2021 International Conference on Information and Communication Technology
for Sustainable Development (ICICT4SD), 27-28 February, Dhaka

Deep Learning based Detection and Segmentation of COVID-19 & Pneumonia on Chest X-ray Image

Md. Jahid Hasan*, Md. Shahin Alom†, Md. Shikhar Ali‡

Department of Electrical and Electronic Engineering, Hajee Mohammad Danesh Science and Technology University
Dinajpur-5200, Bangladesh
jahidnoyon36@gmail.com*, ashahin200@gmail.com†, shikharali13@gmail.com‡

2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD) | 978-1-6654-1460-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICT4SD50815.2021.9396878

***Abstract*—The outbreaks of COVID-19 virus have crossed the limit to our expectation and it breaks all previous records of virus outbreaks. The effect of corona virus causes a serious illness may result in death as a consequence of substantial alveolar damage and progressive respiratory failure. Automatic detection and classification of this virus from chest X-ray image using computer vision technology can be very useful complement with respect to the less sensitive traditional process of detecting COVID-19 i.e. Reverse Transcription Polymerase Chain Reaction (RT-PCR). This automated process offers a great potential to enhance the conventional healthcare tactic for tackling COVID-19 and can mitigate the shortage of trained physicians in remote communities. Again, the segmentation of the infected regions from chest X-ray image can help the medical specialists to view insights of the affected region. So, in this paper we have used deep learning based ensemble model for the classification of COVID-19, pneumonia and normal X-ray image and for segmentation we have used DenseNet based U-Net architecture to segment the affected region. For making the ground truth mask image which is needed for segmenting purpose, we have used Amazon SageMaker Ground Truth Tool to manually crop the activation region (discriminative image regions by which CNN identify a specific class using Grad-CAM algorithm) of the X-ray image. We have found the classification accuracy 99.2% on the available X-ray dataset and 92% average accuracy from the segmentation process.**

Keywords*—*COVID-19, Pneumonia, X-ray image classification, X-ray image segmentation, Deep Learning, Computer Vision.

I. Introduction

The novel corona virus (nCOV) or COVID-19 outbreaks from Wuhan city of China to rest of the world in December 2019. Until 30th November 2020 around 62,195,274 cases are affected by this diseases and 1,453,355 cases are dead which is confirmed by World Health Organization (WHO)[1]. First, it attacks the lungs and respiratory system of a human body which is named after (SARS-COVID-19) i.e. severe acute respiratory system[2]. So, a screening mechanism of this organ like X-ray and Computer Tomography (CT) image can play an important role to recognize the diseases effectively and can be taken proper step[3]. This process can also be helpful to support the design of curative therapy with the help of Computer Vision and Deep Learning technology. Again, the segmentation of the infected region from this organ can help medical personnel's to take immediate treatment, care and isolate to mitigate the spread of the virus. Though a clinical laboratory-based approach namely Reverse Transcription Polymerase Chain Reaction (RT-PCR) has been widely used but it is low sensitive approach and give a high false negative result and also labour-intensive process. On the other hand, X-ray machine is available everywhere and the X-ray image allows further processing's such as segment the affected region that offers more insight view and doctors can take effective steps [4][5][6]. So, in this paper we have proposed a unique approach for the detection and classification of COVID-19, pneumonia and normal status from X-ray image using pretrained deep Convolution Neural Network (CNN) model as well as the segmentation of the affected part from X-ray image using U-Net architecture based DenseNet103 deep learning model. For performing segmentation of the affected part, we need ground truth mask image that contains only the affected region and all other parts are detached. But currently such dataset is not available for COVID-19 and pneumonia X-ray image because of the difficulty to locate the affected area properly. So, we have made this ground truth mask image from the highlighted region of the Grad-CAM[7] algorithm by Amazon SageMaker Ground Truth tool.

Gradient-weighted Class Activation Mapping (Grad-CAM) is an algorithm that is used to visualize the class activation maps of a deep CNN, thereby allows to verify that network is "looking" and "activating" at the correct locations. It ensures whether a CNN model is performing image classification task correctly or not. The output of Grad-CAM is a heatmap visualization for a given class label. We can use this heatmap to visually verify those features of an image that CNN activates for classification. Segmentation is a process of assigning a label to every pixel in an image such that pixels with the same label contain certain characteristics. Instance and semantic are the two favours of image segmentation. The semantic image segmentation task consists of classifying each pixel of an image into an instance, where each instance corresponds to a class.

