



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# A novel deep learning based method for COVID-19 detection from CT image

Seyyed Mohammad Javadi Moghaddam<sup>a,\*</sup>, Hossain Gholamalinejad<sup>b</sup>

<sup>a</sup> Department of Computer Engineering, Bozorgmehr University of Qaenat, Qaen, Iran

<sup>b</sup> Department of Computer Science, Bozorgmehr University of Qaenat, Qaen, Iran

## ARTICLE INFO

### Keywords:

Deep learning model  
Batch normalization  
Mish function  
COVID-19 detection method  
Disease diagnosis

## ABSTRACT

The novel Coronavirus named COVID-19 that World Health Organization (WHO) announced as a pandemic rapidly spread worldwide. Fast diagnosis of the virus infection is critical to prevent further spread of the virus, help identify the infected population, and cure the patients. Due to the increasing rate of infection and the limitations of the diagnosis kit, auxiliary detection tools are needed. Recent studies show that a deep learning model that comes up with the salient information of CT images can aid in the COVID-19 diagnosis. This study proposes a novel deep learning structure that the pooling layer of this model is a combination of pooling and the Squeeze Excitation Block (SE-block) layer. The proposed model uses Batch Normalization and Mish Function to optimize convergence time and performance of COVID-19 diagnosis. A dataset of two public hospitals was used to evaluate the proposed model. Moreover, it was compared to some different popular deep neural networks (DNN). The results expressed an accuracy of 99.03 with a recognition time of test mode of 0.069 ms in graphics processing unit (GPU). Furthermore, the best network results in classification metrics parameters and real-time applications belong to the proposed model.

## 1. Introduction

In the year 2020, a new coronavirus, called COVID-19, was marked as the cause of a pandemic disease by the WHO that first spread in China [1]. COVID-19 is a highly contagious Severe Acute Respiratory Syndrome (SARS). The number of confirmed coronavirus cases by the 27th of October 2020 was 43,341,451, of which 29,300,000 were recovered, and 1,157,509 of the infected died [1]. The standard method to detect COVID-19 cases is the Reverse Transcription-Polymerase Chain Reaction (RT-PCR). However, this detection method has some limitations such as shortage of kits, relentlessness, manual methodology, and is time-consuming. Moreover, the positive rate of the RT-PCR test is only 63% [2].

Recently, many researchers indicate that the abnormalities in the chest radiography images of patients and the computerized tomography (CT) scan are different diagnostic approaches for COVID-19 detection [3–6]. These studies show that visual symptoms in the lungs are different in patients. CT images have some advantages, such as rapid triage of suspected COVID-19 patients, lower risk of transmission, and high availability. Some researchers try to develop a method for COVID-19 detection using CT images [5,6].

Deep learning-based methods can extract the features of images that

are not obvious in the original image [7]. Many studies have applied a Convolution Neural Network (CNN) to detect COVID-19 [8–10]. Ohata [11] et al. combined the CNNs with some techniques based on machine learning, such as Bayes. They showed the proposed architecture using a support vector machine (SVM) and a linear kernel with an F1-score of 98.5%. Researchers [12] suggested a method using the CNN model and class decomposition. Abraham [13] proposed a model using a multi-CNN and correlation feature selection method. They achieved an accuracy of 97.44%. Several studies have been suggested inserting a pre-processing phase using the Visual Geometry Group (VGG) model to modality the images [14–16]. The precision of these methods is up to 86%. Azemin et al. [17] presented an approach based on ResNet-101 CNN architecture. The accuracy of the method is 77.3%. The model of Mangal [18] includes a pre-train layer with a 121-layer dense convolutional network and a full-connected layer with a 90.5% accuracy.

Many approaches use transfer learning with CNN architecture [19–21]. Loey introduced a method based on Generative Adversarial Network (GAN) with deep transfer learning. He used three deep transfer techniques and achieved 85.2% in testing accuracy [22].

Elaziz et al. [23] introduced a parallel framework for automatic COVID-19 diagnosis. This model uses a fraction multi-channel for feature selection.

\* Corresponding author.

E-mail address: [smjavadim@buqaen.ac.ir](mailto:smjavadim@buqaen.ac.ir) (S. Javadi Moghaddam).

<https://doi.org/10.1016/j.bspc.2021.102987>

Received 8 December 2020; Received in revised form 13 May 2021; Accepted 17 July 2021

Available online 21 July 2021

1746-8094/© 2021 Elsevier Ltd. All rights reserved.

