Knee Osteoarthritis Severity Classification using SVM

Shilpa H L[1*], *M. N.* Veena[1], and *M. L.* Anitha[2]

[1]Master of Computer Applications, PES College of Engineering, Mandya, India
[2]Department of Computer Science and Engineering, PES College of Engineering, Mandya, India

Abstract. Numerous ailments have adversely affected people's lifestyles in recent years. Among these, osteoarthritis (OA) of the knee has been identified as the main source of disability and restriction of activity, especially in senior citizens. It is essential to create fast and cost-efficient methods for diagnosing knee OA. In this paper, we analyze severity of OA in knee X-ray images by employing two machine learning (ML) models: decision tree (DT), as well as support vector machine (SVM). For capturing characteristics of the knee, we utilized histogram of oriented gradients (HOG) technique. Classifiers benefit from an enriched feature space. According to empirical findings, the SVM classifier achieved a 70% accuracy rate, outperforming the other.

1 Introduction

A knee osteoarthritis is a chronic degenerative joint disease. It is illustrated by progressive loss of cartilage between joints and development of new bone at the surface and margins of joint. Over time, OA is likely to deteriorate, leading to joint discomfort, stiffness, and significant impairments in knee movement. Studies show that obesity, advancing age, occupation, trauma, female gender, family history, metabolic illnesses, nutrition, smoking, bone density, and ethnicity are the main risk factors for knee OA. Injuries to the joints, overuse, and genetic factors may also contribute. A mismatch between cartilage deterioration and chondrocytes' healing causes structural problems in the joint.

Primary and secondary osteoarthritis are the two categories into which it falls. Primary osteoarthritis is an age-related joint condition. Age causes the cartilage's water content to diminish, increasing its vulnerability to deterioration. Secondary arthritis often affects the joints as a result of certain factors including obesity, diabetes, and trauma.

Millions of people worldwide suffer from OA, most common form of arthritis. In India, its prevalence ranges from 22% to 39%, making osteoarthritis the second most widespread musculoskeletal condition, according to the

National Health Portal. Around 15% of all musculoskeletal issues are caused by OA, which ranks ninth globally among all disorders [1]. India is predicted to lead the globe in chronic diseases till 2025 and has the highest rate of OA proliferation in the world [2]. Of people over 65, 70% have radiographic evidence of OA, and nearly 45% have symptoms. The World Health Organization (WHO) reports that symptomatic OA affects 18.0% of females over 60 and 9.6% of males. Compared to Western nations, OA affects a far higher number of people, especially in rural India. Lack of medical facilities and understanding of dietary behaviors may be the cause. [3].

Osteoarthritis (OA) is widely acknowledged as an independent risk factor that increases the likelihood of mortality. This disease has a significant impact on healthcare services, with the cost of OA to these services anticipated to double by 2020 and again by 2030 [4]. Other than total knee replacement (TKR) surgery at the advanced stage, there is currently no effective treatment for knee OA. As a result, early detection may still be the only option for slowing its course and mitigating the burden of eventual disability. If it can be resolved, the challenging and significant problem of predicting the advancement of OA may speed up the discovery of disease-modifying drugs and eventually prevent the millions of joint replacement procedures that are performed annually. From an economic standpoint, OA incurs substantial expenses for society, and the costs of these operations are high.

The diagnostic foundation for characterizing OA is radiodiagnosis and clinical symptoms. Reduced range of motion, joint discomfort, and stiffness are all signs of clinical OA. The diagnosis of radiographic OA is primarily based on the assessment of osteophyte formation and extent of joint space narrowing observed in imaging studies. One of most widely utilized systems in grading severity of osteoarthritis, Kellgren-Lawrence (KL) grading system. This approach assigns 5 grades to severity of OA, varying from Grade0, indicating no OA radiographic characteristics, to Grade4, which signifies severe OA characterized by significant marked joint space narrowing as well as osteophyte development.

The key contributions of this research are summarized as follows:

* Corresponding author: shilpa@pesce.ac.in

- Development of a cost-effective and rapid OA severity assessment framework that operates on knee X-ray images, enabling automated and scalable diagnosis.
- Integration of the Histogram of Oriented Gradients (HOG) feature extraction method to capture distinctive morphological patterns of the knee joint relevant to disease grading.
- Comparative evaluation of Decision Tree (DT) and Support Vector Machine (SVM) classifiers, with experimental results demonstrating that SVM achieves superior performance in an enhanced feature space.

