Gezira J. of Engineering and applied Sciences
ISSN Linking: 1858-5698
http://journals.uofg.edu.sd/index.php/gjeas/

Classification of COVID-19 in Chest X-Ray Images using Deep Transfer Learning

Roaa Abdalrhman[1], Murtada Elbashir[2], Gais Alhadi Babikir*[3]

[1] *Department of Computer Science, University of Gezira, Wad Madani, Sudan, roaavisions123@gmail.com*
[2] *Department of Computer Science, University of Gezira, Wad Madani, Sudan, murtadabashir@yahoo.com*
[3] *Department of Computer Science, University of Gezira, Wad Madani, Sudan, Gais.alhadi@uofg.edu.sd*

INFORMATION:

Submission: **13/12/2021**

Accepted: **11/9/2022**

Publication: **17/10/2022**

ABSTRACT

The coronavirus emerged in Wuhan, China, and has become a critical public health problem around the world. In fact, the coronavirus caused a devastating effect on daily lives, public health, the global economy and still threatening the lives of billions of people. Therefore, a fast and accurate method will be needed for diagnosing COVID-19 infection to prevent the spread of the disease and to quickly treat affected patients. In this paper, we used a deep transfer learning approach to classify chest X-ray images of COVID-19. Clearly, we proposed a deep learning model for classifying covid-19 chest X-ray images into six classes. However, the main challenges are there is no large enough covid-19 dataset in the public domain compared to other classes. Hence, it is not easy to distinguish the similarities between categories and detailed features. Therefore, to counteract the problem of insufficient annotated images of covid-19 compatible with other classes, transfer learning is used which is also an effective deep feature extractor to extract similarity features between these classes. In fact, we trained three pre-training models [RestNet50, MobileNet, ResNet101] to classify covid-19 X-ray images into six classes. The experimental results showed the validity and efficiency of our proposed model which exceeds all proposed models in the literature.

KEYWORDS

Deep Learning, Convolutional Neural Networks, COVID-19, Coronavirus, Transfer Learning, Chest X-ray images

1. INTRODUCTION

In December 2019 novel coronavirus appeared in Wuhan China [1], and has become a critical public health problem worldwide. Severe acute respiratory syndrome coronavirus 2 is the virus that causes the corona pandemic [2]. Therefore, this virus was named SARS-CoV-2, and recently the World Health Organization (W.H.O) called this virus Coronavirus disease 2019 (COVID-19). It should be noted that fever, dyspnea, cough, myalgia, and headache are the most common symptoms of COVID-19 [1]. Transmission of the virus via small droplets produced by coughing, sneezing, and talking led to the rapid spread of the virus [3]. The total number of people infected with COVID-19 worldwide is 177,842,616 whereas the numbers of reported deaths and recoveries are 3,849,768 and 162,350,700 respectively as of June 17, 2021. The disease has spread widely throughout the world and has become an international concern. Therefore, a rapid and accurate method of diagnosing COVID-19 infection is critical. A diagnosis of COVID-19 is reported by the Chinese Government which has been confirmed through real-time polymerase chain reaction (RT-PCR). However, RT-PCR has high false-negative rates, time-consuming and limited availability in some countries [5]. For detecting the disease at an early stage and instantly quarantining the infected people due to the unavailability of specific drugs for COVID-19, an efficient method is needed. Chest x-ray (CXR) has proved it’s efficient and has been used as an alternative tool to detect the infection caused by Covid-19 [19]. Additionally, Due to the exponential spread of the COVID-19 disease, lack of experts, and automatic detection from the CXR image, A Computer-Aided Diagnosis (CAD) method is required. CAD achieves an excellent diagnosis in a short period of time [6]. However, the accuracy of this approach depends highly on the method used for feature extraction which is a hand-crafted features extractor from the image using feature descriptors [7]. For this reason, we can use the deep learning (DL) approach, in which deep convolutional neural networks (CNNs) are used to automatically extract a mass feature. In fact, DL has gained tremendous research interest due to its excellent ability to learn underlying patterns and features from image datasets and subsequently make predictions on new and unseen data. It should be noted that CNNs are the most favorite and popular deep learning models with superior achievements in the medical imaging domain since it provides high accuracy and impressive results compared with other models [8]. It is worth mentioning that in order to choose the best CNN models for our work, we had to choose between several CNNs that have been applied for CXR. As a result, we chose the ResNet50, mobileNet, and ResNet101 models which achieved state-of-the-art results. Noteworthy, Residual Networks (ResNet) is a convolutional neural network that is trained on more than a million images from the ImageNet database that can classify images into 1000 object categories. The basic idea behind ResNet is to skip blocks of convolutional layers using shortcut connections [17]. Depending on number of neuron network layers, ResNet has a lot of architecture (for example, ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-110, and ResNet-152.). In this paper, we will use Resnet50, which consists of five phases, each with convolution and an identity block, and each block has three convolution layers (See Figure 1). It is noteworthy that Resnet50 has more than 23 million trainable parameters, and Resnet101, which has the same architecture as Resnet50 but is deeper.

