16 Model.	Model are deep models used high accuracy in object detection.	Deep learning Models require high accuracy
7	Usman Yunus (2023)[8]	CNN Model	Speeds up KOA diagnosis, reducing manual work.	Deep learning Models require high accuracy
8	N. Hema Rajini(2023) [9]	PSO-DNN	High accuracy in OA X-ray classification.	Deep learning Models require high accuracy
9	Tayyaba Tariq (2023) [10]	Method I: Classification in Binary Technique II: Multiclass Classification Using an Ordinal Method	For the Binary Classification, six tests are carried out with several machine learning classifiers.	Inaccurate labels Can Affect model performance
10	Suliman Aladhadh (2023) [11]	DenseNet 201	Identifying& localizing specific features or objects in images,	The Performance of the model affects data.

Table 1. Summary of the Different Papers From Previous Studies and Analysis

Tiwari A et al. [2], studied the characteristics of OA visible on X-rays and used deep learning especially Convolutional Neural Networks (CNNs) to diagnose and grade OA severity deep learning. Author Worked on Data Acquisition, Preprocessing, Model development, Training and Testing.

Shivanand S. Gornale et al.[3] proposed, The deterioration of cartilage caused by osteoarthritis results in the bones rubbing against one another, creating severe pain. Conventional approaches to treating such patients are expensive and time-consuming. The study concludes that machine vision-based techniques for osteoarthritis identification

Khalid H et al.[5]explained AI in Osteoarthritis Review of the Current Landscape. In this compressive studies author explained how artificial intelligence (AI) is changing from detection, to grading, and flow of osteoarthritis (OA) detection, through the analysis of medical images. Major challenges include the inconsistency and lack of good-quality OA datasets, issues with model applicability, and basic considerations regarding AI implementation in healthcare.

R Mahum et al[6] Artificial Neural Networks play an important role for OA detection and classification by automating feature extraction and providing efficient and fast assessments. Artificial Neural Network models

have provided good accuracy; with the challenges like data scarcity, model understandability, and extension. Tackling issues through high-quality data collection, multi-faceted methods, and research studies are important in introducing ANN-based OA detection closer to clinical practice.

Phat Nguyen Hue et al[7] work demonstrates the potential of leveraging deep learning models for KOA detection, offering valuable insights into how these models can be applied to clinical environments to automate diagnosis and enhance the quality of care.

Usman et al. [8] discuss established grading systems, such as the Kellgren-Lawrence (KL) scale, which grades OA severity from 0 to 4 based on features extracted from radiological images. Author explain about CNNs are effectiveness in recognizing patterns and features in radiological X-ray images. The metrics are summarized and used to assess model performance. Lastly authors suggested the need for other models which can also handle rare OA images.

N. Hema Rajini et al.[9] demonstrates how well deep learning methods work for medical picture analysis, especially when it comes to OA detection. The PSO-DNN model optimizes the neural network's parameters to improve classification performance. This method shows a great deal of promise for raising healthcare diagnostic precision. The study emphasizes how crucial high-accuracy models are to medical image processing.

Suliman et al.[11] When compared to conventional techniques, the suggested hybrid model, which combines CNN features with machine learning classifiers, performs better. It approaches meaningful features for medical image analysis.

The following points are the main contribution of this paper :

1. A novel model that illustrates their potential application in this field is presented in this paper.
2. The suggested model outperforms the current approaches in terms of accuracy.

3. METHODOLOGY –

- Proposed approach is shown in fig.2, The dataset is split into three sets. train, evaluate, The and verify. This dataset was selected because, as demonstrated by earlier research, it is among the most widely used datasets for predicting knee osteoarthritis.
- The available dataset contains total 4000 images for class normal & OA

Class	Total Number of Images
Normal	3500
osteoarthritis	500

Figure 1—An example of a normal knee & OA knee

In this section pre-processing, image segmentation, image enhancement, feature extraction, and classification of computed features are the five primary processes of the suggested methodology shown in figure 2.

The X-ray image(fig.1) from dataset is given to preprocessing & this preprocessed image with feature extraction is given to different types of classifier for accuracy of image.

In this section pre-processing, image segmentation, image enhancement, feature extraction, and classification of computed features are the five primary processes of the suggested methodology shown in figure 2.

The X-ray image(fig.1) from dataset is given to preprocessing & this preprocessed image with feature extraction is given to different types of classifier for accuracy of image.

