RESEARCH ARTICLE

A Novel Fusion Model of Hand-Crafted Features With Deep Convolutional Neural Networks for Classification of Several Chest Diseases Using X-Ray Images

HASSAAN MALIK[1], **TAYYABA ANEES**[2], **MUHAMMAD UMAR CHAUDHRY**[3], **RADOMIR GONO**[4], (Senior Member, IEEE), **MICHAŁ JASIŃSKI**[4,5], (Member, IEEE), **ZBIGNIEW LEONOWICZ**[4,5], (Senior Member, IEEE), AND **PETR BERNAT**[4]

[1]Department of Computer Science, School of Systems and Technology, University of Management and Technology, Lahore 54000, Pakistan
[2]Department of Software Engineering, School of Systems and Technology, University of Management and Technology, Lahore 54000, Pakistan
[3]Department of Computer Science, MNS-University of Agriculture, Multan 60000, Pakistan
[4]Department of Electrical Power Engineering, Faculty of Electrical Engineering and Computer Science, VSB–Technical University of Ostrava, 708-00 Ostrava, Czech Republic
[5]Department of Electrical Engineering Fundamentals, Faculty of Electrical Engineering, Wrocław University of Science and Technology, 50-370 Wrocław, Poland

Corresponding author: Muhammad Umar Chaudhry (umar.chaudhry@mnsuam.edu.pk)

ABSTRACT With the continuing global pandemic of coronavirus (COVID-19) sickness, it is critical to seek diagnostic approaches that are both effective and rapid to limit the number of people infected with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The results of recent research suggest that radiological images include important information related to COVID-19 and other chest diseases. As a result, the use of deep learning (DL) to assist in the automated diagnosis of chest diseases may prove useful as a diagnostic tool in the future. In this study, we propose a novel fusion model of hand-crafted features with deep convolutional neural networks (DCNNs) for classifying ten different chest diseases such as COVID-19, lung cancer (LC), atelectasis (ATE), consolidation lung (COL), tuberculosis (TB), pneumothorax (PNET), edema (EDE), pneumonia (PNEU), pleural thickening (PLT), and normal using chest X-rays (CXR). The method that has been suggested is split down into three distinct parts. The first step involves utilizing the Info-MGAN network to perform segmentation on the raw CXR data to construct lung images of ten different chest diseases. In the second step, the segmented lung images are fed into a novel pipeline that extracts discriminatory features by using hand-crafted techniques such as SURF and ORB, and then these extracted features are fused to the trained DCNNs. At last, various machine learning (ML) models have been used as the last layer of the DCNN models for the classification of chest diseases. Comparison is made between the performance of various proposed architectures for classification, all of which integrate DCNNs, key point extraction methods, and ML models. We were able to attain a classification accuracy of 98.20% for testing by utilizing the VGG-19 model with a softmax layer in conjunction with the ORB technique. Screening for COVID-19 and other lung ailments can be accomplished using the method that has been proposed. The robustness of the model was further confirmed by statistical analyses of the datasets using McNemar's and ANOVA tests respectively.

INDEX TERMS COVID-19, deep learning, pneumonia, TB, X-rays, DCNN, feature extraction.

The associate editor coordinating the review of this manuscript and approving it for publication was Jon Atli Benediktsson.

I. INTRODUCTION

The sickness that is commonly referred to as coronavirus (COVID-19) disease is brought on by the severe acute

respiratory syndrome coronavirus 2 (SARS-CoV-2), which is also the virus that is responsible for the ongoing pandemic. A dry cough, myalgia, a sore throat, a headache, a fever, and chest discomfort are just some of the typical symptoms that are associated with COVID-19, which may be categorized as a respiratory disease [1]. After around 14 days, infected individuals often present their full complement of symptoms. As of July 2021, more than 200 countries and territories had reported a combined total of 190 million COVID-19 cases, which resulted in around 4 million fatalities [2]. As a consequence of this, there is now a significant basis for concern among members of the international community regarding the current situation of public health. On March 11, 2020, the disease, which began on January 30, 2020, was recognized as a pandemic, and on the same day, the World Health Organization (WHO) classified it as a public health emergency of international concern (PHEIC) [3]. Even if we have access to a wide range of vaccines at the moment, it will take a considerable amount of time for these vaccines to be distributed all over the world. As a result, the use of visual signals as an alternate method for the rapid screening of individuals who are infected with the virus is a realistic option. It is generally accepted that chest radiographs (CXR) and computed tomography (CT) scans include the most reliable visual indicator of a lung infection that is caused by this virus [3]. Lung infection is the most common symptom that patients experience as a result of exposure to this virus. Radiologists must manually examine these images to search for certain visual patterns that may indicate the existence of the COVID-19 infection. Even while the accuracy of the traditional way of diagnosis has increased over time, it is still possible for it to put the lives of those who work in the medical field in jeopardy. In addition to this, there is an increase in expenditures due to the requirement of diagnostic test kits for each patient. Screening, on the other hand, can be performed with medical imaging techniques such as CXR and CT scans, which, in comparison to other methods, are significantly more time efficient, substantially safer, and easily accessible. CXR image screening is preferred to CT scan screening for COVID-19 not only because it is more freely available, but also because it requires less spending of financial resources to acquire [4], [5].

A manual diagnosis of the virus using a CXR image may take a long time. In addition to the mistakes that are caused by people, there is likely a limited quantity of prior experience, which can result in a significant number of inaccurate readings. Because of this, there is an immediate need to automate such procedures on a massive scale, and this automation needs to be available to all individuals to make the process of diagnosis more efficient, accurate, and quick. Computer vision (CV) and artificial intelligence (AI) technology, especially deep learning (DL) models, has recently been included in recent initiatives [6], [7]. In particular, the usage of convolutional neural networks (CNN) is a helpful method for carrying out medical image analysis. CNN has recently been shown to be effective when applied to the process of assisting in the diagnosis of pneumonia in the chest CXR image of a patient [8], [9].

There has been a substantial amount of research done on the COVID-19 diagnostic technique, the most bulk of which has been on the application of DL models to CXR photos. Using CXR images and a DL network that they named DarkCovidNet, Ozturk et al. [10] developed an automated technique for diagnosing COVID-19. An accuracy of 87.02% was achieved by the model when it was used to categorize the data into three distinct categories i.e., COVID-19, normal, and pneumonia. When classifying into only two groups (COVID-19 and healthy images) it attained a classification accuracy of 98.08%. Hemdan et al. [11] took advantage of CXR and built a network that they later referred to as the COVIDX-Net as part of their research. This network was trained by making use of seven different CNN models, and the validation dataset for the model is comprised of fifty CXR pictures (25 normal and 25 COVID-19 cases). The quantity of validation data that was used in the testing of the model produced an accuracy of 90.00%, which is fairly effective taking into consideration the quantity of data that was used in the testing of the model. The advanced CNN model known as COVID-Net was developed by Wang et al. [12]. They put it through its paces and found that it had an accuracy rate of 93.3% after testing it out. In the study [13] they were successful in accomplishing their objective of achieving an accuracy of 95.38% by employing a combination of ResNet-50 and a support vector machine (SVM) classifier on 50 different samples (25 normal and 25 COVID-19 cases). The researchers Nayak et al. [15] devised a novel method that makes use of DL to screen the COVID-19 chest CXR images. They used a technique known as transfer learning in conjunction with eight of the most successful pre-trained CNN models. These models are as follows: MobileNet-V2, AlexNet, VGG-16, GoogleNet, ResNet-34, SqueezeNet, Inception-V3, and ResNet-50. The ResNet-34 model ended up being the most successful one.

Previous research on COVID-19 identification focused mostly on algorithm-specific aspects and did not pay much attention to the regions of interest (ROI) in CXR images. These ROIs are what reveal the distinctive patterns that are associated with the disease. As a result of this, there is an opportunity for research in this area that centers on making use of the CXR images that are supplied with simply the ROI. Even if it is not accurate, it has the potential to lead to a classification of images that is considerably more precise because it is based on true medical terms. To the best of our knowledge, this is the only work that uses ROI of the CXR images to classify ten different chest diseases such as COVID-19, lung cancer (LC), atelectasis (ATE), consolidation lung (COL), tuberculosis (TB), pneumothorax (PNET), edema (EDE), pneumonia (PNEU), pleural thickening (PLT), and normal. This work is based on three steps: segmentation, extracting features, and classification of the CXR into their

respective classes. The preprocessed input CXR images are used in the image segmentation process [16], and information Maximizing GAN is put to good use to retrieve the lung images from those preprocessed photos (Info-MGAN). The feature extraction network uses a combination of deep neural networks (DNNs) and key point extraction algorithms such as Oriented FAST and Rotated BRIEF (ORB) [17] and speeded-up robust features (SURF) [18] to compute the discriminatory features of the several lung images. The DNNs that were taken into consideration for feature extraction in this work include a simple CNN model as well as a total of seven distinct transfer learning (TL) models. These models are as follows: ResNet-101 & ResNet-50 [12], DenseNet-169 [19], Inception-v3 [16], DenseNet-201 [19], VGG-16, and VGG-19 [20]. At last, several distinct machine learning (ML) methods were applied to categorize the retrieved features into ten distinct types of lungs. Combinations of the chosen models that have the highest levels of performance are investigated, assessed, and compared to determine which application of the suggested method will yield the best results. The contributions of the presents work are given below:

1. CXR images were segmented using the Info-MGAN network.
2. The development of a method for the extraction of features that integrates DCNN with ORB and SURF into a single coherent architecture.
3. The TL and ML models were used for the classification of ten different lung diseases using CXR images.
4. Various combinations of feature extraction techniques and ML models were used for classification performance evaluation.

This work has been divided into further sections: Section II contains modern literature. The datasets and methods that are utilized for image segmentation, feature extraction, and classification are described in Section III. Section IV presents the results of the experiments. This work is concluded in section V.

II. RELATED WORKS
Using a variety of medical imaging modalities, such as sonography, CXR, MRI, and CT scans, one of the most significant duties in DL [14] is the classification of respiratory system disorders. It has been suggested in a few studies that CXR pictures could be used to find COVID-19, which would save time and effort for those working in the medical field [5], [16], [17], [18], [19], [20], [21], [22], [23]. In the trials that are being done at the moment, locating COVID-19 at an early stage of the illness is difficult. The following is a selection of the most significant and pertinent studies about the application of DL and ML models to the diagnosis of various chest infections. The most recent studies on COVID-19 are compared to those conducted on a variety of chest conditions in Table 1.

The findings of this study [21] point to the utilization of a one-of-a-kind CNN model as a potential automatic method for detecting COVID-19 through the usage of chest X-ray pictures. The proposed CNN model is meant to serve as a reliable diagnostic instrument to classify the data into two categories: COVID and Normal. The COVID-19 dataset, which contains 13,824 X-ray pictures, is used to evaluate several architectures, such as the pre-trained MobileNetv2 and ResNet-50 models. The accuracy of the suggested model is compared to that of the existing COVID-19 detection techniques. According to the findings of their experiments, the proposed model can diagnose COVID-19 disease in patients with an accuracy of 96.71 percent and an F1- score of 91.89 percent. Ieracitano et al. [22] offer a fuzzy logic-based DL approach to differentiate between CXR images of patients who have pneumonia caused by Covid19 and images of patients who have interstitial pneumonia that is not related to Covid-19. The model that was constructed for this purpose and is referred to as CovNNet can be used to extract certain relevant features from CXR images if it is paired with fuzzy images that have been made using a fuzzy edge detection technique. Experimental results show that using a combination of CXR and fuzzy features within a DL approach by developing a deep network inputted to MLP results in a higher classification performance (with an accuracy rate of up to 81%).

To impose several distinct binary categories, Narin et al. [23] made use of the methodology of five-fold cross-validation in their research. Because it has an accuracy level of 98%, a specificity value of 100%, and a recall level of 96%, the pre-trained ResNet-50 technique offers the maximum level of efficiency possible. Oh et al. [5] used a patch-based technique in their work, which allowed them to develop CNN models. Only a small number of data sets were used in the application of the approach. The conclusion was reached by the algorithm by first counting the votes cast by the vast majority of patched classifiers and then utilizing those combined votes. This experiment was conducted on a total of 15,043 pictures, which comprised 8851 healthy individuals, 6012 patients who were diagnosed with pneumonia, and 180 patients who tested positive for COVID-19. By achieving an accuracy of 88.9%, a precision of 83.4%, a recall of 85.9%, an f1-score of 84.4%, and a specificity of 84.4%, CNN was able to accomplish these astounding achievements. CNN has an overall accuracy rate of 96.4% for its forecasts. The use of an 18-layer residual CNN to CXR images was suggested as a potential diagnostic tool by Zhang et al. [24]. The three important contributions that they made to the field at the time served as the foundation for the work that they did later on. They began by obtaining features with the assistance of the CNN module, after which they moved on to working on the classification, and finally, for the very last phase, they utilized the anomalous module to compute the detection score. To accomplish the goals of this investigation, radiographs from a total of 1531 patients were analyzed. The results of one hundred of those radiographs came back positive for COVID, whereas the results of the other 1431 radiographs showed pneumonia infection. They were able to achieve a sensitivity

of 96% when it came to the detection of COVID-19 while still maintaining a specificity of 70.65%. Apostolopoulos and Mpesiana [25] developed the CNN model and tested its efficacy on a limited number of datasets before publishing their findings. This allowed them to determine the extent to which the model was successful. This experiment was carried out in parallel on both of the databases that were being used at the same time. The initial database had a total of 1427 images, and 224 of those photographs returned positive results when tested for the presence of COVID-19. Although there were also 700 images depicting bacterial diseases, the remaining shots belonged to people who were in good health. In the second set of data that was collected, there was the same number of COVID-19 CXR images, pneumonia-infected patients, and normal cases as there was in the first set of data that was collected. On the second dataset, we utilized the MobileNetV2 approach, which, out of all of the CNN models, provided the best accuracy of 0.966, the highest specificity of 0.964, and the highest sensitivity of 0.986%.

Tsiknakis et al. [26] presented an innovative automated COVID-19 identification technique that was built on a DL model and was given the moniker Inception-v3. This experiment was conducted on a total of 572 patients, 122 of which were found to have a positive COVID result, 150 of which were found to be normal, and 150 were found to have bacterial or viral diseases. Their classifications are correct 76% of the time out of every 100 times they are used. To extract the features from the dataset, which was comprised of 381 chests of CXR pictures, Sethy et al. [27] used nine different TL algorithms on the dataset. They used SVM to make a diagnosis based on the features of COVID-19 that were received from the virus. When it comes to the process of extracting features from datasets, it has been demonstrated that the ResNet-50 approach is the one that has the highest rate of success. Accuracy values of 95.33% and f1-score values of 95.34%, respectively, are displayed by both a ResNet-50 model and an SVM model. The novel methods of COVID-19 detection based on CXR images have been designed by Saha et al. [28]. They came up with the name EMC Net for their proposed system, which was based on a simple CNN architecture and was designed to extract features from images. After the features had been obtained, they were fed into an ensemble of machine learning [29] classifiers to classify the COVID-19-infected cases. The EMC Net was able to attain an impressive level of accuracy, earning an accuracy of 98.91%. New automatic COVID identification algorithms were developed by Mahmud et al. [30] by making use of a DL model called as CovXNet. To accomplish automatic identification, their DL model made use of a phenomenon known as depth-wise convolution. Those who were diagnosed with pneumonia were evaluated with CXR technology in addition to patients who were found to have normal lung conditions. After this, they evaluated their proposed model by classifying CXR images of COVID-19 and pneumonia based on the results of the test. A slack technique for automatic detection is integrated with

the model's gradient-based discriminative localization. The CovXNet model achieve an accuracy of 97.4% when it was employed to analyze both regular occurrences and COVID-19 situations. However, when it was used to analyze all other cases, such as COVID-19 cases, viral infections, and bacterial infections, it only attained an accuracy of 90.2%. To identify COVID-19, Horry et al. [31] applied four well-established TL classifiers to a total of 60,798 photographs taken from different datasets. This number included 60,361 normal cases in addition to 322 patients diagnosed with pneumonia and 115 COVID-19-positive cases. The datasets contained a combined total of 60,798 photos in their entirety. The VGG-16 and VGG-19 models produced solutions to classification issues that were of the highest possible quality consistency in each one of these four models. When it comes to correctly diagnosing COVID-19-positive cases, the VGG-19 demonstrated an accuracy rate of 81%.

