

## Article

# Screening Lung Diseases Using Cascaded Feature Generation and Selection Strategies

Jawad Rasheed <sup>1,\*</sup>  and Raed M. Shubair <sup>2</sup><sup>1</sup> Department of Software Engineering, Nisantasi University, Istanbul 34398, Turkey<sup>2</sup> Department of Electrical and Computer Engineering, New York University (NYU), Abu Dhabi 129188, United Arab Emirates; raed.shubair@nyu.edu

\* Correspondence: jawad.rasheed@nisantasi.edu.tr

**Abstract:** The global pandemic COVID-19 is still a cause of a health emergency in several parts of the world. Apart from standard testing techniques to identify positive cases, auxiliary tools based on artificial intelligence can help with the identification and containment of the disease. The need for the development of alternative smart diagnostic tools to combat the COVID-19 pandemic has become more urgent. In this study, a smart auxiliary framework based on machine learning (ML) is proposed; it can help medical practitioners in the identification of COVID-19-affected patients, among others with pneumonia and healthy individuals, and can help in monitoring the status of COVID-19 cases using X-ray images. We investigated the application of transfer-learning (TL) networks and various feature-selection techniques for improving the classification accuracy of ML classifiers. Three different TL networks were tested to generate relevant features from images; these TL networks include AlexNet, ResNet101, and SqueezeNet. The generated relevant features were further refined by applying feature-selection methods that include iterative neighborhood component analysis (iNCA), iterative chi-square (iChi2), and iterative maximum relevance–minimum redundancy (iMRMR). Finally, classification was performed using convolutional neural network (CNN), linear discriminant analysis (LDA), and support vector machine (SVM) classifiers. Moreover, the study exploited stationary wavelet (SW) transform to handle the overfitting problem by decomposing each image in the training set up to three levels. Furthermore, it enhanced the dataset, using various operations as data-augmentation techniques, including random rotation, translation, and shear operations. The analysis revealed that the combination of AlexNet, ResNet101, SqueezeNet, iChi2, and SVM was very effective in the classification of X-ray images, producing a classification accuracy of 99.2%. Similarly, AlexNet, ResNet101, and SqueezeNet, along with iChi2 and the proposed CNN network, yielded 99.0% accuracy. The results showed that the cascaded feature generator and selection strategies significantly affected the performance accuracy of the classifier.



**Citation:** Rasheed, J.; Shubair, R.M. Screening Lung Diseases Using Cascaded Feature Generation and Selection Strategies. *Healthcare* **2022**, *10*, 1313. <https://doi.org/10.3390/healthcare10071313>

Academic Editors: Alessandro Perrella, Valentina Orlando and Pierpaolo Di Micco

Received: 4 May 2022

Accepted: 13 July 2022

Published: 14 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** COVID-19; diagnostic tool; pneumonia; stationary wavelets transformation; transfer learning

## 1. Introduction

The COVID-19 is a family of viruses that surfaced in China in the last quarter of 2019, and, within a matter of weeks, it affected so many thousands of lives that the World Health Organization (WHO) declared it a pandemic [1]. The devastating effect on the well-being of humankind and the pandemic's complex non-linear nature has made it difficult to analyze the outbreak [2]. Up to today, more than 560 million COVID-19 positive cases have been detected, while 6.37 million humans have lost their lives. Presently, there are two common methods to diagnose COVID-19: taking samples from the respiratory tract for a kind of viral nucleic acid testing called reverse transcription–polymerase chain reaction (RT-PCR) and analyzing chest radiography imaging. However, RT-PCR requires a sophisticated and specialized laboratory with hi-tech machines that can cost up to USD 90,000 [3]. Besides being expensive, RT-PCR testing has a low positivity rate (63%), with high requirements for

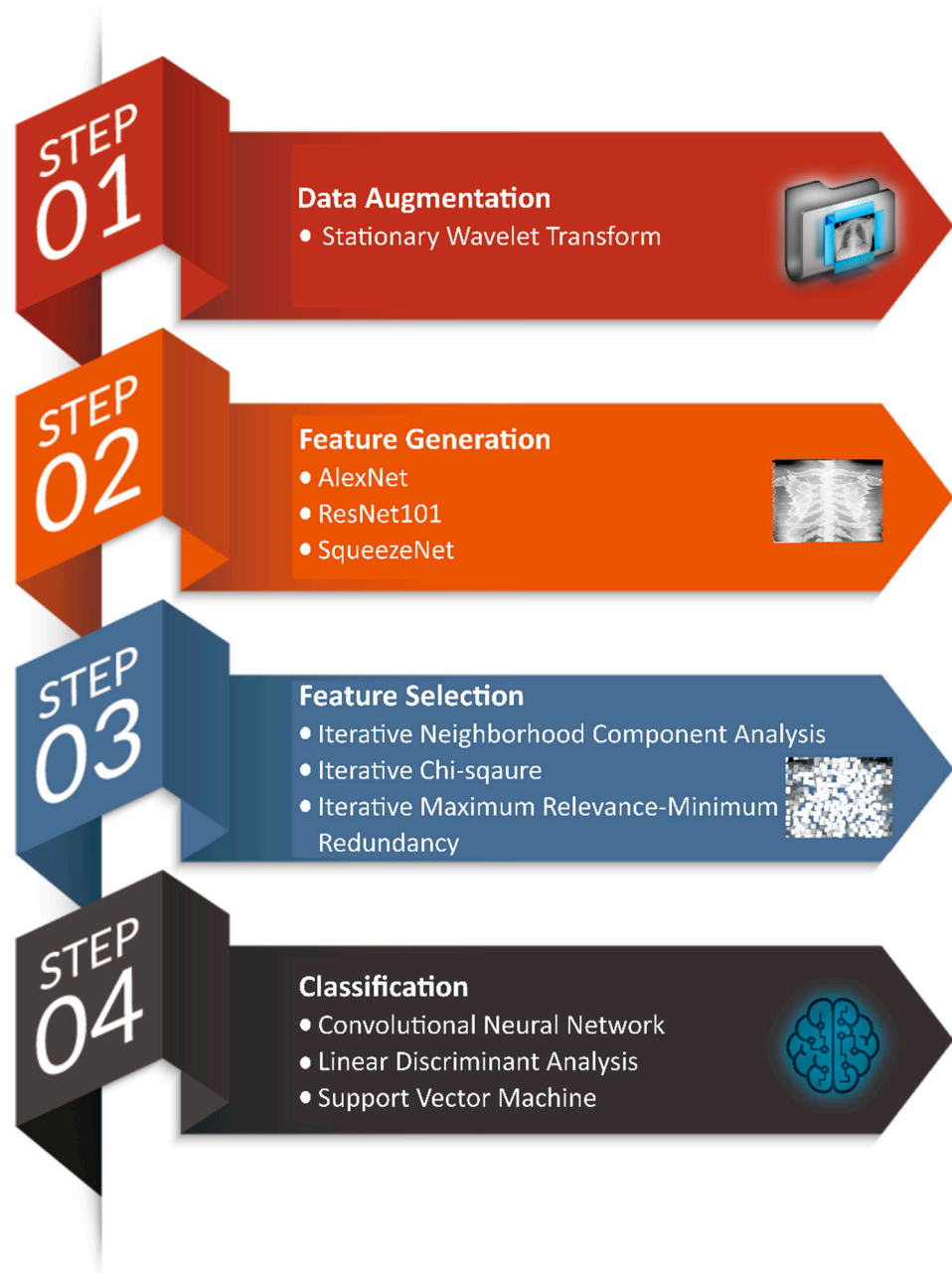
procedural and computing time [4]. Contrarily, testing through chest radiography imaging is cheaper, as it does not require the installation of machinery on a large scale due to high availability and easy accessibility in most parts of world [5]. Expert radiologists infer the diagnosis using chest radiography imaging, such as X-ray or computed tomography (CT) imaging, wherein the presence of COVID-19 is inferred by subtle visual signals. This method has gained popularity as a potent alternative. However, it requires additional radiologists, considering the high reproduction number of the COVID-19 virus [6], to analyze the results and reach a conclusion. Thus, to manage and control this pandemic, the immediate and meticulous screening of individuals is of great significance.

