2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI) | 978-1-6654-1322-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/PRAI53619.2021.9551043

Enhancing Low Quality in Radiograph Datasets Using Wavelet Transform Convolutional Neural Network and Generative Adversarial Network for COVID-19 Identification

Grace Ugochi Nneji
School of Information and Software Engineering
University of Electronic Science and Technology of China
Sichuan, China
ugochinneji@std.uestc.edu.cn

Jingye Cai
School of Information and Software Engineering
University of Electronic Science and Technology of China
Sichuan, China
jycai@uestc.edu.cn

Deng Jianhua
School of Information and Software Engineering
University of Electronic Science and Technology of China
Sichuan, China
jianhua.deng@uestc.edu.cn

Happy Nkanta Monday
School of Computer Science and Engineering
University of Electronic Science and Technology of China
Sichuan, China
mh.nkanta@std.uestc.edu.cn

Ijeoma Amuche Chikwendu
School of Information and Communication Engineering
University of Electronic Science and Technology of China
Sichuan, China
ijeomaamuche@std.uestc.edu.cn

Ariyo Oluwasanmi
School of Information and Software Engineering
University of Electronic Science and Technology of China
Sichuan, China
ariyo@uestc.edu.cn

Edidiong Christopher James
School of Information and Software Engineering
University of Electronic Science and Technology of China
Sichuan, China
edianajames@yahoo.com

Goodness Temofe Mgbejime
School of Computer Science and Engineering
University of Electronic Science and Technology of China
Sichuan, China
temofeerics@gmail.com

***Abstract*—The coronavirus disease of 2019 (COVID-19) pandemic has caused a global public health epidemic since there is no 100% vaccine to cure or prevent the further spread of the virus. With the ever-increasing number of new infections, creating automated methods for COVID-19 identification of Chest X-ray images is critical to aiding clinical diagnosis and reducing the time-consumption for image interpretation. This paper proposes a novel joint framework for accurate COVID-19 identification by integrating an enhanced super-resolution generative adversarial network with a noise reduction filter bank of wavelet transform convolutional neural network on both Chest X-ray and Chest Tomography images for COVID-19 identification. The super-resolution utilized in this study is to enhance the image quality while the wavelet transform Convolutional Neural Network architecture is used to accurately identify COVID-19. Our proposed architecture is very robust to noise and vanishing gradient problem. We used public domain datasets of Chest x-ray images and Chest Tomography to train and check the performance of our COVID-19 identification task. This experiment shows that our system is consistently efficient by accuracy of 0.988, sensitivity of 0.994, and specificity of 0.987, AUC of 0.99, F1-score of 0.982 and 0.989 for precision using the Chest X-ray dataset while for Chest Tomography dataset, an accuracy of 0.978, sensitivity of 0.981, and specificity of 0.979, AUC of 0.985, F1-score of 0.961 and precision of 0.980. These performances have also outweighed other established state-of-the-art learning methods.**

Keywords—chest tomography, chest x-ray, COVID-19, deep learning, identification, super-resolution, wavelet transform

I. INTRODUCTION

Ever since 2019, Coronavirus disease(COVID-19) has become the trending disease causing the death of 3.74 million, 173 million confirmed cases of people within 180 territories as at June 9th, 2021 [1], [2]. Yet, there has not been a 100% vaccine cure for this disease. Therefore, there is a vital need for an automatic system to reduce the further spread of this epidemic as the system detects if an individual is infected with the COVID-19 or not.

There have been different medical trial done in detecting if someone has the COVID-19 or not, one of which is the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) [3] test. Unfortunately, the demerits of this method include the long period of time it takes for the expected results and the low sensitivity of sample results becoming obsolete for medical acceptance. With this downside of RTPCR, it has further the spread of the virus, thereby causing global panic. Meanwhile, other methods have been investigated such as the Chest X-Ray (CXR) and Chest Tomography (CT) for COVID-19 diagnosis and they have performed greatly [4], [5], [6].

