Automatic Detection of Knee Osteoarthritis Disease with the Developed CNN, NCA and SVM Based Hybrid Model

1 author:

Serpil Aslan
Malatya Turgut Özal Üniversitesi
31 PUBLICATIONS **143** CITATIONS

Automatic Detection of Knee Osteoarthritis Disease with the Developed CNN, NCA and SVM Based Hybrid Model

Serpil Aslan

Department of Software Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, Malatya 44210, Turkey

Corresponding Author Email: serpil.aslan@ozal.edu.tr

ABSTRACT

Knee osteoarthritis (Knee-OA) is one of the most common musculoskeletal diseases caused by loss of cartilage and bone changes in the joint. Prediction of early Knee-OA based on early bone tissue analysis is challenging in medical image analysis. If the disease is detected in the later stages, it may cause serious problems, such as the need for knee replacement. Therefore, the detection of Knee-OA disease is essential. With the developing technology, computer-aided systems have been frequently used in the biomedical field in recent years. A deep learning-based hybrid model for the early diagnosis and treatment of Knee-OA disease was developed in this study. In the developed hybrid model, three different CNN architectures were used as the base, and feature extraction was made with these architectures. The features obtained in three different architectures are combined to bring together different features of the same image. After merging, the neighboring component analysis (NCA) size reduction method was used to remove unnecessary features. Since unnecessary features are eliminated from the feature map optimized with NCA, the proposed hybrid model will work faster and produce more successful results. Finally, the feature map optimized with NCA was classified with six different classifiers. The proposed model was also compared to eight different CNN architectures. In comparison to CNN architectures, the proposed hybrid model achieved the highest accuracy performance.

1. INTRODUCTION

Arthritis is an inflammatory condition that develops in the joints against the body's tissue or due to external factors (microorganisms, trauma, etc.). There are two common types of arthritis disease: Osteoarthritis (OA) and Rheumatoid Arthritis (RA). OA, also called joint weakness, is one of the degenerative joint diseases characterized by cell stress and cartilage extracellular matrix degradation due to maladaptive repair responses set in motion by micro and macro traumas [1]. It begins to be seen more frequently in young individuals with advancing life spans, excessive joint use or trauma, family history, external injuries, and obesity. Cartilage loss and inflammation are seen in the cartilage. It is characterized by deterioration, wear and tear in the articular cartilage. As a result, changes may occur in the bone tissue under the articular cartilage. Healthy cartilage prevents the bone from sliding easily within the joint and rubbing against each other [2]. In OA, the top layer of the bone is broken down and eroded, causing the bones to rub against each other. This causes limitation of movement and severe pain [3, 4]. OA is most common in the feet, knees, lower back, hips, and toes [5]. Pain, stiffness, and functional disability are the main clinical features of osteoarthritis. Pain is the most important symptom. Pain that occurs with the use of the involved joint and is relieved by rest is typical. After clinically observing the patient, medical professionals may request a radiographic evaluation (X-Ray), CT (Computed Tomography), or MRI (Magnetic Resonance Imaging). The progression of the disease is prevented by the treatment initiated in patients diagnosed early with biological imaging techniques.

Various parameters are examined to analyze the patient's joint space width in the radiographic images examined in detecting Knee-OA disease [6, 7]. There is no strictly defined grading system for the diagnosis of OA. This situation may vary depending on the subjectivity of the practitioner. Against all these situations, a categorical rating scale called the Kellgren & Lawrence (KL) [8] scale is used to rate the severity of OA through radiographic evaluation, which is widely used. In the KL grading system, the disease is systematized into 5 categorical grades: G-0 (Normal OA), G-1 (Doubtful OA), G-2 (Mild OA), G-3 (Moderate OA), and G-4 (Severe OA). In the framework of KL, which is widely used in clinical settings to make treatment decisions, one joint 0 represents normal, and 4 represents severe OA. Calculated cartilage gap width can assist medical professionals in early diagnosis [9].

Diagnosing the disease in later stages can lead to joint necrosis and disability. Because the symptoms included in the OA classification are continuous, individual ratings are time-consuming, subjective, and prone to error. As a result, medical professionals' assessments of OA may differ from one another.

In addition, the semi-quantitative KL rating scale may cause uncertainty when evaluating medical professionals. These uncertainties will make early diagnosis difficult. A substantial amount of knowledge and experience is required for a valid diagnosis of OA. As a result, OA assessment can be quantified when a low-cost and non-subjective image-based CAD system for Knee-OA is developed. Such a diagnostic system can also

be used for clinical studies such as evaluating drugs that impact the progression of OA, intra-articular injections, or surgical interventions. On the other hand, MRI images of the relevant region may be requested in cases where the findings in the images obtained by X-ray do not provide clear information about joint pain [10]. This will be costly in terms of both raw material and time.

1.1 Motivation and contribution

OA is a joint disease characterized by cartilage loss and bone changes. The most common type of OA is knee OA. In medical image analysis, predicting early knee OA based on early bone tissue analysis is difficult. In recent years, many studies have been conducted to analyze Knee-OA detection and progression using deep learning-based techniques [11]. In this study, a new deep learning-based hybrid model was proposed using knee X-ray images to detect Knee-OA at an early stage and to categorize it according to the KL grading system. A deep learning-based CAD model that can be used by medical professionals as an objective tool in the diagnosis of OA to support their decisions and prevent human misconceptions will prevent the delay in the examination of radiographic images. At the same time, experts will be able to focus more on rare findings than on common findings. The proposed model will significantly accelerate the diagnosis of Knee-OA from radiographs together with clinical evaluations.

The proposed model consists of three stages. The first step is the deep feature extraction stage, in which deep hybrid features are extracted. At this step, feature maps are extracted using the pre-trained DenseNet201 [12], DarkNet53 [13], and ShuffleNet [14] CNN architectures. Then, the obtained feature maps from the three architectures used as feature generators are concatenated. This way, feature maps with higher-quality information were obtained using the advantages of DenseNet201, DarkNet53, and ShuffleNet together. The second step is the feature selection step; in other words, it is the size reduction step, in which the NCA [15] optimization method is applied to reduce the size of the features obtained from the CNN architectures used as feature generators. Thanks to this step, the most informative features were selected, and the features with low information quality were removed from the feature maps, thus reducing the model's runtime. The third step is the classification process through machine learning classifiers using optimized features. Experimental results prove the success of the proposed model.

