

Article

Automated Diagnosis of Childhood Pneumonia in Chest Radiographs Using Modified Densely Residual Bottleneck-Layer Features

Sinan Alkassar ^{1,*}, Mohammed A. M. Abdullah ¹, Bilal A. Jebur ^{1,*}, Ghassan H. Abdul-Majeed ²,
Bo Wei ³ and Wai Lok Woo ^{3,*}

- ¹ Computer and Information Engineering Department, Electronics Engineering College, Ninevah University, Mosul 41002, Iraq; mohammed.abdulguttaleb@uoninevah.edu.iq
² Scientific Affairs Department, University of Baghdad, Karrada, Baghdad 10071, Iraq; ghassan@uob.edu.iq
³ Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK; bo.wei@northumbria.ac.uk
* Correspondence: sinan.alkassar@uoninevah.edu.iq (S.A.); bilal.jebur@uoninevah.edu.iq (B.A.J.); wailok.woo@northumbria.ac.uk (W.L.W.)

Abstract: Pneumonia is a severe infection that affects the lungs due to viral or bacterial infections such as the novel COVID-19 virus resulting in mild to critical health conditions. One way to diagnose pneumonia is to screen prospective patient's lungs using either a Computed Tomography (CT) scan or chest X-ray. To help radiologists in processing a large amount of data especially during pandemics, and to overcome some limitations in deep learning approaches, this paper introduces a new approach that utilizes a few light-weighted densely connected bottleneck residual block features to extract rich spatial information. Then, shrinking data batches into a single vector using four efficient methods. Next, an adaptive weight setup is proposed utilizing Adaboost ensemble learning which adaptively sets weight for each classifier depending on the scores generated to achieve the highest true positive rates while maintaining low negative rates. The proposed method is evaluated using the Kaggle chest X-ray public dataset and attained an accuracy of 99.6% showing superiority to other deep networks-based pneumonia diagnosis methods.

Keywords: pneumonia detection; X-ray images; deep networks; COVID-19



Citation: Alkassar, S.; Abdullah, M.A.M.; Jebur, B.A.; Abdul-Majeed, G.H.; Wei, B.; Woo, W.L. Automated Diagnosis of Childhood Pneumonia in Chest Radiographs Using Modified Densely Residual Bottleneck-Layer Features. *Appl. Sci.* **2021**, *11*, 11461. <https://doi.org/10.3390/app112311461>

Academic Editors: Nikolaos Kourkoumelis and Edgar Guevara

Received: 24 October 2021
Accepted: 30 November 2021
Published: 3 December 2021

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



Copyright: © 2021 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/>).

1. Introduction

Pneumonia is mainly caused by virus pathogens or bacteria pathogens that infect the balloon-shaped air sacks in the human lungs causing inflammation in these sacks and serious implications on patient health. Pneumonia according to the Health World Organization (WHO) [1,2] is the main cause of death in young children under 5 years old, recording an 18% death rate. Furthermore, one of the devastating viruses that affect the lungs and cause pneumonia at an advanced stage is the novel COVID-19 virus which was declared as a global pandemic by WHO in March 2020 [3]. In particular, the new variant of the virus has been shown to affect young children. Some of the common symptoms of bacterial and viral pneumonia include fever, cough, increased breathing rate, and breathing difficulty. However, while bacterial pneumonia can be treated using special antibiotics, viral pneumonia is still challenging. Thus, further diagnosis and supporting techniques are necessary such as radiography imaging. Radiographic images are one of the effective diagnostic methods which are captured by either using Chest X-ray (CXR) radiography or Computed Tomography (CT). Although these images can help radiologists in their diagnosis, however in some viral, bacterial, or other inflammatory lung diseases, CXR images might show similar blurry white areas thus making the diagnosis task rather challenging [4]. Examples of healthy and infected lungs in CXR images are shown in Figure 1.

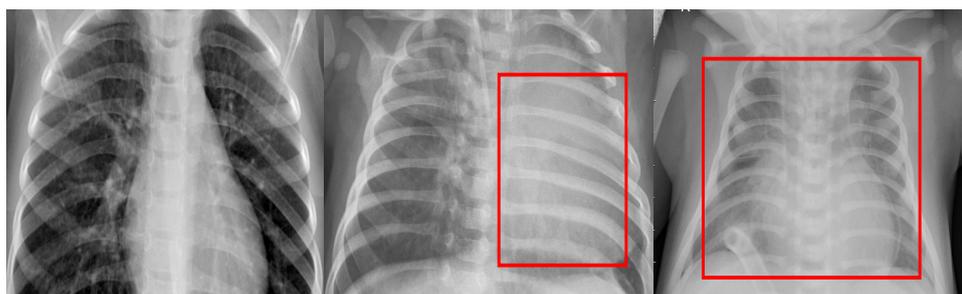


Figure 1. Examples of three CXR images of child human lungs. The left-hand image represents a healthy case, the middle image depicts a viral infection, while a lung with a bacterial infection is shown on the left.

Nowadays, Artificial Intelligence (AI) techniques such as Deep learning Convolutional Neural Network (CNN) techniques have emerged as a supporting tool for distinguishing various features of lung infections as these networks have achieved an outstanding efficiency for feature extraction through representation learning and classification. In essence, these techniques have also made a huge step in the development of disease diagnosis systems via processing an immense number of clinical digital images and interpreting the spatial information of these images, which are prone to various types of noise. However, as these networks went deeper going from few layers such as AlexNet [5] to hundreds such as ResNet [6] as the demand to boost accuracy leading to deeper networks resulting in higher complexity cost and inflated number of parameters [7].

One answer to the aforementioned issues is transfer learning. Transfer learning is the strength of a learning algorithm to utilize similarities between various learning tasks such as similar image representations sharing statistical power and transfer learning information over other tasks. Using these representations has been motivated by the fact that they tend to describe many general priors that are not task-specific but would be beneficial for a machine to solve feature learning tasks [5]. Hence, the power of representation learning through utilizing pre-trained deep networks appears because these networks promote ease of feature reuse and extract more abstract features at higher layers which tend to be invariant to most local changes of the input images [8]. In general, deep-learning networks consist of multiple layers employed to progressively extract high-level features from raw data. The strength of transfer learning feature for medical data is demonstrated for instance through breast cancer classification using histopathology biopsy images [9] and brain tumor semantic segmentation [10]. Yet, there are still challenges that accompany the transfer learning process particularly the negative transfer and task-mapping automation [11]. Hence, a different approach for exploiting transfer learning is necessary to mitigate the aforementioned issues [12].

1.1. Related Work

There has been extensive research on detecting pneumonia using deep learning techniques. As a case in point, Chowdhury et al. [4] tested four different pre-trained deep networks for detecting pneumonia in CXR images through transfer learning and concluded that SqueezeNet [13] outperformed the other three networks. On the other hand, a fine-tuned version of a typical off-the-shelf deep network has been utilized in [14] to automatically classify pneumonia binary images. A CNN network with 10 layers was used by Saraiva et al. [15] to detect infant pneumonia using CXR images. Furthermore, Apostolopoulos and Mpesiana [16] have used transfer learning techniques with a CNN such as VGG19 [17] and MobileNet v2 to classify CXR lung images with pneumonia. Moreover, VGG16 [17], DenseNet [18], Xception [19], and Inception networks [20] along with transfer learning are utilized in [21]. Abdullah et al. [10] proposed a method for brain tumor segmentation using CNN and transfer learning. Kermany et al. [22], have investigated medical diagnosis on treatable diseases such as pneumonia in CXR images utilizing transfer learning with pre-trained deep networks. Togacar et al. [23] on the other

hand have extracted features from deep layers using Alexnet, VGG16, and VGG19 deep networks and applied the minimum Redundancy Maximum Relevance (mRMR) algorithm for feature reduction.

