RESEARCH ARTICLE

The effect of deep feature concatenation in the classification problem: An approach on COVID-19 disease detection

Emine Cengil ⬤ | Ahmet Çınar

Department of Computer Engineering, Faculty of Engineering, Firat University, Elazig, Turkey

Correspondence
Emine Cengil, Department of Computer Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
Email: ecengil@firat.edu.tr

Abstract

In image classification applications, the most important thing is to obtain useful features. Convolutional neural networks automatically learn the extracted features during training. The classification process is carried out with the obtained features. Therefore, obtaining successful features is critical to achieving high classification success. This article focuses on providing effective features to enhance classification performance. For this purpose, the success of the process of concatenating features in classification is taken as basis. At first, the features acquired by feature transfer method are extracted from AlexNet, Xception, NASNETLarge, and EfficientNet-B0 architectures, which are known to be successful in classification problems. Concatenating the features results in the creation of a new feature set. The method is completed by subjecting the features to various classification algorithms. The proposed pipeline is applied to the three datasets: "COVID-19 Image Dataset," "COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset," and "COVID-19 Radiography Database" for COVID-19 disease detection. The whole datasets contain three classes (normal, COVID, and pneumonia). The best classification accuracies for the three datasets are 98.8%, 95.9%, and 99.6%, respectively. Performance metrics are given such as: sensitivity, precision, specificity, and F1-score values, as well. Contribution of paper is as follows: COVID-19 disease is similar to other lung infections. This situation makes diagnosis difficult. Furthermore, the virus's rapid spread necessitates the need to detect cases as soon as possible. There has been an increased curiosity in computer-aided deep learning models to provide the requirements. The use of the proposed method will be beneficial as it provides high accuracy.

KEYWORDS
classification, convolutional neural network (CNN), COVID-19, features concatenation, machine learning algorithms

26 | wileyonlinelibrary.com/journal/ima | *Int J Imaging Syst Technol.* 2022;32:26–40.

1 | INTRODUCTION

1.1 | Motivation

At the end of 2019, a new virus called SARS-CoV-2 (COVID-19) emerged.[1] In the last decade, new Corona viruses have posed a serious threat to human health. The viruses named as (SARS)-CoV and (MERS)-CoV, which were previously encountered as public health threats all over the world, are also known as the previous outbreaks of the corona virus.[2] It has influenced the world. It was declared as an urgent public health problem of international concern by the World Health Organization on January 30, 2020.[3] In August and September 2020, it was reported that the virus had mutated. Viruses can mutate all the time. These may have little or no effect. But there is a possibility that mutations will cause the virus to spread faster. In the region where the new species was found in the United Kingdom, it was observed that the virus spread more rapidly than before. Subsequently, the mutated virus was found in various countries of the world.[4]

Different symptoms are seen in patients infected with the virus. The most prevalent symptoms are dry cough, high fever, and chills. Less prevalent symptoms are headache, loss of smell, skin rash, and sore throat.[5] For individuals with weak immunity and chronic illness, the virus is deadly. Considering the spread rate and serious effects of the virus, it is vital that individuals infected with the virus are detected early. The most efficient method to control COVID-19, which is transmitted by contact and droplets, is to ensure the isolation and treatment of infected patients by early detection.[5]

Diagnosis for patients suspected of infection can be confirmed by many laboratory tests. Full genome sequencing and phylogenetic analysis on bronchoalveolar lavage fluid can verify COVID-2019 infection. Some samples (nasal secretions, blood, sputum, and bronchoalveolar lavage) are taken from suspected patients and subjected to COVID-19-specific serological and molecular tests. In serological tests, enzyme-linked immunosorbent assay (ELISA) or Western blot tests that identify particular COVID-19 proteins are used. In molecular diagnosis, real time transcriptase polymerase chain reaction (RT-PCR) or Northern blot hybridization is used. Furthermore, viral antigens found in clinical samples are detected using the direct immunofluorescence assay (IFA). The gold standard used for diagnosis is the PCR test. It is suggested that nucleic acid tests reach an accuracy of only 30%–50%.[6] The limitations and quality of laboratory tests in the infected area require alternative diagnostic methods.

Recent research has shown that tools such as computed tomography (CT) and X-ray contain important information about the COVID-19 virus and can be utilized as an other diagnostic method.[7] Chest X-ray stands out because of its faster imaging time, low cost, and portability. One of the most common and effective methods used by researchers for the diagnosis of the disease is chest X-ray images.[8,9] However, radiology specialist and time are required to examine the images. This situation is challenging in pandemic conditions. Therefore, as an alternative, automatic, speed screening methods based on artificial intelligence are useful for early diagnosis and to alleviate the workload of medical staff in the field.

1.2 | Literature review

Deep learning is commonly utilized in the medical field. It has been applied in a number of fields, including Breast cancer classification,[10,11] skin cancer classification,[12] anomaly detection,[13] brain tumor detection,[14,15] and lung cancer detection.[16,17] Deep learning is also being used in COVID-19 identification research.[18–20] Various methodologies are suggested, including classification, segmentation, and detection methods.

Verma et al.[21] performed predictive analysis to verify the COVID-19 positive case using the long short-term memory (LSTM). In Ref. [22] a deep learning assisted automated technique utilizing X-ray images for early detection of COVID-19 is proposed. Eight pre-trained convolutional neural network (CNN) architectures were evaluated. The best performance was achieved with ResNet-34 with an accuracy of 98.33%. Xu et al.[23] aimed to differentiate COVID-19, IAVP, and healthy individuals through pulmonary CT images utilizing deep learning methods. The study provided an accuracy of 86.7%. Hemdan et al.[24] used images of 25 positive cases to distinguish COVID-19 from X-ray images. It includes seven different training frameworks such as Covidxnet, ResNet, and GoogleNet. The highest accuracy was achieved with vgg19 as 90%. In Ref. [25], 217 images were used for training and inception migration–learning model was used to create the algorithm. The method obtained an accuracy of 82.9%.

In Song et al.[26] chest CT scans were collected from hospitals of two provinces in China, diagnosed with COVID-19, infected with bacterial pneumonia and healthy persons. Based on the dataset collected, a deep learning-based diagnostic system was implemented to identify COVID-19 patients. The model was able to distinguish patients infected with COVID-19 with an AUC value of 0.95. In the study, ground glass image

localization in the lung was also made. A weakly supervised deep learning technique for detecting and classifying COVID-19 disease from CT images was suggested in Ref. [27]. He et al.[28] created a public dataset containing the CT scan positive for COVID-19. To reduce the risk of over-compliance, a Self-Trans approach that integrates contrast-enhanced self-trans learning with transfer learning has been proposed. The proposed Self-Trans approach reached an F1-score of 0.85 and an AUC of 0.94 in the diagnosis of COVID-19. A fully automated deep learning method for COVID-19 diagnosis and prognostic analysis by CT was proposed by Wang et al.[29] In four external validation sets, the deep learning system performed well in classifying COVID-19 from other pneumonia and viral pneumonia.

In Ref. [30], convolutional CapsNet has been suggested for the classification of COVID-19 infection utilizing capsule networks and chest X-ray images. The proposed strategy is intended to provide diagnosis for binary and multiclass COVID-19 diseases. The recommended method's accuracy for binary and multiclass classifications was 97.24% and 84.22%, respectively. Ozturk et al.[31] made COVID classification from X-ray images using the modified darknet architecture. Authors obtained an accuracy of 87.02%. The Fourier–Bessel series expansion-based dyadic decomposition (FBD) approach for image decomposition was introduced by the authors in Ref. [32]. In the method, the subband images obtained with FBD are given to the pre-trained resnet50 network. Using procedures, deep features from each CNN are combined. The softmax classifier is then fed the community CNN features. The overall accuracy of the model is 98.6%.