In their study [32], they proposed a classification of infected lungs as COVID-19 (+Ve) and non-infected lungs. However, the COVID-19 classification, which is based on CXR, must be performed by a radiology specialist because it takes a considerable amount of time and requires their skill. Because of this, the creation of an automated testing technique is something that ought to be investigated because it will save a significant amount of time for medical professionals. Within the scope of this study, both the theoretical construction and the actual use of a CNN method are dissected and analyzed. In addition to this, the hyperparameters of CNN are modified with the assistance of multi-objective adaptive evolution of differences to produce superior results. Using the ResNet-18 model as a feature vector allowed Zhang et al. [33] to successfully extract acceptable feature representations from the CXR image. This was accomplished by using the ResNet-18 model. After that, those newly formed features were used as input in research on multi-layer perception. This came about as a result of the previous step. To achieve the best possible degree of accuracy, which was 96%, a dataset consisting of one hundred photographs taken by each of seventy different patients was utilized.

Ahmed et al. [34] worked on the raw data, it was improved by deleting several confusing components that had been introduced into it. To prevent new source information from being generalized, however, well-informed models can make use of source-specific confuses to discriminate between COVID-19 and pneumonia. Our models have the potential to provide an AUC of 0.38 in the worst-case scenario, while in the best-case scenario, they have the potential to give an AUC of 1.00 in the data sources. In light of this, it is abundantly evident that additional testing and development are necessary before the clinical deployment on any significant scale. An innovative pipeline for the deep transfer study of COVID diagnosis has been established for 19 patients, as the data that were displayed in the report that was authored by Michail et al. [35] demonstrated. Chest imaging with a CXR and the diagnosis of pneumonia make up the bulk of

the foundation for this pipeline. Both of the models that come together to form our model were developed to add a layer of neural blocks to their architecture. It is feasible that the same technique will also prove effective in other cases, such as those in which two competing networks are complying with extra performance requirements. When analyzing our suggested network, we took into mind the difficulties that come when involving two classes (pneumonia versus healthy), three classes (including COVID-19), and four courses (tuberculosis included). Zhao et al. [36] CT scans are now being developed as prospective testing and screening processes that are not only quick but also cost-effective. In this article, we generate a COVID-CT dataset that contains 275 positive COVID-19 scans and make it available to the public to encourage research and the development of profound learning algorithms that can determine whether or not an individual is involved in COVID-19. Examining a person's CT scans, which can either be positive or negative for COVID-19, is the method that is utilized to determine whether or not these methodologies should be used. They design a CNN on this dataset, and it achieves an F1 of 0.85, which is encouraging but still has to be improved upon.

Thakur and Kumar [37] build a method that, by utilizing extensive prior knowledge, will be able to automatically detect and identify the COVID-19 disease. CNN is possible to implement not just one but two different classification systems, which are alternately referred to as binary and multiclass grades. The binary model was trained using a total of 3877 x-ray and CT scans, with Covid-19 serving as a source for 1917 of those photographs. In total, there were 3877 photos used for training the model. The instruction was carried out utilizing a mix of both of the available imaging modalities. There was a 99.64% overall accuracy for the binary classification, a 99% recall, a 99.56% accuracy, a 99.59% F1 score, and a 100% ROC score. The research that is currently being presented requires a substantial number of CT scans to create accurate and demanding diagnosis models, as indicated by the findings that were obtained by, He et al. [38]. They produce a public data set that is made up of hundreds of CT scans of COVID-19 and develops learning algorithms that are both effective and profound in their scope. Our efforts have led to the development of these algorithms, which have the potential to achieve a high degree of diagnostic accuracy for COVID-19 despite the limited quantity of CT images available for analysis. We propose a method for self-transitioning that combines contrastive auto-monitored learning with transfer education to acquire robust and impartial representations to reduce the likelihood of overfitting.

A. LITERATURE GAP
According to several studies [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], the signs and symptoms of COVID-19 are comparable to those of many different conditions, including normal, LC, ATE, COL, TB, PNET, EDE, PNEU, and PLT. When it came to COVID-19 and other diseases, it was difficult for doctors and healthcare

professionals to correctly diagnose and identify them using CXR. As a result, there was an obvious need to construct an automated framework that was based on DCNN models and capable of automatically diagnosing COVID-19 and several other disorders as mentioned above using CXR. Previous investigations [21], [27], [28], [34], [35], [36], [37], [38] had the primary objective of attempting to differentiate between COVID-19 and non-COVID-19 instances based on CXR images. CXR images have been used in several research studies [5], [22], [23], [24], [25], [26], [27], [28], [29], [30] to identify COVID-19 from pneumonia illnesses including viral and bacterial infections. Because of this, the purpose of this research work is to design a DCNN framework that would identify COVID-19 and differentiate the ten chest diseases based on CXR. This will help researchers overcome the obstacles that were discussed before.

III. MATERIALS AND METHODS
This section consists of an experimental procedure that was conducted for the segmentation and classification of 10 different chest disease CXR images using DL models. For this work, the Info-MGAN network was used for the segmentation. ORB and SURF methods were used to extract the feature. Then, these extracted features were used by DL models for the classification of lung diseases into their respective classes.

A. DATASETS DESCRIPTIONS
CXR images have been analyzed to identify between patients who have been diagnosed with COVID-19, normal, LC, ATE, COL, TB, PNET, EDE, PNEU, and PLT. The images were collected from a variety of various websites that were publically accessible to the researchers. The datasets that were utilized for segmentation and classification in this work are discussed as follows:

1) SEGMENTATION DATASET
We used CXR images for the segmentation collected from the chest radiographs dataset [39], which is obtained from [40], to train the Info-MGAN. Another set of CXR was obtained from [41], which has a total of 200 chest radiographs in its database. From the two different datasets, a combined total of 447 images showing the frontal aspect of the CXR were obtained. The primary intention behind the creation of this database is to simplify the process of dividing lung areas. Because of this, we divided the dataset into three separate subsets of the CXR images: one for training, one for validating, and one for testing. The primary dataset contained a total of images that could be accessed, and the testing subset had 45 of those images. This represents approximately 10% of the total number of images in the primary dataset. Following this step, the remaining 403 CXR images are partitioned into 80% for training and 10% for validation. Furthermore, the number of images is relatively low, we made use of affine transformation [42] methods to increase the number of images. Finally, the segmentation network used to generate lung

TABLE 1. Summary of modern related work for the classification of chest infections.

Ref	Years	Models	Diseases	Classes	Imaging	Accuracy
[21]	2022	CNN	COVID-19 & non-COVID19	2	CXR	96.00%
[22]	2022	Fuzzy logic	COVID-19, Normal, & Pneumonia	3	X-rays	81.00%
[23]	2021	ResNet-50	COVID-19, Normal & Pneumonia	3	X-rays	98.00%
[5]	2020	CNN	COVID-19, Normal & Pneumonia	3	CXR	88.90%
[24]	2020	CNN	COVID-19, Normal & Pneumonia	3	CXR	97.97%
[25]	2020	CNN	COVID-19, Normal & Pneumonia	3	X-rays	96.78%
[26]	2020	CNN	Bacterial & viral infections, COVID-19	3	CXR	94.00%
[27]	2020	SVM	COVID-19	2	X-rays	95.33%
[28]	2021	EMCNET	COVID-19	2	CXR	98.91%
[29]	2021	CovXNet	COVID-19, Normal & Pneumonia	3	CXR	90.20%
[30]	2020	VGG-16 + VGG-19	COVID-19, Normal & Pneumonia	3	X-rays	81.00%
[31]	2020	CNN	COVID-19	2	CXR	90.23%
[32]	2021	VGG-19 + CVV	COVID-19, Normal & Pneumonia	3	CXR	98.10%
[33]	2020	CNN + Inception-v3	COVID-19, Normal, TB, & Pneumonia	4	CXR	93.70%
[34]	2021	CNN	COVID-19	2	CT scans	98.00%
[35]	2021	CNN	COVID-19	2	X-rays	96.64%
[36]	2020	CNN	COVID-19	2	CT scans	94.00%
[37]	2021	CNN	COVID-19	2	CXR	95.30%
[38]	2020	DNN	COVID-19 cough	2	Audio	88.90%
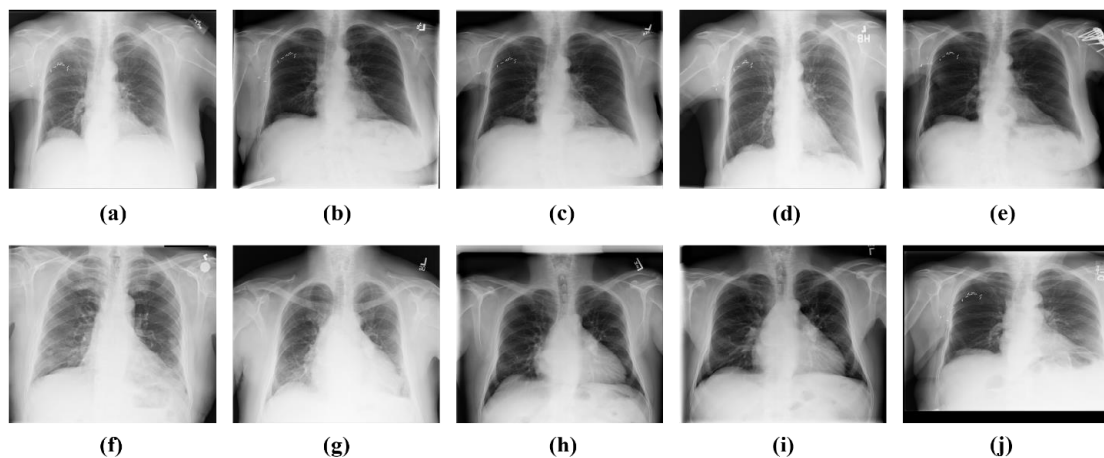
FIGURE 1. CXR images of ten chest infections; (a) COVID-19, (b) lung cancer, (c) atelectasis, (d) consolidation lung, (e) tuberculosis, (f) pneumothorax, (g) edema, (h) pneumonia, (i) pleural thickening, and (j) normal.

masks from chest radiographs has been trained, validated, and tested.

2) CHEST INFECTION CLASSIFICATION DATASETS

A total of ten publically available multiple chest disease datasets were collected from a wide variety of different sources to train and test the DL models. In the beginning, we obtained 930 CXR infected with COVID-19 from a GitHub repository that had been set up by Cohen et al. [43].

This archive contains images that were obtained from a wide variety of hospitals and other public sources.

Although the complete set of metadata information is not going to be presented in this discussion, the patients who had the covid-19 infection were, on average, about 55 years old. From the SIRM database [44], TCIA [45], radiopaedia.org [46], Mendeley [47], and GitHub source [48], a total of 2371 covid-19 positive CXR were obtained. The RSNA [49] was the source from which the dataset of photographs of pneumonia was obtained. This data set
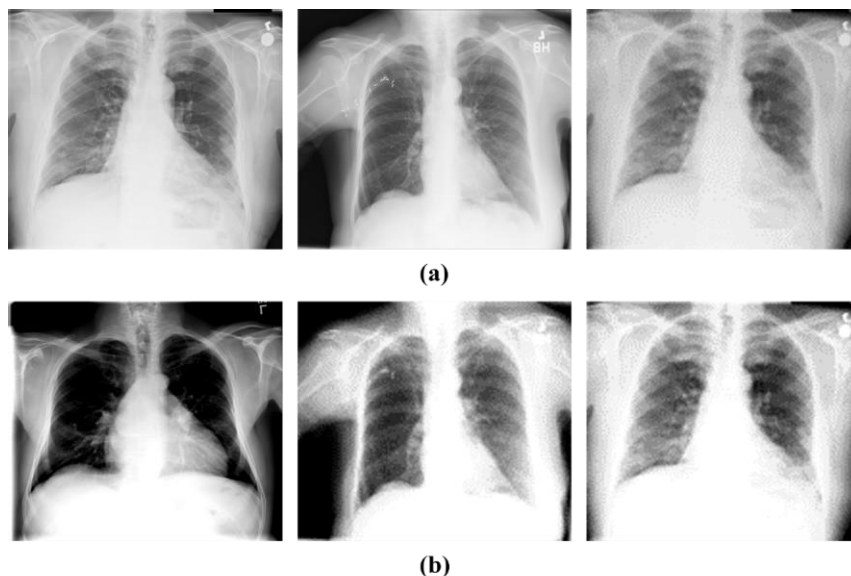
FIGURE 2. Contrast enhancement using AHE; (a) Original CXR and (b) Enhanced.

TABLE 2. Summary of the chest diseases CXR images datasets.

Labels	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal	Total
CXR	3301	5000	5789	6543	700	2793	6331	3867	6046	1349	41,719
Augmentation	7000	7000	7000	7000	7000	7000	7000	7000	7000	7000	70,000
Training (80%)	5600	5600	5600	5600	5600	5600	5600	5600	5600	5600	56,000
Validation (10%)	700	700	700	700	700	700	700	700	700	700	7000
Testing (10%)	700	700	700	700	700	700	700	700	700	700	7000
Total	17301	19000	19789	20543	14700	16793	20331	17867	20046	15349	181719

includes a total of 5216 x-rays, 1349 of which are considered normal, and 3867 of which depict pneumonia. The CXR and CT scan images that make up the lung cancer dataset were obtained from [39]. This dataset contains around 20,000 chest X-rays and CT scans. From the dataset, a total of 5000 CXR images of people with lung cancer were extracted; however, the remaining CT scans were not taken into account for this investigation. The CXR of healthy people was obtained from archives Kaggle [50]. A total of 22,060 CXR images were collected from NIH [51] which included 5789 images of atelectasis, 6543 images of consolidation lung, 2793 of pneumothorax infected CXR, 6331 images of edema, and 6046 images of pleural thickening. At last, a total of 700 TB-infected CXR images were collected [52]. The sample CXR images of all these chest infections are presented in Figure 1.

3) IMAGE PRE-PROCESSING, ENHANCEMENT, AND AUGMENTATION

From the above literature, it is observed that processes for image preprocessing are helpful for improved training of DNNs [47]. The CXR images of chest infections were collected from multiple databases. The resolutions of CXR images varied from 660 to 720 pixels in width and 500 to 672 pixels in height. Thus, we scaled the CXR to the fixed resolutions of 224 × 224 to keep the coherence in this study. Before the segmentation network is built for lung masks, each image was manually reviewed to ensure that it matched the required standards for perfection. After the CXR images that were lacking, in contrast, were found, they were preprocessed with the use of adaptive histogram equalization (AHE) [53] and the thresholding operation is presented in Figure 2.