Rajpurkar et al. [24] developed a deep network named CheXNet which is a 121-layer Dense network applied on the ChestX-ray14 dataset [25]. A contrastive learning with supervised fine tuning was used in the RSNA Kaggle Pneumonia challenge [26,27] and achieved 88% accuracy. Chouhan et al. [28] utilized transfer learning for extracting features from five pre-trained deep networks stated as AlexNet, GoogleNet [29], Inception V3, ResNet18, and DenseNet121. Whereas in [30], four deep pre-trained networks were compared and showed that DenseNet201 outperforms AlexNet, ResNet18, and SqueezeNet in terms of accuracy. Zhang et al. [31] formulated the detection of a viral infection in the lungs as a one-class classification-based anomaly detection problem where an anomaly score is measured for each CXR image and a decision is made based on a threshold. Ayan et al. [32] proposed using transfer learning with seven pre-trained famous deep networks such as VGG16, ResNet50, and SqueezeNet along with an ensemble method based on probabilistic voting to diagnose CXR images. A CNN with Extreme Machine Learning (EML) and Principal Component Analysis (PCA) are utilized in [33] after enhancing the contrast of CXR images using Contrast Limited Adaptive Histogram Equalisation (CLAHE) to diagnose pneumonia. Finally, Gour and Jain [34] proposed Uncertainty-Aware Convolutional neural Network (UA-ConvNet) which is based on the fine-tuned EfficientNet-B3 model for CXR images.

1.2. Limitations and Proposed Work

Although deep learning networks have achieved an efficient disease diagnosis accuracy, their performance has shown some drawback which can be summarized as: (1) increasing network depth does not ensure higher diagnosis accuracy as simply stacking layers together might cause the training error to be higher [35] on top of increasing complexity and computational cost; (2) deep networks are hard to train because of the vanishing gradient problem as the gradient is back-propagated to more initial layers [35]. As a result, when training deep networks, their performance becomes saturated or even begins to degrade rapidly; (3) transfer learning has some problems related to negative learning and the automation of task mapping; (4) features extracted from deep layers lose important spatial information in the CXR images as these deep layers are either identity mappings or copies of the early layers; (5) depending on single descriptor may introduce more classification errors; and (6) most recent work has evaluated their prediction models in only healthy lung vs. either bacterial or viral infected lung CXR images.

Therefore in this paper, our contributions can be summarized as (1) instead of using a complete deep network, we propose a light-weighted bottleneck layer feature descriptors exploiting the residual building blocks suggested in [6] and dense blocks suggested in [18] for building an improved feature descriptor which introduces no extra parameters thereby achieving low training errors and time; (2) the extracted features are reduced based on an efficient method for feature reduction, wherein 3D map of features generated for each image are shrunk into one vector using four methods; (3) adaptive score fusion with learned weights is performed using Adaboost ensemble learning that for each iteration, the decision power for prediction model weights is altered; and (4) the proposed model is examined in three scenarios which are: Normal vs. infected lungs and normal vs. bacterial vs. viral infected lungs.

The organization of this paper is as follows: The proposed method including feature extraction and reduction, the prediction model, and the adaptive score fusion are explained in Section 2. Results and discussion are discussed in Section 3 whereas we draw our conclusion and feature work in Section 4.

2. Proposed Method

2.1. Feature Extraction and Reduction

In general, auto-encoders convert higher dimensional input data into lower dimensional more abstract information through encoding data into another form. Then, the encoded data can be reconstructed to an approximate shape of the original input data depending on the reconstruction error [8,36]. In our work, we utilize the encoding process only and apply post-processing steps to increase the diagnosis accuracy. To avoid learning everything from scratch, we propose an efficient method through feature reuse to extract high spatial information characteristics deduced through abstract features utilizing ResNet auto-encoder bottleneck layers. ResNet has shown exceptional performance in several challenging classification task competitions such as the competition on ImageNet dataset [37].

The encoding process can be defined as a feature-extraction function denoted as ζ_ϕ applied on training data $\{x_1, x_2, \dots, x_N\}$ as:

$$\zeta_\phi(x_n) = \sigma(Wx_n + b), \tag{1}$$

where x represents the input image data vector and N is the number of images producing $c_n = \zeta_\phi(x_n)$ which is the first coded representation of the n input image. The parameters b and W are the encoder bias vector and weights matrices, respectively, whereas σ is the activation function which typically is either sigmoid or Rectified Linear Unit (RELU) function.

Given that a CXR image has a size of $a \times b$, the image is first resized to the size of the ResNet input image layer, i.e., 224×224 , and is then passed to the building blocks of the ResNet network to extract the raw features from that image. Each block, which is denoted as $F_l(c_n)$ where l is the number of bottleneck blocks, consists of 3 weighted convolution layers with Batch Normalization (BN) and RELU as shown in Figure 2. Since a skip connection has some advantages such as no extra parameters, proven to improve deep network performance [6], and can smooth information propagation, therefore, we propose adding more skip connections in a dense-fashion way [18] which allow feature reuse and leading to more accurate and compact learning.

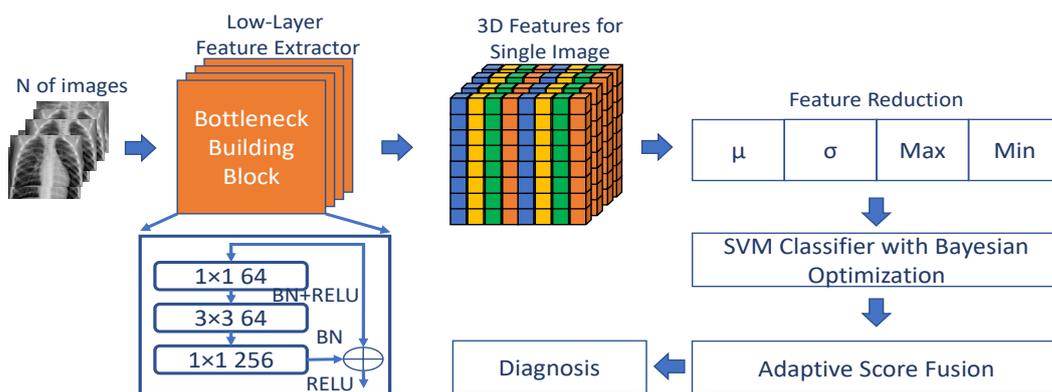


Figure 2. The proposed pneumonia diagnosis system depicting feature extraction via residual building blocks, feature reduction, SVM classifier training with Bayesian optimization, and score fusion.

