IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 June 2023, accepted 9 July 2023, date of publication 12 July 2023, date of current version 18 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294542

RESEARCH ARTICLE

Transfer Learning-Based Smart Features Engineering for Osteoarthritis Diagnosis From Knee X-Ray Images

AMJAD REHMAN[1], (Senior Member, IEEE), ALI RAZA[2], FATEN S. ALAMRI[3], BAYAN ALGHOFAILY[1], AND TANZILA SABA[1], (Senior Member, IEEE)

[1]Artificial Intelligence Data Analytics Laboratory (AIDA), CCIS, Prince Sultan University, Riyadh 11586, Saudi Arabia
[2]Institute of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan
[3]Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

Corresponding author: Faten S. Alamri (fsalamri@pnu.edu.sa)

This research was funded by Princess Nourah bint Abdulrahman University and Researchers Supporting Project number (PNURSP2023R346), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT Osteoarthritis is a deteriorating joint disease affecting millions worldwide. Osteoarthritis is a chronic condition that develops over time due to joint wear and tears. The degeneration of joint cartilage is the underlying cause of osteoarthritis, resulting in bone-to-bone contact and contributing to stiffness, discomfort, and restricted movement. People with osteoarthritis struggle to perform simple tasks such as walking, standing, or climbing stairs. Moreover, osteoarthritis can also cause psychological distress, including depression and anxiety, due to the chronic pain and disability associated with the condition. Improving the quality of life requires the development of efficient methods for early detection. Our study aims to create a model that can effectively diagnose osteoarthritis in knee X-ray images at an early stage. The advanced deep learning-based Convolutional Neural Network (CNN) and several machine learning-based techniques are applied in comparison. A novel transfer learning-based feature engineering technique CRK (CNN Random forest K-neighbors) is proposed to detect osteoarthritis with high performance. Using a 2D-CNN, the proposed CRK smartly extracts the spatial features from the X-ray images. The spatial features are input to the random forest and k-neighbors techniques, creating a probabilistic feature set. The probabilistic feature set is utilized to build the applied machine learning-based techniques. Extensive study experiments demonstrate that the proposed model outperformed with a 99% accuracy score for predicting osteoarthritis. The performance of each applied model is validated using hyperparameter optimization and k-fold-based cross-validation. The proposed study has the potential to revolutionize the prediction of osteoarthritis from X-ray images with high-performance scores.

INDEX TERMS Osteoarthritis, X-ray images, knee X-ray, transfer learning, smart feature engineering, deep learning, machine learning.

I. INTRODUCTION

Osteoarthritis is a deteriorating joint syndrome that affects the knee joint. Osteoarthritis is a form of arthritis, often called "wear and tear" arthritis [1]. Osteoarthritis of the knee occurs when the cartilage cushioning the knee joint wears away over time, resulting in bone-on-bone contact, which causes pain, swelling, stiffness, and reduced mobility. Factors that increase the risk of developing knee osteoarthritis include age, obesity, previous joint injury, and genetic factors [2]. Treatment for knee osteoarthritis may involve a combination of lifestyle modifications such as weight loss, exercise, physical therapy, and medications such as pain relievers and corticosteroid injections. Extreme situations

The associate editor coordinating the review of this manuscript and approving it for publication was Wenbing Zhao.

may require surgical intervention, such as joint replacement, to alleviate discomfort and improve the range of motion. Osteoarthritis disease affects millions of people worldwide [3]. Osteoarthritis significantly impacts individuals, as it causes pain, stiffness, and decreased mobility, which can lead to disability. The cost of osteoarthritis in terms of healthcare and lost productivity is substantial and expected to rise as the population ages. Moreover, the disease can also negatively impact individuals' mental health, leading to depression and anxiety. Therefore, timely prediction of osteoarthritis using advanced artificial intelligence helps develop effective treatment and management strategies and improve the quality of life for those affected by the disease [4].

Neural networks and deep learning have become essential tools for recognising medical images, including X-rays [5]. Deep learning algorithms can learn complex patterns and features within images, allowing them to identify subtle differences that may be difficult for human eyes to detect [6]. Deep learning models are typically trained in X-ray image recognition using large datasets of X-ray images annotated with labels corresponding to different diagnoses. The deep learning models classify X-ray images based on their visual features [7], allowing healthcare professionals to accurately diagnose conditions such as pneumonia, lung cancer, or bone fractures. Although there exist difficulties related to applying deep learning for the recognition of X-ray images, including the large amounts of training data and the potential for biases in the models, these techniques have shown promise in improving the accuracy and efficiency of medical imaging analysis.

Transfer learning-based feature engineering has emerged as a promising technique for improving the accuracy of X-ray image recognition [8]. Transfer learning in image classification tasks involves using pre-trained neural networks to extract best-fit features from image data [9]. The pre-trained network's weights and architecture are usually frozen, and the output of the network's last layer is used as features for a new classifier. In the case of X-ray images, a neural network can extract significant characteristics through a transfer process [10]. Transfer learning feature engineering also allows researchers to customize pre-trained networks for specific tasks, such as identifying lung nodules or detecting pneumonia [11]. Overall, transfer learning feature engineering [12] has shown promising results in improving the accuracy of X-ray image recognition and is likely to continue to be a valuable tool in medical imaging research. The main contributions of our research study related to osteoarthritis prediction are as follows:

- A novel approach CRK is used to extract spatial features from the x-ray images. The extracted spatial features are input to the random forest and k-neighbors techniques to create the probabilistic features set, achieving high performance for predicting osteoarthritis.
- Advanced CNN, random forest, support vector machine, SGD classifier, and k-neighbors methods are applied in comparison to evaluate performance. The outperformed technique achieved high-performance scores compared to the state-of-the-art studies.
- The hyperparameter optimizations and k-fold-based cross-validation techniques are applied to validate the accuracy performance of the machine learning and deep learning methods.

The remaining study follows: Section II comparatively analyzed the related literature for predicting osteoarthritis using X-ray images. Our novel proposed methodology mechanism is described in Section III. The study's experimental evaluations are performed in Section IV. The study findings and conclusions are described in Section V.

II. RELATED WORK

The use of transfer learning-based feature engineering to predict osteoarthritis using knee X-ray images is an active area of research. Several recent studies have focused on developing machine learning models for accurate and early diagnosis of osteoarthritis using X-ray images. Deep learning techniques have been applied to radiographic images to detect and classify osteoarthritis in recent years. This section reviews the related work in deep learning for osteoarthritis diagnosis using knee radiographs, also comparatively analyzed in Table 1.

The knee osteoarthritis detection based on plain radiographs using a deep learning technique was proposed in [13]. The classical deep siamese convolutional neural network was the proposed neural network used for the detection task. The dataset based on 3,000 subjects with 5,960 knees image from the osteoarthritis initiative was utilized to build this study's applied deep learning techniques. The deep learning proposed approach achieved a poor performance of an average multi-class accuracy of 66.71%.

Diagnosing knee osteoarthritis from T2 maps based on a deep learning approach was presented in [14]. A densely connected CNN was trained to diagnose osteoarthritis in this study. The data based on 4,384 subjects from the osteoarthritis initiative was utilized for evaluating the applied neural network. The study results show that the proposed approach achieved low-performance AUC accuracy scores of 83.44%, which can be further improved.

The study [15] proposed a deep-learning technique for predicting pain progression in knee osteoarthritis. The deep learning-based artificial neural network was the proposed technique used for osteoarthritis pain progression prediction. The dataset based on 4674 subjects with a risk of knee osteoarthritis was used to build the applied neural network technique [28]. The proposed neural network achieved an AUC accuracy score of 80%, considered a low score that needs improvement.