Many research papers have reported on the automatic detection of COVID-19 from Chest X-ray and CT using deep neural network. With the effort of some research teams in publicly providing dataset, COVID-19 identification has been carried out mostly in utilizing deep neural network, hence achieving good results [7], [8], [9], [10], [11], [12]. Chen et al. [13] presented a technique of detecting COVID-19 based on

deep neural network from retrospectively CT images received and preprocessed from a hospital in Wuhan. The result reported accuracy of 0.96. Zheng et al. [14] reported a method based on weakly-supervised technique using UNet deep neural network for correctly examining COVID-19 with 530 CT exams and got an accuracy of 0.90. Another medical-based AI framework was implemented by Jin et al [15] to accurately screen COVID-19 from other pneumonia using CT exams. This model achieved an accuracy of 0.92. Tang et al. [16] adopted an automatic learning model of measuring the severity of COVID-19 from CT exams which belongs to 176 patients. This reported accuracy of 0.87.

In our study, we incorporated the technique of super-resolution to handle the quality enhancement of radiograph datasets while the wavelet transform eliminates the noisy effect on the images and execute feature extraction for the identification of COVID-19. The report of our loss, accuracy, ROC AUC, precision, F1-score, sensitivity and specificity are given in the subsequent section. The rest of this paper will be organized as: material and methodology of our work is in Section II, experimental results will be given in Section III. Then, our conclusion will be reported in Section IV.

II. Material and Methodology

This report comprise of three phases. The first phase is the collection of our datasets which are from three sources. Secondly, the improvement of the image quality using the enhanced super-resolution generative adversarial network and finally, we applied our proposed wavelet transform convolutional neural network architecture to extract features of radiographs while eliminating the noisy effect for COVID-19 identification.

A. Dataset

For our experiment, the CXR images of COVID-19 were gotten from the GitHub repository of Cohen et al. [4] whereas the healthy (normal) dataset is collected from the website of the Radiological Society of North America (RSNA) [2]. More so, for the CT COVID-19 and non-COVID-19 datasets, they were acquired from Yang et al. [17]. Table I and II illustrate the collection and the utilization of the dataset while Table III reports the augmentation process and split ratio of our dataset for both Chest Tomography and Chest X-ray. In Table I, we used 162 images each from the two CXR public available dataset, amounting to a total of 324 images of both COVID-19 and normal images. Then for our split ratio as seen in Table III, we utilized 80% for the training set and 20% for the test set. Furthermore, for the CT dataset gotten from Yang et al.[17], we used 324 images each for COVID-19 and normal images, bringing to a total of 648 images. The dataset split is 80% for the training and 20% for the testing sets.

More so, in avoiding overfitting and enabling our dataset to train well, we introduced another aspect of increasing our dataset by employing an image augmentation technique for better efficiency. Our common utilized augmentation techniques were reflection, rescaling, zoom and rotation. Figure 1 reports the image augmentation result in our images. Our augmentation technique increased CXR dataset to 640 images as well as that of CT Images as seen in Table III. Meanwhile, augmentations of data were carefully conducted so as not to produce unrealistic images, thereby not confusing the wavelet transform CNN architecture.

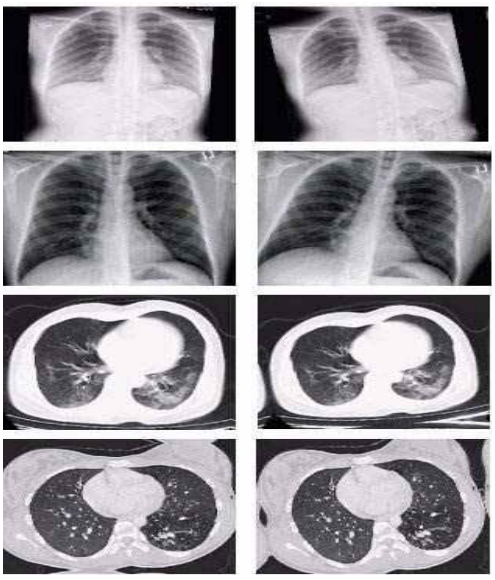

Fig. 1. CT and CXR results after Image augmentation.