The organization of this paper is as follows: Section 2 reviews the related work, Section 3 details the proposed methodology, Section 4 describes the evaluation techniques, Section 5 presents the results and discussion, and Section 6 provides the conclusion and outlines potential future enhancements.

2 Related Work

KOA investigation is one of several studies that use artificial intelligence analytics as a result of increased data collection. The literature describes many techniques that use ML Models to predict the KOA.

Raza et al. [5] increased precision of knee osteoarthritis diagnosis and stage classification using radiographic images by utilizing feature extraction and machine learning. Gaussian Naive Bayes, KNN, SVM, DT, RF, and XGBoost machine learning classifiers they tested using 3154 knee X-ray images. An ensemble model was created to lower the possibility of overfitting Zhao et al. [6] identified structural variables related to pain severity within individuals with knee osteoarthritis by analyzing knee radiographs as well as MRI data from 567 participants in Osteoarthritis Initiative. The most frequent problems, they discovered, comprised effusion/synovitis as well as cartilage loss. For feature extraction, Tariq et al. [7] utilized Haralick Features and Histogram of Oriented Gradients. They investigated machine learning techniques for multi-class classification, binary classification, as well as ordinal regression.

Yoo et al. [8] studied the development rate and outcomes of main knee osteoarthritis (OA) on basis of patient-specific conditions. They collected data from 83,280 individuals with knee pain between 2003 and 2019, focusing on 2,492 knees. The analysis included demographic, radiological, socioeconomic factors, comorbidities, and surgical interventions. Key influences on OA progression included Kellgren-Lawrence (K-L) grade, bone mineral density (BMD), as well as physical activity levels. Among patients who had surgery within the past five years, 49.0% had an initial K-L grade of 2, and 41.3% were employed in physically demanding jobs.

Gornale et. al [9] presents an automated method for early detection and classification of knee osteoarthritis (OA) using Hu's invariant moments to analyze geometric transformations in knee X-ray images. The methodology involves preprocessing to detect knee bone contours and reduce noise, followed by segmentation of the cartilage region, which is the primary area affected by OA. The study utilized a dataset of 2,000 digital knee X-ray images, manually annotated by medical experts, and implemented model using K-NN classifier and Decision Tree classifier.

Kokkotis et al.'s work [10] created a fuzzy ensemble feature selection approach that combines fuzzy logic findings from many fuzzy search methods. The approach yielded a classification accuracy of 73.55%, and an explainability analysis was employed to ascertain selection process for the optimal model.

Charis Ntakolia et al. [11] suggest machine learning techniques to forecast the progress of knee joint space narrowing (JSN) in each knee and both knees together. The techniques include post-hoc interpretation of features' influence on optimal model, robust feature selection to identify risk variables, clustering to identify progressing and non-progressing JSN, and classification algorithms for decision-making. The findings demonstrate that more accurate and robust prediction models are produced when the JSN advancement of both knees is bound.

Abbia Mahum et al. [12] propose an approach, that segments the Region of Interest and preprocesses X-ray images using hybrid feature descriptors like CNN and HOG. HOG is used to record Low-level features, whereas LBP extracts textural qualities. KOA severity is classified using SVM, RF, and KNN. The experimental data indicates that HOG achieves good accuracy for all KL grades, making it a practical tool for improving KOA patients' quality of life. The study [13] used patient data to show that high cholesterol, obesity, diabetes, and hypertension increase the risk of knee osteoarthritis progression.. The XGBoost model performed best, confirming cholesterol as the strongest predictor and highlighting the potential of cholesterol management to slow disease worsening.

Rima Tri et al. [14] used CNN and LSTM algorithms for feature extraction and classification, respectively. It was developed using an OAI dataset. The study achieved good accuracy, but the limitations include manual image cropping. The model could be extended to segmentation tasks and used to optimize techniques to improve classification accuracy. The study suggests that this approach could be extended to predict JSN pixel labels utilizing sparse training data.

Eirini et.al. [15] utilized deep neural network for classifying knee osteoarthritis using OAI dataset It includes 4796 participants' 141 risk variables. They divided dataset into three categories: incidence (no symptomatic knee OA, age eligibility), progression (frequent knee symptoms), and non-exposed control group (no knee symptoms). The model achieved an overall accuracy of 79.39%, but 57.37% of participants has been misclassified to incidence class in class 3.