The prediction of osteoarthritis knee progression using the deep learning technique was proposed in [16]. The deep CNN was the proposed approach for predicting osteoarthritis knee progression. The plain radiographs and clinical data from 2,129 subjects were utilized to build this study's applied deep learning techniques. The results of experiments show that

TABLE 1. **The related osteoarthritis literature summary analysis in image detection for osteoarthritis prediction.**

Ref.	Year	Dataset	Technique	Performance Accuracy
[13]	2018	Image data of 3,000 subjects related to osteoarthritis.	Deep Siamese Convolutional Neural Network	0.66
[14]	2019	Image data of 4,384 subjects from osteoarthritis initiative.	DenseNet	0.83
[15]	2021	Image data of 4674 subjects related to knee osteoarthritis.	Artificial neural network	0.80
[16]	2019	knee images from 2,129 subjects.	Deep Convolutional Neural Network	0.81
[17]	2019	Image data of 5749 subjects.	Deep Neural Network	0.76
[18]	2019	1024 knee X-ray images from database OsteoArthritis Initiative.	Naive Bayes and random forest	0.82
[19]	2021	Gait image dataset CASIA B.	Kernel Extreme Learning Machine	0.96
[20]	2021	9280 knee MRI images of 3268 patients.	Random predictor	0.72
[21]	2022	plain hand radiographs of patients.	Pre-trained VGG-16 network	0.90
[22]	2023	multivariate information-based physiological signals.	Improved deep learning model	0.93
[23]	2023	X-Ray Images data.	EfficientNetB1	0.89
[24]	2023	4403 knee X-rays from 1174 patients.	LR model	0.84
[25]	2023	MRI, X-ray, and the patient's clinical information.	Convolution neural networks	0.76
[26]	2022	X-ray image data.	Fine-tuned ResNet-34	0.61
[27]	2022	X-Ray Osteoarthritis Initiative data	Deep Neural Networks	0.83

the presented deep CNN achieved an AUC accuracy score of 81%, which is not up to the mark.

The deep neural network-based osteoarthritis diagnosis using a deep learning approach was presented in this study [17]. The deep neural network was the proposed technique for detecting osteoarthritis. The image data of 5749 subjects based on health behaviour information was used to build the applied deep learning techniques. The authors also used the principal component analysis for feature engineering. The study results show that the proposed neural network with PCA achieved a poor AUC accuracy score of 76.8%, which must be improved.

The study [18] proposed a machine-learning approach for the early diagnosis of knee osteoarthritis using x-ray images. The data of 1024 x-ray images from the osteoarthritis initiative database were utilized for training and testing the applied machine learning approach. Random forest and naive bayes were used to detect knee osteoarthritis in this study. The feature extraction based on independent component analysis is applied to image data to reduce the dimensionality. The proposed machine learning-based classifier achieved an accuracy score of 82.98%, which is also low in comparison.

The human gait-based osteoarthritis knee prediction using the deep learning framework and kernel extreme learning machine was proposed in [19]. The publicly available image gait dataset CASIA B was utilized to build the applied deep learning methods. The feature extraction technique was also used to enhance the performance. The study results show that the kernel extreme learning machine with the extracted features achieved an acceptable accuracy score of 96%. However, not the highest.

The study [20] proposed a neural network method to predict knee osteoarthritis radiographic progression. The 9280 knee MRI images [29] based on 3268 patients from the osteoarthritis initiative were utilized to build the applied deep learning techniques. The neural network-based attention maps were provided, and the random predictor achieved the 72% for detecting knee osteoarthritis, which is considered low-performance.

The automated deep learning-based diagnosis of osteoarthritis, rheumatoid arthritis, and normal radiographs hand was proposed in [21]. The data from plain radiographs hand of patients were utilized to build the applied technique in this study. The YOLO method obtained hand images from the original radiographs, and a pre-trained VGG-16 network performed classification [30]. The study experiments show that the transfer learning-based proposed technique achieved a performance score of 90%, which are still low.

The study [22] proposed a computer-involved detection method of knee osteoarthritis based on multivariate data using the deep learning approach. The multivariate information-based physiological signals data was used to build the applied deep learning techniques. The study results demonstrate that the improved deep learning method achieved an acceptable performance score of 93% but can be further improved to achieve the highest.

The deep learning-based knee osteoarthritis severity stage prediction was proposed in this study [23]. The Western subjects' x-ray images of knee osteoarthritis were utilized to build the applied deep learning methods. The transfer learning-based EfficientNetB1 was the proposed technique which is trained and tested using the x-ray Images. The study experiments demonstrate that the proposed method achieved and tested an accuracy score of 89% on the data from Indian subjects, which is considered low.

The nomogram method based on signatures radionics data and age to detect knee osteoarthritis was proposed in [24]. The data of 4403 knee x-rays images from 1174 patients was utilized to build the applied machine learning techniques. The support vector machine, logistic regression, adaboost, gradient boosting, and multi-layer perceptron were applied in comparisons. The radionics feature data were extracted from the X-ray data to enhance the performance. The study results show that the applied LR method achieved a low-performance score of 84%.

The knee osteoarthritis classification using deep learning based on the multimodal intermediate fusion of X-ray was presented in [25]. The dataset based on MRI, x-ray and the patient's clinical information was utilized to build this study's applied deep learning approach. The CNN was the proposed approach to classify knee osteoarthritis. The study results show that the applied technique achieved 76% of performance accuracy scores, which are Low and need further improvements.

The classification and detection of osteoarthritis knee using deep neural network techniques was proposed in [26]. The authors proposed a semi-CADx method based on Deep Siamese CNN with the fine-tuned architecture of ResNet-34 for detecting osteoarthritis. The dataset used in this study was imbalanced, which affected the performance scores. The proposed transfer learning based fine-tuned ResNet-34 achieved an abysmal performance score of 61%, which must be improved to the highest.

The anterior part and posterior part detection of osteoarthritis knee using X-Ray Images were presented in this study [27]. The deep learning techniques are applied with random brightness augmentation to enhance the performance. The X-Ray Osteoarthritis Initiative data was utilized to build this study's applied neural network techniques. The proposed deep learning neural network achieved 83% performance accuracy scores for detecting knee osteoarthritis.

In conclusion, several studies have investigated deep-learning techniques for osteoarthritis diagnosis using knee radiographs. Previous studies have shown a central research gap we have identified during related literature analysis:

- The performance accuracy scores in previous studies for detecting osteoarthritis using knee x-ray images are deficient. Classical neural network-based techniques are mainly adopted for osteoarthritis detection from image data. More advanced techniques are needed to detect osteoarthritis with high-performance scores.

III. PROPOSED METHODOLOGY

The proposed study methodology architecture is based on the prediction of osteoarthritis with a high-performance model, as illustrated in Figure 1. The existing knee-based x-ray images dataset is used to build the applied machine learning and deep learning techniques. The transfer learning-based feature engineering technique CRK is applied to extract the best optimal features from x-ray image data, which helps to predict osteoarthritis with performance accuracy scores. The extracted transfer features are then split into two protons of data with a ratio of 80:20. The 80% part of the data is used for training, and the 20% part of data is used for validating the performance of applied machine learning techniques. The outperformed tuned approach is used for detecting osteoarthritis with high performance from knee-based x-ray images.

A. KNEE BASED X-RAY IMAGES

The knee-based x-ray images dataset [31], publicly available at the famous repository Kaggle is utilized to conduct our study experiments. The image dataset contains 3615 files belonging to 2 classes, normal and osteoarthritis, as illustrated in Figure 2. The analysis describes that the osteoarthritis (1) class contains 2026 image samples, and the normal (0) class contains 1589 samples. The dataset distribution analysis concludes that image labels are imbalanced. Furthermore, the x-ray images data analysis with their target label is illustrated in Figure 3.

B. NOVEL TRANSFER LEARNING-BASED FEATURE ENGINEERING

A novel transfer learning-based feature engineering technique is proposed to diagnose osteoarthritis using x-ray image data of the knee. The working mechanism of the proposed feature engineering technique is illustrated in Figure 4. The proposed CRK technique combines three methods to transfer learning experiences and makes a new feature set. First, the convolutional neural network-based spatial features are extracted from the X-ray images dataset. Then, the extracted spatial features are input to the random forest and k-neighbors techniques. The transfer learning-based probabilistic features [32] set is formed from spatial feature data. The probabilistic features set is utilized to build the applied methods for detecting osteoarthritis using x-ray image data of the knee. The novel proposed transfer learning-based feature has revolutionized the prediction of osteoarthritis knee from X-ray images with high-performance scores.

