Article

RED-CNN: The Multi-Classification Network for Pulmonary Diseases

San-Li Yi [1,2,*], Sheng-Lin Qin [1,2], Fu-Rong She [1] and Tian-Wei Wang [1,2,*]

1 School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China
2 Key Laboratory of Computer Technology Application of Yunnan Province, Kunming 650500, China
* Correspondence: 20120024@kust.edu.cn (S.-L.Y.); 20192204062@stu.kust.cn (T.-W.W.)

Abstract: Deep learning is a convenient method for doctors to classify pulmonary diseases such as COVID-19, viral pneumonia, bacterial pneumonia, and tuberculosis. However, such a task requires a dataset including samples of all these diseases and a more effective network to capture the features of images accurately. In this paper, we propose a five-classification pulmonary disease model, including the pre-processing of input data, feature extraction, and classifier. The main points of this model are as follows. Firstly, we present a new network named RED-CNN which is based on CNN architecture and constructed using the RED block. The RED block is composed of the Res2Net module, ECA module, and Double BlazeBlock module, which are capable of extracting more detailed information, providing cross-channel information, and enhancing the extraction of global information with strong feature extraction capability. Secondly, by merging two selected datasets, the Curated Chest X-Ray Image Dataset for COVID-19 and the tuberculosis (TB) chest X-ray database, we constructed a new dataset including five types of data: normal, COVID-19, viral pneumonia, bacterial pneumonia, and tuberculosis. In order to assess the efficiency of the proposed five-classification model, a series of experiments based on the new dataset were carried out and based on 5-fold cross validation, and the results of the accuracy, precision, recall, F1 value, and Jaccard scores of the proposed method were 91.796%, 92.062%, 91.796%, 91.892%, and 86.176%, respectively. Our proposed algorithm performs better than other classification algorithms.

Keywords: pulmonary diseases; COVID-19; Res2Net module; ECA module; Double BlazeBlock module

1. Introduction

Pulmonary diseases such as COVID-19, tuberculosis, bacterial pneumonia, and viral pneumonia are contagious and fatal and therefore greatly threaten people's health. Early detection of these diseases will help doctors to treat them in a timely manner. Therefore, it is quite important to diagnose them effectively and accurately.

Novel coronavirus (COVID-19) is a highly contagious disease which can be detected with nucleic acid testing, chest X-ray examination, etc. Tuberculosis, also a highly contagious disease developed from Mycobacterium tuberculosis infection, can be tested with the sputum smear test, tuberculin test, and chest X-ray examination. Bacterial pneumonia, including Streptococcus pneumonia and Staphylococcus aureus, causes great harm to the health of the elderly and children and can be detected via bacteriological examination, blood examination, and chest X-ray examination. Viral pneumonia, caused by the viral infection of respiratory tract that spreads downward, can be detected with pathogenic examination, serological examination, and chest X-ray examination. Among the various detection methods mentioned above, chest X-ray examination can be used to obtain patients' image data in real time, while the others usually have a longer duration. However, as the diagnosis of the X-ray images depends on the radiologist, it is usually subject to artificial influenced.

Computer-aided diagnosis (CAD) technology, which is immune to artificial influence, has drawn people's attention for its automatically diagnosis and has been widely used in medical diagnosis with the development of deep learning technology, especially convolutional neural networks (CNNs). CNNs, as the feature extractors playing a key role of diagnosis models, have undergone rapid development recently. In 2012, AlexNet [1] was proposed, and many advanced algorithms were subsequently proposed: VGG [2], residual [3], Inceptionv3 [4], NASNetLarge [5], and MobileNetV2 [6] and others [7,8]. In addition to above networks, the following new technologies are of interest: the multibranch Res2Net module proposed in [9], which effectively solved the problem of low utilization of feature information in the process of obtaining image feature information; the ECA module proposed by Wang et al. [10], which can filter the feature information on the channel to strengthen the utilization of useful feature information and greatly improve the performance of the network; and the Double BlazeBlock module proposed by Bazarevsky Valentin et al. using the jump connection structure to extract more image feature information and reduce the number of network parameters through the replacement of ordinary convolution with deep convolution [11].

Deep learning technology has been widely used in the field of medical imaging. Many classification studies for pulmonary diseases have been proposed, which can be categorized as two types, two-classification and multi-classification diagnosis. For the two-classification studies, Singh et al. used anti-aliasing convolutional networks to detect tuberculosis [12]. Abdulfattah et al. used the CNN model for screening tuberculosis disease [13]. Later, more two-classification models on tuberculosis were carried out [14–16]. For COVID-19 and non-COVID-19 detection, studies based on two-classification have also been proposed, such as [17,18]. With the improvement of image quality and number of images in the public pulmonary datasets, multi-classification of pulmonary diseases has attracted more attention. Verma et al. proposed a new neural-network-based classifier to classify tuberculosis, bacterial pneumonia, and viral pneumonia in chest radiographs [19]. Ibrahim et al. used deep learning models to identify COVID-19, pneumonia, and lung cancer [20]. For the classification of normal images, COVID-19, and pneumonia, there are many studies, for example: Wang J et al. proposed a novel task adaptation network (TAN) [21]. Abdrakhmanov et al. used few-shot learning [22]. Shome et al. used Vision Transformer [23] and other methods [24,25]. Altan et al. put forward a hybrid model based on two-dimensional curvelet transform, the chaotic algae algorithm, and deep learning technology to identify COVID-19 disease, viral pneumonia, and normal images from X-ray images [26]. Jin et al. launched a three-step hybrid ensemble model composed of a feature extractor, feature selector, and classifier to identify COVID-19, viral pneumonia, and normal images [27]. Furthermore, many studies have classified COVID-19 and pneumonia with methods based on CNNs [28–31].

The abovementioned pulmonary disease diagnosis technologies based on deep learning networks can successfully alleviate the error rate and time- and labor-intensive nature of manual identification. Nevertheless, two problems remain unsolved: (1) As there are many kinds of pulmonary diseases, algorithms which can identify more pulmonary diseases are especially practical. However, there is currently no public dataset for all of these diseases mentioned above, and each of these methods can only identify two or three kinds of pulmonary diseases. (2) The extraction of image features with neural network algorithms is always affected by the relatively poor image quality of chest X-rays as well as the susceptibility to interference from clothing and superimposed tissues. Therefore, improving the feature extraction ability of the deep learning network for the recognition of pulmonary diseases based on X-ray images is an issue of great concern.

