



Performance Evaluation of Deep Learning in Detection COVID-19 Based on Different Types of Datasets

Rasim Azeez Kadhim^{1*}, Suhad Shakir Jaber²

¹College of Information Technology, University of Babylon, rasimazeez@uobabylon.edu.iq, Babylon, Iraq.

²Technical College/ AL-Mussaib, Al-Furat Al-Awsat Technical University, suhad.jaber@atu.edu.iq, Babylon, Iraq.

*Corresponding author email: rasimazeez@uobabylon.edu.iq mobile: 07739773196

تقييم اداء التعلم العميق في كشف كوفيد 19 بالاعتماد على انواع مختلفة من مجاميع البيانات

راسم عزيز كاظم^{1*}، سهاد شاكر جابر²

1 كلية تكنولوجيا المعلومات، جامعة بابل، rasimazeez@uobabylon.edu.iq ، بابل، العراق

2 الكلية التقنية المسيب، جامعة الفرات الاوسط، suhad.jaber@atu.edu.iq ، بابل، العراق

Accepted:

11/3/2024

Published:

31/3/2024

ABSTRACT

Background:

In this paper, we conducted a comparison of a number of deep learning networks in detecting Covid-19 disease based on X-ray and CT-scan images.

Materials and Methods:

An eight different deep learning network types (ResNet50, Mobilenetv2, Densnet201, SqueezeNet, Efficientb0, Googlenet, VGG19, and Alexnet25) are investigated to detect the COVID-19 based on the X-ray and CT-scan images. A transfer-learning technique is adopted in the training phase for all networks with different datasets. The networks are trained and tested on 70% and 30% of the dataset, respectively.

Results:

The confusion matrix is calculated in the testing phase and the evaluation metrics including F1 score, accuracy, precision, specificity and sensitivity are calculated from the confusion matrix. The comparison results demonstrate that the classification accuracy when using X-ray dataset is better than that of using CT-scan datasets. Moreover, the Mobilenetv2 delivered the greatest results for different datasets in which the accuracy is greater than 99% for X-ray images and the accuracy is less than 80% for CT-scan images .

Conclusion:

Our conclusion is that the using of Mobilenetv2 network with the X-ray images is more suitable than the others.

Keywords: CNN, Deep Learning, COVID-19, Coronavirus, X-ray, CT-scan.



INTRODUCTION

The severe respiratory syndrome coronavirus (COVID-19), which causes the COVID-19 pandemic, remains to have a terrible effect on the human health and general welfare of the world's people (SARS-CoV-2). In order to treat individuals who are affected right away and to prevent the spread of the virus, effective patient screening is crucial in the fight against COVID-19. The main method for identifying COVID-19 infections is the testing of the reverse transcriptase-polymerase chain reaction (RT-PCR), that can identify SARS-CoV-2 RNA from the material of breathing. Testing of RT-PCR is the best standard because it is so precise, but performing it by hand requires a lot of time, effort, and resources [1]. The World Health Organization (WHO) proclaimed the COVID-19 viral outbreak a pandemic on March 11 of 2020, and since that time the virus has spread quickly and killed many people in many different countries. The radiography and imaging techniques used to diagnose pneumonia can be used to identify the signs of COVID-19. When it comes to commercial and widespread utilization, COVID-19 detection, out of the two, uses image testing in a quick and effective manner and can thus be utilized to stop the virus's propagation. Imaging methods like as chest X-rays (CXR) and computed tomography (CT) are crucial in the diagnosis of COVID-19 disease [2]. The high efficiency results obtained by the deep learning in classification of digital images, encouraged many researchers to use it in classification of medical images of COVID-19 [3-9].

In Ref. [10], the authors utilized the pre-trained model of ResNet50 to analyze the classifier. According to degree of similarity to the input images, relevant test images were found using the nearest neighbor technique. The model's accuracy was assessed and determined to be 89.7%. Also, Asif et al. presented employing a pre-trained model named InceptionV3 based on the transfer learning to diagnose the coronavirus in infected people using CXR images. Its categorization accuracy rate is 98% [11]. In Ref. [12], they presented a method depended on a smartphone for warning closed people to an area with COVID-19 infection. The proposed approach employed positional information and distance measurements from nearby signed-up users [12].

In Ref. [13], the authors showed how to use ensembles of successively pruned deep learning models to detect COVID-19 pulmonary symptoms using CXR. The performance of this model was improved based on the weighted average of the top-performing trimmed models where the accuracy is 99.01% [13]. Many frequently occurring thoracic ailments can be recognized and even spatially located, as Wang et al. showed utilizing a tried-and-true unified framework for multi-label classification and localization under weak supervision [14]. Deep learning with DCNN was used by Lakhani and his team, who demonstrated accurate categorization from chest pictures [15]. An accuracy of 95.38%, SVM-based coronavirus identification method was demonstrated by Ref. [16]. In Ref. [17], based on transfer learning, the authors used four types of CNNs (ResNet50, ResNet18, DenseNet121, and SqueezeNet). The trials made use of a database that had 184 COVID-19 photos, 5000 images with no discoveries, and images of pneumonia. According to the information provided, the specificity was 92.9% and the sensitivity was around 98%. The proposed model of COVID-CAPS in [18] based on the dataset of CXR images has been outperform the earlier models of CNN. Although COVID-CAPS had a far fewer number of parameters of training than other models, it still managed to obtain 95.8% specificity, 90% sensitivity, and 95.7% accuracy.