Hussain et al.[33] suggested a 22 layer CNN structure, which has been used for 2, 3, and 4 class classification using X-ray images. They obtained an accuracy of 99.1%, 94.2%, and 91.2%, respectively. In Ref. [34], different learning rates have been tried to eliminate the overfitting problem. In addition, it has been proposed to use the transfer learning method and the two-layer CNN community to identify COVID-19 images. They provided an accuracy of 90.45%. Yıldırım et al.[35] made a COVID-19 classification from X-ray images with the method they created by increasing the number of layers of the ResNet50 structure. The technique illustrated an accuracy of 96.30%.

Most of the studies reported in the literature have achieved high accuracy in binary classification for diagnosing COVID-19. However, the success rate in multiple classification could not reach a high accuracy value in the relevant datasets. It is important to make more accurate multiple classifications in order to differentiate COVID-19 disease from other lung diseases. The study focuses on this subject and works with three-class datasets. The proposed method achieved high accuracy rates for three-class classification. In addition, the proposed method is also efficient in terms of computation time.

1.3 | Contributions of the paper

The aim of this study is to design and evaluate a concatenate based method for automated detection of COVID-19 disease using chest X-ray images. In our study, three COVID-19 X-ray image datasets are examined. In the study, multiclass (COVID, normal, and pneumonia) classification is carried out with proposed method. Simple block diagram of the proposed pipeline is represented in Figure 1.

In the proposed method, firstly, the feature transfer method is concatenated with features from Alexnet,[36] Xception,[37] NASNetLarge,[38] and EfficientNet-B0[39] architectures. The feature sets obtained are given to SVM, LDA, KNN, NB, and DT algorithms. Considering the experimental results, it is seen that the method will be useful in detecting COVID-19. The article's contributions can be listed as follows:

- We recommend feature concatenation for detection of COVID-19 based on the feature transfer method.
- We select the most appropriate classifier by performing the classification step with traditional machine learning algorithms.
- We obtain high classification accuracies on three different datasets containing X-ray images.
- Experiments with datasets consisting of three classes demonstrate that the proposed method is useful in the detection of COVID-19 disease.
- The proposed method contributes to the literature by delivering high accuracy and efficiency for multiclass classification.

1.4 | Organization of paper

The study is organized as follows: Section 2 gives the technical background, which consist of the datasets, CNN architectures, and classification methods. The
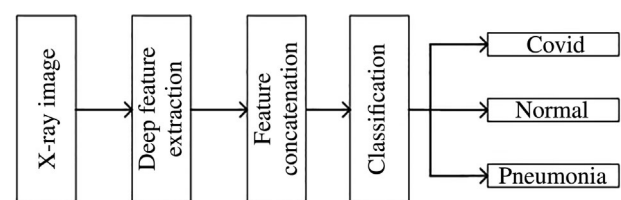

FIGURE 1 Simple block diagram of the proposed pipeline

proposed approach is detailed in Section 3. The experimental results are presented in Section 4. Section 5 gives a discussion and conclusion of the article.

2 | TECHNICAL BACKGROUND

This section explains the dataset used in training and validation processes, CNN utilized for feature extraction, and traditional machine learning methods used in the classification step. Technical background knowledge is required to understand and apply the proposed method.

2.1 | Datasets

The aim of the study is to contribute to the improvement of COVID-19 diagnosis by using datasets containing X-ray images. It has been studied with three different datasets. The first of these is the COVID-19 Image Dataset published by the University of Montreal.[40] The second is the COVID-19 Radiography Database, created in collaboration with a research team from the University of Qatar, Doha, Qatar and the University of Dhaka, Bangladesh, with collaborators in Pakistan and Malaysia.[41,42] Another is COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset.[43] The datasets we use consist of three classes. The number of images they contain is different from each other. The classes and the number of images included in the datasets used in training and validation processes are given in Table 1.

TABLE 1 Class names and number of images included in datasets used in training and validation processes

Dataset name	Class name	Number of image
COVID-19 Image Dataset	COVID	111
	Normal	70
	Viral pneumonia	70
COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset	COVID	1525
	Normal	1525
	Pneumonia	1525
COVID-19 Radiography Database	COVID	3616
	Normal	10 192
	Viral pneumonia	1345

Figure 2 demonstrates some sample images in the datasets. The COVID-19 image dataset utilized in the study contains the least number of images. COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset is balanced in that it contains an equal number of images for each class. The other dataset we prefer is a large dataset and consists of 15 153 images.

2.2 | CNN architectures and feature transfer method

Transfer learning, the process of retraining a particular model for a specific task, pre-trained in some datasets, can be done in two ways. It is known as fine-tuning and feature transfer. Feature transfer is the extraction of features from any layer of the pretrained model.[44] In the study, extraction of features from AlexNet, Xception, NasNet, and Efficient-b0 architectures was performed using the feature transfer method. Each of these architectures has achieved excellent results in the medical field.[45–49] In addition, other CNN models show successful results in the medical field.[8,18] Our aim is to achieve classification with high accuracy by combining features. For this purpose, well-known pre-trained models were preferred, taking into account memory, time, and speed factors.