In this paper, we propose a multi-classification pulmonary disease model with our newly constructed network RED-CNN as the feature extraction backbone. The scope of this study is as follows:

(1)　The diagnosis model which can classify COVID-19, viral pneumonia, bacterial pneumonia, tuberculosis, and normal images, involves pre-processing of input data, feature extraction, and a classifier.

(2)　To extract the features of X-ray images more accurately, a new network, RED-CNN, was created. This network is based on the architecture of a CNN, and its basic function block, the RED block was constructed through a combination of the Res2Net module, ECA module, and Double BlazeBlock module, which are capable of extracting detailed information, providing cross-channel information, and enhancing the extraction of global information.

(3)　As there is no public dataset for the five pulmonary diseases mentioned above, we constructed a new dataset based on the combination of two public datasets: the Curated Chest X-Ray Image Dataset for COVID-19 (curated COVID-19) [32], and the tuberculosis (TB) chest X-ray database [33]. The new dataset contains five types of data, namely COVID-19, virus pneumonia, bacterial pneumonia, tuberculosis, and normal images.

(4)　To assess the efficiency of the RED-CNN network, experiments on the new constructed dataset were performed, and the results of our network were compared with those of the state-of-the-art networks such as VGG19, Inceptionv3, Resnet50, NASNetLarge, and MobileNetV2. To further determine the efficiency of the proposed diagnosis model, we also compared it with methods from the literature mentioned above.

The structure of this article is as follows: The first section introduces the developmental background of pulmonary disease diagnosis, the development of deep learning, and its application in pulmonary disease diagnosis. The second section summarizes the materials and methods, where the multi-classification pulmonary disease model is outlined, and our constructed RED-CNN network and the new constructed dataset including five types of data is introduced. The third section details the classification experiments of our network, which verified its effectiveness by comparing and analyzing the results with the state-of-the-art networks and existing methods in the literature. The fourth section is the conclusion, which summarizes the work of this study and the prospects of the network.

2. Materials and Methods

The multi-classification pulmonary disease model of this work was mainly constructed in 3 steps: the pre-processing of input data, the setup of the RED-CNN network, and the creation of the classifier. The RED-CNN network is the backbone of feature extraction, and its output is the classifier, which includes five categories: normal images, COVID-19, tuberculosis, viral pneumonia, and bacterial pneumonia. The pipeline of the processing steps is shown in Figure 1.
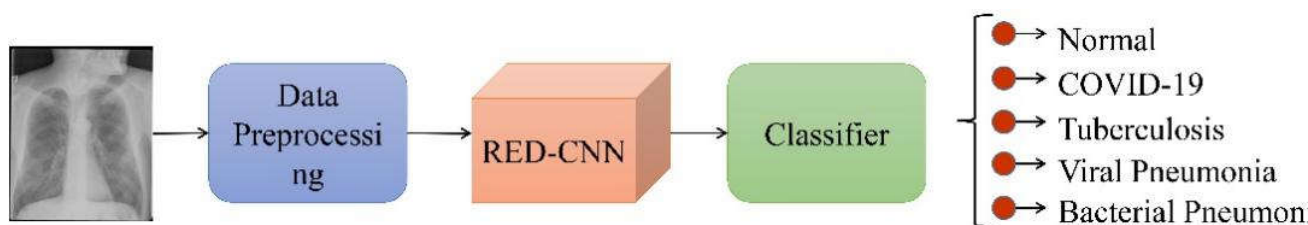
**Figure 1.** The multi-classification pulmonary disease model.

2.1. Dataset

Currently, the widely used datasets of pulmonary diseases are the dataset of chest X-ray images (pneumonia) [34], COVID-19 chest X-ray database [35], selected chest X-ray image dataset for COVID-19 [36], and tuberculosis (TB) chest X-ray database [32]. Some of these datasets are used for the diagnosis of viral pneumonia, bacterial pneumonia, and COVID-19 and some for the diagnosis of tuberculosis. As none of them contains all the four pulmonary diseases described in Section 1, a more powerful dataset containing the data

of five categories (including data of normal images) is needed. To solve this problem, we combined two datasets to construct a new dataset containing patients of four pulmonary diseases and unaffected people. The construction process is detailed as follows:

Firstly, two suitable datasets were selected: (1) The dataset of curated COVID-19obtained from the Kaggle website in 2020, containing 1281 COVID-19 X-rays, 3270 normal X-rays, 1656 viral pneumonia X-rays, and 3001 bacterial pneumonia X-rays; (2) The tuberculosis X-ray chest radiograph database, containing 3500 TB X images and 3500 normal X images, which is also from the 2020 Kaggle collection dataset and was created by a team of researchers from Qatar University in Doha, Qatar, and Dhaka University in Bangladesh, as well as collaborators from Malaysia and doctors from Hamad Medical Corporation and Bangladesh.

Secondly, before the combination of the two datasets, the redundant images for classification were filtered. For this reason, we eliminated the normal X-ray images in the second dataset while retaining those in the first dataset. The remaining images of these two datasets were integrated into a new dataset with five types of lung image features. The images in the new combined dataset were imaged in the same way and were all X-ray images. Various sample images in the new combined dataset are shown in Figure 2. The data distribution of each class is shown in Figure 3.

