A NOVEL COMPARATIVE STUDY FOR AUTOMATIC THREE-CLASS AND FOUR-CLASS COVID-19 CLASSIFICATION ON X-RAY IMAGES USING DEEP LEARNING

Hüseyin Yaşar[1], Murat Ceylan[2]

[1]Ministry of Health of Republic of Turkey, Ankara, Turkey

[2]Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Konya Technical University, Konya, Turkey

E-mail: mirhendise@gmail.com[1] (corresponding author), mceylan@ktun.edu.tr[2]

DOI: https://doi.org/10.22452/mjcs.vol35no4.5

ABSTRACT

The contagiousness rate of the COVID-19 virus, which was evaluated to have been transmitted from an animal to a human during the last months of 2019, is higher than the MERS-Cov and SARS-Cov viruses originating from the same family. The high rate of contagion has caused the COVID-19 virus to spread rapidly to all countries of the world. It is of great importance to be able to detect cases quickly in order to control the spread of the COVID-19 virus. Therefore, the development of systems that make automatic COVID-19 diagnoses using artificial intelligence approaches based on X-ray, CT scans, and ultrasound images are an urgent and indispensable requirement. In order to increase the number of X-ray images used within the study, a mixed data set was created by combining eight different data sets, thus maximizing the scope of the study. In the study, a total of 9,667 X-ray images were used, including 3,405 of COVID-19 samples, 2,780 of bacterial pneumonia samples, 1,493 of viral pneumonia samples and 1,989 of healthy samples. In this study, which aims to diagnose COVID-19 disease using X-ray images, automatic classification has been performed using two different classification structures: COVID-19 Pneumonia/Other Pneumonia/Healthy and COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy. Convolutional Neural Networks (CNNs), a successful deep learning method, were used as a classifier within the study. A total of seven CNN architectures were used: Mobilenetv2, Resnet101, Googlenet, Xception, Densenet201, Efficientnetb0, and Inceptionv3 architectures. The classification results were obtained from the original X-ray images, and the images were obtained by using Local Binary Pattern and Local Entropy. Then, new classification results were calculated from the obtained results using a pipeline algorithm. Detailed results were obtained to meet the scope of the study. According to the results of the experiments carried out, the three most successful CNN architectures for both three-class and four-class automatic classification were Densenet201, Xception, and Inceptionv3, respectively. In addition, it is understood that the pipeline algorithm used in the study is very useful for improving the results. The study results show that up to an improvement of 1.57% were achieved in some comparison parameters.

Keywords: COVID-19, convolutional neural networks, x-ray chest classification, deep learning, local binary pattern, local entropy, densenet201, xception, inceptionv3.

1.0 INTRODUCTION

The contagiousness rate of the COVID-19 virus, which was evaluated to be transmitted from an animal to a human during the last months of 2019, is higher than the MERS-Cov and SARS-Cov viruses originating from the same family. This high rate of contagion has caused the COVID-19 virus to spread rapidly to all countries of the world [1,2]. It is of great importance to be able to detect cases quickly in order to control the spread of the COVID-19 virus. Cases that cannot be detected and quarantined infect other new cases that they are in contact with. Isolation is the main preventive factor for breaking this contagion chain [3].

Detection of the COVID-19 virus is performed by various rapid test kits and the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test [4]. Although rapid kits give much faster results than the RT-PCR test, accuracy and sensitivity levels are limited [4]. However, even though the sensitivity and accuracy of the RT-PCR test are much higher than the rapid test kits, the result can take a few hours. Another disadvantage of the RT-PCR test is that it requires experienced healthcare personnel to take the test sample from the nose and mouth. Also, the fact that the virus has reached the lungs from the mouth and nose is an important factor affecting the accuracy of the test. For all of these reasons, the World Health Organization (WHO) wants to report cases (with the code ICD-10/U07.1) whose RT-PCR test is positive. In addition, although the RT-PCR test is negative, it is also required to report cases (with the code ICD-10/U07.2) found to be COVID-19 with clinical or epidemiological findings [5].

The COVID-19 virus generally manifests itself with severe pneumonia in the lungs [6]. Therefore, radiological imaging of the lung and chest area is the most important clinical data in detecting the presence of the virus. It has been demonstrated in many studies that there are radiological changes, such as interstitial involvement, lung opacities, bilateral ground-glass, and patchy opacity, in the lung and chest region due to the COVID-19 virus [7]. In this context, considering the pressure created by the COVID-19 outbreak in the health system, it is an urgent and important need to establish automatic detection and classification systems that will help radiologists. Studies on the analysis of CT, X-ray, and ultrasound images with current artificial intelligence methods have been started. The information from past studies [8-22], which performed automatic classification in a three-class classification (COVID-19 Pneumonia/Other Pneumonia/Healthy), in which X-ray images were used, the number of images used, classification methods, training, and testing approaches, and study results are given in Table 1 under Appendix section. Similarly, the information from past studies [15,22,23] that performed automatic classification in a four-class classification (namely COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy), the number of images they used, classification methods, training and testing approaches, and study results are shown in Table 2 under Appendix section.

One of the important factors in evaluating the validity of deep learning-based classification studies is the number of sample images used in the study. In this context, the average number of COVID-19 X-ray images used in 15 COVID-19 classification studies (three-class as Pneumonia/Other Pneumonia/Healthy), details of which are shared in Table 1, is 265.3 ± 355.4. Similarly, the average number of COVID-19 X-ray images used in three COVID-19 classification studies (four-class as Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy), details of which are shared in Table 2, is 219.3 ± 130.6. Even though the COVID-19 virus has caused a major epidemic in a short time, it has taken time to create open access data sets and make them available to researchers. For this reason, although the number of images used for initial stage studies is limited, it is at an acceptable level. However, it is still important to carry out studies in larger numbers using real-world data.

In order to increase the number of X-ray images used within the scope of the study and maximize it, a mixed data set was created by combining eight different data sets. In the study, which was automatically classified into three-class as COVID-19 Pneumonia/Other Pneumonia/Healthy and four-class as COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy, a total of 9,667 X-ray images were used, including 3,405 COVID-19, 2,780 bacterial pneumonia, 1,493 viral pneumonia and 1,989 healthy. Therefore, the number of COVID-19 X-ray images used in the study is 12 times and 15 times more, respectively, than the average number of images used in three-class and four-class classification studies. At the same time, the number of COVID-19 X-ray images used in this study is more than twice the number used in any previous study.

When the literature studies in Table 1 and Table 2 are examined, it is seen that many known Convolutional Neural Networks (CNN) architectures are used in the automatic diagnosis of COVID-19 disease via X-ray images. For this reason, some new approaches need to be introduced to improve the classification results. In this context, it is considered that combining the results obtained using different CNN architectures can be an important alternative. In addition, it is considered that for the same CNN architecture, diversifying the input images using texture feature methods may affect the results. Another issue that needs to be revealed is whether it is more successful to use the results obtained by using different CNN architecture and original images in the merging processes, or to use the results obtained using the same CNN architecture and different input images. The study aims to fill the research gaps in question in this context.

In this study, which aims to diagnose COVID-19 disease using X-ray images, automatic classification has been performed under two different titles as COVID-19 Pneumonia/Other Pneumonia/Healthy and COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy. CNN, a successful deep learning method, was used as a classifier within the scope of the study. A total of seven CNN architectures were used: Mobilenetv2, Resnet101, Googlenet, Xception, Densenet201, Efficientnetb0, and Inceptionv3 architectures. The classification results were obtained using the original X-ray images, and the images obtained by using Local Binary Pattern (LBP) and Local Entropy (LE). Then, new classification results were calculated using a pipeline algorithm from the obtained results.

2.0 METHODS

2.1 Used Data

The COVID-19 X-ray images used within the scope of the study were collected by combining five different data sets. First, 462 COVID-19 X-ray images were taken from the data set formed by Cohen et al. [24]. Second, 35 COVID-19 X-ray images from the data set created by Wang et al. [25,26] were included. Third, 243 COVID-19 X-ray images from the data set created by Winther et al. [27,28] were used. The COVID-19 X-ray image collection process was completed by taking 253 images from the data set created by Desai et al. [29,30] and 2,412 images from the data set created by Vayá et al. [31,32]. As a result, a mixed data set containing a total of 3,405 COVID-19 X-ray images from five different data sets was obtained. The images in question are recorded files of different types, such as jpeg, jpg, png, and dicom. Accordingly, they have bit depths ranging from 16-bit to 48-bit. Image sizes vary widely from 154 px × 124 px to 4064 px × 2992 px. A significant portion (more than 90%) of the images in this comprehensive COVID-19 X-ray data set are images taken from the real world.

Bacterial pneumonia and viral pneumonia X-ray images used in the study were taken from the data set created by Kermany et al. [33,34]. In this context, 2,780 bacterial pneumonia and 1,493 viral pneumonia X-ray images were included in the study. The images in question are 24-bit deep and in jpeg format. Image sizes range from 333 px × 127 px to 2292 px × 1552 px.

Healthy X-ray images used in the study were taken from three different data sets. First, 1,583 healthy X-ray images from the data set created by Kermany et al. [33,34] were included in the study. In addition, 80 healthy X-ray images from the Montgomery [35] data set and 326 healthy X-ray images from the Shenzhen [35] data set were used. In this context, a mixed data set containing a total of 1,989 healthy X-ray images was obtained. These images are in jpeg (24-bit) and png (8-bit) format. Image sizes range from 736 px × 536 px to 4892 px × 4020 px.

For the COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy classification, a total of 9,667 X-ray images were used, including 3,405 COVID-19, 2,780 bacterial pneumonia, 1,493 viral pneumonia, and 1,989 healthy. The same images were also used for the COVID-19 Pneumonia/Other Pneumonia/Healthy classification. Other pneumonia images were obtained by combining bacterial pneumonia and viral pneumonia images. As stated earlier, the number of COVID-19 X-ray images used in the study is 12 to 15 times more than the average number of images used in three-class and four-class studies, respectively. Summary information of the X-ray images used within the scope of the study is included in Table 3. Convolutional neural network architectures are used as classifiers within the scope of the study. Due to the working structure of the CNN classifier, the input images must be in a standard form. For this reason, framing was made on the X-ray images to cover the entire chest area. In this way, the area of interest in the image was determined and the remaining unrelated regions were removed from the image. Then, the framed images were re-sized and adjusted to 224 px × 224 px. The bit depths of the images in question were rearranged and standardized to be 8-bit gray-level. After this stage, the image dimensions were rearranged by the CNN architecture to be used.

Table 3: Summary information of the X-ray images used within the scope of the study

Source	**Covid-19 Pneumonia**	**Healthy**	**Other Pneumonia**	
			Bacterial Pneumonia	**Viral Pneumonia**
Cohen et al. [24]	462	X	X	X
Wang et al. [25,26]	35	X	X	X
Winther et al. [27,28]	243	X	X	X
Desai et al. [29,30]	253	X	X	X
Vayá et al. [31,32]	2,412	X	X	X
Kermany et al. [33,34]	X	1,583	2,780	1,493
Montgomery [35]	X	80	X	X
Shenzhen [35]	X	326	X	X
Total	3,405	1,989	2,780	1,493
	3,405	1,989	4,273	
	9,667			

The experiments performed in the study were carried out according to the standard 4-fold cross-validation procedure. In this context, numerical information regarding the division of images for cross validation is given in Table 4.

Table 4: Information on dividing images for cross validation

Class	**1. fold**	**2. fold**	**3. fold**	**4. fold**	**Total**
Covid-19 Pneumonia	851	851	851	852	3,405
Healthy	497	497	497	498	1,989
Bacterial Pneumonia	695	695	695	695	2,780
Viral Pneumonia	373	373	373	374	1,493
Total	2,416	2,416	2,416	2,419	9,667

2.2 Local Binary Pattern

Local Binary Pattern [36] is the process of comparing a processed pixel with neighboring pixels. This comparison provides the new spatial response of the processed pixel. The LBP process is a simple but effective tissue feature analysis method that does not depend on any parameter. Figure 1 describes the general operating structure of the LBP process. Also, it includes the new LBP feature image obtained by applying the LBP process to one of the COVID-19 X-ray images used within the scope of the study. The neighborhood radius value of the LBP operator used in the study was selected as 1. The dimensions of the LBP feature image are smaller than the original image since the LBP process cannot be applied to the starting and ending row and column by the radius size of it. For this reason, the dimensions of the LBP feature image have been rearranged to be 224 px × 224 px.