Several approaches have been proposed for this problem. In a recent study by K. Lee et al. [8] analysed chest X-ray and CT images from nine COVID-19 infected patients by two radiologists to assess the correspondence of abnormal findings on X-rays with those on CT images. M. Mishra[9] et.al. proposed a transfer learning based pneumonia and COVID-19 detection process from chest-X-ray image of open source dataset and achieved 98.2% accuracy in COVID-19 classification. Narin et al. [10] evaluated different CNN for the diagnosis of COVID-19 and achieved an accuracy of 98% using a pre-trained ResNet50 model for single class classification. In another study by T. Ozturk et al. [11], proposed a deep learning model called DarkCovidNet for the automatic diagnosis of COVID-19 based on 125 chest X-ray images to diagnosis i) binary classification (COVID-19 vs no-findings) and ii) multiclass classification (COVID-19 vs no-findings vs pneumonia) and report accuracy 98.08% on binary and 87.02% on multiclass classification task. No segmentation task has been done yet on COVID-19 image. Considering all this issues we have classified COVID-19, pneumonia and normal X-ray image by concatenating three pretrained deep leaning model (Ensemble model) namely DenseNet121, EfficientNetB0 and VGG19. Using Gradient-based heatmap generation methods that is responsible to highlight the important region of an image, we have done the ground truth mask and segmentation task.

II. Dataset Preparation

A collection of total 701 image of COVID-19 were collected from several open source resources namely covid-chestX-ray-dataset that is collected by J. P. Cohen [12, 13] from various publications on COVID-19 topics. Till 30 October, 2020 it contained 319 X-ray images of COVID-19 patients of both AP (Anterioposterior- that is taken from front to back with respect to X-ray film) and PA (Posteroanterior) position and 132 image is collected from AP supine view. Also from a hospital in Spain (80 cases)[14]. Around 1587 normal image and 4283 pneumonia images were collected from [15]. So, total 6581 X-ray images were collected where 20% data were used for validating the proposed model. Table I shows the number of data of all categories in our dataset.

TABLE I. Number of Data in Each Class

Data		
Normal	*Pneumonia*	*COVID-19*
1587	4283	701

III. Materials and Method

In this paper for classifying the 3 types of X-ray images we have used ensemble method. During prediction, each image is passed through the classification layer where we have checked whether an image is COVID-19, pneumonia or normal. Then we have used Grad-CAM which makes the Ensemble model more transparent by visualising the regions of the input image that are important for making predictions. It identifies the important features of the input image which finally contribute in classification. After that, we have manually cropped the highlighted feature region using Amazon SageMaker Ground Truth Tool for the purpose of making ground truth mask. Once the ground truth dataset is ready, then we have trained and tested the U-Net[16] segmentation model based on DenseNet103 architecture as backend to segment the affected region of X-ray image. The following fig. 1 shows the complete process of making ground truth mask image for segmentation.

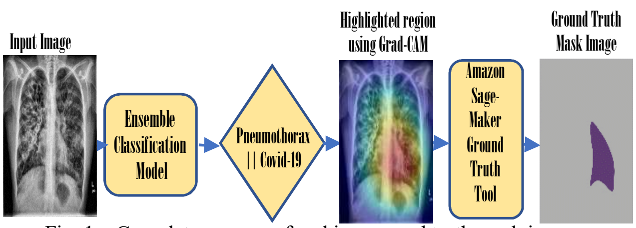

Fig. 1. Complete process of making ground truth mask image.

In the proposed system, the ensemble classification model classifies an input X-ray image and the segmentation model segment the affected region as predicted mask of that input image. Then we have superimpose the predicted mask on input image which give the final predicted segmentation image. The segmentation model will activate if the ensemble classification model detects an X-ray image as COVID-19 or pneumonia so that radiology expert can easily analyse and diagnose this problem. The overview of the proposed model has been shown in fig. 2.

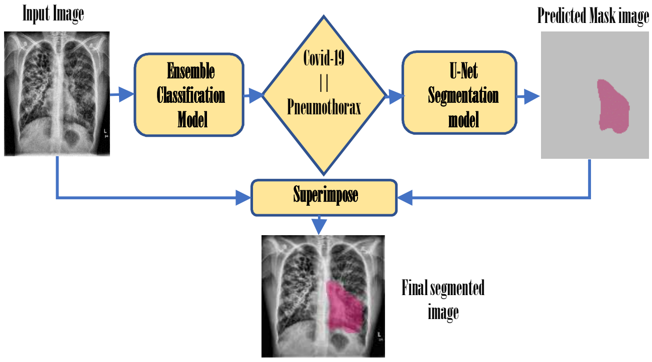

Fig. 2. Overview of the proposed model.