TABLE I. Description of Chest X-Ray (CXR) Dataset

Dataset	COVID-19	Other pneumonia	Healthy
COVID-CXR [4]	162	42	234
RSNA [2]	0	6012	8851
Our Utilization	**162**	**-**	**162**

TABLE II. Description of Chest Tomography (CT) Dataset

Dataset	COVID-19	Non-COVID
Yang et al. [17]	349	397
Our Utilization	**324**	**324**

TABLE III. Description of Our Dataset and the Split Ratio

Dataset	Total Images	Split Ratio (80% : 20%)	
		Train	Test
CT_COVID19	324	520	128
CT_Normal	324		
CXR_COVID19	162	260	64
CXR_Normal	162		
CT_COVID19 with augmentation	640	1024	256
CT_Normal with augmentation	640		
CXR_COVID19 with augmentation	640	1024	256
CXR_Normal with augmentation	640		

B. The Proposed Joint Framwork of Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) with Wavelet Transform Convolutional Neural Network

The proposed network in Figure 3 illustrates a joint model of an enhanced super-resolution improving the image quality and the wavelet transform for minimizing noise and for feature extraction for COVID-19 identification.

- Super-Resolution Network

  In real-life scenario, medical radiographs are characterized with low quality. Therefore, the main idea of our proposed model is to take input of low-quality images and generate them into a higher resolution for better efficiency. In producing this high-resolution from low-resolution counterpart, our joint framework enhanced super-resolution generative adversarial network with wavelet transform CNN has proven to be the best. In the stage of enhanced super-resolution generative adversarial network as seen in Figure 2, our input images were upscaled to a factor of 4. The algorithm uses two neural networks called the generator and discriminator put against each other in order to produce new and synthetic instances of data that are considered real. The network generator produces $w^{k+1} = G_k(w^k)$ images and feature maps are extracted for perceptual loss calculation before the final pass to the activation function. While improving the pixel level quality of the generated image measured by the content loss, the created images are forwarded to the discriminator network to check the distinctive features of the original images $\hat{w}^{k+1}$ and the created one w^{k+1}. For the sake of SR, the discriminatory loss is utilized to ensure the generation of high-resolution images that are more natural and realistic as compared to the original images. Then the adversarial loss checks the classifier to verify either it is real or fake. The entire process continues until there are no difference between the real and the fabricated images.

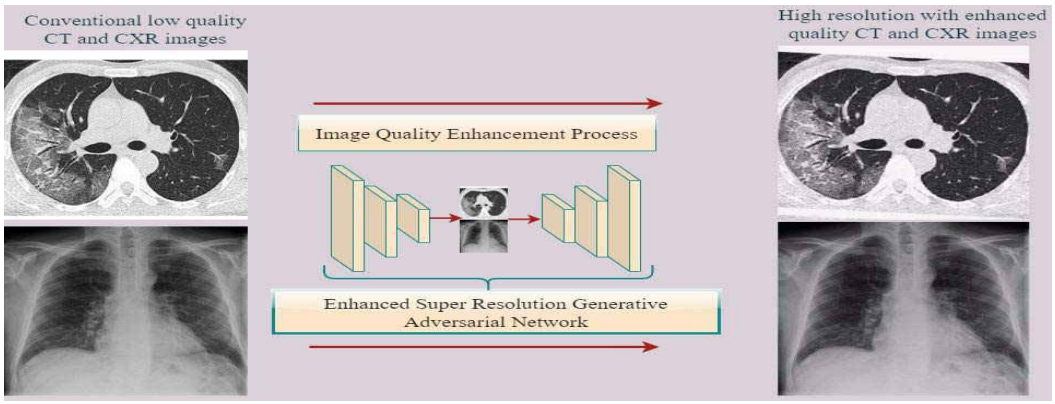

Fig. 2. Enhancement of the image quality using ESRGAN.

The calculation of the summation of the super-resolution network is presented in equation (1).

$$L_{Totalloss} = L_{Gen}\left(L_{Perceptualloss} + \mu L_G^{Ra} + \eta L1\right) + L_{Dis}^{Ra} \quad (1)$$

where $L_{Perceptualloss}$ is the perceptual loss, L_G^{Ra} is the adversarial loss and $L1$ is the content loss, L_{Dis}^{Ra} is the discriminator network loss and μ and η are the coefficient to offsets the losses respectively.