In last two decades, ML-based tools have shown promising results in screening and diagnosing several diseases. For instance, authors provided a decision support system to predict diabetes in individuals using various ML classifiers [7]. Similarly, researchers in [8] exploited random forest (RF), k-nearest neighbor (kNN), and naïve Bayes (NB) to detect breast cancer in patients. In addition, deep learning approaches have been successfully deployed in the domain of medical imaging, such as [9], wherein authors developed a TL-based multipurpose diagnostic system to diagnose pediatric pneumonia and some retinal diseases using chest X-ray images and optical coherence tomographic imaging, respectively. Besides ML techniques, computer vision (CV) experts incorporated feature-extraction and image-filtering techniques to enhance system performance in terms of accuracy and computational time. Others, such as [10], investigated principal component analysis (PCA) with a variant of the artificial neural network (ANN) to analyze and identify Parkinson's disease. Similarly, authors in [11] performed brain tumor classification using wavelet transformation as the feature extractor and SVM as the classifier.

Recently, several artificial-intelligence-based frameworks have been developed to manage the COVID-19 pandemic from different perspectives. Experts in [12] presented a deep-learning-based COVID-19 pneumonia diagnostic tool to distinguish COVID-19 pneumonia from negative cases by utilizing 10,182 chest X-ray radiography images. A textural image-characterization-techniques-based scheme has been proposed in [13] to analyze three ML classifiers (kNN, RF, and SVM) for the identification of COVID-19 positive cases. Their developed scheme achieved 99% accuracy on a test set and 91.3% accuracy on a training set for super-pixel-based histone image characterization. Researchers in [14] classified COVID-19 positive cases by proposing a discrete-wavelet-transform-based neurowavelet capsule network that first minimized the noise present in X-ray images and then performed training for classification. They managed to obtain precision of 99.7%, sensitivity of 99.2%, and accuracy of 99.6%. Another study [15] introduced a deep CNN-based inception model to identify COVID-19 cases by incorporating a Gaussian smoothing filter to enhance image quality and a glowworm swarm optimizer (GSO) for the fine tuning of the hyperparameters. They succeeded in securing 94.29% accuracy and a 93.94% f1-score.

Besides suggesting a single-disease diagnostic tool, computer scientists introduced multi-disease diagnostic systems. For instance, authors in [16] proposed a lung segmentation approach to present a multi-disease diagnostic tool by training a customized CNN-based model to identify tuberculosis, lung opacities, and cancer. Likewise, for the COVID-19 pandemic, researchers in [17] used 6330 images to train 12 well-established pre-trained CNN-based models for discriminating four classes (several pneumonia types including COVID-19). Despite considerable amount of image samples, the system secured an 84.46% average f1-score. A few pre-trained models (Xception, InceptionV3, and ResNeXt) were compared in [18] to identify normal, COVID-19-positive, and pneumonia cases using X-ray images. They obtained an overall accuracy of 86% and a sensitivity of 78%. Another study [19] investigated ResNet50 and VGG16 to classify normal, viral pneumonia and COVID-19 by utilizing 15,153 X-ray images. Their fine-tuned ResNet50 model succeeded in attaining 91.39% accuracy, whereas VGG16 reached the mark of 89.34% accuracy. A more detailed comparison of various studies to manage the COVID-19 pandemic can be found in [1].

As previously published well-established studies hardly attained effective performance in distinguishing COVID-19 cases from other pneumonia and healthy cases, therefore, to overcome the pandemic, auxiliary methods and techniques for the screening of pneumonia cases need further investigation. In addition to viral nucleic acid testing used in COVID-19 case identification, schemes based on artificial intelligence can be exploited for early diagnosis and monitoring. In this study, we investigated a new framework based on TL networks as a feature generator; several feature-selection techniques; and ML classifiers to identify COVID-19, other (viral/bacterial) pneumonia, and normal cases, using X-ray images with high accuracy (see Figure 1).



**Figure 1.** Workflow of the proposed system.

In the first phase of this study, X-ray images from three different online sources are gathered, which are then augmented for the better generalization of the proposed model by exploiting stationary wavelet transform and various other techniques, such as the random translation, rotation, and shearing of images. In the second stage, three pre-trained TL

models (AlexNet, ResNet101, and SqueezeNet) separately extracted the most-relevant 1000 features from each image. The scheme then merges these features to form a vector of 3000 features. In the third stage, three iterative feature-selection techniques were employed separately to select the most-optimal features for better classification. Lastly, one deep-learning-based framework and two ML classifiers are investigated to perform classification using optimal features.

This study contributes to the research community in the following manner:

- Outlines the prior studies (in detail) related to the identification of COVID-19 and other pneumonia from images.
- Introduces stationary wavelet transform as a data-augmentation technique to handle overfitting issues.
- Investigates the impact of different TL-based networks (AlexNet, ResNet101, and SqueezeNet) in diagnosing three clinical states (normal, other pneumonia, and COVID-19-positive).
- Assesses the significance of incorporating iNCA, iChi2, and iMRMR as feature-selection techniques in ML-based diagnostic tool.
- Analyzes the effect of various combinational schemes of feature-generation and feature-extraction techniques in the proposed model.
- Examines the performance and provides comparative analyses of each combinational scheme in terms of performance accuracy, F1-score, precision, and recall.
- Presents a simple but highly accurate ML-based model that can be used with other conventional clinical COVID-19 testing to remove false-alarm probability.
- Provides analysis to indicate that the proposed scheme outperformed prior studies by securing 99.2% accuracy while segregating COVID-19-positive cases from other pneumonia and normal ones.

The order of the remaining sections of this study is as follows. Section 2 outlines the dataset collection and technical background of the frameworks, techniques, and classifiers used for this study. Section 3 presents the detailed results and performance analysis, while Section 4 illustrates the comparison with previous studies. The paper ends with concluding remarks in Section 5.

## 2. Materials and Methods

The below passages present the overview of data collection and its description, as well as exploratory analysis, collectively with in-depth discussion about the selection and architecture design of pre-trained feature-generation frameworks, feature-selection methods and ML-based disease-identification classifiers adopted for this proposed lung-disease segregation system. In particular, it focuses on data-augmentation techniques, TL models, iterative feature selectors, and ML algorithms used for image data enhancement, the generation of useful features, the automatic selection of optimal features, and disease classification, respectively.

