Employing Transfer Learning for Diagnosing COVID-19 Disease

https://doi.org/10.3991/ijoe.v18i15.35761

Lafta Raheem Ali[1], Sabah Abdulazeez Jebur[2(✉)], Mothefer Majeed Jahefer[2], Bassam Noori Shaker[3]
[1]General Directorate of Education of Salahuddin Governorate, Salahuddin, Iraq
[2]Department of Computer Techniques Engineering, Imam Al-Kadhum College (IKC), Baghdad, Iraq
[3]Department of Computer Science, College of Computer Science and Information Technology, University of Al-Qadisiyah, Qadisiyah, Iraq
sabah.abdulazeez@alkadhum-col.edu.iq

Abstract—Corona virus's correct and accurate diagnosis is the most important reason for contributing to the treatment of this disease. Radiography is one of the simplest methods to detect virus infection. In this research, a method has been proposed that can diagnose disease based on radiography (X-ray chest) and deep learning techniques. We conducted a comparative study by using three diagnosis models; the first one was developed by using traditional CNN, while the two others are our proposed models (second and third models). The proposed models can diagnose the COVID-19 infection, normal cases, lung opacity, and Viral Pneumonia according to the four categories in the covid19 radiography dataset. The transfer learning technology had used to increase the robustness and reliability of our model, also, data augmentation was used for reducing the overfitting and to increase the accuracy of the model by scaling rotation, zooming, and translation. The third model showed higher training accuracy of 93.18% compared to the two other models that are dependent on using traditional convolution neural networks with an accuracy of 70.28% of the first model, while the accuracy of the second model that uses data augmentation with traditional convolution neural is 90.1%, while the testing accuracy models was 68.27% for the first model, 87.55% for the second model, and 86.03% for the third model.

Keywords—COVID-19, data augmentation, transfer learning, X-ray, gradient descent

1 Introduction

COVID-19 is a disease that has been declared by The World Health Organization (WHO) as a pandemic spread across overall the world [1]. These pandemics stills infect people increasingly, where the globally daily average of cases and death of it up to march-22-2022 is (470,839,745). Confirmed cases and 6,092,933 Confirmed deaths according to WHO [2]. The virus affects the lungs, which causes several types of severe pneumonia. A valid and accurate COVID-19 pneumonia diagnosis is not an easy task, since is difficult to distinguish among this pneumonia because of the similar symptoms of both COVID-19 infected patients and those who suffer from viral pneumonia, as well as the similarity of their radiological findings [3]. Detection of the virus will contribute to the restriction of the spread of the pandemic and consequently reduces the number of death cases. the cheapest and fastest tool to detect COVID-19 infection is Radiographic imaging while ensuring that the patient is exposed to less radiation compared to computed tomography (CT) [4]. Artificial intelligence (AI) techniques have been employed as a powerful tool for diagnosing COVID-19 infection by using the modalities of medical images like CT and X-ray images. Modern AI systems select carefully the algorithms to be utilized for image classification, segmentation and disease diagnosis [5]. Deep learning is a sub-branch of machine learning Which in turn is a branch of AI. In the case of medical imaging (i.e., X-ray or CT images), deep learning gives great results when used for classification, lesion detection, and segmentation [6]. In this study, a COVID-19 detection system has been proposed by using a deep learning technique based on chest X-ray (CXR) images. where three sub-models of COVID-19 detection systems have been built, they all aim to detect the disease at different times and with accuracy. In the end, the results of these three sub-models have been compared to extract the best model based on time or accuracy, or both. In this study, a COVID-19 detection system using a deep learning technique has been proposed based on chest X-ray (CXR) images. where three sub-models of COVID-19 detection systems have been built, they all aim to detect the disease at different times and with accuracy. In the end, the results of these three sub-models have been compared to extract the best model based on time or accuracy, or both. In the first sub-model, the system is based on building the CNN from scratch, while in the second, transfer learning is used where inception-v3 is used as a pre-trained network and then fine-tuned with the dataset. finally, the third sub-model is also using transfer learning but the dataset has been augmented in different scales, orientations, and other operations. The training time of the three sub-models was 37.33 minutes for the first one, 11.74 hours for the second, and 3 hours for the last one. The accuracy of the three sub-models was as follows: the first one has 70.28% training accuracy and 68.27% testing accuracy, the second one has 90.10% training accuracy and 87.55% testing accuracy, and the last one has 93.18% training accuracy and 86.03 testing accuracy.