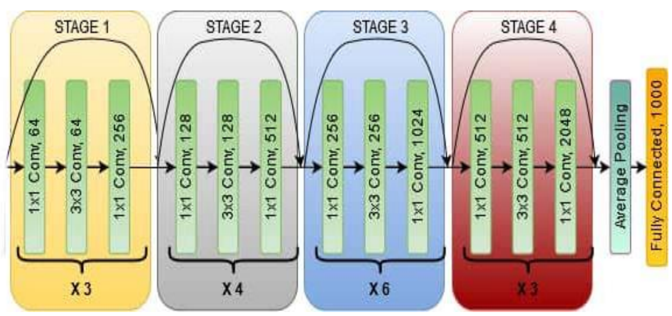

Figure.1 Architectural of Resnet50 model

MobileNets is an efficient and portable CNN architecture that is used in real-world applications. MobileNets is based on a streamlined architecture that uses depthwise separable convolutions to build a lightweight deep neural network. Furthermore, as a lightweight deep neural network, MobileNet has fewer parameters, multiplications, and additions operations. It is worth mentioning that MobileNet has 14 blocks each block consists of a depthwise separable convolution layer and each depthwise separable convolution layer consists of a depthwise convolution and a pointwise convolution.

Although there is a lot of well-proven works on the diagnosis of covid-19 in the binary classification task (covid-19, non-covid-19) [20 - 22]. Getting high accuracy in classifying more than two classes remains a challenge, due to the fact that there is no large enough dataset for covid-19 in the public domain compared to other categories, and also the high similarity between the classes and detailed features are not easily distinguishable [23]. Therefore, in this paper, we proposed a deep learning model based on the use of pre-trained convolutional neural networks which can accurately classify classes that have similarities in features to achieve higher performance in a multi-classes classification task based on CXR images.

The rest of this paper is organized as follows: the next Section 2 presents related works. In Section 3, we provide the materials and methodology. Section 4, presents the main results and discussion. Finally, Section 5 concludes this paper.

2. RELATED WORKS

Given the aim of this work, we recall some representative and related existing works to classify covid-19 from CXRs. For example, Apostolopoulos et al. [12] proposed a method for diagnosing (Covid-19, pneumonia, normal) using transfer learning and the VGG19 model. They observed that the customized VGG16 model achieves an average accuracy of 93.48 and recall of 92.85. In [11], the study was carried out using a chest X-ray image dataset. The DarkNet model was used as a classifier for the YOLO (You Only Look Once) real-time object detection system. Furthermore, the authors applied 17 convolutional layers and introduced a different filter on each layer to classify the images as COVID vs. No-Findings vs. Pneumonia. They showed that the average accuracy of the model was 87.02%. In [8], to detect the classification of COVID-19 chest X-ray images, a deep CNN model was used to transfer learning. The authors used an ImageNet pre-trained ResNet18 model and freeze the weights of the low-level layers and update the weights of the high-level layers. Moreover, various techniques are applied to generate more samples. They also showed that the proposed model obtained an average accuracy of 95.12%. In [13], the authors present COVID‐Net which is one of the first open source networks for COVID‐19 detection, Clearly, a deep convolutional neural network has been designed to detect COVID‐19 cases from chest X-rays (CXR). The proposed model obtained an average accuracy of 93.30%. A deep learning approach based on the pre-trained AlexNet model was used in [9]. Also, CXR scans images were used to classify COVID-19, non-COVID-19 viral pneumonia, bacterial pneumonia, and normal. The authors showed that the model was able to achieve 93.42% accuracy and 89.18% sensitivity. In [10] based on the deep learning model, an improved Snapshot Ensemble technology is proposed for classifying COVID-19 CXR. In fact, the architecture of (ResNet50) was implemented through a transfer learning approach, and different techniques have been used to augment the data. The experimental results showed that the proposed model achieved a precision of 95.23%, 95.63% recall, 95.42% f1-score, and accuracy of 95.18%. 1n [4] a concatenated neural network based on Xception and ResNet50V2 networks was used to classify chest X-ray images into three categories of normal, pneumonia, and COVID-19. The proposed model obtained an average accuracy of 91.4%.