1.2 Organization of paper

The article is organized as follows: Section 2 presents a literature review of relevant studies in this field. The dataset, methods, and technical terminologies used in the proposed model are all described in Section 3. Section 4 compares the experimental results of our method and other CNN architectures. Section 5 concludes the study.

2. RELATED WORKS

OA is a major source of pain and disease that imposes disability, poor quality of life, and a high financial burden on the health system [16, 17]. It is estimated that close to 240 million people worldwide suffer from OA [18]. The knee joint, which has three compartments (medial tibiofemoral, lateral

tibiofemoral, and patellofemoral) and is one of the main weight-bearing joints, is the most common site of OA [1, 19, 20]. People with symptomatic Knee-OA may experience problems that affect their activities of daily living, such as knee pain, joint stiffness, swelling, and physical disability [21]. Since these symptoms can be seen in heterogeneous patterns, it indicates that Knee-OA is a common disorder rather than a simple cartilage problem [1]. OA reduces mobility, quality of life, and productivity while increasing morbidity, healthcare use, and social expenditure [1]. OA creates a significant individual and societal burden [22].

A common cause of morbidity worldwide, Knee-OA can be diagnosed by radiography, which can help determine which patients might benefit from surgery. In recent years, advances in OA diagnosis and grading have been made using reported data from X-ray, MRI, and CT scans of knee-OA [23]. With the use of deep learning and machine learning techniques in medical image processing, more successful methods have been proposed in studies such as the detection of bone, cartilage, meniscus, etc. tissues related to OA, multiple automatic or semi-automatic segmentation and automatic scoring [23, 24]. Shamir et al. [25] proposed a CAD model for the early detection of OA from radiography images. The proposed model, it was used to analyze the difference between KL-0 (Normal OA) and KL-2 (Moderate OA) using tissue and density information in knee joint images. Janvier et al. [26] proposed a fractal tissue analysis method, the DRAE model, to analyze the tissues of the trabecular bone (TB) in radiographic images to predict the progression of Knee-OA. The relevant regions were extracted using the semi-segmentation method first, and then fractal texture analysis was performed using various methods. Experimental results prove that analyzing TB bone structure provides significant success in detecting OA progression. Brahim et al. [27] developed a machine learning-based CAD system to detect Knee-OA early from X-ray images. In the proposed model, radiography images were first preprocessed using the Fourier filter. After preprocessing, the normalization method was applied using the multivariate linear regression method. After applying the independent component analysis method for dimension reduction for feature selection and extraction operations, it is given to machine learning approaches. Riad et al. [28] proposed a method for analyzing OA tissue from knee X-rays based on complex wavelet decomposition. Relative phases and complex coefficients are extracted using wavelet decomposition from preprocessed Knee-OA radiography images. The obtained parameters are used for OA classification and analysis. Antony et al. [29, 30] used a CNN-based method to classify various stages of Knee-OA severity and a Fully Convolutional Network (FCN)-based method to localize the knee joints. Gornale et al. [31] proposed a wavelet filtering-based method for early detection of OA on radiographic images and classification according to the KL grading system. In the proposed method, the cartilage region is automatically detected according to the density and spacing of the wavelet filters in the radiography images. It is then classified through machine learning classifiers (Decision Tree, KNN). Tiulpin et al. [7] proposed a Deep Siamese CNN-based approach for the automatic diagnosis of OA and automatic classification of Knee-OA severity with the KL rating system. Kotti et al. [32] developed a body kinetics-based CAD system and proposed an automated Knee-OA detection system. A dataset of 94 subjects was used to test the proposed system. The authors not only detected the presence of Knee-OA in

their study but also developed a system that generates the specific parameters used in making this decision. Astuto et al. [33] a 3D convolutional neural network model to detect OA abnormalities automatically. The authors extracted 3 ROIs from knee MRI images to reduce dimensionality before multiple 3D CNN preprocessing to classify knee lesions in their study. Raj et al. [34] proposed a new model for knee cartilage segmentation that employs a new 3D CNN called '-Net' in conjunction with a multi-class loss function.

In the literature studies, it has been observed that the methods developed for diagnosis and progression of Knee-OA focus on a single model, and the analysis of hybrid models needs to be adequately examined. As a result, this study proposed a deep learning-based hybrid model that uses knee X-ray images to detect Knee-OA early and classify it using the KL grading system.

3. MATERIALS AND METHODS

In the section of the study, the dataset used, the pre-processes used to design the proposed model, and the proposed model are all explained.

3.1 Dataset

A dataset of publicly available digital knee X-ray images [35] collected from various health centers was used in this study to test the performance of the proposed model. The data set consists of two sub-image folders, "MedicalExpert-I and MedicalExpert-II", labeled with 5 class labels according to the KL grading system for Knee-OA severity by two medical professionals. The data is organized into two subfolders, each with the same image but different tags. The MedicalExpert-I subfolder was used in the study to evaluate the performance of the proposed model. The dataset consists of 1650 images. Figure 1 represents the samples of each data set class labeled according to the KL grading system.
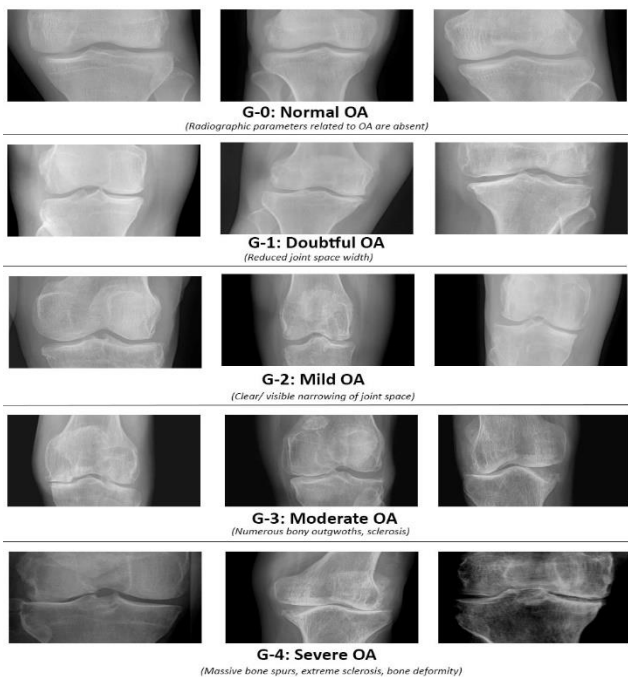