Slim Ben Hassine et. al. [16] study introduced a framework for automatic knee osteoarthritis (KOA) detection and severity grading using radiographic images. The method incorporated a region-of-interest (ROI) localization stage, followed by a convolutional neural network (CNN)–based classification stage aligned with Kellgren–Lawrence (KL) grading

standards. Preprocessing steps included image normalization, histogram equalization, and cropping to enhance joint visibility and feature representation. The approach was evaluated on publicly available datasets, achieving high accuracy and area under the curve (AUC) values across severity grades, and demonstrating robustness to image variations. Key advantages of the framework include its modular design for localization and classification, use of clinically interpretable grading criteria, and strong performance across multiple datasets. However, the study also acknowledged certain limitations, such as reliance on X-ray imaging alone, potential overfitting to specific datasets, and the absence of large-scale multi-center validation to ensure broader clinical applicability. A CNN-based framework was introduced by Sozan Mohammed Ahmed et al. [17] for the automated detection and grading of knee osteoarthritis (KOA) severity from plain radiographs, utilizing Kellgren–Lawrence (KL) grading as the reference standard. The approach incorporated preprocessing steps such as region-of-interest cropping and image enhancement to optimize feature extraction. Trained on a large, annotated dataset, the model achieved high accuracy and area under the curve (AUC), surpassing the performance of several existing methods. Key strengths of the framework include its end-to-end automation, strong predictive capability, and reliance on a clinically recognized grading system. However, its dependence on a single imaging modality, potential dataset bias, and the absence of external multi-center validation limit its generalizability.

Ahmad Almhdie-Imjabbar et al.[18] proposed a combined prediction framework that integrates trabecular bone texture (TBT) descriptors, extracted via variogram-based fractal analysis, with deep learning–derived Kellgren–Lawrence (KL) scores from a Siamese convolutional neural network to predict knee osteoarthritis (KOA) progression. The method was evaluated on 2,740 knees from the OAI cohort and 845 knees from the MOST cohort, considering different image modalities, quality levels, and cross-cohort training scenarios. The TBT-CNN model achieved superior predictive performance, with AUC values up to 0.75 for OAI and 0.81 for MOST, and maintained robustness across acquisition modalities and quality conditions. Its ability to generalize between cohorts (AUC $\geq$ 0.70 in CR and CR&RG cases) highlighted strong transferability. Advantages of this approach include leveraging complementary radiographic texture and shape features, maintaining accuracy across varying image conditions, and demonstrating cross-dataset applicability. Limitations involve the absence of continuous joint space metrics (JSW/JSA), potential gains from additional clinical predictors, and differing follow-up periods between datasets, which restrict time-to-event analyses. Chen et al. [19] developed a predictive framework for KOA using clinical data combined with machine learning techniques. Drawing on a dataset of 2,594 samples, the study applied statistical methods and LASSO regression to identify 44 key clinical indicators. Feature importance was ranked using SHAP values. Several machine learning models—including Random Forest (RF), Logistic Regression, XGBoost, SVM, Naive Bayes, and Decision Tree—were trained using different feature subsets. Among these, the RF model achieved the highest diagnostic accuracy, with area under the curve (AUC) scores in both training and validation datasets. Additionally, recursive partitioning analysis revealed age above 54 as a significant risk factor. The study [20] demonstrates the effectiveness of integrating routine clinical and laboratory parameters with machine learning for early KOA detection, offering a non-imaging alternative for risk stratification in clinical settings.

Numerous studies have presented techniques for detecting knee osteoarthritis employing MRI and X-ray images from OAI dataset. These methods typically involve morphological processing, classification algorithms, as well as feature extraction. This indicates potential for further research utilizing different datasets and methodologies.

3 Proposed methodology

The proposed system is implemented to classify images utilizing different severity grades of knee X-ray images using ML methods. The HOG method is used for feature extraction. The ML algorithms are then utilized to train model. Models are evaluated based on the classification report. Figure 1 illustrates the proposed system.