The perceptual loss expression is given in equation (2). This loss improves the feature level in order to enhance the visual quality of the image.

$$L_{Perceptualloss} = \sum_{x=1}^{F_{ij}} \sum_{y=1}^{H_{ij}} \left(\alpha_{ij}(\hat{w}^{k+1})_{xy} - \alpha_{ij}\left(G_k(w^k)\right)_{xy}\right)^2 \quad (2)$$

The content loss, L1 in equation (3) ensures the pixel-wise quality of the generated image in order to maintain close similarity with the original image.

$$\mathrm{L1} = \sum_x^W \sum_y^H \| G_k(v^k)_{xy} - (\hat{v}^{k+1})_{xy} \|_1 \quad (3)$$

The discriminator network loss is given in equation (4).

$$\mathrm{L}_{Dis}^{Ra} = -E_{\hat{v}^{k+1}}\left[log\left(D_{Ra}\left(\hat{v}^{k+1}, G_k(v^k)\right)\right)\right] - E_{G_k(v^k)}\left[log\left(1 - D_{Ra}(G_k(v^k), \hat{v}^{k+1})\right)\right] \quad (4)$$

Despite this, equation (5) illustrates the adversarial loss for the generator network.

$$\mathrm{L}_G^{Ra} = -E_{\hat{v}^{k+1}}\left[log\left(1 - D_{Ra}\left(\hat{v}^{k+1}, G_k(v^k)\right)\right)\right] - E_{G_k(v^k)}\left[log\left(D_{Ra}(G_k(v^k), \hat{v}^{k+1})\right)\right] \quad (5)$$

The enhancement of image quality can be seen in Figure 2 as compared to the original image.

- Wavelet Transform Convolutional Neural Network

  This stage comprises of 4 blocks of convolutional layers, 6 channel-wise convolutional layers of $1x1$ connecting the four-level of wavelet transform decomposition via concatenation before the last three fully connected layers as seen in Figure 3. In the first level of decomposition, the input image size of 224 x 224 in RGB channel is first decomposed via the wavelet transform to generate detail and approximate coefficients that are further aggregated and fed to the first block of convolutional layers with a kernel size of 64 on each convolutional layer. In the second level decomposition, the approximate coefficient derived from the first level decomposition is further down-sampled to generate another detail and approximate coefficients which are aggregated and concatenated channel-wise by $1x1$ convolutional layer of kernel size 64 to the second convolutional block which consists of

two convolutional layers with kernel size 128. Furthermore, in the third level decomposition, the previous approximate coefficient from the second level decomposition is further down-sampled to generate another detail and approximate coefficients which are aggregated and concatenated channel-wise by two $1x1$ convolutional layer of kernel size 64 and 128 respectively to the third convolutional block which consists of two convolutional layers with kernel size 256. In the fourth level decomposition, the previous approximate coefficient derived from the third level decomposition is further down-sampled to generate another detail and approximate coefficients which are aggregated and concatenated channel-wise by a sequence of three $1x1$ convolutional layer of kernel size 64, 256 and 256 respectively to the fourth convolutional block which consists of three convolutional layers with kernel size 256 each followed by an average pooling. Finally, the extracted features are then passed to the fifth block which comprises of three fully connected layers of which the third FC corresponds to the number of classes with a sigmoid function as a classifier.