A. Training Process of Ensemble Classification Model using Transfer Learning:

The more a CNN model is deep the more accuracy we can gain. But when information pass from input layer until the output layer, gradient becomes so big as gradient is calculated in the opposite direction and they can get vanished before reaching the input side. That causes poor classification result. To mitigate this issue, some networks have been proposed by the researchers such as Inception V3, ResNet, MobileNet, DenseNet, EfficientNet etc. In this paper we have adopted DenseNet121[17], EfficientNetB0[18] and VGG-19[19] to make our ensemble classification model. Densely Connected Convolutional Networks (DenseNets), increases the depth of Convolution layer without vanishing gradient problem. It solves this problem by using features reuse, where output of any previous layer is concatenating with all previous layers. So, each layer has a direct access to the gradient from the loss function and original input image. This also ensures to learn all unique feature only once of the input image and reduces the total parameter size and also increases the features propagation.

EfficientNet is one of the most efficient model in terms of FLOPS(Floating Point Operation Per Second) and inference of general image classification task where a compound scaling method is used that uniformly scales all dimensions of a model attribute i.e depth (number of layers), width (number of channels) and image resolution (image size) in order to improve the accuracy in a more fine grained way. Firstly, the researchers build a new baseline network based on neural architecture search algorithm(a technique for automating the design of neural networks) that optimizes for both accuracy and FLOPS. The researchers then scale up this baseline network to obtain a family of deep learning models, called EfficientNets. In this paper we have used EfficientNet B0. The main building block of this network is mobile inverted bottleneck (MBConv) layer.

Another model VGG-19 which is the most widely used by the deep learning research community because of its simplicity and high accuracy in image classification task. It has different versions. The name of this model was inspired by the name of their research group 'Visual Geometry Group (VGG)'. As this convolutional neural network has 19 layers in its architecture, it was named VGG-19. This model was proposed to reduce the number of parameters during classifying an image. The model use a tiny receptive field or kernel which is 3x3 with stride 1 and 1x1 convolutional filters that make the classification output more nonlinear. This also allows a large number of weight layers which help to gain improved performance. Fig. 3 shows the training process of the ensemble classification model.

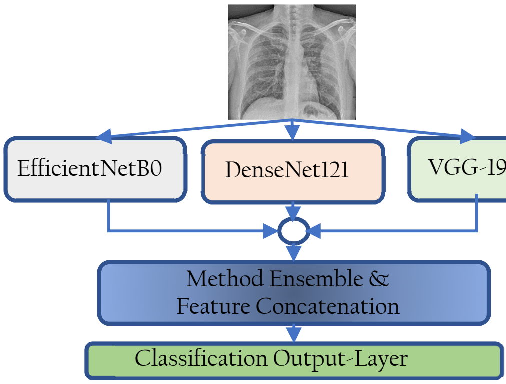

Fig. 3. X-ray image classification process using ensemble model.

B. *Important Area Visualization and Ground Truth Mask:*

During training the ensemble model for X-ray image classification, we have also generated saliency maps using gradient class activation maps (Grad-CAM). It is a visualization technique that helps to understand which features a CNN is looking for when it classifies an image by producing a heat map that highlights the important regions of an image using the class specific gradient information flowing from final convolutional layer for each image. This technique may also be useful as an approximate visual diagnosis for presentation of the radiologists. First, we have depicted the regions of the image that have a positive contribution to the predefined class of COVID-19 and pneumonia using Grad-CAM. Then the highlighted portion has been cropped using Amazon SageMaker Ground Truth Tool to create ground truth mask that is then used for training U-Net segmentation model. Fig. 4 shows the visualition process of important region using Grad-CAM and fig. 5 shows the mask image making procedure.

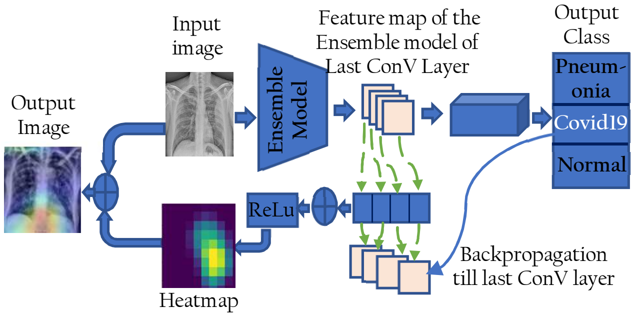

Fig. 4. Important area visualization using Grad-CAM.