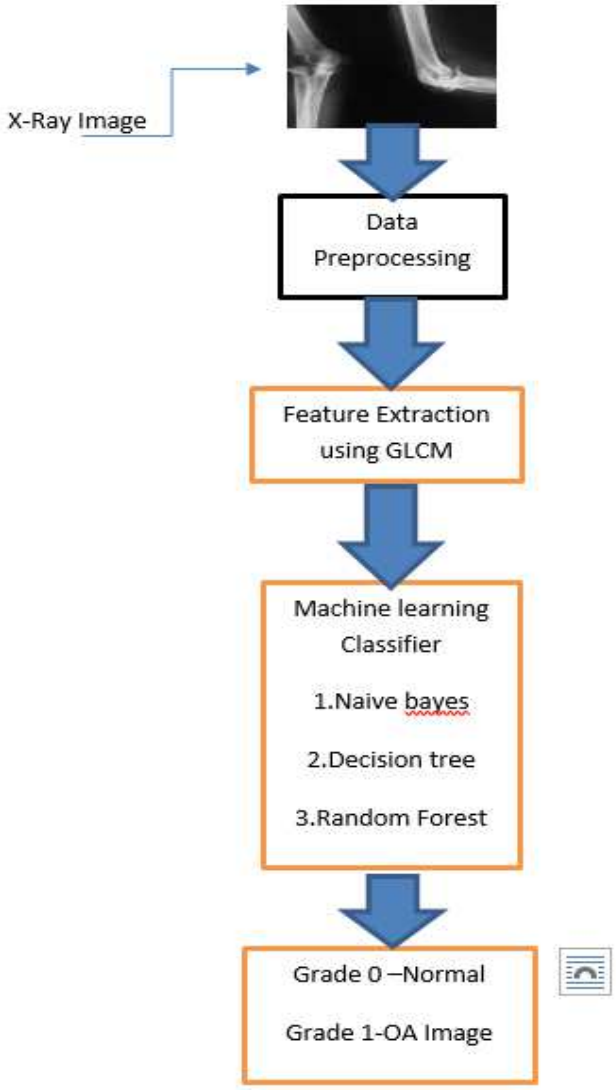

Fig.2 Block diagram of proposed method.

3.1. Pre-processing

Pre-processing is the initial step in this model. In this stage, the quality of image is improved. It can simplify the image and lessen noise. It is essential for increasing image efficiency and accuracy. Function in time analysis for medicine. Because the size of the data set varies, the image is resized.

In preprocessing following are steps-

1. Dataset is loaded. Directory structure is given to load dataset.
2. Image reading is done using cv format
3. Each image is converted into Gray scale using .cv format.
4. The image is cropped & image is resized (224X224).
5. For each image labialization is given

3.2. Feature Extraction (GLCM)

In this methodology GLCM feature extraction with Machine learning classifier is used. Gray-Level Co-Occurrence Matrix (GLCM). It uses concurrence matrix. GLCM is method used to extract texture feature from an image. GLCM features are extracted using RGB channels, each of which has 24 GLCM features. As a result, each image has 24 GLCM features retrieved, for a total of 48 characteristics that are considered for classification The calculation of GLCM attributes for different distances and orientations enables the characterization of distinctive texture features within an image.[12]

Features of GLCM-

1. Correlation –The linear dependency of pixel value of an image.

$$\text{Correlation} = \frac{\sum_{i=1}^{n_g} \sum_{j=1}^{n_g} p(i,j)ij - \mu_x\mu_y}{\sigma_x(i)\sigma_y(j)} \qquad (1)$$

2. Contrast-The intensity difference between pixel & neighboring pixel.

$$\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 P_{(i,j)} \quad (2)$$

3. Homogeneity-The closeness of elements in the signal

3.3. Classification

The features extracted from GLCM are given to machine learning algorithms. These algorithms are given to classify normal and OA image. For this detection three machine learning classifiers include Naïve Bayes, Random forest & Decision tree used.

Naïve Bayes

In machine learning, the naïve bayes classifier is effective method for classification problems. The Bayes theorem is source of naïve algorithm. Each feature pair is ranked independently of each other. It describes probity of the events .The two assumptions of the characteristics of Naïve Bayes are independent and every attribute is handled uniformly.