(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

(d)　　　　　　　　　　　(e)
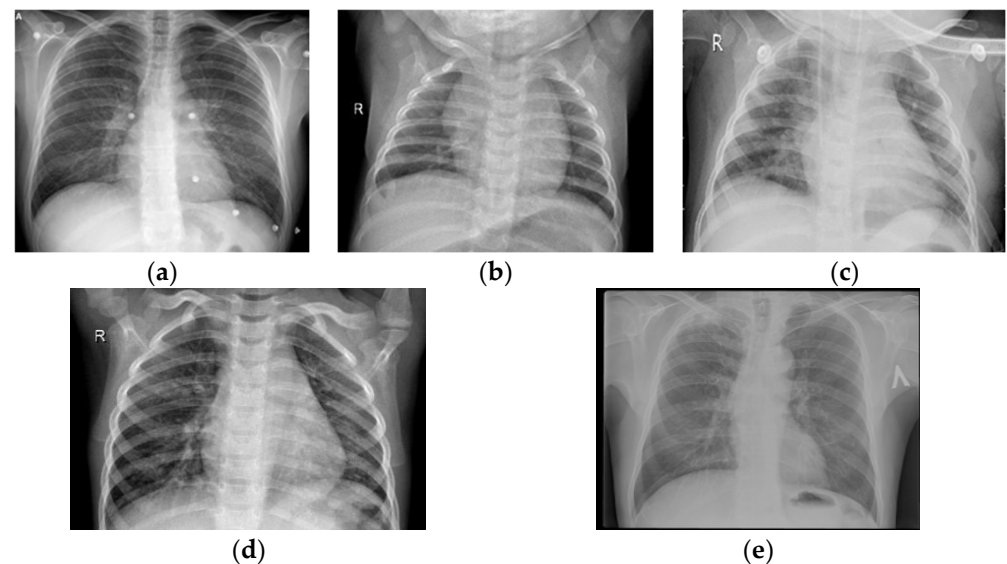
Figure 2. Sample display of each class: (**a**) COVID-19; (**b**) normal; (**c**) bacterial pneumonia; (**d**) viral pneumonia; (**e**) tuberculosis.
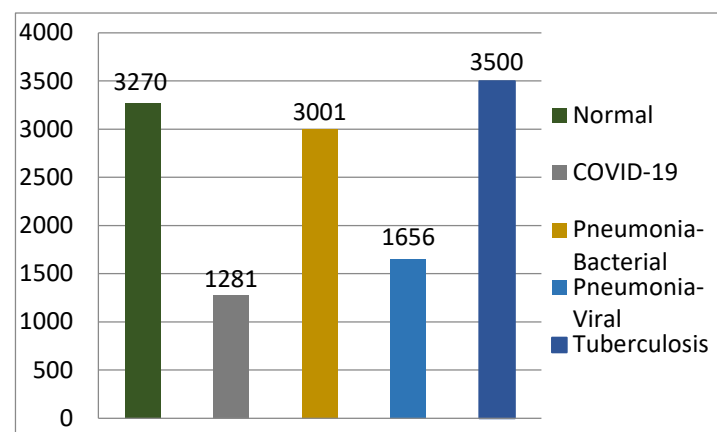

Figure 3. Sample distribution of class data.

2.2. Preprocessing

As different neural networks have different sizes of input images, it is necessary to preprocess the input images to satisfy the input requirements of the network. In this study, the image size was set to 224 × 224 pixels, and all images were normalized and scaled to [0,1].

2.3. RED-CNN

To accurately capture image features, we proposed a new network, RED-CNN, which is based on the CNN architecture and our constructed basic function block, the RED block This block was set up through the combination of 3 function modules: the Res2Net module, ECA module, and Double BlazeBlock module. The architecture of the network is shown in Figure 4a, and RED blocks are shown in Figure 4b. To further present its structure, the details of each layer in the network are listed in Table 1.
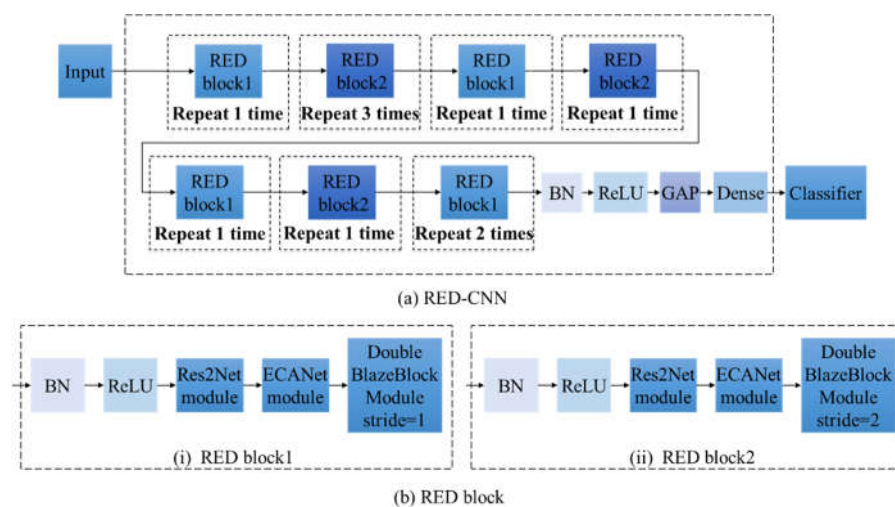
Figure 4. RED-CNN network structure.

Table 1. The layers and layer parameters of the proposed model (for the five-classification task).

Layer Number	Layer Type	Output Shape	Trainable Parameters
1	RED block1	[224, 224, 11]	795
2	RED block2 × 3 *	[28, 28, 88]	26,443
3	RED block1	[28, 28, 176]	36,386
4	RED block2	[14, 14, 352]	254,032
5	RED block1	[14, 14, 704]	487,892
6	RED block2	[7, 7, 1408]	3,873,340
7	RED block1 × 2 *	[7, 7, 5632]	36,922,552
8	BN	[7, 7, 5632]	22,528
9	Activation	[7, 7, 5632]	-
10	GAP	[None, 5632]	-
12	Dense	[None, 5]	28,165

* block × n, where n represents the number of repetitions.

From Figure 4a and Table 1, we can see that the operation of the network is as follows: The input image sequentially goes once through RED block1, 3 times through RED block2, once through RED block1, once through RED block2, once through RED block1, once through RED block2, and twice through RED block1, BN, ReLU, GAP (GlobalAveragePooling2D), and Dense. The last part is the output, which is the result of five categories.

Figure 4b is the RED block We can see that it has two types of structure: RED block1 and RED block2. The these two blocks are composed of the BN, ReLU, Res2Net module, ECA module, and Double BlazeBlock module. The difference between them is the value of

the Double BlazeBlock module's stride, which is 1 for RED block1 and 2 for RED block2. A detailed explanation of the stride is given in Section 2.3.3. Among these compositions, the combination of the Res2Net module, ECA module, and Double BlazeBlock module is the main feature extract section of the network. Such a combination can obtain multi-scale extraction of feature information, filter feature information on the channel, hold useful feature information, and strengthen its use when the network is iterating. BN is the batch normalization layer, which can effectively prevent the network model from overfitting and speed up the model's convergence. ReLU is the activation function layer. The function of each module is detailed as follows.