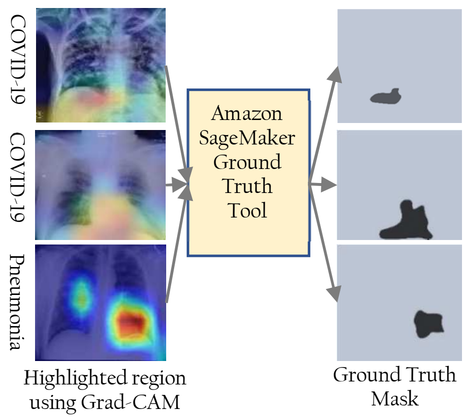

Fig. 5. Grad-CAM output to ground truth mask.

C. *Training Process of U-Net Segemntation Model:*

In image segmentation task U-Net model works very well with respect to traditional CNN model. It is an encoder-decoder based network where the encoder down-sample the image and capture the context in the image by stacking the traditional convolutional and max pooling layers. The decoder is a symmetric expanding path and up-sampler which is used to enable precised localization using transposed convolutions. It helps to understand an image at a much lower level, i.e., the pixel level and widely used in many applications such as medical imaging, self-driving cars, satellite imaging etc. The model being used here is modified U-Net that consists of an DenseNet103 model of both encoder (downsampler) and decoder (upsampler) frame. It takes the input image with ground truth mask and predicts the output mask. Fig. 6 shows the segmentation process using modified U-Net architecture.

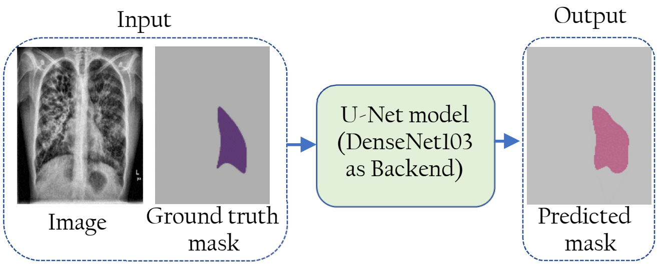

Fig. 6. U-Net based image segmentattion process.

IV. Result and Discussion

A. *Hyper-parameter Setting*

Different hyper-parameter has been tuned for both classification and segmentation. Weight initialization technique has been used to tackle the class imbalance problem. It is a process of predefining some values for each class during training time. Lower value is preferred for the class of higher number of data in a dataset. We have tested the model for different values and determined the value that gives the best classification result that is shown in table II. Again we have used pretrained deep learning model that was previously trained with ImageNet [20] dataset except the DenseNet121 that was trained with Medical X-ray images to classify around 14 disease which was related to lungs. We have ensemble them as some initial layers as non-trainable and rest of the layer as trainable. The other hyper-parameter such as optimizer, batch size is included in following table.

TABLE II. Final Values of All-User Defined Parameters of Proposed Model

Ensemble Model		**Segmentation Model**		**Class Weight**	
Parameter	***Value***	***Parameter***	***Value***	***Class-label***	***Value***
Learning rate	1xe-4	Learning rate	1xe-4	COVID-19	4283
Batch-size	30	Batch-size	30	Pneum-onia	1587
Number of epoch	60	Number of epoch	40	Normal	701

B. Experimental Setting

The proposed approach was trained and tested in python programming language with Keras deep learning framework. Before training all image are resized to specific size using OpenCV tool. Then Amazon SageMaker Ground Truth tool was used to create mask image of specific portion. All necessary package and library was installed in windows operating system with Intel(R) Core (TM) i7 CPU @3.20GHz, 16GB RAM and NVIDIA GeForce GTX-1070.

C. Performance evaluation for Classification

Accuracy is a commonly used classification metrics that indicates how well a classification algorithm can discriminate the available classes of the test set. The accuracy can be defined as the proportion of the predicted correct labels to the total number (predicted and actual) of labels. In this study, accuracy refers to the overall accuracy of the model in distinguishing the three classes (COVID-19, pneumonia, normal). We have split the dataset into 80:20 ratio. During training the ensemble model the train and validation accuracy was recorded. Fig. 7 and fig. 8 respectively shows the training and validation accuracy as well as training and validation loss with respect to number of epoch.