1) SPATIAL FEATURES

Let's assume the input dataset $\mathcal{D}$ consisting of N samples, where each sample is a 2D image represented by a matrix $\mathbf{X} \in \mathbb{R}^{H \times W}$, where H and W denote the height and width of the image, respectively. Our research goal is to extract spatial features using a 2D CNN model.

The 2D CNN model applies convolutional layers and non-linear activation functions to capture local patterns in the input images. Let $\mathbf{X}_l$ denote the output feature map at the l-th layer of the CNN. We can express the convolution operation at layer l as:

$$\mathbf{X}_l = \sigma\left(\mathbf{W}_l * \mathbf{X}_{l-1} + \mathbf{b}_l\right), \tag{1}$$

where $*$ denotes the convolution operation, $\mathbf{W}_l$ represents the convolutional filters or kernels of size $F_l \times F_l$, $\mathbf{X}_{l-1}$ is the

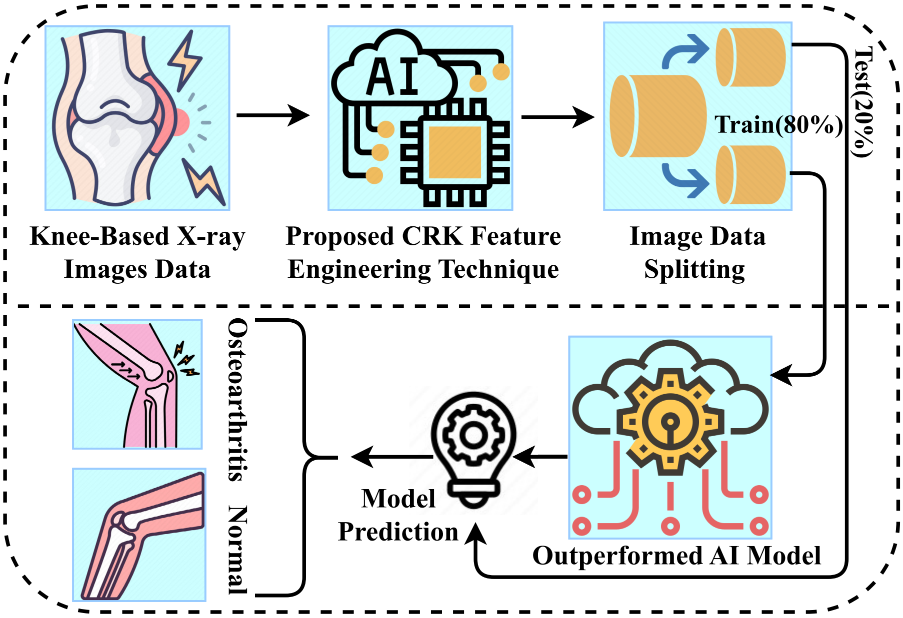


FIGURE 1. The methodology architecture analysis of our study approach for the diagnosis of osteoarthritis disease using knee-based x-ray data.

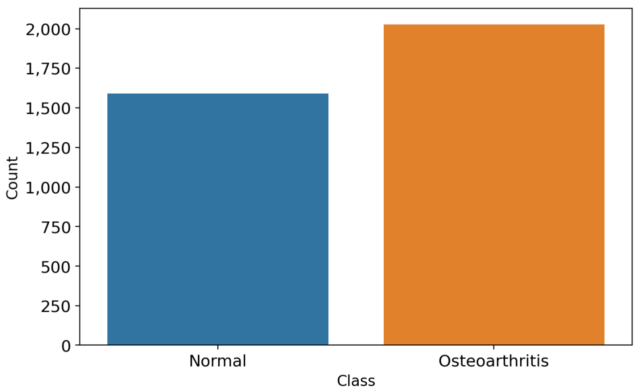


FIGURE 2. The knee x-ray images data distribution analysis with the target label.

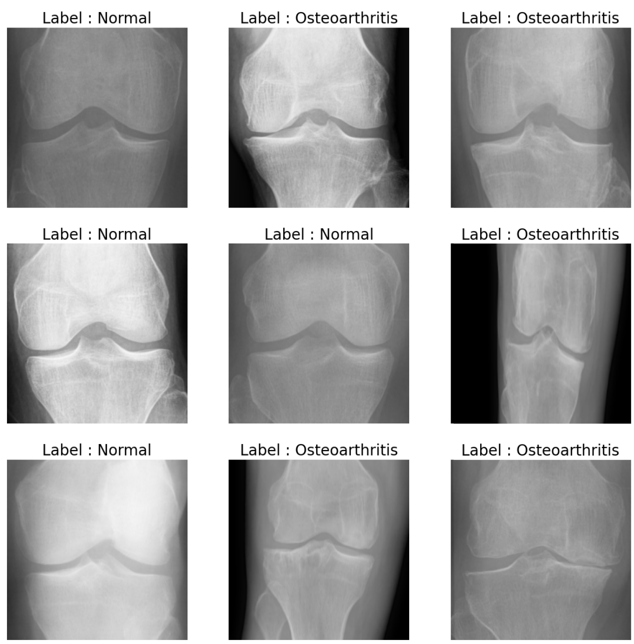


FIGURE 3. The knee x-ray images data analysis with the target label.

input feature map from the previous layer, $\mathbf{b}_l$ is the bias term, and $\sigma(\cdot)$ is the activation function, such as the ReLU function. After applying multiple convolutional layers, we obtain a set of feature maps $\{\mathbf{X}_l\}_{l=1}^{L}$, where L denotes the total number of layers in the CNN model. Each feature map represents a different level of abstraction, capturing increasingly complex spatial patterns.

2) PROBABILISTIC FEATURES

Let X be the input data and Y be the output variable. We can define a probabilistic model $P(Y|X)$ that predicts the output variable Y given the input data X. This model can be represented using the following equation:

$$P(Y|X) = \frac{P(X, Y)}{P(X)} \tag{2}$$

where $P(X, Y)$ is the joint probability distribution of X and Y, and $P(X)$ is the marginal probability distribution of X. The feature extraction technique transforms raw input data X into a set of meaningful and informative features F. This process can be represented as a function $f(X)$:

$$F = f(X) \tag{3}$$

where F represents the extracted features.

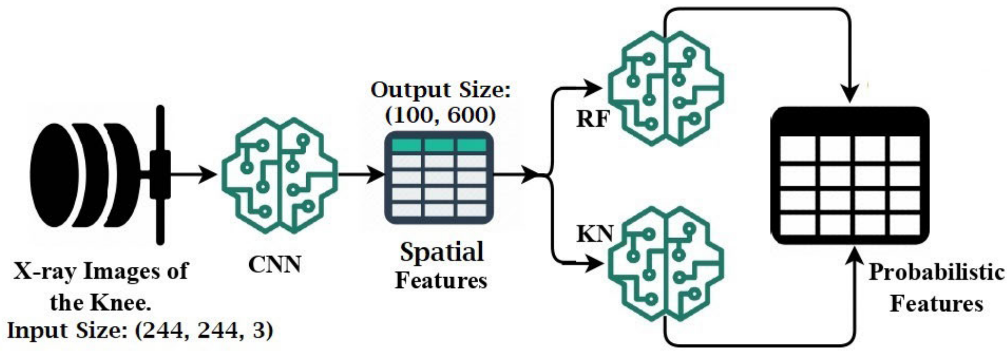


FIGURE 4. The novel proposed transfer learning-based feature engineering working flow architecture.

Algorithm 1 shows the step-by-step flow of the proposed approach.