Alexnet[36] was proposed in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition organized in 2012 by Krizhevsky et al. AlexNet has demonstrated superior success in the dataset to which it was applied compared with other machine learning and computer vision approaches that participated in the competition. Alexnet is made up of eight layers, with the first five being convolutional and the last three being fully connected layers. There are also "pooling" and "activation" layers among these layers. The AlexNet architecture can identify 1000 objects, and the object recognition error rate has been reduced from 26.2% to 15.3%. With this situation, convolutional neural networks (CNN) have gotten a lot of popularity. In the following years, many new architectures were proposed by improving the network structure. One of them, the Xception model,[37] was proposed by Chollet in 2017. The Xception model is based on deeply separable layers of convolution. This model is a modified extension of depthwise separable convolutions with inception modules in the inception[50] architecture. It is a fully CNN architecture. A completely separable hypothesis is the mapping of interchannel and spatial correlations on the property maps of CNN. The architecture that stands for "Extreme Inception" is called Xception since this theory is a stronger version of the Inception architecture's underlying hypothesis.
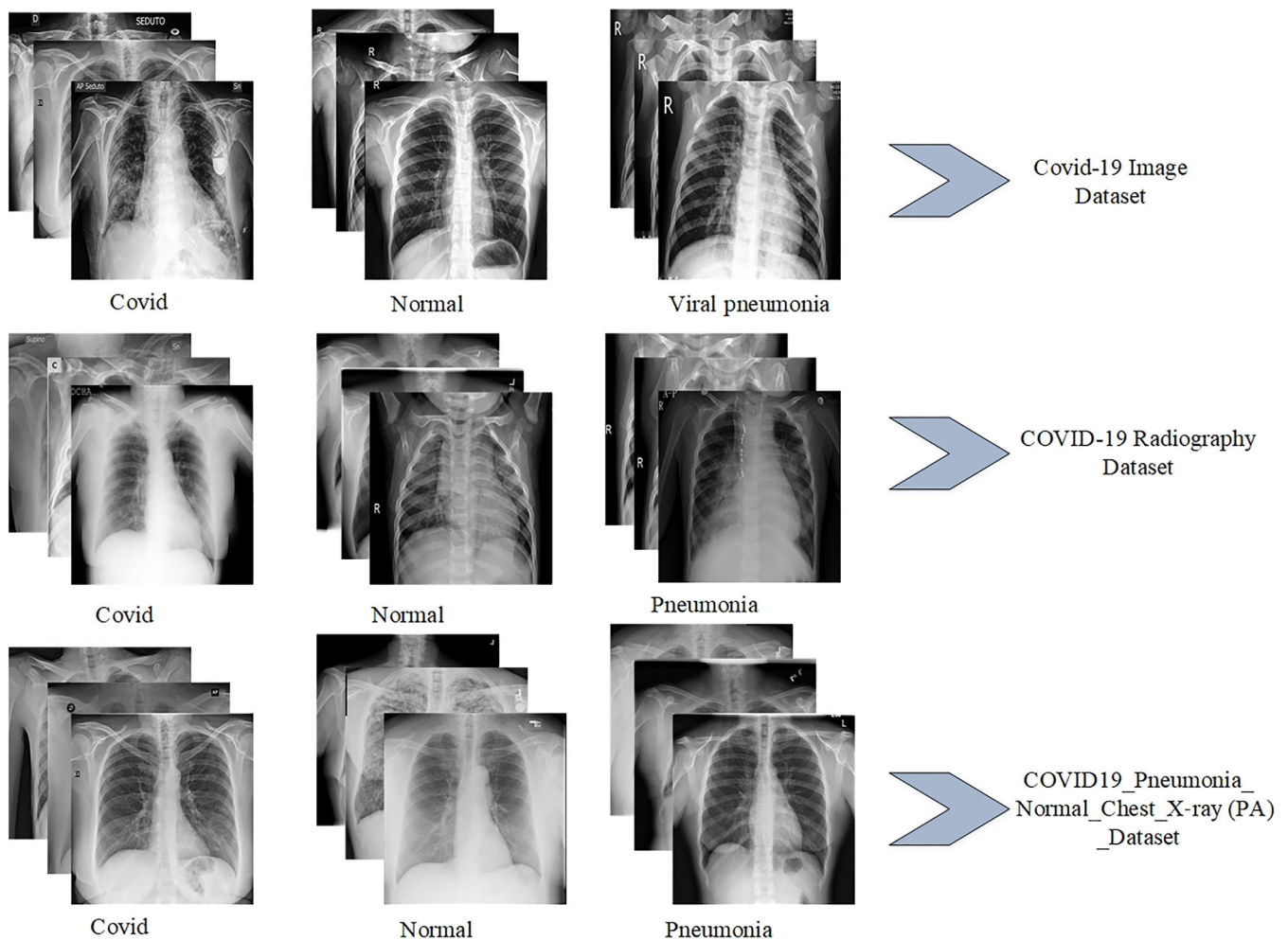
FIGURE 2 Sample images from used datasets

Neural Search Architecture Network (NASNET)[38] was proposed by Google Brain. NASNET's basic principles are different from standard architectures such as AlexNet. The general setup of NAS includes search space, search strategy, and performance estimation components. The authors recommend identifying a structural building block in a small dataset and then converting it to a larger dataset. They specifically search for the best convolutional layer or cell in CIFAR-10, and then add that cell to ImageNet by assembling more copies of it. With a smaller model size and lower complexity, the NASNET model produces cutting-edge performance. EfficientNet, another architecture we use as backbone, was proposed by Tan and Le. Unified scaling method was adopted to achieve better performance. Before EfficientNet is proposed, researchers are trying to scale the width, depth, and resolution of the image to improve network performance. But how to balance all these dimensions is not defined. Based on this observation, a new scaling method has been proposed that equally scales all depth/width/resolution dimensions using a composite coefficient. All EfficientNet models are scaled from the basic EfficientNet-B0 using different composite coefficients φ.[39]

2.3 | Classification methods

Hybrid architectures created by using traditional machine learning techniques in the classification layer of CNN architectures show high performance in classification tasks.[51,52] Therefore, in our method, classification was carried out with traditional machine learning methods. The methods of supervised machine learning that we employ.

2.3.1 | Support vector machine

In machine learning, the support vector machine (SVM) is a strong classifier. Each data item is plotted as a point in m-dimensional space, where m is the number of attributes that define the value of a given coordinate. Then
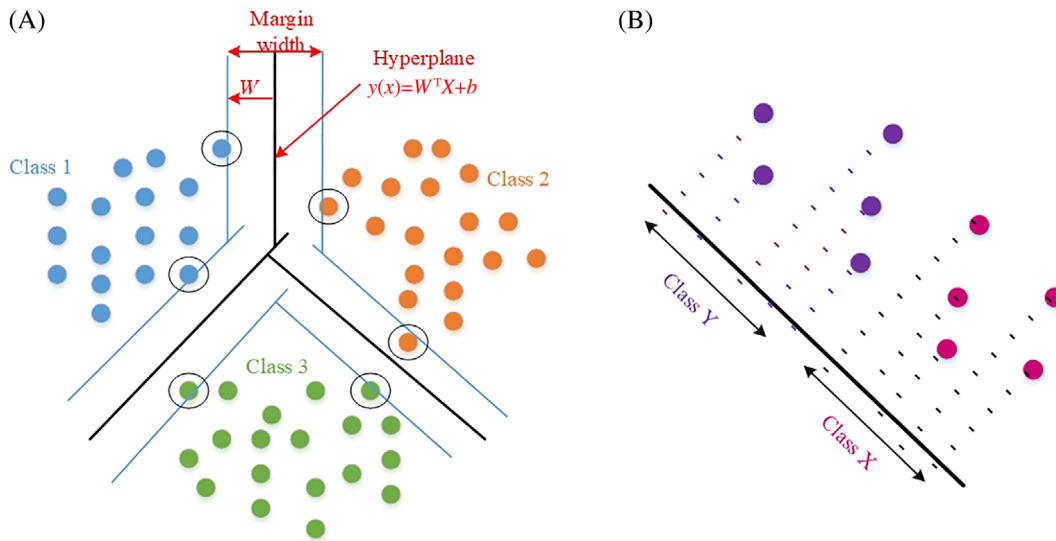
FIGURE 3 (A) SVM for three class classification, (B) The LDA's guiding principle. Two-dimensional data samples are projected into a lower-dimensional space (line)

the classification is made by finding the hyperplane used to separate the classes.[53] Multiple classification with SVM is given in Figure 4. Here, each class pair is trained to separate instances from one class in the other class. Blue, orange, and green points, respectively, reflect the three groups in Figure 3. The hyperplanes and their related support vectors formed between the blue and orange classes, the orange and green classes, and the blue and green classes, respectively, are used to create three binary classifiers to classify them. The test data sample is categorized into the class that earned the most votes after being classified by all three binary classifiers.[54]
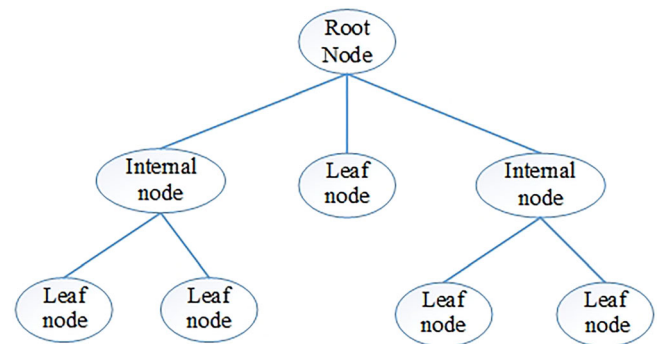
2.3.2 | Linear discriminant analysis