### 2.1. Description of Dataset and Augmentation Approach

The dataset used for this study was created by modifying several publicly available benchmark image repositories to facilitate the evaluation and training of the proposed model. The generated dataset that will be referred as PeN-CoVx throughout this study is a combination of 12,282 images of lungs collected from CRD (COVID-19 Radiography Database | Kaggle: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-data-base>) (accessed on 25 March 2022), CIDC (GitHub COVID-19 Image Data Collection: <https://github.com/ieee8023/covid-chestxray-dataset>) (accessed on 25 March 2022), and CXIP (chest X-ray images (pneumonia) | Kaggle: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>) (accessed on 25 March 2022). In particular, this research collected the correct data for each class (normal, pneumonia, and COVID-19+) from three different repositories to form PeN-CoVx dataset to avoid overfitting issues. It takes advantage of the below-mentioned types of patient cases:

- 4094 randomly selected X-ray images of normal healthy patients from the CRD dataset,
- 3616 X-ray images of confirmed COVID-19-positive cases from the CRD dataset,
- 478 X-ray images of confirmed COVID-19-positive cases from the CIDC dataset,
- 4094 randomly selected X-ray images of viral and bacterial pneumonia-affected patients from the CXIP dataset.

The idea of easily accessible and open source data inspired the preference of the CRD, CIDC, and CXIP datasets to form PeN-CoVx. As the COVID-19 pandemic recently surfaced, a noticeable trend shows a “skewed dataset” for COVID-19 cases, though significantly, enough repositories are publicly available having radiography images of normal and viral or bacterial pneumonia-affected patients to compare to COVID-19-infection cases. Like ML models, deep-learning-based classification frameworks require a huge number of uniform data for proper model training to avoid challenges such as overfitting. In order to resolve such issues, this research employed various pre-processing techniques and a data-augmentation approach by exploiting stationary wavelets.

The course of the collection stage revealed that PeN-CoVx has images of various sizes; therefore, to maintain uniformity in size, each image was resized to  $256 \times 256 \times 3$  in the pre-processing stage. Before proceeding to data enhancement, the PeN-CoVx dataset was divided into training, validation, and testing sets with ratios of 60%, 25%, and 15%, respectively.

## 2.2. SW as Augmentation Approach

Besides using rotational and shear operation to extend the training set, this study exploited SW to perform data augmentation by decomposing images of training set up to three levels. Researchers in [20] introduced a non-sampling wavelet transformation called stationary wavelet transform with the capability of determining a more precise estimation of continuous wavelet transform by decamping a sample image into several levels. Besides this, it also possesses translational invariance and redundancy features.

The stationary wavelet decomposes an image based on the number of levels and the matrices’ dimensionality of the normalized input image. Assuming  $n$  as the number of decomposition levels, then SW transformation will produce 3-dimensional arrays for a 2-dimensional normalized matrix:

$$C_{SW-2D} = [Ver(:, :, 1 : n); Hor(:, :, 1 : n); Dia(:, :, 1 : n); App(:, :, n)], \quad (1)$$

where *Ver* refers to the vertical coefficient; *Hor* represents the horizontal coefficient, and *Dia* denotes the diagonal coefficient, while *App* refers to the approximation coefficient for SW of  $n$ -levels. Similarly, for the 3-dimensional normalized image matrix ( $h \times w \times 3$ ), then SW transformation will produce a coefficient with 4-dimensional arrays ( $h \times w \times 3 \times n$ ) for  $n$ -levels as in Equation (2) [20,21].

$$C_{SW-3D} = [Ver(:, :, 1 : 3, 1 : n); Hor(:, :, 1 : 3, 1 : n); Dia(:, :, 1 : 3, 1 : n); App(:, :, 1 : 3, n)], \quad (2)$$

The present study exploited SW transformation for three levels of wavelet decomposition. Later, the decomposed images are randomly translated, rotated, and sheared to further enhance the PeN-CoVx training set. The experimental setup details about SW transformation and preprocessing steps can be found in the next section of this study.

## 2.3. Feature Generation Using Pre-Trained Frameworks

This section outlines the CNN-based TL models used to extract relevant features from the PeN-CoVx dataset.

### 2.3.1. AlexNet

Back in 2012, Alex Krizhevsky and a handful of researchers presented a new landscape in the field of CV that demolished prior ideas by producing a seismic shift in resolving image classification tasks with AlexNet [22]. It was the winning entry in ILSVRC 2012 by solving an image classification problem for a dataset containing 1.4 million images of



1000 classes. AlexNet has deep architecture, composed of 650,000 neurons that contribute to 62.3 million training parameters. However, it incorporated dropouts [23] and ReLU instead of regularization and *tanh* to deal with issues such as overfitting and linearity, respectively. Moreover, it also utilized overlap pooling to reduce network size but introduced padding to prevent a drastic reduction in the size of the feature map.

It encompasses eight layers (three fully connected layers and five convolutional layers, CLs). The input layer in the network accepts an image with a size of  $227 \times 227$ . Next, the first CL has 96 filters with a size of  $11 \times 11$  and stride 4, while the second CL has 256 filters but each with a size of  $5 \times 5$  and stride 1. Furthermore, padding of 2 is utilized to handle the size of the feature maps. An overlapping max-pooling layer of  $3 \times 3$  size with stride 2 follows each of these CLs. The rest of the CLs have 384 filters of size  $3 \times 3$  with stride and padding of 1, except the last CL, which has 256 filters. The network extracts more features as the number of filters increases in deeper layers, and simultaneously, the feature map shape decreases as the filter size declines. It then has three fully connected layers, each having 4096 neurons except the last, which has 1000. A dropout of 0.5 is performed before and after the first fully connected layer to avoid overfitting.

### 2.3.2. ResNet101

The prior deeper networks generally led to a degradation problem at the time of convergence, thus researchers suggested a residual network (ResNet) [24] to deal with such complications at ILSVRC 2015 competitions. Similarly, it outperformed other models while securing better generalization performance on COCO detection and segmentation in the Common Objects in Context (COCO 2015) competition. Prior to ResNet, deep learning researchers were developing deep networks by stacking more layers to extract complex features in order to attain high accuracy, but as the network converged, it faced a degradation problem. The increase in network depth, first, saturated the accuracy value, which then swiftly degraded due to the vanishing gradient effect. Reference [24] addressed the issue by introducing the ResNet framework, which stacks multiple CLs like any other deep network, but at the same time, it establishes identity connections between different layers. Besides numerous variants of ResNet, this study exploited ResNet101 because of its better performance on the problem at hand compared to the other variants described in the next section of this study. It mitigated the vanishing gradient effect and covariate shift problem through the incorporation of identity connection and batch normalization, respectively.

The ResNet101 architecture consists of four stages. It accepts an instance (input image), having a size of  $224 \times 224 \times 3$  in the case of this study, and passes it to the first CL ( $7 \times 7$  kernel) and then to the max-pooling (MP) layer with a kernel size of  $3 \times 3$ . The resultant is then fed to the first stage, which has three residual blocks with identity connections between them. Each block has three CLs of 64, 64, and 256 kernels, respectively. The second, third, and fourth stages have four, twenty-three, and three blocks, and each block has three CLs of different kernel sizes. After every network stage, the input size decline while the channel width doubles. Lastly, an average-pooling layer (AP) is placed, followed by a dense layer with a thousand neurons.