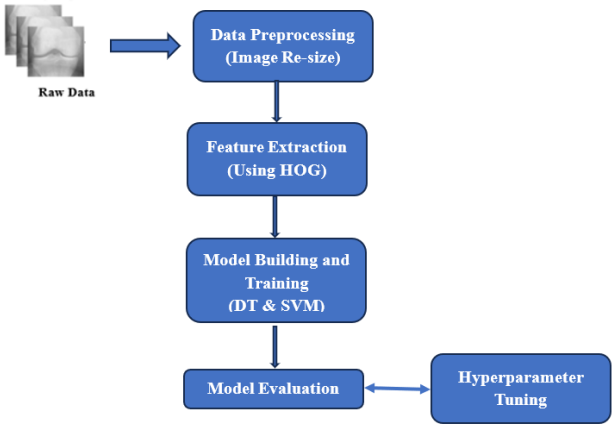


Fig. 1. Proposed system process.

3.1 Data acquisition

The Mendeley website served as the source of the dataset:https://data.mendeley.com/datasets/t9ndx37v5h/1A total of 1650 images in various grades are included. As per KL grading system, medical professionals manually comment on each radiographic knee X-ray image. Every image is categorized into one of 5 categories: KL0(Normal), KL1(Doubtful), KL3(Mild), KL4(Moderate), and KL5(Severe). An X-ray machine PROTEC PRS 500E is used to collect the images.

3.2 Data preprocessing

The dataset is organized into folders representing severity classes (e.g., grades 0, 1, 2, etc.). Manually removing the images that contain both the legs images in the single x-ray image from the dataset resulted in a total of 1465 images. All images are resized to uniform dimension of 128×128 pixels.

3.3 Feature extraction

The HOG method was employed for feature extraction. This method is particularly effective for detecting edges, textures, and other significant features in images.

The steps involved:

- Gradient Calculation: Gradients were computed for the horizontal and vertical directions using Sobel filters. The magnitude and orientation of the gradients were derived for each pixel.
- Cell and Block Formation: The images were divided into 16×16 pixel cells, and gradient orientation was created for each cell. To account for lighting variations, 2×2 cells were grouped into overlapping blocks, and L2-Hys normalization was applied to ensure robustness.
- Feature Vector Creation: The concatenated histograms of all normalized blocks are used as a final feature vector.

3.4 Classification Models

In our approach, we applied two ML algorithms, DT and SVM, provided by the scikit-learn library, to classify knee images into different grades.

3.4.1 DT Classifier

It is a supervised learning technique used in ML for modeling and predicting findings on basis of input data. Structure of DT resembles a tree, where each internal node represents a test of a property, each branch indicates a specific attribute value, and each leaf node signifies the final decision or prediction.

Algorithm Steps:

S1. Start with the Entire Dataset D.

S2. Compute Node Impurity:

Use Entropy to evaluate the purity of the dataset using Equation (1):

$$H(D) = -\sum_{i=1}^{k} p_i \log_2 (p_i) \quad (1)$$

Where p_i is the proportion of samples in the class

S3. Split on the Best Feature:

Iterate over all features and thresholds to find the split that minimizes impurity as represented in Equation (2):

$$Gain = H(D) - \left(\frac{|D_{left}|}{|D|} H(D_{left}) + \frac{|D_{right}|}{|D|} H(D_{right}) \right) \quad (2)$$

S4. Divide the Dataset:

Split the dataset into:

D_{left} : Samples satisfying the split condition.

D_{right} : Remaining samples.

S5. Repeat for Child Nodes:

Continue to apply the splitting process recursively to D_{left} and D_{right} until a stopping criterion is reached.

3.4.2 SVM classifier

It is a kernel-based ML model, in this decision functions are computed directly from training data to optimize the margin between decision boundaries in a highly dimensional feature space. This method reduces training data classification mistakes and improves generalization.

Algorithm Steps:

S1. Input:

- Training data:
  $D = [(x_1, y_1), x_2, y_2), \ldots., (x_n, y_n)]$,
  where and $y_i \in (-1, 1\}$ are class labels.
- Hyperparameters:
  C: Regularization parameter.
  Kernel function $K(x, x')$: For mapping input data to higher dimensions (e.g., linear, RBF, polynomial).