2 Related work

Pedro Silva et al. (2020) proposed a model to diagnose COVID-19 based on CT images, their model had been suggested depending on a cross-dataset analysis and voting system. they tested the model on two huge public datasets; the SARS-CoV-2 CT-scan dataset and COVID-CT, they also used cross-dataset analysis. in the case of cross-dataset, the accuracy for training and testing of the two datasets ranged between 45.25% and 59.12% [7]. Abeer Badawi and Khalid Elgazzar (2021) proposed a high-performance deep learning method by using a huge dataset by applying augmentation to the original images dataset. they used transfer learning, where DenseNet201, VGG16, and VGG19 are used to diagnose COVID-19 with better results. they achieved a great result, where the results were 95.12% when used the three transfer learning models together for multi-class classification [8]. Parag Chatterjee et al. (2021) proposed a method using COVID-Net based on transfer learning (on ImageNet). This network architecture is integrated into a heterogeneous mix of convolution layers and then applied to the COVIDx dataset, the resulting accuracy on that dataset was 93.3% [9]. Y.Pathak et al. (2020) proposed a method based on using transfer learning to construct a COVID-19 detection model. they used cross-validation to avoid overfitting. the dataset was Chest CT images collected from [10] and [11]. They used the ResNet-50 model and the resulting accuracy was 93.01% [12]. Tatiana Chakravorti et al. (2021) proposed a deep learning system based on transfer learning to classify COVID-19 infections as positive cases. Their system is applied to binary and multi-classes. They used a collection of images (545 images) to represent positive cases, normal cases, and pneumonia. Also, pneumonia falls into two classes; bacterial and virus. the accuracy is 95% for binary classification and the classification accuracy for 4-class is 91.3% [13]. Rubina Sarki et al. (2021) developed a deep learning system for COVID-19 detection by using chest radiography. they built the CNN model from scratch, the model has been trained using a public X-Ray dataset. They also used transfer learning, the resulted accuracy with transfer learning for three classes (i.e. positive cases, negative cases, and Pneumonia) was 87.50%, while the accuracy of the built CNN from a scratch model for the same three classes is 93.75% [1]. Fatchul Arifin et al. (2021) proposed a COVID-19 detection system by using the CNN model and they have been hoping to deploy it as a mobile application. the architecture of MobileNet's Single Shot Detection has been used, where two versions of these models are used (V1 and V2). Both models have been able to detect the positive infection of COVID-19, negative cases (normal), and viral pneumonia with 93.24% accuracy. V1 model provides COVID-19 ability detection with 83.7% average accuracy, whereas the V2 model can achieve 87.5% average accuracy of COVID-19 detection [14]. Frikha Hounaida et al. (2022) proposed a method for exploiting the variations of the ST segment perceived on recordings of the ECG signal. Data from two datasets have been used, the first dataset is "physioNet" where they took 300 ECGs, while the second one (100 paper ECGs of pationts) was taken from "cardiology department of hospital X" in Tunis, these data labeled as infected and non-infected samples. They used four artificial intelligence algorithms to train and classify the data, which are Random Forest, CNN-LSTM, ANN, and Xgboost. The testing and evaluation process showed that better accuracy is getting by applying CNN-LSTM and Xgboost with a classification accuracy of 87% and 88.7% respectively [15].

3 Transfer learning

Transfer learning is the process of transferring the model with knowledge from a pre-trained model to the target model, where the target model uses the learned weights of a pre-trained model, this process provisions that the two models have similar tasks to get accurate results. these weights are obtained through training the model using large datasets such as ImageNet [16]. The learned features in the trained model are reused in the target model. With transfer learning less training data is required and the result is a model that requires minimal training time with highly improved accuracy.