Linear discriminant analysis (LDA) is a basic data analysis method suggested by R. Fisher. The method's premise is to define a lower dimensional subspace in comparison with the original data sample size, where the original problem's data points are "separable," as seen in Figure 3B. The line should be selected in such a way that the predicted samples are as separable as possible. One of the benefits of LDA is that it can solve a generalized eigenvalue system to find the solution. The original algorithm was optimized for binary class problems, but there have been multiclass generalizations suggested as well.[55,56]

2.3.3 | K nearest neighbor

K nearest neighbor (KNN) classifies unlabeled objects by placing them to the class of the most similarly labeled

FIGURE 4 A view of a decision tree diagram[58]

instances. The characteristics of the observations are collected for both the training and the test dataset. The KNN, which was proposed in 1970, has been used in many applications since then. In the KNN algorithm, the distance between the training and sample points is calculated, and the point with the shortest distance is referred to as the nearest neighbor. It occurs in two stages in total. In the first step, find the k training instances that are nearest to the unknown instance. The next step is to select the most common class for these k instances. KNN algorithm is mainly utilized when all features are continuous.[57]

2.3.4 | Decision tree

Decision tree (DT) is a type of inductive learning. The aim is to create a model for the process that produces the data for a given dataset. Training is controlled by a

consequence factor known as a target. Iterative binary partitioning is a tool for increasing DTs. A node is a location in a tree where an attribute is checked. The outcomes of a test that leads to another node are called branches. The root node, internal node, and leaf node are the three types of nodes in the tree. The root node is at the top of the tree, and the inner nodes are in between leaf and end nodes. The purity of each node determines whether the test is completed. A node is terminated when it exceeds a predetermined degree of class purity. To classify a new case, a feature values are compared with the decision tree. A path is taken from the root node to the leaf node that contains the class prediction for that case. Figure 4 portrays a schematic view of DT. The main task of creating a DT is to locate a feature to be evaluated on one node repeatedly and then branch to the other node.[58]

2.3.5 | Naïve Bayes

Naïve Bayes (NB) is an algorithm that uses Bayes' rule and an assumption that attributes are conditionally independent by class. Bayes' theorem connects the conditional and marginal probabilities of two random events in probability theory, and it is often used to measure posterior probabilities provided observations. NB often provide competitive classification accuracy. It is often preferred in applications due to its calculation efficiency and many other desired properties.[59] Bayes rule is given in Equation (1).

$$P(b|a) = P(b)P(a|b)/P(a) \tag{1}$$

NB is a form of Bayes Network Classifier based on the Bayes rule with the assumption that the attributes are conditionally independent with respect to the class. In the equations below, a_i is the ith attribute in a. Also, m is

the number of features, n is the number of classes, and c_i is the ith class.

$$P(a|b) = \prod_{i=1}^{m} P(a_i|b) \tag{2}$$

$$P(a) = \prod_{i=1}^{n} P(c_i)P(a|c_i). \tag{3}$$

3 | PROPOSED METHOD

Transfer learning approach is often used to adapt pre-trained deep architectures to a new task. The transfer learning method is applied in two ways. The first is the fine-tuning method and the other is the feature transfer method. The feature transfer approach was used in this study. In the proposed method, the CNN architectures Alexnet, Xception, NASNETLarge, and Efficientnet-B0 are used as backbones to classify COVID-19 images. In the study, 1000 features are extracted from the last fully connected layers of the models using the feature transfer method. Then, the extracted features are combined to create a new feature set (4000). In the last step of the proposed method, feature sets are classified using traditional machine learning algorithms. The classifier that achieves the highest classification accuracy is selected. Figure 5 denotes a graphical abstract of the proposed method.

4 | EXPERIMENTAL RESULTS

To assess classification performance, a variety of metrics are employed. The accuracy (ACC), sensitivity (SN), specificity (SP), precision (PREC), and F1-score parameters
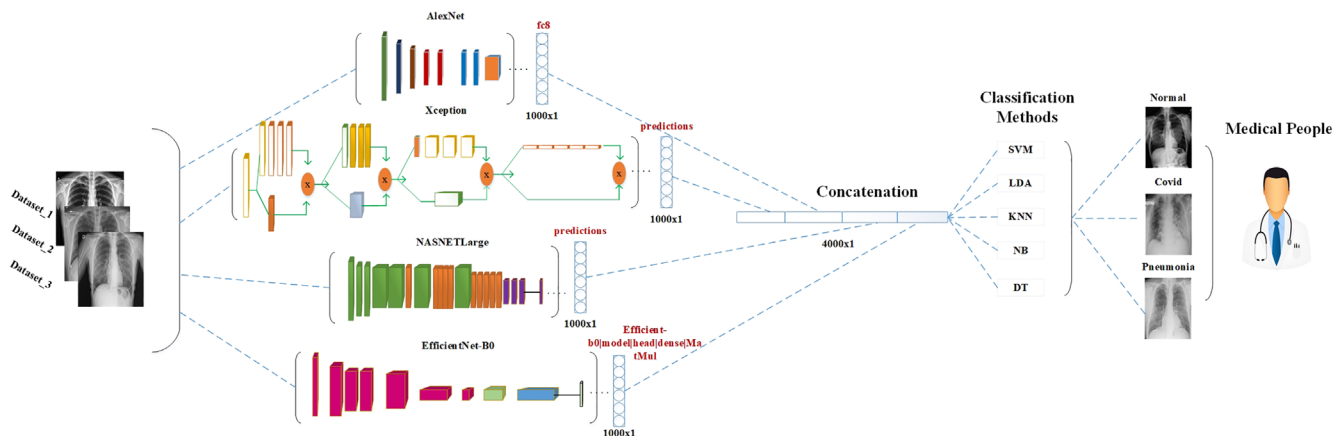

FIGURE 5 Graphical abstract of proposed method

TABLE 2 Formulas for calculating performance measures

ACC	SN	SP	PREC	F1-score
$\frac{TP+TN}{TP+TN+FP+FN}$	$\frac{TP}{TP+FN}$	$\frac{TN}{TN+FP}$	$\frac{TP}{TP+FP}$	$\frac{(1+\beta^2)(PREC \cdot SN)}{(\beta^2 \cdot PREC + SN)}$

were used to assess the models' efficiency. Formulas for the calculation of success measurements are shown in Table 2. The terms used here are True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP). The value of 1 was chosen for the F1-score.[60]

TABLE 3 Classification accuracy (%) of models according to the datasets to which they are applied

Dataset	Features	LDA	SVM	KNN	DT	NB
COVID-19 image Dataset	AlexNet	90.8	89.6	86.1	80.9	89.2
	EfficientNet-b0	94.4	94.4	92	80.9	89.2
	NASNetLarge	89.2	93.2	87.3	78.9	83.7
	Xception	80.5	89.2	87.6	80.5	85.3
COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset	AlexNet	93.4	93.2	89.8	84.2	84.9
	EfficientNet-b0	94.7	93.9	91.8	84.1	87.1
	NASNetLarge	93.0	92.3	87.3	82.8	77.9
	Xception	93.4	93.2	89.8	83.0	85.9
COVID-19 Radiography Database	AlexNet	95.4	92.7	90.5	83.5	71.2
	EfficientNet-b0	96.7	95.7	94.5	85.6	77.9
	NASNetLarge	93.2	92.5	86.8	82.0	72.1
	Xception	93.9	92.6	88.8	81.7	70.3

FIGURE 6 Confusion matrix and performance metric values for COVID-19 image Dataset with EfficientNet-B0 + LDA