In Ref. [19], the dataset of 1428 chest radiographs images of bacterial pneumonia , COVID-19 patients and healthy cases have been examined. The pre-trained VGG16 model was adopted in this paper to successfully training the network depending on comparatively small radiograph images of chest for classification tasks. The three classes classification (normal, COVID, pneumonia,) case, and the two classes (COVID and non-COVID) case, showed accuracy rates of 92.5% and 96%, respectively. In Ref. [1], the COVID-Net architecture was suggested. Even though only 266 of 358 patients were assigned to COVID-19, the COVIDx open resource of 13975 CXR images were used for the model training. The achieved accuracy was 93.3%. The Xception network was utilized by the authors to build a classification model for COVID-19. The used database was aggregated from various sources which consist of 500 pneumonias, 500 controls, and 127 COVID. The achieved accuracy was roughly 97% [20]. For this inquiry, a special an end to end training CNN model was recommended in Ref. [21]. A dataset of 200 regular (healthy) and 180 COVID-19 CXR images was utilized in the experiments of this research. The study's findings led to a classification accuracy performance assessment of 91.6%. The majority of earlier studies only used one kind of dataset. Based on the X-ray and CT-scan datasets, the performances of eight deep learning models, including Densnet201, Efficientb0, Googlenet, Mobilenetv2, Resnet50, VGG19, Alexnet25, and Squeezenet, will be assess in this work.

SYSTEM MODEL

The general block diagram is illustrated in figure 1. The CT-scan or CXR images were inputted to the first stage of preparation where the images are resized and normalized to be suitable for each CNN model. Also, for training and validation, the dataset was divided into two categories. After that, the eight models are trained and tested based on these datasets. Finally, the system was evaluated by calculating the evaluation metrics represented by the precision, accuracy, sensitivity, specificity and f1-score. The details of each CNN will be presented in the following:

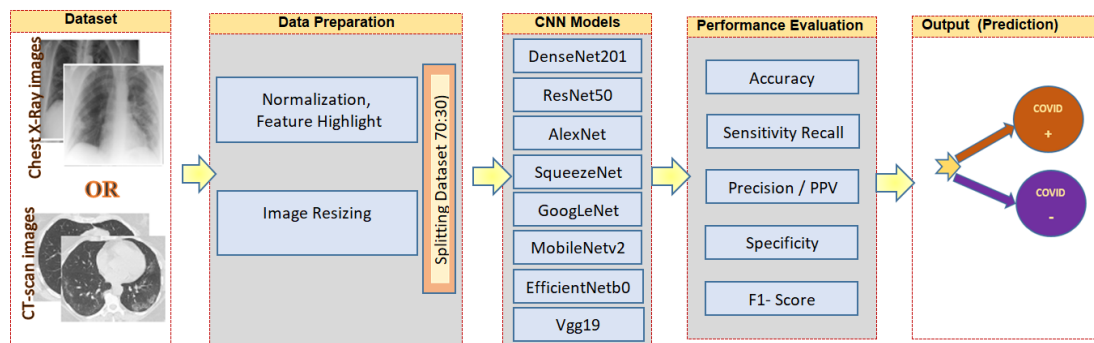


Figure (1): The System Model General Block Diagram.

مجلة جامعة بابل للعلوم الطبية والعلوم الصحية | Volume 32, No. 1, 2024 | ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com | ISSN: 2312-8135 | Print ISSN: 1992-0652

**a- VGG19**

The VGG abbreviation referred to Visual Geometry Group. The construction of VGG19 consist of 33 kernel-sized filters where 11 in the first layer and 5 are integrated in the second layer. A 224×244 RGB image is set up as the input for VGG. More than one million photos from the ImageNet dataset were utilized to train the VGG-19 CNN. The depth of the network is 19 layers and can classify photos from different classifications. The main goal of VGG designs a deep neural network with moderate and constant convolution size.

b- ResNet50

The architecture ResNet50's can be divided into four phases. Images with width, height, and channel numbers that are multiples of 32 can be accepted and fed into the network. Each ResNet design uses 7×7 kernel size for initial convolution and 3×3 kernel size in the layer of max-pooling. By using the block of 3-layers bottleneck in place of each 2-layer block, the network of 34-layers is transformed into a 50-layer ResNet.

c- GoogLeNet

Compared to the Inception architecture, GoogleNet's deep CNN has 22 layers and almost 12 less parameters. However, as the number of parameters increases with more layers, the network may overfit. Images having a resolution of 224×224 are accepted by the pre-trained network. Global average pooling was adopted in GoogLeNet in place of a fully linked layer. The AveragePooling2D, the activation function and Dense layers are utilized by the architecture.

d- MobileNetV2

The MobileNetV2 used the inverted residual structure to build a new module. MobileNetV2 enables contemporary item identification and semantic segmentation. The first layer of the MobileNetV2 architecture is composed of the fully convolutional layer, which has 19 bottleneck layers and 32 filters layers. The number of parameters about 3.4 million and the number of multiply-add operations is 300 million are frequently used by the network. To increase accuracy, ReLU6 is eliminated from each bottleneck module's output.

