CoviNet: Automated COVID-19 Detection from X-rays using Deep Learning Techniques

Samira Lafraxo
LabSIV, Department of Computer Science
Faculty of Science, Ibn Zohr University
BP 8106, 80000 Agadir, Morocco
samira.lafraxo@gmail.com

Mohamed El Ansari
LabSIV, Department of Computer Science
Faculty of Science, Ibn Zohr University
BP 8106, 80000 Agadir, Morocco
melansari@gmail.com

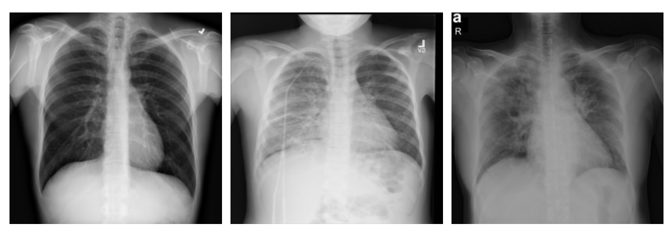
2020 6th IEEE Congress on Information Science and Technology (CiSt) | 978-1-7281-6646-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CIST49399.2021.9357250

***Abstract*—The novel Coronavirus (COVID19) is an infectious epidemic declared in March 2020 as Pandemic. Because of its easy and rapid transmission, Coronavirus has caused thousands of deaths around the world. Thus, developing new systems for accurate and fast COVID19 detection is becoming crucial. X-ray imaging is used by radiology doctors for the diagnosis of coronavirus. However, this process requires considerable time. Therefore, artificial intelligence systems can help in reducing pressure on health care systems. In this paper, we propose CoviNet a deep learning network to automatically detect COVID19 presence in chest X-ray images. The suggested architecture is based on an adaptive median filter, histogram equalization, and a convolutional neural network. It is trained end-to-end on a publicly available dataset. Our model achieved an accuracy of 98.62% for binary classification and 95.77% for multi-class classification. As the early diagnosis may limit the spread of the virus, this framework can be used to assist radiologists in the initial diagnosis of COVID19.**

***Index Terms*—coronavirus, covid19, deep learning, convolutional neural network, chest X-ray, adaptive median filter**

I. Introduction

Since December 2019, coronavirus disease (also named COVID19) has been rapidly spread, causing panic all over the world. By June 12, more than 7.5 million people were worldwide infected and 425.100 of those cases induce to death [1]. This infection is transmitted easily from person to person over sneezing, coughing or respiratory droplets. The virus generally comes with cough, fever and shortness of breath, and it can lead to pneumonia, multi-organ failure and death [2]. Considering the fact of the absence vaccine or therapeutic treatment for the novel coronavirus, early diagnosis is imperative. As they contain useful information for diagnostic, chest radiological imaging including Chest X-ray (CXR) and computed tomography (CT), play an important role in early detection and treatment of this disease [3]. However, the fast increase in patients at the time of the pandemic, make it hard for doctors to complete the diagnoses process in a limited time. In this context, artificial intelligence can present an accurate, fast and low-cost tool for COVID19 diagnostic. Recently, deep learning as a subset of machine learning in artificial intelligence, approved its superiority over classical AI approaches (handcrafted methods), in different medical imaging tasks. It have been used in many problems such as classification, identification and segmentation. Hence, deep learning is becoming an evident choice for CXRs diagnosis. The main objective of this paper is to provide an efficient framework for the automatic diagnosis of COVID19. For this aim, we propose a deep learning model based on an end-to-end convolutional neural network (CNN) [4] architecture without the use of any manual feature extraction method. In our study the preprocessing step is very important to work on images before feeding them to the CNN model. Even if our goal is to detect the presence of COVID19 in the images, we worked on two scenarios, the first one is a binary classification (covid19 vs. normal) and the second one is a multi-class classification (covid19 vs. effusion vs. normal). Examples of images from each class are presented in Fig. 1. In our study, we tried to improve the overall accuracy, to use low computational cost and a lower number of parameters. The rest of this paper will be organized as follows: Section II discusses the state-of-the-art. The methodology is described in Section III. Section IV demonstrates the results . Finally, Section V concludes the paper.



(a) (b) (c)

Fig. 1. Examples of Chest X-Rays. (a) Normal. (b) Effusion affected lungs. (c) COVID-19 positive.