Figure 1. Samples of each class of "KL Grading System" in used dataset

Each image in the dataset is converted to a colored format before applying the proposed method. Class distributions of our dataset labeled by the medical professional according to the KL grading system are shown in Figure 2. For experimental analysis, the dataset is divided into 80% training and 20% testing.
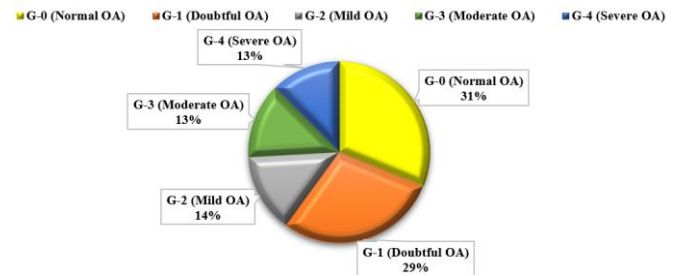

Figure 2. The OA dataset's class distributions

3.2 Background of the proposed hybrid model

CAD systems are vital in medical image processing. Many studies based on deep learning architectures have been proposed. Extracting features with high information potential from each image examined while processing medical images is a very important process for the model's performance. Machine learning-based approaches require a feature selection process, while deep learning-based techniques can be extracted directly from image content without human assistance in feature selections [36]. Therefore, thanks to the use of deep learning architectures, the model can be used effectively without any expert knowledge. CNN structure is a deep learning algorithm that extracts important information from images. It works like convolution, pooling, and a series of sequential layers called ReLu. It is very advantageous for feature extraction in large-scale datasets. For all these reasons, in this study, 8 CNN architectures, commonly used feature generators, were used for feature extraction from the model in which we trained the classes with our predetermined training dataset. The CNN architectures used are: Vgg19 [37], MobileNetV2 [38], ShuffleNet [14], ResNet101 [39], DenseNet201 [12], AlexNet [40], GoogleNet [41] and DarkNet53 [13] are deep feature map generators. 1650 x 1000 features were obtained from each architecture. The best of the extracted key features should be chosen to produce better classification results. To use the architectures used more efficiently in terms of both runtime and accuracy performance, the features with the most information value should be examined instead of examining all the features. Features with high information potential were kept, and redundant features were extracted from feature maps using the NCA dimension reduction method based on NCA. As a result of NCA, the number of features was reduced from 1000 to 600. Feature maps optimized by eight feature generator CNN architectures are then sent to six different machine-learning classifiers.

Figure 3 represents the architecture of the operations carried out to create the proposed model. All of the processes described in this section are used to select the best feature generator CNN architectures and classifiers for use in the hybrid deep learning model.
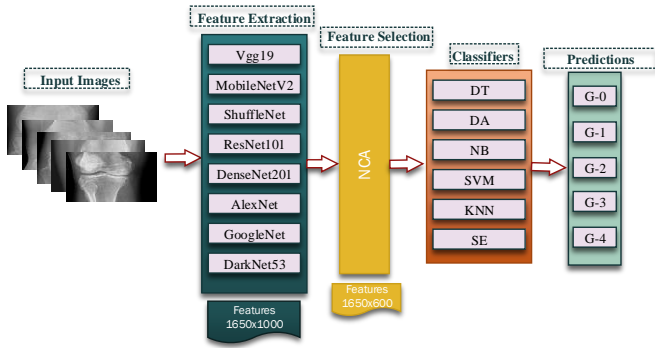
Figure 3. The background architecture of the model

3.3 Proposed hybrid model

A new deep learning-based hybrid model was proposed in this study to detect OA early on and classify it using the KL grading system using knee X-ray images. A deep learning-based CAD model that can be used by medical professionals as an objective tool in the diagnosis of OA to support their decisions and prevent human misconceptions will prevent the delay in the examination of radiographic images. At the same time, experts will be able to focus more on rare findings than on common findings. Thanks to the deep learning architectures used in the study, a hybrid model with maximum classification capability is presented.

Figure 4 represents the basic steps of the proposed model. The proposed model consists of three basic steps: Hybrid Deep Feature Extraction, Feature Selection with NCA, and Classification.
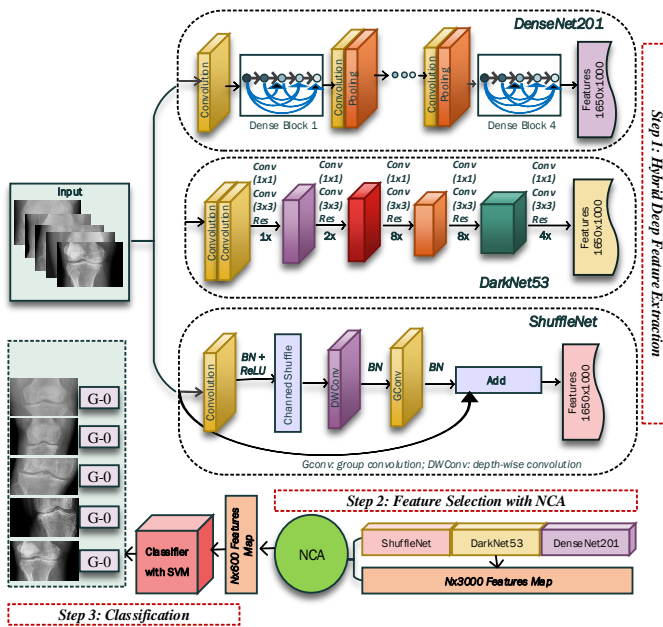

Figure 4. The basic architecture of the proposed model

Step 1: Hybrid Deep Feature Extraction

Hybrid Deep Feature Extraction is the most important step in which the main features of the proposed model are extracted. This step used pre-trained "DenseNet201, DarkNet53, and ShuffleNet"-supervised CNN feature generator architectures as the base.