Transfer learning works based on the mechanism that the model of CNN firstly trained on huge data. then, in the second phase, it is fine-tuned for training on a small dataset [17]. Many popular transfer learning models exist such as VGG, ResNet, Inception, and others. Inception-v3 is a pre-trained CNN model, trained on millions of images on the ImageNet dataset, which classifies the network into 1000 categories [18][19].

3.1 Data augmentation

Data Augmentation is the process of increasing the amount of data by using various techniques (e.g., flipping, rotation, zooming, scaling, translation, etc.) to make other copies modify the existing data and thus the training data will be increased as much as needed. Data Augmentation has a major in building an improved Deep Learning model. Data augmentation is the last stage in the preparing process of the dataset [20][21][22].

3.2 Gradient descent

Gradient descent is the most famous optimization technique that is used to minimize the cost function by calculating the gradients needed for updating the networks' parameters. Gradient descent-based backpropagation is used with deep learning models as the most popular and effective learning algorithm, where the error is adjusted through backward propagation starting from the last layer and going back to the first one. the weights are initialized either randomly or by a probability distribution. There are many optimizations and improving techniques for gradient descent such as stochastic gradient descent (SGD), Adaptive Moment Estimation (Adam), and others [20].

The SGD computes the gradient for one training sample at a time and then updates its weight. SGD can be computed by Equation 1:

$$w = w - \mu .\ \nabla E(w;x(i);y(i)) \tag{1}$$

Where $\nabla E(w;x(i);y(i))$ is the gradient error ∇E w.r.t. the weights w, for training example $\{x(i),y(i)\}$, μ is represent the learning rate.

Adam is faster than the other adaptive techniques and can be computed from Equation 2.

$$w_{t+1} = w_t - \frac{\mu}{\sqrt{\hat{v}_t}} \cdot \hat{m}_t \tag{2}$$

Where m_t and v_t which represent the values of the mean and uncentered variance, respectively, μ is the learning rate, w_t is the current weight, and w_{t+1} is the next weight (updated weight).

3.3 Overfitting

Deep learning models are highly prone to overfitting during the training phase because there is a massive number of parameters involved correlated together in a complex way. thus, the performance of the model will decrease on the test data and consequently, the generalization degraded; model performance will vary and may worsen in various applications [17]. One of the solutions to overcome the problem of overfitting is by adding dropout layers, which work to drop random neurons from the network during training. By using dropout, the model generalization will be improved and the overfitting probability reduced. Figure 1 shows the dropout [20].

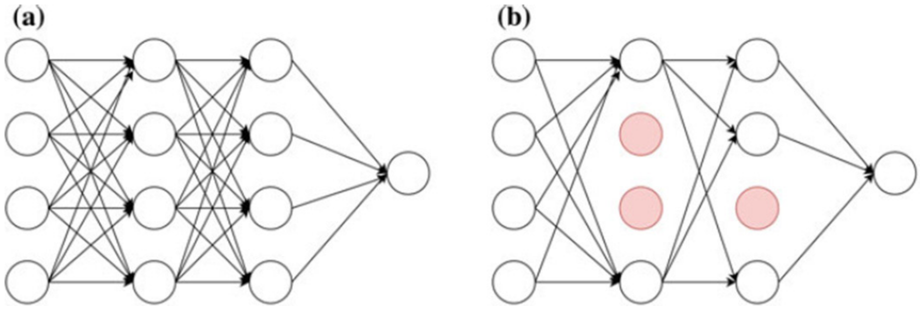


Fig. 1. (a) neural network without dropout, (b) network after dropout

4 Dataset description

The dataset had been collected by a team from both Qatar University and the University of Dhaka, Bangladesh in cooperation with medical doctors from Pakistan and Malaysia. The dataset consists of chest X-ray images that indicate the infection of COVID-19 (positive cases), negative cases (Normal), Lung opacity, and Viral Pneumonia images. The dataset was collected and released with many updates, in the first release, the images that had been detected as positive cases (COVID-19) are 219 images, the normal are 1341 and the viral pneumonia chest X-ray (CXR) images are 1345. The dataset was then updated for the first time by increasing the positive class (COVID-19) to 1200 chest X-ray (CXR) images. In the second update, the dataset had been increased as the following:

- COVID-19 positive cases images: 3616.
- Normal cases images: 10,192.
- Lung Opacity (Non-COVID lung infection) images: 6012.
- Viral Pneumonia images: 1345.