Furthermore, it is observed that the images of all chest infection classes are imbalanced. The Synthetic Minority Oversampling Technique (SMOTE) was used to balance the dataset and prevent the model from producing biased results. A k-nearest neighbor (KNN) method is utilized in the generation of synthetic data by SMOTE. The first step of SMOTE is to select data at random from the minority class (i.e., normal, LC, ATE, COL, TB, PNET, EDE, PNEU, and PLT), and then to determine the data that was selected. Now, the datasets are ready to be fed into the DNN model for the training. Therefore, 80% of the total CXR images are used to train the DNN, 10% of the images are used for validation, and the remaining 10% of the images are preserved for testing purposes. Table 2 provides an all-encompassing summary of the datasets that were utilized for this work.
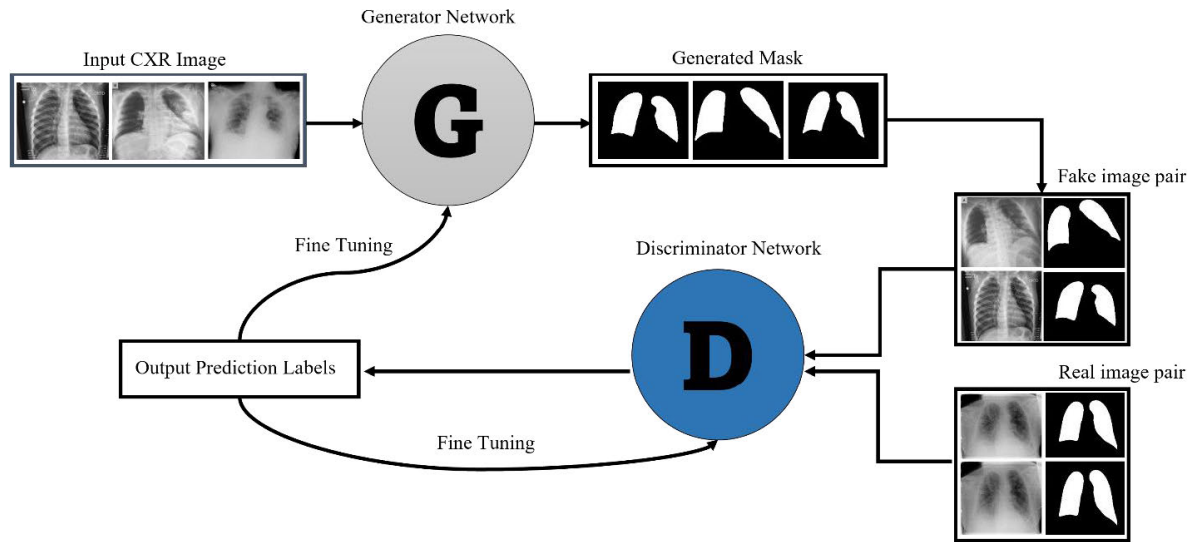
FIGURE 3. Info-MGAN uses CXR images to generate the synthesized segmented mask.

B. LUNG IMAGES SEGMENTATION USING INFO-MGAN

The GAN models that were suggested by [54] have been widely utilized in the field of image processing to translate an input image into its matching output image. The GAN network is made up of two distinct networks: a generator that can produce photorealistic images and a discriminator that can determine whether an image is false or real based on whether or not it was taken from the training dataset. The generator creates images that are incredibly correct to the source material. The image is next analyzed by the discriminator, which determines if it is fake or authentic. The purpose of the generator is to produce photographs that are as lifelike as they possibly may be. The generator Gen will use a min-max strategy in conjunction with the discriminator to transform a set of noise samples Z from distribution P_z into real-world data from distribution P_{data}. Following the completion of the stage before this one, in which the noise samples were loaded into the generator, this transformation will take place.

When the discriminator network is being trained, it makes an effort to differentiate between converted data samples Gen (Z) with distribution P_{data} and actual data samples Y with probability distribution P_Y. To achieve this objective, it is necessary to conduct a comparison of the probability distributions of the two different sets of samples that have been collected. The following Equation (1) is a representation of the mathematical formulation that will be used for the min-max GAN goal function:

$$\min_{gen} \max_B L_{GAN}(Gen, B) = \min_{gen} \max_B A_{p \sim py}[\log B(y)] + A_{z \sim pz}[\log(1 - B(Gen(z)))] \quad (1)$$

The word A used for expectation, while log stands for logarithmic operations. The basic idea of GAN is elaborated upon by Info-MGAN, which consists of a generator called Gen that

generates output data X based on actual data Y and a random noise vector Z. The Info-MGAN is expressed in Equation (2):

$$z, Gen : \{y, z\} \to x \quad (2)$$

The Info-MGAN discriminator B, on the other hand, makes an effort to differentiate between the synthesized data and the actual data (X and Y), both of which it receives as inputs in its operation. The objective function of the Info-MGAN system can be expressed in several different ways, including the following Equation (3):

$$L_{C-GAN}(Gen, B) = A_{x \sim px, y \sim py}[\log B(x, y)] + A_{y \sim py, z \sim pz} \\ \times [\log(1 - B(y, Gen(y, z)))] \quad (3)$$

It is crucial to understand that the Gen serves more than one purpose, and the primary one is not simply to mislead the discriminator. The output of it should, at the same time, correlate with the actual data in an F1 sense as shown in Equations (4-5):

$$L_{F1}(Gen) = A_{x,y,z}[||x - Gen(y, z)||_{F1}] \quad (4)$$
$$Gen^* = \arg \min_{gen} \max_B L_{C-GAN}(Gen, B) + \lambda L_{F1}(Gen) \quad (5)$$

where the F1 parameter was used for F1 regularization weight while B represents the training discriminator B.

The segmentation of the ten chest diseases CXR images was carried out with the assistance of Info-MGAN. In exchange for the CXR that we send to the generator and expect to receive a lung mask that is an exact fit for the patient. The actual CXR images of ten multiple chest diseases and their lung mask that corresponds to the ground truth make up the authentic pair for the discriminator. As can be seen in Figure 3, the Gen, and the D both go through training in an adversarial fashion.
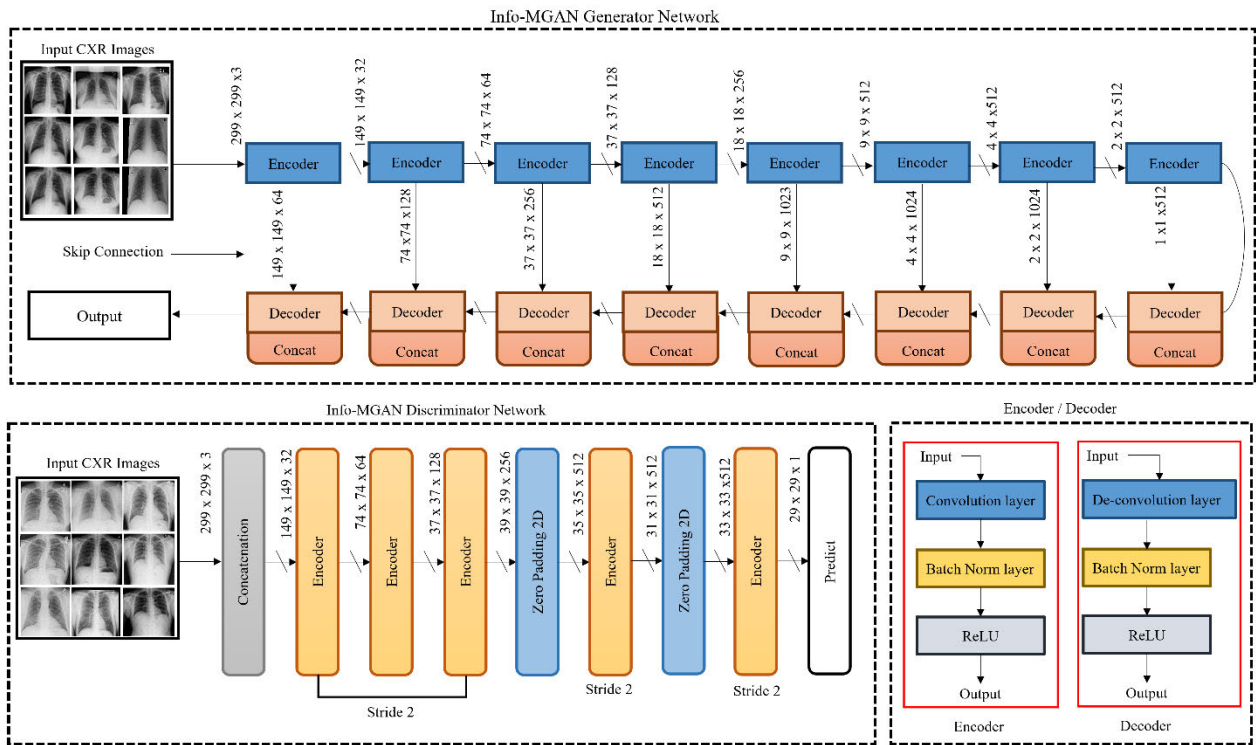
FIGURE 4. Architecture of Info-MGAN used in this study.
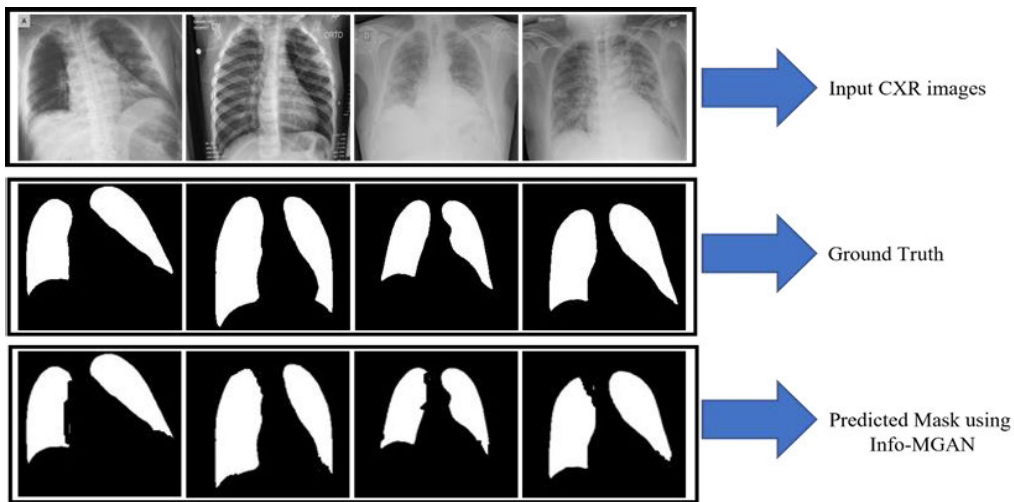
FIGURE 5. Chest CXR image with ground truth and predicted mask.

To conduct the segmentation of different chest diseases using CXR images, we make use of the info-MGAN model that was presented by Chen et al. [55]. The generator (Gen) in Info-MGAN stands in for the encoder-decoder network, while the discriminator (D) is a Patch-GAN. Figure 4 also includes an architecture design depicting the generator, discriminators, encoder, and decoder blocks that are utilized by Info-MGAN. Regardless the block is Gen and D, the convolutional (Conv), batch normalization layer (BNL), and

ReLU make up each encoder block individually. Additionally, each decoder block incorporates deconvolutional, BNL, and ReLU into its architecture Encoder blocks are responsible for producing condensed versions of the data, in contrast to decoder blocks, which produce more detailed versions. The initial three blocks of the decoder are designed to be dropouts throughout both training and testing. This provides the generator with the required amount of random noise so that it can function correctly. In addition to this, the generator
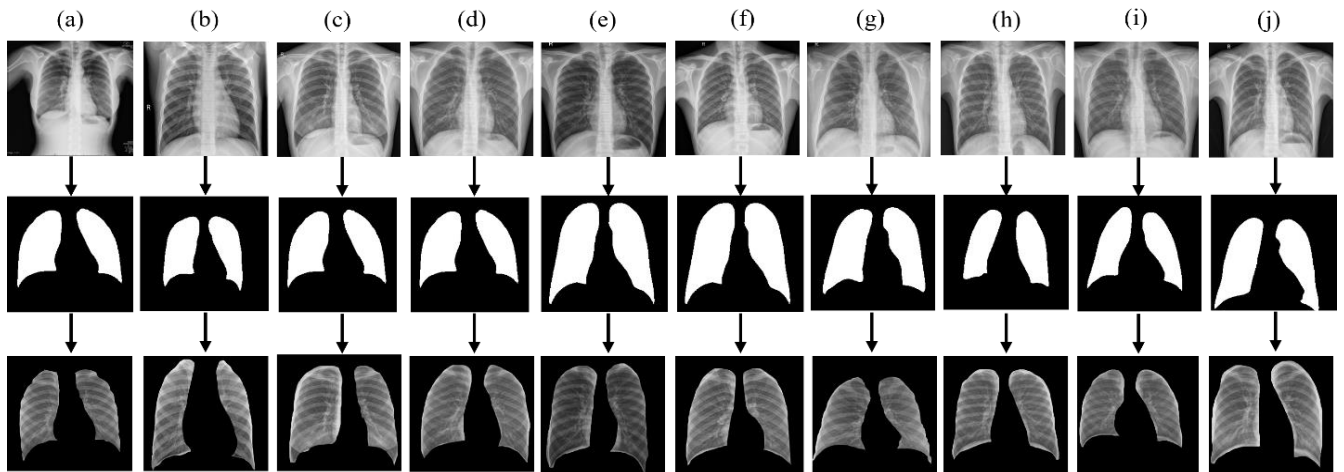
FIGURE 6. Info-MGAN is used for segmentation of CXR images of lungs; (a) COVID-19, (b) LC, (c) ATE, (d) COL, (e) TB, (f) PNET, (g) EDE, (h) PNEU, (i) PLT, and (j) Normal. From top to bottom, the first row shows the CXR images of ten chest diseases, the second row represents the predicted mask, and the last row shows the ROI obtained from the predicted mask using Info-MGAN.
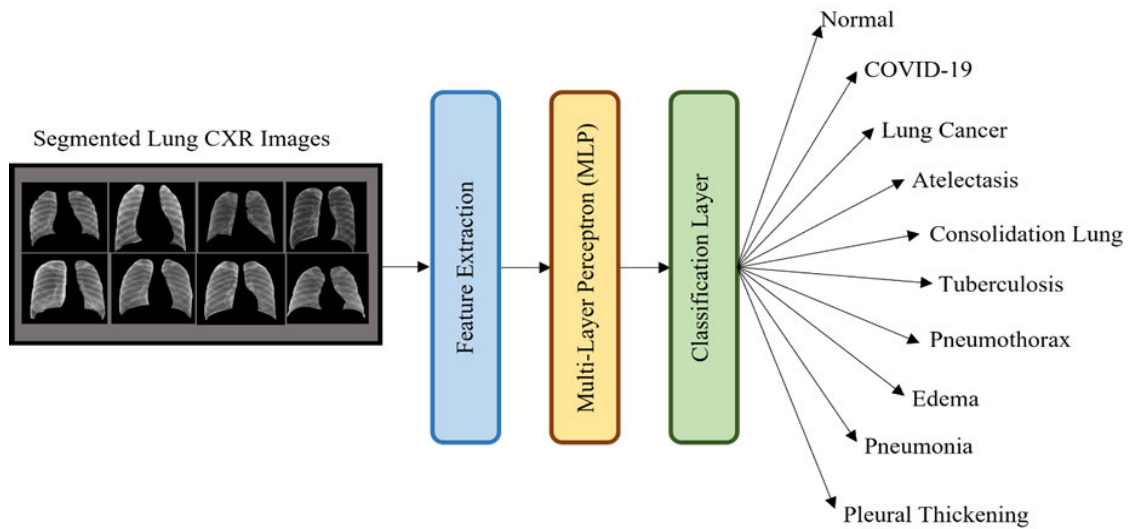
FIGURE 7. Proposed block diagram for the classification of ten chest diseases using CXR images.

network includes skipping connections that run between the encoder and the decoder to facilitate pattern reinforcement on several levels. The output that is produced by the final block of the discriminator is a shape or patch that has dimensions of 29 × 29, and each of the patch's pixels is responsible for classifying a separate portion of the input image. Training an Info-MGAN, which is effectively a U-net architecture [39] on a more general level, is the method that is used to complete the process of CXR image segmentation, as was covered in a previous section of this article. For the training of the generator network, the pix-to-pix approach was applied, and for the training of the discriminator network, the patch GAN-like classifier was put to use. After the training has been completed, the generator will be able to produce lung masks

that are suitable for the CXR pictures that have been loaded into it. Figure 5 displays the input CXR together with the ground truth mask and the lung mask that was produced by the Info-MGAN.