2.3.1. Res2Net Module

The Res2Net [9] module structure can expand the receptive field of each layer of the network and enable the network to extract feature information at multiple scales. Therefore, we used it as a part of our proposed RED-CNN network. The structure of the Res2Net module in our proposed network is shown in Figure 5:
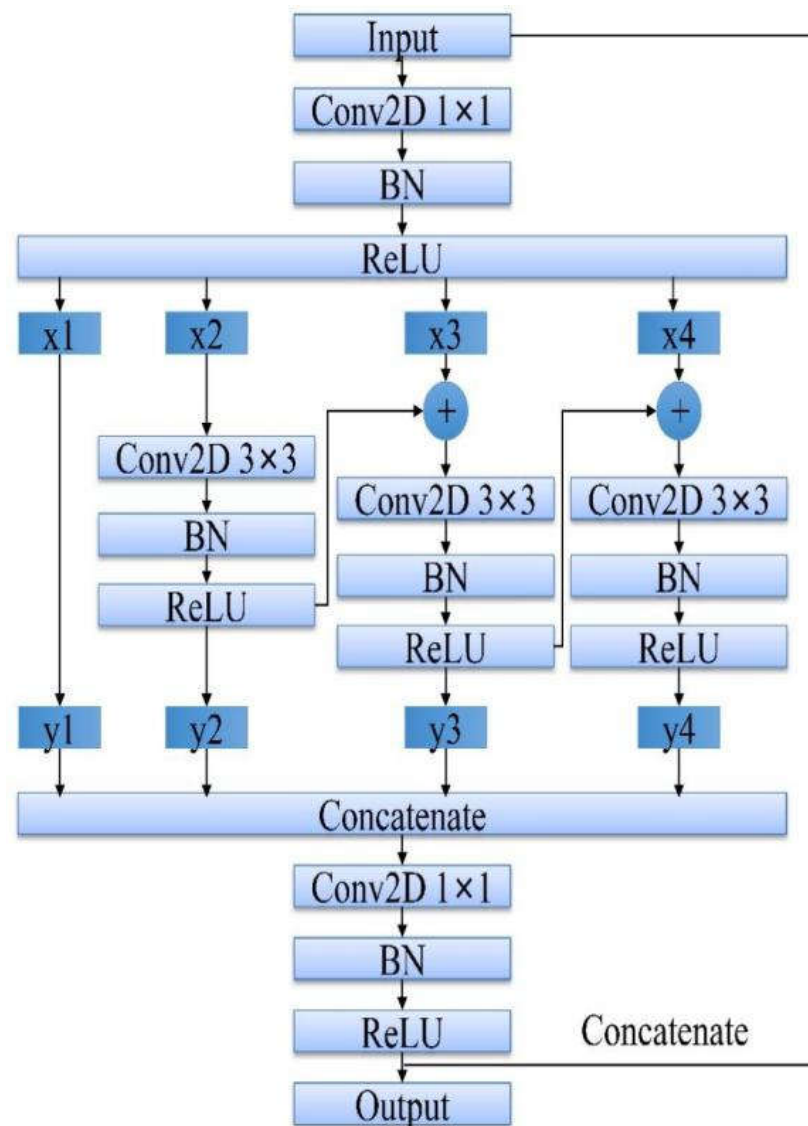
Figure 5. Structure of Res2Net module.

In Figure 5, the input data passes through the Conv2D 1×1, BN, and ReLU layers in sequence, and the output of the ReLU layer is divided into s subsets, which are defined as $x_i, i \in \{1, 2, \ldots, s\}$. Here, we set s as 4, and each subset x_i is of the same size, i.e., the

channel number is 1/s of the input. From Figure 5, we can see that x_1 is the input of the first subset, which is directly output as y_1; x_2 is the input of the second subset, which is transmitted through the Conv2D 3 × 3, BN, and ReLU layers to obtain the output y_2; the input of the third subset is the sum of x_3 and y_2, which then goes through the Conv2D 3 × 3, BN, and ReLU layers to obtain the output y_3; the input of the fourth subset is the sum of x_4 and y_3, which then passes through the Conv2D 3 × 3, BN, and ReLU layers to obtain the output y_4. Finally, the feature fusion operation is performed, i.e., the output of the four subsets: y_1, y_2, y_3, and y_4, are concatenated. The output of the concatenated layer is subsequently input into the Conv2D 1 × 1, BN, and ReLU layer for further feature extraction and is then concatenated with the original input of the module to strengthen the use of feature information for the final output.

2.3.2. ECA Module

The ECA [10] module is a channel attention module that can realize cross-channel information exchange and efficiently extract image features on the channel, thus improving the performance of the network. Therefore, the ECA module was adopted in our RED-CNN. The detailed design is shown in Figure 6:
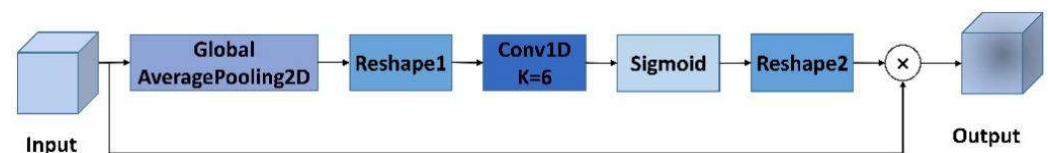

Figure 6. Structure of ECA module.

In Figure 6, the input goes through the layers of GlobalAveragePooling2D, Reshape1, Conv1D (k = 6), sigmoid, and Reshape2 in sequence. Then, the output of Reshape2 layer is multiplied with the input to realize the feature information extraction. GlobalAverage-Pooling2D, a global average pooling layer, is used to perform channel-by-channel global average pooling. The Reshape layers are used to reshape the size of the feature map. The Conv1D layer captures the cross-channel information interaction, and its parameter k is the number of adjacent channels, which is set to 6, indicating that 6 adjacent channels participate in a channel attention prediction. The sigmoid layer is the activation layer with a sigmoid activation function.

2.3.3. Double BlazeBlock Module

The Double BlazeBlock [11] module structure can expand the receptive field of the image and extract more feature information with only a small number of parameters being increased. Therefore, it was used as the functional module of this network. The design is shown in Figure 7:

Figure 7 shows that the Double BlazeBlock module has two types of block structure with stride = 1 and 2. The stride is step size. When stride = 1, the input data sequentially goes through DepthwiseConv2D 5 × 5, Conv2D 1 × 1, ReLU, DepthwiseConv2D 5 × 5, and Conv2D 1 × 1 layers, and then the output of the Conv2D 1 × 1 layer is combined with the original input data. Such an operation is called as information fusion. After that, the output of the information fusion step is fed into the ReLU layer and then output. When the stride is 2, the block structure is almost the same as that of stride = 1; the difference is that instead of directly being fused, before the fusion step, two new operations are added to the input data: MaxPool2D 2 × 2 and Conv2D 1 × 1. The two added operations can help better extract feature information and reduce the image resolution. Among all the operations in Figure 7, DepthwiseConv2D is a depthwise separable convolution, which can not only extract image features but also effectively reduce the network parameters. Conv2D 1 × 1 is a general convolution for feature extraction. MaxPool2D 2 × 2 is the maximum pooling layer, which reduces the image resolution during the sampling process.