DenseNet201 has a 201-layer deep model in which each feed-forward layer is merged into another layer. Each layer uses the feature maps from the previous layers as input, while the new feature maps produced as output are used as input for all subsequent layers. This architecture has advantages such as reducing the vanishing gradient problem, strengthening the feature layer, and reducing the number of bursts.

DarkNet53 is a CNN architecture that has been pre-trained on ImageNet [42]. DarkNet53 has a deep model with 53 layers consists of convolution layers in 1x1 and 3x3 dimensions. Each convolution layer is followed by a batch normalization layer and a LeakyReLU layer.

ShuffleNet is a computationally efficient CNN architecture pre-trained on ImageNet [42]. Compared to other CNN architectures, ShuffleNet has less complexity and fewer parameters. This architecture employs new operations such as group convolution, depth-wise convolution, and channeled shuffle to maintain accuracy while reducing the computational cost.

Each pre-trained architecture produces a feature map with a size of 1650x1000. These three architectures' feature maps are then concatenated to create a new 1650x3000 feature map. Three distinct features of the same image are extracted in this manner. In hybrid architectures, another architecture has a high potential to detect a feature that one model misses. This is the most important advantage of using strong architectures together. This step produces a feature map with a higher information potential. This advantage will have a significant impact on the proposed model's performance. In addition to all these advantages, using a high-dimensional feature map will increase the number of unnecessary features in its content, which will increase the analysis time of the proposed model. All these reasons lead to the need to optimize the new hybrid feature map obtained.

In order to make the proposed model work faster, after the feature merging step, unnecessary features are eliminated by using the NCA method. The size of the feature map, which was 1650 x 3000, became 1650 x 600 at this step. This feature count is lower than the 1650 x1000 feature map obtained in pre-trained models. Finally, the optimized feature map was classified in the SVM classifier and high-performance values were achieved.

Step 2: Feature Selection with NCA

NCA is one of the supervised feature selection methods used in recent years. NCA is a feature selection approach developed using the KNN algorithm. The NCA learns a feature weighting vector by maximizing the expected target classification accuracy with the normalization term. The primary benefit is that no information is lost during the NCA size reduction process, and it generates positive weights for each feature [43, 44].

Feature Selection with NCA is the step in which feature maps obtained from hybrid deep feature extraction are optimized. The best of the extracted key features should be chosen to achieve faster and higher performance classification results. The proposed model preserves high quality features while removing redundant features from feature maps using the NCA dimension reduction method. This yields a hybrid feature map with a high information potential. The feature map becomes 1650 x 600 after the NCA size was reduced. The number of features (600) used after optimization is less than the number of features (1000) obtained from an architecture.

Step 3: Classification

Classification is the step in which KL classes are detected

using machine learning classifiers of the optimized hybrid feature map. Because the SVM classification algorithm achieved the highest success in the proposed model's background design, it was chosen as the classifier in the proposed new model. Here, the classifier determines which of the 5 classes, G-0, G-1, G-2, G-3, and G-4, is closer to each other. During the design of the classification process, experimental results were obtained using six different algorithms in order to determine the best classifier algorithm.

4. EXPERIMENTAL RESULTS

During the design of the proposed model, the experimental results of 8 different CNN feature generator architectures and the proposed model are examined to determine the highest performing feature map generator. In the experiment, 80% of the image data in the MedicalExpert-I subfolder of the digital knee X-ray images dataset collected from various health centers was used for training, while the remainder was used for testing. Table 1 shows the training parameters that were used in the experimental results.

Table 1. Training Hyperparameter settings

Hyperparameter	
MiniBatchSize	16
Max Epochs	7
LearnRate	0.0001
ValidationFrequency	9

In this study, confusion matrices were used to evaluate the performance of the proposed model and other CNN architectures. 8 different evaluation metrics obtained using confusion matrices: "F1-score (F1), Accuracy (Acc), Specificity (Sp), Sensitivity (Se), False Discovery Rate (FDR), False Positive Rate (FPR), and False Negative Comprehensive evaluation has been made using Rate (FNR)". The calculation formulas of performance metrics are provided in Table 2.

Table 2. The calculation formulas of performance metrics

Measure	Formula
Accuracy (Acc)	(TP+TN)/(TP+TN+FN+FP)
Sensitivity (Se)	TP/(TP+FN)
Specificity (Sp)	TN/(FP+TN)
False Positive Rate (FPR)	FP/(FP+TN)
False Discovery Rate (FDR)	FP/(FP+TP)
False Negative Rate (FNR)	FN/(FN+TP)
F1-Score (F1)	F1=2TP/(2TP+FP+FN)

where the number of predictions that the TP classifier correctly recognizes as belonging to the class, the number of incorrect predictions that the FP classifier incorrectly assigned to the class, the number of predictions that do not belong to the class that the TN classifier correctly recognizes, and the number of incorrect predictions that the FN classifier does not recognize as class samples [45].