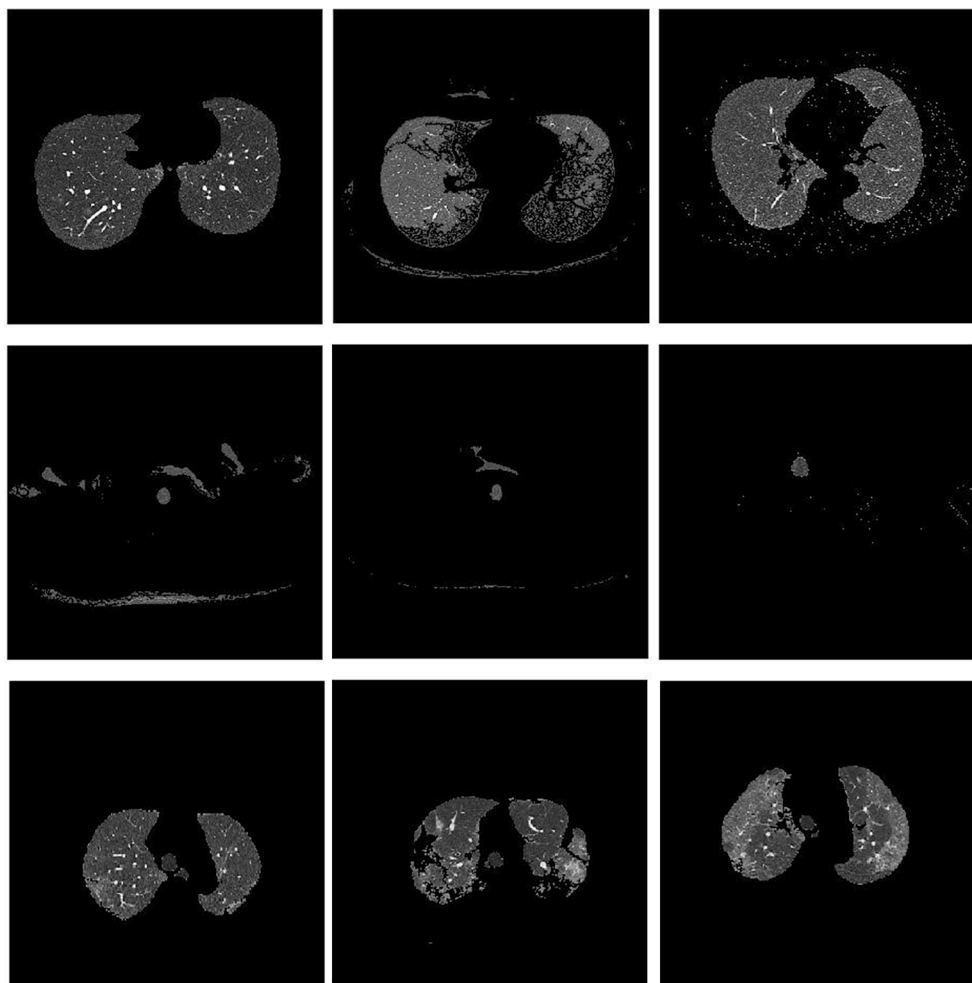


Fig. 1. A Sample image from covid-19 CT dataset.

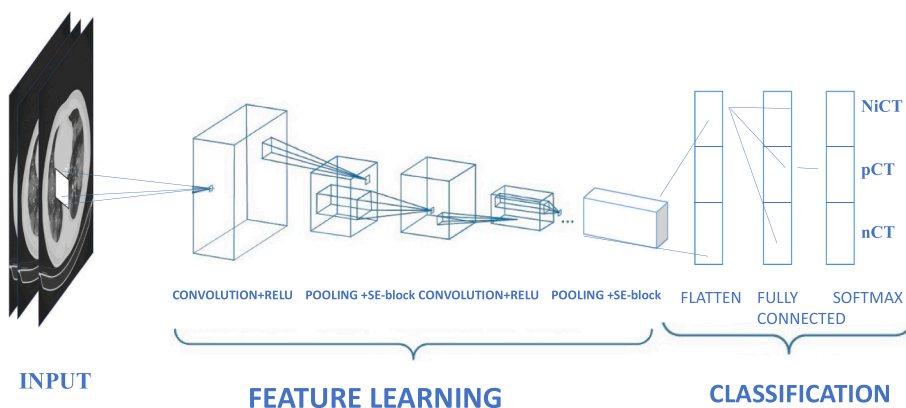


Fig. 2. Proposed model architecture.

Some researchers proposed COVID-19 detection methods for multi-class classification [24]. Khan [25] suggested a model using the deep CNN structure to COVID-19 infection diagnosis. This method uses an architecture based on Xception. The classification accuracy of the approach is 95% for the 3-class model. Karim [26] proposed to use human explanations and a deep learning model for prediction.

Yoo et al. [27] have combined three binary decision trees with a CNN architecture based on the PyTorch frame. The maximum accuracy was 98% for the first decision tree.

In common deep learning classification structures, the feature extraction process is performed only in the convolution layer. The polishing layer only plays the role of minimizing the dimensions of the feature maps and the network. This paper proposed a new pooling layer that performs reducing the network dimensions and feature extraction simultaneously. The feature extraction is performed using Haar wavelet [28]. Moreover, the model introduces a network structure based on Batch Normalization (BN) and Mish Function[29] to reduce the convergence time and achieve better performance.