The Double BlazeBlock module can not only enhance the use of feature information but also effectively reduce network parameters.
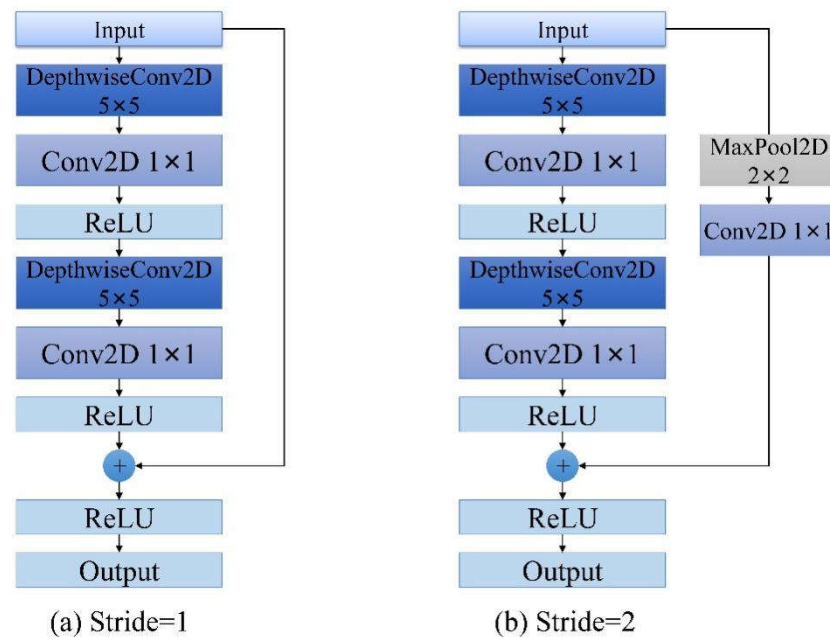
Figure 7. Structure of Double BlazeBlock module.

3. Experiment

Based on the new constructed dataset, two experiments were carried out: In the first experiment, the RED-CNN network was compared with some state-of-the-art networks, such as VGG19 [2], ResNet50 [3], Inceptionv3 [4], NASNetLarge [5], MobileNetV2 [6]. For experiment two, methods from recent studies were compared with the proposed diagnosis model to assess its performance. The results of two experiments are analyzed and discussed. In this section, the experimental settings, evaluation indicators, selection of the key hyperparameter k, experimental results, and analysis are detailed as follows.

3.1. Experimental Setup

The experiments were conducted on a computer equipped with an Intel Core™ i9-10900k CPU@3.70 GHz, 32.0 GB of RAM, an NVIDIA GeForce RTX3080 graphics card, and Windows 10 64-bit operating system. The integrated development environment was PyCharm, the programming language was Python, and the experimental frameworks were Keras and Tensorflow.

In the experiments, the network's optimizer was set to Adam to update the parameters, the loss function was set to sparse_categorical_crossentropy, the learning rate was set to 0.001, the number of epochs was set to 60, and the batch size was set to 8. Based on our newly constructed dataset, the proportion of the training set and test set was set to 8:2, i.e., 10166 images were used in the training set and 2542 images in the test set. The division of the dataset is shown in Table 2.

Table 2. Division of dataset.

Dataset Format	COVID-19	Normal	Bacterial Pneumonia	Viral Pneumonia	Tuberculosis	Total
Training set	1046	2591	2400	1314	2815	10,166
Test set	235	679	601	342	685	2542
Total	1281	3270	3001	1656	3500	12,708

The number of images in the five categories of data in our new constructed dataset was not equal because they have more training weights than other categories [37], which leads to the tendency of the classification model to enhance the classification of the larger categories. To avoid this problem, we used weighted averaging to present the experiment results.

3.2. Selection of Hyperparameter k

The hyperparameter k is the number of adjacent channels, reflecting the channel's correlation, which can be used to improve the module's ability to extract feature information. The correlation information is related to the adjacent distance k of the channel, and different values of k reflect different ranges of correlation, which can affect the performance of the network. Therefore, experiments that involved selecting the k value were conducted to obtain optimized performance. The results are shown in Table 3:

Table 3. Selection of parameter k.

Evaluation Index	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9
Accuracy	92.21	92.72	92.17	92.91	92.68	92.68	92.05
Precision	92.28	92.82	92.16	93.00	92.74	92.77	92.09
Recall	92.21	92.72	92.17	92.91	92.68	92.68	92.05
F1	92.24	92.76	92.16	92.95	92.70	92.72	92.07
Jaccard	86.76	87.50	86.58	87.63	87.30	87.32	86.40

As is shown in Table 3, when k = 6, the accuracy, precision, recall, F1, and Jaccard values of the model are almost optimal. Therefore, the value of k in this experiment was selected as 6.

3.3. Experimental Results and Analysis

3.3.1. Experiment 1

The experiment was performed on the constructed dataset, and the results for the networks mentioned above were obtained. To ensure the stability of the results, a 5-fold cross validation approach was used, and average values were adopted as the final values of each network. Table 4 lists the comparison of the average results of different networks based on 5-fold cross validation. Figure 8 shows the performance metrics of different networks, presenting a comparison of different networks in realizing classification of five types of pulmonary images.

Table 4. Comparison of weighted average results of different networks.

Method	Accuracy	Precision	Recall	F1	Jaccard
VGG19	88.886	91.304	88.886	89.374	82.508
ResNet50	90.036	90.556	90.036	90.148	83.524
NASNetLarge	90.348	90.484	90.348	90.226	84.274
Inceptionv3	90.272	89.814	90.272	90.608	84.514
MobileNetV2	89.32	89.772	89.32	89.10	82.052
RED-CNN	91.796	92.062	91.796	91.892	86.176

It can be seen from Table 4 and Figure 8 that the average evaluation values of the RED-CNN network for accuracy, precision, recall, F1, and Jaccard values are 91.796%, 92.062%, 91.796%, 91.892%, and 86.176%, respectively, better than other networks. This means that our new network can identify pulmonary diseases more efficiently than the other networks.
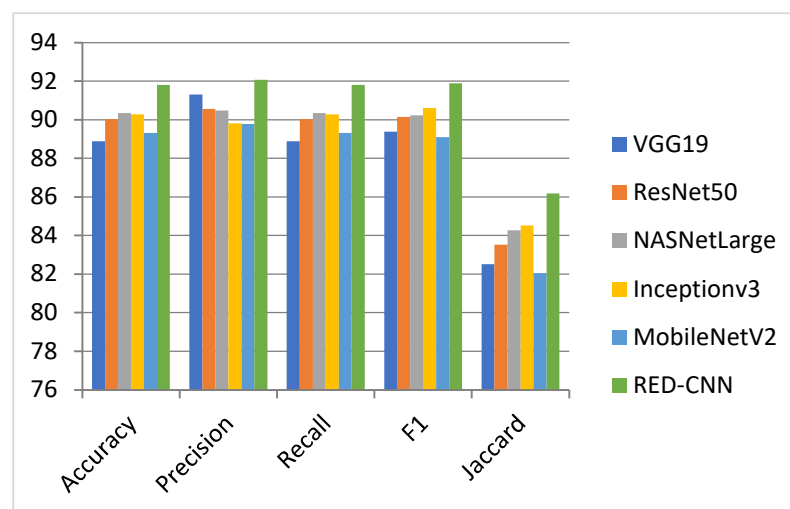
Figure 8. The evaluation index for different networks of different types.