From the presented related works, we can conclude that obtaining high accuracy in the classification of more than two classes remains a major challenge. Hence, the comparison of our proposed model and the state-of-the-art proved the validity and efficiency of our proposed model which exceeds all proposed models' results and demonstrated its robustness in coping with the limited availability of training data. In fact, in the multiclass classification, there are many studies that have classified covid19 into three or four classes, but there are many diseases that have a lot of similarities in features. Therefore, in this paper, we have expanded the classes into six classes.

3. MATERIALS AND METHODS

In this section, the multi-class classification method will be introduced (See Figure 2). In the rest of this Section, we describe the main datasets, the employed deep learning techniques, as well as performance evaluation of the proposed model.

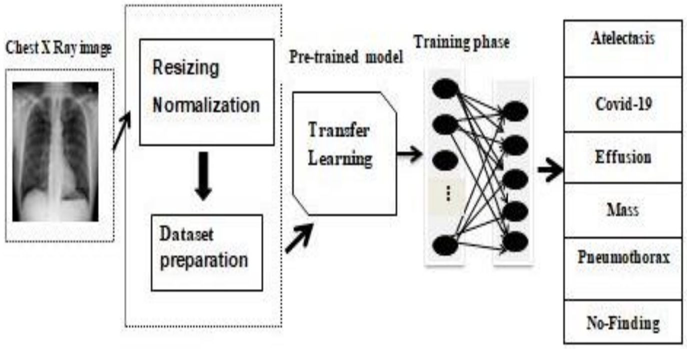

Figure.2 Schematic representation of pre-trained models for detection multi-class classification

A. Data Collection

In this paper, we used a combination of two datasets, described as follows. The first is a hospital-scale chest X-ray image database namely "ChestX-ray8" and COVID-19 dataset [16]. Chest X-ray images of COVID-19 patients have been obtained from the GitHub repository (shared by Dr. Joseph Cohen) [15]. The dataset includes more than 900 X-ray images, of which 500 were infected with the Covid-19 virus, and the rest of the X-ray images are infected with other diseases such as MERS, SARS, and ARDS. It should be noted that most of the images are from males (about 60/40% of males and females, respectively) and they are between 50 and 80 years old. The second chest X-ray8 dataset contains 108,948 X-ray images obtained from 32,717 patients in the years 1992 to 2015 diagnosed with eight common diseases. It should be noted that all the images are Portable Graphics Format (PNG) and have a size of 1024×1024.

From the two mentioned sets, we created a new dataset consisting of six classes, namely: Effusion, Covid19, Pneumothorax, Atelectasis, Mass, no-finding. In the new dataset, there are 500 cases of COVID-19 class, the rest of the classes have more cases compared to covid-19 cases. To avoid unbalancing of data a random selection of 500 cases was used to construct the Effusion, Pneumothorax, Atelectasis, Mass, no-finding cases. Thus, the final total of the new dataset became 3000 images. In Table 1 we provide a description of the newly created dataset, and in Figure 3 samples of the dataset are depicted.