**Table 1**  
Configuration of the proposed network.

layer	Output size	Kernel	Stride
Image input	256×256×3	–	–
Conv	256×256×32	5×5×3×32	1
Mish + Pool + BN + Drop out (0.5) + SEBlock	128×128×32	–	2
Conv	128×128×32	5×5×3×32	1
Mish + Pool + BN + Drop out (0.5) + SEBlock	64×64×32	–	2
Conv	64×64×64	5×5×3×64	1
Mish + Pool + BN + Drop out (0.5) + SEBlock	32×32×64	–	2
Conv	32×32×128	5×5×3×128	1
Mish + Pool + BN + Drop out (0.5) + SEBlock	16×16×128	–	2
Conv	16×16×256	5×5×3×256	1
Fully connected	65,536×256	–	–
Fully connected	256×(number of classes)	–	–
SoftMax	(number of classes) ×1	–	–

**Table 2**  
Training parameters.

Parameter	Value
Batch size	64
Epochs	50
Momentum rate	0.9
Learning rate	0.01
Weight decay	1e-3
Epsilon	1e-10
Sampler	Weighted random sampler

## 2. Material and methods

### 2.1. CT scan image dataset

This paper has used a CT image dataset [30] to evaluate the proposed model. This dataset was gathered from two Union Hospital (HUST-UH) and Liyuan Hospital (HUST-LH) [31]. The individual CT images have been classified into three categories:

- (i) The first category includes 5705 non-informative CT (NiCT) images without lung parenchyma,

- (ii) The second category includes 4001 positive CT (pCT) images with features related to COVID-19 pneumonia, and
- (iii) The third category includes 9979 negative CT (nCT) images with irrelevant features to COVID-19 pneumonia.

Fig. 1 is a sample image of the dataset.

### 2.2. The proposed model

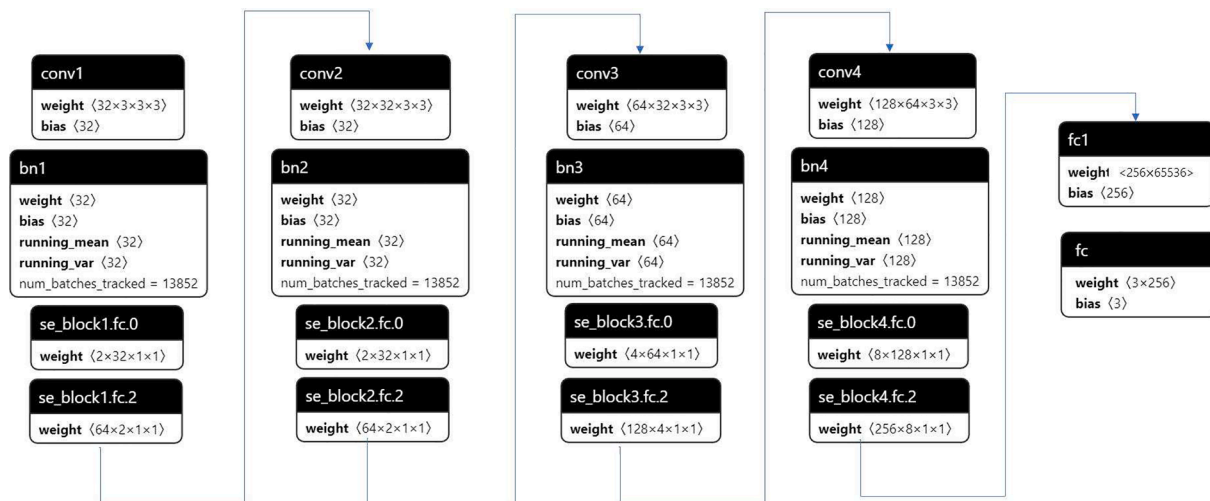
Fig. 2 shows the proposed model in which the feature extraction process is performed in three layers:

1. In the convolution layer, in the proposed polishing layer, and in the SE block layer, how to extract a feature in convolution layers similar to other structures.
2. The proposed pooling layer is prepared using HAAR wavelet filters. In this layer, minimizing the dimensions and feature extraction are performed using wavelet HAAR simultaneously. The proposed method extracts better features than other classifying CT scan image data methods, as the test results show.

The proposed model includes a novel network structure in which batch normalization (BN) [32] is used to shorten the convergence time and achieve better performance. The activation function is Mish Function [33] to improve the classification capacity in nonlinear cases. The

**Table 3**  
Metric results of the proposed network.

Test Cohen kappa score	Test mean precision	Test mean recall	loss	Test accuracy	optimizer
95.33	98.06	95.48	0.1471	97.15	SGDM [35]
96.68	98.42	96.86	0.0869	97.97	GC-SGDM [36]
95.17	97.98	95.59	0.0958	97.05	Adam [37]
95.16	98.10	95.73	0.0715	97.06	GC-Adam [36]
93.74	96.99	94.04	0.0993	96.19	NAdam [38]
72.22	84.03	80.99	0.3780	82.26	GC-NAdam [36]
98.43	98.71	98.91	0.0338	99.03	RAdam [39]
96.27	98.18	97.02	0.0790	97.71	GC-RAdam [36]