A confusion matrix is a visualization tool that can display the network performance in diagnosing each category intuitively. The main diagonal value in the confusion matrix indicates the number of correct classifications. The larger the total value of the main diagonal, the better the classification model performs. In Figure 9, the overlapped confusion matrix computed from each fold visually demonstrates each network's capability of classifying lung images.

The subgraphs in Figure 9 are the overlapped confusion matrices of different networks. From Figure 9 we can see that among the total number of 2542 tested images including five categories (12710 for a 5-fold overlap), the number of images correctly identified by RED-CNN is 11668, higher than 11298 of VGG19, 11444 of ResNet50, 11484 of NASNetLarge, 11474 of Inceptionv3, and 11353 of MobileNetV2. This means that the RED-CNN network has better performance in the classification of pulmonary diseases than the other networks.

The value of AUC for different classes of images reflects the performance of the classifier. It refers to the area under the ROC curve, and the larger the value, the better the network performs. In Figure 10, the average of the ROC curve presents the classification performance of our proposed network in each class of the pulmonary image.

As is shown in Figure 10, based on the classification of the RED-CNN network, the average AUC values of the five categories, i.e., COVID-19, normal, bacterial pneumonia, viral pneumonia, and tuberculosis, are 0.9791, 0.9864, 0.9116, 0.8301, and 0.9929, respectively. This demonstrates that the RED-CNN network has excellent classification performance for each category.

3.3.2. Experiment 2

In this experiment, the classification efficiency of the proposed network was compared with the existing methods in the literature. As far as we know, there are no five-classification studies on pulmonary diseases. However, in recent years, based on the tuberculosis (TB) chest X-ray database, some two-classification studies have been proposed, and based on the curated COVID-19 dataset, three-classification studies have recently been proposed. In order to compare with these studies, three kinds of comparisons were carried out: Firstly, we performed the two-classification experiment by using RED-CNN on the tuberculosis (TB) chest X-ray database, and the results of comparison are shown in Table 5. Secondly, the three-classification experiment was performed on the curated COVID-19 dataset using the RED-CNN, and the results are shown in Table 6. Finally, to further assess the efficiency of the proposed network, the results of its three-classification were also compared with the studies based on different datasets, which can be seen in Table 7.
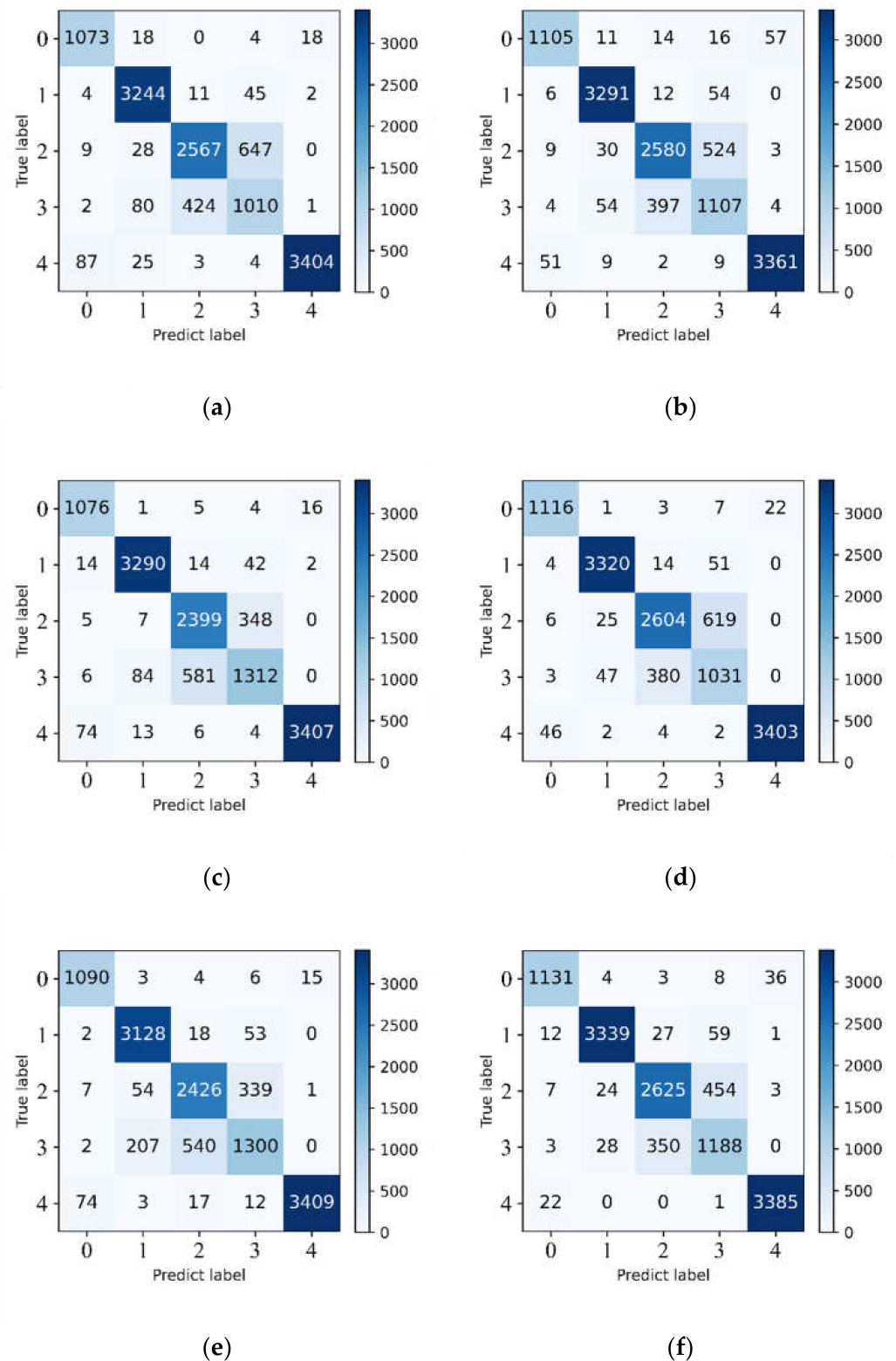
Figure 9. Overlapped confusion matrix of different algorithms: (**a**) VGG19; (**b**) Resnet50; (**c**) NAS-NetLarge; (**d**) Inceptionv3; (**e**) MobileNetV2; (**f**) RED-CNN. In the graph: 0 represents COVID-19, 1 represents normal, 2 represents bacterial pneumonia, 3 represents viral pneumonia, and 4 represents pulmonary tuberculosis.
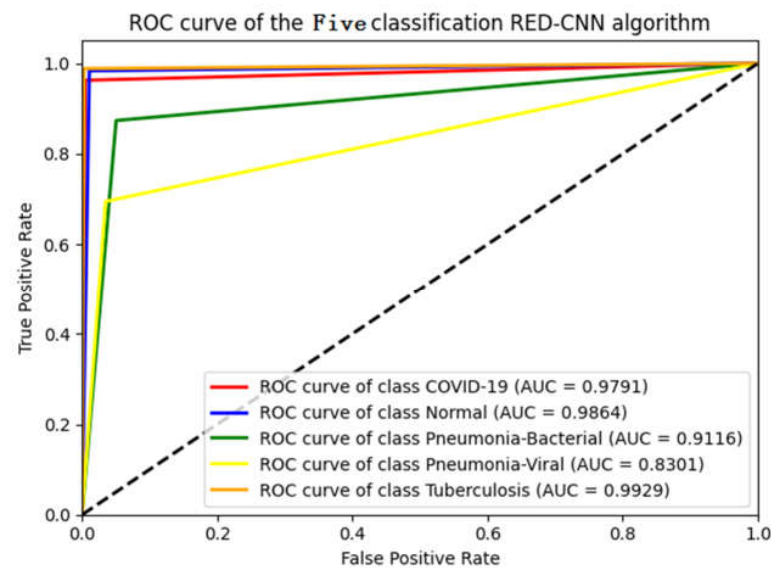
Figure 10. Five-category average value of AUC curve.