Thus, we can define the resulted code for one image after F_l blocks with skip connections known as identity mapping from previous blocks as:

$$c_{n,l} = \sum_{i=2}^l (h(c_{n,1:i-1}), F_i(c_n)), \tag{2}$$

where $h(c_n)$ represents the identity mapping. As depicted in Figure 3, each residual block has an identity mapping as shown in sub-Figure 3a and each layer in the dense block is

fed with information from all previous layers as shown in sub-Figure 3b, whereas the proposed feature extraction descriptor takes advantage of both paradigms where each residual block is fed with information from previous blocks through skip connections as shown in sub-Figure 3c. For instance, $c_{n,3}$ is the sum of $h(c_{n,1})$, $h(c_{n,2})$, and $F_2(c_n)$.

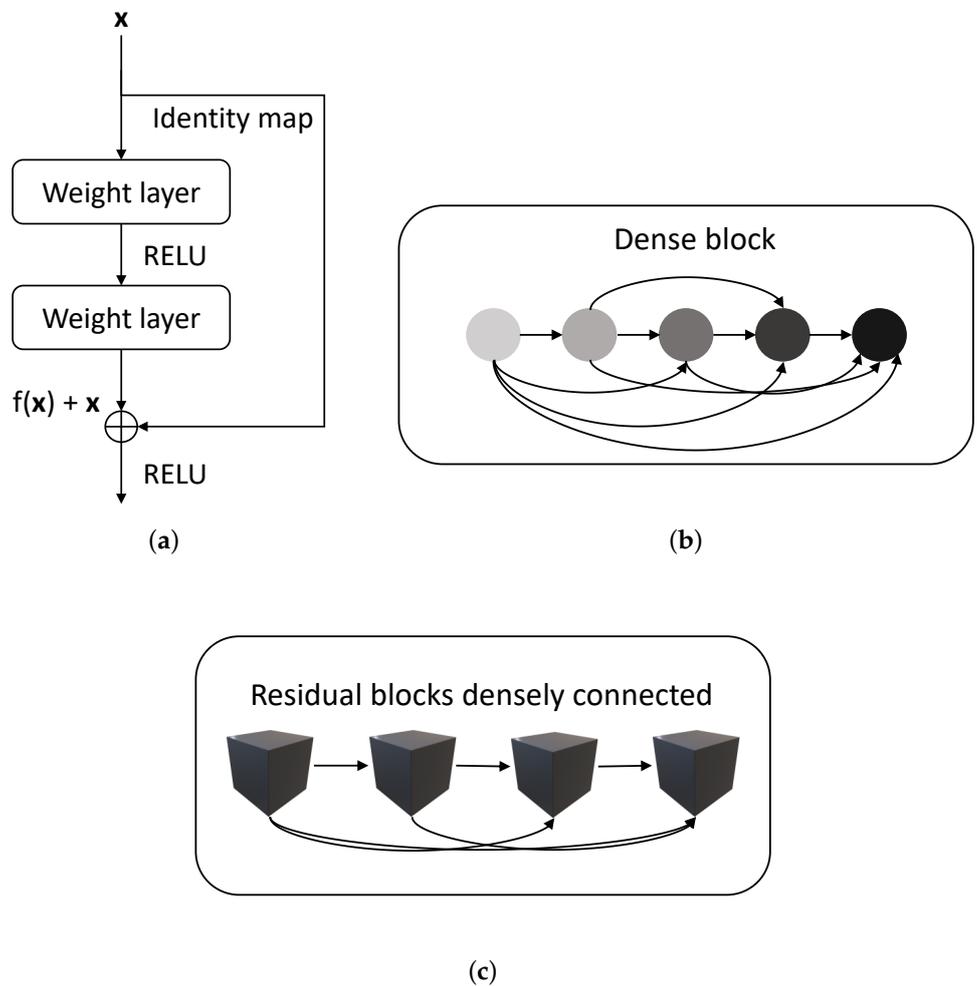


Figure 3. The modified connections of residual units showing each block are densely connected to previous blocks. (a) Residual block, (b) dense block, and (c) the modified residual blocks densely connected.

As a consequence of many input images being processed through the ResNet layers each having different convolution and batch size, stride, and padding, leading to a higher data dimensionality. To alleviate this problem, feature reduction techniques are performed on each slide of the 3D data resulting in reduced feature points for each input image. These techniques include calculating one point for each slide based on four rules, i.e., $\{\min, \max, \sigma, \mu\}$, where μ is the mean and σ is the standard deviation. This will result in four feature vectors of size $1 \times h$, i.e., $c_{n_{\max}}$, $c_{n_{\min}}$, $c_{n_{\sigma}}$, and $c_{n_{\mu}}$ where h represents the feature vector length.

2.2. Prediction Model

Next for classification, each matrix (i.e., $C_{N_{\max}}$, $C_{N_{\min}}$, $C_{N_{\sigma}}$, and $C_{N_{\mu}}$) is divided into training and testing observations and utilized to create a prediction model based on Support Vector Machine (SVM). SVM has proven to perform well in many classification tasks although its hyper-parameters and regularization term must be tuned carefully during the training process to achieve the highest accuracy resulting in high computational

cost [38]. Therefore, Bayesian optimization is used to find the best hyper-parameters and thus improving the prediction model [39].

Given N observations $\{\mathbf{c}_n, \mathbf{y}_n\}_{n=1}^N$ where $\mathbf{c}_n \in \mathbb{R}^N$ and \mathbf{y}_n is the corresponding category vector for each \mathbf{c}_n , the optimal score function f is found by SVM through solving the regularized risk minimization objective using hinge loss [40,41] as an optimization problem such:

$$\arg \min_{f(x)} \sum_{n=1}^N \max(0, 1 - \mathbf{y}_n f(\mathbf{c}_n)) + \gamma R(f) \tag{3}$$

where $\max(0, 1 - \mathbf{y}_n f(\mathbf{c}_n))$ is the hinge loss for the classification function $f(\mathbf{c}_n)$, γ is the hyper-parameter to fine-tune f training error versus complexity, and R is the regularize function. For nonlinear classifier, a kernel method is used in SVM which transform $f(\mathbf{c}_n)$ into higher transformed data-point function $\varphi(\mathbf{c}_n)$ using a kernel function $k(\mathbf{c}_n, \mathbf{c}_m) = \varphi(\mathbf{c}_n) \cdot \varphi(\mathbf{c}_m)$. The hyper-parameters are obtained by maximizing:

$$\max_{\alpha_{n=1:N}} = \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{m=1}^N \mathbf{c}_n \mathbf{y}_n k(\mathbf{c}_n, \mathbf{c}_m) \mathbf{c}_m \mathbf{y}_m. \tag{4}$$

Various kernel functions, scales, and constrained are examined such as linear, quadratic, cubic, and Gaussian kernels.

2.3. Adaptive Score Fusion

To make the final prediction, an adaptive score fusion technique is proposed based on Adaboost ensemble learning [42] where the weights for each classifier are updated based on previous errors made by these classifiers. This is shown in Algorithm 1. First, each classifier is given a weight equally. Second, for several iterations, each classifier is trained and classification error is calculated, and based on that error, each classifier weight is updated where the higher the weighted error, the less the decision power is given to the corresponding classifier. Finally, a weighted sum rule is exploited where the final score is determined via summing the four scores such that:

$$S_f = \sum_{i=1}^4 S_i \cdot \hat{w}_i, \tag{5}$$

where S_f is the final score, S_i is the binary score generated from one of the four prediction models where the value 1 refers to positive pneumonia diagnosis.