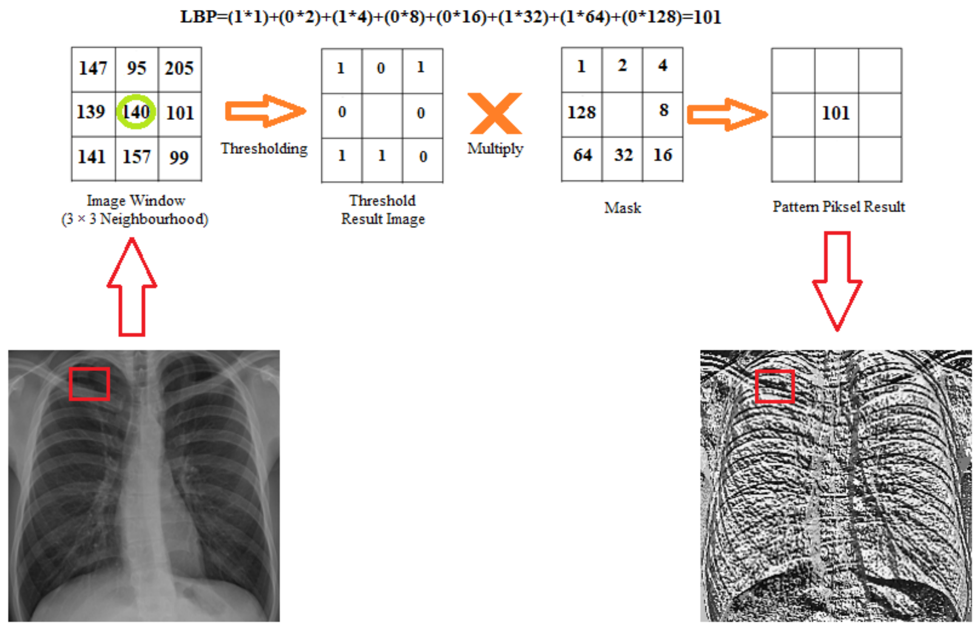

Fig. 1: The general structure of the LBP operator and the image obtained by applying LBP to a COVID-19 X-ray image

2.3 Local Entropy

Local Entropy is used to reveal the amount of uncertainty or randomness from local histograms of an image [37]. The LE feature image can be obtained by applying the entropy filter to the original image. In the scope of the study, the *entropyfilt* function of Matlab 2020 (b) software was used to obtain LE feature images. In the *entropyfilt* function, neighborhood (nhood) parameter is chosen as default (true [9]). It is possible to examine more detailed information about the neighborhood parameter on the Mathworks [38] page.

2.4 Convolutional Neural Network

A Convolutional Neural Network is a deep learning architecture formed by the combination of sub-layers, such as the convolution layer, activation function, pooling layer, flattening, and fully connected layer. The convolution layer is the layer where the convolution process is done by dividing the image into sections. Activation functions, on the other hand, are architectural components that generate new outputs from their inputs, in accordance to their function types. The pooling layer is the layer on which pooling processes are performed in order to reduce the increased image size by the convolution process. The image whose convolution processes have been completed must be converted from matrix form to vector form before entering the classification layer. Flattening is where feature matrices are translated into feature vectors. The fully connected layer is the classification process using feature vectors and machine learning. As machine learning, one of the alternatives, such as support vector machine and artificial neural network can be selected.

Within the scope of the study, a total of seven CNN architectures were processed. These architectures, Mobilenetv2 [39], Resnet101 [40], Googlenet [41], Xception [42], Densenet201 [43], Efficientnetb0 [44], and Inceptionv3 [45], are modified versions made suitable for use in the study. The input image sizes are 224 × 224 × 3 for Mobilenetv2, Resnet101, Googlenet, Densenet201, and Efficientnetb0 architectures, 299 × 299 × 3 for Xception and Inceptionv3 architectures, respectively. Since the images used within the scope of the study are in 8-bit gray-scale format, the inputs in question were rearranged as 224 × 224 × 1 and 299 × 299 × 1, respectively. In addition, the fully connected layer output sizes of these architectures are 1000. The fully connected layer output was rearranged to be 3 and 4, as three-class and four-class classification, respectively was made within the scope of the study.

Information on CNN training options used within the scope of the study is included in Table 5. Matlab 2020 (b) software was used, and it is possible to examine the descriptions of the parameters and the information about the parameters used by default on the Mathworks [46] page. The training options in question were used so as to be the same in all architectures and all experiments performed.

Table 5: CNN training options

Solver for training network	sgdm (stochastic gradient descent with momentum)
Maximum number of epochs	30 (default)
Size of mini-batch	16
Option for data shuffling	every-epoch
Initial learning rate	0.01 (default for sgdm)
Other parameters	default

2.5 Evaluation Criteria of Classification Results

Within the scope of the study, two different multi-class classifications as three-class and four-class were made. The parameters obtained from the confusion matrix elements were used to evaluate the results. It is possible to examine detailed descriptions of confusion matrix elements (TP, FP, TN, and FN) from these studies [15,47,48]. In general, in the multi-class classification given in Table 6, the formulas for calculating the TP, FP, TN, and FN values for each class that make up the classification are given in Equation (1) - (4). In the aforementioned formulas, *i* parameters represent the calculated class, and *n* represents the total number of classes in the confusion matrix.

Table 6: Multi-class confusion matrix

		Actual Class				
		Class-1	Class-2	Class-3	Class-4	Class-n
Predicted Class	Class-1	C_{11}	.	.	.	C_{1n}
	Class-2	.	C_{22}			
	Class-3	.		C_{33}		
	Class-4	.			C_{44}	
	Class-n	C_{n1}				C_{nn}

$$TP_i = C_{ii} \tag{1}$$

$$FN_i = \sum_{l=1}^{n} C_{li} - TP_i \tag{2}$$

$$FP_i = \sum_{l=1}^{n} C_{il} - TP_i \tag{3}$$

$$TN_i = \sum_{l=1}^{n}\sum_{k=1}^{n} C_{lk} - TP_i - FP_i - FN_i \tag{4}$$

Also, within the scope of the study, sensitivity (SEN), specificity (SPE), accuracy (ACC), and F-1 score (F-1) parameters were calculated for each class. The calculation of these parameters is shown between Equation (5) - (8). Also, Receiver Operating Characteristic (ROC) and Area Under the ROC Curve (AUC) values were compared for each class. Detailed descriptions of SEN, SPE, ACC, F-1, and AUC parameters can be examined from these studies [15,47,48].

$$SEN_i = TP_i / (TP_i + FN_i) \tag{5}$$

$$SPE_i = TN_i / (TN_i + FP_i) \tag{6}$$

$$ACC_i = (TP_i + TN_i) / (TP_i + TN_i + FP_i + FN_i) \tag{7}$$

$$F_1 - Score_i = (2 \times TP_i) / (2 \times TP_i + FP_i + FN_i) \tag{8}$$

The calculation of the weighted overall SEN, SPE, ACC, F-1, and AUC values was carried out by weighting the results obtained in each group according to the number of images in the group. In addition, overall accuracy (Overall-ACC) was calculated by proportioning the total number of correctly classified images within the scope of the study to the total number of images used in the study. The mathematical calculation of Overall-ACC can be performed with Equation (9).

$$Overall - ACC = \frac{\sum_{i=1}^{n} C_{ii}}{\sum_{l=1}^{n}\sum_{k=1}^{n} C_{lk}} \tag{9}$$

2.6 Pipeline Algorithm Used Within the Scope of the Study

The results of the study obtained separately using the original images, LBP feature images, and LE feature images were combined using a pipeline algorithm. In this context, the results obtained from the original images were combined with the results from the LBP feature images using a mixing rate of 50%-50%, creating new results. That is, the percentage results produced by the CNN architecture for each class were multiplied by 0.5 and summed. A similar process was performed by combining the original images and the results obtained using LE images with a mixing rate of 50%-50%. The general block diagram of this process is shown in Figure 2. The time required to run the pipeline algorithm is less than a thousandth of a second. Therefore, the time required to generate results for an image using the pipeline algorithm is equal to the sum of the individual result generation times of the original image and the LBP (or LE) feature image. This pipeline algorithm was used by Yasar and Ceylan [49,50] in the two-class classification of COVID-19 and healthy X-ray and CT images and provided successful results.

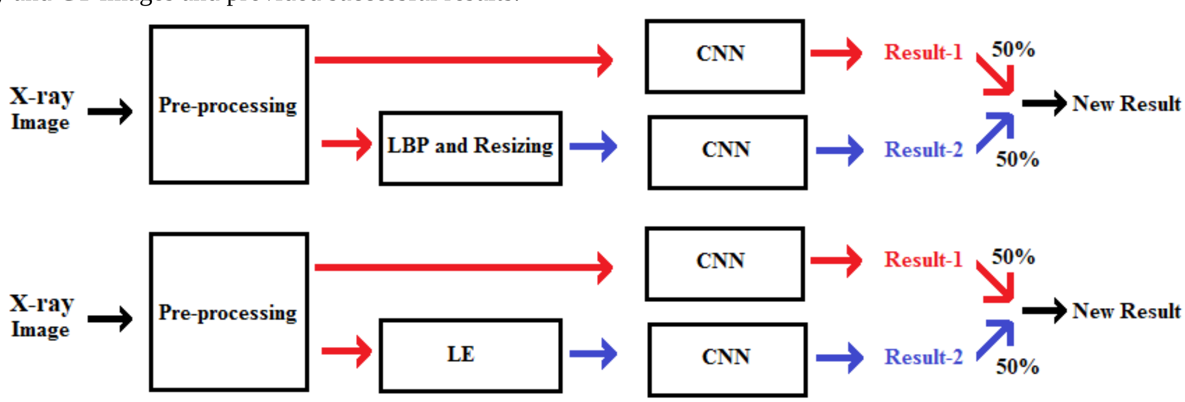

Figure 2: Block diagram of the pipeline approach used in the study

3.0 RESULTS

3.1 COVID-19 Pneumonia/Other Pneumonia/Healthy Classification Results

The experiments within the scope of the study were conducted according to the 4-fold cross-validation approach. As a result of cross validation, the classification results obtained for all images and the real labels were compared to create a confusion matrix. No transfer of weight was assigned for any CNN architecture used. All training started using randomly assigned initial weights. The training options are the same for the seven CNN architectures used. The software used was created and run on the Matlab 2020 (b) platform. In the experiments, the total processing time per image was measured as CPU time in seconds. The experiments in the study were run on Intel (R) Xeon (R) CPU E5-2680 2.7 GHz (32 CPUs) hardware. The hardware in question has 64 GB RAM.

For the COVID-19 Pneumonia/Other Pneumonia/Healthy classification, a total of 9,667 X-ray images were used, including 3,405 COVID-19, 4,273 other pneumonia (2,780 bacterial pneumonia and 1,493 viral pneumonia), and 1,989 healthy. The input image sizes used in the study are 224 × 224 × 1 for Mobilenetv2, Resnet101, Googlenet, Densenet201, and Efficientnetb0 architectures, 299 × 299 × 1 for Xception and Inceptionv3, respectively. The results obtained for the original, LBP, and LE input images using Mobilenetv2, Resnet101, Googlenet, Xception, Densenet201, Efficientnetb0, and Inceptionv3 CNN architectures are listed between Table 7 and Table 13, respectively under Appendix section. In addition, these tables contain new classification results obtained using pipeline algorithms. For the CNN architectures used within the scope of the study, the result acquisition times per image (CPU time/second) are included in Table 14.

Table 14: Comparison of run time (CPU time/ second) for three-class classification

CNN Architecture	Original	LBP	LE	Pipeline (Original-LBP)	Pipeline (Original-LE)
Mobilenetv2	1.7175	1.7149	1.7112	3.4324	3.4286
Resnet101	4.6348	4.6434	4.6470	9.2782	9.2818
Googlenet	0.8115	0.8454	0.8344	1.6570	1.6460
Xception	3.4998	3.4652	3.4572	6.9651	6.9570
Densenet201	11.3934	11.4319	11.3858	22.8253	22.7792
Efficientnetb0	5.1678	5.1760	5.1980	10.3438	10.3658
Inceptionv3	4.5589	4.5833	4.4531	9.1422	9.0120

3.2 COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy Classification Results

For the COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy classification, a total of 9,667 X-ray images were used, including 3,405 COVID-19, 2,780 bacterial pneumonia and 1,493 viral pneumonia, and 1,989 healthy. The input image sizes used in the study are 224 × 224 × 1 for Mobilenetv2, Resnet101, Googlenet, Densenet201, and Efficientnetb0 architectures, 299 × 299 × 1 for Xception and Inceptionv3, respectively. The results obtained for the original, LBP, and LE input images using Mobilenetv2, Resnet101, Googlenet, Xception, Densenet201, Efficientnetb0, and Inceptionv3 CNN architectures are listed between Table 15 and Table 21, respectively under Appendix section. In addition, these tables contain new classification results obtained using pipeline algorithms. For the CNN architectures used within the scope of the study, the result acquisition times per image (CPU time/second) are included in Table 22.

Table 22: Comparison of run time (CPU time/second) for four-class classification

CNN Architecture	Original	LBP	LE	Pipeline (Original-LBP)	Pipeline (Original-LE)
Mobilenetv2	1.3685	1.3698	1.3709	2.7382	2.7394
Resnet101	4.6275	4.6449	4.6448	9.2724	9.2723
Googlenet	0.7231	0.7265	0.7239	1.4495	1.4469
Xception	2.9742	2.9788	2.9805	5.9531	5.9547
Densenet201	11.2532	11.3975	11.3028	22.6506	22.5560
Efficientnetb0	4.1183	4.1328	4.1424	8.2510	8.2606
Inceptionv3	3.5685	3.5727	3.5508	7.1412	7.1193

4.0 DISCUSSION

The COVID-19 Pneumonia/Other Pneumonia/Healthy classification results obtained within the scope of the study are given between Table 7 and Table 13, and the COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy classification results are given between Table 15 and Table 21. However, since very comprehensive results were obtained in the study, more simplified results should be created in order to facilitate comprehension. This section contains these summary results and evaluations.