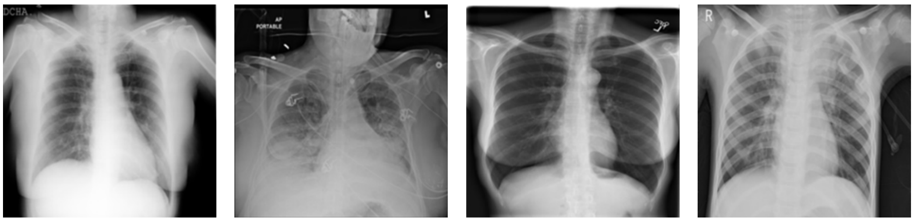
The dataset will continue to update as new CXR images are obtained for COVID-19 pneumonia patients. The dataset is found online on [23].

5 Proposed model

The system has been done in three different ways; where the dataset has been trained and tested in three models, firstly, we built our model and applied it to the dataset without any modification, while in the other two models transfer learning has been applied.

5.1 The first model

In the first model, a convolution neural network (CNN) has been used to build the network. The dataset is divided into two parts, the first part is the training set which occupies 70% of the dataset, and the remaining 30% represents the testing set. The CNN network consists of two main stages, the former is feature extraction, while the latter is classification [20]. The feature extraction stage falls into three main parts; convolutional layers, activation function layers, and pooling layers. While the fully connected layers represent the classification stage. We used rectified linear unit as an activation function in the hidden layer, while the SoftMax function is used as a classification function in the output layer. Here, SGD has been used as an optimization technique to minimize the loss function through computing the gradient necessary to update the weights of the network. Figure 2 shows images that represent the four classes in the dataset.



COVID-19 | Lung Opacity | Normal | Viral Pneumonia

Fig. 2. Samples of the source image dataset

The input is two-dimensional images at 299 × 299 (height and width), firstly the images are pre-processed by rescaling (1/255), flipped horizontally, zooming (0.2), and shearing range (0.2). secondly, the training and testing were done with a batch size of 16. The CNN model consists of six convolutional layers each one followed by an activation function (ReLU), four pooling layers, and a fully connected layer which consists of one flatten and three dense layers, the first two dense layers followed by the ReLU activation function, the last dense layer achieve the classification process by using SoftMax function. Dropout is used here four times, the first two times before flattening and the last two after it with dense layers. Figure 3 shows a diagram of this phase.

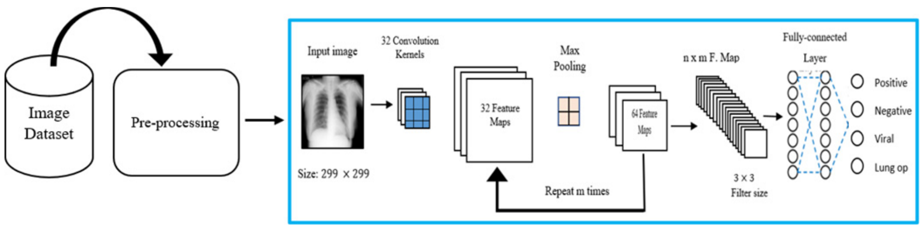


Fig. 3. Samples of the source image dataset

5.2 The second model

In the second model, we used transfer learning, where the Inception-v3 has been used as a development of CNN architecture. This network is similar to the first only differs in using transfer learning. Here the pre-trained network that was trained on Inception-v3 is used first, then the network is fine-tuned with our dataset after applying the same pre-processing steps as in the first model. During the fine-tuning phase, the last layer (classification layer) of the pre-trained network is removed, and Instead, the classification layer of the target model has been connected. The batch size is 32 for both training and testing, Adam is used as a gradient descent optimizer. Figure 4 shows the process of transfer learning.