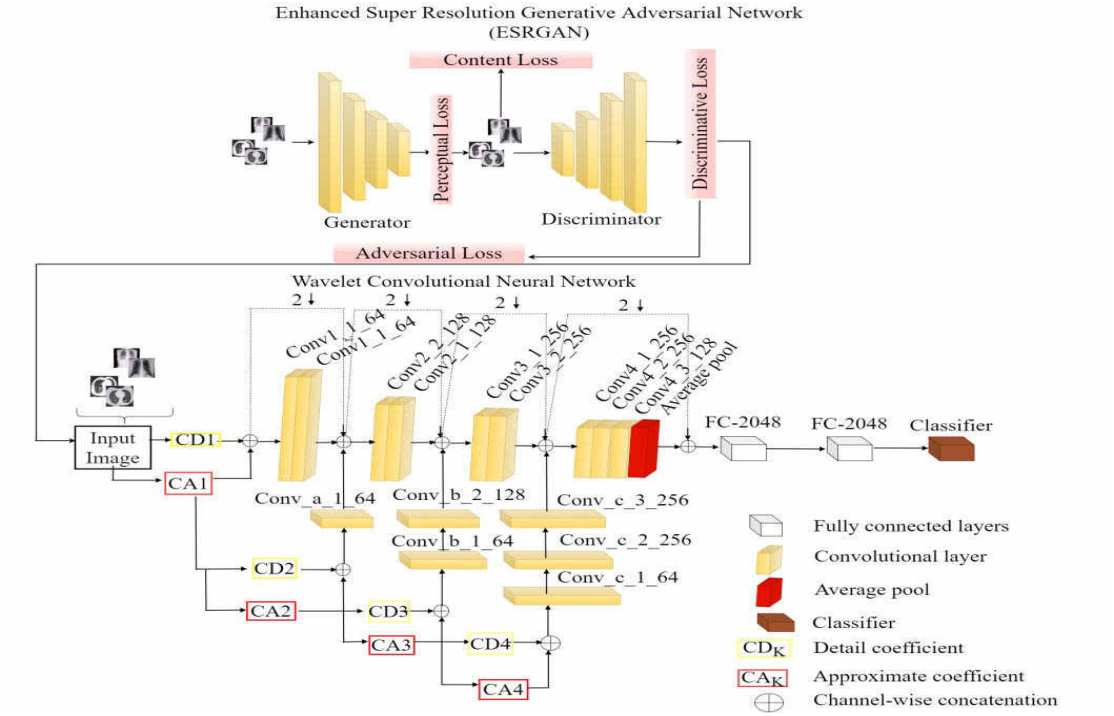

Fig. 3. The proposed joint framework of enhanced super generative adversarial network with wavelet transform convolutional neural network.

III. Experimental Results

A. Model Hyper-Parameters

In this study, we trained our model for COVID-19 identification based on the wavelet transform CNN using a data split of 80% for training and 20% for testing set with an epoch of 20. The batch size was set to 16, and Adam optimizer is used in training with a learning rate of 0.0001. With the different variations of the dataset collected from the public domains, we resized the images to 224 x 224 before being fed to the enhanced super-resolution generative adversarial network and the wavelet transform CNN. All implementations were carried out in Keras framework.

B. Evaluation Metrics

This paper uses the major metrics for evaluating the performance of the COVID-19 identification based on the CT and CXR dataset used individually. These metrics include accuracy, loss, sensitivity, precision, F1-score, AUC and specificity. For the identification of COVID-19, the equations for explaining the aforementioned metrics can be seen in equation (6), (7) and (8).

$$\text{Accuracy} = (TP + TN) / (TP - TN + FP + FN) \quad (6)$$

$$\text{Sensitivity} = TP / (TP + FN) \quad (7)$$

$$\text{Specificity} = TN / (TN + TP) \quad (8)$$

Where True Positives (TP) is the correctly identified COVID-19, False Negative is the incorrectly identified COVID-19, True Negative is the correctly identified normal instances and False Positive is the incorrectly identified normal instances.

1) Model predicted score for the CXR dataset: Initially, we used a total number of 324 images of both normal patient and COVID-19 patients. With a split ratio of 80% for the training set and 20% for the testing set, our model achieved a loss of 0.113, an accuracy (ACC) of 0.963, specificity(SPE) of 0.969 and sensitivity (SEN) of 0.958, AUC of 0.97, f1-score (F1) of 0.961 and precision (PRE) of 0.97. In improving the efficiency, we introduced the augmented CXR dataset with

a total of 1280 images for both normal and COVID-19 patients as seen in Table III in our proposed network while maintaining the hyperparameters. Thus, there was a slight improvement of 0. 025 in our accuracy and our SEN, SPE, ACC, F1 and PRE achieved the result of 0.994, 0.987, 0.99, 0.982 and 0.989 respectively with a lesser loss of 0.022 as reported in Table IV.

2) Model predicted score for the CT dataset: Checking the performance of our proposed model on CXR dataset, we also introduced the same algorithm in our CT dataset. In this case, we used a total number of 648 images for both normal and COVID-19 patients. With a split ratio of 80% for the training set and 20% for the testing set, our model achieved a loss of 0.115, an accuracy (ACC) of 0.937, specificity(SPE) of 0.954 and sensitivity (SEN) of 0.934, AUC of 0.951, f1-score (F1) of 0.955 and precision (PRE) of 0.941. To avoid overfitting and achieving a better efficiency, we incorporated our augmented CT dataset with a total of 1280 images for both normal and COVID-19 patients as seen in Table III in our proposed network while maintaining the hyperparameters. Hence, there was a slight improvement of 0. 041 in accuracy and our SEN, SPE, ACC, F1 and PRE achieved the result of 0.981, 0.979, 0.985, 0.961 and 0.98 respectively with a lesser loss of 0.049 as seen in Table IV.