Table 23 shows the highest weighted parameters, Overall-ACC values, and CPU time obtained without and after using pipeline algorithms for the COVID-19 Pneumonia/Other Pneumonia/Healthy classification. When Table 23 is examined, it is seen that the highest results are obtained by using the original images as the input image without using the pipeline algorithms. In the detailed analysis made in Table 7 and Table 13, it is understood that the results obtained using LE input images are higher than the results obtained using LBP input images, although there are some exceptions (Inceptionv3).

When Table 23 is examined, it is understood that the top five most successful CNN architectures are Densenet201, Xception, Inceptionv3, Resnet101, and Efficientnetb0, respectively. The first five CNN architectures with the slowest results per image, including training and testing, are Densenet201, Efficientnetb0, Resnet101, Inceptionv3, and Xception, respectively. In the case of using pipeline algorithms, these rankings are generally preserved.

When Table 23 is examined, it is seen that the results are improved for all CNN architectures with pipeline algorithms. In this context, an improvement between 0.51% and 0.94% was achieved in the weighted SEN values. Similarly, there was an increase in the weighted SPE values by between 0.16% and 0.64% and in the weighted ACC values by between 0.33% and 0.63%. An improvement between 0.50% and 0.94% was achieved in the weighted F-1 values and an improvement between 0.04% and 0.13% in the weighted AUC values. An increase between 0.51% and 0.94% was achieved in the overall ACC parameter. In pipeline algorithms, it is seen that original image-LE matching generally provides better results than original image-LBP matching.

Table 23: Summary of the results obtained for the COVID-19 Pneumonia/Other Pneumonia/Healthy classification within the scope of the study

Method	**Overall (weighted)**					**Overall**	**CPU**
	SEN	**SPE**	**ACC**	**F-1**	**AUC**	**ACC**	**Time**
Before Pipeline (Mobilenetv2/Original)	0.9602	0.9821	0.9739	0.9605	0.9940	0.9602	1.7175
After Pipeline (Mobilenetv2/Original-LE)	0.9669	0.9837	0.9784	0.9669	0.9949	0.9669	3.4286
Before Pipeline (Resnet101/Original)	0.9607	0.9798	0.9743	0.9608	0.9950	0.9607	4.6348
After Pipeline (Resnet101/Original-LE)	0.9674	0.9833	0.9787	0.9675	0.9957	0.9674	9.2818
Before Pipeline (Googlenet/Original)	0.9559	0.9731	0.9711	0.9556	0.9945	0.9559	0.8115
After Pipeline (Googlenet/Original-LE)	0.9647	0.9796	0.9771	0.9646	0.9958	0.9647	1.6460
Before Pipeline (Xception/Original)	0.9667	0.9846	0.9783	0.9669	0.9961	0.9667	3.4998
After Pipeline (Xception/Original-LE)	0.9728	0.9865	0.9823	0.9728	0.9965	0.9728	6.9570
Before Pipeline (Densenet201/Original)	0.9701	0.9849	0.9802	0.9702	0.9968	0.9701	11.3934
After Pipeline (Densenet201/Original-LE)	0.9753	0.9868	0.9837	0.9753	0.9974	0.9753	22.7792
Before Pipeline (Efficientnetb0/Original)	0.9607	0.9801	0.9743	0.9608	0.9943	0.9607	5.1678
After Pipeline (Efficientnetb0/Original-LE)	0.9658	0.9819	0.9776	0.9657	0.9951	0.9658	10.3658
Before Pipeline (Inceptionv3/Original)	0.9638	0.9812	0.9761	0.9639	0.9962	0.9638	4.5589
After Pipeline (Inceptionv3/Original-LBP)	0.9732	0.9865	0.9825	0.9732	0.9970	0.9732	9.1422

In order to reveal the advantages of using texture feature images in the pipeline algorithm in three-class classification, another experiment was conducted within the scope of the study. The results of the Xception and Densenet201 architectures, which ensure the highest results using the original images, are combined using the same pipeline algorithm. The coupling results obtained in this experiment are shown in Table 24. In addition, the original and LE image results for the Densenet201 architecture and the results obtained by combining these result sets with the pipeline algorithm are given in the same table for an easier understanding of the comparison. The overall accuracy was increased to 0.9732 by combining two result sets with an overall accuracy of 0.9701 and 0.9667 for the Densenet201 and Xception architectures, ensuring the highest results using the original images. By combining the original input image results of the Densenet201 architecture and the LE input image results of the Densenet201 architecture, the overall accuracy is increased to 0.9753. Although a result set with lower overall accuracy (0.9643) is used in this merging process, it is seen that the merging result is higher.

Table 25 shows the highest weighted parameters, Overall-ACC values, and CPU time obtained without and after using pipeline algorithms for four-class (COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy) classification. When Table 25 is examined, it is seen that the highest results are obtained by using the original images as the input image without using the pipeline algorithms. In the detailed analysis made in Table 15 and Table 21, it is understood that the results obtained using LE input images are higher than the results obtained using LBP input images, although there are some exceptions (Densenet201 and Inceptionv3).

When Table 25 is examined, it is seen that the results are improved for all CNN architectures with pipeline algorithms. In this context, an improvement between 0.68% and 1.57% was achieved in the weighted SEN values. Similarly, there was an increase in the weighted SPE values by between 0.23% and 0.48% and in the weighted ACC values by between 0.34% and 0.70%. An improvement between 0.63% and 1.59% was achieved in the weighted F-1 values and an improvement between 0.22% and 0.55% in the weighted AUC values. An increase between 0.68% and 1.57% was achieved in the overall ACC parameter. In pipeline algorithms, it is seen that original image-LE matching generally provides better results than original image-LBP matching.

Table 24: The effect of using texture feature images in the pipeline algorithm on the COVID-19 Pneumonia/Other Pneumonia/Healthy classification results

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall
Densenet201 (Original)	Covid-19	3373	32	6227	35	0.9906	0.9944	0.9931	0.9902	0.9995	0.9701
	Other Pn.	4121	152	5280	114	0.9644	0.9789	0.9725	0.9687	0.9954	
	Healthy	1884	105	7538	140	0.9472	0.9818	0.9747	0.9389	0.9954	
	Overall (weighted)					0.9701	0.9849	0.9802	0.9702	0.9968	
Xception (Original)	Covid-19	3378	27	6239	23	0.9921	0.9963	0.9948	0.9927	0.9993	0.9667
	Other Pn.	4086	187	5283	111	0.9562	0.9794	0.9692	0.9648	0.9946	
	Healthy	1881	108	7490	188	0.9457	0.9755	0.9694	0.9271	0.9940	
	Overall (weighted)					0.9667	0.9846	0.9783	0.9669	0.9961	
Pipeline (Densenet201 (Original)-Xception (Original))	Covid-19	3384	21	6238	24	0.9938	0.9962	0.9953	0.9934	0.9996	0.9732
	Other Pn.	4122	151	5304	90	0.9647	0.9833	0.9751	0.9716	0.9964	
	Healthy	1902	87	7533	145	0.9563	0.9811	0.9760	0.9425	0.9962	
	Overall (weighted)					0.9732	0.9874	0.9824	0.9733	0.9975	
Densenet201 (Original)	Covid-19	3373	32	6227	35	0.9906	0.9944	0.9931	0.9902	0.9995	0.9701
	Other Pn.	4121	152	5280	114	0.9644	0.9789	0.9725	0.9687	0.9954	
	Healthy	1884	105	7538	140	0.9472	0.9818	0.9747	0.9389	0.9954	
	Overall (weighted)					0.9701	0.9849	0.9802	0.9702	0.9968	
Densenet201 (LE)	Covid-19	3385	20	6227	35	0.9941	0.9944	0.9943	0.9919	0.9994	0.9643
	Other Pn.	4117	156	5236	158	0.9635	0.9707	0.9675	0.9633	0.9937	
	Healthy	1820	169	7526	152	0.9150	0.9802	0.9668	0.9190	0.9918	
	Overall (weighted)					0.9643	0.9810	0.9768	0.9643	0.9953	
Pipeline (Densenet201 (Original)-Densenet201 (LE))	Covid-19	3393	12	6239	23	0.9965	0.9963	0.9964	0.9949	0.9996	0.9753
	Other Pn.	4157	116	5283	111	0.9729	0.9794	0.9765	0.9734	0.9964	
	Healthy	1878	111	7573	105	0.9442	0.9863	0.9777	0.9456	0.9958	
	Overall (weighted)					0.9753	0.9868	0.9837	0.9753	0.9974	

Table 25: Summary of the results obtained for the COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy classification within the scope of the study

Method	Overall (weighted)					Overall	CPU
	SEN	SPE	ACC	F-1	AUC	ACC	Time
Before Pipeline (Mobilenetv2/Original)	0.8581	0.9558	0.9374	0.8546	0.9651	0.8581	1.3685
After Pipeline (Mobilenetv2/Original-LE)	0.8738	0.9606	0.9445	0.8705	0.9705	0.8738	2.7394
Before Pipeline (Resnet101/Original)	0.8573	0.9589	0.9370	0.8568	0.9683	0.8573	4.6275
After Pipeline (Resnet101/Original-LE)	0.8708	0.9615	0.9431	0.8690	0.9722	0.8708	9.2723
Before Pipeline (Googlenet/Original)	0.8657	0.9604	0.9408	0.8651	0.9724	0.8657	0.7231
After Pipeline (Googlenet/Original-LE)	0.8799	0.9627	0.9470	0.8772	0.9757	0.8799	1.4469
Before Pipeline (Xception/Original)	0.8766	0.9609	0.9458	0.8730	0.9735	0.8766	2.9742
After Pipeline (Xception/Original-LE)	0.8834	0.9634	0.9492	0.8793	0.9757	0.8834	5.9547
Before Pipeline (Densenet201/Original)	0.8824	0.9648	0.9481	0.8814	0.9745	0.8824	11.2532
After Pipeline (Densenet201/Original-LBP)	0.8954	0.9687	0.9545	0.8934	0.9787	0.8954	22.6506
Before Pipeline (Efficientnetb0/Original)	0.8631	0.9574	0.9396	0.8598	0.9662	0.8631	4.1183
After Pipeline (Efficientnetb0/Original-LE)	0.8750	0.9601	0.9451	0.8706	0.9709	0.8750	8.2606
Before Pipeline (Inceptionv3/Original)	0.8752	0.9628	0.9445	0.8743	0.9730	0.8752	3.5685
After Pipeline (Inceptionv3/Original-LBP)	0.8885	0.9671	0.9511	0.8876	0.9785	0.8885	7.1412

In order to reveal the advantages of using texture feature images in the pipeline algorithm in four-class classification, another experiment was conducted within the scope of the study. The results of the Xception and Densenet201 architectures, which ensure the highest results using the original images, are combined using the same pipeline algorithm. The coupling results obtained in this experiment are shown in Table 26. In addition, the original and LBP image results for the Densenet201 architecture and the results obtained by combining these result sets with the pipeline algorithm are given in the same table for an easier understanding of the comparison. The overall accuracy was increased to 0.8906 by combining two result sets with an overall accuracy of 0.8824 and 0.8766 for the Densenet201 and Xception architectures respectively, ensuring the highest results using the original images. By combining the original input image results of the Densenet201 architecture and the LBP input image results of the Densenet201 architecture, the overall accuracy is increased to 0.8954. Although a result set with lower overall accuracy (0.8687) is used in this merging process, it is seen that the merging result is higher.

Table 26: The effect of using texture feature images in the pipeline algorithm on the COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy classification results

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall
Densenet201 (Original)	Covid-19	3377	28	6234	28	0.9918	0.9955	0.9942	0.9918	0.9994	0.8824
	Bacterial	2336	444	6386	501	0.8403	0.9273	0.9022	0.8318	0.9591	
	Viral	947	546	7720	454	0.6343	0.9445	0.8966	0.6545	0.9205	
	Healthy	1870	119	7524	154	0.9402	0.9799	0.9718	0.9320	0.9940	
	Overall (weighted)					0.8824	0.9648	0.9481	0.8814	0.9745	
Xception (Original)	Covid-19	3385	20	6234	28	0.9941	0.9955	0.9950	0.9930	0.9996	0.8766
	Bacterial	2415	365	6257	630	0.8687	0.9085	0.8971	0.8292	0.9604	
	Viral	809	684	7781	393	0.5419	0.9519	0.8886	0.6004	0.9112	
	Healthy	1865	124	7536	142	0.9377	0.9815	0.9725	0.9334	0.9939	
	Overall (weighted)					0.8766	0.9609	0.9458	0.8730	0.9735	
Pipeline (Densenet201 (Original)-Xception (Original))	Covid-19	3388	17	6235	27	0.9950	0.9957	0.9954	0.9935	0.9996	0.8906
	Bacterial	2445	335	6339	548	0.8795	0.9204	0.9087	0.8470	0.9648	
	Viral	891	602	7822	352	0.5968	0.9569	0.9013	0.6513	0.9261	
	Healthy	1885	104	7547	131	0.9477	0.9829	0.9757	0.9413	0.9955	
	Overall (weighted)					0.8906	0.9654	0.9519	0.8878	0.9774	
Densenet201 (Original)	Covid-19	3377	28	6234	28	0.9918	0.9955	0.9942	0.9918	0.9994	0.8824
	Bacterial	2336	444	6386	501	0.8403	0.9273	0.9022	0.8318	0.9591	
	Viral	947	546	7720	454	0.6343	0.9445	0.8966	0.6545	0.9205	
	Healthy	1870	119	7524	154	0.9402	0.9799	0.9718	0.9320	0.9940	
	Overall (weighted)					0.8824	0.9648	0.9481	0.8814	0.9745	
Densenet201 (LBP)	Covid-19	3393	12	6226	36	0.9965	0.9943	0.9950	0.9930	0.9997	0.8687
	Bacterial	2263	517	6436	451	0.8140	0.9345	0.8999	0.8238	0.9585	
	Viral	894	599	7651	523	0.5988	0.9360	0.8839	0.6144	0.9080	
	Healthy	1848	141	7419	259	0.9291	0.9663	0.9586	0.9023	0.9892	
	Overall (weighted)					0.8687	0.9623	0.9430	0.8672	0.9715	
Pipeline (Densenet201 (Original)-Densenet201 (LBP))	Covid-19	3398	7	6237	25	0.9979	0.9960	0.9967	0.9953	0.9999	0.8954
	Bacterial	2408	372	6454	433	0.8662	0.9371	0.9167	0.8568	0.9672	
	Viral	952	541	7799	375	0.6376	0.9541	0.9052	0.6752	0.9306	
	Healthy	1898	91	7500	178	0.9542	0.9768	0.9722	0.9338	0.9946	
	Overall (weighted)					0.8954	0.9687	0.9545	0.8934	0.9787	

Table 27 and Table 28 include the comparison of the results obtained in the previous three-class and four-class classification studies and the first five best results obtained after and without using pipeline algorithms within the scope of the study.