S2. Model training steps:

- Initialization
  Define the optimization problem using Equation (3) to find weights W, bias b, and Lagrange multipliers α:

$$min \frac{1}{2} ||w||^2 + C \sum_{i=1}^{n} \xi_i \quad (3)$$

  Subject to:
  $y_i (w \cdot xi + b) \geq 1 - \xi_i \ , \xi \geq 0, \forall i$
  Where, ξ_i are snack variables for soft margins.
- Kernal Transformation
  Apply a kernel function $K(x, x')$ to handle non-linear separable data using Equation (4):

$$K(x, x') = exp(-\gamma || x - x' || 2) \quad (4)$$

- Optimization:
  Solve the dual problem using Lagrange multipliers using Equation (5):

$$max \sum_{i=1}^{n} \propto_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \propto_i \propto_j y_i y_j K(x_i, x_j) \quad (5)$$

Subject to:

$$0 \leq x_i \leq C, \sum^{n} \propto_i y_i = 0$$

Decision Boundary

Compute W and b using Equations (6) and (7) :

$$w = \sum_{i=1}^{n} \propto_i y_i x_i \quad (6)$$

$$b = h_k - w \cdot x_k \quad (7)$$

for any k with $\alpha k > 0$

4 Evaluations

Model evaluation ensures that the selected model is reliable, interpretable, and capable of meeting demands of application. It serves as a checkpoint to validate whether model may be deployed in real-world scenarios or requires further tuning. The classification report is generated to verify the model performance. On basis of result of confusion matrix values precision, recall, accuracy, and f1-scores metrics might be measured, the Equations (8), (9), (10) and (11) are used to calculate the same.

$$Accuracy = \frac{Number\ of\ Correct\ Prediction}{Total\ Number\ of\ Prediction} \quad (8)$$

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \quad (9)$$

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (10)$$

$$F1 - score = 2 * \frac{Precision * Recall}{recision + Recall} \quad (11)$$

5 Results and discussion

This section reports experimental findings of DT and SVM classifiers. For model analysis, dataset was divided into training and validation sets at a ratio of 80:20. Applied 5 fold cross validation.

Initially, the DT classifier has been developed using default settings, achieving an overall accuracy of 44.03% when applied to the test data. The model was then fine-tuned by exploring various combinations of parameters. The parameter grid included: `max_depth`:{5, 10, 15, 20}; `criterion`:{gini, entropy}; `min_samples_split`:{2, 5, 10}; `min_samples_leaf`:{1, 2, 5}; `max_features`:{None, sqrt, log2}. A grid search was executed using this parameter grid to determine optimal parameters. Best parameters found for dataset were:{ `criterion`: entropy, `min_samples_split`: 10, `max_depth`:15, `min_samples_leaf`:2,`max_features`: None }. Using these ideal parameters, a new model was developed, which achieved overall accuracy 45% on test data, average precision 44.8%, recall of 43%, and average F1-score 43.6%.

Since DT classifier model accuracy is less than 50%, we chose to apply another classification technique, the support vector machine. The initial SVM model was generated using just two parameters: kernel type and regularization with 'linear' and 1.0 values, respectively, and attained an accuracy of 65.53%. Then, to modify the model, the optimum search technique, GridSearchCV, is performed on the parameter grid. The given parameter grid consists of kernel types {'linear,' 'rbf,' 'poly'}, regularization parameters (C){0.1,1,10,100}, kernel coefficient (gamma) values {'scale,' 'auto, 0.001, 0.1, 1}, and degrees of polynomial kernel values {2, 3, 4}. The best parameters determined after the optimum search are {'C': 10, 'degree': 2, 'gamma': 'auto', 'kernel': 'rbf'}. Once again, the model is constructed using ideal parameters and has accomplished an accuracy of 69.97% on test data. An average recall, precision, and f1-score were 69.8%, 68%, and 67.8% respectively.

The classification report for all five distinct classes of both models is presented in Table 1.

Table 1. Classification Report of DT and SVM Classifiers.

Measures	KL Grades				
	KL0	***KL1***	***KL2***	***KL3***	***KL4***
	DT Classification Model				
Precision	.54	.43	.29	.55	.43
Recall	.56	.45	.35	.40	.39
F1-Score	.55	.44	.32	.46	.41
Accuracy	45%				
	SVM Classification Model				
Precision	.76	.62	.64	.79	.68
Recall	.78	.71	.35	.77	.79
F1-Score	.77	.66	.45	.78	.73
Accuracy	70%				

For the DT classifier, Grade 0 (Normal), performs relatively well with an f1-score of 0.55, correctly identifying healthy cases in most instances. Similarly, Grade 3 (Moderate) achieves an f1-score of 0.46, indicating some reliability. However, performance is weaker for Grade 1 (Doubtful), Grade 2 (Mild), and Grade 4 (Severe), with f1-scores of 0.44, 0.32, and 0.41, respectively. These grades exhibit significant misclassification, likely due to overlapping features and subtle differences in severity.