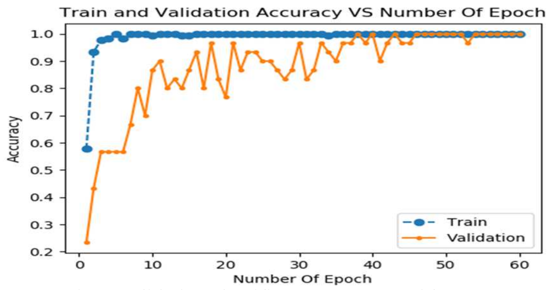

Fig. 7. Train & validation classificaton accuracy with respect to number of epoch.

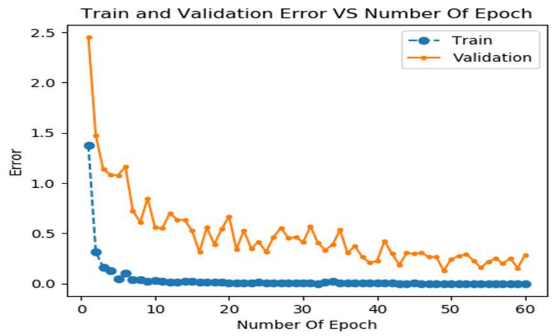

Fig. 8. Train & validation classificaton loss with respect to number of epoch.

D. Performance evaluation for Segmentation

Loss function is one of the most important ingredients in deep learning-based medical image segmentation method. We have used region based accuracy function named Dice coefficient and IoU (Intersection over Union) score which is considered as the best evaluation matrices for segmentation task in imbalance dataset. This method checks how much the output mask is similar to ground truth mask. After 40 epochs of training we were able to get 0.90% average IoU score and 0.92% Dice coefficient score which indicates that the model is very reliable and convincing for COVID-19 and pneumonia X-ray image segmentation task. The Dice coefficient score in terms of train and validation accuracy is shown in fig. 9.

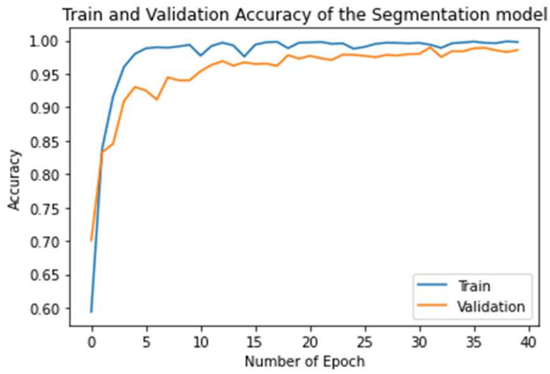

Fig. 9. Train & validation segmentation accuracy.

The proposed segmentation model was tested with the test dataset. Some of the predicted mask with superimpose image is shown in fig. 10.

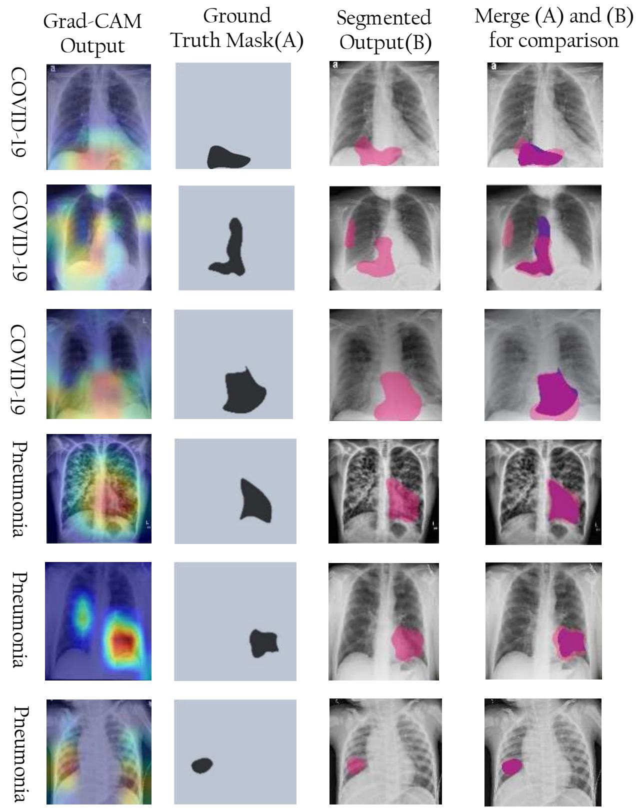

Fig. 10. Some of the segmentation output

An comparison study has been performed to gauge the execution of the proposed system with other recent works. A table has been offered underneath to analyze the highlights of the existing works with the proposed framework.