For continuous data, Gaussian Naïve Bayes is employed; for binary data, Bernoulli Naïve Bayes; and for count data, Polynomial Naïve Bayes is employed.[13]

Random Forest

An algorithm for supervised learning is called a random forest. A forest is a collection of decision trees & supervised algorithm. The basic logic of the random forest algorithm is the decision tree. Decision trees essentially learn a hierarchy of if– then–else type problems that ultimately lead to classification. A decision tree asks the top node a series of questions with the highest priority. Then, several decision trees arerun from the forest, and a majority vote is to give the final classification decision [14]

Decision Tree

A decision tree is simple & effective algorithm which is used for classification. Based on the feature values, this supervised learning model divides the data into subgroups. To create a model that predicts the value of a target variable, the goal is to learn fundamental decision rules based on the attributes. The decision tree maximizes the purity of the resulting subsets by dividing the dataset according to specific criteria, typically a feature threshold. [15]

4. Experimental Results

4.1. Implementation Details

In experimental Analysis, a computer system with an i3 processor running Windows 10 was used. As part of the data pre-processing procedure, RGB photos were converted to grayscale while maintaining the contrast of the original image. The resulting grayscale photos from this conversion have less contrast, which made the following procedures more difficult. All RGB photos were converted to grayscale using the rgb2gray function because each one had different sizes & color information. We use MATLAB and Google Collaborator for feature extraction and pre-processing. Python modules are utilized in Google Colab, and the Weka tool implements machine learning categorization.

4.2. Performance Measures

This method's performance can be evaluated using a wide range of measures, such as F1 score, recall, accuracy, and precision. Medical applications make advantage of these measurements.nts can be summed

Accuracy (A): - The percentage of accurate predictions a model makes out of all the forecasts is known as accuracy (A) [16]

1. Accuracy=Number of correct Predictions/Total Number of prediction

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Where TP = True positives & FP is False positive

TN= True Negatives & FN is False Negative

2. Precision(P) :- It is computed as the number of accurate predictions a model makes from all of the data's true positives and false positives.[17]

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

3. *Recall (R):* - It is computed as the percentage of all positive cases in the provided data that have the correct distribution of positives. It assesses the model's ability to detect positive instances. [18]

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

F1 Score (F1): - It is computed as the mean of recall and precision.

$$\text{F1 Score} = 2 \times \frac{P \times R}{P + R} \quad (7)$$

Particularly when different kinds of errors (false positives or false negatives) have different outcomes, precision, recall, F1-Score, specificity, and AUC provide a more comprehensive picture of model performance [19]

5. Results and Discussions

Combining several feature extraction techniques with classification techniques yields results**.**

Dataset-I Results

GLCM Feature Extraction for Dataset-I [20]

In dataset there are different Knee Osteoarthritis X-ray images The random forest classifier gives high accuracy of 89%. The following subsections describe these experimental results.

GLCM_Machine_Learning_Classifiers_Test –

The Gray-Level Co-occurrence Matrix (GLCM) has been a key technique in recent texture classification advances for identifying spatial information that characterize an image's texture.

The below chart shows comparison of test classifier. The findings show that classifier selection has a major influence on classification accuracy, with SVM performing best because of its capacity to manage high-dimensional GLCM information. This study emphasizes how crucial feature selection and classifier tuning are to improving texture recognition systems' Random forest classifier have high accuracy, precision, recall & f1-score.