Table.1 Description of the created dataset

Class	No. of Cases
Effusion	500
Covid19	500
Pneumothorax	500
Atelectasis	500
Mass	500
No-finding	500
Total	**3000**

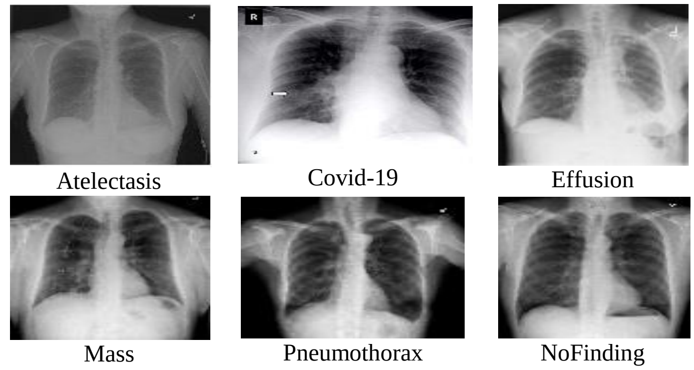

Figure.3 Representative sample chest X-ray images from different classes

B. Pre- Processing

To use the pre-training model, we need to make all images the same size to make them compatible with the input size (of CNN Pre-Trained model), so, all images were resized to 224×224. Then, a converting of any greyscale images to RGB images. Finally, data normalization was applied by dividing each pixel by the maximum value of 255.

C. Dataset Preparation

In this phase, a one-hot encoding will be applied to the labels of image data to indicate the cases of the classes. The dataset was randomly divided into two independent datasets for training and testing. As cross-validation method, k-fold was chosen and the results were obtained according to 10 different k values (See Figure 4).

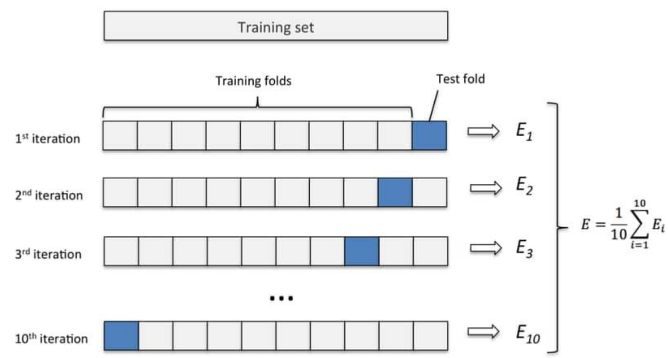

Figure.4 A schematic illustration of K-fold cross-validation for K = 10

D. Transfer learning

The idea behind this approach is to use the CNN models, which have already been trained on the ImageNet database. In the proposed model, the higher layers are fine-tuned to our specific domain. A flat layer is added that converts the 2D feature map into a 1D feature vector. A dense layer with 64 nodes is added, which is a part of a fully connected layer, in which the various features of the pre-trained model are converted to provide an output of 64 nodes. Then, a dropout layer with a rate of 0.5 is added to reduce over-fitting, followed by the dense output layer with six nodes, which correspond to six different CXR classes. It should be noted that all models were trained using the same parameters that were determined after several experiments as follows: optimizer, adm; initial learning rate, 0.000001; batch size, 32; softmax as activation function; and Categorical Cross-entropy as the loss function, which calculates the average difference between the actual and predicted probability distributions for all classes in the problem, and can be illustrated by the following equation:

$$E\ (y_j,a_j) = -\frac{1}{n}\sum_x \sum_j [\, y_j \ln a_j + (1-y_j)\ln(1-a_j)] \quad (1)$$

Where y_i is the truth label while a_j is the predicted output from the softmax function.

E. Performance Evaluation of the Model

In this phase, performance evaluation will be performed using the following metrics: Accuracy (ACC), Recall (or) Sensitivity (SE), Precision, f1-score (or) F-measure, and the area under the receiver operating characteristic (ROC) curve (AUC). Hence, to evaluate the above-mentioned metrics, we have to calculate the following metrics for the test rating.