Table 5. Comparison of RED-CNN with other studies based on tuberculosis (TB) chest X-ray database.

Study	Accuracy	Precision	Recall	F1
[13]	98.81	-	98.86	-
[14]	99.36	98.00	99.00	99.00
[15]	99.92	99.85	1.00	99.92
[16]	99.55	99.55	99.54	-
RED-CNN	99.92	99.92	99.92	99.92

Table 6. Comparison of RED-CNN with other studies based on curated COVID-19 dataset.

Study	Accuracy	Precision	Recall	F1
[26]	97.99	98.38	98.15	98.26
[21]	-	92.0	99.1	95.4
[22]	97.70	-	-	-
[23]	92.00	93.00	89.00	91.00
[27]	96.61	-	-	-
RED-CNN	98.47	98.48	98.47	98.47

Table 7. Comparison of RED-CNN with other studies based on different datasets.

Study	Accuracy	Precision	Recall	F1
[28]	87.02	89.96	85.35	87.37
[29]	92.48	89.02	88.93	88.76
[30]	94.20	94.04	93.10	91.32
[31]	97.05	94.44	96.96	95.38
RED-CNN	98.47	98.48	98.47	98.47

As can be seen from Table 5, the accuracy, precision, recall, and F1 of RED-CNN reached 99.92%, indicating superiority to all methods in the literature for all values except recall. From Table 6, we can see that the accuracy, precision, recall, and F1 of RED-CNN, reached 98.47%, 98.48%, 98.47%, and 98.47%, respectively. Compared with the methods in the literature, the performance of our network is better. From Table 7 we can see that in the three-classification task, RED-CNN has better performance than other existing studies. Our network achieved the highest accuracy of 98.47%, precision of 98.48%, recall of 98.47%, and F1 of 98.47%.

4. Discussion and Conclusions

In this paper, we focus on the effective classification of five pulmonary diseases, namely COVID-19, viral pneumonia, bacterial pneumonia, and tuberculosis. To this end, we propose a pulmonary disease diagnostic model including the pre-processing of input data, the RED-CNN network, and a classifier. In the model: (1) To extract the features of X-ray images more accurately, we propose a new network, RED-CNN, which is based on the architecture of a CNN and was constructed with three function modules: the Res2Net module, ECA module, and Double BlazeBlock module. The Res2Net module makes the network wider, enabling it to extract more detailed information. The ECA module can improve the network's performance through providing cross-channel information. The Double BlazeBlock module can strengthen the extraction of global information by expanding the image's receptive field. (2) We also constructed a new dataset based on the combination of two public datasets, i.e., the dataset of curated COVID-19 and the tuberculosis (TB) chest X-ray database. Experiments were conducted based on the constructed dataset. Compared with the state-of-the-art networks mentioned in these experiments, our proposed model obtained the best results in the case of 5-fold cross-validation averaging with accuracy, precision, recall, F1, and Jaccard scores of 91.796%, 92.062%, 91.796%, 91.892%, and 86.176%, respectively, indicating the effectiveness of the proposed model. Furthermore, our model was compared with existing methods in the literature. In the two-classification experiment involving the tuberculosis (TB) chest X-ray database, the RED-CNN network achieved the best results, with accuracy, precision, recall, and F1 value being 99.92%. In the three-classification experiment involving the curated COVID-19 dataset, the performance of our network is better than other existing studies, with accuracy of 98.47%, precision of 98.48%, recall of 98.47%, and F1 value of 98.47%. In addition, three-classification studies based on different datasets were compared with our model, indicating that our model has better results.

The superiority of our model can be outlined as follows: (1) The model realized a five-classification method of pulmonary diseases, proving its feasibility for the detection of COVID-19, viral pneumonia, bacterial pneumonia, tuberculosis, and normal images. (2) As the feature extractor of the model is designed for chest X-ray images, the new network RED-CNN can effectively realize a five-classification method of pulmonary diseases with more accuracy and speed than the state-of-the-art networks. Furthermore, the proposed network is more efficient than those in the literature mentioned above. (3) The newly constructed dataset in this study can provide five types of pulmonary disease data for the classification model, while no public dataset available includes all these types of data.

In future work, the aim will be to enhance the performance of the proposed network by reducing the parameters of the network, increasing the number of categories of pulmonary disease data of the dataset, and using other networks, such as EfficientNet, in the feature extraction backbone.