**Fig. 3.** Block diagram of proposed network.

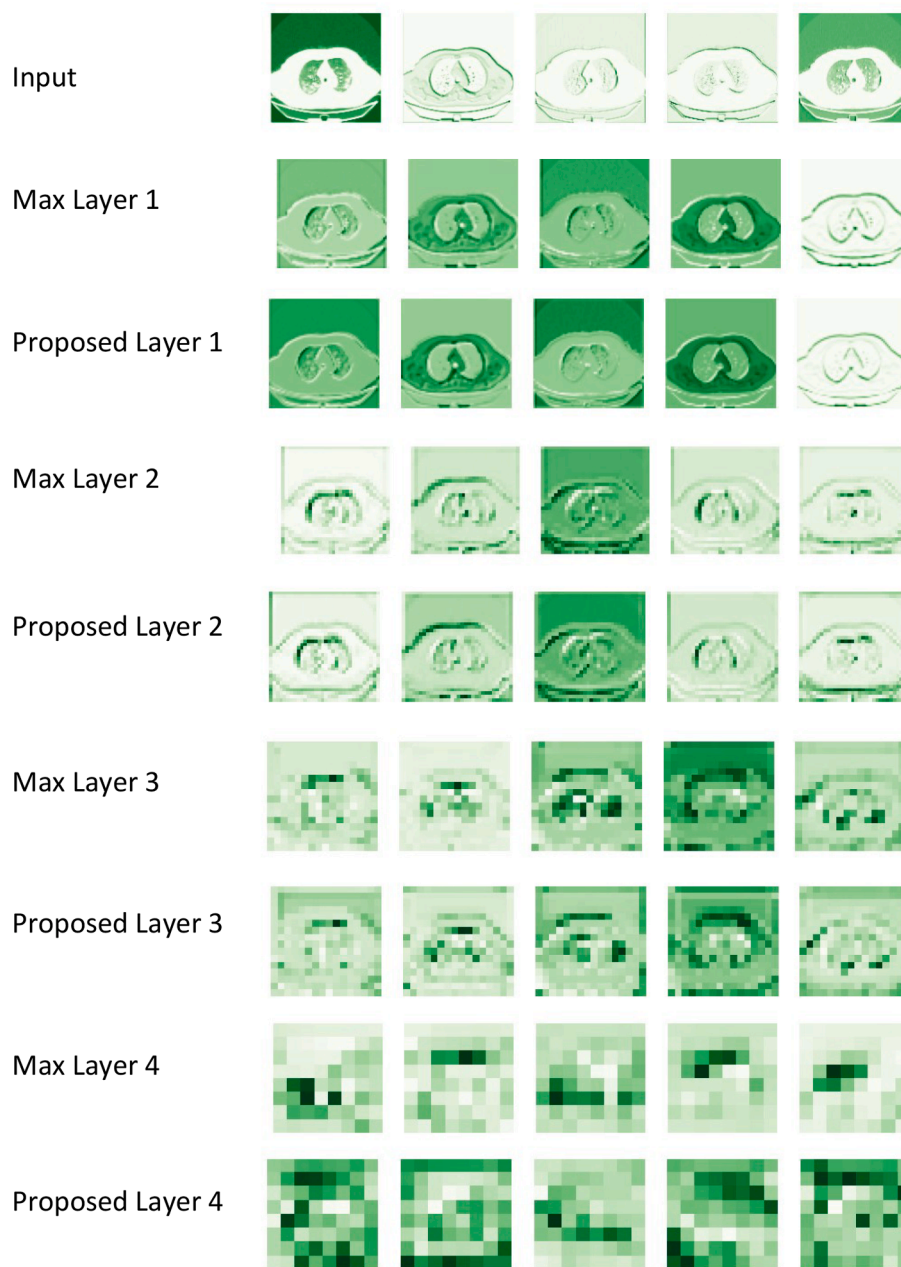


Fig. 4. Comparison of max pooling and the proposed pooling layers.

Table 4

Comparison of the proposed method with some popular DNNs. Prediction times do not include image loading.

	Model	Number of Parameters	Predictiontime GPU	Accuracy	Loss	Cohen kappa score	Mean precision	mean recall
Popular DNNs	VGG11 + BN	128,792,325	5.186 ms	92.73	0.9402	87.87	95.35	89.42
	ResNet18	11,179,077	4.262 ms	93.39	0.9594	88.98	95.79	89.49
	ResNet50	23,518,277	5.554 ms	95.98	0.5697	93.40	97.01	93.98
	Inception-v3	24,351,719	6.844 ms	98.11	0.1604	96.94	98.49	97.61
Proposed network	WCNN4	4,610,531	0.069 ms	99.03	0.0338	98.43	98.71	98.91

dropout layer with  $p = 0.5$  is used. Finally, an SE block is added after each dropout layer.

The pooling layer is the Haar wavelet transform layer [28] to produce better features. Table 1 shows the network’s configuration in detail. In the first stages, there are some convolutional layers so that the size of their kernel is according to table 2. Then, the SE block is

embedded after any activation function. Finally, the features are fed into the Fully-Connected and SoftMax layers. Fig. 3 depicts the block diagram of the proposed network. Because the proposed model contains Wavelet and four convolution layers, it is named Wavelet CNN-4 (WCNN4).