Algorithm 1: Adaboost ensemble weights learning.

```

initialize classifier weights as  $w_i = \frac{1}{\text{No.of classifiers}}$ ;
while While No. of iterations is not reached do
    train each classifier using SVM with Bayesian optimization;
    calculate the error of each classifier ( $e_i$ );
    calculate each classifier weight in the ensemble as  $w_{e_i} = l_r \times \log \frac{1-e_i}{e_i}$ ;
    update  $w_i$  as  $\hat{w}_i = w_i * \exp^{w_{e_i}}$ ;
    normalize  $\hat{w}_i \leftarrow \frac{\hat{w}_i}{\sum_{i=1}^{\text{no of classifiers}} w_i}$ 
end

```

3. Results and Discussion

The proposed method is evaluated using a publicly available dataset named Kaggle collected and labeled by Kermany et al. [22,43]. The dataset includes 5232 CXR images collected from a cohort of pediatric patients of age from one to five years old. Image resolution varies from 400 to 2000 pixels with common noise factors introduced in the CXR

imaging such as the position of the patient during image capture, the screening device, medical sensors and tools fixed on the patient while screening, and other inherited health conditions. In particular, the dataset consists of 3883 pneumonia-diagnosed CXR images including bacterial and viral infection and 1349 normal CXR images. As shown in Table 1, we trained the proposed system utilizing 75% of the dataset with 10-K fold validation to avoid over-fitting achieving a training time of 96.57 s whereas the rest of the data are used for testing with image augmentation including rotation, scaling, and translation.

Table 1. Training, Validation, and testing CXR image number using Kaggle CXR dataset.

Category	No. of Images	Training	Testing
Normal	1349	1012	337
Bacterial	2538	1903	635
Viral	1345	1008	337
Total	5232	3922	1309

The experiment is conducted on a PC with core i7 and 16 GB of RAM under the Matlab 2020b environment and we utilized ResNet50 as the backbone network which was modified using Matlab Deep Learning Design Application. The input size of each image is changed to 224×224 as the ResNet50 input layer with batch normalization utilized before the activation layer and before the convolution layer. The list of layers before the $l = 5$ residual bottle-neck building blocks depicted in Figure 2 can be listed as the first convolution layer with 64 (7×7) filters with stride = 2, batch normalization, RELU, and 3×3 max pooling. Stochastic Gradient Descent (SGD) with a mini-batch size of 256 is used. The learning rate is set to 0.1 and is divided by 10 when the error is high whereas the weight decay is set to 0.0001 and a momentum of 0.9.

For evaluating the proposed method, three metrics are employed namely: *sensitivity*, *F1 score*, and *accuracy*. These metrics can be defined as:

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (6)$$

$$F1 \text{ score} = \frac{2TP}{2TP + FP + FN} \quad (7)$$

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

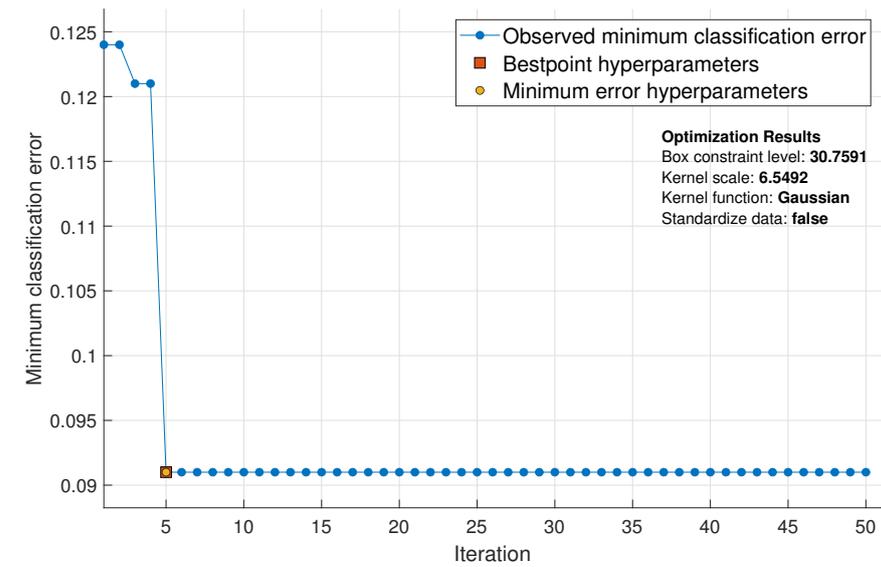
where TP and TN are the true positives and negatives, respectively, and FP and FN are the false positives and negatives consequently. As shown in Table 2, only $l = 5$ bottleneck blocks are needed, which is neither deep nor early layers to achieve the best performance (99.6%) as the accuracy degraded after $l = 6$ although the number of features is the same from $l = 4$. This is an evidence that going deeper using CNN can complicate the classification system while degrading its accuracy.

Table 2. Performance of the number of modified ResNet bottleneck layers.

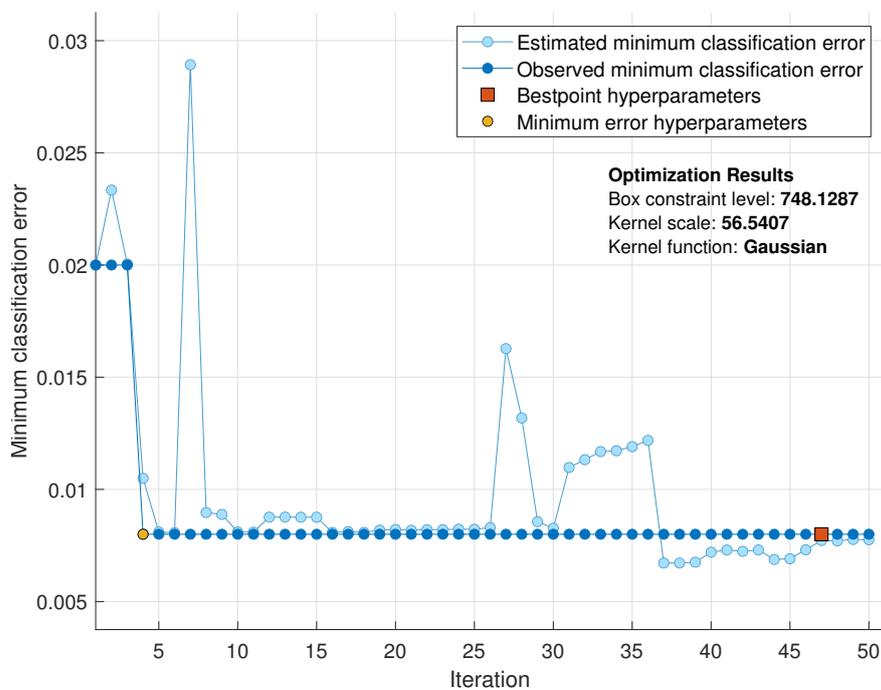
Number of Bottleneck Blocks	Accuracy	Number of Features
1	95.6	1×256
2	96.2	1×256
3	96.8	1×256
4	96.8	1×512
5	99.6	1×512
6	97.6	1×512
7	97.2	1×512

3.1. SVM Optimization Using Bayesian Optimization

First, 50 iterations were set for the Bayesian optimizer of the SVM classifier as shown in Figure 4 where minimum classification error is plotted at each iteration. Figure 4 depicts the optimized hyper-parameters for the SVM including kernel function, kernel scale, and box constraint level. As shown in sub-Figure 4b, the optimization needed only 4 iterations to achieve the optimal SVM hyper-parameters achieving lower than 1% of minimum classification error. In comparison, the random search optimization of the SVM reached higher than 9% of the minimum classification error as shown in sub-Figure 4a.