TABLE III. COMPARISON WITH EXISTING WORKS

Ref. No.	**No. of Covid-19 image**	**Others image**	**Validation accuracy**	**Method**
[9]	369	309	98.2%	Ensemble Model
[10]	341	2700	93.3%	ResNet50
[11]	300	500	87.8%	17 layered CNN Model
Proposed method	701	3500	99.2%	Ensemble Model

V. Conclusion

In this paper we have classified an X-ray image as COVID-19, pneumonia and normal. X-ray machines are widely available and provide images for quickly diagnosis. So, chest X-ray images can be very useful in early diagnosis of COVID-19 and pneumonia. The result of the segmentation can then be used to obtain further diagnostic insights. The goal was to detect and segment those affected area of COVID-19, and pneumonia with the help of semantic image segmentation method, so that we can help the radiologist by giving the results with higher precision. In first phase we have developed a classification model to classify COVID-19, pneumonia or normal and in second phase we have built a model for segmentation task on given image. The segmentation model will activate only when the classification model classifies an X-ray image as COVID-19 or pneumonia. The recognizing of COVID-19 and pneumonia from X-ray images with other X-ray images is very challenging because of the high variation in infection characteristics, and low intensity contrast between infections and normal tissues. The affected area has been segmented properly which vastly depend on the output of Grad-CAM algorithm. In future we have a great fascination to work with more deep learning algorithms to find out the right affected area and work with more X-ray image class through different algorithm for both classification and segmentation as well as to make API for that.

References

[1] COVID-19 situation reports-world health organization. https://www.who.int/publications/m/item/weekly-epidemiological-update---30-november-2020. [Accessed: 30-Nov-2020].

[2] R. N. Jennifer, L.W. Fink, "What COVID-19 Does to Your Body", Healthgrades, 18-Nov-2020. [Online]. Available: https://www.healthgrades.com/right-care/coronavirus/what-covid-19-does-to-your-body. [Accessed: 30-Nov-2020].

[3] W. Kong, P.P. Agarwal, "Chest imaging appearance of COVID-19 infection," Radiology:Cardiothoracic Imaging, vol. 2, no. 1, 2020.

[4] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, et al., "Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," Radiology, vol. 296, no. 2, 2020.

[5] D. Cozzi, M. Albanesi, E. Cavigli et al.,"Chest X-ray in new Coronavirus Disease 2019 (COVID-19) infection: findings and correlation with clinical outcome," Radiol med, vol. 125, no. 1, 2020, pp730–737.

[6] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, et al., "Sensitivity of chest CT for COVID-19: comparison to RT-PCR," Radiology, vol 296, no. 2, 2020.

[7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 618-626.

[8] L. Wang and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images," arXiv:2003.09871, 2020.

[9] M. Mishra, V. Parashar and R. Shimpi, "Development and evaluation of an AI System for early detection of Covid-19 pneumonia using X-ray (Student Consortium) ," 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), New Delhi, India, 2020, pp. 292-296.

[10] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," arXiv:2003.10849, 2020.

[11] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Comput Biol Med:103792, vol 121, Jun 2020.

[12] J. P. Cohen, P. Morrison, L. Dao, "COVID-19 image data collection," arXiv:2003.11597, 2020.

[13] C. Imaging, "This is a thread of COVID-19 CXR (all SARS-CoV-2 PCR+) from my hospital (Spain)," 28-Mar-2020. [Online]. Available: https://twitter.com/ChestImaging/status/1243928581983670272. [Accessed: 12-Aug-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," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 3462-3471.

[15] P. Mooney, "Chest X-Ray Images (Pneumonia)," *Kaggle*, 24-Mar-2018. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia. [Accessed: 12- Aug-2020].

[16] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks forbiomedical image segmentation," International Conference on Medical Image Computing and Computer-Assisted Intervention – (MICCAI), vol 9351, 2015, pp 234-241.

[17] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261-2269.

[18] M. Tan, Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proceedings of the 36th International Conference on Machine Learning (PMLR), vol 97, 2019, pp 6105-6114.

[19] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv: 1409.1556.

[20] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255.