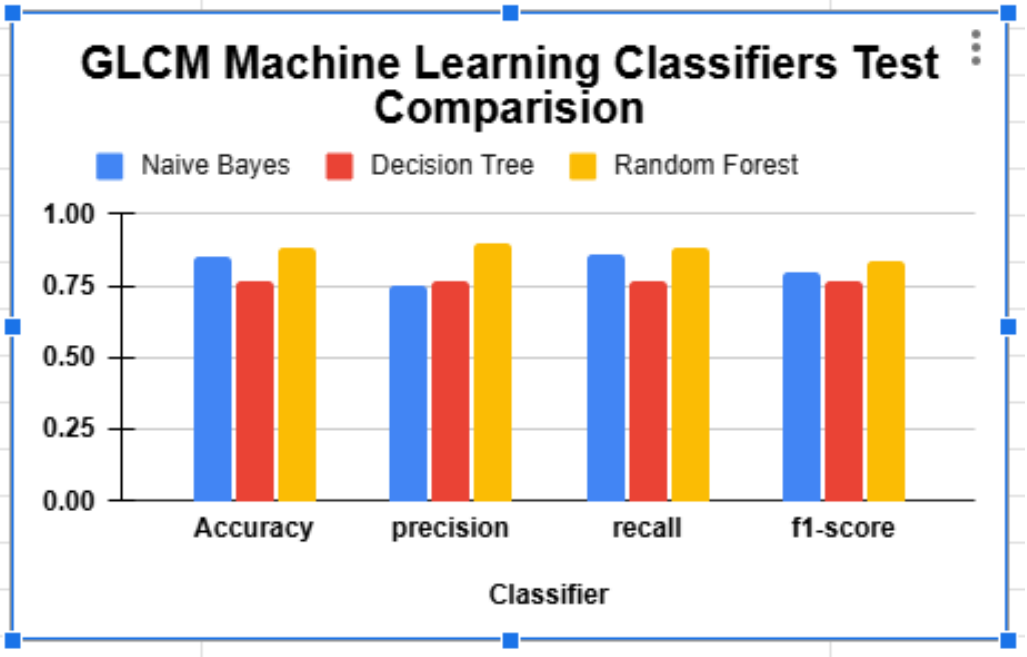

Fig.3 Comparative study of classifier

5.1. Result of Naive Bayes-

In Glcm test Classifier the accuracy of Naïve bayes is 85%. Python modules are utilized in Google Colab, and the Weka tool implements machine learning categorization.

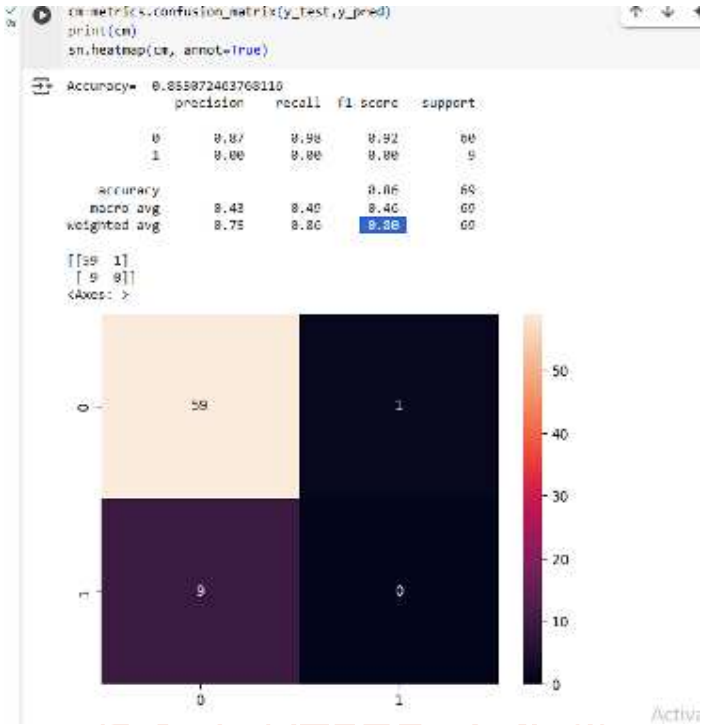

Fig.4 -Naive Bayes

5.2. Result of Decision Tree-

Fig.5 shows the accuracy, precision recall & F1 score of decision tree. A confusion matrix shows the correct & incorrect predictions for each class. The accuracy of decision tree is 76%