e- AlexNet25

Five convolutional layers, three max-pooling layers, two normalization layers, two fully connected layers, and one softmax layer make up AlexNet25. Convolutional filters and a ReLU nonlinear activation function make up each convolutional layer. The pooling layers are used to achieve maximum pooling. The input size of 224x224x3 is fixed because there are entirely linked layers. In order to convert a grayscale image to RGB, the single channel is duplicated, yielding a three-channel RGB image. The total number of parameters in AlexNet is 60 million, and its batch size is 128.

f- EfficientNetB0

The AutoML MNAS framework was used to generate a new baseline network to improve performance, which increases both accuracy and effectiveness (FLOPS). The mobile inverted bottleneck convolution (MBCConv) is used in this architecture, the architecture is the same of the MnasNet and MobileNetV2, however it is a little greater because of increasing the budget of the FLOP. The



fundamental network is then expanded to give the EfficientNets models. There are no trained weights included in EfficientNetB7.

g- DenseNet201

Every layer is directly connected with other layers in dense connected CNN. Between "M" levels, the direct connections are equal to $M(M + 1)/2$. Instead of being averaged, the feature maps from earlier levels are considered as inputs for each layer. Hence, DenseNets uses less number of parameters than the traditional CNN, which allowing the reuse of features by rejecting the redundant feature maps. Dense Blocks use constant number of feature maps and variable number of filters. A 1×1 convolution, batch normalization, and 2×2 pooling are used in the intermediate layers to downscale the image.

h- SqueezeNet

The SqueezeNet CNN was developed to obtain a competitive accuracy with a small number of parameters in the structure. The structure of SqueezeNet based on the Fire module which made up of an extend layer that combines 1×1 and 3×3 convolution filters and a squeeze convolution layer that only includes 1×1 filters [22].

DATASETS

There are many datasets of coronavirus that collected by the medicine organization and universities. The following two types of datasets were used in this study:

A- The CT-scan dataset

The digital CT-scan images dataset for the COVID-19 that utilized in this work was collected from the Kaggle repository [23]. The collection is arranged into a single folder with subfolders for each type of image (COVID/NonCOVID). There are 2481 CT-scan images for COVID-19 and NonCOVID arranged in two categories. There are 1229 photos in the NonCOVID class and 1252 in the COVID-19 class.

B- The X-ray dataset

The dataset of the X-ray images of this work was downloaded from the repository of Kaggle, and it contains the images of pneumonia and Covid-19 afflicted individuals as well as normal images [24]. This dataset contains 5228 samples of CXR images in which it divided into three groups or classes (1626 COVID, 1802 NORMAL, and 1800 PNEUMONIA) make up this data set. Each group was separated in sub-folder. For comparison purposes, the third class (PNEUMONIA) is eliminated.

SIMULATION RESULTS AND DISCUSSION

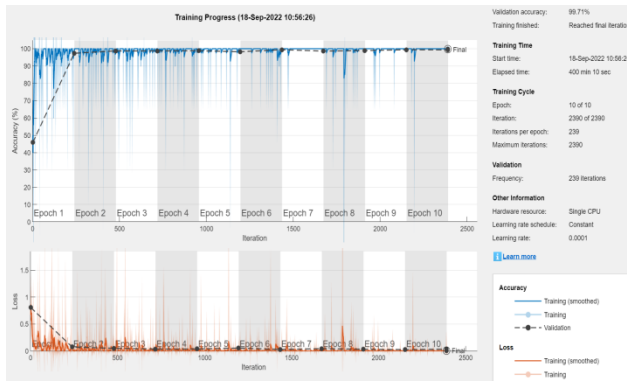
Densnet201, Efficientb0, Googlenet, Mobilenetv2, Resnet50, VGG19, Alexnet25, and Squeezenet are eight newly developed deep learning models that are trained individually on two different datasets. These networks' training and testing phases are constructed in MATLAB on an I7 core of laptop running at 1.8 GHz and 16GB of RAM. The first dataset is made up of CT-scan images, and the second dataset is made up of X-ray data. All networks employed transfer learning, with the last three layers altered to support the classification of only two classes. Ten epochs are set for all cases. The eight models' training process on the X-ray dataset is depicted in Figure 2. Figure 3 also displays the training phase for the eight models using the CT-scan dataset.



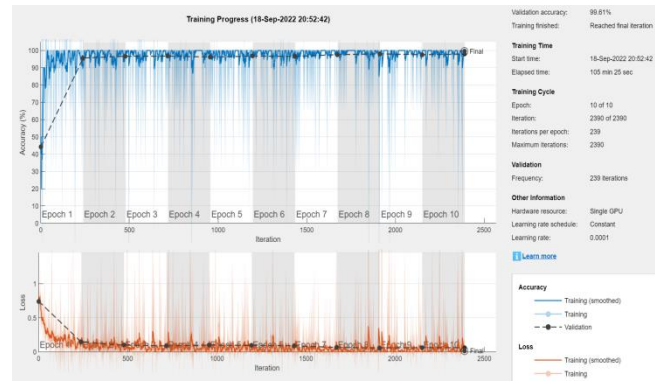
The two classes (COVID/NonCOVID) were presented in all test experiment cases, we selected randomly 70% of the images in dataset for training phase and 30% are used in test phase. In the training phase, 2399 images are used, and in the validation phase, 1029 X-ray images are employed. Additionally, the training and testing phases use 1736 and 745 CT-scan pictures, respectively. The confusion matrices of all eight models based on the X-ray and CT-scan datasets are illustrated in figure 4 and figure 5, respectively. Based on the confusion matrix, the evaluation metrics represented by the Accuracy, Sensitivity, F1-score, Precision, and Specificity are computed according to the following formulas in Table (1).