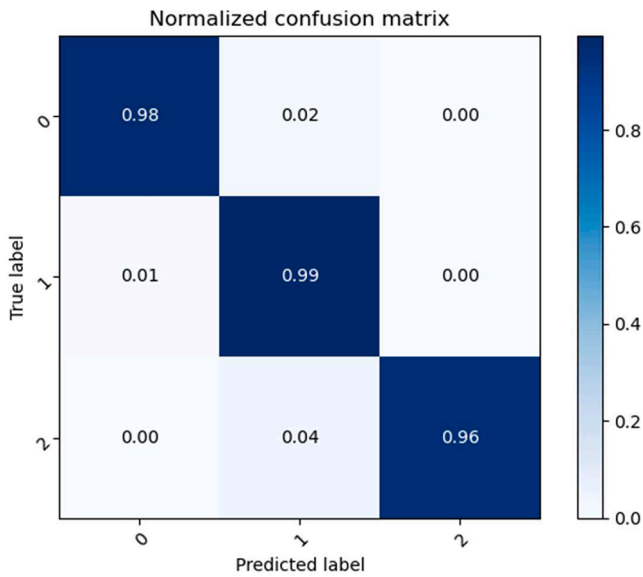


Fig. 5. Confusion matrix of Inception network.

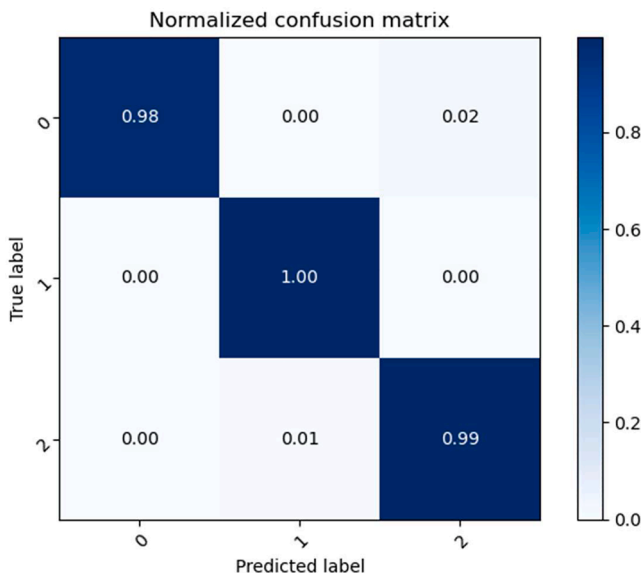


Fig. 6. Confusion matrix of proposed network.

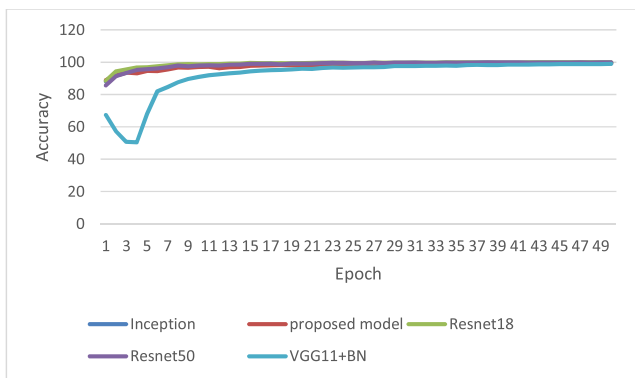


Fig. 7. Accuracy in the train phase for different DNNs.

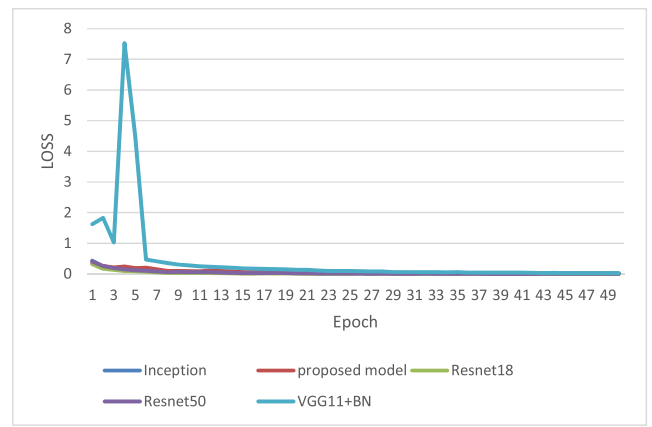


Fig. 8. Loss in train phase for different DNNs.

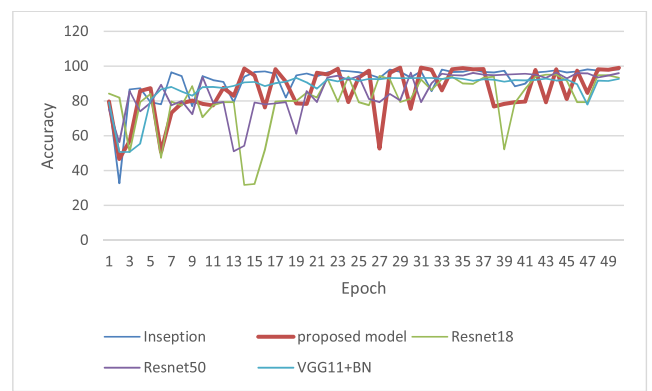


Fig. 9. Accuracy in the train phase with test data for different DNNs.