Table 27: Comparison of the results obtained for the three-class classification within the scope of the study with the results obtained in previous studies

Study	SEN	SPE	ACC/Overall ACC	F-1	AUC
Islam et al. [8]	X	X	0.9940	X	X
Yildirim and Cinar [9]	X	X	0.8889-0.9630	X	X
Rahimzadeh and Attar [10]	X	X	0.8979-0.9140	X	X
Nour et al. [11]	0.8753-0.9461	0.9554-0.9975	0.9209-0.9897	0.8769-0.9672	X
Narayan Das et al. [12]	0.970921	0.972973	0.974068	0.969697	X
Ozturk et al. [13]	0.8535	0.9218	0.8702	0.8737	X
Toraman et al. [14]	0.8422	0.9179	0.8919/0.8422	0.8421	X
Khan et al. [15]	0.969	0.975	0.950	0.956	X
Toğaçar et al. [16]	X	X	0.9781-0.9927	X	X
Ucar and Korkmaz [17]	0.6921-0.9826	0.7993-0.9913	0.7637-0.9826	0.6689-0.9825	X
Civit-Masot et al. [18]	0.85-0.86	0.92-0.93	0.85-0.86	0.85-0.86	0.949
Singh et al. [19]	0.956	X	0.958	0.9588	X
Shorfuzzaman and Masud [20]	0.9565-1.000	0.9767-0.9889	0.9411-0.9926	0.9573-0.9889	0.9547-0.9944
Pandit and Banday [21]	0.867	0.951	0.9253	X	X
Loey et al. [22]	0.8148-0.8519	X	0.8148-0.8519	0.8146-0.8519	X
Before Pipeline (Efficientnetb0/Original)	0.9607	0.9801	0.9743/0.9607	0.9608	0.9943
After Pipeline (Efficientnetb0/Original-LE)	0.9658	0.9819	0.9776/0.9658	0.9657	0.9951
Before Pipeline (Resnet101/Original)	0.9607	0.9798	0.9743/0.9607	0.9608	0.9950
After Pipeline (Resnet101/Original-LE)	0.9674	0.9833	0.9787/0.9674	0.9675	0.9957
Before Pipeline (Inceptionv3/Original)	0.9638	0.9812	0.9761/0.9638	0.9639	0.9962
After Pipeline (Inceptionv3/Original-LBP)	0.9732	0.9865	0.9825/0.9732	0.9732	0.9970
Before Pipeline (Xception/Original)	0.9667	0.9846	0.9783/0.9667	0.9669	0.9961
After Pipeline (Xception/Original-LE)	0.9728	0.9865	0.9823/0.9728	0.9728	0.9965
Before Pipeline (Densenet201/Original)	0.9701	0.9849	0.9802/0.9701	0.9702	0.9968
After Pipeline (Densenet201/Original-LE)	0.9753	0.9868	0.9837/0.9753	0.9753	0.9974

Table 28: Comparison of the results obtained for the four-class classification within the scope of the study with the results obtained in previous studies

Study	SEN	SPE	ACC/Overall ACC	F-1	AUC
Khan et al. [15]	0.8992	0.964	0.896	0.898	X
Mahmud et al. [23]	0.899	0.891	0.902	0.904	0.911
Loey et al. [22]	0.6667-0.8056	X	0.6667-0.8056	0.6566-0.8232	X
Before Pipeline (Efficientnetb0)	0.8631	0.9574	0.9396/0.8631	0.8598	0.9662
After Pipeline (Efficientnetb0)	0.8750	0.9601	0.9451/0.8750	0.8706	0.9709
Before Pipeline (Googlenet)	0.8657	0.9604	0.9408/0.8657	0.8651	0.9724
After Pipeline (Googlenet)	0.8799	0.9627	0.9470/0.8799	0.8772	0.9757
Before Pipeline (Inceptionv3)	0.8752	0.9628	0.9445/0.8752	0.8743	0.9730
After Pipeline (Inceptionv3)	0.8885	0.9671	0.9511/0.8885	0.8876	0.9785
Before Pipeline (Xception)	0.8766	0.9609	0.9458/0.8766	0.8730	0.9735
After Pipeline (Xception)	0.8834	0.9634	0.9492/0.8834	0.8793	0.9757
Before Pipeline (Densenet201)	0.8824	0.9648	0.9481/0.8824	0.8814	0.9745
After Pipeline (Densenet201)	0.8954	0.9687	0.9545/0.8954	0.8934	0.9787

5.0 CONCLUSION

An automatic classification study of COVID-19 Pneumonia/Other Pneumonia/Healthy and COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/Healthy was performed using X-ray images. A total of seven CNN architectures (Mobilenetv2, Resnet101, Googlenet, Xception, Densenet201, Efficientnetb0, and Inceptionv3) were used in the experiments. In this respect, it can be said that the study is one of the most comprehensive comparisons in the literature. The classification results were calculated by giving the original images, LBP, and LE feature images as separate inputs to the mentioned CNN architectures. Also, using a pipeline algorithm, the results were combined and further improved. Another feature of the study is that a larger number of COVID-19 images were used than in the previous studies in the literature. In this context, the number of COVID-19 X-ray images used in the study is 12 times and 15 times more, respectively, than the average number of images used in three-class and four-class classification studies. Moreover, the number of COVID-19 X-ray images used in this study is more than twice the number ever used in a study before.

Within the scope of the study, it is seen that using the original images directly provides higher classification results than using LBP and LE feature images. In other words, using LBP and LE feature images as input images alone are not effective in increasing the classification results. In fact, it negatively affects the classification results. However, the introduction of the pipeline algorithm increases the classification results. Within the scope of the study, the pipeline algorithm has been used to combine the original image results with the LBP feature image results for the same CNN architecture. Similarly, original and LE feature image results were also combined. The said pipeline algorithm can also be used to combine the original image results of different CNN architectures. However, in this case, it shows a lower performance than when used in the study. The main reason for this situation is that when the original images are used as input, the set of misclassified images that can be considered and named stubborn, even if they are classified with different CNN architectures, does not change much. However, when the input image type is changed, the incorrectly classified image set changes, even if the CNN architecture is not changed. This allows the pipeline algorithm to produce better results.

In the case of using the pipeline algorithm, there is an approximately two-fold increase in CPU time cost. In order to understand the CPU time cost in question, it would be helpful to examine CPU times if different CNN architectures are used for classification. For example, using the original images for a three-class classification, the second-highest overall accuracy result was 0.9667 with the Xception architecture. The CPU time for the architecture in question is 3.4998 seconds. Using the original images in the same classification title, the highest overall accuracy result was 0.9701 with the Densenet201 architecture. The CPU time for the architecture in question is 11.3934 seconds. For an overall accuracy increase of 0.034, the time cost was approximately 3.25 times higher. When the results obtained using the Original and LE images for the Xception architecture are combined with the pipeline algorithm, the overall accuracy increases to 0.9728 and the CPU time to 6.9570 seconds. This result is higher than the result obtained using the original images for the Densenet201 architecture. However, in terms of CPU runtime, it is almost lower by half. The same is true for the four-class classification.

When the classification successes of the seven CNN architectures used within the scope of the study are compared, the first three CNN architectures that stand out are Densenet201, Xception, and Inceptionv3, respectively. However, these CNN architectures are generally costlier in terms of CPU time than other CNN architectures. When pipeline algorithms are used, the results obtained for this CNN architecture are again the highest.

When the results obtained within the scope of the study are compared with the results obtained in previous studies in the literature, it is seen that high results are obtained with the contribution of the pipeline algorithm used. Almost all CNN architectures used in previous studies are included in the study. It is not healthy to make a complete comparison owing to the differences between training-test procedures and the number of images used in the studies. Many CNN architectures used in past studies were included in the study. By using the pipeline algorithm, the results obtained with the CNN architectures in question have been further enhanced. For this reason, the results of the study are more successful than the previous studies.

The studies to be carried out after this stage will aim to reveal the COVID-19 Pneumonia/Other Pneumonia/Healthy classification results for CT images, as for X-ray images. As in this study, future studies will aim to use multiple comprehensive CNN architectures, input image types and pipeline algorithms. Another important alternative is to test the success of 3D-CNN architectures for input image combinations.

Compliance with Ethical Standards

Conflict of Interest: Dr. Ceylan declares that he has no conflict of interest. Mr. Yasar declares that he has no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] T. P. Velavan and C. G. Meyer, “The COVID- 19 epidemic”, Tropical Medicine & International Health, Vol. 25, No. 3, 2020, pp. 278-280, doi: 10.1111/tmi.13383.

[2] S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng and T. Wu, “The effect of large-scale anti-contagion policies on the COVID-19 pandemic”, Nature, Vol. 584, No. 7820, 2020, pp. 262-267, doi: 10.1038/s41586-020-2404-8.

[3] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, S. Funk and R. M. Eggo, “Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts”, The Lancet Global Health, Vol. 8, No. 4, 2020, pp. e488-e496, doi: 10.1016/S2214-109X(20)30074-7.

[4] E. Kriegova, R. Fillerova and P. Kvapil, “Direct-RT-qPCR Detection of SARS-CoV-2 without RNA Extraction as Part of a COVID-19 Testing Strategy: From Sample to Result in One Hour”, Diagnostics, Vol. 10, No. 8, 2020, pp. 605, doi: 10.3390/diagnostics10080605.

[5] World Health Organization, https://www.who.int/classifications/icd/covid19/en/, 2020.

[6] L. Gattinoni, D. Chiumello and S. Rossi, “COVID-19 pneumonia: ARDS or not?”, Critical Care, Vol. 24, No.1, 2020, pp. 1-3, doi: 10.1186/s13054-020-02880-z.

[7] D. Zhao, F. Yao, L. Wang, L. Zheng, Y. Gao, J. Ye, F. Guo, H. Zhao and R. Gao, “A comparative study on the clinical features of COVID-19 pneumonia to other pneumonias”, Clinical Infectious Diseases, Vol. 71, No. 15, 2020, pp. 756-761, doi: 10.1093/cid/ciaa247.

[8] M. Z. Islam, M. M. Islam and A. Asraf, “A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images”, Informatics in Medicine Unlocked, Vol. 20, 2020, pp. 100412, doi: 10.1016/j.imu.2020.100412.

[9] M. Yildirim and A. Cinar, “A deep learning based hybrid approach for COVID-19 disease detections”, Traitement du Signal, Vol. 37, No. 3, 2020, pp. 461-468, doi: 10.18280/ts.370313.

[10] M. Rahimzadeh and A. Attar, “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”, Informatics in Medicine Unlocked, Vol. 19, 2020, pp. 100360, doi: 10.1016/j.imu.2020.100360.

[11] M. Nour, Z. Cömert and K. Polat, "A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization", Applied Soft Computing, Vol. 97, No. Part A, 2020, pp. 106580, doi: 10.1016/j.asoc.2020.106580.

[12] N. N. Das, N. Kumar, M. Kaur, V. Kumar and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays", IRBM, Vol. 43, No. 2, 2022, pp. 114-119, doi: 10.1016/j.irbm.2020.07.001.

[13] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images", Computers in Biology and Medicine, Vol. 121, 2020, pp. 103792, doi: 10.1016/j.compbiomed.2020.103792.

[14] S. Toraman, T. B. Alakus and I. Turkoglu, "Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks", Chaos, Solitons & Fractals, Vol. 140, 2020, pp. 110122, doi: 10.1016/j.chaos.2020.110122.

[15] A. I. Khan, J. L. Shah, and M. M. Bhat, "Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images", Computer Methods and Programs in Biomedicine, Vol. 196, 2020, pp. 105581, doi: 10.1016/j.cmpb.2020.105581.