Algorithm 1 CRK Algorithm

Input: X-ray images data of the knee.
Output: Transfer learning-based features.

initiate;
1- $F_{cnn} \longleftarrow CNN_{prediction}(XiS)$ // $XiS \in X - ray\, images\, set$, here XiS is original image data and F_{cnn} is extracted spatial features set.
2- $F_{rf} \longleftarrow RF_{probabilities\, prediction}(F_{cnn})$ // here F_{rf} is the extracted probabilistic based features set.
3- $F_{kn} \longleftarrow KN_{probabilities\, prediction}(F_{cnn})$ // here F_{kn} is the extracted probabilistic based features set.
4- $F_{Prob} \longleftarrow F_{rf} + F_{kn}$ // here F_{Prob} is final probabilistic based features set used for osteoarthritis detection.
end;

C. IMAGE DATA SPLITTING

Dataset splitting is an essential step in machine learning tasks that involve supervised learning. The goal of splitting a dataset is to divide it into two subsets, one for training the model and the other for evaluating its performance. In this study, we used a 80:20 splitting ratio of data, where 80% of the image data is used for training the model, and 20% of the image data is used for testing the model. The training part is used to fit the model parameters, while the test set is used to validate the model's performance on unseen data. The dataset splitting helps prevent overfitting, a common problem when a model is trained on too little data. In summary, dataset splitting is essential in building robust and reliable machine learning models.

D. APPLIED MACHINE LEARNING AND DEEP LEARNING METHODS

Machine and deep learning methods have shown remarkable success in image classification [33], [34]. CNNs are a popular deep learning approach that has achieved outstanding performance in image detection tasks [35]. The CNN use multiple layers of filters to extract spatial features from image data and learn complex data representations. The extracted features from CNN can be used to train machine learning-based techniques for image classification [36], resulting in high performance.

In this study, we have applied CNN to classify and extract the spatial features from x-ray image data of the knee. The extracted spatial features used to build the machine-learning methods includes random forest classifier, support vector machines, SGD classifiers, and k-neighbors. The probabilistic-based features are extracted from spatial features using high-performer methods. The advanced machine learning techniques achieved high performance for predicting osteoarthritis from X-ray images. The layered architecture of the applied 2D-CNN model for feature extraction is analyzed in Table 2.

TABLE 2. Applied CNN model layered architecture analysis.

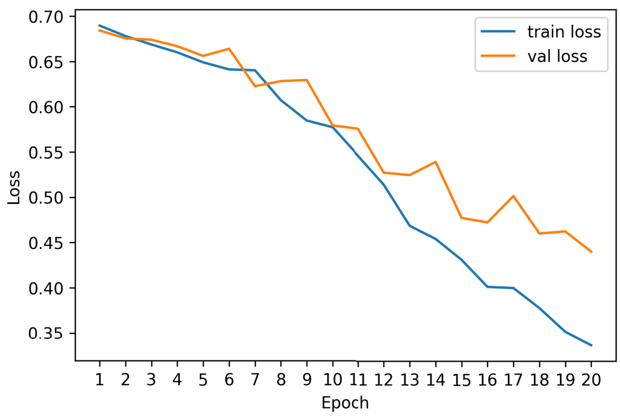
Layer (Type)	Output Shape	Parameters
Conv2D Layer	(100, 222, 222, 64)	1792
MaxPooling2D Layer	(100, 111, 111, 64)	0
Conv2D layer	(100, 109, 109, 128)	73856
MaxPooling2D Layer	(100, 54, 54, 128)	0
Flatten Layer	(100, 373248)	0
Dense Layer	(100, 600)	223949400
Total Parameter		224,025,048

E. HYPERPARAMETER TUNING

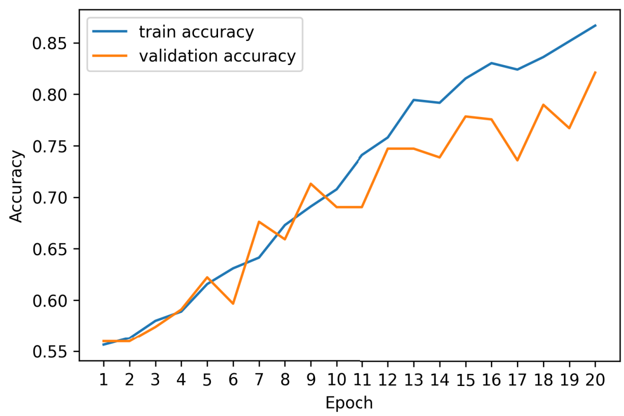
Hyperparameter tuning is a crucial step in machine learning that involves selecting the optimal set of hyperparameters for a given model. The hyperparameter tuning process involves trying different combinations of hyperparameters and evaluating the model's performance on a testing set. Hyperparameter tuning aims to find the set of hyperparameters that produces the best performance on the test set. Hyperparameter tuning is essential because it can significantly improve the performance of applied machine learning methods. The hyperparameter optimization of our applied techniques is analyzed in Table 3.

TABLE 3. The hyperparameter optimization analysis uses applied deep learning and machine learning techniques.

Technique	Hyperparameters Values
CNN	activation= 'sigmoid', optimizer='adam', loss = 'binary_crossentropy'
RF	max_depth=300, n_estimators=300, criterion="gini", max_features="sqrt", bootstrap=True
SVM	kernel='linear', degree=3, C=1.0, tol=1e-3, cache_size=200
SGD	max_iter=1000, tol=1e-3, loss='hinge', penalty='l2', tol=1e-3
KN	n_neighbors=2, weights='uniform', leaf_size=30, metric='minkowski'



(a)



(b)

FIGURE 5. The time series-based performance scores analysis of applied convolutional neural network during training.

IV. RESULTS AND DISCUSSIONS

The performance results validation and discussions of applied artificial intelligence-based techniques are analyzed in this section. The experimental setup and results with x-ray image data are comparatively described. The performance metrics, such as accuracy, recall, f1-score, and precision, are used to evaluate the model's effectiveness. Overall, the results and discussions section is essential in comprehensively evaluating the proposed machine learning and deep learning methods and their potential for solving real-world problems.

A. EXPERIMENTAL SETUP

In this study, machine learning models are built using popular Python libraries, including sklearn, keras, and tensorflow. The study experiments are conducted in a Google Colab environment using a high-end GPU backend with 13 GB RAM and 90 GB of disk space. The system model used in the experiments is an Intel(R) Xeon(R) processor. The experiments involved using Python 3 programming language to train and validate the performance of the models. The evaluation metrics used in the experiments are accuracy, precision, recall, f1 score, and time complexity to assess the performance of machine learning models.

B. CONVOLUTIONAL NEURAL NETWORK RESULTS

The classical CNN is applied to X-ray images, and performance is evaluated for a fair comparison. The time series-based analysis is performed while training an applied CNN with 20 epochs, as illustrated in Figure 5. Figure 5(a) study shows that the train and validations loss score values decrease as the epochs increase. This analysis demonstrates that the neural network learns more patterns from the X-ray image data with each epoch and adjusts the network weights to find optimal performance. Figure 5(b) analysis demonstrates that the accuracy score for training and validation is also increased with the increase in training epochs. The highest training accuracy score achieved at epoch 20 is 85%, and the validation accuracy score is above 80%. This analysis concludes that the classical convolutional neural network achieved acceptable scores during training. However, not the highest performance.

The performance comparison of applied deep learning-based convolutional neural networks on unseen data is illustrated in Figure 6. The analysis shows that the highest recall score is 89%, achieved by the convolutional neural network during testing. The performance metrics scores for accuracy, precision, and f1 are 78%, 75%, and 81%, respectively. This analysis concludes that the applied convolutional neural network achieved poor performance scores for detecting osteoarthritis using the knee's x-ray images.

C. RESULTS WITH SPATIAL FEATURES

The spatial features are extracted from the knee's x-ray image data using the deep learning-based CNN, and performance results are evaluated. The advanced machine learning-based techniques are trained and tested using the extracted spatial features.

TABLE 4. The comparative performance analysis of machine learning technique using spatial features on unseen test data.