The SVM classifier demonstrates strong performance for Grade 0, 3, and 4 classes, with f1-scores of 0.77, 0.78, and 0.73, correspondingly, showing appropriate balance between precision as well as recall. However, it struggles with Grade 2 class, achieving only a 0.45 f1-score due to low recall (35%), and the Grade 1 class performance is moderate with an F1-score of 0.66, affected by lower precision (62%).

Overall, the model performs well on distinct categories but faces challenges with overlapping or subtle distinctions between adjacent severity levels.

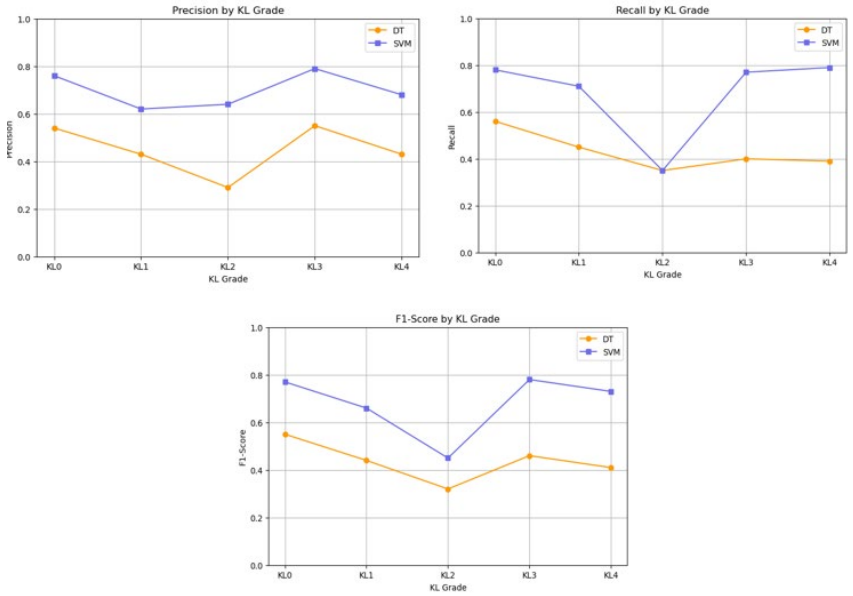


Fig. 2. Precision, Recall, F1-score of DT and SVM models.

Figure 2 shows the graphical comparison of Precision, recall and F1-score across KL grades (KL0-KL4) for both DT and SVM models. We can observe that SVM outperformed DT across all metrics and KL grades. The performance gap is most notable in KL2 and KL4, where DT struggled significantly. SVM shows higher consistency, especially in KL3 and KL4 grades—critical for late-stage KOA diagnosis.

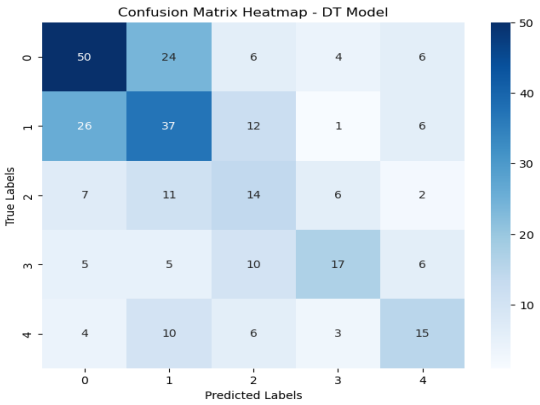
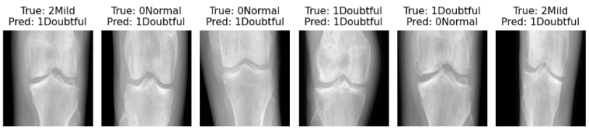


Fig. 3. Confusion matrix heatmap of the DT classifier.

Figure 3. is a confusion matrix resulted from DT model. It performs well for Grades 0, indicating better classification at the extreme ends of the KL scale. There is substantial confusion in grades 1 to 4, which could be due to overlapping of images. Grade 1 significantly overlaps with Grade 0, suggesting the need for better feature extraction or more robust criteria to distinguish doubtful OA from healthy joints. Figure 4 shows sample images of predicted classes.