[16] M. Toğaçar, B. Ergen and Z. Cömert, "COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches", Computers in Biology and Medicine, Vol. 121, 2020, 103805, doi: 10.1016/j.compbiomed.2020.103805.

[17] F. Ucar and D. Korkmaz, "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based Diagnostic of the Coronavirus Disease 2019 (COVID-19) from X-Ray Images", Medical Hypotheses, Vol. 140, 2020, pp. 109761, doi: 10.1016/j.mehy.2020.109761.

[18] J. Civit-Masot, F. Luna-Perejón, M. Domínguez Morales and A. Civit, "Deep learning system for COVID-19 diagnosis aid using X-ray pulmonary images", Applied Sciences, Vol. 10, No. 13, 2020, pp. 4640, doi: 10.3390/app10134640.

[19] K. K. Singh, M. Siddhartha and A. Singh, "Diagnosis of Coronavirus Disease (COVID-19) from Chest X-ray images using modified XceptionNet", Romanian Journal of Information Science and Technology, Vol. 23, No. 657, 2020, pp. 91-105.

[20] M. Shorfuzzaman and M. Masud, "On the detection of covid-19 from chest x-ray images using cnn-based transfer learning", CMC-Computers Materials & Continua, Vol. 64, No. 3, 2020, pp. 1359-1381, doi: 10.32604/cmc.2020.011326.

[21] M. K. Pandit and S. A. Banday, "SARS n-CoV2-19 detection from chest x-ray images using deep neural networks", International Journal of Pervasive Computing and Communications, Vol. 16, No. 5, 2020, pp. 419-427, doi: 10.1108/IJPCC-06-2020-0060.

[22] M. Loey, F. Smarandache and N. E. M. Khalifa, "Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning", Symmetry, Vol. 12, No. 4, 2020, pp. 651, doi: 10.3390/sym12040651.

[23] T. Mahmud, M. A. Rahman and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization", Computers in Biology and Medicine, Vol. 122, 2020, pp. 103869, doi: 10.1016/j.compbiomed.2020.103869.

[24] J. P. Cohen, P. Morrison and L. Dao, "COVID-19 image data collection", arXiv, 2020, doi: 10.48550/arXiv.2003.11597.

[25] L. Wang, Z. Q. Lin and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images", Scientific Reports, Vol. 10, No. 1, 2020, pp. 1-12, doi: 10.1038/s41598-020-76550-z.

[26] GitHub, https://github.com/agchung/Figure1-COVID-chestxray-dataset, 2020.

[27] H. B. Winther, H. Laser, S. Gerbel, S. K. Maschke, J. B. Hinrichs, J. Vogel-Claussen, F. K. Wacker, M. M. Höper and B. C. Meyer, "Dataset: COVID-19 Image Repository", 2020, doi: 10.25835/0090041.

[28] GitHub, https://github.com/ml-workgroup/covid-19-image-repository, 2020.

[29] S. Desai, A. Baghal, T. Wongsurawat, S. Al-Shukri, K. Gates, P. Farmer, M. Rutherford, G. D. Blake, T. Nolan, T. Powell, K. Sexton, W. Bennett and F. Prior, "Data from Chest Imaging with Clinical and Genomic Correlates Representing a Rural COVID-19 Positive Population [Data set]", The Cancer Imaging Archive, 2020, doi: 10.7937/tcia.2020.py71-5978.

[30] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox and F. Prior "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository", Journal of Digital Imaging, Vol. 26, No. 6, 2013, pp. 1045-1057, doi: 10.1007/s10278-013-9622-7.

[31] M. D. L. I. Vayá, J. M. Saborit, J. A. Montell, A. Pertusa, A. Bustos, M. Cazorla, J. Galant, X. Barber, D. Orozco-Beltrán, F. García-García, M. Caparrós, G. González and J. M. Salinas, "BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients", arXiv, 2020, doi: 10.48550/arXiv.2006.01174.

[32] Medical Imaging Databank of the Valencia Region (BIMCV), https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/, 2020.

[33] D. S. Kermany, M. Goldbaum, W. Cai, C. C. S. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, M. K. Prasadha, J. Pei, M. Y. L. Ting, J. Zhu, C. Li, S. Hewett, J. Dong, I. Ziyar, A. Shi, R. Zhang, L. Zheng, R. Hou, W. Shi, X. Fu, Y. Duan, V. A. N. Huu, C. Wen, E. D. Zhang, C. L. Zhang, O. Li, X. Wang, M. A. Singer, X. Sun, J. Xu, A. Tafreshi, M. A. Lewis, H. Xia and K. Zhang, "Identifying medical diagnoses and treatable diseases by image-based deep learning", Cell, Vol. 172, No. 5, 2018, pp. 1122-1131. doi: 10.1016/j.cell.2018.02.010.

[34] Kaggle, https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia, 2020.

[35] S. Jaeger, S. Candemir, S. Antani, Y. X. J. Wáng, P. X. Lu and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases", Quantitative Imaging in Medicine and Surgery, Vol. 4, No. 6, 2014, pp. 475-477, doi: 10.3978/j.issn.2223-4292.2014.11.20.

[36] T. Ojala, M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on featured distributions", Pattern Recognition, Vol. 29, No. 1, 1996, pp. 51-59, doi: 10.1016/0031-3203(95)00067-4.

[37] W. Yang, L. Cai and F. Wu, "Image segmentation based on gray level and local relative entropy two dimensional histogram", Plos One, Vol. 15, No. 3, 2020, pp. e0229651, doi: 10.1371/journal.pone.0229651.

[38] MathWoks, https://www.mathworks.com/help/images/ref/entropyfilt.html, 2020.

[39] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2018, pp.4510-4520, doi: 10.1109/CVPR.2018.00474.

[40] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[41] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[42] F. Chollet, "Xception: Deep learning with depthwise separable convolutions", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2017, pp. 1251-1258, doi: 10.1109/CVPR.2017.195.

[43] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2017, pp. 4700-4708, doi: 10.1109/CVPR.2017.243.

[44] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks", in International Conference on Machine Learning, 2019, pp. 6105-6114.

[45] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the inception architecture for computer vision", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, IEEE, 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.

[46] MathWorks, https://www.mathworks.com/help/deeplearning/ref/trainingoptions.html, 2020.

[47] H. Yasar and M. Ceylan, "A novel comparative study for detection of Covid-19 on CT lung images using texture analysis, machine learning, and deep learning methods", Multimedia Tools and Applications, Vol. 80, No. 4, 2021, pp. 5423-5447, doi: 10.1007/s11042-020-09894-3.

[48] F. Hardalac, H. Yasar, A. Akyel and U. Kutbay, "A novel comparative study using multi-resolution transforms and convolutional neural network (CNN) for contactless palm print verification and identification", Multimedia Tools and Applications, Vol. 79, No. 31, 2020, pp. 22929-22963, doi: 10.1007/s11042-020-09005-2.

[49] H. Yasar and M. Ceylan, "A new deep learning pipeline to detect Covid-19 on chest X-ray images using local binary pattern, dual tree complex wavelet transform and convolutional neural networks", Applied Intelligence, Vol. 51, No. 5, 2021, pp. 2740-2763, doi: 10.1007/s10489-020-02019-1.

[50] H. Yasar and M. Ceylan, "Deep Learning–Based Approaches to Improve Classification Parameters for Diagnosing COVID-19 from CT Images", Cognitive Computation, 2021, pp. 1-28, doi: 10.1007/s12559-021-09915-9.

APPENDIX

Table 1: Information from previous studies for COVID-19 Pneumonia/Other Pneumonia/ Healthy classification using X-ray Images

Study	Number of X-Ray Images	Methods	Train-Test Methods	Results
Islam et al. [8]	4,575 images (1,525 Covid-19, 1,525 Other Pneumonia, and 1,525 Healthy)	Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM)	Train: 3,660 images (1,220 Covid-19, 1,220 Other Pneumonia, and 1,220 Healthy); Test: 915 images (305 Covid-19, 305 Other Pneumonia, and 305 Healthy)	**Covid-19:** Sen: 0.990-0.993; Spe: 0.982-0.992; Acc: 0.985-0.992; F-1 Score: 0.977-0.989; AUC: 0.953-0.999; **Other Pneumonia:** Sen: 0.964-0.980; Spe: 0.997-0.998; Acc: 0.986-0.992; F-1 Score: 0.978-0.988; **Healthy:** Sen: 1.000; Spe: 0.997-0.998; Acc: 0.998-0.999; F-1 Score: 0.997-0.998; **Overall:** Acc: 0.994
Yildirim and Cinar [9]	543 images (136 Covid-19, 162 Other Pneumonia, and 245 Healthy)	Convolutional Neural Network (Alexnet, Resnet50, Googlenet, VGG16, and Developed Hybrid Architectures)	80% Train-20% Test	**Covid-19:** Sen: 0.8438-0.9630; Spe: 0.9737-1.000; Acc: 0.9505-0.9811; F-1 Score: 0.9153-0.9643; **Other Pneumonia:** Sen: 0.8596-1.000; Spe: 0.9032-1.000; Acc: 0.9245-0.9811; F-1 Score: 0.9245-0.9800; **Healthy:** Sen: 0.9286-1.000; Spe: 0.8721-0.9650; Acc: 0.8889-0.9630; F-1 Score: 0.7778-0.9355; **Overall:** Acc: 0.8889-0.9630
Rahimzadeh and Attar [10]	15,085 images (180 Covid-19, 6,054 Other Pneumonia, and 8,851 Healthy)	Modified Deep Convolutional Neural Network (Based on the Concatenation of Xception and Resnet50v2)	Train: 3,783 images (149 Covid-19, 1,634 Other Pneumonia, and 2,000 Healthy); Test: 11,302 images (31 Covid-19, 4,420 Other Pneumonia, and 6,851 Healthy) and 5-fold	**Covid-19:** Sen: 0.7335-0.8053; Spe: 0.9933-0.9956; Acc: 0.9926-0.9950; **Other Pneumonia:** Sen: 0.8554-0.8895; Spe: 0.9298-0.9432; Acc: 0.9007-0.9160; **Healthy:** Sen: 0.9260-0.9406; Spe: 0.8664-0.8963; Acc: 0.9025-0.9171; **Overall:** Acc: 0.8979-0.9140
Nour et al. [11]	2,905 images (219 Covid-19, 1,345 Other Pneumonia, and 1,341 Healthy)	Deep Features, Bayesian Optimization, Support Vector Machine, Decision Tree, and k-Nearest Neighbor	Train: 2,033 images (153 Covid-19, 941 Other Pneumonia, and 939 Healthy); Test: 872 images (66 Covid-19, 404 Other Pneumonia, and 402 Healthy)	**Overall:** Sen: 0.8753-0.9461; Spe: 0.9554-0.9975; Acc: 0.9209-0.9897; F-1 Score: 0.8769-0.9672
Narayan Das et al. [12]	1,125 images (125 Covid-19, 500 Other Pneumonia, and 500 Healthy)	Transfer Learning with Convolutional Neural Networks (Xception)	70% Train-10% Validation-20% Test	**Overall:** Sen: 0.970921; Spe: 0.972973; Acc: 0.974068; F-1Score: 0.969697
Ozturk et al. [13]	1,125 images (125 Covid-19, 500 Other Pneumonia, and 500 Healthy)	Convolutional Neural Network (Darknet)	5-fold	**Overall:** Sen: 0.8535; Spe: 0.9218; Acc: 0.8702; F-1 Score: 0.8737
Toraman et al. [14]	2,331 images (231 Covid-19, 1,050 Other Pneumonia, and 1,050 Healthy)	Convolutional Neural Network (Capsnet)	10-fold	**Overall:** Sen: 0.8422; Spe: 0.9179; Acc: 0.8919; F-1 Score: 0.8421 (Note: The results show the average fold.); Avagare Acc: 0.8422
Khan et al. [15]	1,251 images (284 Covid-19, 657 Other Pneumonia, and 310 Healthy)	Convolutional Neural Network (Coronet (Xception))	4-fold	**Overall:** Sen: 0.969; Spe: 0.975; Acc: 0.95; F-1 Score: 0.956
Toğaçar et al. [16]	458 images (295 Covid-19, 98 Other Pneumonia, and 65 Healthy)	Convolutional Neural Network (Squeezenet and Mobilenetv2), Social Mimic Optimization, and Support Vector Machines (SVM)	70% Train-30% Test and 5-fold	**Covid-19:** Sen: 0.9932-1.000; Spe: 0.9937-1.000; Acc: 0.9926-1.000; F-1 Score: 0.9944-1.000; **Other Pneumonia:** Sen: 0.9655-1.000; Spe: 0.9815-0.9907; Acc: 0.9781-0.9927; F-1 Score: 0.9491-0.9831; **Healthy:** Sen: 0.90-0.95; Spe: 0.9914-1.000; Acc: 0.9781-0.9927; F-1 Score: 0.9231-0.9743; **Overall:** Acc: 0.9781-0.9927
Ucar and Korkmaz [17]	5,949 images (76 Covid-19, 4,290 Other Pneumonia, and 1,583 Healthy)	Convolutional Neural Network (Deep Bayes-Squeezenet)	Train: 5,310 images (66 Covid-19, 3,895 Other Pneumonia, and 1,349 Healthy) Test: 639 images (10 Covid-19, 395 Other	**Covid-19:** Sen: 0.7-1.000; Spe: 0.9904-0.9967; Acc: 0.7-1.000; F-1 Score: 0.6087-0.9967; **Other Pneumonia:** Sen: 0.9673-0.9873; Spe: 0.4098-0.9902; Acc: 0.9673-0.9873; F-1 Score: 0.8396-0.9737; **Healthy:**