II. Related work

The use of machine learning approaches for automatic diagnostic in the medical area have recently become an important tool for clinicians [5]–[8]. Various recent studies based on deep learning, have been widely applied on chest X-rays to detect the novel coronavirus. In this section, we will describe some notable works presented in the literature concerning this topic. In [9], Khan et al. have implemented the CoroNet, a convolutional neural network based on Xception (Extreme Inception) which contains 71 layers pretrained on ImagNet dataset. The proposed model obtained an accuracy of 87% for COVID19 detection. Sethy and Behra [10] proved in their paper that ResNet50 model combined with SVM classifier provide better performance in identifying COVID19. [11] presented a comparison between seven various well-known deep learning networks architectures. They used a small dataset of 50 images in their experiments and they reported that VGG19 and DensNet201 were the best performing. In [12], Authors proposed a new architecture of CNN (COVID-Net), which was designed for CXR images classification into pneumonia, normal and COVID19. The model was validated on a large dataset containing 13.800 images. It achieved an overall accuracy of 92.4%. Wang and Wong [12] suggested a deep model for COVID19 recognition (COVID-Net), which achieved 92.4% accuracy in detecting normal, non-COVID pneumonia, and COVID-19 classes. Ioannis et al. [13] implemented a deep learning model using 224 labled COVID-19 images. Their model obtained 98.75% and 93.48% accuracy rates for two and three classes, respectively.

III. Methods

We propose a novel deep learning framework designed for automatically detecting of COVID19 virus in 2D chest X-ray images. Our system is also dedicated to distinguishing between COVID19, effusion, and normal cases. The network is based on Convolutional Neural Networks which receives already processed and filtered images (using Adaptive Median Filter and Histogram Equalization). We evaluate the performance of the model and trained it again changing architecture and hyper-parameters, until getting the best possible accuracy score. Fig. 2 illustrates the overall work-flow of the proposed CoviNet.

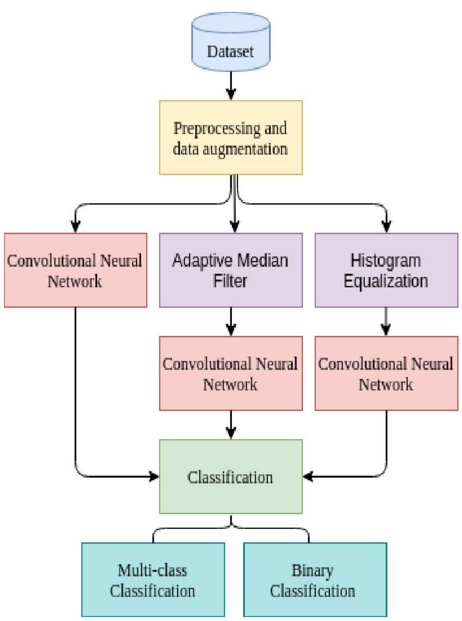


Fig. 2. The overall framework of the proposed CoviNet.

A. Preprocessing

The preprocessing step involve many operations including :

- Labeling the dataset images : 0 for normal, 1 for covid19 virus and 2 for effusion disease.
- Resizing all of the images to one dimension size 256x256.
- Saving the images in one dimension scheme (gray-scale).
- Solving the problem of imbalances of dataset classes by data augmentation techniques (mirroring, vertical flip and random rotation). Statistics of the new obtained dataset are shown in Table I.

TABLE I
Dataset distribution.

Dataset	Normal	Effusion	COVID19
Original dataset	1000	107	308
Balanced dataset	1000	1000	1000

B. Adaptive Median Filter

The adaptive median filter [14] is a non-linear digital filtering technique, often applied on images to remove their noise and make images more clear and easier to distinguish. The adaptive median filter (AMF) changes the noise area window to detect noised pixels in the neighborhood. If noise points are detected, they are replaced by median pixels value and if not, the original pixel value is kept. The advantage of an AMF is the possibility of preserving details when smoothing non-pulsed noise which is not done with traditional median filters.

C. Network Architecture

CoviNet is based on a CNN architecture that includes different layers and consists of two stages: automatic feature extraction and a classification part. As depicted in Fig. 3, the proposed CNN has four convolutional layers, with a filter size of 3x3, the number of filters in the three convolutional layers is 32. In the fourth layer, 64 filters are used. After each convolution layer, Relu is used as an activation function. To reduce the computational complexity and the spacial size of the obtained feature maps, max-pooling layers are used after the two convolutional layers with a window size of 2x2 and 2 pixels stride. The pooled output is then flattened and fed to the first fully connected layer which has 128 neurons. Dropout [15] is used as a regularization technique to avoid overfitting problems.