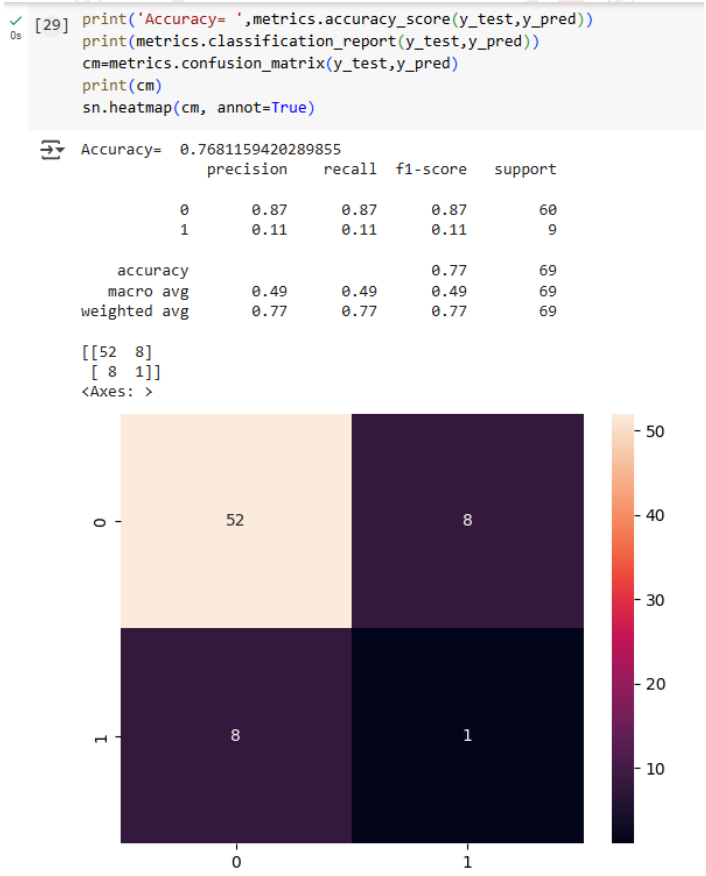

Fig.5 –Decision Tree

5.3. Result of random forest classifier[20]

```
print(cm)
sn.heatmap(cm, annot=True)
```

```
Accuracy=  0.8840579710144928
              precision    recall  f1-score   support

           0       0.88      1.00      0.94        60
           1       1.00      0.11      0.20         9

    accuracy                           0.88        69
   macro avg       0.94      0.56      0.57        69
weighted avg       0.90      0.88      0.84        69

[[60  0]
 [ 8  1]]
<Axes: >
```

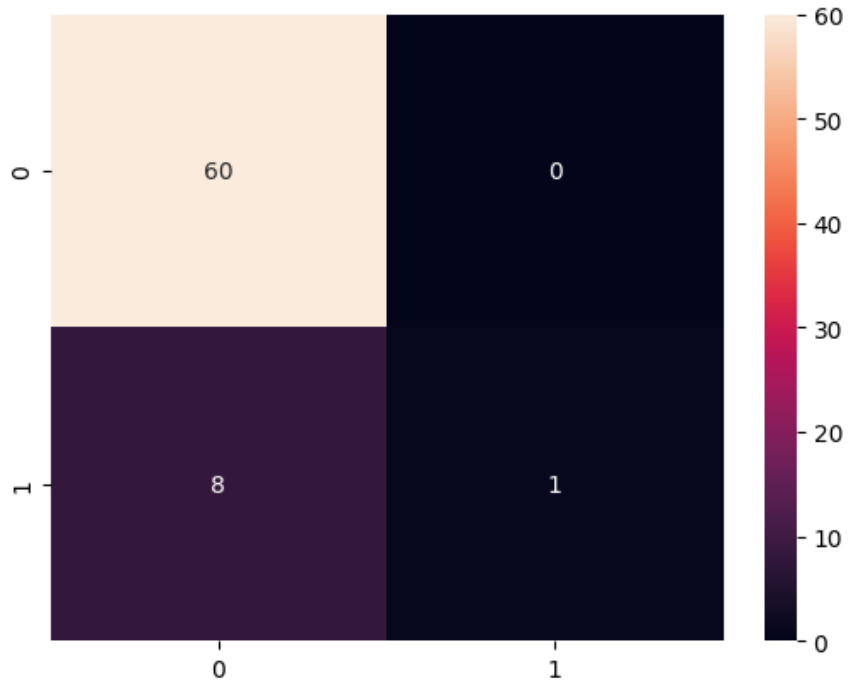

Fig.6 –Random forest Tree.

Random forest classifier has highest accuracy than decision and naive Bayes algorithm.

6. Conclusion

In this paper GLCM is used for image feature extraction & classification methods . The main part of this paper is dataset which is used detection of OA. In this paper GLCM achieves accuracy of 89% for feature fusion. For achieving high accuracy random forest classifier is used. Future research, however, might examine customizing deep learning models to improve performance even more as dataset sizes increase. This opens the door to more accurate and broadly applicable diagnostic instruments for individuals with osteoarthritis. To further enhance classification performance and resilience, future research should concentrate on optimizing feature selection, validating the model using a variety of external datasets, and investigating cutting-edge machine learning approaches.

References-

[1] Liu, X., Virk, S., Fedorova, T., Oo, W. M., & Hunter, D. J. (2023). The effect of pentosan polysulfate sodium for improving dyslipidaemia and knee pain in people with knee osteoarthritis: A pilot study. Osteoarthritis and Cartilage Open, 5(2), 100343. https://doi.org/10.1016/j.ocarto.2023.100343

[2] Gornale, Shivanand, "Detection of Osteoarthritis using Knee X -Ray Image Analyses: A MachineVision based Approach ", International Journal of Computer Applications (0975 – 8887) Volume 145 No.1, July 2016.

[3] Tiwari A, Poduval M, Bagaria V. Evaluation of artificial intelligence models for osteoarthritis of the knee using deep learning algorithms for orthopedic radiographs. World J Orthop. 2022

Jun 18;13(6):603-614. doi:10.5312/wjo.v13.i6.603. PMID: 35949704; PMCID: PMC9244962.