### 2.3.3. SqueezeNet

In 2016, scientists at Stanford University and the University of California proposed a replacement for AlexNet, called SqueezeNet [25]. The devised compact network performs three times faster than AlexNet due to reduced parameters, as it replaces  $3 \times 3$  filters with  $1 \times 1$  filters to capture relationships among its channels. In addition, they used squeeze layers that reduces input channels to  $3 \times 3$  filters. These strategies judiciously decrease the number of parameters in a CNN while pursuing the preservation of accuracy. Besides this, researchers introduced delayed down sampling in the network to attain large activation maps that maximizes the accuracy on a defined budget of parameters.

A fire module acts as a building block in the SqueezeNet architecture and is composed of *expand* and *squeeze* layers. The output of squeeze layers, each having a convolutional

filter of size  $1 \times 1$ , is passed to the expand layer that contains  $3 \times 3$  and  $1 \times 1$  convolutional filters. The network contains eight fire modules, a global average-pooling layer, three max-pooling layers, and two standalone convolutional layers. Unlike other state-of-the-art networks, SqueezeNet does not have any dense layer.

The SqueezeNet has variants based on its architecture. Inspired by ResNet, the SqueezeNet architecture has several bypass connections to increase the filters. This study exploited SqueezeNet with a deep compression approach that uses  $1 \times 1$  convolutional filters as the complex bypass connection designed in [25].

#### 2.4. Feature-Selection Techniques

Feature selection is a crucial task in ML, as it has numerous options available but must choose useful features to design an efficient network. This study exploited three different feature-selection techniques, described below.

##### 2.4.1. Iterative Neighborhood Component Analysis (iNCA)

NCA is a technique to identify and segregate multivariate data by learning a distance metric [26]. It linearly transforms the original features to maximize classification performance. Functionally, it uses stochastic nearest neighbors (SNN) to differentiate the objective function by considering the whole transformed set instead of the  $k$ -nearest neighbors (kNN) at each transformed point for leave-one-out classification.

Instead of picking a fixed number of kNNs and taking a majority vote, it randomly picks a single neighbor point and determines the expected votes for each class. So each point  $i$  chooses another point  $j$  probabilistically by exploiting the softmax of the Euclidean distance between these two points using [26]:

$$p_{ij} = \frac{\exp(-d_{i,j})}{\sum_{k \neq i} \exp(-d_{i,k})}, \quad p_i = 0, \quad (3)$$

where  $d_{i,j}$  refers to the Euclidean distance between a neighboring point and the LOO point. Thus, for better classification, NCA aims to maximize the objective function that can be defined as:

$$f(A) = \sum_i \sum_{j \in S_i} p_{ij} = \sum_i p_i, \quad (4)$$

where  $S_i$  is the set of nearest neighbors.

##### 2.4.2. Iterative Chi-Square (iChi2)

Chi2 is a statistical technique used to analyze the dependency of events. Chi2 determines the deviation between the observed count,  $O$ , and the expected count,  $E$ , for data on two variables, using the following:

$$x_{fr}^2 = \sum \frac{(O_i + E_i)^2}{E_i}, \quad (5)$$

where  $fr$  refers to the degree of freedom. As in feature selection, the objective is to consider features that highly depend on the response, which can be achieved with a higher value of Chi2, if  $O$  is not close to  $E$ . Thus, a lower Chi2 value indicates that features are independent, while a higher value suggests that those representations are response-reliant and can be considered to train the model.

##### 2.4.3. Iterative Maximum Relevance–Minimum Redundancy (iMRMR)

Feature-selection algorithms can be classified into two groups, namely all relevant and minimal-optimal. Generally, all-relevant-based algorithms give individual assessment for each feature as they find statistical dependency with target variable. However, such techniques might be too indulgent in cases of many features with high redundancy. Therefore,

ref. [27] discovered and proved that the iMRMR technique can achieve maximum accuracy even when selecting fewer features out of thousands for predicting a disease.

It works iteratively to identify an optimal set of features by calculating the iMRMR score for each feature using:

$$x_i(f) = \frac{Rev(f | TV)}{Reu(f | \text{features selected until } i - 1)}, \quad (6)$$

where  $f$ ,  $Rev$ ,  $Reu$ , and  $TV$  refer to features under consideration, relevance, redundancy, and target variable, respectively. The feature that attains the highest score at the  $i$ th iteration is selected. In fact, it chooses a feature that has less redundancy with regard to features chosen at previous iteration,  $i - 1$ , but has higher relevance with regard to the target variable. Researchers suggested multiple variants of MRMR for discrete and continuous variables, but this study exploited a combine F-test with correlation using a quotient (FCQ) based on its popularity for various classification models in terms of robust accuracy and computational time.

## 2.5. Models for Classification

### 2.5.1. Convolutional Neural Network

CNN is a widely used deep learning classification tool that has multiple perceptron layers placed next to each other in a sequence [28]. It is a chain of simple layers connected together to implement a form of progressive data distillation, by taking the input, transforming it to a meaningful representation, and predicting the final output. It usually encompasses three main forms of layers: convolutional, dense, and pooling layers. CL acts as a core component that consists of neurons in one layer allied with limited neurons of the neighboring layers, which shares weights having similar characteristics.

The relevant features are fed to CL (hidden layers) as an input feature map in a matrix or vector structure. The network determines a weight,  $w$ , for every connection between neurons of the CL and neurons from the first layer (input layer). According to first layer, the model then determines 'a' (weighted sums of all activations):

$$w_{0,0}a_0^{(0)} + w_{0,1}a_1^{(0)} + \dots + w_{0,n}a_n^{(0)} \quad (7)$$

A function (such as  $\sigma$  sigmoid) is used to normalize the resultant, and the neuron activation is adjusted through the inclusion of appropriate bias, 'b', to represent the second neuron of the first hidden layer. In the same way, the network determines the weights and biases of every hidden layer to produce activation maps. Later, a pooling layer plays the role of a fuzzy filter by gradually downscaling the feature dimensionality to subsequently reduce computational cost. Lastly, a fully connected layer combines all the extracted representations of the preceding layers for classification.

### 2.5.2. Linear Discriminant Analysis

Unlike logistic regression, the LDA traditional classification algorithm models differences in multi-classes by projecting data low dimensional-space. It determines the statistical properties from data to form a new axis of lower dimensional space that reduces intra-class variation while keeping greater differences in the means of classes. It calculates the mean value of each input  $x$  for every class:

$$m_z = \frac{1}{n_z} \sum_i x_i, \quad (8)$$

where  $n_z$  is the number of instances within class  $z$ . It determines the variance across all classes using:

$$v^2 = \frac{1}{n - c} \sum_i (x_i - m)^2, \quad (9)$$



where  $c$  equates to the number of classes. Lastly, it uses Bayes theorem to determine the probability for the new input set.

### 2.5.3. Support Vector Machine (SVM)

SVM is widely adopted as a discriminative classifier to deal with linear as well as non-linear tasks. In order to correctly classify instances, it aims to define a decision boundary that maximizes the separating distance between the training data. Thus, it incorporates the first feature, known as the maximal margin classifier. It tries to find support vectors, instances located on the edge of class-descriptors, so that it can appropriately tag the samples into two classes. It is achieved by dividing the 2-dimensional space with a line such that data points falling on the left side of boundary are segregated into a different class than the ones on the other side. This division can be achieved with an infinite number of lines, but what makes SVM outstanding compared to the others (like k-nearest neighbors) is its ability to discover the best optimal possible separation line that has maximum margin between support vectors. This separating line is also known as a hyperplane if it is for a more-than-3-dimensional space by performing a dot product (kernel trick) on the transformed input vector.