To compare the proposed model's performance metrics, 8 different CNN architectures are used. Figure 5 represents the confusion matrices of these models. Figure 5 demonstrates the correct and incorrect predictions of each architecture on 329 (20%) test images. Here, the highest performing DarkNet53 architecture predicted 233 correctly and 96 incorrectly out of 329 test images. Table 3 represents the accuracy values obtained in CNN architectures. As can be seen Table 3, the

highest accuracy value is 70.82% in the DarkNet53 architecture and the lowest accuracy value is 49.54% in the ShuffleNet architecture.
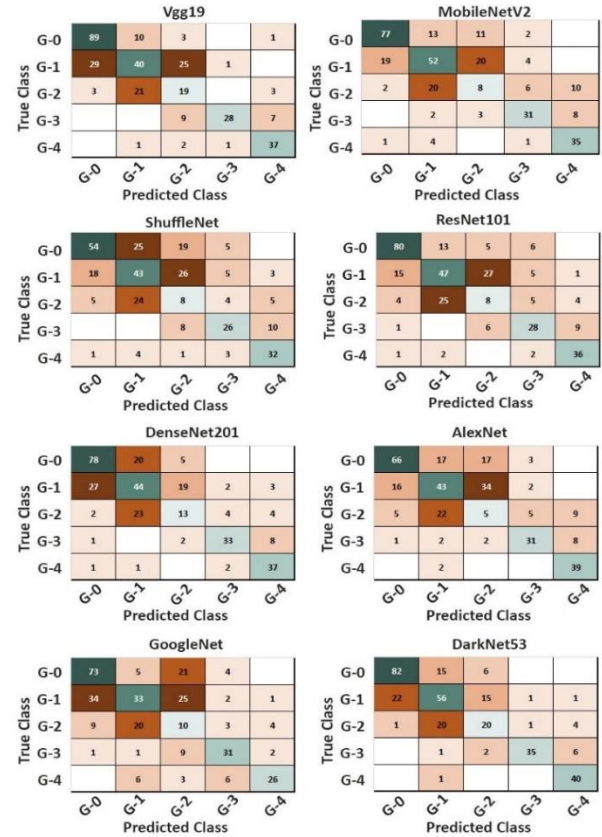

Figure 5. Confusion matrix of CNN architectures

Features were extracted from each of the eight different architectures used in the study to determine which should be used as a base. The obtained features were used in six different classifiers after being optimized with the NCA method. Then, using the layers given in Table 4 of the CNN architectures, feature maps were obtained for each architecture.

The number of features obtained from each of the 8 different architectures is 1650 x 1000. For these features to work faster and more effectively, unnecessary features are eliminated by using the NCA size reduction method. As a result, the size of the feature map has been reduced from 1650x1000 with NCA to 1650x600. After the features obtained from 8 different CNN architectures were optimized with the NCA method; Decision Tree (DT) [46], Discriminant Analysis (DA) [47], Naive Bayes (NB) [48], Support Vector Machine (SVM) [49], K-Nearest Neighbors (KNN) [50], and Ensemble Subspace (ES) [51] were classified using machine learning classifiers. The details of this stage are shown in Figure 4. Then, the performance of each architecture according to the accuracy evaluation metric in 6 different classifiers is given in Table 5. As seen in Table 5, the highest performance was obtained in SVM. The SVM machine learning algorithm was chosen as the classifier in the proposed model as a result of its success.

In addition, as seen in Table 5, all of the compared CNN architectures performed the highest-performing in the SVM classifier. Among these architectures, the highest success rate was achieved with DenseNet201, DarkNet53, and ShuffleNet in the SVM classifier, with 79.3%, 78.1%, and 77.9%, respectively. For this reason, DenseNet201, DarkNet53, and ShuffleNet machine learning models were selected for the

proposed model's Hybrid Deep Feature Extraction stage. As a result, confusion matrices in the NCA+SVM classifier of each architecture are as given in Figure 6.
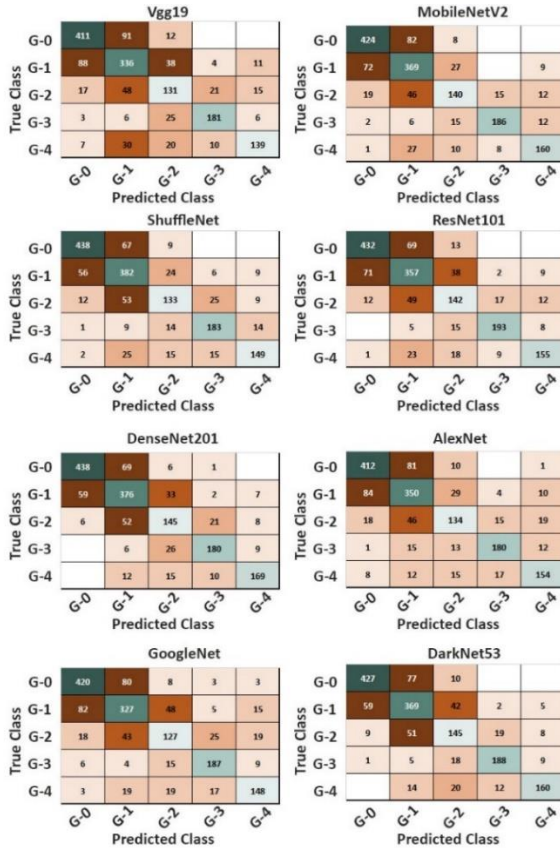

Figure 6. Confusion matrixes of CNN architectures+NCA+SVM

Our three-step model was created after we have collected all of the experimental results from the proposed model's design phase. The details of the proposed model are represented in Figure 4. Each 1650x1000 feature map generated by DenseNet201, DarkNet53, and ShuffleNet feature generators is concatenated in the proposed model. The feature map obtained after concatenating is reduced to 1650x3000 size.