			Pneumonia, and 234 Healthy)	Sen: 0.3889-0.9804; Spe: 0.9869-0.9975; Acc: 0.3889-0.9804; F-1 Score: 0.5583-0.9772; **Overall:** Sen: 0.6921-0.9826; Spe: 0.7993-0.9913; Acc: 0.7637-0.9826; F-1 Score: 0.6689-0.9825
Civit-Masot et al. [18]	396 images (132 Covid-19, 132 Other Pneumonia, and 132 Healthy)	Convolutional Neural Network (VGG16)	Train: 316 images (105 Covid-19, 106 Other Pneumonia, and 105 Healthy) Test: 80 images (27 Covid-19, 26 Other Pneumonia, and 27 Healthy)	**Covid-19:** Sen: 0.96-1.000; F-1 Score: 0.91-0.92; AUC: 0.989; **Other Pneumonia:** Sen: 0.69-0.73; F-1 Score: 0.78-0.81; AUC: 0.897-0.902; **Healthy:** Sen: 0.81-0.93; F-1 Score: 0.81-0.88; AUC: 0.941; **Overall (Macro):** Sen: 0.85-0.86; Spe: 0.92-0.93; Acc: 0.85-0.86; F-1 Score: 0.85-0.86; AUC: 0.949
Singh et al. [19]	1,419 images (132 Covid-19, 619 Other Pneumonia, and 668 Healthy)	Convolutional Neural Network	Train: 1,135 images (106 Covid-19, 495 Other Pneumonia, and 534 Healthy); Test: 284 images (26 Covid-19, 124 Other Pneumonia, and 134 Healthy)	**Covid-19:** Sen: 0.96; Acc: 0.9894; F-1 Score: 0.94; **Other Pneumonia:** Sen: 0.94; Acc: 0.9613; F-1 Score: 0.96; **Healthy:** Sen: 0.97; Acc: 0.9648; F-1 Score: 0.96; **Overall:** Sen: 0.956; Acc: 0.958; F-1 Score: 0.9588
Shorfuzzaman and Masud [20]	678 images (226 Covid-19, 226 Other Pneumonia, and 226 Healthy)	Transfer Learning with Convolutional Neural Networks (VGG16, Resnet50v2, Mobilenet, Xception, Densenet121 and Ensemble)	5-fold	**Overall:** Sen: 0.9565-1.000; Spe: 0.9767-0.9889; Acc: 0.9411-0.9926; F-1 Score: 0.9573-0.9889; AUC: 0.9547-0.9944
Pandit and Banday [21]	1,428 images (224 Covid-19, 700 Other Pneumonia, and 504 Healthy)	Transfer Learning with Convolutional Neural Networks (VGG16)	70% Train-30% Test	**Overall:** Sen: 0.867; Spe: 0.951; Acc: 0.9253
Loey et al. [22]	227 images (69 Covid-19, 79 Other Pneumonia, and 79 Healthy)	Transfer Learning with Convolutional Neural Networks (Alexnet, Googlenet, and Resnet18)	Train: 200 images (60 Covid-19, 70 Other Pneumonia, and 70 Healthy); Test: 27 images (9 Covid-19, 9 Other Pneumonia, and 9 Healthy)	**Covid-19:** Acc: 0.818-1.000; **Other Pneumonia:** Acc: 0.643-0.875; **Healthy:** Acc: 0.75-1.000; **Overall:** Sen: 0.8148-0.8519; Acc: 0.8148-0.8519; F-1 Score: 0.8146-0.8519

Table 2: Information from previous studies for COVID-19 Pneumonia/Bacterial Pneumonia/Viral Pneumonia/ Healthy classification using X-ray Images

Study	Number of X-Ray Images	Methods	Train-Test Methods	Results
Khan et al. [15]	1,251 images (284 Covid-19, 330 Bacterial Pneumonia, 327 Viral Pneumonia and 310 Healthy)	Convolutional Neural Network (Coronet (Xception)	4-fold	**Overall:** Sen: 0.8992; Spe: 0.964; Acc: 0.896; F-1 Score: 0.898
Mahmud et al. [23]	1,220 images (305 Covid-19, 305 Bacterial Pneumonia, 305 Viral Pneumonia and 305 Healthy)	Transfer Learning with Convolutional Neural Networks (Stacked Multi-Resolution Covxnet)	5-fold	**Overall:** Sen: 0.899; Spe: 0.891; Acc: 0.902; F-1 Score: 0.904; AUC: 0.911
Loey et al. [22]	306 images (69 Covid-19, 79 Bacterial Pneumonia, 79 Viral Pneumonia, and 79 Healthy)	Transfer Learning with Convolutional Neural Networks (Alexnet, Googlenet, and Resnet18)	Train: 270 images (60 Covid-19, 70 Bacterial Pneumonia, 70 Viral Pneumonia, and 70 Healthy); Test: 36 images (9 Covid-19, 9 Bacterial Pneumonia, 9 Viral Pneumonia, and 9 Healthy)	**Covid-19:** Acc: 1.000; **Bacterial Pneumonia:** Acc: 0.444-0.70; **Viral Pneumonia:** Acc: 0.40-0.667; **Healthy:** Acc: 0.643-1.000; **Overall:** Sen: 0.6667-0.8056; Acc: 0.6667-0.8056; F-1 Score: 0.6566-0.8232

Table 7: Comparison of results obtained using Mobilenetv2 for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3360	45	6229	33	0.9868	0.9947	0.9919	0.9885	0.9991	0.9602
	Other Pn.	4045	228	5274	120	0.9466	0.9778	0.9640	0.9588	0.9914	
	Healthy	1877	112	7446	232	0.9437	0.9698	0.9644	0.9161	0.9908	
	Overall (weighted)					0.9602	0.9821	0.9739	0.9605	0.9940	
LBP	Covid-19	3364	41	6197	65	0.9880	0.9896	0.9890	0.9845	0.9993	0.9413
	Other Pn.	4016	257	5140	254	0.9399	0.9529	0.9471	0.9402	0.9873	
	Healthy	1720	269	7430	248	0.8648	0.9677	0.9465	0.8693	0.9793	
	Overall (weighted)					0.9413	0.9689	0.9618	0.9412	0.9899	
LE	Covid-19	3368	37	6194	68	0.9891	0.9891	0.9891	0.9847	0.9990	0.9461
	Other Pn.	4044	229	5161	233	0.9464	0.9568	0.9522	0.9460	0.9881	
	Healthy	1734	255	7458	220	0.8718	0.9713	0.9509	0.8795	0.9834	
	Overall (weighted)					0.9461	0.9712	0.9649	0.9459	0.9910	
Pipeline (Original-LBP)	Covid-19	3390	15	6232	30	0.9956	0.9952	0.9953	0.9934	0.9997	0.9663
	Other Pn.	4095	178	5270	124	0.9583	0.9770	0.9688	0.9644	0.9929	
	Healthy	1856	133	7506	172	0.9331	0.9776	0.9684	0.9241	0.9902	
	Overall (weighted)					0.9663	0.9835	0.9781	0.9663	0.9948	
Pipeline (Original-LE)	Covid-19	3391	14	6231	31	0.9959	0.9950	0.9953	0.9934	0.9995	0.9669
	Other Pn.	4096	177	5271	123	0.9586	0.9772	0.9690	0.9647	0.9930	
	Healthy	1860	129	7512	166	0.9351	0.9784	0.9695	0.9265	0.9913	
	Overall (weighted)					0.9669	0.9837	0.9784	0.9669	0.9949	

Table 8: Comparison of results obtained using Resnet101 for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3353	52	6229	33	0.9847	0.9947	0.9912	0.9875	0.9979	0.9607
	Other Pn.	4098	175	5230	164	0.9590	0.9696	0.9649	0.9603	0.9942	
	Healthy	1836	153	7495	183	0.9231	0.9762	0.9652	0.9162	0.9920	
	Overall (weighted)					0.9607	0.9798	0.9743	0.9608	0.9950	
LBP	Covid-19	3370	35	6196	66	0.9897	0.9895	0.9896	0.9852	0.9989	0.9366
	Other Pn.	4008	265	5103	291	0.9380	0.9461	0.9425	0.9351	0.9813	
	Healthy	1676	313	7422	256	0.8426	0.9667	0.9411	0.8549	0.9737	
	Overall (weighted)					0.9366	0.9656	0.9588	0.9363	0.9859	
LE	Covid-19	3353	52	6185	77	0.9847	0.9877	0.9867	0.9811	0.9983	0.9467
	Other Pn.	4002	271	5203	191	0.9366	0.9646	0.9522	0.9454	0.9874	
	Healthy	1797	192	7431	247	0.9035	0.9678	0.9546	0.8911	0.9833	
	Overall (weighted)					0.9467	0.9734	0.9648	0.9468	0.9904	
Pipeline (Original-LBP)	Covid-19	3389	16	6231	31	0.9953	0.9950	0.9951	0.9931	0.9995	0.9644
	Other Pn.	4118	155	5228	166	0.9637	0.9692	0.9668	0.9625	0.9940	
	Healthy	1816	173	7531	147	0.9130	0.9809	0.9669	0.9190	0.9917	
	Overall (weighted)					0.9644	0.9807	0.9768	0.9643	0.9954	
Pipeline (Original-LE)	Covid-19	3378	27	6240	22	0.9921	0.9965	0.9949	0.9928	0.9994	0.9674
	Other Pn.	4117	156	5256	138	0.9635	0.9744	0.9696	0.9655	0.9941	
	Healthy	1857	132	7523	155	0.9336	0.9798	0.9703	0.9283	0.9925	
	Overall (weighted)					0.9674	0.9833	0.9787	0.9675	0.9957	

Table 9: Comparison of results obtained using Googlenet for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3347	58	6222	40	0.9830	0.9936	0.9899	0.9856	0.9989	0.9559
	Other Pn.	4154	119	5135	259	0.9722	0.9520	0.9609	0.9565	0.9934	
	Healthy	1740	249	7551	127	0.8748	0.9835	0.9611	0.9025	0.9894	
	Overall (weighted)					0.9559	0.9731	0.9711	0.9556	0.9945	
LBP	Covid-19	3270	135	6206	56	0.9604	0.9911	0.9802	0.9716	0.9977	0.9288
	Other Pn.	4041	232	5049	345	0.9457	0.9360	0.9403	0.9334	0.9861	
	Healthy	1668	321	7391	287	0.8386	0.9626	0.9371	0.8458	0.9748	
	Overall (weighted)					0.9288	0.9609	0.9537	0.9288	0.9879	
LE	Covid-19	3330	75	6194	68	0.9780	0.9891	0.9852	0.9790	0.9980	0.9459
	Other Pn.	4036	237	5177	217	0.9445	0.9598	0.9530	0.9468	0.9905	
	Healthy	1778	211	7440	238	0.8939	0.9690	0.9536	0.8879	0.9855	
	Overall (weighted)					0.9459	0.9720	0.9645	0.9460	0.9921	
Pipeline (Original-LBP)	Covid-19	3380	25	6238	24	0.9927	0.9962	0.9949	0.9928	0.9995	0.9614
	Other Pn.	4155	118	5163	231	0.9724	0.9572	0.9639	0.9597	0.9938	
	Healthy	1759	230	7560	118	0.8844	0.9846	0.9640	0.9100	0.9896	
	Overall (weighted)					0.9614	0.9766	0.9748	0.9611	0.9950	
Pipeline (Original-LE)	Covid-19	3373	32	6230	32	0.9906	0.9949	0.9934	0.9906	0.9993	0.9647
	Other Pn.	4156	117	5206	188	0.9726	0.9651	0.9684	0.9646	0.9948	
	Healthy	1797	192	7557	121	0.9035	0.9842	0.9676	0.9199	0.9921	
	Overall (weighted)					0.9647	0.9796	0.9771	0.9646	0.9958	