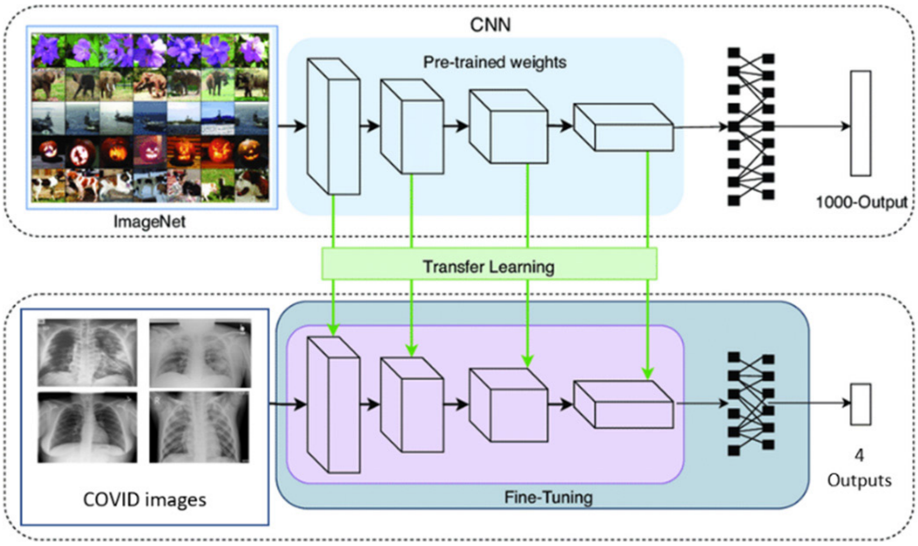


Fig. 4. Transfer learning process

5.3 The third model

In the third model, in addition to using transfer learning, we used augmentation techniques, where for each image in the dataset, many duplication images have been generated in different rotations, zooming in/out, scaling, etc. Figure 5 shows the images after augmentation.

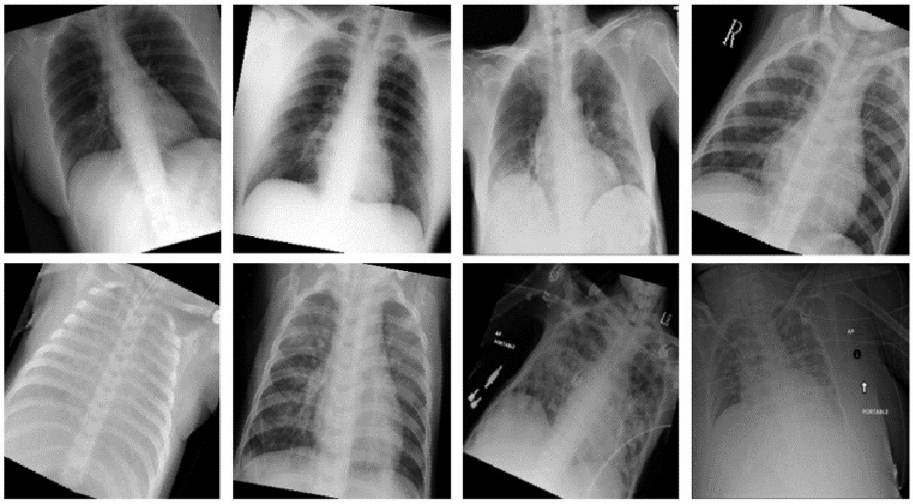


Fig. 5. Images augmentation

The dataset is duplicated several times with rotation in different angles and other augmentation techniques are also applied, thus the model during the fine tuning phase is trained with the dataset that contains the source and augmented images. Hence Adam optimizer has been used as an optimization technique to minimize the loss function through computing the gradient necessary to update the weights of the network. Figure 6 shows the model that used transfer learning that trained with augmented images.