IV. Experimental results

In this section, we describe the main dataset used in this study. We also present the experimental setup and results for both binary and multi-class classification, we conclude the section with a comparison with the state-of-the-art methods.

A. Dataset

In this study, two different publicly available sub-databases were used to build one dataset :

1) COVID-19 Radiography database : Chowdhury et al. [16] have created a public dataset on Kaggle [17] by collecting 219 COVID-19 positive chest X-ray images.

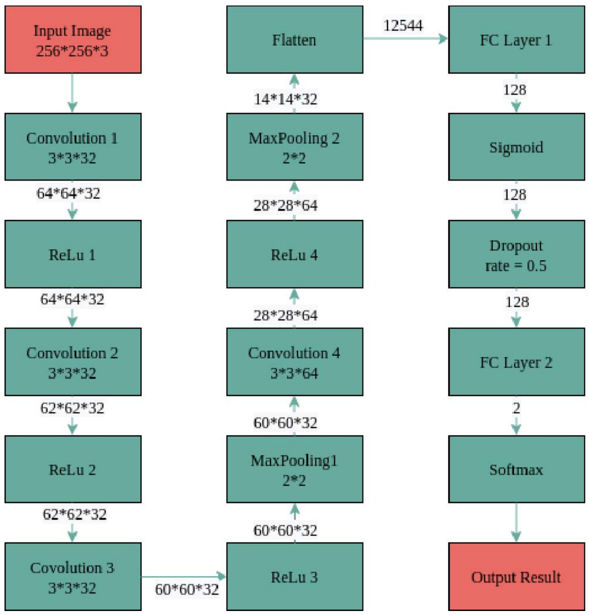


Fig. 3. The CNN model architecture.

All the images are in PNG file format with resolution 1024x1024.

2) CXR data : This dataset was created by Murali Kummitha [18], it contains two categories , the first is effusion with 107 CXRs and the second is normal with a total number of 1000 images. All the images in PNG format with different resolution size.

B. Experimental Environment

Our model has been developed using python and keras library with tensorflow backend on an Intel (R) core (TM) i7 Gen GHz processor. Experiments were conducted using Graphical Processing Unit (GPU) NVIDIA GEFORCE GTX 1050 Ti and RAM with 8 GB and 4 GB respectively. To train the models, we set the number of epochs, batch size, and learning rate to 120, 400, 0.001 respectively. We used cross-entropy as a loss function (binary, categorical) and Adam as an optimizer for the cross-entropy function.

C. Evaluation Process

The performance of our trained model was evaluated by calculating accuracy, sensitivity, specificity, precision, and F1-score. The TP is the number of true-positive patterns, FP represents the number of false-positive patterns, TN is the number of true-negative patterns and FN is the number of false-negative patterns.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$F1 - score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{5}$$

Sensitivity and specificity give the portion of positives and negatives that well classified, respectively.

D. Model Performance

We conducted experiments to classify and detect COVID19 using chest X-ray radiographs in two different scenarios. We, first have trained the CoviNet deep learning network to classify CXRs into two classes covid19 and normal. In the second scenario, the model was trained to distinguish between three classes covid19, normal, and effusion.

1) Binary classification results: The results are illustrated in Table II. From this table, we observe that feeding images to the CNN model after applying histogram equalization (H-CNN) gives better results for almost most of the metrics. We note that the best model has achieved an average accuracy of 98.62% and obtained sensitivity, specificity, and F1-score of 98.52%, 98.72%, and 98.65%. Fig. 4, Fig. 5, Fig. 6 represent respectively accuracy and loss curve, confusion matrix and ROC curve for the ConviNet model concerning binary classification.

2) Multi-class classification: The triple classification performance of the two models has been evaluated and the accuracy, precision, recall metrics were calculated. The results are depicted in Table III, where we can clearly see that the model where we applied histogram equalization on the images before feeding them to the convolutional neural network achieved better results than the two other models in terms of different metrics, achieving an average accuracy of 95.77% in just 940 seconds. To better validate the efficiency of this model, we plot the accuracy curve in Fig. 7. The confusion matrix is also shown in Fig. 8 enabling us to have an idea about the accuracy of each class. Based on this figure, we can notice that the COVID19 cases are predicted to have much higher probabilities than the nonCOVID cases, which is good as it allows us to detect the positive cases precisely.