The 1650x600 feature map obtained through the NCA optimization method was then sent to the SVM classifier, and the classification process for the detection of Knee-OA was carried out. Figure 7 represents the proposed model's confusion matrix. When compared to the confusion matrix of other CNN architectures used in the performance comparison, the proposed model clearly outperforms them.
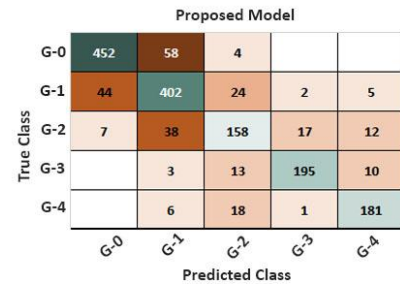

Figure 7. The confusion matrix of proposed model ((DenseNet201+DarkNet53+ShuffleNet)+NCA+SVM)

The proposed model's performance was evaluated using multiple performance evaluation metric. Table 6 shows the obtained values. As shown in Table 6, the proposed method achieved the highest accuracy performance in the G-3 class. While the success of the proposed model in classes with distinctive features such as G-0 (87.93%) and G-4 (87.86) is expected, the highest success rate of 88.23% in the G-3 class proves the success of the proposed model. The proposed hybrid model correctly classified 452 Normal OA images, 402 Doubtful OA images, 158 Mild OA images, 195 Moderate OA images, and 181 Severe OA images. Out of 1650 Knee-OA images, it predicted 1388 correctly and incorrectly predicted 262. In addition, when Table 5 and Table 6 are compared, the highest success among 8 CNN architectures is achieved in DenseNet201 architecture with 79.3%. Furthermore, the proposed model's accuracy performance in the SVM classifier is 84.12%.

Figure 8 shows the AUC (Area Under Curve) / ROC (Receiver Operating Characteristic) curves for each class in the proposed model.

Table 3. Accuracy rates of CNN architectures (%)

Vgg19	MobileNetV2	ShuffleNet	ResNet101	DenseNet201	AlexNet	GoogleNet	DarkNet53
64.74	61.70	49.54	60.49	62.31	55.93	52.58	**70.82**

Table 4. Layers used for feature extraction

Vgg19	MobileNetV2	ShuffleNet	ResNet101	DenseNet201	AlexNet	GoogleNet	DarkNet53
Fc8	Logits	node_202	Fc1000	Fc1000	Fc8	loss3-classifier	Conv53

Table 5. The accuracy performance comparisons with CNN architectures + NCA + Classifiers

CNN Models	Accuracy Rate of Classifiers					
	DT	DA	NB	SVM	KNN	ES
Vgg19	49.2	70.7	51.8	**72.6**	68.9	70.3
MobileNetV2	52.2	71.3	59.7	**77.5**	75.5	75.8
ShuffleNet	53.9	75	55	**77.9**	73.1	73.8
ResNet101	56.2	74.5	56.5	**77.5**	74.5	74.1
DenseNet201	56.3	73.4	58.6	**79.3**	74.1	74.8
AlexNet	51.2	71.1	51	**74.5**	71.6	72.2
GoogleNet	51.8	69.8	51.8	**73.3**	69.6	70.4
DarkNet53	55.7	75.5	62.7	**78.1**	75.6	75.6

Table 6. The proposed model's performance metrics

Classes	Accuracy	Sensitivity	Specificity	FPR	FDR	FNR	F1
G-0	87.93	93.38	94.68	5.31	12.06	6.61	90.58
G-1	84.27	79.28	93.43	6.56	15.72	20.71	81.70
G-2	68.10	72.81	94.83	5.16	31.89	27.18	70.37
G-3	88.23	90.69	98.18	1.81	11.76	9.30	89.44
G-4	87.86	87.01	98.26	1.73	12.13	12.98	87.43

Figure 8. The AUC / ROC curves for each class in the proposed model. (a: G-0 (Normal OA), b: G-1 (Doubtful OA), c: G-2 (Mild OA), d: G-3 (Moderate OA) and e: G-4 (Severe OA))

5. DISCUSSIONS

This study used a public dataset of 1650 knee X-ray images pre-labeled by a medical professional according to the knee OA KL grading system for experiments. The proposed method's primary goal is to present a preliminary diagnosis system to medical professionals. For this purpose, a comprehensive analysis was carried out to select the features with the highest information value from the analyzed images. The performance of eight different CNN architectures was evaluated using experimental results. Knee X-ray images may show geometric distortions in the cartilage shadow due to the progression of OA. In such cases, extracting the features of important regions in the knee region may be difficult. For this reason, two steps, deep feature extraction, and feature selection, were applied to extract the features with the highest information content in the study. This way, selected pre-trained DenseNet201, DarkNet53, and ShuffleNet feature maps were extracted and combined, then optimized with the NCA algorithm. In the final step, it was classified using the SVM classifier algorithm. This method is a multi-level and flexible method. From this point of view, although this method produces low, medium, and high-level features, the most distinctive of these features can be selected.

Figure 9. Comparative analysis of proposed model, DenseNet201, DarkNet53 and ShuffleNet with SVM classifier and Medical Expert-I opinion

The success comparisons of the proposed model and three other architectures with the SVM classifier on the MedicalExpert-I database are shown in Figure 9. As shown in Figure 9, the closest estimates to expert opinion for the five class labels were estimated alongside the proposed model.

An accuracy of 84.12% was achieved using the proposed method and 1650 knee X-ray images. The success of the proposed model is more competitive and promising compared to existing architectures in the literature.

6. CONCLUSIONS

Orthopedists use various imaging techniques to diagnose knee OA as part of their routine practice. These diseases are widespread throughout the world. Intelligent medical assistants are needed to speed up and automate this process. A new hybrid three-step model was proposed in this study to detect Knee-OA at an early stage. It consists of feature extraction, feature selection, and automatic Knee-OA classification steps. The experimental results in this study were run on a publicly available dataset of digital knee images collected from various health centers. Knee-OA disease is a serious condition that psychologically impacts individuals, patients, and their families. The proposed model will significantly speed up the diagnosis of Knee-OA using radiographs and clinical evaluations. The proposed model performed with an accuracy of 84.12%. Experimental results prove the success of the proposed model. A deep learning-based CAD model that can be used by medical professionals as an objective tool in the diagnosis of OA to support their decisions and prevent human misconceptions will prevent the delay in the examination of radiographic images. At the same time, experts will be able to focus more on rare findings than on common findings.

REFERENCES

[1] Teoh, Y.X., Lai, K.W., Usman, J., Goh, S.L., Mohafez, H., Hasikin, K., Qian, P.J., Jiang, Y.Z., Zhang, Y., Dhanalakshmi, S. (2022). Discovering knee osteoarthritis imaging features for diagnosis and prognosis: review of manual imaging grading and machine learning approaches. Journal of Healthcare Engineering, Article ID: 4138666. https://doi.org/10.1155/2022/4138666

[2] Bindushree, R., Kubakaddi, S., Urs, N. (2015). Detection of knee osteoarthritis by measuring the joint space width in knee X ray images. International Journal of Electronics & Communication, 3(4): 18-21.