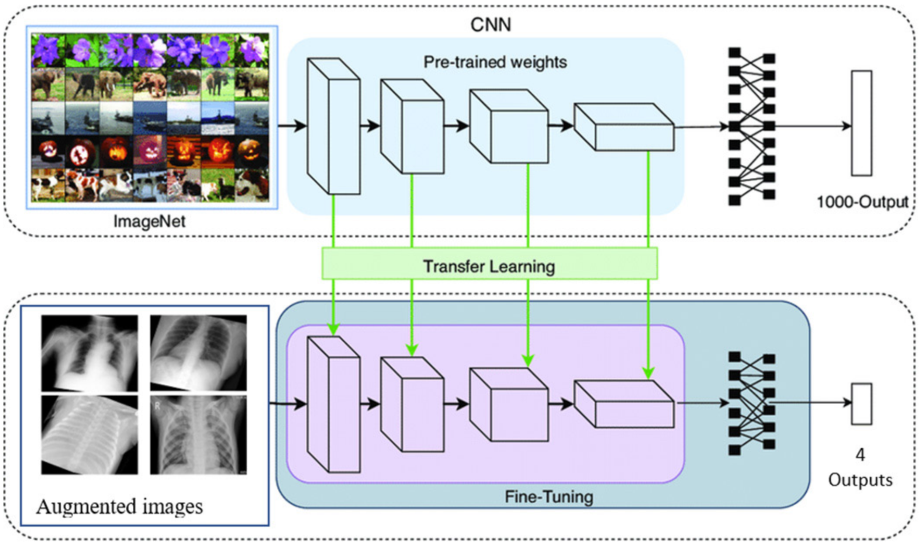


Fig. 6. Transfer learning model trained with augmented images

6 Results

The system has been evaluated by measuring the training accuracy and testing accuracy. The evaluation for the three models is as follows: In the first model, the training accuracy is 70.28% while the testing accuracy is 68.27% and because of the overfitting, we can observe that testing accuracy is lower than training accuracy. Figure 7 shows the training and testing accuracy with their loss for this model.

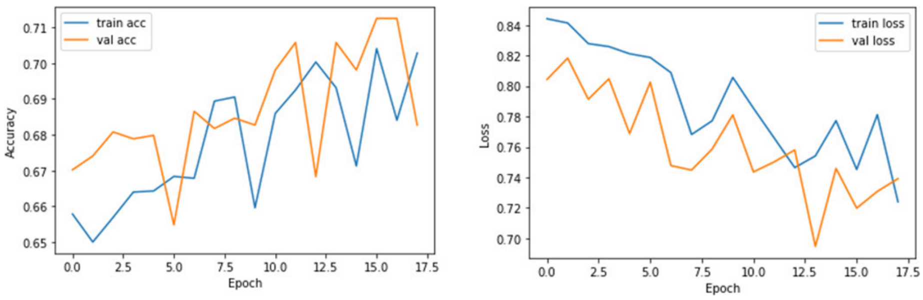


Fig. 7. Training and testing accuracy and loss for the first model

In the second model, the training accuracy is 90.10% while the testing accuracy is 87.55%, and also testing accuracy is lower than the training accuracy because of the overfitting. Figure 8 shows the training and testing accuracy with their loss for this model.

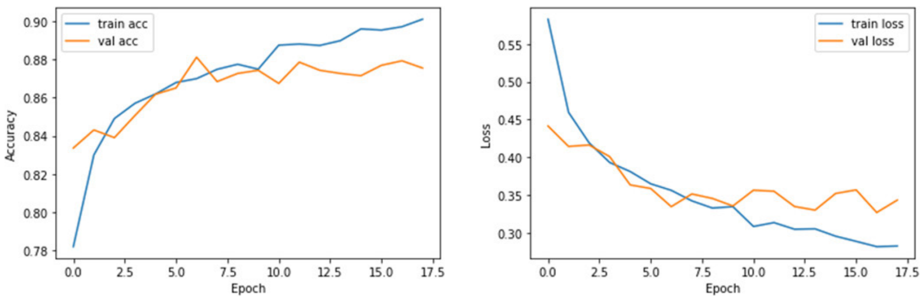


Fig. 8. Training and testing accuracy and loss for the second model

In the third model, the training accuracy is 93.18% while the testing accuracy is 86.03%, and also testing accuracy is lower than the training accuracy because of the overfitting. Figure 9 shows the training and testing accuracy with their loss for this model.