Table-1: The Performance Metrics

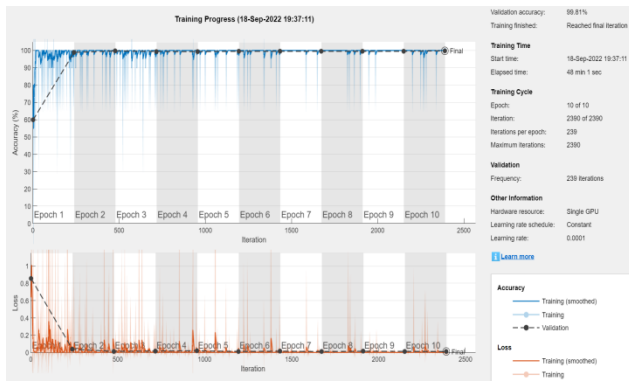
Metrics	Formula
Sensitivity	$\text{Sensitivity} = \frac{TP}{FN + TP}$
Precision	$\text{Precision} = \frac{TP}{TP + FP}$
Accuracy	$\text{Accuracy} = \frac{TN + TP}{TP + FP + TN + FN}$
Specificity	$\text{Specificity} = \frac{TN}{TN + FP}$
F1-score	$\text{F1 - Score} = \left(\frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \right) \times 2$



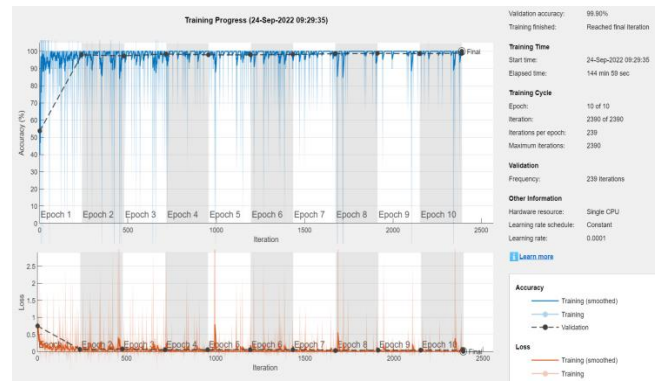
(a)



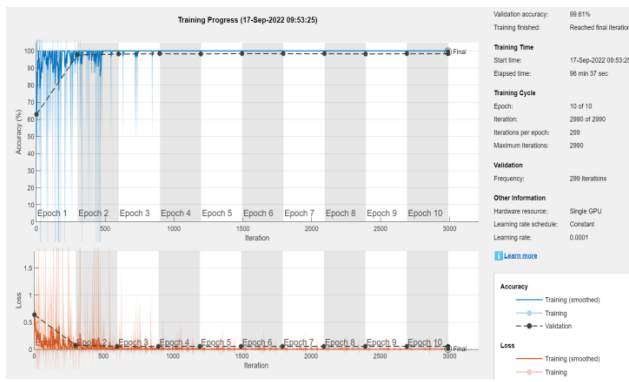
(b)



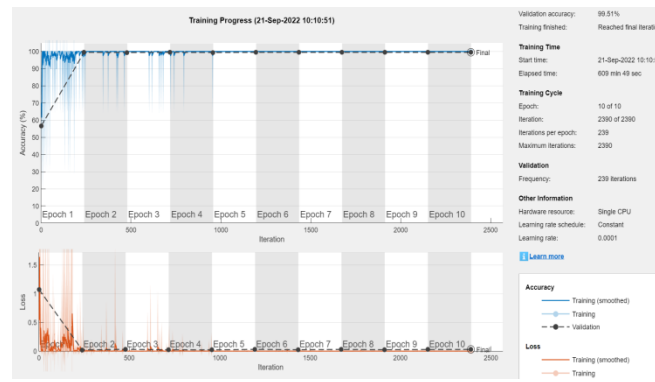
(c)



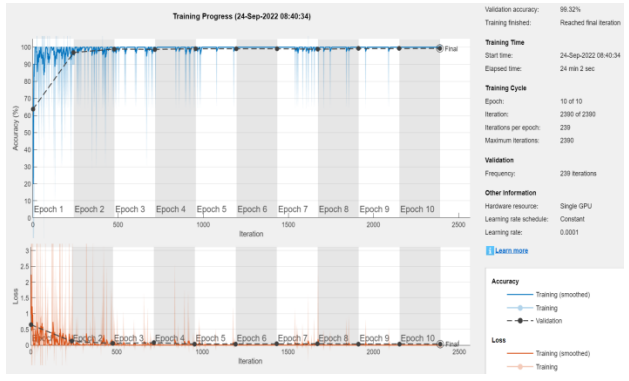
(d)