1. True Positive (TP): Refers to the number of predictions where the classifier correctly predicts the positive category as positive.
2. True Negative (TN): Refers to the number of predictions where the classifier correctly predicts the negative category as negative.
3. False Positive (FP): Refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.
4. False Negative (FN): Refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

The equation of Accuracy, Sensitivity, Precision, and F1-score will be described as follows [18]:
Accuracy is part of the correct predictions obtained by the model, and it can be calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (2)$$

The proportion of actual positive samples given is correctly calculated by sensitivity measures as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP+FN})} \quad (3)$$

Precision is a fraction of the predictions as the positive class was actually positive, and it can be calculated as follows:

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (4)$$

F-score is the weighted average of Precision and Recall, which is calculated as:

$$\text{F1-score} = \frac{2*(\text{Recall}*\text{Precision})}{(\text{Recall}+\text{Precision})} \quad (5)$$

Under the receiver operating characteristic (ROC), the curve is a two-dimensional plot that illustrates how well a classier model works as the dis-crimination cut-off value is changed over the range of the predictor variable. The x-axis or independent variable is the false positive. The y-axis or dependent variable is the true positive rate.

4. RESULTS AND DISCUSSION

Python programming language and Keras packages were used to train the proposed deep transfer learning models. All experiments were performed on Kaggle. Noteworthy, Kaggle is a platform that provides notebook-based cloud services for this disseminating Knowledge and works on machine learning, it completely optimizes the running time of deep learning and free access to the stable GPU.

The prediction performance results obtained from different pre-trained CNN models are depicted in Table 2. We can see that the pre-trained ResNet50 model achieves the best performance of an Accuracy of 96.5%, Sensitivity of 96.4%, Precision of 96.5%, and F1-score of 96.4%. We can also see that the pre-trained mobileNet model achieves the lowest performance of an Accuracy of 94.2%, Sensitivity of 94%, Precision of 94.9%, and F1-score of 94.1%.

Table.2 Performance measurements for all models

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-Score (%)
MobileNet	94.2	94	94.9	94.1
Resnet101	95.8	95.8	95.9	95.7
Resnet50	96.5	96.4	96.5	96.4

As a result, the ResNet50 model provides superior performance compared to other models. Indeed, in Table3 we provide comparisons between our three models using Receiver Operating Characteristic (ROC) curve plots, and Area Under Curve (AUC) for each model are given in Figures 5, 6, and 7.