A linear kernel is defined by the dot product of data  $X$  (which is to be categorized) with weight vector  $w$  (which the user wants to minimize) and the addition of a linearly estimated coefficient,  $b$ :

$$\text{Kernel Function } (X) = w^T X + b. \quad (10)$$

The most used kernel for non-linear data is RBF, which computes distances between specific features with all others to generate new features. Therefore, this study used Gaussian RBF, which can be defined for  $X_1$  and  $X_2$  data by:

$$\text{Kernel Function } (X_1, X_2) = e^{(-\gamma * \|X_1 - X_2\|^2)}. \quad (11)$$

## 3. Experimental Results and Analysis

The aim of this study is to design an efficient diagnostic tool that can accurately classify X-ray images of human lungs as COVID-19-infected, pneumonia or normal by enhancing a dataset using stationary wavelet transformation, generating relevant features via TL models, selecting optimal representations, and finally classifying using ML-based classifiers. In order to segregate the three best pre-trained TL models for feature generation, this study first tested several TL state-of-the-art networks with seven various ML and deep learning classifiers, tabulated in Table 1. Numerous experiments were conducted, wherein each TL network extracted useful features from images of the PeN-CoVx dataset; these images were later fed to the given classifiers for the diagnosis of the three clinical states. Table 1 puts forth the classification accuracies attained by CNN, DT, KNN, LDA, LR, NB, RF, and SVM against each pre-trained TL model on the PeN-CoVx dataset. Table 1 evidently shows that CNN, LDA, and SVM outperformed other classifiers (DT, KNN, LR, and RF) by effectively utilizing the features generated by TL networks. Similarly, AlexNet, ResNet101, and SqueezeNet successfully generated the most relevant features compared to other networks for the problem at hand, therefore; this research exploited AlexNet, ResNet101, and SqueezeNet as feature generators and CNN, LDA, and SVM for classification in the proposed scheme.

The PeN-CoVx dataset gathered for this research contained 4094 X-ray images for each clinical state (normal, COVID-19+, other pneumonia) and was further split into 60% training, 25% validation, and 15% testing sets, as listed in Table 2. As limited training dataset can lead to overfitting problem in TL-based models, this research employed various pre-processing techniques and data-augmentation approaches by exploiting SW transform, random rotation, translation, and shear operations. For this reason, the images are first normalized before decomposing them into various levels of stationary wavelets.

**Table 1.** Comparison analysis of pre-trained transferred models as feature-extractor and machine-learning-classification algorithms.

Feature-Generation Models	Classification Accuracy (%)						
	CNN	DT	KNN	LDA	LR	RF	SVM
AlexNet	95.6	90.7	88.8	95.0	91.9	90.4	<b>95.9</b>
DenseNet-121	95.0	90.4	87.5	94.8	91.9	91.0	94.8
DenseNet-169	95.1	91.0	89.5	95.1	92.0	91.0	95.3
DenseNet-201	95.1	90.1	87.6	94.5	90.7	88.9	95.3
DenseNet-263	94.8	88.8	85.0	93.8	90.1	87.5	95.0
EfficientNet-B0	92.9	86.6	8.9	90.7	86.6	86.8	91.5
GoogleNet	93.5	90.1	82.2	92.5	87.1	88.8	92.5
InceptionNetV3	95.0	91.0	83.1	93.3	87.3	89.3	95.1
MobileNetV2	90.2	87.5	80.1	89.9	85.0	87.5	90.1
ResNet18	90.7	88.8	82.2	90.7	86.7	85.0	90.6
ResNet50	92.5	86.6	82.6	91.9	87.1	85.5	92.9
ResNet101	<b>96.1</b>	91.9	85.0	95.9	90.7	87.8	<b>96.1</b>
ResNet152	95.1	91.2	85.3	95.0	90.7	90.4	95.3
SqueezeNet	96.1	93.5	85.0	96.1	91.5	92.2	<b>96.6</b>
VGG16	93.5	90.1	84.4	91.9	88.6	91.9	93.0
VGG19	92.9	90.7	83.1	92.9	88.6	91.5	92.5
XceptionNet	91.0	87.5	80.5	88.8	87.5	86.8	91.0

CNN: convolutional neural network; DT: decision tree; KNN: k-nearest neighbors; LDA: linear discriminant analysis; LR: logistic regression; RF: random forest; SVM: support vector machine.

**Table 2.** Data set information.

Clinical State	Number of Instances			
	Training Set without Augmentation	Training Set with Augmentation	Validation Set	Testing Set
Normal	2456	22,104	1024	614
Other pneumonia	2456	22,104	1024	614
COVID-19+	2456	22,104	1024	614
<b>Total</b>	<b>7368</b>	<b>66,312</b>	<b>3072</b>	<b>1842</b>

In the data-augmentation phase, the normalized PeN-CoVx images were decomposed to three levels using SWT. For this proposed study, each normalized image passed through several high-pass and low-pass filters. It exploited Daubechies's (db2) orthogonal filter to decompose 2-D SW. For first-level decomposition, it determined the detail coefficients ( $Ver_1$ ,  $Hor_1$ , and  $Dia_1$ ) and the approximate coefficient ( $App_1$ ) of high and low frequencies, respectively. Similarly, for next-level decomposition, it determined the detail coefficients ( $Ver_2$ ,  $Hor_2$ , and  $Dia_2$ ) and the approximate coefficient ( $App_2$ ) of high and low frequencies, respectively. Likewise, the SWT data-augmentation phase produced three images for each corresponding image in the training set. Table 3 tabulates the detailed output coefficient for each level in 2-D SW.

**Table 3.** Output coefficient for each level in 2-D stationary wavelet (SW) transform.

Decomposition Level	Down Sampling	Approximate Coefficient (Low Frequency)	Detail Coefficient (High Frequency)
1	Yes (by 2)	$App_1$	$Ver_1, Hor_1, Dia_1$
2	-	$App_2$	$Ver_2, Hor_2, Dia_2$
3	-	$App_3$	$Ver_3, Hor_3, Dia_3$

$App$ : approximate coefficient (image);  $Ver$ : vertical coefficient;  $Hor$ : horizontal coefficient;  $Dia$ : diagonal coefficient of the SW transformed image.

In order to enhance the training dataset further, techniques including shear, random translation, and rotation were applied to each wavelet-decomposed image. Table 4 represents the parametric details of augmentation techniques exploited to overcome the overfitting problem. As a result, the number of images for each clinical state in training set was increased nine times; thus, the training data now contains 22,104 images for each class (normal, pneumonia, COVID-19+).

**Table 4.** Parametric values of data-augmentation techniques.

Augmentation Technique	Parametric
Translation	−10, 10
Rotation	−90, 90
Shear	−30, 30

Later, three pre-trained TL models (described in the previous section) utilized these augmented training datasets to extract the 3000 most-relevant features for better classification. Before feeding images to the TL model, the size of the images was adjusted according to the input layer of each TL model. Each TL model (AlexNet, ResNet101, SqueezeNet) generated 1000 features for each image, which were then merged as shown in Algorithm 1. Thus, the merging accumulated  $66,312 \times 3000$  features for all of the training sets.