(a)



(b)

Figure 4. Minimum classification error plot of the optimized SVM classifier for the pneumonia diagnosis system using Bayesian optimization vs. random search optimization. (a) Random search optimization, and (b) Bayesian optimization.

Then, for weight learning using Adaboost, we set the number of trials of testing to 20 and the weights for each classifier is set equally for the $S_i(\mu)$, $S_i(\sigma)$, $S_i(\max)$, and $S_i(\min)$, respectively, where Algorithm 1 is used to obtain the optimum final score S_f . We plotted the ROC curve for each sub-classifier (min, max, σ , μ) compared to S_f as shown in Figure 5. The feature reduction techniques using mean and standard deviation for each batch have achieved the highest weights, subsequently these two feature vectors contribute in 75% of the decision power. The S_f performed better than other classifiers achieving an area under the curve of 99.6% supporting that the proposed method has significance in terms of adaptive weight setting over a single decision process.

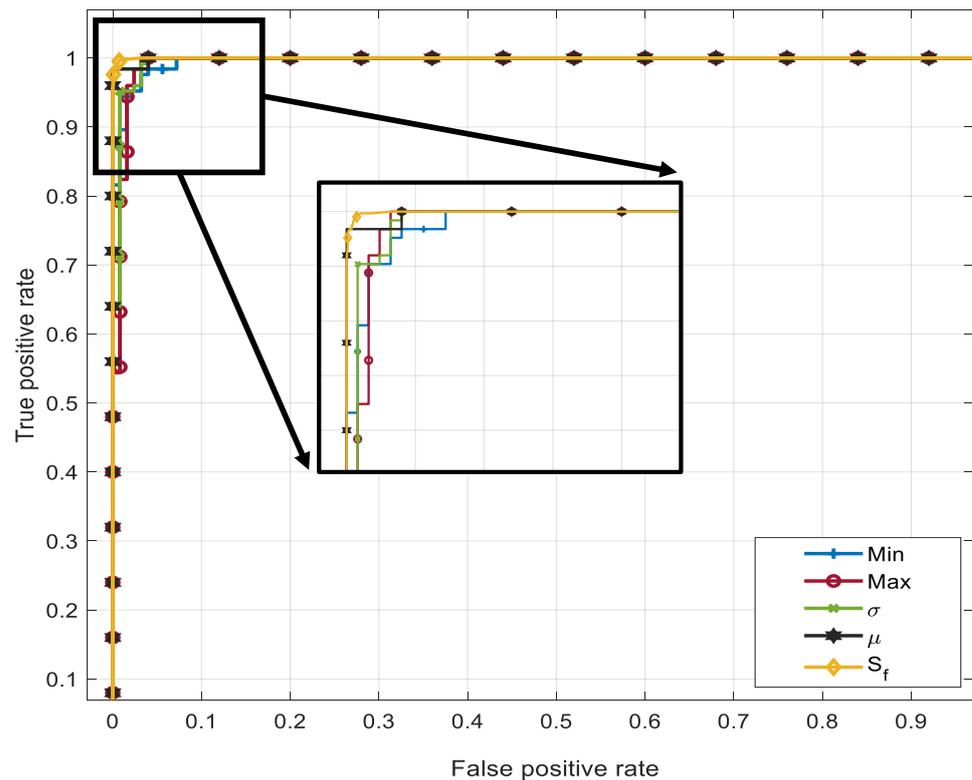


Figure 5. ROC curves of four sub-classifiers depicting $S_i(\mu)$, $S_i(\sigma)$, $S_i(\max)$, and $S_i(\min)$ scores compared to the accuracy of the proposed score fusion S_f .

3.2. Normal vs. Bacterial vs. Viral Pneumonia Infected Lungs

The most challenging scenario is when testing the prediction model utilizing bacterial vs. viral CXR lung images. According to a study conducted by Swingler [44] on the differentiation between viral and bacterial pneumonia in children using CXR images, many radiologists have agreed on distinguishing bacterial cases in an accuracy ranging from 26% to 70%. The review findings raise intriguing questions regarding the accuracy of lung pneumonia diagnosis and highlighted the complexity of deciding by either a human expert or a prediction model. This is evident in the feature space shown in sub-Figure 6a where the feature points of the viral and bacterial lung infections are overlapping. Next, a similar procedure, where SVM with Bayesian optimization and adaptive score fusion, is applied as shown in sub-Figure 4, Figure 7 and in Figure 8. The minimum classification error plot shows that even after 50 iterations, the classification error has not improved reaching 9%. In addition, the confusion matrix in Figure 8 shows that the prediction model has performed better in detecting viral pneumonia. Nevertheless, the overall prediction model accuracy is lower when classifying normal vs. viral and bacterial pneumonia compared to healthy vs. infected lung CXR images scoring 89.5% yet higher when compared to deep features.