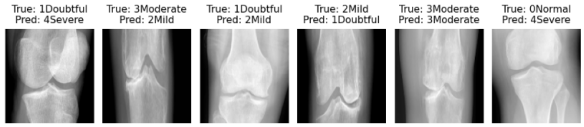


Fig. 4. Sample output images after applying the DT model.

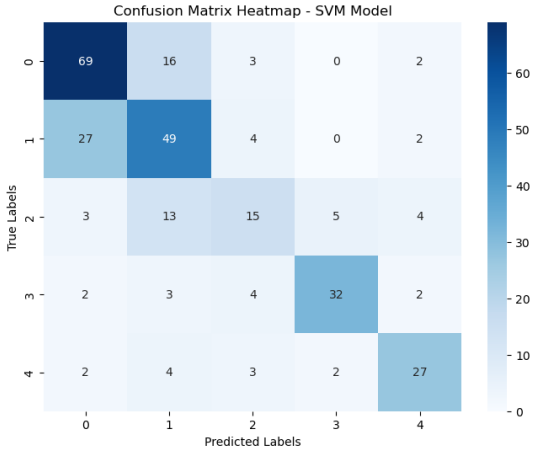


Fig. 5. Confusion matrix heatmap of the SVM classifier.

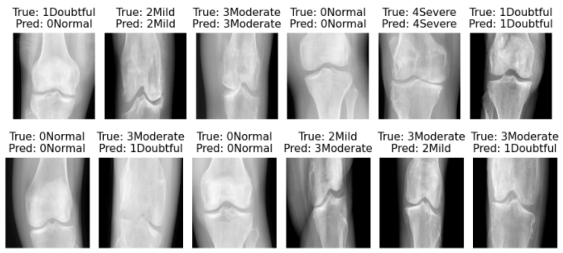


Fig. 6. Sample output images after applying the SVM model.

Figure 5. shows confusion matrix of the SVM model and sample output images. It performs well in classifying Grade 0, (69 correct), Grade 3 (32 correct), and Grade 4 (27 correct) cases, as evidenced by high true positive counts. However, there is significant misclassification between adjacent classes, such as Grade 1 being confused with Grade 0 (27 cases) and

Grade 3 being confused with Grade 2 (13 cases). Additionally, Grade 3 exhibits the most confusion overall, with predictions spread across multiple other classes. This suggests that the model handles distinct categories effectively but struggles with overlapping or subtle boundaries between severity levels. Figure 6 Shows the sample output images of SVM model.

6 Conclusion and future enhancement

This study employs ML models to utilize X-ray images to categorize the level of knee osteoarthritis. Two models, SVM and DT, are utilized, incorporating histogram of oriented gradients method to extract features. Both classifiers improve knee OA classification performance by using enriched feature space. The SVM classifier outperforms with a 70% accuracy rate. Since there is room to enhance the accuracy achieved, our further work will concentrate on implementing advanced feature engineering, addressing data imbalances, and utilizing ensemble approaches to enhance classification performance.