Table.3 Performance result of AUC for all models

Model	AUC (%)
MobileNet	96
Resnet101	97
Resnet50	98

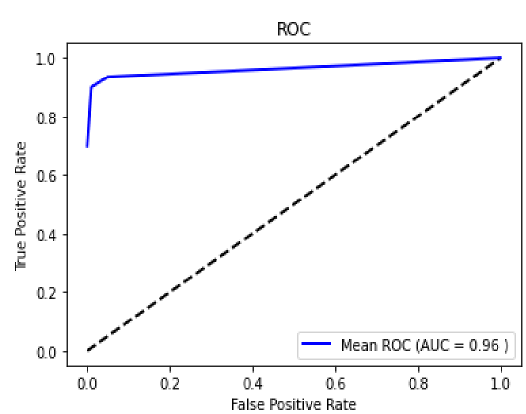

Figure 5 The AUC results and ROC curves obtained by Mobilenet

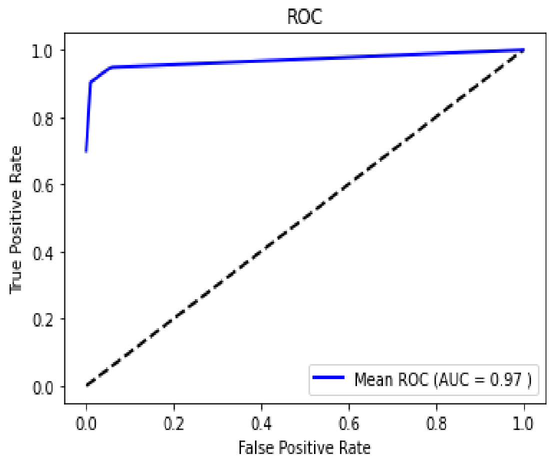

Figure.6 The AUC results and ROC curves obtained by Resne101

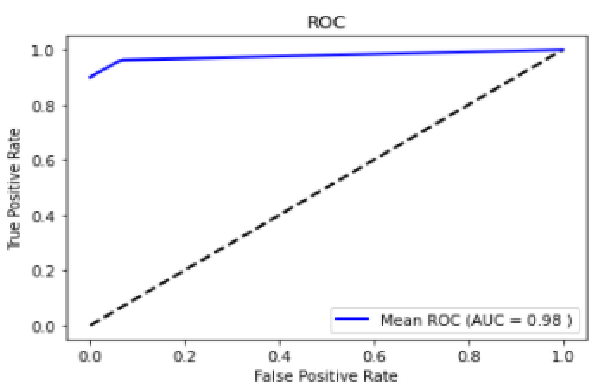

Figure.7 The AUC results and ROC curves obtained by Resnet50

Table 4 provides detailed comparisons of the proposed model with recent baseline models. All comparisons are made only for multi-class data. Although many proposed models in the literature showed effective results while classifying CXR images, our proposed model obtained the best performance (of an Accuracy of 96.5%, Sensitivity of 96.4%, Precision of 96.5%, and F1-score of 96.4%) compared to other methods applied to CXR images.

Table. 4 Comparing our results with the state-of-the-art

Study	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-Score (%)
Ref. [11]	87.02	85.35	89.96	87.37
Ref. [14]	91.40	80.53	35.27	49.05
Ref. [12]	93.48	92.80	93.20	93.00
Ref. [9]	93.42	89.18	-	-
Ref. [8]	95.12	97.91	-	
Ref. [10]	92.25 91.00 95.18	92.47 92.00 95.63	92.35 92.00 95.23	92.41 92.00 95.42
This paper	96. 50	96. 40	96.50	96.40

In fact, most of the results in the literature have unbalanced results. Therefore, to be fair, we will compare our model with the same CNN pre-training model in [10], where the authors used the same Resnet50 pre-trained model. The authors used Resnet50 to capture generic features. Then, a dropout layer and two FCL have been added. Finally, a dense layer with three nodes, which correspond to three different classes of CXR, and the ReLU function is used as an activation function. This setting has been used in three strategies: First, Resnet50 with balancing data, which achieved an Accuracy of 92%. Second, Resnet50 without balancing, which achieved an Accuracy of 91%. And Third, Resnet50 + Improved Snapshot Ensemble + Data Balance, which achieved the best Accuracy of 95%. Clearly, they divided the dataset into a training and test set using a random subsampling technique, and this method is not recommended because the result depends on the split, where a new split can give you a new result, therefore, it is very difficult to speculate that the new output will be true. Hence, in our proposed model we used a softmax as an activation function, which interprets the output activation a_j^l as the network's estimate of the probability that the correct output is j, and by using k fold cross-validation which makes our model more reliable and has great results for classifying six classes. Furthermore, for faster computation time a flatten layer has been added which converts a 2D feature map into a 1D feature vector. Importantly, this comparison validates the validity and efficiency of our proposed model which exceeds all the proposed models with an Accuracy of 96.5% and demonstrated its robustness in coping with the limited availability of training data.

5. CONCLUSION

In this paper, a deep learning model based on transfer learning is proposed to efficiently classify COVID-19 CXR images. We used two open-source datasets, all images in these datasets were resized to a size of 224×224 pixels. To efficiently extract the deep features, we used three different pre-training models [RestNet50, MobileNet, ResNet101]. The proposed models demonstrated efficient performance in classifying COVID-19, Effusion, Pneumothorax, Atelectasis, Mass, and No-finding CXR images. To prove the efficiency of our proposed model we compared our model with the state-of-the-art. The experimental results showed that our proposed model outperforms all the models in the literature. In fact, our proposed (Resnet50 model) obtained the best performance of an Accuracy of 96.5%, Sensitivity of 96.4%, Precision of 96.5%, and F1-score of 96.4%. Furthermore, the Resnet50 model provides a high AU-ROC value of 98%, and demonstrated its robustness in coping with the limited availability of training data.

In future work, we look to the use of multi-classification on chest X-ray images in patients with more than one disease.