3) Comparison with benchmarks: In this subsection, we compare the performance of the proposed approach presented in this work with the state of the art methods based on accuracy. Table IV depicts the results of this comparison, where to types of images are used to detect COVID19: CT and CXR images. As it can be seen, our method shows better results in both multi-class and binary classification. We can explain this by the importance of choosing the appropriate hyper-parameters for the problem and also well preprocessing the images before feeding them to the network.

V. Conclusion

In this study, we have presented a deep learning-based model to classify CXR images into COVID19, normal, or effusion. Our model is entirely automated with an end-to-end architecture that does not use any handcrafted method.

TABLE II
PERFORMANCE OF THE TWO APPROACHES UNDER THE SAME DATASET FOR THE BINARY CLASSIFICATION.

Model	Acc %	AUC %	Spe %	Sen %	Pre %	F1-Score %	Time (s)
CNN	96.0	0.99	93.53	98.49	93.77	96.07	223
AMF-CNN	95.5	0.99	95.52	95.47	95.47	95.47	149
H-CNN	98.62	0.99	98.72	98.52	98.77	98.65	705

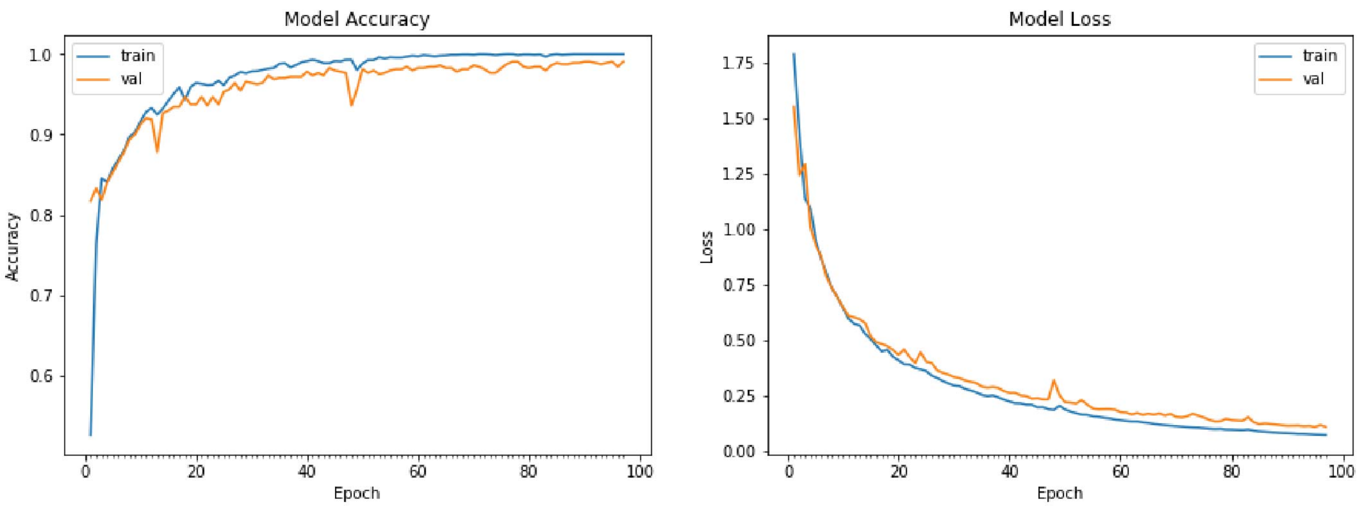


Fig. 4. Accuracy and loss model curves for the binary classification.

TABLE III
PERFORMANCE OF THE TWO APPROACHES FOR THE SAME DATASET (3 CLASSES).

Model	Acc %	Precison %	Recall %	Time (s)
CNN	93.05	89.88	89.66	537
AMF-CNN	95.47	93.39	93.16	1176
H-CNN	95.77	93.69	93.66	940

CoviNet system is designed to perform both, binary and multi-class classification with an accuracy of 98.62% and 95.77% respectively. As a future work, we plan to use larger clinical datasets to better evaluate the effectiveness of our model, and also to work on the segmentation stage.

REFERENCES

[1] [Online]. Available: https://www.worldometers.info/coronavirus/

[2] C.-C. Lai, Y. H. Liu, C.-Y. Wang, Y.-H. Wang, S.-C. Hsueh, M.-Y. Yen, W.-C. Ko, and P.-R. Hsueh, "Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (sarscov-2): facts and myths," *Journal of Microbiology, Immunology and Infection*, 2020.