References

1. A. Tiulpin, S. Klein, S.M.A. Bierma-Zeinstra, J. Thevenot, E. Rahtu, J. van Meurs, E.H.G. Oei, S. Saarakkala, Multimodal machine learning-based knee osteoarthritis progression prediction from plain radiographs and clinical data. Sci. Rep. **9**, 20038 (2019). https://doi.org/10.1038/s41598-019-56527-3
2. C.S. Azad, A.K. Singh, P. Pandey, M. Singh, P. Chaudhary, N. Tia, A. Rastogi, Osteoarthritis in India: an epidemiologic aspect. Int. J. Recent Sci. Res. **8**, 20918-20922 (2017)
3. P.P. Kawathekar, K.J. Karande, Severity analysis of osteoarthritis of knee joint from X-ray images: A literature review, in Proceedings of the International Conference, Signal Propagation and Computer Technology (ICSPCT 2014), IEEE, Ajmer, India, August 28 (2014), 648-652
4. A. Jamshidi, J.-P. Pelletier, J. Martel-Pelletier, Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nat. Rev. Rheumatol. **15**, 49-60 (2019). https://doi.org/10.1038/s41584-018-0130-5
5. A. Raza, T.-L. Phan, H.-C. Li, N.V. Hieu, T.T. Nghia, C.T.S. Ching, A comparative study of machine learning classifiers for enhancing knee osteoarthritis diagnosis. Information **15**, 183 (2024). https://doi.org/10.3390/info15040183
6. Z. Zhao, M. Zhao, T. Yang, J. Li, C. Qin, B. Wang, L. Wang, B. Li, J. Liu, Identifying significant structural factors associated with knee pain severity in patients with osteoarthritis using machine learning. Sci. Rep. **14**, 14705 (2024). https://doi.org/10.1038/s41598-024-65613-0
7. T. Tariq, Z. Suhail, Z. Nawaz, Machine learning approaches for the classification of knee osteoarthritis, in Proc. 3rd Int. Conf. Electrical, Computer, Communications and Mechatronics Engineering (ICECCME 2023), IEEE, Tenerife, Canary Islands, Spain, July 19-21 (2023), 1-6
8. H.J. Yoo, H.W. Jeong, S.B. Park, S.J. Shim, H.S. Nam, Y.S. Lee, Do individualized patient-specific situations predict the progression rate and fate of knee osteoarthritis? Prediction of knee osteoarthritis. J. Clin. Med. **12**, 1204 (2023). https://doi.org/10.3390/jcm12031204
9. S.S. Gornale, P.U. Patravali, P.S. Hiremath, Automatic detection and classification of knee osteoarthritis using Hu's invariant moments, Front. Robot. AI **7**, 591827 (2020).
10. C. Kokkotis, C. Ntakolia, S. Moustakidis, G. Giakas, D. Tsaopoulos, Explainable machine learning for knee osteoarthritis diagnosis based on a novel fuzzy feature selection methodology. Phys. Eng. Sci. Med. **45**, 219-229 (2022). https://doi.org/10.1007/s13246-022-01106-6
11. C. Ntakolia, C. Kokkotis, S. Moustakidis, D. Tsaopoulos, Prediction of joint space narrowing progression in knee osteoarthritis patients. Diagnostics **11**, 285 (2021). https://doi.org/10.3390/diagnostics11020285
12. R. Mahum, S.U. Rehman, T. Meraj, H.T. Rauf, A. Irtaza, A.M. El-Sherbeeny, M.A. El-Meligy, A novel hybrid approach based on deep CNN features to detect knee osteoarthritis. Sensors **21**, 6189 (2021). https://doi.org/10.3390/s21186189
13. Li, Hong-bo, Yong-jun Du, Guy Romeo Kenmegne, and Cheng-wei Kang, Machine learning analysis of serum cholesterol's impact on knee osteoarthritis progression, Scientific Reports **14**, 18852. (2024)
14. R.T. Wahyuningrum, L. Anifah, I.K.E. Purnama, M.H. Purnomo, A new approach to classify knee osteoarthritis severity from radiographic images based on CNN-LSTM method in Proceedings of the 10th IEEE International Conference, Awareness Science and Technology (iCAST 2019), IEEE, Morioka, Japan, December 05 (2019), 1-6
15. E. Christodoulou, S. Moustakidis, N. Papandrianos, D. Tsaopoulos, E. Papageorgiou, Exploring deep learning capabilities in knee osteoarthritis case study for classification, in Proceedings of the 10th IEEE International Conference, Information, Intelligence, Systems and Applications (IISA 2019), IEEE, Patras, Greece, November 14 (2019), 1-6
16. S. Ben Hassine, A. Balti, S. Abid, M.M. Ben Khelifa, M. Sayadi, Markerless vision-based knee osteoarthritis classification using machine learning and gait videos. Front. Signal Process. **4**, 1479244 (2024). https://doi.org/10.3389/frsip.2024.1479244
17. S.M. Ahmed, 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 **12**, 2939

(2022).
https://doi.org/10.3390/diagnostics12122939

18. A.I. Ahmad, N. Khac-Lan, T. Hechmi, J. Rachid, L. Eric, Prediction of knee osteoarthritis progression using radiological descriptors obtained from bone texture analysis and Siamese neural networks: data from OAI and MOST cohorts. Arthritis Res. Ther. **24**, 1-13 (2022).
19. W. Chen, H. Zheng, B. Ye, T. Guo, Y. Xu, Z. Fu, Q. Deng, Identification of biomarkers for knee osteoarthritis through clinical data and machine learning models. Sci. Rep. **15**, 1703 (2025). https://doi.org/10.1038/s41598-025-85945-9
20. L. Shen, S. Yue, A clinical model to predict the progression of knee osteoarthritis: data from Dryad, J. Orthop. Surg. Res. **18**, 628 (2023). https://doi.org/10.1186/s13018-023-04118-4