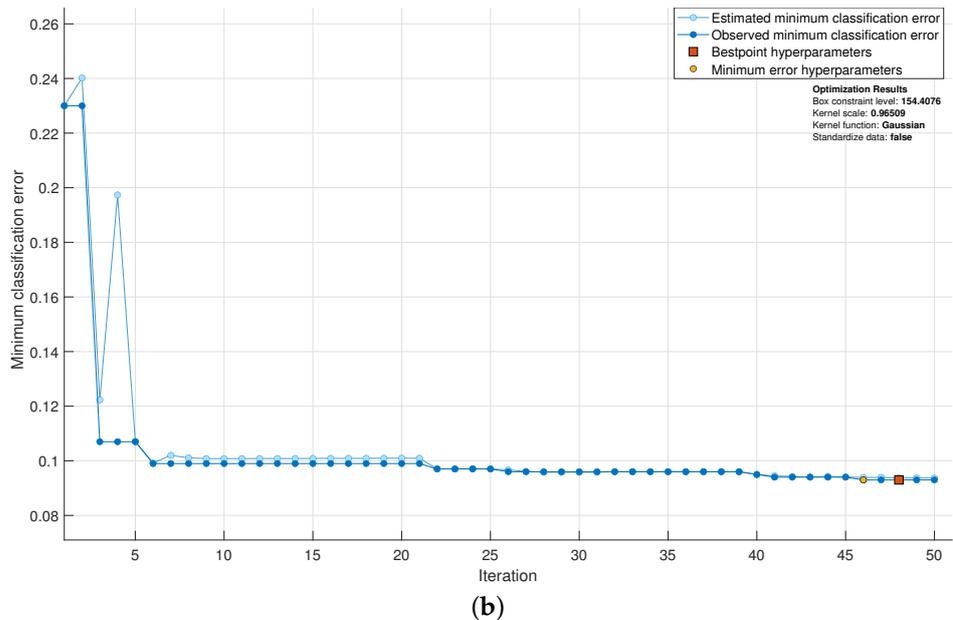
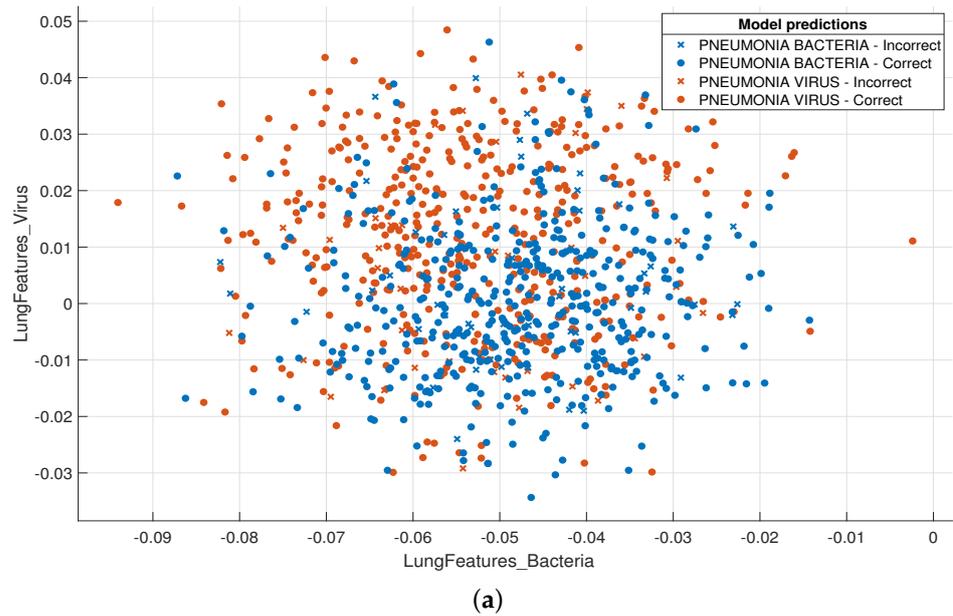


Figure 6. Performance of proposed model using bacterial vs. viral pneumonia scenario. (a) is the feature space of bacterial vs. viral infected lung CXR images, and (b) is the SVM with Bayesian optimization of the proposed scenario.

3.3. Normal vs. Pneumonia Infected Lungs

The first scenario conducted is to test the proposed algorithm in a normal vs. infected lungs fashion including a mix of bacterial and viral infection CXR images. First, the confusion matrix of the proposed feature extraction method compared to deep-layer features is shown in Figure 7. The sensitivity, F1 score and accuracy are (99.4%, 99.2%, and 99.6% respectively compared to features extracted from deep bottleneck layers, i.e., when l is larger than 5. It is evident that in our case, going deeper in some applications and in particular using medical images for diagnosis could have a negative impact on system performance. Additionally, to illustrate the advantage of using the modified residual building blocks as an efficient feature extractor, we conducted the same experiment on most off-the-shelf deep networks as shown in Table 3. We extracted low-layer features

from these networks and our proposed method has the upper hand in terms of accuracy achieving 99.6%.

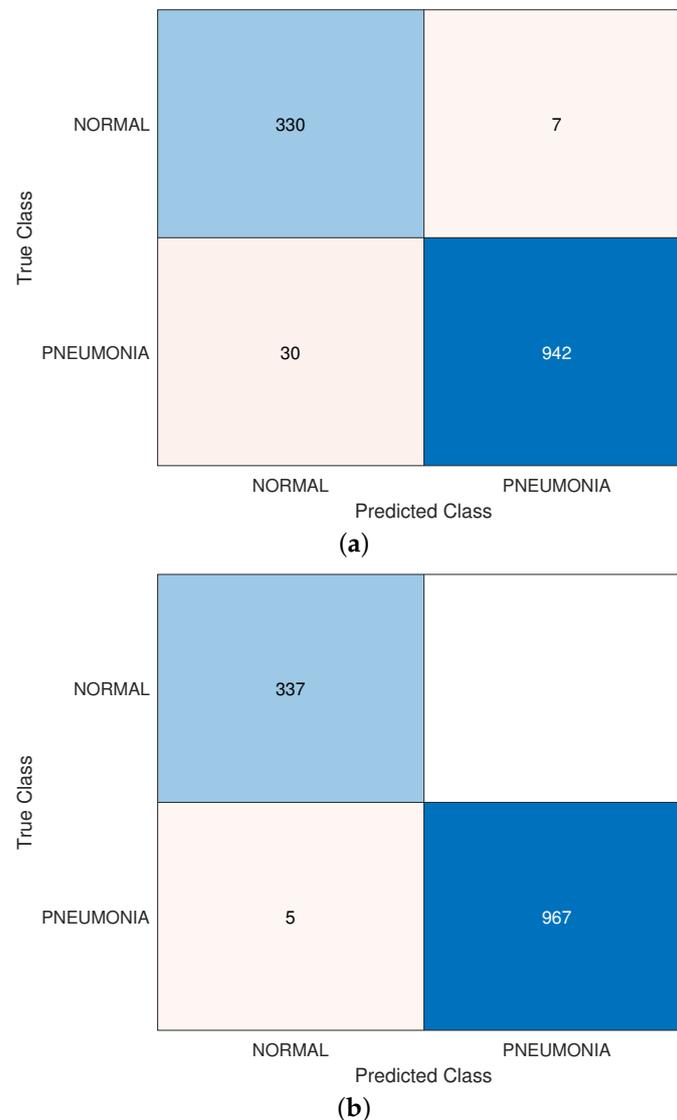


Figure 7. The confusion matrices showing a comparison between (a) deep-layer features and (b) features extracted using proposed method.

Next, we compared our work with recent state-of-the-art methods for pneumonia diagnosis using CXR images as shown in Table 4. It is evident that the proposed method has outperformed recent transfer learning methods in terms of accuracy, training time, and complexity as only a few modified residual blocks are needed and trained to achieve higher accuracy on a simple PC providing a feasible solution. It is worth pointing out that our proposed method is slightly better than the work in [23] by only 0.02% and lower than the work in [33] by only 0.04%, our proposed method however utilizes only five residual building blocks densely connected to extract features. Whereas the method in [23] involves transfer learning and re-train three different deep networks extracting an abundance amount of features; the classification method in [33] uses mix of various algorithms such as PCA, CNN, CLAHE, and EML. As a consequence, more complex feature reduction and more processing time are required which subsequently increase system complexity. Besides, our proposed method uses an adaptive weight setting method to assign the decision power effectively to the best feature descriptor.