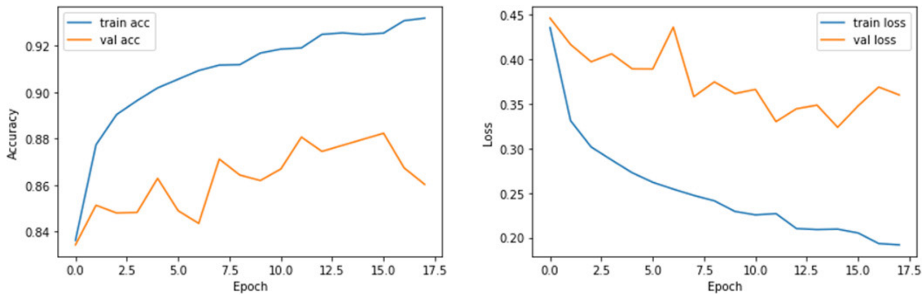


Fig. 9. Training and testing accuracy and loss for the third model

7 Conclusion

Recently, many modern applications of covid-19 detection have appeared as a result of their critical role in classifying patient lungs was infected by covid-19 and non-infected. In this proposed model, a new model of the human lung classification model is proposed based on the proposed deep transfer learning architecture, which occurred in two stages in two phases. The first phase, pre-processing the X-ray chest images dataset to acquire the lungs area. In the second phase of training and validation, pre-processed X-ray chest images dataset by using proposed model this stage is done by three cases classification are unaffected by covid-19, vaccinated and unvaccinated by covid-19 vaccine and then validated training X-ray chest images dataset. Finally, it's possible to improve the new architecture to increase the accuracy of the classification.

8 References

[1] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, and K. Wang, "Automated detection of COVID-19 through convolutional neural network using chest x-ray images," *PLoS One*, vol. 17, no. 1 January, pp. 1–26, 2022, https://doi.org/10.1371/journal.pone.0262052

[2] "Coronavirus Disease," *Coronavirus Disease (COVID-19) Pandemic*, 2022. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[3] M. A. and S. E. Petersen, "Lightweight neural network for COVID-19 detection from chest X-ray images implemented on an embedded system," *Front. Artif. Intell.*, vol. 4, 2021, https://doi.org/10.3390/technologies10020037

[4] D. M. Elsheakh et al., "Rapid detection of coronavirus (Covid-19) using microwave immunosensor cavity resonator," *Sensors*, vol. 21, no. 7021, pp. 1–14, 2021, https://doi.org/10.3390/s21217021

[5] M. Abdulkareem and S. E. Petersen, "The promise of AI in detection, diagnosis, and epidemiology for combating COVID-19: beyond the hype," *Front. Artif. Intell.*, vol. 4, no. May, pp. 1–18, 2021, https://doi.org/10.3389/frai.2021.652669

[6] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Anal. Appl.*, vol. 24, pp. 1207–1220, 2021, https://doi.org/10.1007/s10044-021-00984-y