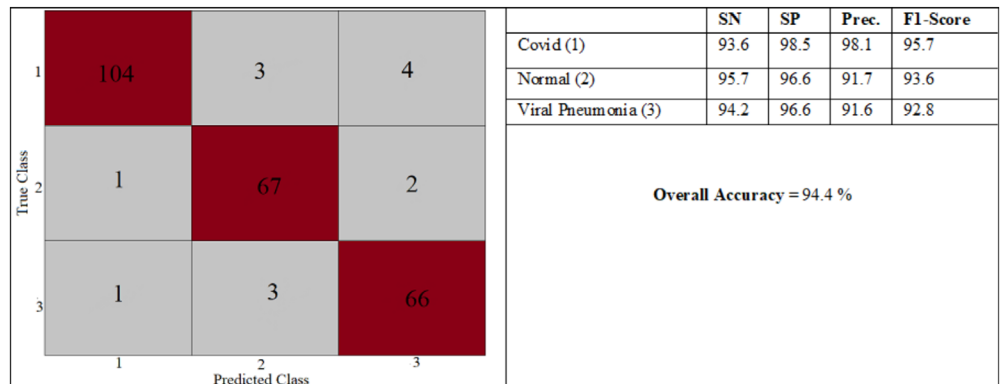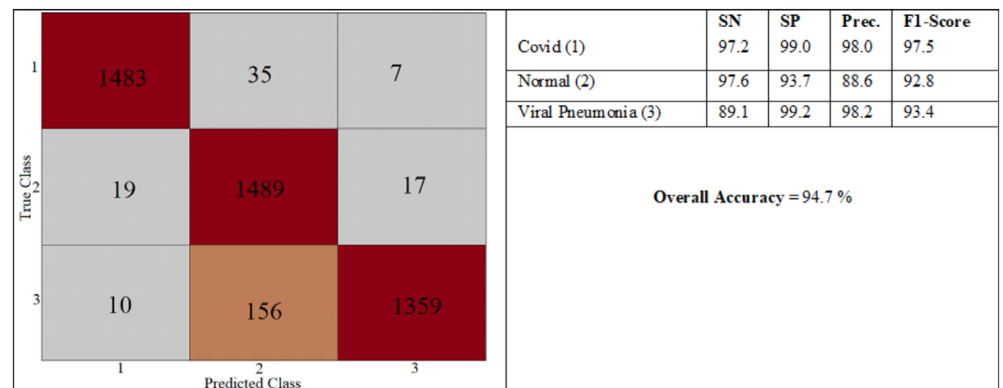|	SN	SP	Prec.	F1-Score
Covid (1)	93.6	98.5	98.1	95.7
Normal (2)	95.7	96.6	91.7	93.6
Viral Pneumonia (3)	94.2	96.6	91.6	92.8

Overall Accuracy = 94.4 %

FIGURE 7 Confusion matrix and performance metric values for COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset with EfficientNet-B0 + LDA

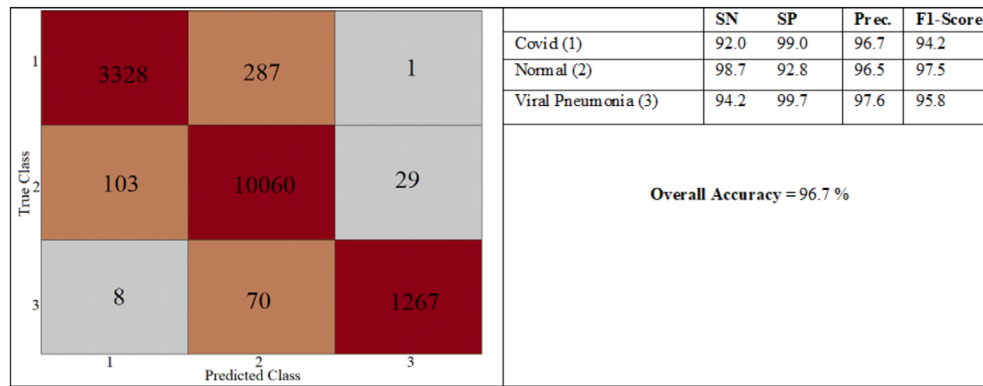|	SN	SP	Prec.	F1-Score
Covid (1)	97.2	99.0	98.0	97.5
Normal (2)	97.6	93.7	88.6	92.8
Viral Pneumonia (3)	89.1	99.2	98.2	93.4

Overall Accuracy = 94.7 %

FIGURE 8 Confusion matrix and performance metric values for COVID-19 radiography database with EfficientNet-B0 + LDA

TABLE 4 Experimental results of the feature set created by concatenating features from two architectures (COVID-19 image Dataset)

Feature extractor	Classifier	Acc. (%)	SN (%)	SP (%)	PREC (%)	Time(s) (%)
AlexNet + EfficientNet-b0	LDA	98.0	99.0	100	100	4.651
	SVM	97.2	98.0	99.2	99.0	4.589
	KNN	94.0	91.8	100	100	3.864
	DT	81.3	88.2	90.0	87.5	4.958
	NB	94.0	98.1	97.1	96.4	10.757
AlexNet + NASNetLarge	LDA	94.0	94.5	98.5	98.1	4.003
	SVM	96.0	98.1	98.5	98.1	3.791
	KNN	90.4	94.5	97.1	96.3	3.542
	DT	80.5	90.0	90.7	88.4	3.661
	NB	93.2	96.3	96.4	95.5	10.538
AlexNet + Xception	LDA	92.8	94.5	100	100	4.053
	SVM	96.8	98.1	100	100	3.997
	KNN	92.0	92.7	100	100	3.883
	DT	83.9	90.9	95.7	94.3	3.854
	NB	95.6	97.2	100	100	11.171
EfficientNet-b0 + NASNetLarge	LDA	97.2	98.1	99.2	99.0	4.052
	SVM	97.2	97.2	99.2	99.0	4.321
	KNN	91.6	92.7	98.5	98.0	4.038
	DT	76.5	84.6	87.8	84.6	3.956
	NB	93.6	97.2	97.8	97.2	11.139
EfficientNet-b0 + Xception	LDA	96.0	97.2	100	100	4.217
	SVM	97.2	97.2	100	100	4.153
	KNN	96.4	95.4	100	100	4.019
	DT	80.1	87.3	93.3	90.6	4.010
	NB	93.6	95.4	100	100	11.475
NASNetLarge + Xception	LDA	94.8	95.4	97.8	97.2	4.760
	SVM	97.6	99.0	99.2	99.0	4.485
	KNN	90.8	91.8	97.1	96.2	3.862
	DT	80.1	90.9	90.0	87.8	3.942
	NB	94.8	97.2	98.5	97.2	12.314

GPU support is provided while compiling models. Before being utilized as input for CNNs, images are resized to suit the input size of the networks. The system environment is 64-bit Windows 10, NVIDIA GeForce 950M graphics, Intel © i7-7500U @ 2.7 GHz processor, 16 GB RAM. Experiments using three datasets containing X-ray images were carried out in Matlab (2020a) environment. Data consisting of three classes are tested with 5-cross fold.

The datasets are implemented in the first step of the experiments by feeding features from pre-trained CNN architectures to machine learning algorithms. The models' performance results are presented in Table 3. In all three datasets, the model using the LDA classifier with the highest performance EfficientNet-B0 features is shown. The method provides accuracy for COVID-19 image Dataset, COVID-19 Pneumonia Normal Chest Xray PA Dataset, and COVID-19 Radiography Database, as 94.4%, 94.7%, and 96.7%, respectively.