References

1. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Processing Syst.* **2012**, *25*, 84–90. [CrossRef]
2. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
3. 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.
4. 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.
5. 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, 18–23 June 2018; pp. 8697–8710.
6. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
7. 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.
8. 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.
9. Gao, S.H.; Cheng, M.M.; Zhao, K.; Zhang, X.Y.; Yang, M.H.; Torr, P. Res2net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 652–662. [CrossRef] [PubMed]
10. Wang, Q.L.; Wu, B.G.; Zhu, P.F.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 11531–11539.
11. Bazarevsky, V.; Kartynnik, Y.; Vakunov, A.; Raveendran, K.; Grundmann, M. Blazeface: Sub-millisecond neural face detection on mobile gpus. *arXiv* **2019**, arXiv:1907.05047.
12. Singh, J.; Tripathy, A.; Garg, P.; Kumar, A. Lung tuberculosis detection using anti-aliased convolutional networks. *Procedia Comput. Sci.* **2020**, *173*, 281–290. [CrossRef]
13. Alawi, A.E.B.; Al-basser, A.; Sallam, A.; Al-sabaeei, A.; Al-khateeb, H. Convolutional Neural Networks Model for Screening Tuberculosis Disease. In Proceedings of the 2021 International Conference of Technology, Science and Administration (ICTSA), Taiz, Yemen, 22–24 March 2021; pp. 1–5. [CrossRef]
14. Faruk, O.; Ahmed, E.; Ahmed, S.; Tabassum, A.; Tazin, T.; Bourouis, S.; Monirujjaman Khan, M. A Novel and Robust Approach to Detect Tuberculosis Using Transfer Learning. *J. Healthc. Eng.* **2021**, *2021*, 1002799. [CrossRef]
15. Rahman, M.; Cao, Y.; Sun, X.; Li, B.; Hao, Y. Deep pre-trained networks as a feature extractor with XGBoost to detect tuberculosis from chest X-ray. *Comput. Electr. Eng.* **2021**, *93*, 107252. [CrossRef]
16. Dey, S.; Roychoudhury, R.; Malakar, S.; Sarkar, R. An optimized fuzzy ensemble of convolutional neural networks for detecting tuberculosis from Chest X-ray images. *Appl. Soft Comput.* **2022**, *114*, 108094. [CrossRef]
17. Panahi, A.H.; Rafiei, A.; Rezaee, A. FCOD: Fast COVID-19 Detector based on deep learning techniques. *Inform. Med. Unlocked* **2021**, *22*, 100506. [CrossRef]
18. Turkoglu, M. COVID-19 detection system using chest CT images and multiple kernels-extreme learning machine based on deep neural network. *IRBM* **2021**, *42*, 207–214. [CrossRef]
19. Verma, D.; Bose, C.; Tufchi, N.; Pant, K.; Tripathi, V.; Thapliyal, A. An efficient framework for identification of Tuberculosis and Pneumonia in chest X-ray images using Neural Network. *Procedia Comput. Sci.* **2020**, *171*, 217–224. [CrossRef]
20. Ibrahim, D.M.; Elshennawy, N.M.; Sarhan, A.M. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. *Comput. Biol. Med.* **2021**, *132*, 104348. [CrossRef]
21. Wang, J.; Feng, W.; Liu, C.; Yu, C.; Du, M.; Xu, R.; Qin, T.; Liu, T.Y. Learning Invariant Representations across Domains and Tasks. *arXiv* **2021**, arXiv:2103.05114.
22. Abdrakhmanov, R.; Altynbekov, M.; Abu, A.; Shomanov, A.; Viderman, D.; Lee, M.H. Few-Shot Learning Approach for COVID-19 Detection from X-Ray Images. In Proceedings of the 2021 16th International Conference on Electronics Computer and Computation (ICECCO), Kaskelen, Kazakhstan, 25–26 November 2021; pp. 1–3. [CrossRef]
23. Shome, D.; Kar, T.; Mohanty, S.N.; Tiwari, P.; Muhammad, K.; AlTameem, A.; Zhang, Y.; Saudagar, A.K.J. COVID-Transformer: Interpretable COVID-19 Detection Using Vision Transformer for Healthcare. *Int. J. Environ. Res. Public Health.* **2021**, *18*, 11086. [CrossRef]
24. Jesmar, F.; Montalbo, P. Diagnosing COVID-19 chest x-rays with a lightweight truncated DenseNet with partial layer freezing and feature fusion. *Biomed. Signal Processing Control* **2021**, *68*, 102583.
25. Srivastava, G.; Chauhan, A.; Jangid, M.; Chaurasia, S. CoviXNet: A novel and efficient deep learning model for detection of COVID-19 using chest X-Ray images. *Biomed. Signal Processing Control* **2022**, *78*, 103848. [CrossRef]
26. Altan, A.; Karasu, S. Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique. *Chaos Solitons Fractals* **2020**, *140*, 110071. [CrossRef]

27. Jin, W.; Dong, S.; Dong, C.; Ye, X. Hybrid ensemble model for differential diagnosis between COVID-19 and common viral pneumonia by chest X-ray radiograph. *Comput. Biol. Med.* **2021**, *131*, 104252. [CrossRef]

28. Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Acharya, U.R. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* **2020**, *121*, 103792. [CrossRef]

29. Uçar, E.; Atila, Ü.; Uçar, M.; Akyol, K.l. Automated detection of COVID-19 disease using deep fused features from chest radiography images. *Biomed. Signal Processing Control* **2021**, *69*, 102862. [CrossRef]

30. Hussain, E.; Hasan, M.; Rahman, M.A.; Lee, I.; Tamanna, T.; Parvez, M.Z. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos Solitons. Fractals* **2021**, *142*, 110495. [CrossRef]

31. Gopatoti, A.; Vijayalakshmi, P. CXGNet: A tri-phase chest X-ray image classification for COVID-19 diagnosis using deep CNN with enhanced grey-wolf optimizer. *Biomed. Signal Processing Control* **2022**, *77*, 103860. [CrossRef]

32. Available online: https://www.kaggle.com/unaissait/curated-chest-xray-image-dataset-for-covid19 (accessed on 1 May 2022).

33. Available online: https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset (accessed on 1 May 2022).

34. Available online: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia (accessed on 1 May 2022).

35. Available online: https://www.kaggle.com/datasets/bachrr/covid-chest-xray (accessed on 1 May 2022).

36. Available online: https://www.kaggle.com/datasets/jonathanchan/dlai3-hackathon-phase3-covid19-cxr-challenge (accessed on 1 May 2022).

37. Chowdhury, N.K.; Rahman, M.; Kabir, M.A. PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. *Health Inf. Sci. Syst.* **2020**, *8*, 27. [CrossRef]