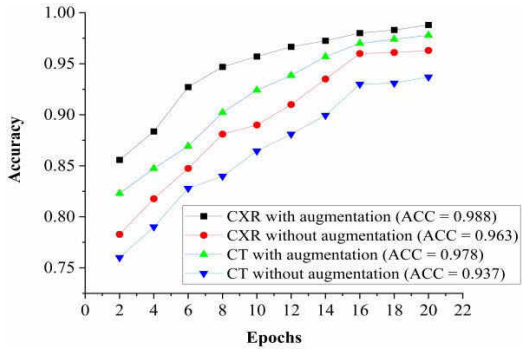

Fig. 4. The proposed model accuracy.

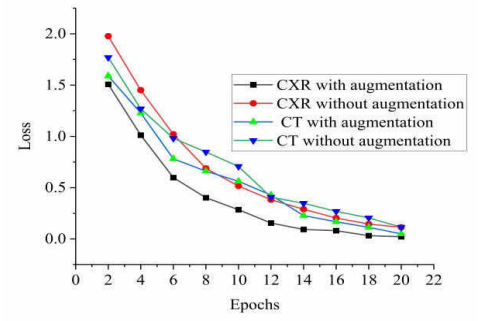

Fig. 5. The proposed model loss.

Figure 4, 5, and 6 show the performance metrics of the proposed model for the images in the test sets of the four aspects of our datasets (Chest X-ray dataset without augmentation, Chest x-ray with augmentation, Chest Tomography without augmentation and Chest Tomography with augmentation). Additionally, the black line which is the CXR dataset with augmentation in Figure 6 achieved the best ideal curve performance in our ROC AUC; the CT without augmentation had the worst performance while the others, CXR without augmentation and the CT with augmentation had moderate performances. As we can see, our model improves with sufficient dataset even when those images are more difficult to differentiate.

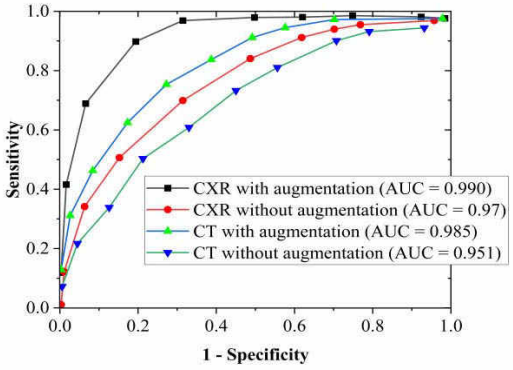

Fig. 6. The ROC AUC curves for both CXR and CT datasets.

TABLE IV. SOTA RESEARCH COMPARISON BASED ON THE IDENTIFICATION OF COVID-19

Dataset	ACC	SEN	SPE
Chen et al. [13]	0.960	0.980	0.940
Zheng et al. [14]	0.901	0.907	0.911
Jin et al. [15]	0.952	0.974	0.923
Tang et al. [16]	0.875	0.933	0.746
Our Proposed method	**0.988**	**0.994**	**0.987**

TABLE V. PERFORMANCE EVALUATION METRICS

Dataset	LOSS	ACC	SEN	SPE	AUC	F1	PRE
CT with augmentation	0.049	0.978	0.981	0.979	0.985	0.961	0.980
CT without augmentation	0.115	0.937	0.934	0.954	0.951	0.955	0.941
CXR with augmentation	**0.022**	**0.988**	**0.994**	**0.987**	**0.990**	**0.982**	**0.989**
CXR without augmentation	0.113	0.963	0.958	0.969	0.970	0.961	0.970

To further investigate the best result which is the CXR with augmentation in Table IV, the confusion matrix is presented in Figure 7. In this scenario, 126 cases were correctly confirmed as COVID-19 and 125 cases were correctly identified as normal cases while the incorrectly predicted were just 5 cases as illustrated in Figure 7. Thus, it is proven that our proposed model is effective for the identification of COVID-19. In Table V, to prove the effectiveness of our proposed network, we compared our proposed method with other state of the art (SOTA) models using the implementation task of the identification COVID-19 from healthy patients or other pneumonia either using Chest Tomography dataset or Chest X-ray dataset. Notwithstanding, our proposed model surpassed the SOTA model by at least

0.0284 in accuracy. Hence, our model is still a promising AI-based method for the identification of COVID-19.