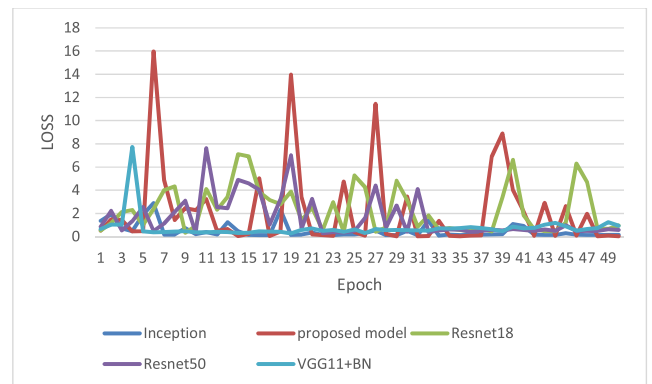


Fig. 10. Loss in train phase with test data for different DNNs.

### 3. Experimental results

The experiments have been run on a PC with GeForce Turbo RTX-2080 GPU and Corei3-9100f CPU running at 3600 MHz. The implementation of the model was in Python 3.7 using the Pytorch library. All experiments have been done in GPU. Because of the difference in class numbers in any dataset, the weighted random sampler [34] was used as a Sampler in the training step. Table 2 shows all training parameters.

The proposed Deep Neural network was trained using Covid-19 CT Scan images. The training phase used eight different optimizers. Table 3 shows the metric results of the proposed approach. This table expresses the best results in metrics achieved using the RAdam optimizer.

### 3.1. Comparison with popular DNNs

Fig. 4 depicts a comparison between the proposed structure and the standard structure using the most common pooling called max-pooling [40]. The result shows there is not much difference in the extracted features in the network's initial layers. However, in the final layer, the difference between the extracted features is clear.

To further evaluate, some popular DNNs, including VGG [41], ResNet [42], and Inception [43], have been trained on the test dataset.

VGG architecture is one of the first popular structures that produced outstanding results in the ImageNet large-scale visual recognition challenge (ILSVRC-2014) competition [44]. Its original structure cannot be trained from scratch due to the vanishing gradients problem [45]. However, today, with batch normalization, its modified version can be trained from scratch. For comparison, VGG11 with batch normalization has been trained on the test dataset.

The second structure is ResNet. It has a novel structure, and having the new residual connections, can be trained on every dataset. It supports very deep layers and exists in different versions from ResNet18 to ResNet152 and even more. To compare the structures, ResNet18 and ResNet50 have been trained on the test dataset.

Inception is another network that concatenates the sparse layers to make dense layers [46]. This structure reduces dimension to achieve more efficient computation and deeper networks as well as overfitting. Inception architecture takes multiple kernel filter sizes in a convolutional neural network.

Table 4 presents the classification metric parameters for popular DNNs and the proposed network. Moreover, Figs. 5 and 6 depict the confusion matrix for the proposed network and inception-V3. The accuracy and loss diagrams in the training phase of the tests are presented in Figs. 7 and 8. The final relationship of these networks is similar to each other. In the next step, the accuracy and loss for test data were measured for each epoch. The results depict that the proposed network has the best performance (see Figs. 9 and 10).

## 4. Discussion

This work proposes a deep model based on batch normalization and Mish function to detect COVID-19 cases from CT images. Two real datasets were used to evaluate the proposed model. The metric results of the proposed network showed an accuracy of 99.03.

This study tries to introduce a network with kernel-based machine learning applications [47]. Therefore, some networks have been selected for the test phase suitable for real-time and online applications. According to Table 4, the worst result in popular DNNs belongs to VGG11 with a 92.73% accuracy. This network is too deep and has 128 million parameters. The best performing DNN was Inception-v3 with a 98.11% accuracy. As shown in Table 4, the best recognition time in popular DNNs is achieved using Resnet18 and is 4.262 ms. This network has more than 11 million trainable parameters and has about 93.39% accuracy on the test dataset. According to the Cohen Kappa score rate for Resnet18, it is about 88.98%; the classification is not good.

Although all of the popular DNNs in our work are powerful networks in classification problems for RGB images [48], they do not have good CT image behavior. Comparing to these networks, the proposed model has four novel innovations; using wavelet [49] for pooling the feature maps, Mish [33] activation Function, Mini-Batch Normalization [32], and SE-blocks [50]. Using wavelet transform into network structure is a reason for extracting the correct and appropriate information for recognition tasks. Mish Function is a novel activation function used in powerful detection and classification networks, such as yolov4 [51], which has an essential role in the appropriate information extraction. One of the essential parts of the proposed model is using mini-batch normalization in any layer. Squeeze-and-Excitation blocks (SE block) is an architectural unit that can be plugged into a CNN structure to improve performance with only a slight increase in the total number of

parameters. Squeeze-and-Excitation blocks explicitly model channel relationships and channel interdependencies and include a form of self-attention on channels. Consequently, it results in the best performance in classification metrics parameters and recognition time.