Figure 6 gives the confusion matrix and performance measurement values of the model in which the COVID-19 image dataset is used in training and validation. The proposed model only rendered 14 misclassifications in the dataset, as seen in the Figure 6. In addition, SN, SP, PREC, and F1 score values of the classes show that the classification is highly accurate. Likewise, the results obtained according to the other two datasets are as in Figures 7 and 8, respectively.

In the second step of the study, new feature sets were obtained by concatenating the models in pairs and triples. Then these feature sets are given to the classification algorithms. Table 4 portrays the results obtained by combining the features from only two of the four models used. Models have been applied to the COVID-19 image dataset. The model in which the feature set obtained by concatenating the features extracted from AlexNet and EfficientNet networks is given to the LDA classifier achieved the highest accuracy of 98.0%.

Table 5 reveals the results obtained by combining the features from only three of the four models. In accordance with the results, the highest accuracy is obtained with the features provided by AlexNet, EfficientNet-b0, and Xception architectures. LDA and SVM both reached the same accuracy (98.4%) for the respective feature set as classifiers. Tables 4 and 5 also demonstrate the obtained SN, SP, PREC (for COVID class), and training time.

In the third step of the study, features from the four architectures were concatenated. Then, the obtained feature sets were given to various traditional machine learning classifiers as in the first experiments. The results gathered were given in Table 6. According to the table, the classification is achieved with 98.8% accuracy by using the LDA classifier in the experiment performed with COVID-19 Image Dataset. In the experiment using the features extracted with COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset, SVM classification achieved an accuracy of 95.9%. Finally, among the algorithms applied to the deep feature set created with the COVID-19 Radiography Database, the highest performance was achieved with 99.6% accuracy with LDA.

TABLE 5 Experimental results of the feature set created by concatenating features from three architectures (COVID-19 image Dataset)

Feature extractor	Classifier	Acc. (%)	SN (%)	SP (%)	PREC (%)	Time(s)
AlexNet + EfficientNet-b0 + NASNetLarge	LDA	98.0	99.0	100	100	8.044
	SVM	98.0	99.0	99.2	99.0	7.576
	KNN	94.4	95.4	100	100	7.180
	DT	84.1	91.8	90.0	87.9	7.027
	NB	95.2	97.2	97.1	96.4	20.428
AlexNet + EfficientNet-b0 + Xception	LDA	98.4	99.0	100	100	8.673
	SVM	98.4	99.0	100	100	7.886
	KNN	96.4	97.2	100	100	7.301
	DT	83.7	86.4	95.7	94.1	7.719
	NB	97.2	97.2	100	100	20.97
EfficientNet-b0 + NASNetLarge + Xception	LDA	97.6	99.0	100	100	8.830
	SVM	98.0	99.0	99.2	99.0	8.057
	KNN	94.8	95.4	100	100	8.360
	DT	80.5	86.4	92.8	90.5	7.560
	NB	96.0	98.1	98.5	98.1	22.17

TABLE 6 Experimental results of the feature set created by concatenating features from four architectures

| Dataset	Feature extractor	Classifier	Acc. (%)	SN (%)	SP (%)	PREC (%)	Time (s)
COVID-19 image Dataset	AlexNet + EfficientNet-b0 + NASNetLarge + Xception	LDA	98.8	99.0	99.2	99.0	10.827
		SVM	98.8	98.1	100	100	11.1
		KNN	97.2	97.2	99.2	99.0	9.8717
		DT	84.5	93.6	92.8	91.2	10.779
		NB	98.0	97.2	100	100	28.869
COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset	AlexNet + EfficientNet-b0 + NASNetLarge + Xception	LDA	72.6	83.5	90.5	81.5	192.42
		SVM	95.9	98.6	99.9	99.8	55.071
		KNN	92.0	91.4	99.9	99.7	223.39
		DT	84.8	86.5	93.9	87.6	82.697
		NB	93.2	93.2	99.0	93.2	43.496
COVID-19 Radiography Database	AlexNet + EfficientNet-b0 + NASNetLarge + Xception	LDA	99.6	98.7	99.9	99.8	418.93
		SVM	99.3	97.8	99.8	99.5	227.02
		KNN	93.2	76.3	99.1	96.5	2195.9
		DT	88.4	74.3	95.5	84.0	201.49
		NB	91.1	92.5	94.8	84.9	102.12
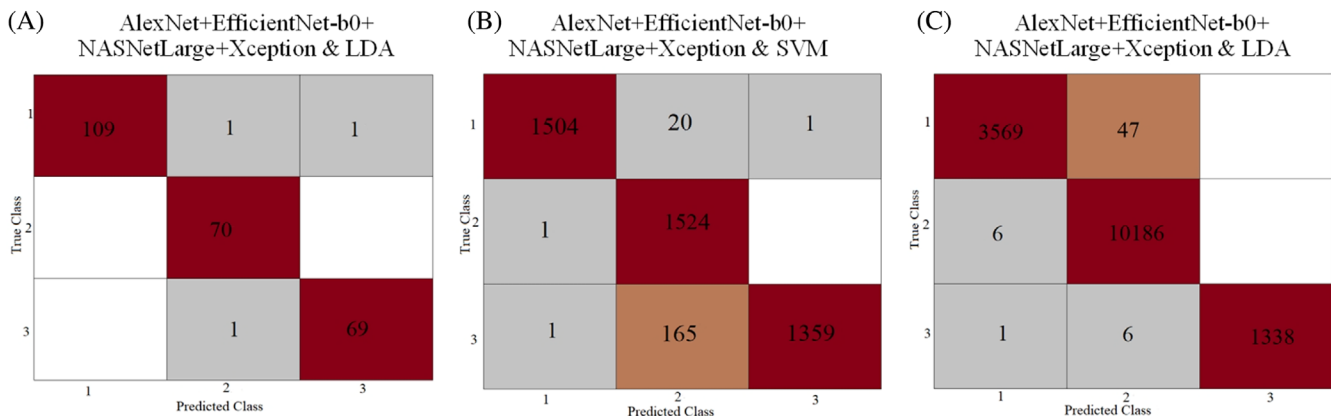
FIGURE 9 Confusion matrices of models that achieve the best results in (A) COVID-19 image Dataset, (B) COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset, and (C) COVID-19 Radiography Database

Confusion matrixes of the models that reach the highest accuracy are given in Figure 9. In Figure 9A), while 14 misclassifications were made in the model that provided the highest performance in the first experiments for COVID image dataset, the number of misclassified images was reduced to three with the improved method. Figure 9B shows COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset with 1325 images in each class was improved to 95.7%. Finally, as shown in Figure 9C, the number of misclassified images decreased and the accuracy rate improved from 96.7% to 99.6% for the COVID-19 Radiography Database, which contains 15 153 images and is the largest dataset we used. When compared with the results obtained with the models we applied without concatenating the features, it is seen that the accuracy values increased in all three datasets.