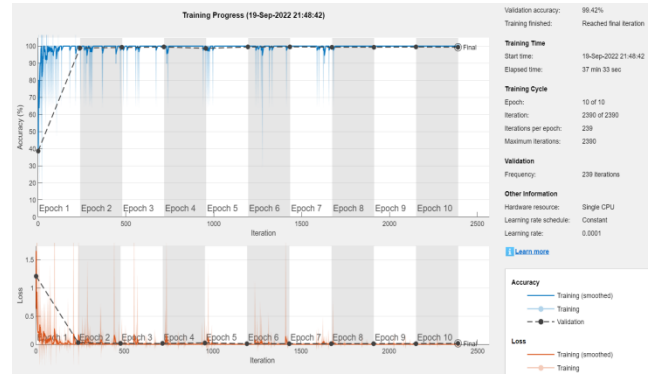
(e)



(f)



(g)

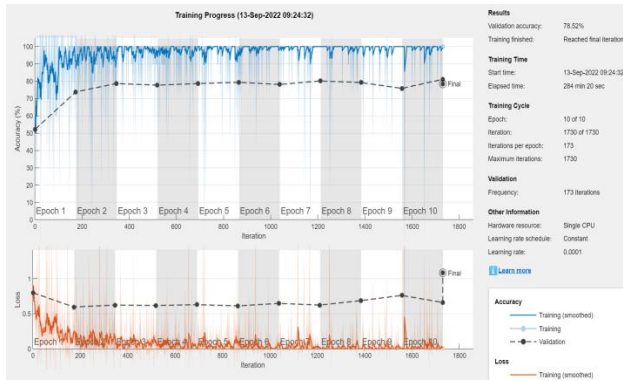


(h)

Figure (2): Training progress on X-ray dataset of (a) Densnet201, (b)Efficientb0, (c)Googlenet, (d)Mobilenetv2, (e) Resnet50, (f) VGG19, (g) Alexnet25, (h) Squeezenet.



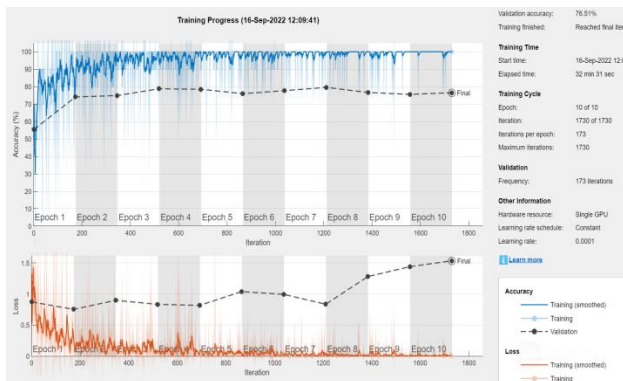
تحقیق
فنی
مجموعه
فصلنامه
علوم
پایه
و
کاربردی
(JUBPAS)



(a)



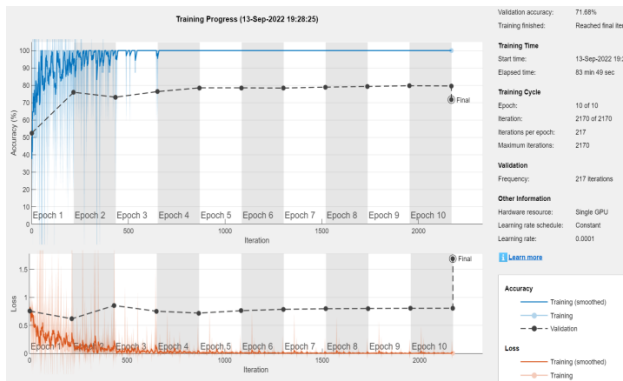
(b)



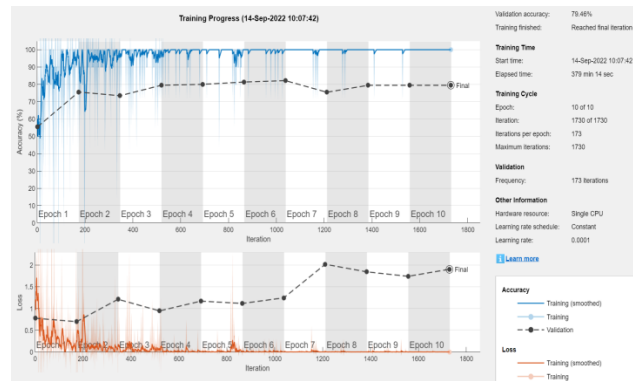
(c)



(d)



(e)



(f)