Both the ground truth lung mask and the predicted lung mask have a morphological form that is relatively comparable to one another, as is evident from the comparison that can be made between the two. The capture of images of the lungs is made possible by applying masks that were produced based on the CXR that were used as input. Figure 6 is an illustration of the image masks that are used in Posterior Anterior (PA) view CXRs to separate the lungs into their respective sections. Following this, the cleaned-up collection of lung images is put to use in the training of the proposed classification process.
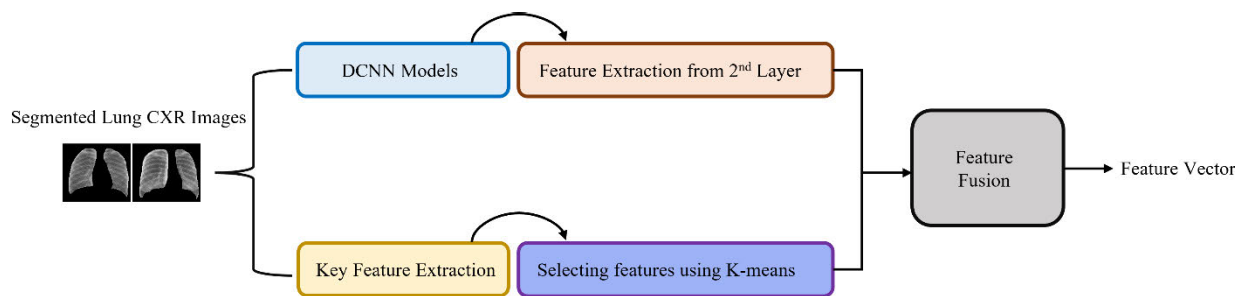
FIGURE 8. Process of extracting features using segmented CXR images.

C. EXTRACTION OF HIDDEN PATTERNS AND CLASSIFICATION OF CXR IMAGES OF CHEST DISEASES

In this section, we will investigate the process of automatically extracting crucial features from segmented lung images to arrive at an appropriate diagnosis of ten different chest diseases. The technique of extracting features from an image is one of the most critical processes in the image classification process [56]. It has been observed from the literature that conventional hand-crafted features can be useful for image classification. DL techniques greatly improve classification accuracy in contrast to standard feature extraction approaches. For the present study, a pipeline for in-hand image classification utilizes both DL architectures and more conventional feature extraction algorithms [57]. When deep architectures are employed separately, their classification performance is analyzed in detail. Figure 7 represents the pipeline that is utilized for the classification process. It is based on three blocks; the first block represents the feature extraction layer, the second block contains the MLP layer, and the last block is the final classification layer.

D. EXTRACTING FEATURES FOR TRAINING PURPOSES

Two key components are used for extracting features such as deep CNN and handcrafted features. The process of extracting features from the segmented lung images is shown in Figure 8. For this study, a total of eight DL models including simple CNN, ResNet-101, ResNet-50, DenseNet-169, DenseNet-201, Inception-v3, VGG-16, and VGG-19. The DenseNet models were chosen for this work because of the advantages discussed below. All layers are directly connected, the feature map's size is preserved, identity mapping attributes are integrated organically, it offers both shallow and deep supervision, and it may recycle previously used features. Conversely, VGG models with a depth of 16-19 weight layers and very small size (3 × 3) convolutional kernel showed a significant improvement over the state-of-the-art (SOTA) models in terms of classification accuracy and validation error [58]. Table 3 presents a detailed summary of the parameters used in implementing the DCNN models and Figure 9 illustrates the architecture of these models. In addition, we also designed a simple CNN model which contains five 2D convolutional layers (ConvL), five 2D max-pooling layers (MPL), and three flattened layers.

TABLE 3. Hyperparameters used for implementing DCNN models.

Parameters	Value
No. of Epochs	100
Learning rate (LR)	0.001
Activation Function	ReLU, Softmax
Loss Function	Cross entropy
Batch Size	32
Callbacks	Model validation, LR reduction on the plateau

The purpose of this research is to evaluate the effectiveness of simple CNN and TL models in the process of extracting discriminatory characteristics from lung image segments. The simple CNN, on the other hand, was trained from the scratch, in contrast to the TL models. After the DL model has been trained, the fully connected layer (FCL) is used to compute the features that will be output by the model. Note that the dimension of the computed feature vector expands to 128×1 when all 128 units in the FCL are taken into consideration. In the second stage of the feature extraction process, which is depicted in Figure 8, computer vision methods are applied to separate the most significant aspects of the image. These foci are often the elements of the image that have a blobby local appearance. ORB [50] and SURF [52] are two algorithms from the pipeline that we used individually to identify these properties in the segmented lung CXR.

The ORB is a method that is utilized extensively in the field of computer vision for key-point extraction and the development of feature descriptors. To comprehend the ORB algorithm [51], there are four phases involved. To begin, a difference-of-Gaussian function is applied to the complete CXR image to locate the areas of interest over the whole of the image. The subsequent step is known as "key point localization," and it involves the application of a precise model to determine the specific coordinates and size of each prospective key point [56], [57], [58]. By examining the stability metrics of the nodes, crucial nodes are identified. In the final step, directions of local CXR image gradients are employed to establish an orientation to each key point position [58]. In the fourth stage, we compute the local image gradients at the scale that you have selected at each critical point. Figure 10 (a) is a representation of the ORB method after it was applied to the masked lung CXR images as input.

FIGURE 9. Architecture of eight DCNN models.

TABLE 4. MLP model's layers were used for this study.

Layer Name	Output Shape	Activation Method
Dense_layer_01	1024	ReLU
Dense_layer_02	750	ReLU
Dense_layer_03	512	ReLU
Dense_layer_04	256	ReLU
Dense_layer_05	128	ReLU
Dense_layer_06	64	ReLU
Output_Layer_11	10	Softmax

FIGURE 10. Key feature extraction; (a) ORB method, (b) SURF.

SURF is a feature detector method [26] and is used to determine not just the relevance of features on a geographical scale but also on a pixel level. We can ascertain the dimensions of the key points in the continuous domain and locate them precisely due to the application of the quadratic function fitting technique. A sample pattern is used to describe the neighborhood surrounding each key point (points on scaled concentric rings).

Figure 10 (b) displays an example of the output that the SURF algorithm produces when it is given a masked lung X-ray picture as input. The k-means clustering approach is utilized to partition the important locations that were extracted from the images of the lungs into 64 separate categories. After that, we obtain a feature vector with a size of 64×1 by averaging the values of the descriptors that are contained within each class. After that, the feature vector (128×1) that was obtained from the FCL is attached to the feature vector that was previously described to raise its dimension to the maximum that the DL model will allow for it. The feature vector that is produced once the lungs have been masked has a size of 192×1. The aforementioned process is depicted as a block diagram in Figure 8. Next, as a component of our primary pipeline, we make use of a dense MLP model to feed the augmented feature vector (see Figure 7).

E. MLP MODEL
An MLP model is the following constructing element in this pipeline; it is the one that receives as input the combined characteristics from the model that came before it. Following this, the MLP model is trained to make predictions for the unseen right labels by using the labels that have been assigned to the input attributes as training data. The MLP model accepts as input a 192-element vector that is comprised of 128 features derived from the TL model's most recent iteration and 64 features derived from the mean points of the clusters in which the features were extracted utilizing either the ORB or SURF technique. The MLP model is comprised of seven deep layers, and there is no need for any additional hyper-parameters. In the first six layers of the network, ReLU activation functions were utilized. In the seventh and final layer, however, softmax activation functions were utilized. The principles of the MLP network's structure are discussed in Table 4.
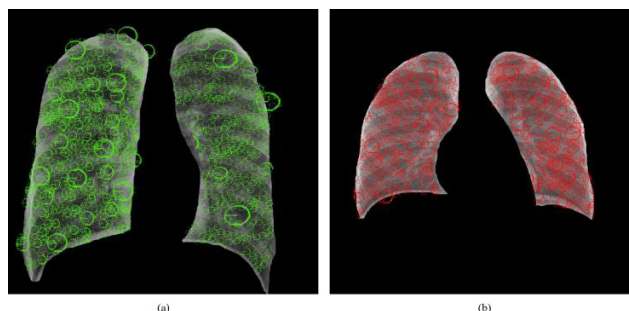
F. LAST LAYERS FOR PREDICTING THE OUTPUT
In this investigation, in addition to the softmax layer, we also made use of the quadratic support vector machine (Q-SVM) kernel [57], the AdaBoost [58], and the random forest (RF) [59]. These ML classifiers were utilized in a broad variety of investigations [55], [56], [57], [58], [59] as a result of their efficiency, ease, and affluence of execution in terms of health application. Therefore, it is believed that they are capable of exploiting the features obtained from the Dense layer (see Table 4) of the MLP network to classify lung images into the ten different classes that have been established.

G. PERFORMANCE EVALUATION
Using segmented CXR images, determine the classification performance of the deep features extraction (DFE) approaches to identify the 10 different chest diseases. Once all of the models have been trained, the confusion matrix-based performance parameters are computed using data from each stage of the proposed approach. The methodology designed in this study effectively integrates the handmade and DFE methods. On the testing dataset, the identification performance of the DCNN models is evaluated in terms of many parameters such as accuracy (ACU) [60], recall (REC) [61], false positive rate (FPR), F1-measure, specificity (SPF), precision (PRE), negative predicted value (NPV) [62], and false negative rate (FNR) [63]. The following equations (6-13) are used to measure these parameters.

$$ACU = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (6)$$

$$PRE = \frac{TP}{(TP + FP)} \quad (7)$$

$$REC = \frac{TP}{(TP + FN)} \quad (8)$$

$$SPF = \frac{TN}{(TN + FP)} \quad (9)$$

$$F1 - Score = \frac{2 \times (PRE \times REC)}{PRE + REC} \quad (10)$$

$$FPR = \frac{FP}{(FP + TN)} \quad (11)$$

$$FNR = \frac{FN}{(TP + FN)} \quad (12)$$

$$NPV = \frac{TN}{(TN + FN)} \qquad (13)$$

IV. RESULTS AND DISCUSSIONS

In this section, we show that the proposed approaches for extracting deep features efficaciously identified ten different chest diseases segmented CXR images. After the training of the model has been completed, the proposed method successfully incorporates both the handcrafted and the DFE methods by computing performance parameters based on the confusion matrix.

A. EXPERIMENTAL SETUP

The proposed method was implemented with the assistance of the Keras library [64]. Python [65] was used as the programming language for the procedures that did not have a direct connection to the CNN. The experiment was carried out using a computer equipped with a Windows operating system, an NVIDIA GeForce GTX GPU with 11 GB of memory, and 32 GB of total RAM.

B. TRAINING OF DCNN MODELS

To build the DCNN, we first trained a segmentation network on previously acquired CXR images, then utilized the generator model [66], [67], [68], [69] to segment the new data. Image augmentation was undertaken, which involved making the affine transformation to the segmented CXR images (such as rotations and shears), so that deep classification models could be adequately trained to make use of a large training dataset. Eight different DCNN models were investigated as DFE methods, and their performance was evaluated by combining these models with ORB and SURF-based feature extraction algorithms. The entire dataset was analyzed by using these eight different DCNN models, and the process used to do so was called 10-fold cross-validation. The last layer of classifications was made using several different ML methods, including Q-SVM, AdaBoost, and RF. Loss function [70], [71], [72] convergence for the DCNN models is displayed in Figure 11. We observe from the plots (see Figure 11) that the six DCNN models converge more quickly than the simple CNN model trained from scratch. The present work reports and analyses the training-phase execution times of the DFE models (shown in Table 5) to evaluate their computational complexity. The training of the ResNet-50 model took a total of 1349 seconds (s), with an average training time of 152.3 milliseconds (ms) per step. The DenseNet-169 model training was finished in a total of 2157s, with an average training time per step of 199.5 ms. For inception-v3, the overall training duration of 1239s and an average per-step training time of 141.5 ms.

The longest amount of time spent training a model was 2399s for the DenseNet-201, with an average training time per step of 201.7 ms. The average training time per step for the VGG-16 model was 122.6 ms, and the total training time was 811s. At last, the simple CNN model contains the lowest

TABLE 5. Amount of time required for training DCNN models.

Sr#	Models	Model Training Time	Average time per step
1	ResNet-50	1349s	152.3 ms
2	DenseNet-169	2157s	199.5 ms
3	Inception-v3	1239s	141.5 ms
4	DenseNet-201	2399s	201.7 ms
5	VGG-16	811s	122.6 ms
6	Simple CNN	405s	69.7 ms
7	VGG19	991s	133.2 ms
8	ResNet-101	789s	165.7 ms

training time of 405s, with an average training time per step of 69.7 ms.

C. RESULTS ANALYSIS OF DFE BY DCNN WITH SEVERAL ML MODELS

After extracting features, the last layers of eight DCNN models were replaced with softmax and several ML models such as Q-SVM, AdaBoost, and RF. To measure the classification evaluation of these models, several metrics such as ACU, REC, NPV, F1-measure, SPF, and PRE were considered for the diagnosis of ten chest diseases. The results of the DFE from eight different DCNN models are presented in Table 6, and the feature matrices generated by these models were categorized by several different ML models. From Table 6, we can see that the six DCNN models outperformed the simple CNN model in terms of ACU, REC, NPV, F1-measure, SPF, and PRE. Further, while comparing the AdaBoost classifier to the other ML classifiers, it is clear that it has the highest obtained classification performance characteristics. The best classification performance is seen when deep characteristics that are taken from the VGG-19 model are fed to the softmax. In this scenario, the values for ACU, REC, NPV, F1-measure, SPF, and PRE are obtained as follows: 96.97%, 96.92%, 96.93%, 96.92%, 96.97%, and 96.97%, respectively. The simple CNN model that used AdaBoost as the last layer had an ACU score of 94.30%, a REC score of 94.42%, an NPV score of 94.26%, an F1-measure score of 94.25%, an SPF score of 94.29%, and a PRE score of 94.28%. It has been discovered that the performance of some other DCNN models is on par with that of the best-performing VGG-19 model.