Table 10: Comparison of results obtained using Xception for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3378	27	6239	23	0.9921	0.9963	0.9948	0.9927	0.9993	0.9667
	Other Pn.	4086	187	5283	111	0.9562	0.9794	0.9692	0.9648	0.9946	
	Healthy	1881	108	7490	188	0.9457	0.9755	0.9694	0.9271	0.9940	
	Overall (weighted)					0.9667	0.9846	0.9783	0.9669	0.9961	
LBP	Covid-19	3387	18	6217	45	0.9947	0.9928	0.9935	0.9908	0.9995	0.9515
	Other Pn.	4046	227	5190	204	0.9469	0.9622	0.9554	0.9494	0.9905	
	Healthy	1765	224	7458	220	0.8874	0.9713	0.9541	0.8883	0.9848	
	Overall (weighted)					0.9515	0.9749	0.9685	0.9514	0.9925	
LE	Covid-19	3368	37	6225	37	0.9891	0.9941	0.9923	0.9891	0.9991	0.9588
	Other Pn.	4089	184	5220	174	0.9569	0.9677	0.9630	0.9581	0.9918	
	Healthy	1812	177	7491	187	0.9110	0.9756	0.9623	0.9087	0.9892	
	Overall (weighted)					0.9588	0.9786	0.9732	0.9589	0.9939	
Pipeline (Original-LBP)	Covid-19	3393	12	6241	21	0.9965	0.9966	0.9966	0.9952	0.9996	0.9698
	Other Pn.	4113	160	5277	117	0.9626	0.9783	0.9713	0.9674	0.9950	
	Healthy	1869	120	7524	154	0.9397	0.9799	0.9717	0.9317	0.9933	
	Overall (weighted)					0.9698	0.9851	0.9803	0.9698	0.9962	
Pipeline (Original-LE)	Covid-19	3396	9	6247	15	0.9974	0.9976	0.9975	0.9965	0.9995	0.9728
	Other Pn.	4129	144	5285	109	0.9663	0.9798	0.9738	0.9703	0.9954	
	Healthy	1879	110	7539	139	0.9447	0.9819	0.9742	0.9379	0.9940	
	Overall (weighted)					0.9728	0.9865	0.9823	0.9728	0.9965	

Table 11: Comparison of results obtained using Densenet201 for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3373	32	6227	35	0.9906	0.9944	0.9931	0.9902	0.9995	0.9701
	Other Pn.	4121	152	5280	114	0.9644	0.9789	0.9725	0.9687	0.9954	
	Healthy	1884	105	7538	140	0.9472	0.9818	0.9747	0.9389	0.9954	
	Overall (weighted)					0.9701	0.9849	0.9802	0.9702	0.9968	
LBP	Covid-19	3392	13	6223	39	0.9962	0.9938	0.9946	0.9924	0.9994	0.9589
	Other Pn.	4063	210	5236	158	0.9509	0.9707	0.9619	0.9567	0.9919	
	Healthy	1815	174	7478	200	0.9125	0.9740	0.9613	0.9066	0.9887	
	Overall (weighted)					0.9589	0.9795	0.9733	0.9590	0.9939	
LE	Covid-19	3385	20	6227	35	0.9941	0.9944	0.9943	0.9919	0.9994	0.9643
	Other Pn.	4117	156	5236	158	0.9635	0.9707	0.9675	0.9633	0.9937	
	Healthy	1820	169	7526	152	0.9150	0.9802	0.9668	0.9190	0.9918	
	Overall (weighted)					0.9643	0.9810	0.9768	0.9643	0.9953	
Pipeline (Original-LBP)	Covid-19	3395	10	6241	21	0.9971	0.9966	0.9968	0.9955	0.9996	0.9727
	Other Pn.	4124	149	5289	105	0.9651	0.9805	0.9737	0.9701	0.9961	
	Healthy	1884	105	7540	138	0.9472	0.9820	0.9749	0.9394	0.9955	
	Overall (weighted)					0.9727	0.9865	0.9821	0.9727	0.9972	
Pipeline (Original-LE)	Covid-19	3393	12	6239	23	0.9965	0.9963	0.9964	0.9949	0.9996	0.9753
	Other Pn.	4157	116	5283	111	0.9729	0.9794	0.9765	0.9734	0.9964	
	Healthy	1878	111	7573	105	0.9442	0.9863	0.9777	0.9456	0.9958	
	Overall (weighted)					0.9753	0.9868	0.9837	0.9753	0.9974	

Table 12: Comparison of results obtained using Efficientnetb0 for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3371	34	6223	39	0.9900	0.9938	0.9924	0.9893	0.9995	0.9607
	Other Pn.	4083	190	5239	155	0.9555	0.9713	0.9643	0.9595	0.9924	
	Healthy	1833	156	7492	186	0.9216	0.9758	0.9646	0.9147	0.9894	
	Overall (weighted)					0.9607	0.9801	0.9743	0.9608	0.9943	
LBP	Covid-19	3374	31	6172	90	0.9909	0.9856	0.9875	0.9824	0.9986	0.9402
	Other Pn.	3969	304	5175	219	0.9289	0.9594	0.9459	0.9382	0.9861	
	Healthy	1746	243	7409	269	0.8778	0.9650	0.9470	0.8721	0.9777	
	Overall (weighted)					0.9402	0.9698	0.9608	0.9402	0.9888	
LE	Covid-19	3371	34	6208	54	0.9900	0.9914	0.9909	0.9871	0.9993	0.9522
	Other Pn.	4060	213	5193	201	0.9502	0.9627	0.9572	0.9515	0.9890	
	Healthy	1774	215	7471	207	0.8919	0.9730	0.9563	0.8937	0.9850	
	Overall (weighted)					0.9522	0.9749	0.9689	0.9521	0.9918	
Pipeline (Original-LBP)	Covid-19	3389	16	6229	33	0.9953	0.9947	0.9949	0.9928	0.9996	0.9654
	Other Pn.	4113	160	5246	148	0.9626	0.9726	0.9681	0.9639	0.9923	
	Healthy	1831	158	7525	153	0.9206	0.9801	0.9678	0.9217	0.9888	
	Overall (weighted)					0.9654	0.9819	0.9775	0.9654	0.9942	
Pipeline (Original-LE)	Covid-19	3386	19	6232	30	0.9944	0.9952	0.9949	0.9928	0.9997	0.9658
	Other Pn.	4115	158	5242	152	0.9630	0.9718	0.9679	0.9637	0.9933	
	Healthy	1835	154	7529	149	0.9226	0.9806	0.9687	0.9237	0.9910	
	Overall (weighted)					0.9658	0.9819	0.9776	0.9657	0.9951	

Table 13: Comparison of results obtained using Inceptionv3 for three-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3350	55	6224	38	0.9838	0.9939	0.9904	0.9863	0.9992	0.9638
	Other Pn.	4115	158	5243	151	0.9630	0.9720	0.9680	0.9638	0.9952	
	Healthy	1852	137	7517	161	0.9311	0.9790	0.9692	0.9255	0.9933	
	Overall (weighted)					0.9638	0.9812	0.9761	0.9639	0.9962	
LBP	Covid-19	3375	30	6222	40	0.9912	0.9936	0.9928	0.9897	0.9994	0.9558
	Other Pn.	4028	245	5248	146	0.9427	0.9729	0.9596	0.9537	0.9923	
	Healthy	1837	152	7437	241	0.9236	0.9686	0.9593	0.9034	0.9888	
	Overall (weighted)					0.9558	0.9793	0.9712	0.9560	0.9941	
LE	Covid-19	3362	43	6228	34	0.9874	0.9946	0.9920	0.9887	0.9992	0.9536
	Other Pn.	4043	230	5210	184	0.9462	0.9659	0.9572	0.9513	0.9918	
	Healthy	1813	176	7447	231	0.9115	0.9699	0.9579	0.8991	0.9889	
	Overall (weighted)					0.9536	0.9768	0.9696	0.9537	0.9938	
Pipeline (Original-LBP)	Covid-19	3389	16	6237	25	0.9953	0.9960	0.9958	0.9940	0.9996	0.9732
	Other Pn.	4139	134	5288	106	0.9686	0.9803	0.9752	0.9718	0.9961	
	Healthy	1880	109	7550	128	0.9452	0.9833	0.9755	0.9407	0.9945	
	Overall (weighted)					0.9732	0.9865	0.9825	0.9732	0.9970	
Pipeline (Original-LE)	Covid-19	3387	18	6244	18	0.9947	0.9971	0.9963	0.9947	0.9996	0.9728
	Other Pn.	4137	136	5280	114	0.9682	0.9789	0.9741	0.9707	0.9960	
	Healthy	1880	109	7547	131	0.9452	0.9829	0.9752	0.9400	0.9946	
	Overall (weighted)					0.9728	0.9861	0.9821	0.9728	0.9970	

Table 15: Comparison of results obtained using Mobilenetv2 for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3372	33	6207	55	0.9903	0.9912	0.9909	0.9871	0.9992	0.8581
	Bacterial	2337	443	6247	640	0.8406	0.9071	0.8880	0.8119	0.9488	
	Viral	774	719	7702	472	0.5184	0.9423	0.8768	0.5652	0.8851	
	Healthy	1812	177	7473	205	0.9110	0.9733	0.9605	0.9046	0.9897	
	Overall (weighted)					0.8581	0.9558	0.9374	0.8546	0.9651	
LBP	Covid-19	3352	53	6170	92	0.9844	0.9853	0.9850	0.9788	0.9985	0.8292
	Bacterial	2166	614	6325	562	0.7791	0.9184	0.8783	0.7865	0.9428	
	Viral	726	767	7582	592	0.4863	0.9276	0.8594	0.5165	0.8688	
	Healthy	1772	217	7273	405	0.8909	0.9473	0.9357	0.8507	0.9769	
	Overall (weighted)					0.8292	0.9493	0.9248	0.8258	0.9580	
LE	Covid-19	3361	44	6193	69	0.9871	0.9890	0.9883	0.9835	0.9989	0.8419
	Bacterial	2159	621	6293	594	0.7766	0.9138	0.8743	0.7804	0.9392	
	Viral	844	649	7570	604	0.5653	0.9261	0.8704	0.5740	0.8861	
	Healthy	1775	214	7417	261	0.8924	0.9660	0.9509	0.8820	0.9831	
	Overall (weighted)					0.8419	0.9529	0.9296	0.8409	0.9611	
Pipeline (Original-LBP)	Covid-19	3396	9	6214	48	0.9974	0.9923	0.9941	0.9917	0.9994	0.8701
	Bacterial	2372	408	6318	569	0.8532	0.9174	0.8989	0.8292	0.9566	
	Viral	784	709	7791	383	0.5251	0.9531	0.8870	0.5895	0.8995	
	Healthy	1859	130	7422	256	0.9346	0.9667	0.9601	0.9059	0.9897	
	Overall (weighted)					0.8701	0.9594	0.9432	0.8652	0.9697	
Pipeline (Original-LE)	Covid-19	3389	16	6220	42	0.9953	0.9933	0.9940	0.9915	0.9994	0.8738
	Bacterial	2369	411	6303	584	0.8522	0.9152	0.8971	0.8264	0.9550	
	Viral	831	662	7762	412	0.5566	0.9496	0.8889	0.6075	0.9059	
	Healthy	1858	131	7496	182	0.9341	0.9763	0.9676	0.9223	0.9913	
	Overall (weighted)					0.8738	0.9606	0.9445	0.8705	0.9705	

Table 16: Comparison of results obtained using Resnet101 for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3359	46	6217	45	0.9865	0.9928	0.9906	0.9866	0.9984	0.8573
	Bacterial	2184	596	6380	507	0.7856	0.9264	0.8859	0.7984	0.9505	
	Viral	883	610	7567	607	0.5914	0.9257	0.8741	0.5920	0.9028	
	Healthy	1862	127	7458	220	0.9361	0.9713	0.9641	0.9148	0.9907	
	Overall (weighted)					0.8573	0.9589	0.9370	0.8568	0.9683	
LBP	Covid-19	3369	36	6181	81	0.9894	0.9871	0.9879	0.9829	0.9987	0.8358
	Bacterial	2284	496	6174	713	0.8216	0.8965	0.8749	0.7907	0.9417	
	Viral	683	810	7651	523	0.4575	0.9360	0.8621	0.5061	0.8649	
	Healthy	1744	245	7408	270	0.8768	0.9648	0.9467	0.8713	0.9778	
	Overall (weighted)					0.8358	0.9486	0.9275	0.8311	0.9573	
LE	Covid-19	3364	41	6164	98	0.9880	0.9844	0.9856	0.9798	0.9985	0.8484
	Bacterial	2163	617	6359	528	0.7781	0.9233	0.8816	0.7907	0.9419	
	Viral	871	622	7574	600	0.5834	0.9266	0.8736	0.5877	0.8852	
	Healthy	1803	186	7438	240	0.9065	0.9687	0.9559	0.8943	0.9834	
	Overall (weighted)					0.8484	0.9547	0.9323	0.8473	0.9616	
Pipeline (Original-LBP)	Covid-19	3390	15	6215	47	0.9956	0.9925	0.9936	0.9909	0.9995	0.8700
	Bacterial	2384	396	6285	602	0.8576	0.9126	0.8968	0.8269	0.9566	
	Viral	778	715	7777	397	0.5211	0.9514	0.8850	0.5832	0.9073	
	Healthy	1858	131	7467	211	0.9341	0.9725	0.9646	0.9157	0.9915	
	Overall (weighted)					0.8700	0.9591	0.9430	0.8653	0.9713	
Pipeline (Original-LE)	Covid-19	3389	16	6202	60	0.9953	0.9904	0.9921	0.9889	0.9993	0.8708
	Bacterial	2267	513	6401	486	0.8155	0.9294	0.8967	0.8194	0.9568	
	Viral	895	598	7672	502	0.5995	0.9386	0.8862	0.6194	0.9117	
	Healthy	1867	122	7477	201	0.9387	0.9738	0.9666	0.9204	0.9926	
	Overall (weighted)					0.8708	0.9615	0.9431	0.8690	0.9722	