Technique	Accuracy	Target	Precision	Recall	F1-score	Support
RF	0.76	0	0.79	0.66	0.72	335
		1	0.74	0.85	0.79	388
		Average	**0.77**	**0.76**	**0.76**	**723**
SVM	0.54	0	0.00	0.00	0.00	335
		1	0.54	1.00	0.70	388
		Average	**0.29**	**0.54**	**0.37**	**723**
SGD	0.54	0	0.00	0.00	0.00	335
		1	0.54	1.00	0.70	388
		Average	**0.29**	**0.54**	**0.37**	**723**
KN	0.63	0	0.56	0.87	0.68	335
		1	0.79	0.42	0.55	388
		Average	**0.69**	**0.63**	**0.61**	**723**

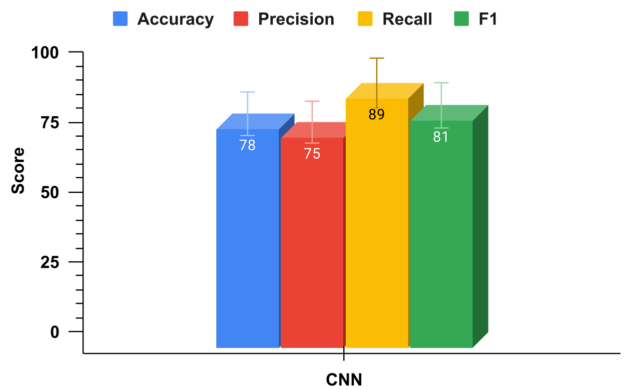


FIGURE 6. The bar chart-based performance metrics scores analysis of applied convolutional neural networks on unseen test data.

The performance comparison analysis of applied machine learning methods with spatial features is analyzed in Table 4. The study demonstrates that the applied SVM and SGD techniques achieved a poor accuracy performance score of 0.54 in comparison. The precision score achieved by the SVM and SGD techniques is 0.29, which is the lowest in comparison. The metric performance results with each class show that the SVM and SGD techniques also achieved low scores. The applied KN achieved a performance score of 0.63 which is better than the SVM and SGD techniques. The KN method achieved a precision score of 0.69. However, not the highest. Only the tree-based RF technique achieved acceptable scores in comparison. The RF technique achieved a 0.76 accuracy performance score and 0.77 precision score. The analysis concludes that the RF technique achieved good accuracy performic score of 0.76 using the extracted spatial features. The performance scores of applied machine learning techniques still need significant improvements to detect osteoarthritis using the knee's x-ray images.

The bar chart-based comparative performance analysis of used machine learning methods with spatial features is illustrated in Figure 7. The visual investigation shows that machine learning methods achieved poor performance scores with spatial features extracted from the original X-ray images. The RF technique achieves the maximum accuracy score of 76% in this analysis which is low. In conclusion, performance enhancement is needed for detecting osteoarthritis based on knee X-ray images.

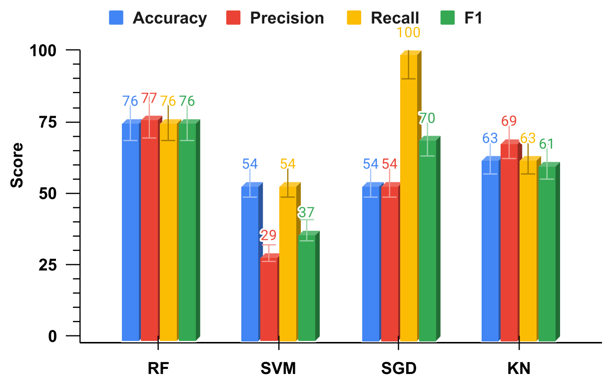


FIGURE 7. The bar chart-based performance metrics analysis of machine learning techniques with spatial features.

The confusion matrix analysis of machine learning techniques with spatial features is illustrated in Figure 8. The confusion matrix is a table that summarizes the predictions made by the model against the ground truth labels of a dataset. By analyzing the metrics in the confusion matrix, we can gain insights into the strengths and weaknesses of the model and identify areas for improvement. This confusion matrix analysis concludes that machine learning models achieved low-performance scores with spatial features. Applied methods achieved a high error rate for X-ray image classification.

D. RESULTS WITH PROBABILITY-BASED FEATURES

The probabilistic features are extracted from the spatial features using machine learning-based techniques, and performance results are evaluated. The best performer machine learning methods on spatial features are used to extract the probabilistic features. The applied machine learning-based techniques are then trained and tested using the extracted probabilistic features to evaluate the performance.

The performance comparison analysis of applied machine learning methods with probabilistic features is described in Table 5. The research demonstrates that using the probabilistic features, each applied machine learning technique

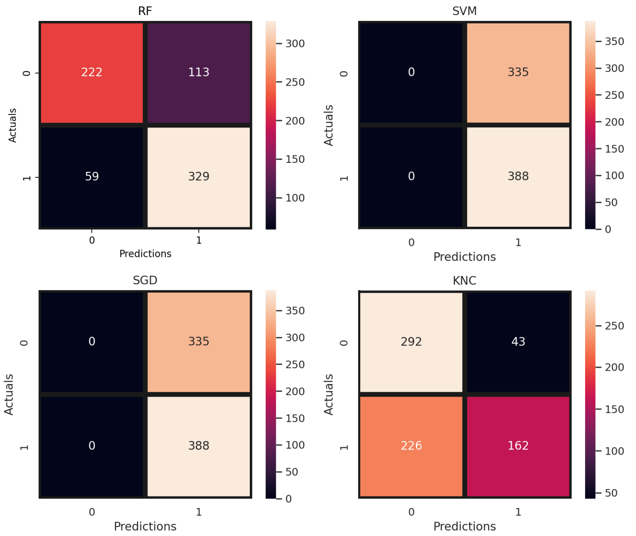


FIGURE 8. **The confusion matrix analysis of machine learning techniques with spatial features.**

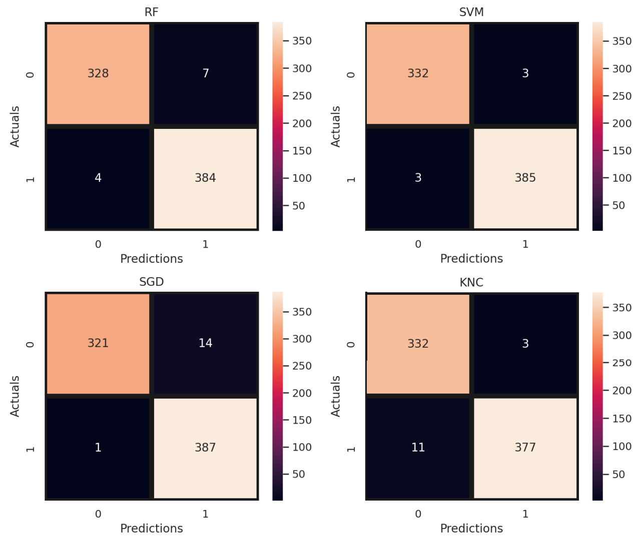


FIGURE 10. **The confusion matrix analysis of machine learning techniques with probabilistic features.**

scores the highest performance in comparison. The applied RF, SGD, and KN achieved a performance accuracy score of 0.98, which is well compared. The performance metrics score for each class also achieved the best scores. The applied SVM technique achieved the highest accuracy performance score of 0.99. This analysis concludes that the highest performance score is achieved using the extracted probabilistic features for detecting osteoarthritis using probabilistic features of x-ray images.

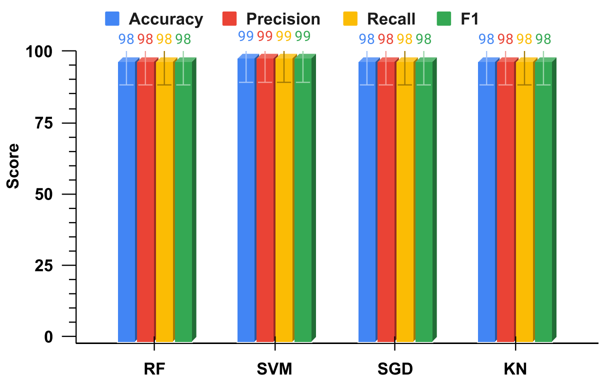


FIGURE 9. **The bar chart-based performance metrics analysis of machine learning techniques with probabilistic features.**

The bar chart-based comparative performance analysis of used machine learning methods with probabilistic features is illustrated in Figure 9. The research demonstrates a sudden increase in the performance accuracy scores with probabilistic features extracted from the spatial features. The proposed SVM technique achieves the highest accuracy score of 99% in this analysis. This analysis concludes that all applied machine methods achieved the best performance for detecting osteoarthritis based on knee x-ray images.