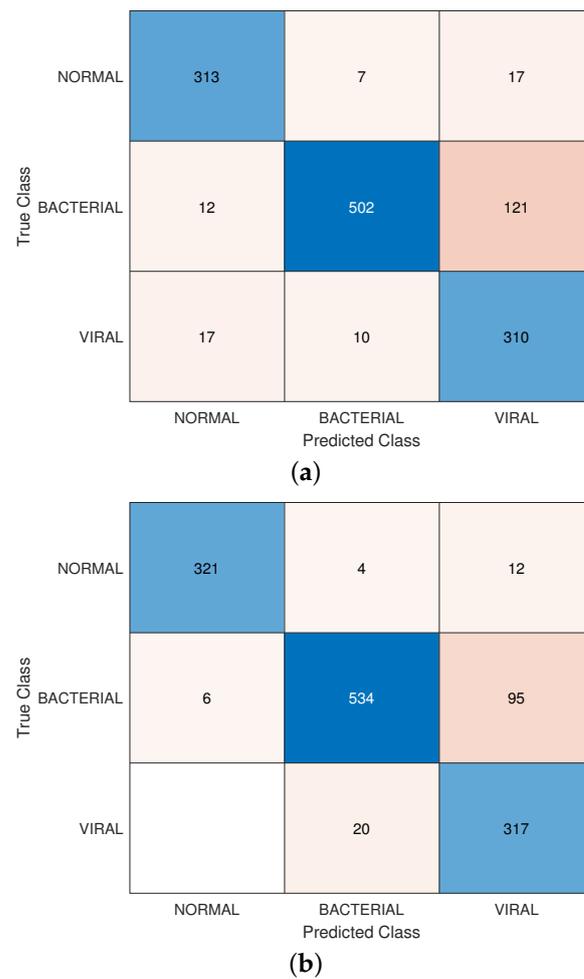


Figure 8. The confusion matrix of the proposed model in a normal vs. bacterial vs. viral CXR images classification fashion using (a) deep features and (b) proposed features.

Table 3. Comparison of early-layer and deep-layer accuracy for pneumonia diagnosis using most off-the-shelf deep networks vs. proposed densely-connected residual blocks.

Deep Network	Number of Layers	Accuracy (%)	
		Deep Layer Features	Early Layer Features
AlexNet [5]	8	96.0	97.3
VGG [17]	19	96.8	98.4
SqueezeNet [13]	14	96.7	96.0
GoogleNet [29]	27	96.2	97.7
ShuffleNet [45]	20	96.5	96.8
NASNetMobile [46]	913	95.8	96.9
DenseNet [18]	201	98.0	98.4
Xception [19]	36	96.4	98.4
ResNet [6]	50	97.6	97.1
Proposed method	35	na	99.6

Table 4. Comparison of proposed pneumonia diagnosis system with recent state-of-the-art methods.

Pneumonia Diagnosis Method	Deep Learning Technique	Accuracy (%)
Chowdhury et al. [4]	Transfer Learning with SqueezeNet	99.00
Asnaoui et al. [14]	Transfer Learning with ResNet50	96.61
Saraiva et al. [15]	CNN 10 Layers	95.30
Apostolopoulos et al. [16]	Transfer Learning with MobileNetv2	96.78
Liang and Zheng [21]	CNN with 49 Residual Blocks	95.30
Kermany et al. [22]	Transfer Learning with AlexNet	92.80
Toğaçar et al. [23]	Deep Features Fused from AlexNet, VGG16, and VGG19	99.41
Rajpurkar et al. [24]	Transfer Learning with ChexNet	82.83
Han et al. [26]	Contrastive Learning with ResNetAttention	88.00
Chouhan et al. [28]	Transfer Learning with 5 Deep Networks	96.40
Rahman et al. [30]	Transfer Learning with DenseNet201	98.00
Zhang et al. [31]	One-Class Classification Based Anomaly Detection	83.61
Ayan et al. [32]	Transfer Learning with Ensemble Voting	95.21
Nahiduzzaman et al. [33]	CNN with EML and PCA	99.83
Gour and Jain [34]	fine-tuned EfficientNet-B3	99.83
Proposed Method	Bottleneck Layer Features with 5 Densely-Connected Residual Building Blocks	99.60

3.4. Discussion

This work set out with the aim of assessing the importance of utilizing features extracted from modified deep network layers to detect viral and bacterial pneumonia in CXR lung images. A light-weighted bottleneck layer feature descriptors with an adaptive score fusion with learned weights are applied in two scenarios which are: Normal vs. infected lungs and normal vs. bacterial vs. viral infected lungs.

One interesting finding is only four iterations were required to optimize the hyper-parameters of the SVM classification utilizing Bayesian optimization and achieving a classification error lower than 0.1% as shown in Figure 4 compared to Random search optimization in terms of performance and processing time. Furthermore, the proposed weight learning method as depicted in Algorithm 1 has showed a significant benefit of score fusion of four descriptors as S_f achieved 99.6% of area under the curve in Figure 5 compared to scores acquired individually (i.e., $S_i(\mu)$, $S_i(\sigma)$, $S_i(\max)$, and $S_i(\min)$).

Another interesting finding for normal vs. infected lungs fashion including a mix of bacterial and viral infection CXR images scenario was that going deeper in some classification tasks could have a negative impact on accuracy. This is evident in Table 2 when the value of $l > 5$, the accuracy starts degrading as well as the complexity increases as more layers are used to extract features. Nevertheless, the proposed method has the upper hand compared to recent state-of-the-art methods shown in Tables 3 and 4.

Despite these promising results, questions remain. For instance, the accuracy of determining viral or bacterial pneumonia by either human experts or prediction models and how the process is prone to errors affecting patients' health. For instance, the observed results when viral vs. bacterial vs. normal lung CXR images are used confirm the challenging

task as the overall accuracy degraded to 89.5%. We recommend more future studies on the aforementioned issue.

4. Conclusions

We proposed an efficient method for pneumonia diagnosis using features extracted from bottleneck layers using modified densely-connected residual building blocks. Four types of features using four feature descriptors were extracted for each CXR image and the adaptive score fusion with Adaboost ensemble learning weight set-up was employed for selecting the best weight for each descriptor. The proposed method was evaluated using Kaggle's chest X-ray public dataset which contains more than 5000 images using two scenarios. The achieved accuracy compared to deep-layer features extracted from the most famous deep networks using state-of-the-art methods was promising to reach 99.6% and 90.2% for healthy vs. infected lungs classification scenario and bacterial vs. viral infected lungs classification scenario, respectively.

Furthermore, this work proposed a feasible pneumonia approach that can be utilized for COVID-19 chest X-ray images where the viral infection is present, we have not used images of real COVID-19 cases as to our knowledge, there is no sufficient public dataset for COVID-19 chest X-ray or CT images. Nevertheless, this case will be tested in our future work as soon as it becomes available.

Author Contributions: Conceptualization, S.A., M.A.M.A., B.A.J. and G.H.A.-M.; Data Curation, S.A.; Formal Analysis, S.A.; Methodology, S.A. and G.H.A.-M.; Project Administration, S.A., M.A.M.A., W.L.W.; Resources, B.A.J., B.W. and W.L.W.; Supervision, W.L.W.; Visualization, B.W.; Writing—Original Draft, S.A.; Writing—Review and Editing, M.A.M.A., B.W. and W.L.W. 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: This dataset is available online and can be used by anyone.