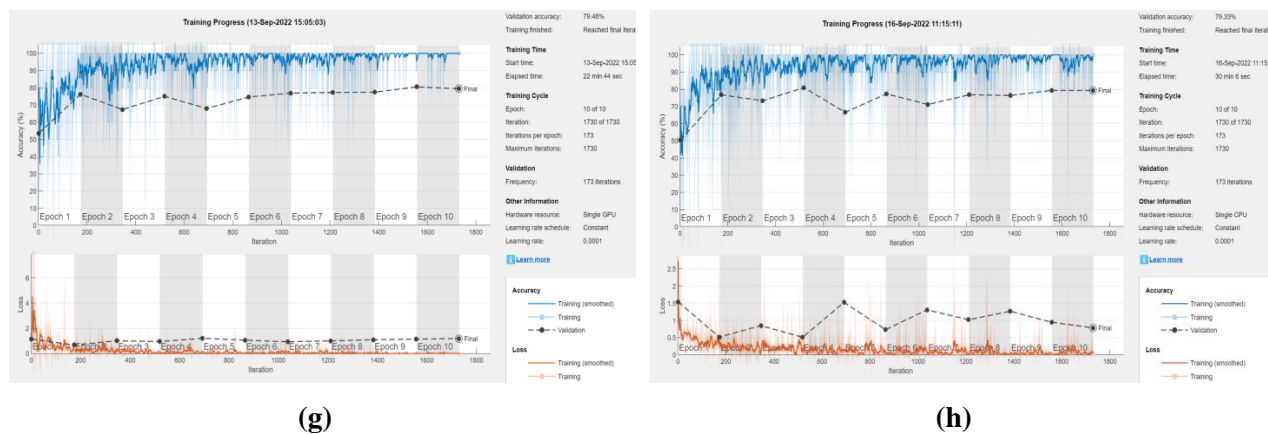


Figure (3): Training progress on CT-scan dataset of (a) Densnet201, (b)Efficientb0, (c)Googlenet, (d)Mobilenetv2, (e) Resnet50, (f) VGG19, (g) Alexnet25, (h) Squeezenet.

The sensitivity values for the eight models and the two different dataset types are shown in Figure. 6.a. The models using the X-ray dataset, such as Googlenet and Mobilenetv2, produce results that are obviously greater than those using the CT-scan dataset. The specificity values are shown in figure 6.b in a similar way. For every model, there are significant discrepancies between the two datasets. Figure 6.c, 6.d and 6.e show the values for precision, accuracy, and f1-score, respectively. The metrics for the eight models with the X-ray dataset are greater than those for the models with the CT-scan dataset, as seen in all figures.