Furthermore, the number of trainable parameters is more suitable for CT images than more deep structures that suffer from overfitting. The proposed network is the best in classification metric parameters comparing to the popular DNNs. Moreover, WCNN4 has the least prediction time. Therefore, it is a helpful network for real-time application for the recognition of covid-19 using CT images.

Figs. 5-8 show the networks are similar in the final Epochs in the training phase. However, there are differences between them in the network test phase. This difference is generally due to the complexity of the network structure. Bigger structures need more data in the training phase. As shown in Fig. 7 and Fig. 8, vgg11 has the most considerable accuracy entropy in the training phase, and the proposed network has a smooth curve. Fig. 9 and Fig. 10 are the accuracy and loss for test data after each epoch. As shown, the suggested model has decreasing trend entropy in loss and accuracy.

Finally, an important achievement of the proposed network's evaluation is that the test mode's recognition time was 0.069 ms in GPU.

## 5. Conclusion

Rapid and accurate diagnosis is necessary as the number of patients with COVID-19 infection is increasing. This work proposes a deep learning-based model to detect COVID-19 disease from CT images. The model is an automated method without any feature extraction phase. The proposed model has four novel innovations, including wavelet for pooling the feature maps, Mish activation function, mini-batch normalization, and SE-blocks. For evaluation, the model was trained and tested on the datasets of two public hospitals. The results show the suggested model achieved excellent metric parameters in classifying COVID-19 cases, such as accuracy of 99.03. Moreover, it decreases trend entropy in loss and accuracy. Finally, the recognition time in test mode was 0.069 ms in GPU. The experimental results show the proposed model is a helpful network for real-time application for recognition of covid-19 using CT images.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] (15 October 2020). WHO Coronavirus Disease. Available: <https://covid19.who.int/>.
- [2] W. Wang, et al., Detection of SARS-CoV-2 in different types of clinical specimens, *JAMA* 323 (18) (2020) 1843–1844.
- [3] A. Bernheim, et al., Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection, *Radiology* (2020), 200463.
- [4] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, J. Liu, Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing, *Radiology* (2020), 200343.
- [5] S. Ahuja, B. K. Panigrahi, N. Dey, V. Rajinikanth, T. K. Gandhi. Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. 2020.
- [6] J. Zhang, Y. Xie, Y. Li, C. Shen, Y. Xia. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. arXiv preprint arXiv:2003.12338, 2020.
- [7] K. Elasnoui, Y. Chawki, Using X-ray images and deep learning for automated detection of coronavirus disease, *J. Biomol. Struct. Dyn.* no. just-accepted (2020) 1–22.
- [8] S. Asif, Y. Wenhui, H. Jin, Y. Tao, S. Jinhai. Classification of covid-19 from chest x-ray images using deep convolutional neural networks. medRxiv, 2020.
- [9] K. Purohit, A. Kesarwani, D. R. Kisku, M. Dalui. Covid-19 detection on chest x-ray and ct scan images using multi-image augmented deep learning model. BioRxiv, 2020.
- [10] T. Majeed, R. Rashid, D. Ali, A. Asaad. Covid-19 detection using CNN transfer learning from X-ray Images. MedRxiv. 2020.