D. RESULTS ANALYSIS OF DCNN MODELS WITH FEATURES EXTRACTED BY ORB AND SURF

After fusing deep features with the hand-crafted features produced by the ORB and SURF algorithms respectively, the resulting performance parameters are given in Tables 7 and 8. According to Table 7, the incorporated SURF algorithm-based features did not result in a significant shift in any of the acquired classification performance parameters. With Q-SVM as the classification layer, the ACU, REC, NPV, F1-measure, SPF, and PRE scores were 97.11%, 96.81%, 96.48%, 96.23%, 96.51, and 96.25%, respectively, when DFE from the DenseNet-201 model were fused with SURF-based features. Furthermore, while utilizing softmax
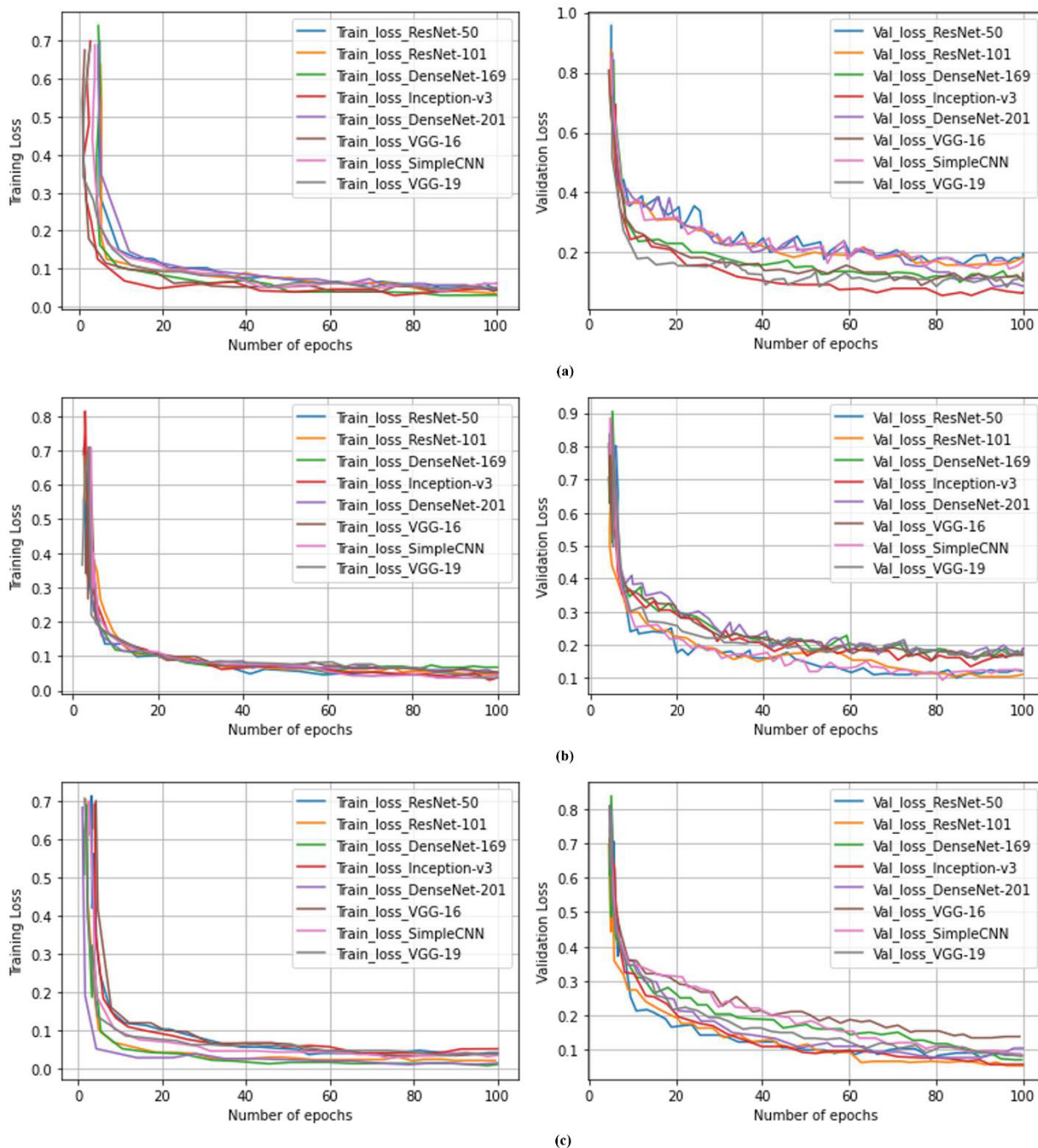
FIGURE 11. DCNN models loss function convergence plots; (a) DCNN models with several ML classifiers; (b) DCNN models using SURF-based handcrafted features and several ML models; (c) DCNN models using ORB-based handcrafted features and several ML models.

as the classification layer, the ACU, REC, NPV, F1-measure, SPF, and PRE scores obtained from combining DFE from the VGG-19 model with SURF-based features were 96.90%, 96.54%, 96.42%, 96.51, 96.52, and 96.50%, respectively.

Table 8 indicates that classification performance was significantly enhanced by combining deep features with ORB-based key-point features and then feeding these features to several ML algorithms. The VGG-19 model-based DFE

and the ORB-based features increased ACU, REC, NPV, F1-measure, SPF, and PRE scores by 98.20%, 97.98%, 97.96%, 97.95%, 97.99, and 98.04%, respectively, using soft-max as the final layer. These values are approximately 2% higher than they were when a softmax layer was only used to classify deep features that were based on VGG-19. The ACU, REC, NPV, F1-measure, SPF, and PRE scores for the DenseNet-169 model-based deep features, the ORB features,

TABLE 6. Performance of the DFE from DCNN models and several ML models.

Sr#	Models	Last Layer	ACU (%)	REC (%)	NPV (%)	F1-measure (%)	SPF (%)	PRE (%)
1	ResNet-50	Softmax	95.39	95.85	95.96	95.65	95.15	95.57
		Q-SVM	94.31	94.81	94.76	94.29	94.35	94.75
		AdaBoost	95.29	95.71	95.72	95.36	95.32	95.57
		RF	94.31	94.81	94.76	94.29	94.35	94.75
2	ResNet-101	Softmax	95.40	95.87	95.97	95.60	95.25	95.27
		Q-SVM	94.29	94.75	94.76	94.31	94.34	94.75
		AdaBoost	95.29	95.75	95.79	95.36	95.32	95.57
		RF	94.34	94.81	94.75	94.22	94.23	94.75
3	DenseNet-169	Softmax	96.10	95.99	95.94	95.97	96.00	95.85
		Q-SVM	95.92	95.96	95.92	95.96	95.99	95.95
		AdaBoost	94.99	95.15	95.24	95.34	95.58	95.58
		RF	95.36	95.86	95.81	95.34	95.40	95.80
4	Inception-v3	Softmax	93.99	93.85	93.95	93.99	93.91	93.96
		Q-SVM	94.01	94.15	94.05	94.19	94.21	94.19
		AdaBoost	94.08	94.22	94.17	94.29	94.32	94.33
		RF	94.10	94.20	94.21	94.26	94.30	94.31
5	DenseNet-201	Softmax	94.21	94.37	94.41	94.43	94.45	94.44
		Q-SVM	94.19	94.23	94.25	94.26	94.22	94.26
		AdaBoost	94.07	93.99	93.87	93.97	93.99	94.01
		RF	94.15	94.12	94.11	94.10	94.14	94.13
6	VGG-16	Softmax	95.09	95.12	95.13	95.19	95.20	95.21
		Q-SVM	95.05	95.01	95.09	95.07	95.10	95.11
		AdaBoost	94.93	95.08	95.12	95.22	95.37	95.48
		RF	95.18	95.64	95.62	95.17	95.21	95.39
7	Simple CNN	Softmax	94.10	93.10	93.88	93.89	93.99	93.96
		Q-SVM	94.09	94.10	94.38	94.39	94.49	94.46
		AdaBoost	94.30	94.42	94.26	94.25	94.29	94.28
		RF	94.25	94.35	94.16	94.15	94.18	94.16
8	VGG-19	Softmax	96.97	96.92	96.93	96.92	96.97	96.97
		Q-SVM	96.92	96.88	96.86	96.84	96.89	96.87
		AdaBoost	96.94	96.90	96.91	96.92	96.95	96.94
		RF	96.90	96.85	96.82	96.79	96.83	96.80

and the softmax classifier in the final layer were reported to be 96.15%, 96.04%, 95.99%, 96.02%, 96.05, and 95.90%, respectively.

ACU, REC, NPV, F1-measure, SPF, and PRE scores were recorded at 94.27%, 94.49%, 94.51%, 94.63, and 94.52%, respectively, while using the DenseNet-201 model's deep features, together with the ORB features and the softmax classifier in the final layer. Out of all the models tested, it was the ORB VGG-19 model with softmax layer that achieved the highest accuracy rates in classification. The detailed results are presented in Table 8.

Figures 12-14 depict average FPR and FNR bar plots, which allow for further investigation into the categorization accuracy of the approach. To get higher classification performance, it is necessary to significantly lower the FPR and FNR values.

Figure 12 shows that the lowest FPR value (3.85%) was achieved by VGG-16 when the Q-SVM layer was used as the last one. The VGG-16 softmax function applied to the last layer produces an FNR that is 6.35% lower than any other. The FPR and FNR for VGG-19 with softmax are both significantly lower than average, coming in at 1.81 and 4.33 percent,

respectively. According to Figure 13, the combination of ResNet-101 and VGG-16 with SURF and Q-SVM in the final layer achieved the lowest FPR and FNR of 4.85% and 5.71%, respectively. Figure 14 represents that VGG-19 with ORB achieved the lowest FPR and FNR of 1.10% and 2.33%, respectively, when using softmax as the final layer, making it the best combination of the proposed feature extraction models. To further illustrate the class-specific performance of our proposed feature extraction models, we present the confusion matrices for a total of 7000 test CXR images in Figure 15.

E. STATISTICAL ANALYSIS

The feasibility of the suggested model was evaluated using McNemar's statistical test [67] and the analysis of variance (ANOVA) test [68]. This was done by comparing the proposed model to the base classifiers, the probability scores of which were utilized in the process of determining the development of the suggested model. Table 9 presents the findings of McNemar's, and ANOVA tests carried out on the multi-chest disease CXR dataset that was used in the present analysis. To conclude that the alternative hypothesis is

TABLE 7. Experimental outcomes of DCNN models using SURF-based handcrafted features and several ML models.

Sr#	Models	Last Layer	ACU (%)	REC (%)	NPV (%)	F1-measure (%)	SPF (%)	PRE (%)
1	SURF ResNet-50	Softmax	95.44	95.91	95.97	95.65	95.21	95.63
		Q-SVM	95.43	95.90	95.96	95.65	95.21	95.60
		AdaBoost	95.44	95.91	95.97	95.65	95.21	95.63
		RF	95.44	95.91	95.97	95.65	95.21	95.63
2	SURF ResNet-101	Softmax	95.54	95.93	95.96	95.61	95.31	95.53
		Q-SVM	95.23	95.80	95.86	95.55	95.01	95.50
		AdaBoost	95.45	95.90	95.96	95.56	95.21	95.63
		RF	95.44	95.91	95.97	95.56	95.19	95.53
3	SURF DenseNet-169	Softmax	96.13	96.02	95.97	96.01	96.05	96.00
		Q-SVM	96.12	96.02	95.97	96.01	96.05	96.00
		AdaBoost	96.13	96.01	95.96	96.01	96.05	96.00
		RF	96.11	96.02	95.97	96.01	96.05	96.00
4	SURF Inception-v3	Softmax	94.40	91.56	95.18	92.13	94.56	95.27
		Q-SVM	94.40	91.57	95.19	92.15	94.53	95.26
		AdaBoost	94.50	91.56	95.19	92.13	94.49	95.26
		RF	94.50	91.56	95.20	92.14	94.52	95.26
5	SURF DenseNet-201	Softmax	97.10	96.79	96.46	96.23	96.52	96.25
		Q-SVM	97.11	96.81	96.48	96.23	96.51	96.25
		AdaBoost	97.09	96.80	96.48	96.24	96.51	96.25
		RF	97.06	96.80	96.48	96.23	96.51	96.25
6	SURF VGG-16	Softmax	94.0	94.26	94.30	95.30	94.89	95.10
		Q-SVM	94.10	94.26	94.35	95.30	94.89	95.10
		AdaBoost	94.11	94.27	94.35	95.31	94.89	95.10
		RF	94.10	94.27	94.35	95.31	94.90	95.10
7	SURF Simple CNN	Softmax	95.27	95.81	95.79	95.78	95.99	95.62
		Q-SVM	95.26	95.81	95.81	95.78	95.99	95.62
		AdaBoost	95.26	95.82	95.79	95.78	95.92	95.60
		RF	95.30	95.82	95.79	95.82	95.92	95.60
8	SURF VGG-19	Softmax	96.90	96.54	96.42	96.51	96.52	96.50
		Q-SVM	96.63	96.55	96.42	96.53	96.52	96.51
		AdaBoost	96.87	96.55	96.42	96.53	96.52	96.50
		RF	96.87	96.55	96.45	96.52	96.55	96.51

more likely to be true, McNemar's test and the ANOVA need to have P-Values that are less than 0.05. The results shown in Table 9 show that the p values for all of the dataset samples fall below the 0.05 threshold. The null hypothesis cannot be maintained, according to the results of both of the statistical tests. This demonstrates that the proposed model contains more information from the base classifiers and that its predictions are better, ensuring that it is statistically distinct from the other contributing models. Additionally, this demonstrates that the proposed model contains more information from the base classifiers.

F. COMPARISON WITH MODERN EXISTING METHODS
As shown in Table 10, we compared our best-performing model (ORB & VGG-19 with a softmax layer) to other existing state-of-the-art DCNN methods to classify 10 different chest diseases using CXR images. ACU, REC, and SPF values of 95.11 percent, 93.15 percent, and 96.5 percent, respectively, were obtained through the fusion of features derived from local binary patterns (LBP) with the DFE by utilizing the Inception V3 architecture [56]. These values were derived from images that had been filtered with a Gaussian filter. In terms of classification, features based on MobileNet-v2 have an ACU of 93.33 percent, a REC of 90.66 percent, and an SPF of 95.22 percent. The study [57] introduced a novel model for the classification of COVID-19 using CXR images and they achieved the ACU, REC, and SPF of 95.72, 93.59, and 96.78 percent, respectively. In contrast, our suggested framework integrating the VGG-19 model, the ORB feature extraction method, and the softmax classifier outperformed the existing methods for classifying the 10 different chest diseases including COVID-19, LC, ATE, COL, TB, PNET, EDE, PNEU, PLT, and normal (see Table 10). In addition, it is important to note that the proposed DFE framework classifies ten distinct chest disorders using segmented lung pictures as input. In contrast to the approaches already in use that are summarized in Table 10 and compute features straight from the CXR images, this approach does not work.

G. DISCUSSIONS
In this work, we investigated all three stages such as image segmentation, feature extraction, and classification that are

FIGURE 12. FNR and FPR results obtained after DFE from DCNN with different ML classifiers; a) VGG-16, b) ResNet-50, c) ResNet-101, d) Inception-v3, e) DenseNet-169, f) DenseNet-201, g) Simple CNN, and h) VGG-19.