Table 17: Comparison of results obtained using Googlenet for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3384	21	6182	80	0.9938	0.9872	0.9896	0.9853	0.9987	0.8657
	Bacterial	2257	523	6432	455	0.8119	0.9339	0.8988	0.8219	0.9594	
	Viral	933	560	7623	551	0.6249	0.9326	0.8851	0.6268	0.9142	
	Healthy	1795	194	7466	212	0.9025	0.9724	0.9580	0.8984	0.9890	
	Overall (weighted)					0.8657	0.9604	0.9408	0.8651	0.9724	
LBP	Covid-19	3315	90	6189	73	0.9736	0.9883	0.9831	0.9760	0.9972	0.8128
	Bacterial	1950	830	6448	439	0.7014	0.9363	0.8687	0.7545	0.9419	
	Viral	920	573	7230	944	0.6162	0.8845	0.8431	0.5481	0.8723	
	Healthy	1672	317	7324	354	0.8406	0.9539	0.9306	0.8329	0.9721	
	Overall (weighted)					0.8128	0.9502	0.9178	0.8168	0.9568	
LE	Covid-19	3354	51	6173	89	0.9850	0.9858	0.9855	0.9796	0.9983	0.8465
	Bacterial	2294	486	6235	652	0.8252	0.9053	0.8823	0.8013	0.9470	
	Viral	809	684	7635	539	0.5419	0.9341	0.8735	0.5695	0.8957	
	Healthy	1726	263	7474	204	0.8678	0.9734	0.9517	0.8808	0.9824	
	Overall (weighted)					0.8465	0.9521	0.9316	0.8446	0.9645	
Pipeline (Original-LBP)	Covid-19	3385	20	6207	55	0.9941	0.9912	0.9922	0.9890	0.9993	0.8697
	Bacterial	2218	562	6486	401	0.7978	0.9418	0.9004	0.8216	0.9610	
	Viral	960	533	7609	565	0.6430	0.9309	0.8864	0.6362	0.9146	
	Healthy	1844	145	7439	239	0.9271	0.9689	0.9603	0.9057	0.9898	
	Overall (weighted)					0.8697	0.9631	0.9429	0.8693	0.9732	
Pipeline (Original-LE)	Covid-19	3392	13	6199	63	0.9962	0.9899	0.9921	0.9889	0.9993	0.8799
	Bacterial	2367	413	6385	502	0.8514	0.9271	0.9053	0.8380	0.9628	
	Viral	892	601	7772	402	0.5975	0.9508	0.8962	0.6401	0.9238	
	Healthy	1855	134	7484	194	0.9326	0.9747	0.9661	0.9188	0.9922	
	Overall (weighted)					0.8799	0.9627	0.9470	0.8772	0.9757	

Table 18: Comparison of results obtained using Xception for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3385	20	6234	28	0.9941	0.9955	0.9950	0.9930	0.9996	0.8766
	Bacterial	2415	365	6257	630	0.8687	0.9085	0.8971	0.8292	0.9604	
	Viral	809	684	7781	393	0.5419	0.9519	0.8886	0.6004	0.9112	
	Healthy	1865	124	7536	142	0.9377	0.9815	0.9725	0.9334	0.9939	
		Overall (weighted)				0.8766	0.9609	0.9458	0.8730	0.9735	
LBP	Covid-19	3382	23	6208	54	0.9932	0.9914	0.9920	0.9887	0.9995	0.8448
	Bacterial	2274	506	6277	610	0.8180	0.9114	0.8846	0.8030	0.9488	
	Viral	699	794	7711	463	0.4682	0.9434	0.8700	0.5266	0.8784	
	Healthy	1812	177	7305	373	0.9110	0.9514	0.9431	0.8682	0.9813	
		Overall (weighted)				0.8448	0.9527	0.9322	0.8391	0.9625	
LE	Covid-19	3384	21	6205	57	0.9938	0.9909	0.9919	0.9886	0.9992	0.8610
	Bacterial	2289	491	6331	556	0.8234	0.9193	0.8917	0.8139	0.9537	
	Viral	795	698	7726	448	0.5325	0.9452	0.8815	0.5811	0.8959	
	Healthy	1855	134	7395	283	0.9326	0.9631	0.9569	0.8990	0.9874	
		Overall (weighted)				0.8610	0.9575	0.9388	0.8570	0.9678	
Pipeline (Original-LBP)	Covid-19	3397	8	6236	26	0.9977	0.9958	0.9965	0.9950	0.9998	0.8818
	Bacterial	2457	323	6289	598	0.8838	0.9132	0.9047	0.8422	0.9627	
	Viral	777	716	7864	310	0.5204	0.9621	0.8939	0.6023	0.9150	
	Healthy	1893	96	7469	209	0.9517	0.9728	0.9684	0.9254	0.9926	
		Overall (weighted)				0.8818	0.9621	0.9485	0.8761	0.9745	
Pipeline (Original-LE)	Covid-19	3395	10	6237	25	0.9971	0.9960	0.9964	0.9949	0.9997	0.8834
	Bacterial	2433	347	6323	564	0.8752	0.9181	0.9058	0.8423	0.9635	
	Viral	828	665	7823	351	0.5546	0.9571	0.8949	0.6198	0.9196	
	Healthy	1884	105	7491	187	0.9472	0.9756	0.9698	0.9281	0.9939	
		Overall (weighted)				0.8834	0.9634	0.9492	0.8793	0.9757	

Table 19: Comparison of results obtained using Densenet201 for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3377	28	6234	28	0.9918	0.9955	0.9942	0.9918	0.9994	0.8824
	Bacterial	2336	444	6386	501	0.8403	0.9273	0.9022	0.8318	0.9591	
	Viral	947	546	7720	454	0.6343	0.9445	0.8966	0.6545	0.9205	
	Healthy	1870	119	7524	154	0.9402	0.9799	0.9718	0.9320	0.9940	
		Overall (weighted)				0.8824	0.9648	0.9481	0.8814	0.9745	
LBP	Covid-19	3393	12	6226	36	0.9965	0.9943	0.9950	0.9930	0.9997	0.8687
	Bacterial	2263	517	6436	451	0.8140	0.9345	0.8999	0.8238	0.9585	
	Viral	894	599	7651	523	0.5988	0.9360	0.8839	0.6144	0.9080	
	Healthy	1848	141	7419	259	0.9291	0.9663	0.9586	0.9023	0.9892	
		Overall (weighted)				0.8687	0.9623	0.9430	0.8672	0.9715	
LE	Covid-19	3383	22	6211	51	0.9935	0.9919	0.9924	0.9893	0.9993	0.8751
	Bacterial	2331	449	6359	528	0.8385	0.9233	0.8989	0.8267	0.9572	
	Viral	892	601	7750	424	0.5975	0.9481	0.8940	0.6351	0.9086	
	Healthy	1854	135	7474	204	0.9321	0.9734	0.9649	0.9162	0.9916	
		Overall (weighted)				0.8751	0.9616	0.9447	0.8728	0.9716	
Pipeline (Original-LBP)	Covid-19	3398	7	6237	25	0.9979	0.9960	0.9967	0.9953	0.9999	0.8954
	Bacterial	2408	372	6454	433	0.8662	0.9371	0.9167	0.8568	0.9672	
	Viral	952	541	7799	375	0.6376	0.9541	0.9052	0.6752	0.9306	
	Healthy	1898	91	7500	178	0.9542	0.9768	0.9722	0.9338	0.9946	
		Overall (weighted)				0.8954	0.9687	0.9545	0.8934	0.9787	
Pipeline (Original-LE)	Covid-19	3396	9	6236	26	0.9974	0.9958	0.9964	0.9949	0.9998	0.8938
	Bacterial	2415	365	6404	483	0.8687	0.9299	0.9123	0.8507	0.9653	
	Viral	931	562	7808	366	0.6236	0.9552	0.9040	0.6674	0.9297	
	Healthy	1898	91	7526	152	0.9542	0.9802	0.9749	0.9398	0.9949	
		Overall (weighted)				0.8938	0.9674	0.9535	0.8915	0.9780	

Table 20: Comparison of results obtained using Efficientnetb0 for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3375	30	6213	49	0.9912	0.9922	0.9918	0.9884	0.9992	0.8631
	Bacterial	2336	444	6265	622	0.8403	0.9097	0.8897	0.8142	0.9511	
	Viral	796	697	7720	454	0.5332	0.9445	0.8809	0.5804	0.8873	
	Healthy	1837	152	7480	198	0.9236	0.9742	0.9638	0.9130	0.9901	
		Overall (weighted)				0.8631	0.9574	0.9396	0.8598	0.9662	
LBP	Covid-19	3370	35	6138	124	0.9897	0.9802	0.9836	0.9770	0.9986	0.8313
	Bacterial	2296	484	6168	719	0.8259	0.8956	0.8756	0.7924	0.9397	
	Viral	646	847	7719	455	0.4327	0.9443	0.8653	0.4981	0.8518	
	Healthy	1724	265	7345	333	0.8668	0.9566	0.9381	0.8522	0.9761	
		Overall (weighted)				0.8313	0.9455	0.9249	0.8243	0.9544	
LE	Covid-19	3370	35	6191	71	0.9897	0.9887	0.9890	0.9845	0.9990	0.8517
	Bacterial	2308	472	6252	635	0.8302	0.9078	0.8855	0.8066	0.9484	
	Viral	776	717	7690	484	0.5198	0.9408	0.8758	0.5637	0.8831	
	Healthy	1779	210	7434	244	0.8944	0.9682	0.9530	0.8868	0.9845	
		Overall (weighted)				0.8517	0.9538	0.9344	0.8483	0.9636	
Pipeline (Original-LBP)	Covid-19	3387	18	6216	46	0.9947	0.9927	0.9934	0.9906	0.9996	0.8661
	Bacterial	2435	345	6220	667	0.8759	0.9032	0.8953	0.8279	0.9546	
	Viral	716	777	7834	340	0.4796	0.9584	0.8845	0.5618	0.8944	
	Healthy	1835	154	7437	241	0.9226	0.9686	0.9591	0.9028	0.9892	
		Overall (weighted)				0.8661	0.9567	0.9413	0.8596	0.9683	
Pipeline (Original-LE)	Covid-19	3389	16	6223	39	0.9953	0.9938	0.9943	0.9920	0.9996	0.8750
	Bacterial	2422	358	6270	617	0.8712	0.9104	0.8991	0.8324	0.9574	
	Viral	796	697	7810	364	0.5332	0.9555	0.8902	0.6001	0.9046	
	Healthy	1852	137	7490	188	0.9311	0.9755	0.9664	0.9193	0.9906	
		Overall (weighted)				0.8750	0.9601	0.9451	0.8706	0.9709	

Table 21: Comparison of results obtained using Inceptionv3 for four-class classification

Method	Class	TP	FN	TN	FP	SEN	SPE	ACC	F-1	AUC	Overall ACC
Original	Covid-19	3341	64	6228	34	0.9812	0.9946	0.9899	0.9855	0.9993	0.8752
	Bacterial	2317	463	6377	510	0.8335	0.9259	0.8993	0.8265	0.9578	
	Viral	929	564	7698	476	0.6222	0.9418	0.8924	0.6411	0.9148	
	Healthy	1874	115	7492	186	0.9422	0.9758	0.9689	0.9257	0.9929	
		Overall (weighted)				0.8752	0.9628	0.9445	0.8743	0.9730	
LBP	Covid-19	3377	28	6208	54	0.9918	0.9914	0.9915	0.9880	0.9993	0.8621
	Bacterial	2216	564	6424	463	0.7971	0.9328	0.8938	0.8119	0.9557	
	Viral	961	532	7547	627	0.6437	0.9233	0.8801	0.6238	0.9121	
	Healthy	1780	209	7489	189	0.8949	0.9754	0.9588	0.8994	0.9890	
		Overall (weighted)				0.8621	0.9607	0.9395	0.8629	0.9712	
LE	Covid-19	3362	43	6193	69	0.9874	0.9890	0.9884	0.9836	0.9986	0.8566
	Bacterial	2282	498	6294	593	0.8209	0.9139	0.8871	0.8071	0.9497	
	Viral	793	700	7728	446	0.5311	0.9454	0.8815	0.5805	0.8978	
	Healthy	1844	145	7400	278	0.9271	0.9638	0.9562	0.8971	0.9897	
		Overall (weighted)				0.8566	0.9555	0.9362	0.8528	0.9672	
Pipeline (Original-LBP)	Covid-19	3391	14	6229	33	0.9959	0.9947	0.9951	0.9931	0.9997	0.8885
	Bacterial	2344	436	6444	443	0.8432	0.9357	0.9091	0.8421	0.9663	
	Viral	978	515	7725	449	0.6551	0.9451	0.9003	0.6699	0.9317	
	Healthy	1876	113	7525	153	0.9432	0.9801	0.9725	0.9338	0.9945	
		Overall (weighted)				0.8885	0.9671	0.9511	0.8876	0.9785	
Pipeline (Original-LE)	Covid-19	3384	21	6229	33	0.9938	0.9947	0.9944	0.9921	0.9997	0.8835
	Bacterial	2380	400	6360	527	0.8561	0.9235	0.9041	0.8370	0.9622	
	Viral	870	623	7810	364	0.5827	0.9555	0.8979	0.6381	0.9241	
	Healthy	1907	82	7476	202	0.9588	0.9737	0.9706	0.9307	0.9946	
		Overall (weighted)				0.8835	0.9638	0.9486	0.8802	0.9762	