**Algorithm 1.** Pseudo code of the proposed scheme.

**Proposed Algorithm**

```

01 ImP = Load (PeN-CoVx) // Read PeN-CoVx image dataset
02 ImNorm = (ImP − min(ImP)) / (max(ImP) − min(ImP)) // Normalize images within range [0, 1]
03 ImSWC = CSW(ImNorm) // Data augmentation: 3-levels decomposition by stationary wavelet transformation
04 ImAug = DataAugmentation(ImSWC) // Performing random rotation, translation, and shear operation
05 for i = 1 to size-of ImAug do // Loop to extract relevant features for each image using transfer-learning model
06     fg1 = AlexNet(ImAug) // Extract 1000 features via AlexNet
07     fg2 = ResNet101(ImAug) // Extract 1000 features via ResNet101
08     fg3 = SqueezeNet(ImAug) // Extract 1000 features via SqueezeNet
09     for j = 1 to 3 do // Loop to merge extracted feature
10         X(i, 1000 × j + 1: 1000 × (j + 1)) = fgj // merge extracted features
11     end for loop // end loop to merge extracted feature
12 end for loop
13 fs1 = iNCA(X, Y) // Determine optimal features via iterative Neighborhood Component Analysis
14 fs2 = iChi2(X, Y) // Determine optimal features via iterative Chi-square
15 fs3 = iMRMR(X, Y) // Determine optimal features via iterative Maximum Relevance Minimum Redundancy
16 PLk = CNN(fsk, Y, 3) // Predict clinical-state using optimal features by Convolutional Neural Network
17 PLk = LDA(fsk, Y, 3) // Predict clinical-state using optimal features by Linear Discriminant Analysis
18 PLk = SVM(fsk, Y, 3) // Predict clinical-state using optimal features by Support Vector Machine

```

Subsequently, the merged features were fed separately to three different feature-selection algorithms (iNCA, iChi2, and iMRMR). These feature selectors select optimal features to form vector by defining the upper and lower bound of the determined indices. Moreover, the feature-selection scheme iteratively calculates loss against each feature and finally selects features having low loss value as optimal features. Thus, iNCA, iChi2, and iMRMR successfully obtained 1588, 1620, and 1468 optimal features, respectively, that will contribute decisively in the classification of the three clinical states.

Lastly, one deep learning (CNN)- and two ML (LDA and SVM)-based classifiers were trained separately with 10-fold cross-validation for each optimal feature set obtained by iNCA, iChi2, and iMRMR. Later, each model was evaluated on an unseen test data set composed of 614 images of each class. As each classifier was trained separately for each iNCA feature vector, this study computed nine different classification schemes overall. For

a better comparison, this study evaluated all trained models with the metrics accuracy, F1-score, recall, and precision.

After extensive trials, the most suitable CNN architecture was devised, having a few CLs and a fully connected layer. This CNN architecture contained three CLs with ReLU, stride of 1, and kernel size of 2. The first, second, and last CL had 256, 128, and 64 filter sizes, respectively. Averaging pooling followed the first and second CL for dimensionality reduction. Lastly, a fully connected layer with three neurons was utilized to classify an X-ray image into one of three-clinical states. Table 5 represents the classification performance achieved by CNN against each feature-selection technique.

**Table 5.** Performance metrics for the convolutional neural network against each feature-selection technique.

Feature Selector	Statistics	Precision	Recall	F1-Score	Accuracy
iNCA	Minimum	98.536	98.534	98.535	98.534
	Maximum	99.078	99.077	99.077	99.077
	Average	98.810	98.806	98.808	98.806
iChi2	Minimum	98.915	98.914	98.914	98.914
	Maximum	99.133	99.132	99.132	99.131
	<b>Average</b>	<b>99.026</b>	<b>99.023</b>	<b>99.024</b>	<b>99.023</b>
iMRMR	Minimum	98.489	98.154	98.322	98.154
	Maximum	98.914	98.914	98.914	98.914
	Average	98.535	98.534	98.534	98.534

iNCA: iterative neighborhood component analysis; iChi2: iterative chi-square; iMRMR: iterative maximum relevance–minimum redundancy.

Similarly, for each optimal feature set obtained from feature selectors, a separate LDA model was trained over 10-fold cross-validation. The model was trained by setting the gamma value as null with a linear discriminant type. Table 6 represents the classification performance achieved by LDA against each feature-selection technique.

**Table 6.** Performance metrics for linear discriminant analysis against each feature-selection technique.

Feature Selector	Statistics	Precision	Recall	F1-Score	Accuracy
iNCA	Minimum	97.790	97.768	97.779	97.774
	Maximum	98.752	98.751	98.751	98.751
	<b>Average</b>	<b>98.271</b>	<b>98.263</b>	<b>98.267</b>	<b>98.263</b>
iChi2	Minimum	97.674	97.666	97.670	97.666
	Maximum	98.207	98.209	98.208	98.208
	Average	97.535	97.515	97.525	97.515
iMRMR	Minimum	96.526	96.526	96.526	96.526
	Maximum	97.927	97.720	97.724	97.720
	Average	97.124	97.123	97.123	97.123

iNCA: iterative neighborhood component analysis; iChi2: iterative chi-square; iMRMR: iterative maximum relevance–minimum redundancy.

Another ML classifier, SVM was separately trained over 10-fold cross-validation for each optimal feature set. RBF was used as a kernel function with auto scaling to classify the input X-ray image into three clinical states. Table 7 represents the classification performance achieved by SVM against each feature-selection technique.

**Table 7.** Performance metrics for the support vector machine against each feature-selection technique.

Feature Selector	Statistics	Precision	Recall	F1-Score	Accuracy
iNCA	Minimum	98.756	98.751	98.754	98.751
	Maximum	99.191	99.186	99.189	99.186
	Average	98.969	98.969	98.969	98.969
iChi2	Minimum	99.024	99.023	99.023	99.023
	Maximum	99.462	99.457	99.460	99.457
	<b>Average</b>	<b>99.241</b>	<b>99.240</b>	<b>99.241</b>	<b>99.240</b>
iMRMR	Minimum	98.534	98.534	98.534	98.534
	Maximum	99.078	99.077	99.078	99.077
	Average	98.806	98.806	98.806	98.806

iNCA: iterative neighborhood component analysis; iChi2: iterative chi-square; iMRMR: iterative maximum relevance–minimum redundancy.

#### 4. Discussion

Besides analyzing the performance of the proposed scheme using accuracy, precision, sensitivity, and f1-score, we also determined the error of omission (see Table 8). This measures the false negatives of each classifier for all classes against each feature-selection technique to represent the fraction of samples belonging to a class but predicted to be in a different class. It is noteworthy that all proposed combinational schemes have less error of omission for COVID-19 cases than other classes (pneumonia and healthy individuals), except CNN with iNCA. The structural combination (CNN and iNCA) performed well for pneumonia classification, as it obtained only 0.81% error of omission. However, SVM with iChi2 outperformed all other schemes, as it has low error of omission on average for all classes.

**Table 8.** Comparing the error of omission of the convolutional neural network (CNN), linear discriminant analysis (LDA), and support vector machine (SVM) for each class against the exploited feature-selection techniques.