5 | DISCUSSION AND CONCLUSION

COVID-19 has a worldwide impact in a short time. Its transmission by touch and droplets makes it difficult to control the virus. The way to control the virus is through the speed detection of infected people. In this way, people can be isolated and treated. The RT-PCR test, which is the gold standard for the identification of cases, is not

TABLE 7 Performance comparison of the proposed method with other approaches

Ref. no/year	Images	Method	Dataset source	Performance	Class
2020[28]	CT images	Self-Trans	COVID-CT[28]	0.85 F1 0.94 AUC	COVID-19 Non-COVID-19
2020[29]	CT images	DenseNet121-FPN	From six cities in China	0.86 AUC	COVID-19 Bacterial pneumonia Mycoplasma pneumonia Viral pneumonia Fungal pneumonia
2020[30]	X-ray images	CapsNet	COVID-Chest X-ray dataset	Acc. 84.22 (Multi) Acc. 97.24 (Binary)	COVID-19 No-findings Pneumonia
2020[25]	CT images	Transfer learning on inception	Three hospitals from China	0.93 AUC Specificity 88.0 Sensitivity 87.0	Viral pneumonia
2020[31]	X-ray images	DarkNet	COVID-Chest X-ray dataset Chest X-ray	Sensitivity 85.53 Acc. 87.02	SARS-COV-2 No finding Pneumonia
2020[32]	X-ray images	FBD + ResNet50 + Softmax	750 Images	Acc. 98.6% Specificity 94% Sensitivity 96%	Pneumonia COVID-19 Healthy
2020[26]	CT images	ARENET (ResNet50 and FPN)	Hospitals of two provinces in China	Precision 93% Recall 93% Acc. 93%	COVID-19 Bacterial pneumonia Healthy
2021[22]	X-ray images	ResNet34	Collected from two sources[18]	Acc. 98.33%	Normal COVID-19
2021[33]	X-ray images	22 layer CNN	COVID-R[29]	Acc. 99.1% (2 class) Acc. 94.2% (3 class) Acc. 91.2% (4 class)	COVID-19 Normal Pneumonia (viral + bacteria)
2021[34]	X-ray images	Transfer learning on CNN	COVIDGR 1.0[30]	Acc. 90.45%	COVID Non-COVID
2021[61]	X-ray images	InstacovNet-19	3150 images	Acc. 99.53% (2 class) Acc. 99.08% (3 class)	COVID-19 Pneumonia Normal
2020[62]	X-ray images	COVID-Net	COVIDx Dataset	Acc. 93.3%	COVID Non-COVID Normal
2020[63]	X-RAY images	Transfer learning on VGG19	1424 images	Acc.98.75 (2 class) Acc. 93.48 (3 class)	COVID-19 Pneumonia

(Continues)

TABLE 7 (Continued)

Ref. no/year	Images	Method	Dataset source	Performance	Class
					Normal
The proposed method	X-ray images	AlexNet + EfficientNet-B0 + NASNetLarge + Xception and LDA	COVID-19 image Dataset	Acc. 98.4%	COVID / Normal / Viral pneumonia
The proposed Method	X-ray images	AlexNet + EfficientNet-B0 + NASNetLarge + Xception and SVM	COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset	Acc. 95.7%	COVID / Normal / Viral pneumonia
The proposed method	X-ray images	AlexNet + EfficientNet-B0 + NASNetLarge + Xception and LDA	COVID-19 Radiography Database	Acc. 99.6%	COVID / Normal / Pneumonia

sufficient by itself. In addition, the test is difficult to obtain, especially when cases are on the rise. There are other tests and methods for detecting cases. However, in pandemic conditions, sufficient health personnel and devices may not be possible. For such reasons, a need has arisen for computer aided systems that can produce fast and accurate results. Especially deep learning-based methods have been developed. Some of these methods are as in Table 7.

According to Table 7, the suggested method gains a more accuracy rate than other methods. Using Capsnet to detect COVID-19 using Chest X-ray images,[30] achieved 97.24% accuracy for binary classification. However, this rate was reported as 84.2% for three class classification. Poonam et al.[34] made a COVID, non-COVID classification with the method they applied using the transfer learning approach. The classification accuracy achieved was 90.45% in the respective dataset. Wang et all.[62] In their triple classification study, they suggested a CNN-based model called COVID-Net. Authors using COVIDx Dataset with this model achieved 93.3% classification accuracy.

In Ref. [63] the authors have retrained the pre-trained CNN architectures VGG19, MobileNetV2, Inception, Xception, and Inception ResNet V2, using the transfer learning method. The best result of these architectures has been achieved with VGG19. Classification accuracy provided as 98.75% for two classes could reach 93.48% for three classes. In another study, Ref. [61], it was studied with a limited number of data. There are 270 images in each class for training. In the relevant dataset, an accuracy of 99.08% is reported for three classes.

When the studies are examined, methods should be developed for fast and accurate detection of cases. For this aim, three different datasets containing X-ray images from normal, COVID, and pneumonia classes were used as inputs in our study. In the first stage of the suggested method, the feature set is created by concatenating the features taken from the last layers of AlexNet, EfficientNet-B0, NASNETLarge, and Xception architectures. In CNNs, the first layers give general information that does not contribute much to the classification. The last layers contribute to the classification by giving more specific information. Therefore, the last layer is preferred while taking the features. In the second step of the method, the classification process was carried out using traditional machine learning algorithms. Among these algorithms, LDA provided the highest performance for COVID-19 Image Dataset and COVID-19 Radiography Database. The highest performance for COVID-19 Pneumonia Normal Chest X-ray (PA) Dataset was achieved with the SVM classifier. According to the results, it has been observed that extracting and concatenated features

from CNN architectures with the feature transfer method is effective for COVID-19 classification. The key novelty of the study is the contribution made to the literature in terms of providing high accuracy and efficiency with the proposed pipeline.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable - no new data generated.