[7] D. M. Pedro Silva, Eduardo Luz, Guilherme Silva, Gladston Moreira, Rodrigo Silva, Diego Lucio, "COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis," *Informatics Med. Unlocked*, vol. 20, 2020, https://doi.org/10.1016/j.imu.2020.100427
[8] A. Badawi and K. Elgazzar, "Detecting coronavirus from chest X-rays using transfer learning," *Covid*, vol. 1, no. 1, pp. 403–415, 2021, https://doi.org/10.3390/covid1010034
[9] P. Chatterjee, M. Biswas, and A. K. Das, "Specialized COVID-19 detection techniques with machine learning," in *Journal of Physics: Conference Series, IOP Puplishing*, 2021, vol. 1797, no. 1, https://doi.org/10.1088/1742-6596/1797/1/012033
[10] M. E. H. Chowdhury et al., "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020, https://doi.org/10.1109/ACCESS.2020.3010287
[11] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 7, pp. 1379–1389, 2020, https://doi.org/10.1007/s10096-020-03901-z
[12] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, and S. Singh, "Deep transfer learning based classification model for COVID-19 disease," *Irbm*, vol. 1, pp. 1–6, 2020, https://doi.org/10.1016/j.irbm.2020.05.003
[13] T. Chakravorti, V. K. Addala, and J. S. Verma, "Detection and classification of COVID 19 using convolutional neural network from chest X-ray images," in *2021 6th International Conference for Convergence in Technology, I2CT 2021*, 2021, pp. 2–7, https://doi.org/10.1109/I2CT51068.2021.9418221
[14] F. Arifin, H. Artanto, Nurhasanah, and T. S. Gunawan, "Fast COVID-19 detection of chest X-Ray images using single shot detection MobileNet convolutional neural networks," *J. Southwest Jiaotong Univ.*, vol. 56, no. 2, pp. 235–248, 2021, https://doi.org/10.35741/issn.0258-2724.56.2.19
[15] F. Hounaida, O. Fokapu, C.-A. Larbi, M. M. Amel, and Z. Faouzi, "ST-based deep learning analysis of COVID-19 patients," *Int. J. Biol. Biomed. Eng.*, vol. 16, no. July, pp. 321–329, 2022, https://doi.org/10.46300/91011.2022.16.39
[16] L. Alzubaidi et al., "Novel transfer learning approach for medical imaging with limited labeled data," *Cancers (Basel).*, vol. 13, no. 7, pp. 1–22, 2021, https://doi.org/10.3390/cancers13071590
[17] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, 2021, https://doi.org/10.1186/s40537-021-00444-8
[18] M. Madanan and B. T. Sayed, "Designing a deep learning hybrid using CNN and inception V3 transfer learning to detect the aggression level of deep obsessive compulsive disorder in children," *Int. J. Biol. Biomed. Eng.*, vol. 16, no. February, pp. 207–220, 2022, https://doi.org/10.46300/91011.2022.16.27
[19] M. Yani, B. Irawan, and C. Setiningsih, "Application of transfer learning using convolutional neural network method for early detection of Terry's Nail," in *International Conference on Electronics Representation and Algorithm (ICERA)*, 2019, vol. Conf. Seri, no. 1, https://doi.org/10.1088/1742-6596/1201/1/012052
[20] M. A. Wani, F. A. Bhat, Khan, S. Afzal, A. Iqbal, and E. Al., "Basics of supervised deep learning," in *Advances in Deep Learning*, vol. 57, Springer Nature., 2020, pp. 13–30, https://doi.org/10.1007/978-981-13-6794-6_2

[21] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, 2019, https://doi.org/10.1186/s40537-019-0197-0
[22] B. Jabir, N. Falih, and K. Rahmani, "Accuracy and efficiency comparison of object detection open-source models," *Int. J. online Biomed. Eng.*, vol. 17, no. 5, pp. 165–184, 2021, https://doi.org/10.3991/ijoe.v17i05.21833
[23] "https://www.kaggle.com/tawsifurrahman/covid19-radiography-database." [Online]. Available: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

9 Authors

Lafta Raheem Ali received the BSc degree in computer science from Al Rafidain University College, he received the degree of higher diploma in Information Technology and MSc degree in Computer Science in 2013 and 2018, respectively from Iraqi Commission for Computers & Informatics. Now, he is a Ph.D. candidate in the department of computer science, University of Technology. His research interests in Artificial Intelligence, Image Processing, Pattern Recognition and Computer Networks.

Sabah Abdulazeez Jebur received the BSc degree in Computer Science from Mustansiriyah University, Baghdad, Iraq in 2003. He completed his MSc study at Middle East University (MEU), Amman, Jordan in 2015 with a thesis title Enhanced Solutions for Misuse Network Intrusion Detection System using SGA and SSGA. Presently he is a PhD student of computer science in University of Technology, Baghdad, Iraq. His research interests include Artificial Intelligence and Image Processing.

Mothefer Majeed Jahefer received the BSc degrees in Computer Sciences in 2005 from the Baghdad College of Economic Sciences University (Iraq) and MSc degrees in Information technology in 2017 form University of Lodz (Poland). his research interests are IoT and IoNT. Data structure algorithms, database analyses, networks.

Bassam Noori Shaker received the BSc degree in computer science from Baghdad University; he received the degree of higher diploma in multimedia Technology and MSc degree in Computer Science in 2013 and 2018, respectively from Iraqi Commission for Computers & Informatics. Now, he is a Ph.D. candidate in the department of computer science, University of Kufa. His research interests in Artificial Intelligence, computer network and Networks security.

Article submitted 2022-09-10. Resubmitted 2022-10-11. Final acceptance 2022-10-12. Final version published as submitted by the authors.