- [11] E. F. Ohata et al. Automatic detection of COVID-19 infection using chest X-ray images through transfer learning. *IEEE/CAA Journal of Automatica Sinica*. 2020.
- [12] A. Abbas, M. M. Abdelsamea, M. M. Gaber. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. arXiv preprint arXiv:2003.13815. 2020.
- [13] B. Abraham, M.S. Nair, Computer-aided detection of COVID-19 from X-ray images using multi-CNN and Bayesnet classifier, *Biocybernetics and Biomedical Engineering* 40 (4) (2020) 1436–1445.
- [14] M.J. Horry, et al., Covid-19 detection through transfer learning using multimodal imaging data, *IEEE Access* 8 (2020) 149808–149824.
- [15] M. Taresh, N. Zhu, T. A. A. Ali. Transfer learning to detect COVID-19 automatically from X-ray images, using convolutional neural networks. *MedRxiv*. 2020.
- [16] K. H. Shibly, S. K. Dey, M. T. U. Islam, M. M. Rahman. COVID Faster R-CNN: A Novel Framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray Images. *MedRxiv*. 2020.
- [17] M.Z. Che Azemin, R. Hassana, M.I. Mohd Tamrin, M.A. Md Ali, COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest X-ray images as training data: preliminary findings, *Int. J. Biomed. Imaging* 2020 (2020) 1–7.
- [18] A. Mangal et al., "CovidAID: COVID-19 Detection Using Chest X-Ray," arXiv preprint arXiv:2004.09803, 2020.
- [19] 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.
- [20] B. Sekeroglu I. Ozsahin "<? covid19?> Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks," *SLAS TECHNOLOGY: Translating Life Sciences Innovation* 2020 2472630320958376.
- [21] R. Mohammadi, M. Salehi, H. Ghaffari, A. Rohani, R. Reiazi, Transfer Learning-Based Automatic Detection of Coronavirus Disease 2019 (COVID-19) from Chest X-ray Images, *Journal of Biomedical Physics and Engineering* 10 (5) (2020) 559–568.
- [22] M. Loey, F. Smarandache, 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* 12 (4) (2020) 651.
- [23] M.A. Elaziz, K.M. Hosny, A. Salah, M.M. Darwish, S. Lu, A.T. Sahlol, R. Damasevicius, New machine learning method for image-based diagnosis of COVID-19, *PLoS ONE* 15 (6) (2020) e0235187, <https://doi.org/10.1371/journal.pone.0235187>.
- [24] 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*, p. 103792, 2020.
- [25] A.I. Khan J.L. Shah 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* 2020 105581.
- [26] M. Karim, T. Döhmen, D. Rebholz-Schuhmann, S. Decker, M. Cochez, and O. Beyan, "Deepcovidexplainer: Explainable covid-19 predictions based on chest x-ray images," arXiv preprint arXiv:2004.04582, 2020.
- [27] S.H. Yoo, H. Geng, T.L. Chiu, S.K. Yu, D.C. Cho, J. Heo, M.S. Choi, I.H. Choi, C. Cung Van, N.V. Nhung, B.J. Min, H. Lee, Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging, *Frontiers in medicine* 7 (2020), <https://doi.org/10.3389/fmed.2020.00427>.
- [28] R.S. Stanković, B.J. Falkowski, The Haar wavelet transform: its status and achievements, *Comput. Electr. Eng.* 29 (1) (2003) 25–44.
- [29] D. Misra, "Mish: A self regularized non-monotonic neural activation function," arXiv preprint arXiv:1908.08681, 2019.
- [30] COVID-19 CT Scan Images. Available: <https://www.kaggle.com/azaemon/preprocessed-ct-scans-for-covid19?select=Original+CT+Scans>.
- [31] W. Ning et al., "iCTCF: an integrative resource of chest computed tomography images and clinical features of patients with COVID-19 pneumonia," 2020.
- [32] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
- [33] D. Misra "Mish, A Self Regularized Non-Monotonic Activation Function," arXiv preprint arXiv:1908.08681 2020 1 14.
- [34] P. Efraimidis and P. Spirakis, "Weighted Random Sampling," in *Encyclopedia of Algorithms*, M.-Y. Kao, Ed. Boston, MA: Springer US, 2008, pp. 1024–1027.
- [35] R. Sutton Two problems with back propagation and other steepest descent learning procedures for networks in *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* 1986, 1986, 823 832.
- [36] H. Yong J. Huang X. Hua L. Zhang Gradient Centralization: A New Optimization Technique for Deep Neural Networks 2020 arXiv preprint arXiv:2004.01461.
- [37] D.P. Kingma J. Ba Adam: A method for stochastic optimization arXiv preprint arXiv:1412.6980 2014.
- [38] T. Dozat, "Incorporating nesterov momentum into adam," 2016.
- [39] L. Liu et al. On the variance of the adaptive learning rate and beyond arXiv preprint arXiv:1908.03265 2019.
- [40] V. Christlein, L. Spranger, M. Seuret, A. Nicolaou, P. Král, A. Maier, in: *Deep generalized max pooling*, *IEEE*, 2019, pp. 1090–1096.
- [41] K. Simonyan A. Zisserman Very deep convolutional networks for large-scale image recognition arXiv preprint arXiv:1409.1556 2014.
- [42] K. He X. Zhang S. Ren J. Sun Deep residual learning for image recognition in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2016 770 778.
- [43] C. Szegedy et al. Going deeper with convolutions in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2015 1 9.
- [44] A. Khan, A. Sohail, U. Zahoor, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, *Artif. Intell. Rev.* 53 (8) (2020) 5455–5516.
- [45] C. Francois, "Deep learning with Python," ed: Manning Publications Company, 2017.
- [46] A. Kapoor, R. Shah, R. Bhuvan, and T. Pandit, "UNDERSTANDING INCEPTION NETWORK ARCHITECTURE FOR IMAGE CLASSIFICATION".
- [47] L. Le, Y. Xie, Deep embedding kernel, *Neurocomputing* 339 (2019) 292–302.
- [48] H. Gholamalinejad, H. Khosravi, IRVD: A Large-Scale Dataset for Classification of Iranian Vehicles in Urban Streets," *Journal of AI and Data Mining* (2020).
- [49] A. Barbhuiya and K. Hemachandran, "Wavelet Transformations and Its Major Applications in Digital Image Processing," *International Journal of Engineering Research & Technology (IJERT)*, ISSN, pp. 2278-0181.
- [50] J. Hu L. Shen G. Sun Squeeze-and-excitation networks in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2018 7132 7141.
- [51] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934. 2020.