[3] Galvan-Tejada, J.I., Treviño, V., Celaya-Padilla, J.M., Tamez-Pena, J.G. (2014). Knee osteoarthritis pain prediction from X-ray imaging: Data from osteoarthritis Initiative. In 2014 International Conference on Electronics, Communications and Computers (CONIELECOMP), pp. 194-199. https://doi.org/10.1109/CONIELECOMP.2014.6808590

[4] Deokar, D.D., Patil, C.G. (2015). Effective feature extraction based automatic knee osteoarthritis detection and classification using neural network. International Journal of Engineering and Techniques, 1(3): 134-139.

[5] Minciullo, L., Cootes, T. (2016). Fully automated shape analysis for detection of Osteoarthritis from lateral knee radiographs. In 2016 23rd international conference on pattern recognition (ICPR), pp. 3787-3791. https://doi.org/10.1109/ICPR.2016.7900224

[6] Gornale, S.S., Patravali, P.U., Uppin, A.M., Hiremath, P.S. (2019). Study of segmentation techniques for assessment of osteoarthritis in knee X-ray images. International Journal of Image, Graphics and Signal Processing (IJIGSP), 11(2): 48-57. https://doi.org/10.5815/ijigsp.2019.02.06

[7] Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., Saarakkala, S. (2018). Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach. Scientific Reports, 8(1): 1-10. https://doi.org/10.1038/s41598-018-20132-7

[8] Emrani, P.S., Katz, J.N., Kessler, C.L., Reichmann, W.M., Wright, E.A., McAlindon, T.E., Losina, E. (2008). Joint space narrowing and Kellgren–Lawrence progression in knee osteoarthritis: An analytic literature synthesis. Osteoarthritis and Cartilage, 16(8): 873-882. https://doi.org/10.1016/j.joca.2007.12.004

[9] Kellgren, J.H., Lawrence, J. (1957). Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases, 16(4): 494. https://doi.org/10.1136/ard.16.4.494

[10] Gornale, S.S., Patravali, P.U., Manza, R.R. (2016). A survey on exploration and classification of osteoarthritis using image processing techniques. International Journal of Scientific & Engineering Research, 7(6): 334-355.

[11] Abdullah, S.S., Rajasekaran, M.P. (2022). Automatic detection and classification of knee osteoarthritis using deep learning approach. La radiologia medica, 127(4): 398-406. https://doi.org/110.1007/s11547-022-01476-7

[12] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700-4708.

[13] Yildirim, K., Yildirim, M., Eryesil, H., Talo, M., Yildirim, O., Karabatak, M., Ogras, M.S., Artas, H., Acharya, U.R. (2022). Deep learning-based PI-RADS score estimation to detect prostate cancer using multiparametric magnetic resonance imaging. Computers and Electrical Engineering, 102: 108275. https://doi.org/10.1016/j.compeleceng.2022.108275

[14] Rezaee, K., Mousavirad, S.J., Khosravi, M.R., Moghimi, M.K., Heidari, M. (2021). An autonomous UAV-assisted distance-aware crowd sensing platform using deep ShuffleNet transfer learning. IEEE Transactions on Intelligent Transportation Systems, 23(7): 9404-9413. https://doi.org/10.1109/TITS.2021.3119855

[15] Goldberger, J., Hinton, G.E., Roweis, S., Salakhutdinov, R.R. (2004). Neighbourhood components analysis. Advances in Neural Information Processing Systems, 17.

[16] Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., Kaur, M. (2021). Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. Journal of Biomolecular Structure and Dynamics, 39(15): 5682-5689. https://doi.org/10.1080/07391102.2020.1788642

[17] Eroglu, Y., Yildirim, K., Çinar, A., Yildirim, M. (2021). Diagnosis and grading of vesicoureteral reflux on voiding cystourethrography images in children using a deep hybrid model. Computer Methods and Programs in Biomedicine, 210: 106369. https://doi.org/10.1016/j.cmpb.2021.106369

[18] Ribas, L.C., Riad, R., Jennane, R., Bruno, O.M. (2022). A complex network based approach for knee Osteoarthritis detection: Data from the Osteoarthritis initiative. Biomedical Signal Processing and Control, 71: 103133. https://doi.org/10.1016/j.bspc.2021.103133

[19] Sethy, P.K., Behera, S.K. (2020). Detection of coronavirus disease (covid-19) based on deep features. Preprints 2020, 2020030300. https://doi.org/10.20944/preprints202003.0300.v1

[20] Koli, J., Multanen, J., Kujala, U.M., Häkkinen, A., Nieminen, M.T., Kautiainen, H., Lammentausta, E., Jämsä, T., Ahola, R., Selänne, H., Kiviranta, I., Heinonen, A. (2015). Effects of exercise on patellar cartilage in women with mild knee osteoarthritis. Med Sci Sports Exerc, 47(9): 1767-74. https://doi.org/10.1249/mss.0000000000000629

[21] Van Spil, W.E., Kubassova, O., Boesen, M., Bay-Jensen, A.C., Mobasheri, A. (2019). Osteoarthritis phenotypes and novel therapeutic targets. Biochemical Pharmacology, 165: 41-48. https://doi.org/10.1016/j.bcp.2019.02.037

[22] Ölmez, E., Akdoğan, V., Korkmaz, M., Er, O. (2020). Automatic segmentation of meniscus in multispectral MRI using regions with convolutional neural network (R-CNN). Journal of Digital Imaging, 33(4): 916-929. https://doi.org/10.1007/s10278-020-00329-x

[23] Katz, J.N., Arant, K.R., Loeser, R.F. (2021). Diagnosis and treatment of hip and knee osteoarthritis: A review. Jama, 325(6): 568-578. https://doi.org/10.1001/jama.2020.22171

[24] Heidari, B. (2011). Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. Caspian Journal of Internal Medicine, 2(2): 205.