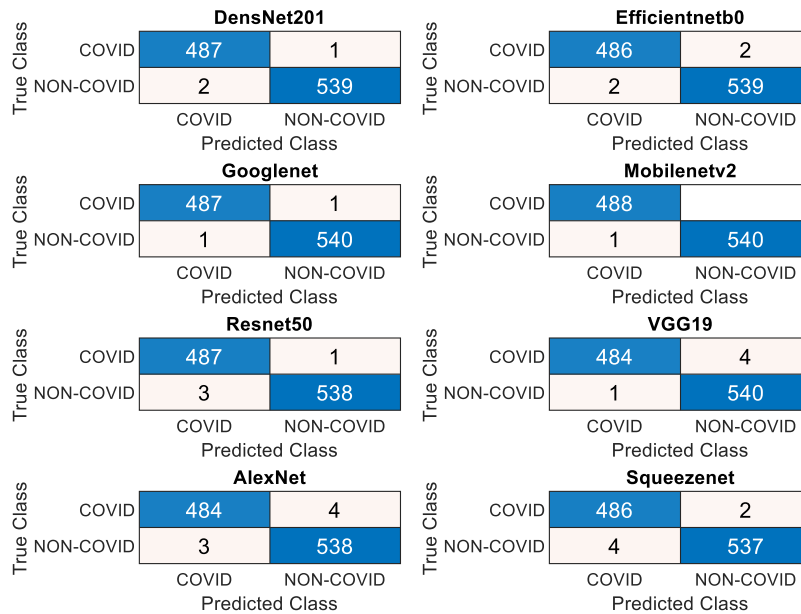


Figure (4): The Eight Networks confusion matrices based on X-ray dataset

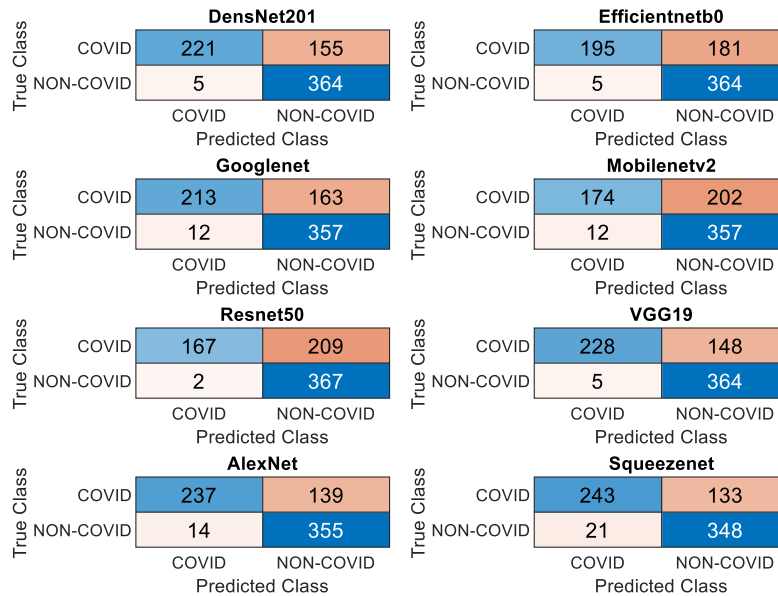


Figure (5): The Eight Networks Confusion matrices based on CT-scan dataset

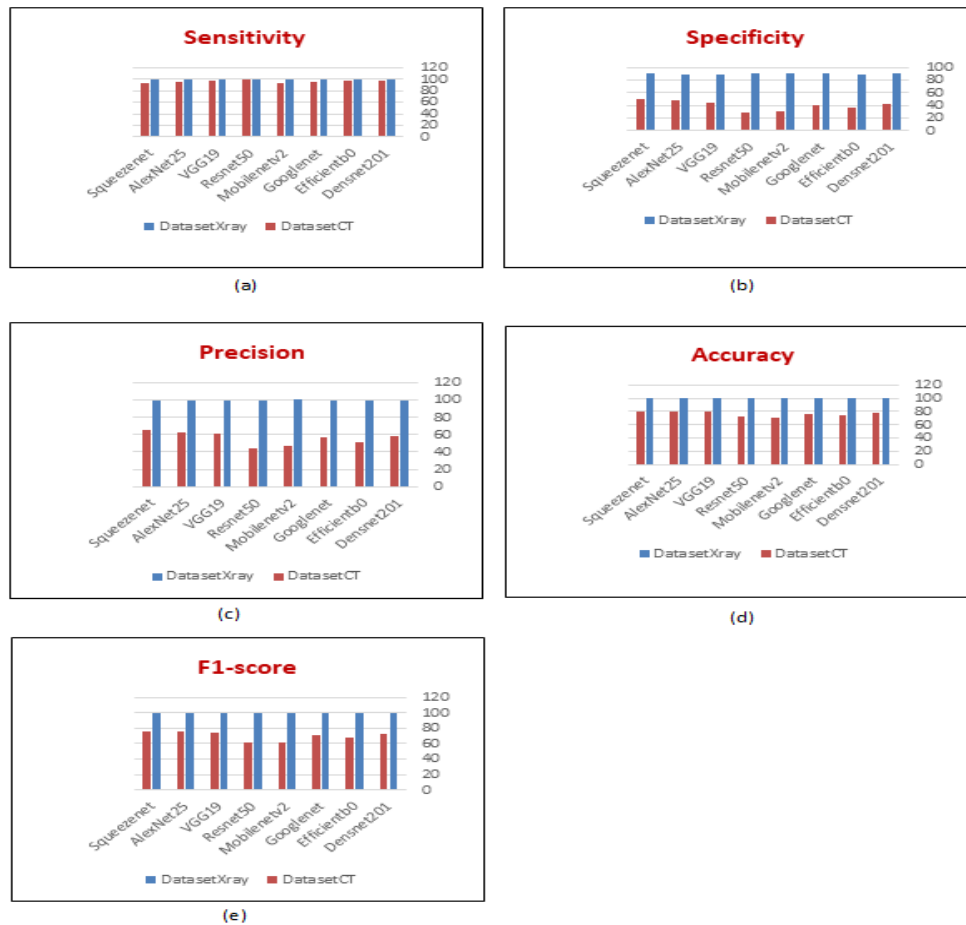


Figure (6): The performance measurement for different datasets: (a) Sensitivity, (b) Specificity, (c) Precision, (d) Accuracy, (e) F1-Score.

CONCLUSION

The radiographic techniques were used for the chest region to detect the corona virus infections in all hospitals. As a result, hospitals have created a database of patients who have had imaging tests like X-rays or CT scans. Researchers used these photos to swiftly and automatically perform the diagnosing procedure utilizing computers and deep learning. Eight different deep learning network types were used in this work, including ResNet50, MobileNetV2, DenseNet201, SqueezeNet, EfficientB0, GoogLeNet, VGG19, and Alexnet25. They were trained using a transfer-learning technique on two different types of data. For comparisons, the metrics of performance including the accuracy, F1 score, precision, specificity, and sensitivity are employed. The comparison results showed that the detection of COVID based on the CXR image dataset is better than the CT-scan dataset.



Conflict of interests

There are non-conflicts of interest.