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

References

1. Viral vs. Bacterial Pneumonia: Understanding the Difference. 2020. Available online: https://www.pfizer.com/news/hot-topics/viral_vs_bacterial_pneumonia_understanding_the_difference (accessed on 28 April 2021).
2. Popovsky, E.Y.; Florin, T.A. Community-Acquired Pneumonia in Childhood. *Ref. Modul. Biomed. Sci.* **2020**. [CrossRef]
3. WHO Director Generals Opening Remarks at the Media Briefing on COVID-19 2020. 2020. Available online: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (accessed on 28 April 2021).
4. Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Mahbub, Z.B.; Islam, K.R.; Khan, M.S.; Iqbal, A.; Al-Emadi, N.; et al. Can AI help in screening viral and COVID-19 pneumonia? *arXiv* **2020**, arXiv:2003.13145.
5. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
6. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
7. Saikia, P.; Baruah, R.D.; Singh, S.K.; Chaudhuri, P.K. Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models. *Comput. Geosci.* **2020**, *135*, 104357. [CrossRef]
8. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [CrossRef]
9. Alkassar, S.; Jebur, B.A.; Abdullah, M.A.; Al-Khalidy, J.H.; Chambers, J. Going deeper: Magnification-invariant approach for breast cancer classification using histopathological images. *IET Comput. Vis.* **2021**, *15*, 151–164. [CrossRef]
10. Abdullah, M.A.; Alkassar, S.; Jebur, B.; Chambers, J. LBTS-Net: A fast and accurate CNN model for brain tumour segmentation. *Healthc. Technol. Lett.* **2021**, *8*, 31. [CrossRef]
11. Torrey, L.; Shavlik, J. Transfer learning. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*; IGI Global: Hershey, PA, USA, 2010; pp. 242–264.

12. Fooladgar, F.; Kasaei, S. Lightweight residual densely connected convolutional neural network. *Multimed. Tools Appl.* **2020**, *79*, 25571–25588. [[CrossRef](#)]
13. 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**, arXiv:1602.07360.
14. Asnaoui, K.E.; Chawki, Y.; Idri, A. Automated methods for detection and classification pneumonia based on X-ray images using deep learning. *arXiv* **2020**, arXiv:2003.14363.
15. Saraiva, A.; Ferreira, N.; Sousa, L.; Carvalho da Costa, N.; Sousa, J.; Santos, D.; Soares, S. Classification of Images of Childhood Pneumonia using Convolutional Neural Networks. In Proceedings of the 6th International Conference on Bioimaging, Prague, Czech Republic, 22–24 February 2019; pp. 112–119.
16. Apostolopoulos, I.D.; Mpesiana, T.A. COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [[CrossRef](#)]
17. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
18. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
19. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
20. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
21. Liang, G.; Zheng, L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Comput. Methods Programs Biomed.* **2019**, *187*, 104964. [[CrossRef](#)] [[PubMed](#)]
22. 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. [[CrossRef](#)]
23. Toğaçar, M.; Ergen, B.; Cömert, Z. A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models. *IRBM* **2019**, *41*, 212–222. [[CrossRef](#)]
24. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv* **2017**, arXiv:1711.05225.
25. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. Chestx-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2097–2106.
26. Han, Y.; Chen, C.; Tewfik, A.; Ding, Y.; Peng, Y. Pneumonia detection on chest X-ray using radiomic features and contrastive learning. In Proceedings of the 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13–16 April 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 247–251.
27. Shih, G.; Wu, C.C.; Halabi, S.S.; Kohli, M.D.; Prevedello, L.M.; Cook, T.S.; Sharma, A.; Amorosa, J.K.; Arteaga, V.; Galperin-Aizenberg, M.; et al. Augmenting the National Institutes of Health chest radiograph dataset with expert annotations of possible pneumonia. *Radiol. Artif. Intell.* **2019**, *1*, e180041. [[CrossRef](#)]
28. Chouhan, V.; Singh, S.K.; Khamparia, A.; Gupta, D.; Tiwari, P.; Moreira, C.; Damaševičius, R.; De Albuquerque, V.H.C. A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl. Sci.* **2020**, *10*, 559. [[CrossRef](#)]
29. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
30. Rahman, T.; Chowdhury, M.E.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Appl. Sci.* **2020**, *10*, 3233. [[CrossRef](#)]
31. Zhang, J.; Xie, Y.; Pang, G.; Liao, Z.; Verjans, J.; Li, W.; Sun, Z.; He, J.; Li, Y.; Shen, C.; et al. Viral pneumonia screening on chest x-rays using Confidence-Aware anomaly detection. *IEEE Trans. Med. Imaging* **2020**, *40*, 879–890. [[CrossRef](#)]
32. Ayan, E.; Karabulut, B.; Ünver, H.M. Diagnosis of Pediatric Pneumonia with Ensemble of Deep Convolutional Neural Networks in Chest X-ray Images. *Arab. J. Sci. Eng.* **2021**, 1–17. [[CrossRef](#)] [[PubMed](#)]
33. Nahiduzzaman, M.; Goni, M.O.F.; Anower, M.S.; Islam, M.R.; Ahsan, M.; Haider, J.; Gurusamy, S.; Hassan, R.; Islam, M.R. A Novel Method for Multivariant Pneumonia Classification based on Hybrid CNN-PCA Based Feature Extraction using Extreme Learning Machine with Chest X-ray Images. *IEEE Access* **2021**, *9*, 147512–147526. [[CrossRef](#)]
34. Gour, M.; Jain, S. Uncertainty-aware convolutional neural network for COVID-19 X-ray images classification. *Comput. Biol. Med.* **2021**, *140*, 105047. [[CrossRef](#)]
35. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**, *1*, 541–551. [[CrossRef](#)]
36. Chen, M.; Shi, X.; Zhang, Y.; Wu, D.; Guizani, M. Deep features learning for medical image analysis with convolutional autoencoder neural network. *IEEE Trans. Big Data* **2017**, *7*, 750–758. [[CrossRef](#)]
37. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
38. Czarnecki, W.M.; Podlowska, S.; Bojarski, A.J. Robust optimization of SVM hyperparameters in the classification of bioactive compounds. *J. Cheminform.* **2015**, *7*, 1–15. [[CrossRef](#)] [[PubMed](#)]

39. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. *arXiv* **2012**, arXiv:1206.2944.
40. Wenzel, F.; Deutsch, M.; Galy-Fajou, T.; Kloft, M. Scalable Approximate Inference for the Bayesian Nonlinear Support Vector Machine. In Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, 5–10 December 2016.
41. Wenzel, F.; Galy-Fajou, T.; Deutsch, M.; Kloft, M. Bayesian nonlinear support vector machines for big data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 307–322.
42. Freund, Y.; Schapire, R.; Abe, N. A short introduction to boosting. *J.-Jpn. Soc. Artif. Intell.* **1999**, *14*, 1612.
43. Kermany, D.; Zhang, K.; Goldbaum, M. Labeled optical coherence tomography (oct) and chest X-ray images for classification. *Mendeley Data* **2018**, *2*. [[CrossRef](#)]
44. Swingler, G.H. Radiologic differentiation between bacterial and viral lower respiratory infection in children: A systematic literature review. *Clin. Pediatr.* **2000**, *39*, 627–633. [[CrossRef](#)] [[PubMed](#)]
45. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.
46. Zoph, B.; Vasudevan, V.; Shlens, J.; Le, Q.V. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8697–8710.