[25] Shamir, L., Ling, S.M., Scott, W., Hochberg, M., Ferrucci, L., Goldberg, I.G. (2009). Early detection of radiographic knee osteoarthritis using computer-aided analysis. Osteoarthritis and Cartilage, 17(10): 1307-1312. https://doi.org/10.1016/j.joca.2009.04.010

[26] Janvier, T., Jennane, R., Valery, A., Harrar, K., Delplanque, M., Lelong, C., Loeuille, D., Toumi, H., Lespessailles, E. (2017). Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative: Tibial bone texture & knee OA progression. Osteoarthritis and Cartilage, 25(2): 259-266. https://doi.org/10.1016/j.joca.2016.10.005

[27] Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H., Lespessailles, E. (2019). A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. Computerized Medical Imaging and Graphics, 73: 11-18. https://doi.org/10.1016/j.compmedimag.2019.01.007

[28] Riad, R., Jennane, R., Brahim, A., Janvier, T., Toumi, H., Lespessailles, E. (2018). Texture analysis using complex wavelet decomposition for knee osteoarthritis detection: Data from the osteoarthritis initiative. Computers & Electrical Engineering, 68: 181-191. https://doi.org/10.1016/j.compeleceng.2018.04.004

[29] Antony, J., McGuinness, K., Moran, K., O'Connor, N.E. (2017). Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks. In International Conference on Machine Learning and Data Mining in Pattern Recognition, pp. 376-390. https://doi.org/10.1007/978-3-319-62416-7_27

[30] Antony, J., McGuinness, K., O'Connor, N.E., Moran, K. (2016). Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In 2016 23rd International Conference on Pattern Recognition (ICPR), pp. 1195-1200. https://doi.org/10.1109/ICPR.2016.7899799

[31] Gornale, S.S., Patravali, P.U., Hiremath, P.S. (2020). Osteoarthritis detection in knee radiographic images using multiresolution wavelet filters. In International Conference on Recent Trends in Image Processing and Pattern Recognition, pp. 36-49. https://doi.org/10.1007/978-981-16-0493-5_4

[32] Kotti, M., Duffell, L.D., Faisal, A.A., McGregor, A.H. (2017). Detecting knee osteoarthritis and its discriminating parameters using random forests. Medical Engineering & Physics, 43: 19-29. https://doi.org/10.1016/j.medengphy.2017.02.004

[33] Astuto, B., Flament, I., K. Namiri, N., Shah, R., Bharadwaj, U., M. Link, T., Bucknor, M.D., Pedoia, V., Majumdar, S. (2021). Automatic deep learning–assisted detection and grading of abnormalities in knee MRI studies. Radiology: Artificial Intelligence, 3(3): e200165. https://doi.org/10.1148/ryai.2021200165

[34] Raj, A., Vishwanathan, S., Ajani, B., Krishnan, K., Agarwal, H. (2018). Automatic knee cartilage segmentation using fully volumetric convolutional neural networks for evaluation of osteoarthritis. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 851-854. https://doi.org/10.1109/ISBI.2018.8363705

[35] Gornale, S., Patravali, P. (2020). Digital knee X-ray images. Mendeley Data, 1. https://doi.org/10.17632/t9ndx37v5h.1

[36] Aslan, S. (2022). A novel TCNN-Bi-LSTM deep learning model for predicting sentiments of tweets about COVID-19 vaccines. Concurrency and Computation: Practice and Experience, 34(28): e7387. https://doi.org/10.1002/cpe.7387

[37] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://arxiv.org/abs/1409.1556

[38] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. https://doi.org/10.48550/arXiv.1704.04861

[39] Ghosal, P., Nandanwar, L., Kanchan, S., Bhadra, A., Chakraborty, J., Nandi, D. (2019). Brain tumor classification using ResNet-101 based squeeze and excitation deep neural network. In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), pp. 1-6. https://doi.org/10.1109/ICACCP.2019.8882973

[40] Heravi, E.J., Aghdam, H.H., Puig, D. (2016). Classification of foods using spatial pyramid convolutional neural network. In CCIA, pp. 163-168. https://doi.org/10.3233/978-1-61499-696-5-163

[41] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-9. https://doi.org/10.48550/arXiv.1409.4842

[42] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Li, F.F. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision

and Pattern Recognition, pp. 248-255. https://doi.org/10.1109/CVPR.2009.5206848

[43] Tuncer, T., Ertam, F. (2020). Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. Physica A: Statistical Mechanics and its Applications, 540: 123143. https://doi.org/10.1016/j.physa.2019.123143

[44] Raghu, S., Sriraam, N. (2018). Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms. Expert Systems with Applications, 113: 18-32. https://doi.org/10.1016/j.eswa.2018.06.031

[45] Praveen, S.V., Ittamalla, R., Deepak, G. (2021). Analyzing the attitude of Indian citizens towards COVID-19 vaccine–A text analytics study. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 15(2): 595-599. https://doi.org/10.1016/j.dsx.2021.02.031

[46] Safavian, S.R., Landgrebe, D. (1991). A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, 21(3): 660-674. https://doi.org/10.1109/21.97458

[47] Lachenbruch, P.A., Goldstein, M. (1979). Discriminant analysis. Biometrics, 45(4): 69-85. https://doi.org/10.3102/00346543045004543

[48] Jain, R., Nagrath, P., Kataria, G., Kaushik, V.S., Hemanth, D.J. (2020). Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. Measurement, 165: 108046. https://doi.org/10.1016/j.measurement.2020.108046

[49] Cortes, C., Vapnik, V. (1995). Support-vector networks. Machine le3arning, 20(3): 273-297. https://doi.org/10.1007/BF00994018

[50] Peterson, L.E. (2009). K-nearest neighbor. Scholarpedia, 4(2): 1883. https://doi.org/10.4249/scholarpedia.1883

[51] Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1): 217-222. https://doi.org/10.1016/j.isprsjprs.2016.01.011