The confusion matrix analysis of machine learning techniques with probabilistic features is illustrated in Figure 10. The machine learning methods achieved high-performance scores with probabilistic features by analyzing the confusion matrix. The analysis concludes that applying the probabilistic features methods achieved a low error rate for x-ray image classification.

E. K-FOLD CROSS-VALIDATION-BASED PERFORMANCE ANALYSIS

The k-fold cross-validation analysis to evaluate the generalization of each applied model is performed in Table 6. Both feature data are validated using the ten folds for each applied technique in this analysis. The analysis shows that poor performance scores with high standard deviation scores are achieved during the cross-validation mechanism using extracted spatial features. The highest k-fold accuracy score is 0.75, acquired by the RF technique. The k-fold results comparisons demonstrate that the best accuracy scores are achieved using the probability features with minimum standard deviation scores. The highest k-fold accuracy of 0.99 is completed by our proposed SVM approach using the probability features. The analysis concludes all applied techniques are validated and are in generalized form for detecting osteoarthritis using x-ray images data.

The bar chart-based performance validation of applied methods based on 10-fold of data with both features is illustrated in Figure 11. The cross-validation analysis demonstrates that the highest performance scores are achieved using probabilistic features extracted from the spatial features. Our proposed approach achieved a 99% cross-validation accuracy score. The analysis concludes that a low-performance score is achieved using spatial features extracted from original X-ray images. However, with the probabilistic

TABLE 5. The comparative performance analysis of applied machine learning technique using probability-based features on unseen test data.

Technique	Accuracy	Target	Precision	Recall	F1-score	Support
RF	0.98	0	0.99	0.98	0.98	335
		1	0.98	0.99	0.99	388
		Average	**0.98**	**0.98**	**0.98**	**723**
SVM	**0.99**	0	0.99	0.99	0.99	335
		1	0.99	0.99	0.99	388
		Average	**0.99**	**0.99**	**0.99**	**723**
SGD	0.98	0	1.00	0.96	0.98	335
		1	0.97	1.00	0.98	388
		Average	**0.98**	**0.98**	**0.98**	**723**
KN	0.98	0	0.97	0.99	0.98	335
		1	0.99	0.97	0.98	388
		Average	**0.98**	**0.98**	**0.98**	**723**

TABLE 6. The performance validation analysis of applied machine learning techniques with 10-fold cross-validation.

Technique	With spatial Features		With Probability Features	
	K-fold Accuracy (%)	Standard Deviation (+/-)	K-fold Accuracy (%)	Standard Deviation (+/-)
RF	0.75	0.024	0.98	0.005
SVM	0.56	0.022	0.99	0.005
SGD	0.56	0.026	0.98	0.006
KN	0.62	0.031	0.94	0.025

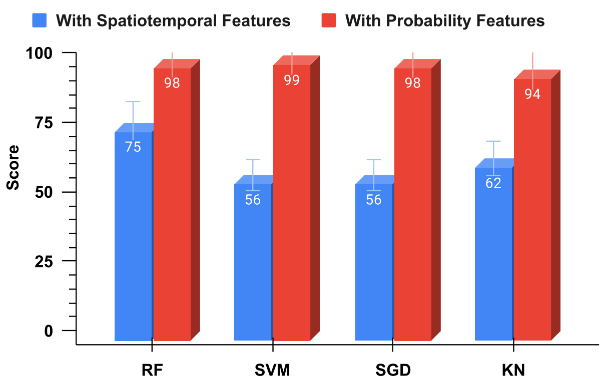


FIGURE 11. The bar chart-based k-fold cross-validation analysis of machine learning techniques.

features, we reached the highest performance score above 98% for each applied technique, showing the generalization for detecting osteoarthritis.

F. ANALYSIS OF RUNTIME COMPUTATIONAL COMPLEXITY

The runtime computations comparative analysis of applied machine learning techniques during the training is analyzed in Table 7. The computational complexity scores show that the applied machine learning techniques face high runtime computations cost with spatial features data. However, very minimum runtime computations cost is achieved by using the probability features. This analysis concludes that spatial feature data is expensive as computationally. RF and SVM have the highest time of 13.07 seconds and 2.037 seconds, respectively. Applied machine learning models can be trained in less time than other methods using probability features. The minimum training time is needed from the SGD model. However, SGD also has less accuracy of 56% with spatial features. In contrast, the proposed SVM method requires a longer training time than SGD. However, SVM provides the highest level of accuracy when it comes to making predictions.

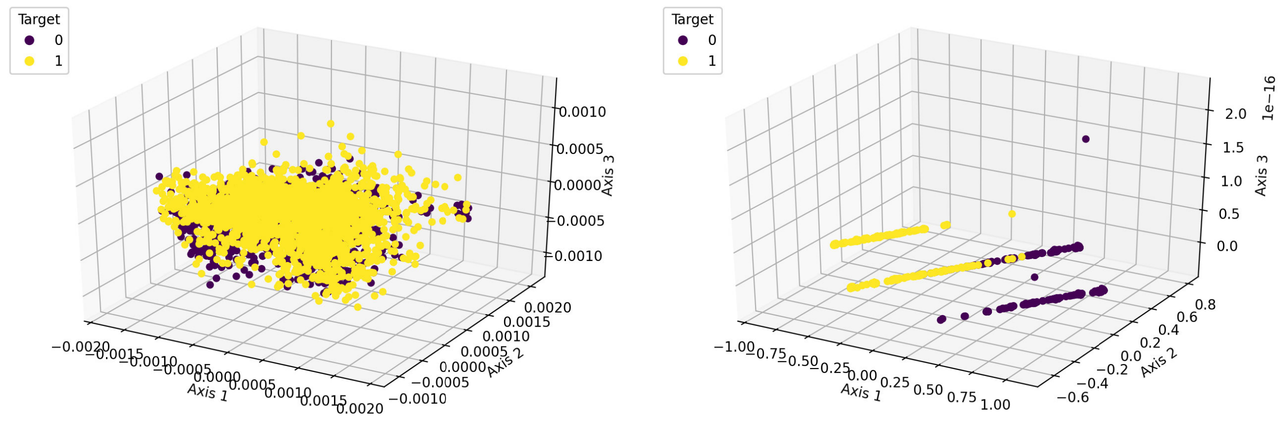
TABLE 7. The runtime computational complexity analysis of applied machine learning techniques.

Technique	Runtime Computational Cost (Seconds)	
	With Spatial features	With Probability Features
RF	13.07	0.494
SVM	2.037	0.014
SGD	0.144	0.005
KN	0.012	0.008

G. COMPARISON OF SPATIAL AND PROBABILITY-BASED FEATURES

The extracted spatial features by the CNN using the original knee's x-ray images are not linearly separable to a significant degree. The non-linearly separable behavior of spatial features results in the low performance of applied machine learning methods. The proposed CRK approach uses the RF and KN techniques to create a probability-based feature set. The feature set produced through probability-based generation displays a higher degree of linear separability, resulting in greater discrimination between the target classes with a larger margin.

The study experiments demonstrate that our proposed CRK approach significantly enhances the performance of applied machine learning techniques. The feature space analysis for both feature's representations to validate the improved performance is visualized in Figure 12. Figure 12(a) shows



(a) (b)

FIGURE 12. The analysis of the feature space based on class distribution using spatial and probabilistic features.

TABLE 8. The performance comparison of the proposed study with other state-of-the-art studies in detecting osteoarthritis using knee x-ray images.