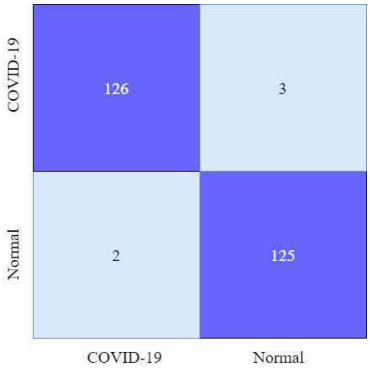

Fig. 7. The confusion matrix of our best result (cxr with augmentation).

IV. Conclusion

We reported a joint framework of an enhanced super-resolution generative adversarial network with a wavelet transform convolutional neural network on our training set. We utilized datasets from both Chest Tomography and Chest X-ray in optimizing the performance of our proposed algorithm. We investigated a detailed experimental analysis to evaluate the performance of our algorithm with and without the use of image augmentation technique using the evaluation metrics such as loss, accuracy, sensitivity and specificity, F1-score and precision. Thus, in all our result the Chest x-ray dataset which is augmented achieved a better performance of 0.988 in identification accuracy, sensitivity of 0.994 and specificity of 0.987, AUC of 0.99, F1-score of 0.982 and precision of 0.989 with a total of 1280 images on 80% is to 20% split ratio for the training and testing sets. Meanwhile, as a result of the limited number of datasets on the other experiments, our algorithm is still able to identify COVID-19 cases. Therefore, with the experimental comparison carried out, we could say that that the augmentation of images can produce better efficiency.

References

[1] "WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data." https://covid19.who.int/ (accessed May 03, 2021).

[2] "RSNA Pneumonia Detection Challenge | Kaggle." https://www.kaggle.com/c/rsna-pneumonia-detection-challenge (accessed Apr. 29, 2021).

[3] W. Wang et al., "Detection of SARS-CoV-2 in different types of clinical specimens," Jama, vol. 323, no. 18, pp. 1843–1844, 2020.

[4] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," arXiv Prepr. arXiv2006.11988, 2020.

[5] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images," Pattern Recognit. Lett., vol. 138, pp. 638–643, 2020.

[6] E. Luz et al., "Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images," Res. Biomed. Eng., pp. 1–14, 2021.

[7] 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," Comput. Methods Programs Biomed., vol. 196, p. 105581, 2020.

[8] F. Ucar and D. Korkmaz, "COVIDiagnosis-Net: deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images," Med. Hypotheses, vol. 140, p. 109761, 2020.

[9] E. Tartaglione, C. A. Barbano, C. Berzovini, M. Calandri, and M. Grangetto, "Unveiling covid-19 from chest x-ray with deep learning: a hurdles race with small data," Int. J. Environ. Res. Public Health, vol. 17, no. 18, p. 6933, 2020.

[10] J. C. Souza, J. O. B. Diniz, J. L. Ferreira, G. L. F. da Silva, A. C. Silva, and A. C. de Paiva, "An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks," Comput. Methods Programs Biomed., vol. 177, pp. 285–296, 2019.

[11] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," Phys. Eng. Sci. Med., vol. 43, no. 2, pp. 635–640, 2020.

[12] 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," Sci. Rep., vol. 10, no. 1, pp. 1–12, 2020.

[13] J. Chen et al., "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography," Sci. Rep., vol. 10, no. 1, pp. 1–11, 2020.

[14] C. Zheng et al., "Deep learning-based detection for COVID-19 from chest CT using weak label," MedRxiv, 2020.

[15] C. Jin et al., "Development and evaluation of an artificial intelligence system for COVID-19 diagnosis," Nat. Commun., vol. 11, no. 1, pp. 1–14, 2020.

[16] Z. Tang et al., "Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images," arXiv Prepr. arXiv2003.11988, 2020.

[17] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-Dataset: a CT scan dataset about COVID-19," 2020.