Feature Selector	Class Label	Error of Omission (%)		
		CNN	LDA	SVM
iNCA	Normal	1.63	2.12	1.14
	Pneumonia	0.81	1.63	1.14
	COVID-19	1.14	1.47	0.81
iChi2	Normal	1.14	2.12	0.98
	Pneumonia	0.98	2.44	0.65
	COVID-19	0.81	1.63	0.65
iMRMR	Normal	1.47	3.09	1.30
	Pneumonia	1.79	3.26	1.30
	COVID-19	1.14	2.28	0.98

iNCA: iterative neighborhood component analysis; iChi2: iterative chi-square; iMRMR: iterative maximum relevance–minimum redundancy.

Experts had developed numerous diagnostic tools to manage infectious diseases efficiently that can handle several class-classification tasks. For instance, ref. [29] blended conservative smoothing filtering, PCA, and SVM to secure 99.93% accuracy; however, it only segregated COVID-19-positive cases from negative cases. Contrarily, for a three-class classification task, ref. [30] used ResNet50 as the feature-extraction technique, with a SVM, to attain 95.33% accuracy. Similarly, ref. [31] exploited ResNet152 with various classifiers to reach an overall accuracy of 97.70% using X-ray images. Table 9 presents a comparative analysis of the proposed study with prior works.



**Table 9.** Comparative analysis of the proposed study with prior works.

Study	Techniques	Accuracy (%)
[30]	ResNet50 feature extractor with SVM	95.33
[31]	SMOTE and ResNet152 with XGBoost and random forest	97.70
[32]	Customized CNN-based network	84.22
[33]	VGG-16-based scheme	97.0
[34]	Customized Xception Net	95.0
[35]	CNN with transfer multireceptive feature optimizer	95.1
[36]	Cascaded ResNet50V2 and Xception Net	91.4
[37]	Customized CNN-based model	93.30
[38]	Pre-trained deep learning models with GAN	85.2
<b>Proposed</b>	<b>SWT + (AlexNet, ResNet101, and SqueezeNet) + iChi2 + SVM</b>	<b>99.24</b>

SVM: support vector machine; CNN: convolutional neural network; GAN: generative adversarial network; SWT: stationary wavelet transform; iChi2: iterative chi-square.

This study suggested a robust smart diagnostic tool based on a modern TL framework and iterative feature-selector techniques to rapidly screen and accurately diagnose COVID-19, healthy individuals and other pneumonia using radiography images. It tested various TL networks and finally selected three TL architectures (AlexNet, ResNet101, and SqueezeNet) to generate the most-relevant features from radiography images. Later, it employed three feature-selection techniques (iNCA, iChi2, and iMRMR) and analyzed their effects on the proposed deep-learning-based CNN framework, LDA, and SVM classifiers to classify COVID-19+, other pneumonia, and healthy cases.

For this experimental task, 12,282 radiography images were examined, of which 4094 belonged to COVID-19 infected patients and 4094 belonged to other pneumonia cases, while the rest pertained to normal patients or healthy humans. Extensive experiments were performed to ensure that the proposed model scheme generalized well using 10-fold cross-validation to cater to unseen data and does not perform overfitting or underfitting. By practicing feature generators (AlexNet, ResNet101, and SqueezeNet) along iChi2 with SVM, the network achieved an overall accuracy of 99.2% with minimal computational time. Likewise, combinations of feature generators with iChi2 and CNN also achieved 99.0%. Evidently, the proposed scheme ((AlexNet, ResNet101, SqueezeNet) + iChi2 + SVM) surpassed prior studies.

Contrary to previously well-established studies that utilized fewer samples of COVID-19 cases for training, this proposed study analyzed 4094 X-ray images labeled as COVID-19 positive and also practiced stationary-wavelet-transform-based data augmentation and the k-fold validation method to prevent overfitting. Undoubtedly, prior studies adopted pre-trained networks as a base classifier without implementing any feature generator and selector technique together. In contrast, this work is composed of various TL-based feature generators and several iterative feature-selector techniques that eliminates redundant features and significantly affects the accuracy of the proposed diagnosis decision-making system. Unlike previous studies, the experimental work of this study showed that selecting a relatively shallow network even produced optimal results with less computational time when trained with an augmented data set. Such schemes can be exploited for tasks discussed in [39,40] to manage health services more efficiently.

## 5. Conclusions

The recent outbreak of coronavirus disease has severely affected more than 560 million individuals around the globe, causing more than 6.3 million deaths to date. Besides smart diagnostic tools, presently, the only standard practiced to identify COVID-19 positive cases is by taking samples of the respiratory tract through viral nucleic acid testing. In this paper, we investigated a framework for COVID-19 and other viral/bacterial pneumonia identification from X-ray images by employing a combination of TL models (AlexNet, ResNet101, and SqueezeNet) as a feature generator, feature selectors (iNCA, iChi2, iMRMR), and ML classifiers. The TL networks were helpful for generating the relevant features

from X-ray images; the features were refined by applying a feature-selection approach using iNCA, iChi2, and iMRMR. The feature-selection process not only helped identify highly discriminative features, but it also reduced the processing speed for training the classifiers by removing unwanted or redundant features. Finally, CNN, LDA, and SVM were investigated for the detection of COVID-19-positive cases among viral/bacterial pneumonia and normal cases. In addition to this, the study exploited stationary wavelet transform to enhance a limited training dataset by decomposing each training image up to three levels. Furthermore, it also incorporated random rotation, translation, and shear to augment the data set by a factor of nine. The results presented in this study indicate that the proposed scheme can be used as a smart diagnostic tool for the identification and monitoring of COVID-19 cases, as well as other pneumonia cases, as it achieved an overall accuracy of 99.24%. The results indicated that using a combination of computer vision and deep learning techniques on X-ray images of lungs can help with the identification of COVID-19 cases to reduce the burden on healthcare workers in the time of the pandemic. In the future, the integration of clinical results along with radiography imaging can be exploited for the better detection and diagnosis of COVID-19-positive cases, with the addition of tuberculosis and lung cancer.

**Author Contributions:** For Conceptualization, J.R. and R.M.S.; methodology, J.R.; software, J.R. and R.M.S.; validation, J.R.; formal analysis, J.R.; investigation, R.M.S.; resources, J.R. and R.M.S.; data curation, J.R. and R.M.S.; writing—original draft preparation, J.R.; writing—review and editing, J.R. and R.M.S.; visualization, J.R.; supervision, J.R. and R.M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data is generated for this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Rasheed, J.; Jamil, A.; Hameed, A.A.; Aftab, U.; Aftab, J.; Shah, S.A.; Draheim, D. A survey on artificial intelligence approaches in supporting frontline workers and decision makers for the COVID-19 pandemic. *Chaos Solitons Fractals* **2020**, *141*, 110337. [CrossRef] [PubMed]
2. Chahar, S.; Roy, P.K. COVID-19: A Comprehensive Review of Learning Models. *Arch. Comput. Methods Eng.* **2022**, *29*, 1915–1940. [CrossRef] [PubMed]
3. Mahmud, F. Bangladesh Scientists Create \$3 Kit. Can It Help Detect COVID-19? Aljazeera. 2020. Available online: <https://www.aljazeera.com/news/2020/3/24/bangladesh-scientists-create-3-kit-can-it-help-detect-covid-19> (accessed on 30 December 2021).
4. Wang, W.; Xu, Y.; Gao, R.; Lu, R.; Han, K.; Wu, G.; Tan, W. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* **2020**, *323*, 1843–1844. [CrossRef] [PubMed]
5. Gostic, K.; Gomez, A.C.; Mummah, R.O.; Kucharski, A.J.; Lloyd-Smith, J.O. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife* **2020**, *9*, e55570. [CrossRef]
6. Wang, S.; Kang, B.; Ma, J.; Zeng, X.; Xiao, M.; Guo, J.; Cai, M.; Yang, J.; Li, Y.; Meng, X.; et al. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). *medRxiv* **2020**. [CrossRef]
7. Yahyaoui, A.; Jamil, A.; Rasheed, J.; Yesiltepe, M. A Decision Support System for Diabetes Prediction Using Machine Learning and Deep Learning Techniques. In Proceedings of the 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara, Turkey, 6–7 November 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.
8. Nallamala, S.H.; Mishra, P.; Koneru, S.V. Breast Cancer Detection using Machine Learning Way. *Int. J. Recent Technol. Eng.* **2019**, *8*, 1402–1405. [CrossRef]
9. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131.e9. [CrossRef]
10. Rasheed, J.; Hameed, A.A.; Ajlouni, N.; Jamil, A.; Özyavaş, A.; Orman, Z. Application of Adaptive Back-Propagation Neural Networks for Parkinson’s Disease Prediction. In Proceedings of the 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 26–27 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.