REFERENCES

[1]. A.A. Ardakani, A.R. Kanafi, U.R. Acharya, N. Khadem, and A. Mohammad, 2020. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. Computers in biology and medicine, 121, p.103795.

[2]. A. Narin, C. Kaya, and Z. Pamuk, 2021. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Analysis and Applications, pp.1-14.

[3]. X. He, S. Wang, S. Shi, X. Chu, J. Tang, X. Liu, C. Yan, J. Zhang, and G. Ding, 2020. Benchmarking deep learning models and automated model design for covid-19 detection with chest ct scans. medRxiv, 2020 Jan 1.

[4]. Coronavirus. (2021, june 17). Retrieved from worldometer: https://www.worldometers.info/coronavirus/

[5]. D. Singh, V. Kumar, and M. Kaur, 2020. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. European Journal of Clinical Microbiology & Infectious Diseases, 39(7), pp.1379-1389.

[6]. S. Vaid, R. Kalantar, and M. Bhandari, 2020. Deep learning COVID-19 detection bias: accuracy through artificial intelligence. International Orthopaedics, 44, pp.1539-1542.

[7]. I. Allaouzi, and M.B. Ahmed, 2019. A novel approach for multi-label chest X-ray classification of common thorax diseases. IEEE Access, 7, pp.64279-64288.

[8]. A. Abbas, M.M. Abdelsamea, and M.M. Gaber, 2021. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Applied Intelligence, 51(2), pp.854-864.

[9]. A.U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P.S. Yakoi, 2021. Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive Computation, pp.1-13.

[10]. C.S.R. Annavarapu, 2021. Deep learning-based improved snapshot ensemble technique for COVID-19 chest X-ray classification. Applied Intelligence, 51(5), pp.3104-3120.

[11]. T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, and U.R. Acharya, 2020. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in biology and medicine, 121, p.103792.

[12]. I.D. Apostolopoulos, and T.A. Mpesiana, 2020. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine, 43(2), pp.635-640.

[13]. L. Wang, Z.Q. Lin, and A. Wong, 2020. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports, 10(1), pp.1-12.

[14]. M. Rahimzadeh, and A. Attar, 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics in Medicine Unlocked, 19, p.100360.

[15]. J. P. Cohe, (2020). covid-chestxray-dataset. Retrieved from ieee8023: https://github.com/ieee8023/covid-chestxray-dataset

[16]. R. Summers, (2017). NIHCC. Retrieved from ChestXray-NIHCC: https://nihcc.app.box.com/v/ChestXray-NIHCC

[17]. E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. De Geus, 2017. Malicious software classification using transfer learning of resnet-50 deep neural network. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1011-1014). IEEE.

[18]. L. Yan, H.T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, and X. Huang, 2020. An interpretable mortality prediction model for COVID-19 patients. Nature machine intelligence, 2(5), pp.283-288.

[19]. X. Qi, L.G. Brown, D.J. Foran, J. Nosher, and I. Hacihaliloglu, 2021. Chest X-ray image phase features for improved diagnosis of COVID-19 using convolutional neural network. International journal of computer assisted radiology and surgery, 16(2), pp.197-206.

[20]. A. Muguli, L. Pinto, N. Sharma, P. Krishnan, P.K. Ghosh, R. Kumar, S. Bhat, S.R. Chetupalli, S. Ganapathy, S. Ramoji, and V. Nanda, 2021. DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics. arXiv preprint arXiv:2103.09148.

[21]. A.S. Elkorany, and Z.F. Elsharkawy, 2021. COVID etection-Net: A tailored COVID-19 detection from chest radiography images using deep learning. Optik, 231, p.166405.

[22]. R.C. Joshi, S. Yadav, V.K. Pathak, H.S. Malhotra, H.V.S. Khokhar, A. Parihar, N. Kohli, D. Himanshu, R.K. Garg, M.L.B. Bhatt, and R. Kumar, 2021. A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images. Biocybernetics and Biomedical Engineering, 41(1), pp.239-254.

[23]. J. Shuja, E. Alanazi, W. Alasmary, and A. Alashaikh, 2021. COVID-19 open source data sets: a comprehensive survey. Applied Intelligence, 51(3), pp.1296-1325.