References

- [1] L. Wang, ZQ. 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* 10, no. 1, pp. 1-12, Nov. 2020.
- [2] A. Saxena, and S. P. Singh. "A Deep Learning Approach for the Detection of COVID-19 from Chest X-Ray Images using Convolutional Neural Networks", *arXiv preprint arXiv:2201.09952*, Jan. 2022.
- [3] A. H. Ahmed, M. N. Al-Hamadani, and I. A. Satam. "Prediction of COVID-19 disease severity using machine learning techniques", *Bulletin of Electrical Engineering and Informatics*, Vol. 11, no. 2, pp. 1069-1074, Apr. 2022.
- [4] M. Alagarsamy, K. Anbalagan, Y. Thangavel, J. Sakkarai, J. Pauliah, and K. Suriyan. "Classification of covid patient image dataset using modified deep convolutional neural network system", *Bulletin of Electrical Engineering and Informatics*, Vol. 11, no. 4, pp. 2273-2279, Aug. 2022.
- [5] H. A. Owida, H. S. Migdadi, O. S. M. Hemied, N. F. F. Alshdaifat, S. F. A. Abuowaida, and R. S. Alkhaldeh. "Deep learning algorithms to improve COVID-19 classification based on CT images", *Bulletin of Electrical Engineering and Informatics*, Vol. 11, no. 5, pp. 2876-2885, Dec. 2022.
- [6] E. El. 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*, Mar. 2020.
- [7] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai et al. "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)", *European radiology*, Vol. 31, no. 8, pp. 6096-6104, Aug. 2021.
- [8] E. Ayan, and H. M. Ünver. "Diagnosis of pneumonia from chest X-ray images using deep learning", In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, IEEE, 2019.
- [9] A. Narin, C. Kaya, and Z. Pamuk. "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks", *arXiv preprint arXiv:2003.10849*, 2003.
- [10] E. Goldstein, D. Keidar, D. Yaron, Y. Shachar, A. Blass, L. Charbinsky, and I. Aharony. "Covid-19 classification of x-ray images using deep neural networks", *arXiv preprint arXiv:2010.01362*, 2020.
- [11] S. Asif, Y. Wenhui, H. Jin, and S. Jinhai. "Classification of COVID-19 from chest X-ray images using deep convolutional neural network", In *2020 IEEE 6th international conference on computer and communications (ICCC)*, IEEE, 2020.
- [12] H. Maghded, K. Z. Ghafoor, A. S. Sadiq, K. Curran, D. B. Rawat, and K. Rabie. "A novel AI-enabled framework to diagnose coronavirus COVID-19 using smartphone embedded sensors: design study", In *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, IEEE, 2020.
- [13] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio, and S. K. Antani. "Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays", *IEEE Access*, Vol. 8, pp.115041-115050, Jun. 2020.
- [14] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases", In *Proceedings of the IEEE conference on computer vision and*



- pattern recognition*, IEEE, 2017.
- [15] P. Lakhani, and B. Sundaram. "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks", *Radiology*, Vol. 284, no. 2, pp. 574-582, Aug. 2017.
- [16] P. K. Sethy, and S. K. Behera. "Detection of coronavirus disease (covid-19) based on deep features and Support Vector Machine", *International Journal of Mathematical, Engineering and Management Sciences*, Apr. 2020.
- [17] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi. "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning", *Medical image analysis*, Vol. 65, pp. 101794, Oct. 2020.
- [18] 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 Recognition Letters*, Vol. 138, pp. 638-643, Oct. 2020.
- [19] M. K. Pandit, S. A. Banday, R. Naaz, and M. A. Chishti. "Automatic detection of COVID-19 from chest radiographs using deep learning", *Radiography*, Vol. 27, no. 2, pp. 483-489, May 2021.
- [20] 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, pp. 114-119, Apr. 2020.
- [21] A. M. Ismael, and A. Şengür. "Deep learning approaches for COVID-19 detection based on chest X-ray images", *Expert Systems with Applications*, Vol. 164, pp. 114054, Feb. 2021.
- [22] F. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size", *arXiv preprint arXiv:1602.07360*, Feb. 2016.
- [23] E. Soares, and; P. Angelov. <i>SARS-COV-2 Ct-Scan Dataset</i> [Dataset]. Kaggle. <https://doi.org/10.34740/KAGGLE/DSV/1199870>, 2020.
- [24] S. Kumar. COVID19+PNEUMONIA+NORMAL Chest X-Ray Image Dataset [Dataset]. Kaggle. <https://doi.org/10.34740/KAGGLE/DSV/1857760>, 2021.

الخلاصة

مقدمة:

في هذا البحث اجرينا مقارنة لعدد من شبكات التعلم العميق في كشف مرض كوفيد19 بالاعتماد على صور اشعة اكس وصور المسح المقطعي.

طريقة العمل:

تم فحص ثمانية أنواع مختلفة من شبكات التعلم العميق (ResNet50، و Mobilenetv2، و Densnet201، و SqueezeNet، و Efficientb0، و Googlenet، و VGG19، و Alexnet25) للكشف عن COVID-19 بناءً على صور الأشعة السينية والمسح المقطعي. تم اعتماد تقنية نقل التعلم في مرحلة التدريب لجميع الشبكات التي تحتوي على مجموعات بيانات مختلفة. تم تدريب الشبكات واختبارها على 70% و 30% من مجموعة البيانات، على التوالي.

النتائج:

تم حساب مصفوفة الارتباك في مرحلة الاختبار وحساب مقاييس التقييم بما في ذلك درجة F1 والضبط والدقة والنوعية والحساسية من مصفوفة الارتباك. توضح نتائج المقارنة أن دقة التصنيف عند استخدام مجموعة بيانات الأشعة السينية أفضل من استخدام مجموعات بيانات التصوير المقطعي. علاوة على ذلك، حقق Mobilenetv2 أفضل النتائج لمجموعات البيانات المختلفة حيث حققت دقة تجاوزت 99% بالنسبة لصور الأشعة السينية وكانت الدقة اقل من 80% بالنسبة لصور المسح المقطعي.

الاستنتاج:

الاستنتاج الذي توصلنا إليه هو أن استخدام شبكة Mobilenetv2 مع صور الأشعة السينية هو أكثر ملاءمة من غيرها.

الكلمات المفتاحية: CNN، التعلم العميق، كوفيد-19، فيروس كورونا، الأشعة السينية، الأشعة المقطعية.