Ref.	Year	Learning Type	Technique	Performance accuracy (%)
[37]	2022	Deep learning	CNN+ SVM	90
[23]	2023	Transfer Learning	EfficientNetB1	89
[38]	2021	Deep learning	CNN+ YOLO	95
[27]	2022	Deep learning	DCNN	83
[39]	2023	Deep learning	CNN	91
[40]	2020	Machine learning	LR	78
[41]	2022	Transfer Learning	EfficientNetB1	96
[42]	2022	Machine learning	KNN	90
Our study	**2023**	**Transfer Learning**	**CRK+SVM**	**99**

that the spatial features extracted using the original knee's x-ray images are not linearly separable. However, by applying our CRK feature engineering approach, the feature space becomes more conducive to linear separation, as Figure 12(b) visualizes. The performance improvement achieved through the CRK approach in this study is confirmed by the feature space analysis.

H. COMPARISONS WITH OTHER STATE-OF-THE-ART STUDIES

Table 8 examines the comparative analysis of the performance of our proposed study with other state-of-the-art studies. We have taken the recently published studies from the year 2020 to 2023 for a fair comparison. Researchers in previously published studies primarily used deep learning techniques and achieved the maximum performance score of 96% in this analysis, which is low. Our proposed research used transfer learning-based feature engineering techniques to enhance performance. This analysis concludes that our proposed SVM with transfer features achieved the highest performance scores of 99% for detecting osteoarthritis based on x-ray images of the knee.

I. DISCUSSION

The X-ray images of the knee are used to detect osteoarthritis with advanced neural network-based feature engineering techniques. Extensive results experiments are conducted to validate the performance of each applied model. The spatial and probabilistic features extracted from the knee's X-ray images are mainly utilized to compare the performance of applied machine-learning techniques. The result comparison shows that the applied machine learning techniques achieved a low-performance score using the spatial features extracted from original x-rays images. However, using the probabilistic features, the applied method performed the highest score.

In addition, we have applied the feature space analysis to show that the proposed CRK approach significantly enhances the performance of applied machine learning techniques. The study results demonstrate that the proposed SVM model outperformed the state-of-the-art studies. Furthermore, we have also analyzed the computational complexity of applied machine-learning techniques to test the latency for predicting osteoarthritis. In conclusion, our proposed research has the potential to revolutionize osteoarthritis prediction using x-rays images of the knee.

V. CONCLUSION

The diagnosis of osteoarthritis based on x-rays images of the knee using advanced transfer learning and machine learning techniques is proposed in this study. A novel transfer learning-based feature engineering technique CRK is proposed to detect osteoarthritis with high performance.

Classical CNN and several machine learning-based techniques are applied in comparison. Using a 2D-CNN, the proposed CRK extracts the spatial features from the X-ray images. Then probabilistic feature set is formed using the random forest and k-neighbors techniques. The extracted probabilistic feature set is used to build the applied machine learning-based methods. The support vector machine is our proposed model with high-performance scores of 99% for accuracy, recall, f1, and precision metrics. The performance validation of each applied model is conducted using hyperparameter optimization and k-fold-based cross-validation.

FUTURE DIRECTION

In the future, we will apply image augmentation techniques to balance the X-ray image data to enhance the performance for osteoarthritis detection. More advanced image processing techniques and neural networks will be implemented.

ACKNOWLEDGMENT

This research was funded by Princess Nourah bint Abdulrahman University and Researchers Supporting Project number (PNURSP2023R346), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

REFERENCES

[1] K. D. Allen, L. M. Thoma, and Y. M. Golightly, "Epidemiology of osteoarthritis," *Osteoarthritis Cartilage*, vol. 30, no. 2, pp. 184–195, Feb. 2022.

[2] A. Mahmoudian, L. S. Lohmander, A. Mobasheri, M. Englund, and F. P. Luyten, "Early-stage symptomatic osteoarthritis of the knee—Time for action," *Nature Rev. Rheumatol.*, vol. 17, no. 10, pp. 621–632, Oct. 2021.

[3] M. Herrera-Pérez, D. González-Martín, M. Vallejo-Márquez, A. L. Godoy-Santos, V. Valderrabano, and S. Tejero, "Ankle osteoarthritis aetiology," *J. Clin. Med.*, vol. 10, no. 19, p. 4489, Sep. 2021.

[4] R. Javed, M. S. M. Rahim, T. Saba, and A. Rehman, "A comparative study of features selection for skin lesion detection from dermoscopic images," *Netw. Model. Anal. Health Informat. Bioinf.*, vol. 9, no. 1, pp. 1–13, Dec. 2020.

[5] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet," *Chaos, Solitons Fractals*, vol. 138, Sep. 2020, Art. no. 109944.

[6] M. A. Khan, T. Akram, M. Sharif, K. Javed, M. Raza, and T. Saba, "An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection," *Multimedia Tools Appl.*, vol. 79, nos. 25–26, pp. 18627–18656, Jul. 2020.

[7] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio, and S. K. Antani, "Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays," *IEEE Access*, vol. 8, pp. 115041–115050, 2020.

[8] N. N. Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays," *IRBM*, vol. 43, no. 2, pp. 114–119, Apr. 2022.

[9] V. Chouhan, S. K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damasevicius, and V. H. C. De Albuquerque, "A novel transfer learning based approach for pneumonia detection in chest X-ray images," *Appl. Sci.*, vol. 10, no. 2, p. 559, Jan. 2020.

[10] M. Nazir, M. A. Khan, T. Saba, and A. Rehman, "Brain tumor detection from MRI images using multi-level wavelets," in *Proc. Int. Conf. Comput. Inf. Sci. (ICCIS)*, Apr. 2019, pp. 1–5.

[11] T. Saba, S. U. Khan, N. Islam, N. Abbas, A. Rehman, N. Javaid, and A. Anjum, "Cloud-based decision support system for the detection and classification of malignant cells in breast cancer using breast cytology images," *Microsc. Res. Technique*, vol. 82, no. 6, pp. 775–785, Jun. 2019.

[12] I. M. Wani and S. Arora, "Osteoporosis diagnosis in knee X-rays by transfer learning based on convolution neural network," *Multimedia Tools Appl.*, vol. 82, no. 9, pp. 14193–14217, Apr. 2023.

[13] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, "Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, Jan. 2018.

[14] V. Pedoia, J. Lee, B. Norman, T. M. Link, and S. Majumdar, "Diagnosing osteoarthritis from T_2 maps using deep learning: An analysis of the entire osteoarthritis initiative baseline cohort," *Osteoarthritis Cartilage*, vol. 27, no. 7, pp. 1002–1010, Jul. 2019.

[15] B. Guan, F. Liu, A. H. Mizaian, S. Demehri, A. Samsonov, A. Guermazi, and R. Kijowski, "Deep learning approach to predict pain progression in knee osteoarthritis," *Skeletal Radiol.*, vol. 51, pp. 363–373, Feb. 2022.

[16] A. Tiulpin, S. Klein, S. M. A. Bierma-Zeinstra, J. Thevenot, E. Rahtu, J. V. Meurs, E. H. G. Oei, and S. Saarakkala, "Multimodal machine learning-based knee osteoarthritis progression prediction from plain radiographs and clinical data," *Sci. Rep.*, vol. 9, no. 1, p. 20038, Dec. 2019.

[17] J. Lim, J. Kim, and S. Cheon, "A deep neural network-based method for early detection of osteoarthritis using statistical data," *Int. J. Environ. Res. Public Health*, vol. 16, no. 7, p. 1281, Apr. 2019.

[18] A. Brahim, R. Jennane, R. Riad, T. Janvier, L. Khedher, H. Toumi, and E. Lespessailles, "A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis initiative," *Computerized Med. Imag. Graph.*, vol. 73, pp. 11–18, Apr. 2019.

[19] M. A. Khan, S. Kadry, P. Parwekar, R. Damasevicius, A. Mehmood, J. A. Khan, and S. R. Naqvi, "Human gait analysis for osteoarthritis prediction: A framework of deep learning and kernel extreme learning machine," *Complex Intell. Syst.*, vol. 9, pp. 2665–2683, Jan. 2021.