ORCID

Emine Cengil [ID] https://orcid.org/0000-0003-4313-8694

REFERENCES

1. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* 2020;395(10223):497-506.
2. WHO. Middle East respiratory syndrome coronavirus (MERS-CoV). November, 2019. Accessed February 13, 2021. http://www.who.int/emergencies/merscov/en/
3. World Health Organization. Novel Coronavirus (2019-nCoV) Situation Report—10 [online]. Website. Accessed February 13, 2020. https://www.who.int/docs/defaultsource/coronaviruse/situation-reports/20200130-sitrep-10-ncov.pdf?sfvrsn=d0b2e480_2
4. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. *Bull World Health Organ.* 2020;98(7):495-504.
5. Sağlik Bakanliği TC. COVID-19 (SARS-CoV2 enfeksiyonu) rehberi. Accessed July 16, 2020. https://covid19bilgi.saglik.gov.tr/depo/rehberler/COVID-19_Rehberi.pdf
6. Kutlu R. What we have learned about the new coronavirus pandemic, current diagnostic and therapeutic approaches and the situation in Turkey. *TJFMPC.* 2020;14(2):329-344.
7. Qi X, Brown LG, Foran DJ, Nosher J, Hacihaliloglu I. Chest X-ray image phase features for improved diagnosis of COVID-19 using convolutional neural network. *Int J Comput Assist Radiol Surg.* 2021;16(2):197-206.
8. Khan MA, Kadry S, Zhang YD, et al. Prediction of COVID-19-pneumonia based on selected deep features and one class kernel extreme learning machine. *Comput Electr Eng.* 2021;90:106960.
9. Rauf HT, Lali MIU, Khan MA, et al. Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks. *Pers Ubiquitous Comput.* 2021;25:1-18.
10. Emine C, Ahmet Ç. Göğüs verileri metrikleri üzerinden kanser siniflandirilmasi. *Dicle Univ Mühendis Fak Mühendis Derg.* 2020;11(2):513-519.
11. Tsang J, Tse GM. Molecular classification of breast cancer. *Adv Anat Pathol.* 2020;27(1):27-35.
12. Chaturvedi SS, Gupta K, Prasad PS. Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet. *International Conference on Advanced Machine Learning Technologies and Applications.* Springer; 2020.
13. Han C, Rundo L, Murao K, et al. MADGAN: unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction. *BMC Bioinformatics.* 2021;22:1-20.
14. Çinar A, Yildirim M. Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. *Med Hypotheses.* 2020;139:109684.
15. Waghmare VK, Kolekar MH. Convolutional neural network-based automatic brain tumor detection. In: Singh PK, Noor A, Kolekar MH, Tanwar S, Bhatnagar RK, Khanna S, eds. *Evolving Technologies for Computing, Communication and Smart World.* Springer; 2020:463-474.
16. Bonavita I, Rafael-Palou X, Ceresa M, Piella G, Ribas V, González Ballester MA. Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline. *Comput Methods Programs Biomed.* 2020;185:105172.
17. Cengil E, Cinar A. A deep learning based approach to lung cancer identification. In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP); 2018:1-5.
18. Khan MA, Hussain N, Majid A, et al. Classification of positive COVID-19 CT scans using deep learning. *Comput Mater Contin.* 2021;66(3):3483-3498.
19. Zhang YD, Khan MA, Zhu Z, Wang SH. Pseudo zernike moment and deep stacked sparse autoencoder for COVID-19 diagnosis. *Comput Mater Contin.* 2021;69:3145-3162.
20. Majid A, Khan MA, Nam Y, et al. COVID19 classification using CT images via ensembles of deep learning models. *Comput Mater and Contin.* 2021;69:319-337.
21. Verma P, Negi CS, Panth M, Saxena, A. A. Deep_Saliency: A Deep Learning based Saliency Approach to detect Covid-19 through x-ray images. Research Square; 2021. https://doi.org/10.21203/rs.3.rs-65967/v2
22. Nayak SR, Nayak DR, Sinha U, Arora V, Pachori RB. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. *Biomed Signal Process Control.* 2021;64:102365.
23. Xu X, Jiang X, Ma C, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering.* 2020;6(10):1122-1129.
24. Hemdan EED, Shouman MA, Karar ME. Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in x-ray imagesHemdan, arXiv preprint arXiv:2003.11055 (2020)
25. Wang S, Kang B, Ma J, et al. A deep learning algorithm using ct images to screen for corona virus disease (COVID-19). *MedRxiv.* 2020;31:6096–6104.
26. Song Y, Zheng S, Li L. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021.
27. Hu S, Gao Y, Niu Z, et al. Weakly supervised deep learning for covid-19 infection detection and classification from ct images. *IEEE Access.* 2020;8:118869-118883.
28. He X, Yang X, Zhang S, et al. (2020). Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. medrxiv. *MedRxiv.* 2020.
29. Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J.* 2020;56(2):2000775.
30. Toraman S, Alakus TB, Turkoglu I. Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Soliton Fract.* 2020;140:110122.

31. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med.* 2020;121:103792.

32. Chaudhary PK, Pachori RB. Automatic diagnosis of COVID-19 and pneumonia using FBD method. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2020:2257–2263.

33. Hussain E, Hasan M, Rahman MA, Lee I, Tamanna T, Parvez MZ. CoroDet: a deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos Soliton Fract.* 2021;142:110495.

34. Verma P, Tripathi V, Pant B. Comparison of different optimizers implemented on the deep learning architectures for COVID-19 classification. *Mater Today Proc.* 2021;46:11098-11102.

35. Yildirim M, Cinar A. A deep learning based hybrid approach for COVID-19 disease detections. *Trait du Signal.* 2020;37(3): 461-468.

36. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst.* 2012;25:1097-1105.

37. Chollet, F. Xception: deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017.

38. Zoph B, Vasudevan V, Shlens J, Le Q. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018.

39. Tan, M, Le Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks. In International Conference on Machine Learning, PMLR; 2019.

40. Covid-19 Image Dataset. [accessed 2021 Apr 08]. https://www.kaggle.com/pranavraikokte/covid19-image-dataset

41. Chowdhury MEH, Rahman T, Khandakar A, et al. Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access.* 2020;8:132665-132676.

42. Rahman T, Khandakar A, Qiblawey Y, et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine.* 2021;132:104319.

43. COVID19_Pneumonia_Normal_Chest_Xray(PA)_Dataset. [accessed 2021 April 08]. https://www.kaggle.com/amanullahasraf/covid19-pneumonia-normal-chest-xraypa-dataset

44. Zhu W, Braun B, Chiang L, Romagnoli J. Investigation of transfer learning for image classification and impact on training sample size. *Chemom Intel Lab Syst.* 2021;211:104269.

45. Igarashi S, Sasaki Y, Mikami T, Sakuraba H, Fukuda S. Anatomical classification of upper gastrointestinal organs under various image capture conditions using AlexNet. *Comput Biol Med.* 2020;124:103950.

46. Chetoui M, Akhloufi MA. Explainable diabetic retinopathy using EfficientNET. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2020:1966–1969.

47. Martínez F, Martínez F, Jacinto E. Performance evaluation of the NASNet convolutional network in the automatic identification of COVID-19. *Int J Adv Sci Eng Inform Technol.* 2020;10(2):662.

48. Rahimzadeh M, Attar A. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inform Med Unlocked.* 2020;19:100360.

49. Togacar M, Ergen B, Sertkaya ME. Subclass separation of white blood cell images using convolutional neural network models. *Elektron Elektrotech.* 2019;25(5):63-68.

50. Szegedy C, Ioffe S, Vanhoucke V. Inception-v4, inception-resnet and the impact of residual connections on learning. (2016)arXiv Preprint. arXiv:1602.07261

51. Ahlawat S, Choudhary A. Hybrid CNN-SVM classifier for handwritten digit recognition. *Procedia Comput Sci.* 2020;167: 2554-2560.

52. Bhandary A, Prabhu GA, Rajinikanth V, et al. Deep-learning framework to detect lung abnormality–a study with chest X-ray and lung CT scan images. *Pattern Recognit Lett.* 2020;129: 271-278.

53. Koranga M, Pandey R, Joshi M, et al. Analysis of white wine using machine learning algorithms. *Mater Today Proc.* 2021;46: 11087-11093.

54. Wu H, Wang L, Zhao Z, Shu C, Lu C. Support vector machine based differential pulse-width pair brillouin optical time domain analyzer. *IEEE Photon J.* 2018;10(4):1-11.

55. Xanthopoulos P, Pardalos PM, Trafalis TB. Linear discriminant analysis. *Robust data mining.* Springer; 2013:27-33.

56. Tharwat A, Gaber T, Ibrahim A, Hassanien AE. Linear discriminant analysis: a detailed tutorial. *AI Commun.* 2017;30(2):169-190.

57. Deekshatulu BL, Chandra P. Classification of heart disease using k-nearest neighbor and genetic algorithm. *Procedia Technol.* 2013;10:85-94.

58. Vanfretti L, Arava VSN. Decision tree-based classification of multiple operating conditions for power system voltage stability assessment. *Int J Electr Power Energy Syst.* 2020;123:106251.

59. Webb GI. Naïve Bayes. In: Sammut C, Webb GI, eds. *Encyclopedia of Machine Learning.* Vol 15. Springer; 2010:713-714.

60. Luque A, Carrasco A, Martín A, De las Heras A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognit.* 2019;91:216-231.

61. Gupta A, Gupta S, Katarya R. InstaCovNet-19: a deep learning classification model for the detection of COVID-19 patients using chest X-ray. *Appl Soft Comput.* 2021;99:106859.

62. Wang L, Lin ZQ, Wong A. Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci Rep.* 2020;10(1):1-12.

63. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med.* 2020;43(2):635-640.