[3] Z. Y. Zu, M. D. Jiang, P. P. Xu, W. Chen, Q. Q. Ni, G. M. Lu, and L. J. Zhang, "Coronavirus disease 2019 (covid-19): a perspective from china," *Radiology*, p. 200490, 2020.

[4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[5] S. Charfi, M. El Ansari, and I. Balasingham, "Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy images," *IET Image Processing*, vol. 13, no. 6, pp. 1023–1030, 2019.

[6] S. Charfi and M. El Ansari, "Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images," *Multimedia Tools and Applications*, vol. 77, no. 3, pp. 4047–4064, 2018.

[7] M. Souaidi, A. A. Abdelouahed, and M. El Ansari, "Multi-scale completed local binary patterns for ulcer detection in wireless capsule endoscopy images," *Multimedia Tools and Applications*, vol. 78, no. 10, pp. 13 091–13 108, 2019.

[8] M. Souaidi and M. El Ansari, "Multi-scale analysis of ulcer disease detection from wce images," *IET Image Processing*, vol. 13, no. 12, pp. 2233–2244, 2019.

[9] 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*, p. 105581, 2020.

[10] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid-19) based on deep features," *Preprints*, vol. 2020030300, p. 2020, 2020.

[11] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[12] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images," *arXiv*, pp. arXiv–2003, 2020.

[13] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.

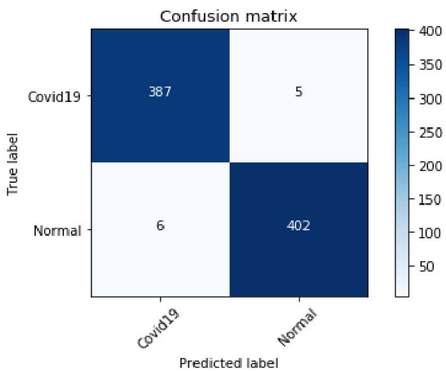
[14] H. Ibrahim, N. S. P. Kong, and T. F. Ng, "Simple adaptive median filter for the removal of impulse noise from highly corrupted images," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1920–1927, 2008.

[15] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

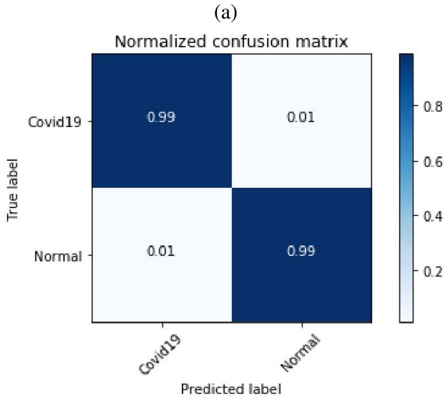
[16] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *arXiv preprint arXiv:2003.13145*, 2020.

TABLE IV
COMPARISON OF THE PROPOSED COVINET WITH BENCHMARKS.

Authors	**Method used**	**Type of images**	**Number of classes**	**Accuracy %**
Wang et al. [19]	M-Inception	Chest CT	2	82.9
Song et al. [20]	DRE-Net	Chest CT	2	86.0
Xu et al. [21]	ResNet + Location Attention	Chest CT	3	86.7
Hemdan et al. [11]	COVIDX-Net	Chest X-ray	2	90.0
Zheng et al. [22]	UNet + 3D deep network	Chest CT	2	90.8
Wang and Wong [12]	COVID-Net	Chest X-ray	3	92.4
Ioannis et al. [13]	VGG-19	Chest X-ray	3	93.48
Sethy and Behra [10]	ResNet50 + SVM	Chest X-ray	2	95.38
Our method	CoviNet	Chest X-ray	2	98.62
Our method	CoviNet	Chest X-ray	3	95.77



(a)



(b)

Fig. 5. Confusion Matrix and normalized confusion matrix for the binary classification.

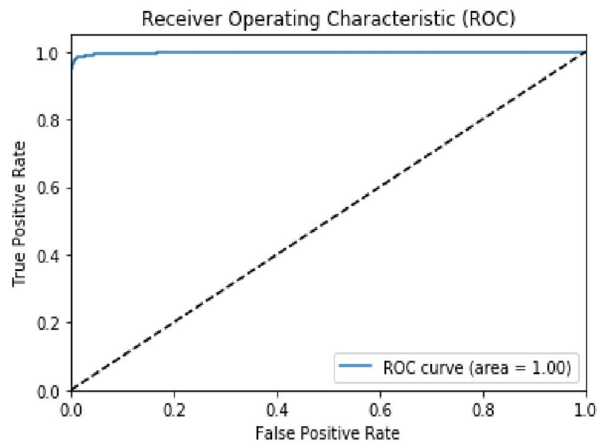


Fig. 6. ROC curve for the binary classification.