[20] J.-B. Schiratti, R. Dubois, P. Herent, D. Cahané, J. Dachary, T. Clozel, G. Wainrib, F. Keime-Guibert, A. Lalande, M. Pueyo, R. Guillier, C. Gabarroca, and P. Moingeon, "A deep learning method for predicting knee osteoarthritis radiographic progression from MRI," *Arthritis Res. Therapy*, vol. 23, no. 1, pp. 1–10, Dec. 2021.

[21] K. Üreten and H. H. Maras, "Automated classification of rheumatoid arthritis, osteoarthritis, and normal hand radiographs with deep learning methods," *J. Digit. Imag.*, vol. 35, no. 2, pp. 193–199, Apr. 2022.

[22] J. Song and R. Zhang, "A novel computer-assisted diagnosis method of knee osteoarthritis based on multivariate information and deep learning model," *Digit. Signal Process.*, vol. 133, Mar. 2023, Art. no. 103863.

[23] B. C. Dharmani and K. Khatri, "Deep learning for knee osteoarthritis severity stage detection using X-ray images," in *Proc. 15th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2023, pp. 78–83.

[24] W. Li, J. Feng, D. Zhu, Z. Xiao, J. Liu, Y. Fang, L. Yao, B. Qian, and S. Li, "Nomogram model based on radiomics signatures and age to assist in the diagnosis of knee osteoarthritis," *Experim. Gerontol.*, vol. 171, Jan. 2023, Art. no. 112031.

[25] C. Guida, M. Zhang, and J. Shan, "Improving knee osteoarthritis classification using multimodal intermediate fusion of X-ray, MRI, and clinical information," *Neural Comput. Appl.*, vol. 35, pp. 9763–9772, Jan. 2023.

[26] J. H. Cueva, D. Castillo, H. Espinós-Morató, D. Durán, P. Díaz, and V. Lakshminarayanan, "Detection and classification of knee osteoarthritis," *Diagnostics*, vol. 12, no. 10, p. 2362, Sep. 2022.

[27] E. M. Yuniarno and M. H. Purnomo, "Classification anterior and posterior of knee osteoarthritis X-ray images grade KL-2 using deep learning with random brightness augmentation," in *Proc. Int. Conf. Comput. Eng., Netw., Intell. Multimedia (CENIM)*, Nov. 2022, pp. 1–5.

[28] A. Sarkar, M. Z. Khan, M. M. Singh, A. Noorwali, C. Chakraborty, and S. K. Pani, "Artificial neural synchronization using nature inspired whale optimization," *IEEE Access*, vol. 9, pp. 16435–16447, 2021.

[29] I. Abunadi, "Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease," *Connection Sci.*, vol. 34, no. 1, pp. 2395–2430, Dec. 2022.

[30] M. M. Eltahir, I. Abunadi, F. N. Al-Wesabi, A. M. Hilal, A. Yousif, A. Motwakel, M. Al Duhayyim, and M. A. Hamza, "Optimal hybrid feature extraction with deep learning for COVID-19 classifications," *Comput., Mater. Continua*, vol. 71, no. 3, pp. 6257–6273, 2022.

[31] F. Kabir, "Osteoarthritis prediction | Kaggle," Tech. Rep., 2023.

[32] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, "Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction," *PLoS ONE*, vol. 17, no. 11, Nov. 2022, Art. no. e0276525.
[33] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognit. Lett.*, vol. 141, pp. 61–67, Jan. 2021.
[34] I. Abunadi and E. M. Senan, "Multi-method diagnosis of blood microscopic sample for early detection of acute lymphoblastic leukemia based on deep learning and hybrid techniques," *Sensors*, vol. 22, no. 4, p. 1629, Feb. 2022.
[35] V. Terziyan and O. Vitko, "Causality-aware convolutional neural networks for advanced image classification and generation," *Proc. Comput. Sci.*, vol. 217, pp. 495–506, Jan. 2023.
[36] A. Raza, K. Munir, and M. Almutairi, "A novel deep learning approach for deepfake image detection," *Appl. Sci.*, vol. 12, no. 19, p. 9820, Sep. 2022.
[37] S. M. Ahmed and R. J. Mstafa, "Identifying severity grading of knee osteoarthritis from X-ray images using an efficient mixture of deep learning and machine learning models," *Diagnostics*, vol. 12, no. 12, p. 2939, Nov. 2022.
[38] Y. Wang, X. Wang, T. Gao, L. Du, and W. Liu, "An automatic knee osteoarthritis diagnosis method based on deep learning: Data from the osteoarthritis initiative," *J. Healthcare Eng.*, vol. 2021, pp. 1–10, Sep. 2021.
[39] A. D. Goswami, "Automatic classification of the severity of knee osteoarthritis using enhanced image sharpening and CNN," *Appl. Sci.*, vol. 13, no. 3, p. 1658, Jan. 2023.
[40] C. Ntakolia, C. Kokkotis, S. Moustakidis, and D. Tsaopoulos, "A machine learning pipeline for predicting joint space narrowing in knee osteoarthritis patients," in *Proc. IEEE 20th Int. Conf. Bioinf. Bioengineering (BIBE)*, Oct. 2020, pp. 934–941.
[41] M. Y. Kardawi, K. Muchtar, Y. Nurdin, N. Maulina, and S. Muchallil, "A comparative analysis of deep learning models for detection of knee osteoarthritis disease through mobile apps," in *Proc. 8th Int. Conf. Wireless Telematics (ICWT)*, Jul. 2022, pp. 1–6.
[42] D. A. Zebari, S. S. Sadiq, and D. M. Sulaiman, "Knee osteoarthritis detection using deep feature based on convolutional neural network," in *Proc. Int. Conf. Comput. Sci. Softw. Eng. (CSASE)*, Mar. 2022, pp. 259–264.

AMJAD REHMAN (Senior Member, IEEE) received the Ph.D. degree from the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia, specializing in information security using image processing techniques, in 2010. He is currently an Assistant Professor with the CCIS, Prince Sultan University, Riyadh, Saudi Arabia. He is also a PI in several funded projects and also completed projects funded by MoHE Malaysia, Saudi Arabia. His current research interests include health informatics, big data mining, and machine learning. He received the Rector Award for the 2010 Best Student from UTM Malaysia.

ALI RAZA received the Bachelor of Science and M.S. degrees in computer science from the Department of Computer Science, Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan, in 2021 and 2023, respectively. His current research interests include data science, artificial intelligence, data mining, natural language processing, machine learning, deep learning, and image processing.

FATEN S. ALAMRI received the Ph.D. degree in system modeling and analysis in statistics from Virginia Commonwealth University, USA, in 2020. Her Ph.D. research included Bayesian dose-response modeling, experimental design, and nonparametric modeling. She is currently an Assistant Professor with the Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University. Her research interests include spatial area, environmental statistics, and brain imaging.

BAYAN ALGHOFAILY received the master's and Ph.D. degrees in computer science from Toronto Metropolitan University, Toronto, Canada. During that period, she was a member of the Distributed Applications and Broadband Networks Laboratory (DABNEL). She focused on studying how the performance of machine learning models is affected by dataset features. She is currently an Assistant Professor with the Department of Information System, CCIS, Prince Sultan University (PSU). She is also a member of the Artificial Intelligence and Data Analytics (AIDA) Laboratory, CCIS, PSU. Her research interests include AI, NLP, ML, and neural networks. She continues to explore this further in her research.

TANZILA SABA (Senior Member, IEEE) received the Ph.D. degree in document information security and management from the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia, in 2012. She is currently a Full Professor with the College of Computer and Information Sciences, Prince Sultan University (PSU), Riyadh, Saudi Arabia, and also the Leader of the AIDA Laboratory. She has published over 300 publications in high-ranked journals. Her primary research interests include bioinformatics, data mining, and classification using AI models. She received the Best Student Award from the Faculty of Computing, UTM, in 2012. She received the Best Research of the Year Award from PSU, from 2013 to 2016. Due to her excellent research achievement, she is included in Marquis Who's Who (S & T) 2012. She is the editor of several reputed journals and on a panel of TPC of international conferences.