FIGURE 13. FNR and FPR results obtained after DFE from DCNN and SURF feature with different ML; a) VGG-16, b) ResNet-50, c) ResNet-101, d) Inception-v3, e) DenseNet-169, f) DenseNet-201, g) Simple CNN, and h) VGG-19.
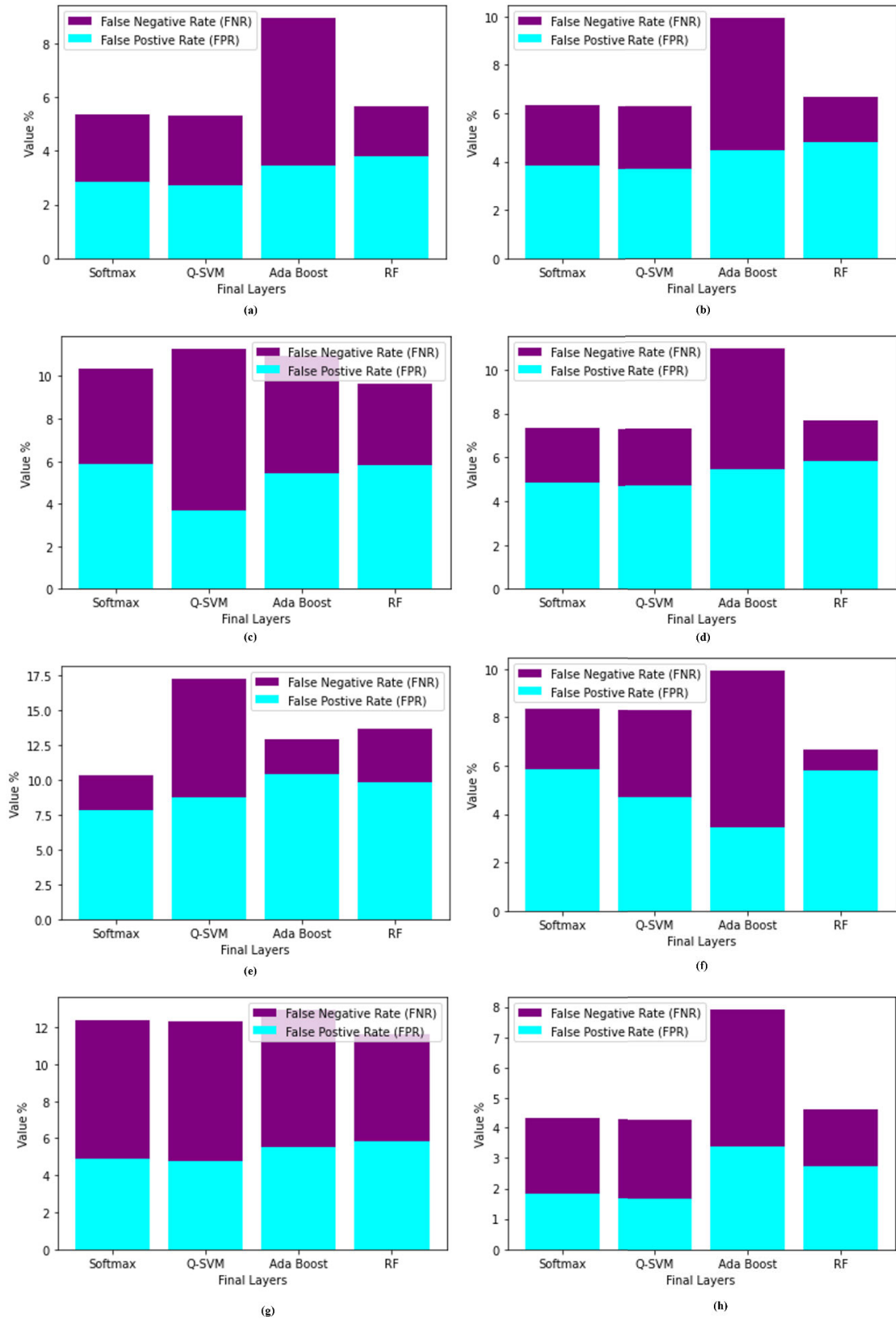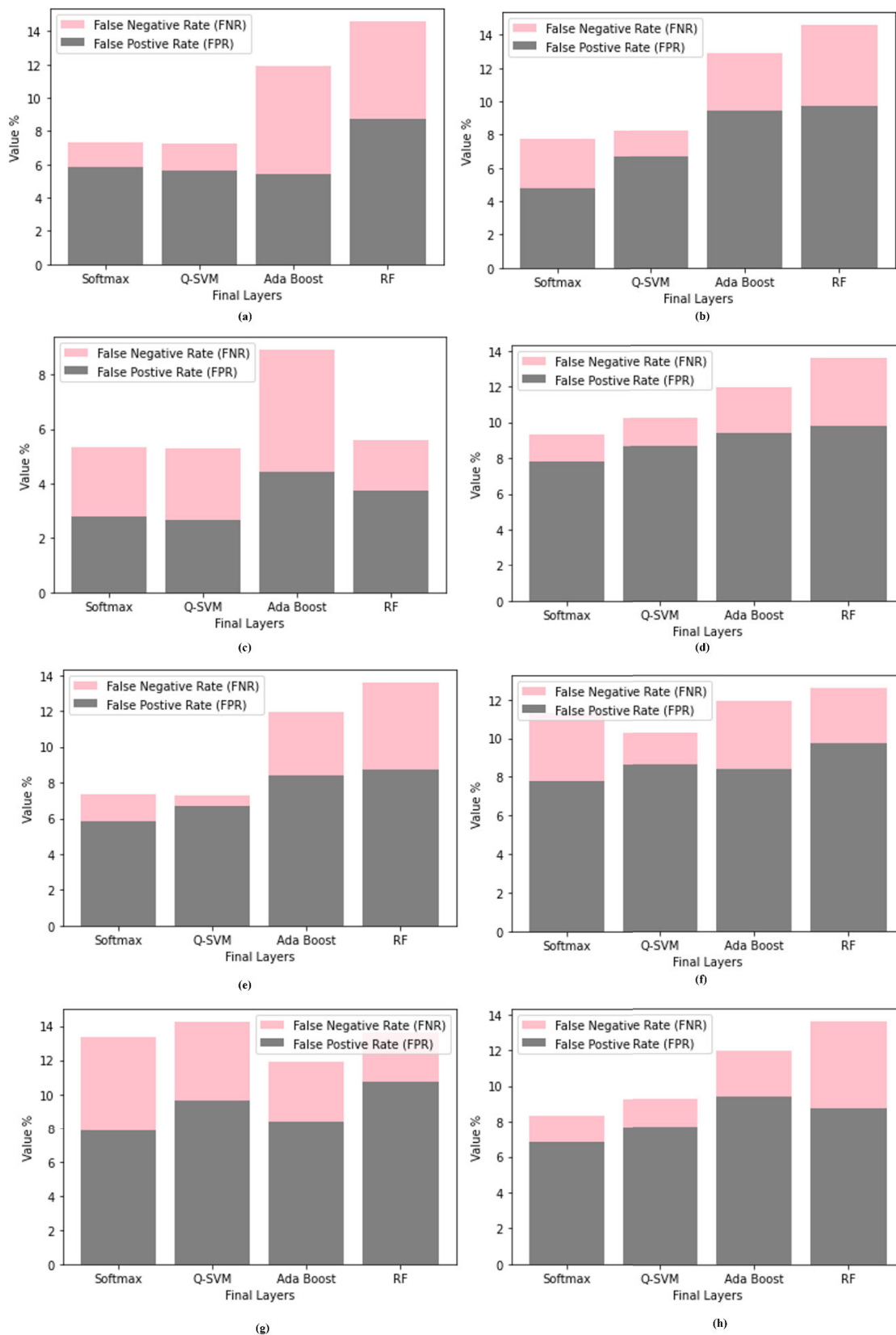
FIGURE 14. FNR and FPR results obtained after DFE from DCNN and ORB feature with different ML; a) VGG-16, b) ResNet-50, c) ResNet-101, d) Inception-v3, e) DenseNet-169, f) DenseNet-201, g) Simple CNN, and h) VGG-19.
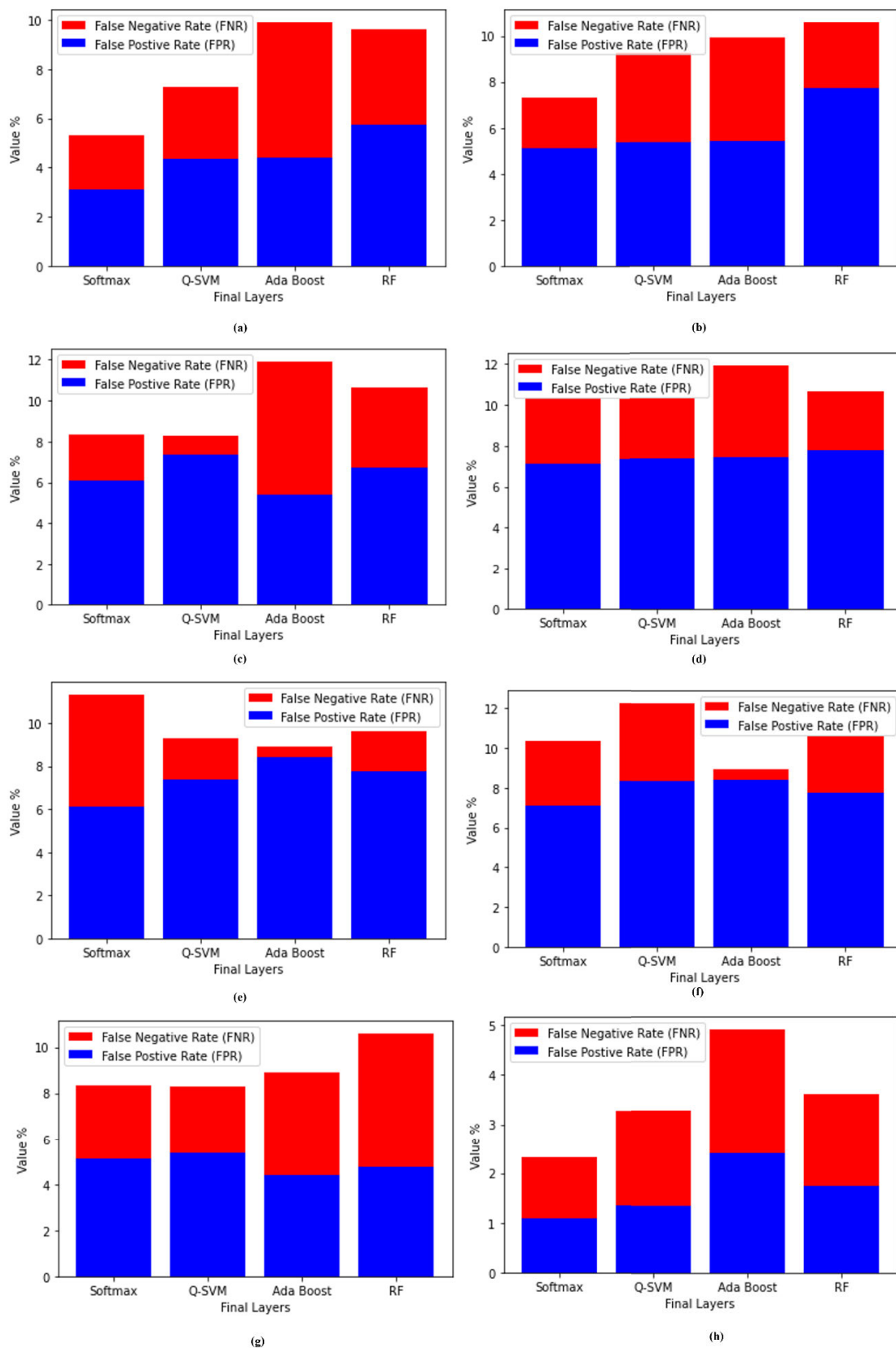
TABLE 8. Results of experiments using DCNN models that combine ORB-based features with a variety of ML models.

Sr#	Models	Last Layer	ACU (%)	REC (%)	NPV (%)	F1-measure (%)	SPF (%)	PRE (%)
1	ORB ResNet-50	Softmax	95.41	95.87	95.98	95.67	95.17	95.59
		Q-SVM	94.34	94.84	94.79	94.32	94.38	94.78
		AdaBoost	95.33	95.75	95.76	95.40	95.36	95.61
		RF	94.35	94.85	94.80	94.29	94.35	94.75
2	ORB ResNet-101	Softmax	95.31	95.81	95.95	95.65	95.21	95.61
		Q-SVM	94.30	94.82	94.75	94.31	94.29	94.76
		AdaBoost	95.29	95.74	95.70	95.29	95.38	95.71
		RF	94.39	94.81	94.79	94.31	94.39	94.74
2	ORB DenseNet-169	Softmax	96.15	96.04	95.99	96.02	96.05	95.90
		Q-SVM	95.97	96.01	95.97	96.01	96.04	96.00
		AdaBoost	95.04	95.20	95.29	95.39	95.63	95.63
		RF	95.41	95.91	95.86	95.39	95.45	95.85
3	ORB Inception-v3	Softmax	94.05	93.91	94.01	94.05	93.97	93.02
		Q-SVM	94.07	94.21	94.11	94.25	94.27	94.25
		AdaBoost	94.14	94.28	94.23	94.35	94.38	94.39
		RF	94.16	94.26	94.27	94.32	94.36	94.37
4	ORB DenseNet-201	Softmax	94.27	94.45	94.49	94.51	94.63	94.52
		Q-SVM	94.27	94.31	94.33	94.26	94.22	94.26
		AdaBoost	94.07	93.99	93.87	93.97	93.99	94.01
		RF	94.15	94.15	94.11	94.10	94.14	94.13
5	ORB VGG-16	Softmax	95.19	95.15	95.23	95.19	95.20	95.21
		Q-SVM	95.15	95.13	95.19	95.07	95.10	95.11
		AdaBoost	94.83	95.19	95.19	95.22	95.37	95.48
		RF	95.28	95.64	95.78	95.17	95.21	95.39
6	ORB Simple CNN	Softmax	94.10	93.10	93.88	93.89	93.99	93.96
		Q-SVM	94.09	94.10	94.38	94.39	94.49	94.46
		AdaBoost	94.30	94.42	94.26	94.25	94.29	94.28
		RF	94.25	94.35	94.16	94.15	94.18	94.16
7	ORB VGG-19	Softmax	98.20	97.98	97.96	97.95	97.99	98.04
		Q-SVM	97.99	96.85	96.84	96.33	96.61	96.65
		AdaBoost	98.01	96.85	96.88	96.95	96.87	96.94
		RF	97.69	97.01	97.00	96.99	96.98	96.99

TABLE 9. Results of McNemar's and ANOVA test of the proposed model.

Sr#	Statistical Analysis	P-Value
1	ANOVA Test	0.0019
2	McNemar's Test	0.0228

necessary for accurate image classification [5], [6], [7], [8], [12], [16], [73], [74], [75], [76], [77]. The segmentation process began with the use of fundamental supervised learning models, during which we investigated a variety of U-net designs. Despite this, the Info-MGAN model produced the most significant outcomes when compared to the other methods of supervised learning. A more extensive dataset of TB and COVID-19 CXR images was not readily available, so the SMOTE technique was utilized to increase the number of images belonging to the minority groups. The development of a feature extraction pipeline (see Figure 8) has been made by the integration of handcrafted feature [77] extraction algorithms with DCNN models. It was shown that key point descriptor techniques are effective at extracting object intensities from a CXR image [64], [66], [71], [72], [73], [74], [75]. To accomplish this goal, we make use of

the key point descriptors (namely SURF and ORB) to locate candidate key intensity points that significantly contribute to the classification of CXR segments. Every single one of these models has demonstrated impressive accuracy, in addition to a convergence of nature and noticeably deeper coverage. DL function (such as softmax) and other ML techniques (such as Q-SVM, AdaBoost, & RF are used to categorize the computed features in the last layer. Both of the offered ways of segmentation and classification have been fine-tuned using data that is publically available [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72]. We have offered a comprehensive evaluation of the usefulness of the proposed structure, examining its performance both with and without the concentration key point elements being taken into consideration. The findings of the experiments reveal that the DCNN models work more effectively than the simple CNN model consistently [78]. The results of this study substantially indicate the effectiveness of the suggested method (ORB & VGG-19 with softmax) in classifying ten different chest diseases using CXR images.

This technique can be utilized by radiologists as a supplementary resource during the diagnostic process of patient
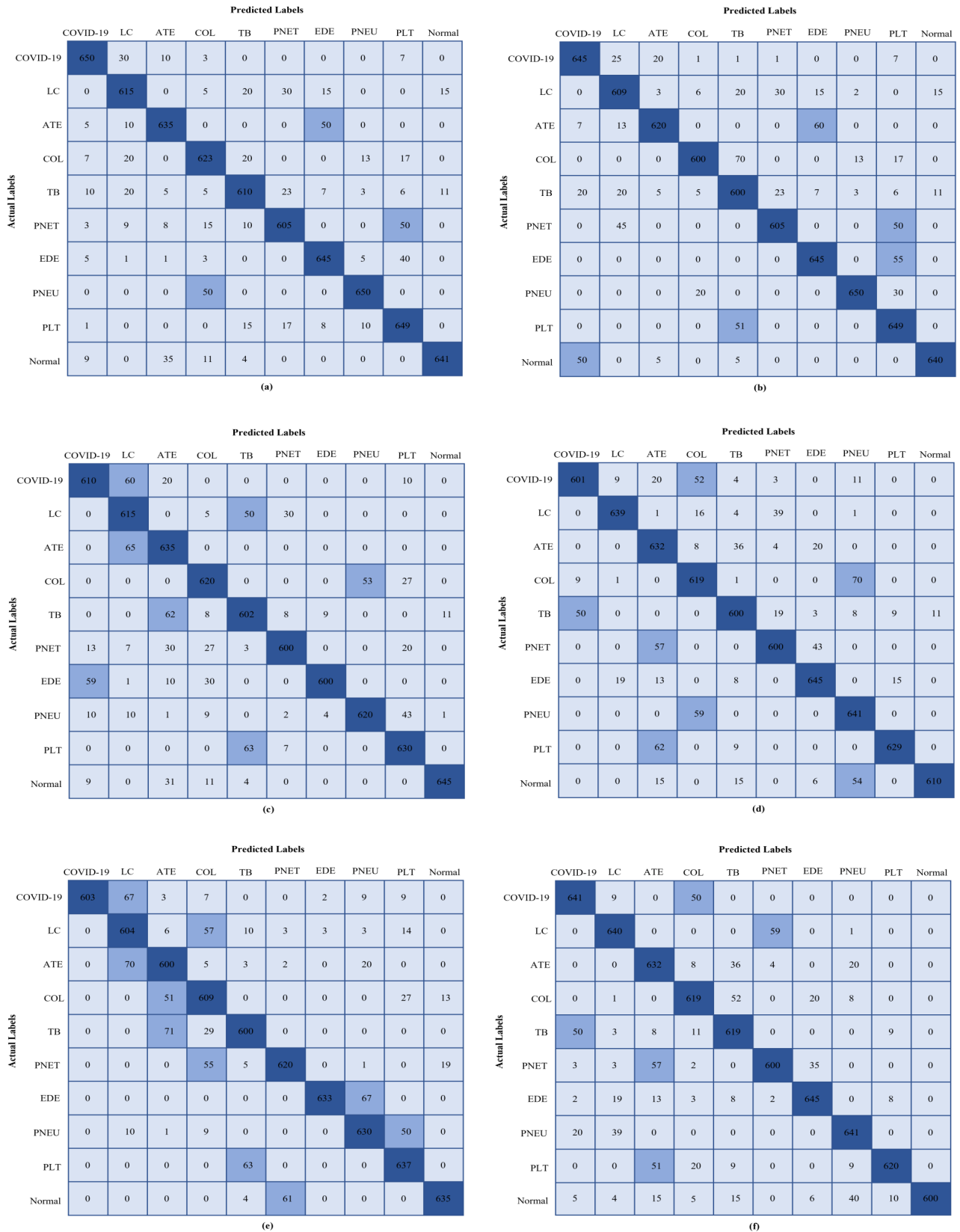
(a) ResNet-50 and ORB — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	650	30	10	3	0	0	0	0	7	0
LC	0	615	0	5	20	30	15	0	0	15
ATE	5	10	635	0	0	0	50	0	0	0
COL	7	20	0	623	20	0	0	13	17	0
TB	10	20	5	5	610	23	7	3	6	11
PNET	3	9	8	15	10	605	0	0	50	0
EDE	5	1	1	3	0	0	645	5	40	0
PNEU	0	0	0	50	0	0	0	650	0	0
PLT	1	0	0	0	15	17	8	10	649	0
Normal	9	0	35	11	4	0	0	0	0	641