11. Gurbina, M.; Lascu, M.; Lascu, D. Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines. In Proceedings of the 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 1–3 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 505–508.
12. Li, D.; Li, S. An artificial intelligence deep learning platform achieves high diagnostic accuracy for COVID-19 pneumonia by reading chest X-ray images. *iScience* **2022**, *25*, 104031. [[CrossRef](#)]
13. Ortiz-Toro, C.; García-Pedrero, A.; Lillo-Saavedra, M.; Gonzalo-Martín, C. Automatic detection of pneumonia in chest X-ray images using textural features. *Comput. Biol. Med.* **2022**, *145*, 105466. [[CrossRef](#)]
14. Monday, H.N.; Li, J.; Nneji, G.U.; Nahar, S.; Hossin, M.A.; Jackson, J. COVID-19 Pneumonia Classification Based on NeuroWavelet Capsule Network. *Healthcare* **2022**, *10*, 422. [[CrossRef](#)]
15. Abunadi, I.; Albraikan, A.A.; Alzahrani, J.S.; Eltahir, M.M.; Hilal, A.M.; Eldesouki, M.I.; Motwakel, A.; Yaseen, I. An Automated Glowworm Swarm Optimization with an Inception-Based Deep Convolutional Neural Network for COVID-19 Diagnosis and Classification. *Healthcare* **2022**, *10*, 697. [[CrossRef](#)] [[PubMed](#)]
16. Trung Huynh, H.; Anh, V.N.N. A Deep Learning Method for Lung Segmentation on Large Size Chest X-ray Image. In Proceedings of the 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF), Danang, Vietnam, 20–22 March 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5.
17. Avola, D.; Bacciu, A.; Cinque, L.; Fagioli, A.; Marini, M.R.; Taiello, R. Study on transfer learning capabilities for pneumonia classification in chest-X-rays images. *Comput. Methods Programs Biomed.* **2022**, *221*, 106833. [[CrossRef](#)] [[PubMed](#)]
18. Jain, R.; Gupta, M.; Taneja, S.; Hemanth, D.J. Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Appl. Intell.* **2020**, *51*, 1690–1700. [[CrossRef](#)] [[PubMed](#)]
19. Sri Kavya, N.; Shilpa, T.; Veeranjaneyulu, N.; Divya Priya, D. Detecting COVID-19 and pneumonia from chest X-ray images using deep convolutional neural networks. *Mater. Today Proc.* **2022**. [[CrossRef](#)]
20. Nason, G.P.; Silverman, B.W. The Stationary Wavelet Transform and some Statistical Applications. In *Wavelets and Statistics*; Springer: Berlin, Germany, 1995; pp. 281–299.
21. Pesquet, J.-C.; Krim, H.; Carfantan, H. Time-invariant orthonormal wavelet representations. *IEEE Trans. Signal Process.* **1996**, *44*, 1964–1970. [[CrossRef](#)]
22. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
23. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv* **2012**. [[CrossRef](#)]
24. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 770–778.
25. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**. [[CrossRef](#)]
26. Goldberger, J.; Roweis, S.; Hinton, G.; Salakhutdinov, R. Neighbourhood components analysis. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2005.
27. Ding, C.; Peng, H. Minimum Redundancy Feature Selection from Microarray Gene Expression Data. *J. Bioinform. Comput. Biol.* **2005**, *3*, 185–205. [[CrossRef](#)]
28. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
29. Rasheed, J. Analyzing the Effect of Filtering and Feature-Extraction Techniques in a Machine Learning Model for Identification of Infectious Disease Using Radiography Imaging. *Symmetry* **2022**, *14*, 1398. [[CrossRef](#)]
30. Sethy, P.K.; Behera, S.K.; Ratha, P.K.; Biswas, P. Detection of coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine. *Int. J. Math. Eng. Manag. Sci.* **2020**, *5*, 643–651. [[CrossRef](#)]
31. Kumar, R.; Arora, R.; Bansal, V.; Sahayasheela, V.J.; Buckchash, H.; Imran, J.; Narayanan, N.; Pandian, G.N.; Raman, B. Classification of COVID-19 from chest X-ray images using deep features and correlation coefficient. *Multimed. Tools Appl.* **2022**. [[CrossRef](#)] [[PubMed](#)]
32. Toraman, S.; Alakus, T.B.; Turkoglu, I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* **2020**, *140*, 110122. [[CrossRef](#)]
33. Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, A. Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays. *Comput. Methods Programs Biomed.* **2020**, *196*, 105608. [[CrossRef](#)]
34. Khan, A.I.; Shah, J.L.; Bhat, M.M. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. *Comput. Methods Programs Biomed.* **2020**, *196*, 105581. [[CrossRef](#)]
35. Mahmud, T.; Rahman, M.A.; Fattah, S.A. 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.* **2020**, *122*, 103869. [[CrossRef](#)] [[PubMed](#)]
36. Rahimzadeh, M.; Attar, A. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inform. Med. Unlocked* **2020**, *19*, 100360. [[CrossRef](#)] [[PubMed](#)]

37. Wang, L.; Lin, Z.Q.; Wong, A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci. Rep.* **2020**, *10*, 19549. [[CrossRef](#)]
38. Loey, M.; Smarandache, F.M.; Khalifa, N.E. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. *Symmetry* **2020**, *12*, 651. [[CrossRef](#)]
39. Giovannini, S.; Onder, G.; Lattanzio, F.; Bustacchini, S.; Di Stefano, G.; Moresi, R.; Russo, A.; Bernabei, R.; Landi, F. Selenium Concentrations and Mortality among Community-Dwelling Older Adults: Results from ilSIRENTE Study. *J. Nutr. Health Aging* **2018**, *22*, 608–612. [[CrossRef](#)]
40. Onder, G.; Giovannini, S.; Sganga, F.; Manes-Gravina, E.; Topinkova, E.; Finne-Soveri, H.; Garms-Homolová, V.; Declercq, A.; van der Roest, H.G.; Jónsson, P.V.; et al. Interactions between drugs and geriatric syndromes in nursing home and home care: Results from Shelter and IBenC projects. *Aging Clin. Exp. Res.* **2018**, *30*, 1015–1021. [[CrossRef](#)] [[PubMed](#)]