[4] Hegadi, R. S., et al. (2019). "Osteoarthritis Detection and Classification from Knee X-ray Images Based on Artificial Neural Network." *Recent Trends in Image Processing and Pattern Recognition*, 1036.

[5] Khalid, H., et al. (2020). "A Comparative Systematic Literature Review on Knee Bone Reports from MRI, X-rays, and CT Scans Using Deep Learning and Machine Learning Methodologies." *Diagnostics*,10(8), 518.

[6] Mahum, R., et al. (2021). "A Novel Hybrid Approach Based on Deep CNN Features to Detect Knee Osteoarthritis." *Sensors*, 21, 6189.

[7] P. Nguyen Huu, D. Nguyen Thanh, T. le Thi Hai, H. Chu Duc, H. Pham Viet and C. Nguyen Trong, "Detection and Classification Knee Osteoarthritis Algorithm using YOLOv3 and VGG-16 Models," 7th National Scientific Conference on Applying New Technology in Green Buildings (ATiGB), Da Nang, Vietnam, pp. 31-36, (2022).

[8] Usman Yunus(2022). " Recognition of Knee Osteoarthritis (KOA) Using YOLOv2 and Classification Based on Convolutional Neural Network," Volume 12

[9] Hema Rajini, N., Anton Smith, A., "Osteoarthritis Detection and Classification in Knee XRay Images Using Particle Swarm Optimization with Deep Neural Network", In: Kose, U., Gupta. D, Khanna, A., Rodrigues, J.J.P.C. (eds) Interpretable Cognitive Internet of Things for Healthcare. Internet of Things. Springer, Cham, (2023)

[10] T. Tariq, Z. Suhail and Z. Nawaz, "Machine Learning Approaches for the Classification of Knee Osteoarthritis," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 2023, pp. 1-6, doi:10.1109/ICECCME57830.2023.10252236.

[11] S. Aladhadh, "Identifying and localizing specific features or objects in images using DenseNet: The impact of model performance on data," 2023.

[12] Soh, L.K.; Tsatsoulis, C. Texture Analysis of SAR Sea Ice Imagery Using Gray-level Co-Occurrence Matrices. IEEE Trans. Geosci. Remote Sens. 1999, 37, 780–795.

[13] Chung, M.K. Introduction to logistic regression. arXiv 2020, arXiv:2008.13567

[14] L. Breiman, "Random Forests," Machine Learning, (2001).

[15] Chen, Pingjun , "Knee Osteoarthritis Severity Grading Dataset", Mendeley Data, V1, (2018).

[16] S. U. Rehman and V. Gruhn, "A Sequential VGG16+CNN-Based Automated Approach With Adaptive Input for Efficient Detection of Knee Osteoarthritis Stages," in *IEEE Access*, vol. 12, pp. 62407-62415, 2024, doi:10.1109/ACCESS.2024.3395062.

[17] Touahema, S.; Zaimi, I.; Zrira, N.; Ngote, M.N.; Doulhousne, H.; Aouial, M. MedKnee: A New Deep Learning-Based Software for Automated Prediction of Radiographic Knee Osteoarthritis. Diagnostics 2024, 14, 993. https://doi.org/10.3390/diagnostics14100993

[18] Y. Wang et al., "Learning From Highly Confident Samples for Automatic Knee Osteoarthritis Severity Assessment: Data From the Osteoarthritis Initiative," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 1239-1250, March 2022, doi:10.1109/JBHI.2021.3102090.

[19] Ijaz, M.F.; Woźniak, M. Editorial: Recent Advances in Deep Learning and Medical Imaging for Cancer Treatment. *Cancers* 2024, *16*, 700. https://doi.org/10.3390/cancers16040700.

[20] N. Kumar Kar, S. Jana, A. Rahman, P. Rahul Ashokrao, I. G and R. Alarmelu Mangai, "Automated Intracranial Hemorrhage Detection Using Deep Learning in Medical Image Analysis," *2024International Conference on Data Science and Network Security (ICDSNS)*, Tiptur, India, 2024, pp. 1-6, doi:10.1109/ICDSNS62112.2024.10691276.

[21] El-Shafai, W., Mahmoud, A.A., Ali, A.M. et al. Efficient classification of different medical image multimodalities based on simple CNN architecture and augmentation algorithms. J Opt 53, 775–787(2024). https://doi.org/10.1007/s12596-022-01089-3.