(b) ResNet-101 and SURF — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	645	25	20	1	1	1	0	0	7	0
LC	0	609	3	6	20	30	15	2	0	15
ATE	7	13	620	0	0	0	60	0	0	0
COL	0	0	0	600	70	0	0	13	17	0
TB	20	20	5	5	600	23	7	3	6	11
PNET	0	45	0	0	0	605	0	0	50	0
EDE	0	0	0	0	0	0	645	0	55	0
PNEU	0	0	0	20	0	0	0	650	30	0
PLT	0	0	0	0	51	0	0	0	649	0
Normal	50	0	5	0	5	0	0	0	0	640

(c) Simple CNN and ORB — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	610	60	20	0	0	0	0	0	10	0
LC	0	615	0	5	50	30	0	0	0	0
ATE	0	65	635	0	0	0	0	0	0	0
COL	0	0	0	620	0	0	0	53	27	0
TB	0	0	62	8	602	8	9	0	0	11
PNET	13	7	30	27	3	600	0	0	20	0
EDE	59	1	10	30	0	0	600	0	0	0
PNEU	10	10	1	9	0	2	4	620	43	1
PLT	0	0	0	63	7	0	0	0	630	0
Normal	9	0	31	11	4	0	0	0	0	645

(d) DenseNet-169 — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	601	9	20	52	4	3	0	11	0	0
LC	0	639	1	16	4	39	0	1	0	0
ATE	0	0	632	8	36	4	20	0	0	0
COL	9	1	0	619	1	0	0	70	0	0
TB	50	0	0	0	600	19	3	8	9	11
PNET	0	0	57	0	0	600	43	0	0	0
EDE	0	19	13	0	8	0	645	0	15	0
PNEU	0	0	0	59	0	0	0	641	0	0
PLT	0	0	62	0	9	0	0	0	629	0
Normal	0	0	15	0	15	0	6	54	0	610

(e) DenseNet-201 and SURF — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	603	67	3	7	0	0	2	9	9	0
LC	0	604	6	57	10	3	3	3	14	0
ATE	0	70	600	5	3	2	0	20	0	0
COL	0	0	51	609	0	0	0	0	27	13
TB	0	0	71	29	600	0	0	0	0	0
PNET	0	0	0	55	5	620	0	1	0	19
EDE	0	0	0	0	0	0	633	67	0	0
PNEU	0	10	1	9	0	0	0	630	50	0
PLT	0	0	0	0	63	0	0	0	637	0
Normal	0	0	0	0	4	61	0	0	0	635

(f) Inception-v3 and ORB — Predicted Labels

Actual \ Predicted	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	641	9	0	50	0	0	0	0	0	0
LC	0	640	0	0	0	59	0	1	0	0
ATE	0	0	632	8	36	4	0	20	0	0
COL	0	1	0	619	52	0	20	8	0	0
TB	50	3	8	11	619	0	0	0	9	0
PNET	3	3	57	2	0	600	35	0	0	0
EDE	2	19	13	3	8	2	645	0	8	0
PNEU	20	39	0	0	0	0	0	641	0	0
PLT	0	0	51	20	9	0	0	9	620	0
Normal	5	4	15	5	15	0	6	40	10	600

FIGURE 15. Confusion matrix for the different proposed feature extraction networks: (a) ResNet-50 and ORB (the last layer is Q-SVM), (b) ResNet-101 and SURF with final layer is softmax, (c) Simple CNN and ORB with the last layer are RF, (d) DenseNet-169 and last layer is AdaBoost, (e) DenseNet-201 and SURF with final layer is RF, (f) Inception-v3 and ORB (the last layer is Q-SVM), (g) VGG-16 and SURF with the last layer is softmax, and (h) VGG-19 and ORB with the last layer is softmax.
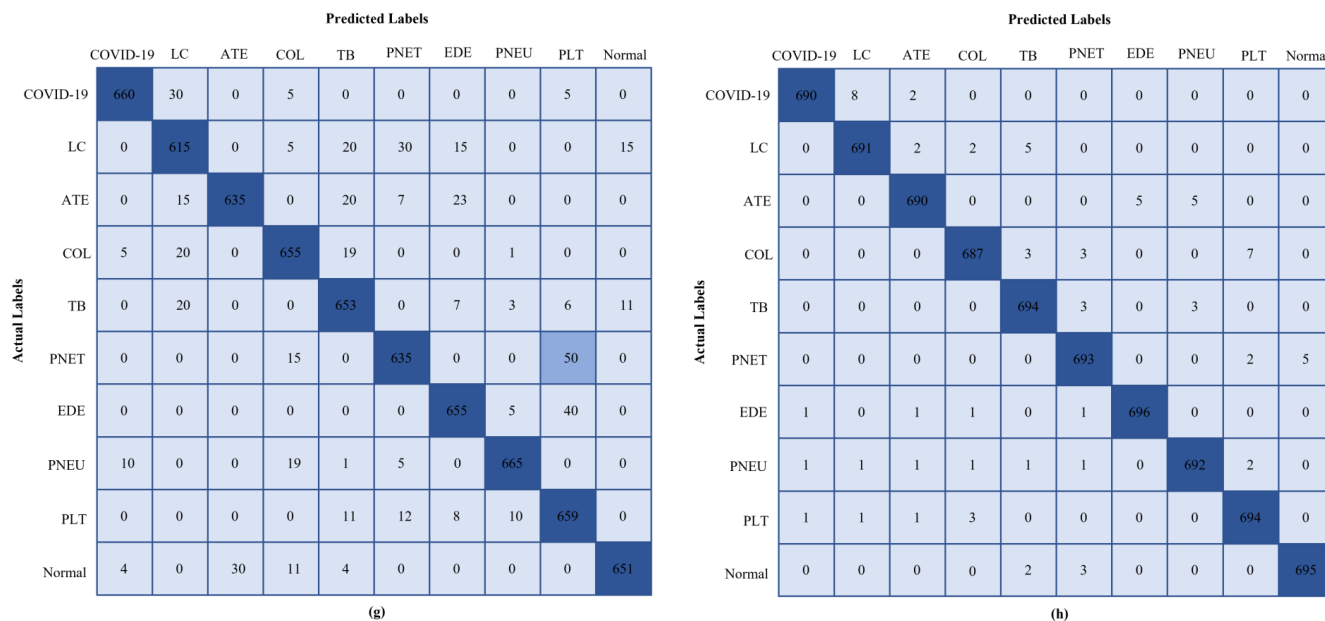
Predicted Labels (g)

Actual Labels	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	660	30	0	5	0	0	0	0	5	0
LC	0	615	0	5	20	30	15	0	0	15
ATE	0	15	635	0	20	7	23	0	0	0
COL	5	20	0	655	19	0	0	1	0	0
TB	0	20	0	0	653	0	7	3	6	11
PNET	0	0	0	15	0	635	0	0	50	0
EDE	0	0	0	0	0	0	655	5	40	0
PNEU	10	0	0	19	1	5	0	665	0	0
PLT	0	0	0	0	11	12	8	10	659	0
Normal	4	0	30	11	4	0	0	0	0	651

Predicted Labels (h)

Actual Labels	COVID-19	LC	ATE	COL	TB	PNET	EDE	PNEU	PLT	Normal
COVID-19	690	8	2	0	0	0	0	0	0	0
LC	0	691	2	2	5	0	0	0	0	0
ATE	0	0	690	0	0	0	5	5	0	0
COL	0	0	0	687	3	3	0	0	7	0
TB	0	0	0	0	694	3	0	3	0	0
PNET	0	0	0	0	0	693	0	0	2	5
EDE	1	0	1	1	0	1	696	0	0	0
PNEU	1	1	1	1	1	1	0	692	2	0
PLT	1	1	1	3	0	0	0	0	694	0
Normal	0	0	0	0	2	3	0	0	0	695

FIGURE 15. *(Continued.)* Confusion matrix for the different proposed feature extraction networks: (a) ResNet-50 and ORB (the last layer is Q-SVM), (b) ResNet-101 and SURF with final layer is softmax, (c) Simple CNN and ORB with the last layer are RF, (d) DenseNet-169 and last layer is AdaBoost, (e) DenseNet-201 and SURF with final layer is RF, (f) Inception-v3 and ORB (the last layer is Q-SVM), (g) VGG-16 and SURF with the last layer is softmax, and (h) VGG-19 and ORB with the last layer is softmax.

TABLE 10. Comparison of the proposed model with recent work.

Authors	Year	Methods	ACU	REC	SPF
Lasker et al. [58]	2022	LWSNet	97.28%	96.50%	-
Gupta et al. [59]	2022	VMD	97.30%	96.97%	-
Khan et al. [60]	2022	EfficientNetB1	92.0%	91.75%	94.50%
Ukwuoma et al. [61]	2022	Dual_Pachi	96.65%	93.90%	93.31%
Li et al. [62]	2021	Mag-SD	92.35%	92.50%	92.20%
Shankar & Perumal [56]	2021	LBP & Inception-v3	95.11%	93.15%	96.5%
Apostolopoulos & Mpesiana [25]	2020	MobileNet-v2	93.33%	90.66%	95.23%
Pham [57]	2021	AlexNet	95.72%	93.59%	96.78%
Proposed Model	-	ORB & VGG-19 with softmax classifier	98.20%	97.98%	97.99%

cases. The fundamental purpose of this work is to establish a diagnostic approach that is both efficient and cost-effective, and which is capable of promptly identifying COVID-19 and other chest disease patients based on their CXR images.

V. CONCLUSION

This study presents a proposed model (ORB & VGG-19 with softmax) with the amalgamation of lung image segmentation and classification to accurately classify the ten chest diseases i.e., COVID-19, LC, ATE, COL, TB, PNET, EDE, PNEU, PLT, and normal using CXR images. With the use of Info-MGAN, we were able to efficiently train the pix-to-pix algorithm that was used in the segmentation of the lung images. This was accomplished by making use of the CXR images as well as the ground truth masks. To finish the task of image segmentation, the trained segmentation network, which is essentially the generator element of the Info-MGAN, is utilized for the segmentation of the preprocessed images. After the lung images had been segmented, they were loaded

into the feature extraction network, which is comprised of DCNN models (such as ResNet-50, ResNet-101, VGG-19, simple CNN, VGG-16, Inception-v3, DenseNet-169, and DenseNet-201) as well as key point identification techniques such as SURF and ORB. The extracted features obtained from DCNN models were then categorized using multiple ML approaches, including AdaBoost, Q-SVM, softmax, and RF, to diagnose ten distinct chest diseases. This study indicated that the classification accuracy was greatest when the VGG-19 model was paired with the ORB key points extraction method, and softmax was applied to the final layer. The model also yielded the lowest average FPR and FNR of all the architectures that were suggested in this study, with respective values of 1.10% and 2.33%. The performance of the suggested method for diagnosing ten chest diseases utilizing CXR images was superior to the state-of-the-art models. In the future, the suggested approach will be trained and tested on CT scan images for the identification of several chest diseases.

REFERENCES

[1] X. Yang, Y. Yu, J. Xu, H. Shu, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, and Y. Wang, "Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: A single-centered, retrospective, observational study," *Lancet Respiratory Med.*, vol. 8, no. 5, pp. 475–481, 2020, doi: 10.1016/S0140-6736(20)30183-5.

[2] WHO. (Jul. 2021). *Weekly Epidemiological Update on COVID-19–20*. [Online]. Available: www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

[3] M. Chung, A. Bernheim, X. Mei, N. Zhang, M. Huang, X. Zeng, J. Cui, W. Xu, Y. Yang, Z. A. Fayad, A. Jacobi, K. Li, S. Li, and H. Shan, "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, Apr. 2020.

[4] J. Zhang, Y. Xie, G. Pang, Z. Liao, J. Verjans, W. Li, Z. Sun, J. He, Y. Li, C. Shen, and Y. Xia, "Viral pneumonia screening on chest X-ray images using confidence-aware anomaly detection," 2020, arXiv:2003.12338.

[5] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2688–2700, Aug. 2020, doi: 10.1109/TMI.2020.2993291.

[6] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[7] N. Ghassemi, A. Shoeibi, M. Khodatars, J. Heras, A. Rahimi, A. Zare, R. B. Pachori, and J. Manuel Gorriz, "Automatic diagnosis of COVID-19 from CT images using CycleGAN and transfer learning," 2021, arXiv:2104.11949.

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

[9] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks," *Radiology*, vol. 284, no. 2, pp. 574–582, Aug. 2017.

[10] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, Jun. 2020, Art. no. 103792, doi: 10.1016/j.compbiomed.2020.103792.

[11] E. E. D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, arXiv:2003.11055.

[12] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, Nov. 2020.

[13] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (COVID-19) based on deep features," 2020.

[14] A. Amirkhani, A. Khosravian, M. Masih-Tehrani, and H. Kashiani, "Robust semantic segmentation with multi-teacher knowledge distillation," *IEEE Access*, vol. 9, pp. 119049–119066, 2021.

[15] S. R. Nayak, D. R. Nayak, U. Sinha, V. Arora, and R. B. Pachori, "Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102365.

[16] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2017, pp. 1125–1134.

[17] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *J. Big Data*, vol. 6, no. 1, pp. 1–18, Dec. 2019.

[18] A. Iqbal, M. Usman, and Z. Ahmed, "An efficient deep learning-based framework for tuberculosis detection using chest X-ray images," *Tuberculosis*, vol. 136, Sep. 2022, Art. no. 102234.

[19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE CVPR*, Jun. 2017, pp. 4700–4708.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556.

[21] M. E. Sahin, "Deep learning-based approach for detecting COVID-19 in chest X-rays," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103977.

[22] C. Ieracitano, N. Mammone, M. Versaci, G. Varone, A.-R. Ali, A. Armentano, G. Calabrese, A. Ferrarelli, L. Turano, C. Tebala, Z. Hussain, Z. Sheikh, A. Sheikh, G. Sceni, A. Hussain, and F. C. Morabito, "A fuzzy-enhanced deep learning approach for early detection of COVID-19 pneumonia from portable chest X-ray images," *Neurocomputing*, vol. 481, pp. 202–215, Apr. 2022.

[23] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1207–1220, Aug. 2021.

[24] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 screening on chest X-ray images using deep learning based anomaly detection," 2020, arXiv:2003.12338.