[17] [Online]. Available: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
[18] [Online]. Available: https://www.kaggle.com/murali0861/cxrdata
[19] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng *et al.*, “A deep learning algorithm using ct images to screen for corona virus disease (covid-19),” *MedRxiv*, 2020.
[20] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, H. Zhao, Y. Jie, R. Wang *et al.*, “Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images,” *medRxiv*, 2020.
[21] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Y. Chen, J. Su, G. Lang *et al.*, “Deep learning system to screen coronavirus disease 2019 pneumonia. arxiv 2020,” *arXiv preprint arXiv:2002.09334*.
[22] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, “Deep learning-based detection for covid-19 from chest ct using weak label,” *medRxiv*, 2020.

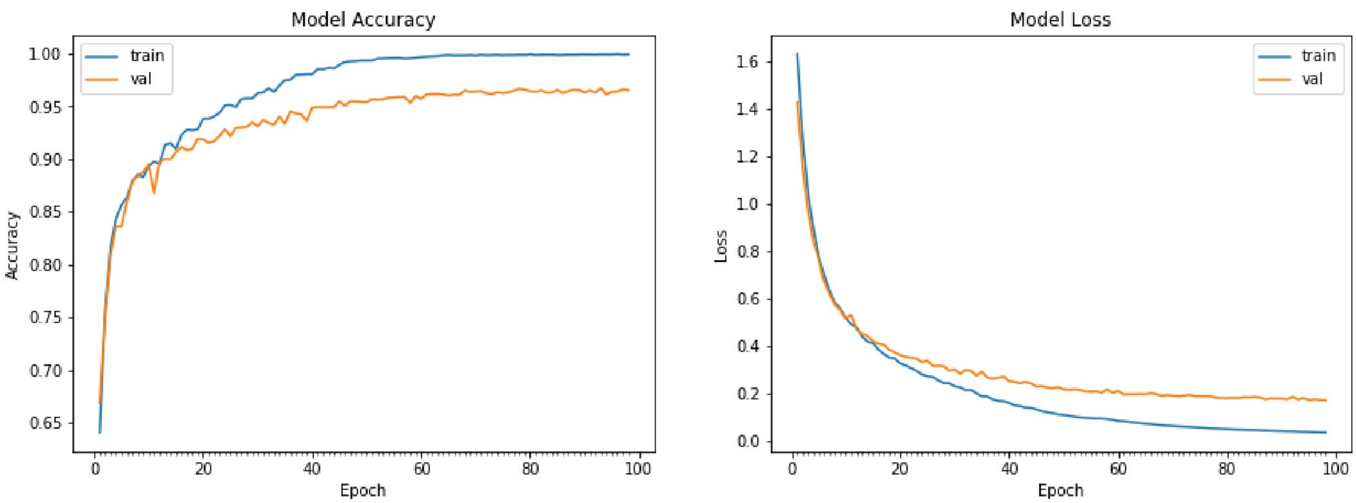
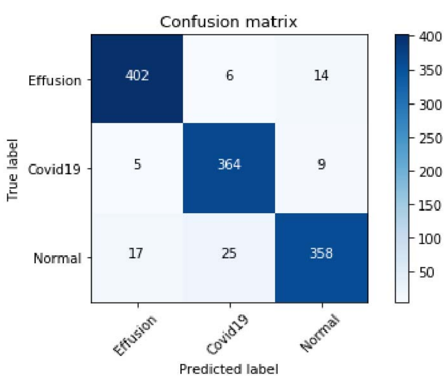
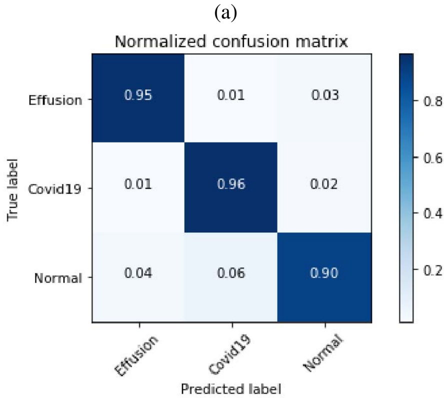


Fig. 7. Accuracy and loss model curves for the triple classification.



(a)



(b)

Fig. 8. Confusion matrix and normalized confusion matrix for the triple classification.