[25] I. D. Apostolopoulos and T. Bessiana, "COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, 2020, doi: 10.1007/s13246-020-00865-4.

[26] N. Tsiknakis, E. Trivizakis, E. Vassalou, G. Papadakis, D. Spandidos, A. Tsatsakis, J. Sánchez-García, R. López-González, N. Papanikolaou, A. Karantanas, and K. Marias, "Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays," *Experim. Therapeutic Med.*, vol. 20, no. 2, pp. 727–735, May 2020, doi: 10.3892/etm.2020.8797.

[27] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of coronavirus disease (COVID-19) based on deep features and support vector machine," *Int. J. Math., Eng. Manag. Sci.*, vol. 5, no. 4, pp. 643–651, Aug. 2020, doi: 10.33889/IJMEMS.2020.5.4.052.

[28] P. Saha, M. S. Sadi, and M. M. Islam, "EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers," *Informat. Med. Unlocked*, vol. 22, Jan. 2021, Art. no. 100505.

[29] H. Singh and S. Bawa, "Predicting COVID-19 statistics using machine learning regression model: Li-MuLi-Poly," *Multimedia Syst.*, vol. 28, no. 1, pp. 113–120, Feb. 2022, doi: 10.1007/s00530-021-00798-2.

[30] T. Mahmud, M. A. Rahman, and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization," *Comput. Biol. Med.*, vol. 122, Jul. 2020, Art. no. 103869, doi: 10.1016/j.compbiomed.2020.103869.

[31] M. J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, and N. Shukla, 2020, "X-ray image based COVID-19 detection using pretrained deep learning models," *engrXiv*, doi: 10.31224/osf.io/wx89s.

[32] D. Singh, V. Kumar, V. Yadav, and M. Kaur, "Deep neural network-based screening model for COVID-19-infected patients using chest X-ray images," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 35, no. 3, Mar. 2021, Art. no. 2151004.

[33] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 screening on chest X-ray images using deep learning based anomaly detection," 2020, arXiv:2003.12338.

[34] K. B. Ahmed, G. M. Goldgof, R. Paul, D. B. Goldgof, and L. O. Hall, "Discovery of a generalization gap of convolutional neural networks on COVID-19 X-rays classification," *IEEE Access*, vol. 9, pp. 72970–72979, 2021.

[35] M. Mamalakis, A. J. Swift, B. Vorselaars, S. Ray, S. Weeks, W. Ding, R. H. Clayton, L. S. Mackenzie, and A. Banerjee, "DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays," 2021, arXiv:2104.04006.

[36] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-dataset: A CT scan dataset about COVID-19," 2020, arXiv:2003.13865.

[37] S. Thakur and A. Kumar, "X-ray and CT-scan-based automated detection and classification of COVID-19 using convolutional neural networks (CNN)," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102920.

[38] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, and E. Xing, "Sample-efficient deep learning for COVID-19 diagnosis based on CT scans," *Medrxiv*, Apr. 2020, doi: 10.1101/2020.04.13.20063941.

[39] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-I. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists detection of pulmonary nodules," *Amer. J. Roentgenol.*, vol. 174, no. 1, pp. 71–74, 2000.

[40] B. van Ginneken, M. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database," *Med. Image Anal.*, vol. 10, no. 1, pp. 19–40, 2006. [Online]. Available: http://www.isi.uu.nl/Research/Databases/SCR/

[41] H. A. Malik and N. A. Tayyaba, "Chest radiographs of COVID-19 infected," *Mendeley Data*, 2020, doi: 10.17632/67dmnmx33v.1.

[42] H. Malik and T. Anees, "BDCNet: Multi-classification convolutional neural network model for classification of COVID-19, pneumonia, and lung cancer from chest radiographs," *Multimedia Syst.*, vol. 28, no. 3, pp. 815–829, Jun. 2022.

[43] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 image data collection: Prospective predictions are the future," 2020, *arXiv:2006.11988*.

[44] Frederick Nat Lab. (2018). *The Cancer Imaging Archive (TCIA) The Cancer Imaging Archive*. [Online]. Available: https:// www. cancerimagingarchive.net/about-the-cancer-imaging-archive-tcia/%0A and https://www.cancerimagingarchive.net/%0A and https://www.cancerimagingarchive.net/%0A and http://www.cancerimagingarchive.net/

[45] P. Mooney. (2018). *Chest X-ray Images (Pneumonia) | Kaggle. Kaggle.com*. Accessed: Mar. 3, 2022. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xraypneumonia%0A and https://data.mendeley.com/datasets/rscbjbr9sj/2

[46] M. Rastgarpour and J. Shanbehzadeh, "Application of AI techniques in medical image segmentation and novel categorization of available methods and tools," in *Proc. Int. Multiconf. Eng. Comput. Scientists*, 2011, pp. 23–519.

[47] A. M. Alqudah and S. Qazan, "Augmented COVID-19 X-ray images dataset," Tech. Rep., 2020, doi: 10.17632/2FXZ4PX6D8.4.

[48] E. J. Hwang, "Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs," *JAMA Netw. Open*, vol. 2, no. 3, 2019, Art. no. e191095, doi: 10.1001/jamanetworkopen.2019.1095.

[49] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, and J. Dong, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018, doi: 10.1016/j.cell.2018.02.010.

[50] H. Mohsen, E. S. A. El-Dahshan, E. S. M. El-Horbaty, and A. B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Comput. Inform. J.*, vol. 3, no. 1, pp. 68–71, 2018, doi: 10.1016/j.fcij.2017.12.001.

[51] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2097–2106.

[52] T. Rahman, A. Khandakar, M. A. Kadir, K. R. Islam, K. F. Islam, R. Mazhar, T. Hamid, M. T. Islam, S. Kashem, Z. B. Mahbub, and M. A. Ayari, "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," *IEEE Access*, vol. 8, pp. 191586–191601, 2020, doi: 10.1109/ACCESS.2020.3031384.

[53] H. Saeed, H. Malik, U. Bashir, A. Ahmad, S. Riaz, M. Ilyas, W. A. Bukhari, and M. I. A. Khan, "Blockchain technology in healthcare: A systematic review," *PLoS ONE*, vol. 17, no. 4, 2022, Art. no. e0266462.

[54] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 2672–2680.

[55] X. Chen, Y. Duan, R. Houthooft, J. Schulman, and I. P. S. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.

[56] K. Shankar and E. Perumal, "A novel hand-crafted with deep learning features based fusion model for COVID-19 diagnosis and classification using chest X-ray images," *Complex Intell. Syst.*, vol. 7, no. 3, pp. 1277–1293, Jun. 2021, doi: 10.1007/s40747-020-00216-6.

[57] T. D. Pham, "Classification of COVID-19 chest X-rays with deep learning: New models or fine tuning?" *Health Inf. Sci. Syst.*, vol. 9, no. 1, pp. 1–11, Dec. 2021.

[58] A. Lasker, M. Ghosh, S. M. Obaidullah, C. Chakraborty, and K. Roy, "LWSNet—A novel deep-learning architecture to segregate COVID-19 and pneumonia from X-ray imagery," *Multimedia Tools Appl.*, Dec. 2022, doi: 10.1007/s11042-022-14247-3.

[59] H. Gupta, N. Bansal, S. Garg, H. Mallik, A. Prabha, and J. Yadav, "A hybrid convolutional neural network model to detect COVID-19 and pneumonia using chest X-ray images," *Int. J. Imag. Syst. Technol.*, vol. 33, no. 1, pp. 39–52, Jan. 2023.

[60] E. Khan, M. Z. U. Rehman, F. Ahmed, F. A. Alfouzan, N. M. Alzahrani, and J. Ahmad, "Chest X-ray classification for the detection of COVID-19 using deep learning techniques," *Sensors*, vol. 22, no. 3, p. 1211, Feb. 2022.

[61] C. C. Ukwuoma, Z. Qin, V. K. Agbesi, B. M. Cobbinah, S. B. Yussif, H. S. Abubakar, and B. D. Lemessa, "Dual_Pachi: Attention-based dual path framework with intermediate second order-pooling for COVID-19 detection from chest X-ray images," *Comput. Biol. Med.*, vol. 151, Dec. 2022, Art. no. 106324.

[62] J. Li, Y. Wang, S. Wang, J. Wang, J. Liu, Q. Jin, and L. Sun, "Multi-scale attention guided network for COVID-19 diagnosis using chest X-ray images," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 5, pp. 1336–1346, May 2021.

[63] T. Liu, E. Siegel, and D. Shen, "Deep learning and medical image analysis for COVID-19 diagnosis and prediction," *Annu. Rev. Biomed. Eng.*, vol. 24, no. 1, pp. 179–201, Jun. 2022.

[64] A. Sedik, M. Hammad, F. E. A. El-Samie, B. B. Gupta, and A. A. Abd El-Latif, "Efficient deep learning approach for augmented detection of Coronavirus disease," *Neural Comput. Appl.*, vol. 34, pp. 11423–11440, Jan. 2021.

[65] E. Luz, P. Silva, R. Silva, L. Silva, J. Guimarães, G. Miozzo, G. Moreira, and D. Menotti, "Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images," *Res. Biomed. Eng.*, vol. 38, no. 1, pp. 149–162, Mar. 2022.

[66] R. Chandra, A. Jain, and D. S. Chauhan, "Deep learning via LSTM models for COVID-19 infection forecasting in India," *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0262708.

[67] V. N. Nyaga, M. Aerts, and M. Arbyn, "ANOVA model for network meta-analysis of diagnostic test accuracy data," *Stat. Methods Med. Res.*, vol. 27, no. 6, pp. 1766–1784, Jun. 2018.

[68] M. Mohammadi, A. A. Atashin, W. Hofman, and Y. Tan, "Comparison of ontology alignment systems across single matching task via the McNemar's test," *ACM Trans. Knowl. Discovery Data*, vol. 12, no. 4, pp. 1–18, Aug. 2018.

[69] G. Dhiman, V. Chang, K. Kant Singh, and A. Shankar, "ADOPT: Automatic deep learning and optimization-based approach for detection of novel coronavirus COVID-19 disease using X-ray images," *J. Biomolecular Struct. Dyn.*, vol. 40, no. 13, pp. 5836–5847, Sep. 2022.

[70] Z. A. A. Alyasseri, M. A. Al-Betar, I. A. Doush, M. A. Awadallah, A. K. Abasi, S. N. Makhadmeh, O. A. Alomari, K. H. Abdulkareem, A. Adam, R. Damasevicius, and M. A. Mohammed, "Review on COVID-19 diagnosis models based on machine learning and deep learning approaches," *Exp. Syst.*, vol. 39, no. 3, 2022, Art. no. e12759.

[71] K. Shankar, E. Perumal, P. Tiwari, M. Shorfuzzaman, and D. Gupta, "Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images," *Multimedia Syst.*, vol. 28, no. 4, pp. 1175–1187, Aug. 2022.

[72] A. M. Ismael and A. Sengur, "Deep learning approaches for COVID-19 detection based on chest X-ray images," *Exp. Syst. Appl.*, vol. 164, Feb. 2021, Art. no. 114054.

[73] S. Lee, E. K. Choe, H. Y. Kang, J. W. Yoon, and H. S. Kim, "The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population," *Skeletal Radiol.*, vol. 49, no. 4, pp. 613–618, Apr. 2020.

[74] P. K. Mall, P. K. Singh, and D. Yadav, "GLCM based feature extraction and medical X-ray image classification using machine learning techniques," in *Proc. IEEE Conf. Inf. Commun. Technol.*, Dec. 2019, pp. 1–6.

[75] N.-A.-A. Alam, M. Ahsan, M. A. Based, J. Haider, and M. Kowalski, "COVID-19 detection from chest X-ray images using feature fusion and deep learning," *Sensors*, vol. 21, no. 4, p. 1480, Feb. 2021.

[76] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, "Pneumonia detection using CNN based feature extraction," in *Proc. IEEE Int. Conf. Electr., Comput. Commun. Technol. (ICECCT)*, Feb. 2019, pp. 1–7.

[77] Y. Yu, P. Hu, J. Lin, and P. Krishnaswamy, "Multimodal multitask deep learning for X-ray image retrieval," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2021, pp. 603–613.

[78] H. Malik, A. Naeem, R. A. Naqvi, and W.-K. Loh, "DMFL_Net: A federated learning-based framework for the classification of COVID-19 from multiple chest diseases using X-rays," *Sensors*, vol. 23, no. 2, p. 743, Jan. 2023.

HASSAAN MALIK received the M.S. degree in computer science from the National College of Business Administration and Economics, Lahore, Pakistan. He is currently pursuing the Ph.D. degree in computer science with the University of Management and Technology, Lahore. He is a Lecturer with the Department of Computer Science, National College of Business Administration and Economics, Multan. He has seven years of professional experience in education and industry. His research interests include blockchain, federated learning, transformers, machine learning, and deep learning.
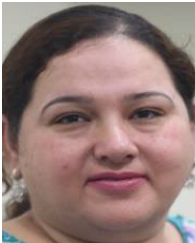
TAYYABA ANEES was born in Pakistan. She received the Ph.D. degree from the Vienna University of Technology, Vienna, Austria, in 2012. Her Ph.D. dissertation is in the area of service-oriented architecture and web services availability domain. For four years, she was a Project Assistant with the Vienna University of Technology. She is currently the Director of the software engineering program/an Assistant Professor with the Software Engineering Department, University of Management and Technology, Lahore. Her research interests include service-oriented architecture, web services, the Web of Things, deep learning, artificial intelligence, semantic web, cloud computing, the IoT, software availability, software safety, software fault tolerance, deep learning, and real-time data warehousing.

MUHAMMAD UMAR CHAUDHRY received the B.S. degree in computer engineering from Bahauddin Zakariya University, Multan, Pakistan, in 2009, and the Ph.D. degree in computer engineering from Sungkyunkwan University, Suwon, South Korea, in 2019. Currently, he is an Assistant Professor with the Department of Computer Science, MNS-University of Agriculture, Multan. His research interests include data science, recommender systems, feature selection, and machine learning.

RADOMIR GONO (Senior Member, IEEE) received the M.Sc., Ph.D., and Habilitate Doctorate degrees in electrical power engineering, in 1995, 2000, and 2008, respectively. In 2019, he was a professor of electrical power engineering. He has been with the Department of Electrical Power Engineering, VSB–Technical University of Ostrava, Czech Republic, since 1999, where he is currently a Professor and the Vice-Head of the Department of Electrical Power Engineering. His current research interests include electric power systems reliability, the optimization of maintenance, and renewable energy sources.

MICHAŁ JASIŃSKI (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the Wrocław University of Science and Technology, in 2016 and 2019, respectively. Since 2018, he has been with the Electrical Engineering Faculty, Wrocław University of Technology, where he is currently an Assistant Professor. He is the author and coauthor of more than 100 scientific publications. His research interests include big data in power systems, especially in point of power quality and optimization in multicarrier energy systems. Currently, he is a Guest Editor of special issues in *Energies*, *Electronics*, *Sustainability*, and *Frontiers in Energy Research*.

ZBIGNIEW LEONOWICZ (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the Wrocław University of Science and Technology, in 1997 and 2001, respectively, and the Habilitation degree from the Bialystok University of Technology, in 2012. Since 1997, he has been with the Electrical Engineering Faculty, Wrocław University of Technology. In 2019, he received the two titles of a Full Professor from the President of Poland, and the President of the Czech Republic. Since 2019, he has been a Professor with the Department of Electrical Engineering, where he is currently the Head of the Chair of Electrical Engineering Fundamentals.

PETR BERNAT received the Ph.D. degree in electrical machines, devices, and electric drives, in 2007. Since 1991, he has been with the Department of Electrical Machines and Apparatus, VSB–Technical University of Ostrava. Since 2011, he has been with the Department of Electrical Power Engineering. He holds a patent in the field of non-contact asynchronous machine diagnostics. His research interests include the reliability of electrical machines and devices, technical diagnostics, the optimization of machine maintenance, and the operation of renewable energy sources.

• • •