

# Deep\_Saliency: A Deep Learning based Saliency Approach to detect Covid-19 through x-ray images

First: Praveen Verma, Roorkee Institute of Technology, Roorkee. Second: Dr. Chetan Singh Negi G.B. PANT UNIVERSITY OF AGRICULTURE & Technology, Pant Nagar Uttarakhand. Third: Maneesh Pant, College of Engineering Roorkee, Roorkee. Fourth: Anuj Saxena, CEO, Institute de informatica, Dehradun, Uttarakhand India.

## Abstract

Covid-19 a small virus has created a havoc in the world. The pandemic has already taken over 4 lakh lives. The tests to detect a Covid-19 positive takes time and is costly. Moreover, the ability of the virus to mutate surprises the doctors every day. Present paper proposes a saliency-based model called Deep\_Saliency. The model works on chest x-rays of healthy, unhealthy, and covid-19 patients. An x-ray repository of Covid-19, available in public domain, is taken for the study. Deep\_Saliency uses visual, disparity, and motion saliency to create a feature dataset of the x-rays. The collected features are tested and trained using Long Short-Term Memory (LSTM) network. A predictive analysis is performed using the x-ray of a new patient to confirm a Covid-19 positive case. The first objective of the paper is to detect Covid-19 positive cases from x-rays. The other objective is to provide a benchmark dataset of biomarkers. The proposed work achieved an accuracy of 96.66%.

Keywords: Covid-19, Virus Detection, Deep learning, LSTM, Saliency, Biomarkers.

## I Introduction

A challenge for all the researchers today is to detect Covid-19 positive. More than 0.2126076 billion patients spread over 213 countries, the pandemic has brought the world to its knees. More than 0.761018 million people have already lost their lives [1]. Scientists who work round the clock also could not find a vaccine for this virus in six months. Preventive measures and a strong immunity have helped recover 35 lakh patients. A vaccine may take another month or so to come but will only be commercially available by year's end. The dynamic nature of the virus has increased the difficulties of the doctors and the scientists. With no medicine available, and no vaccine the only way out is to prevent it from spreading. It is therefore important to detect and isolate the positive patients. A test conducted on a patient is no guarantee that the patient would not turn positive in the next few days or even the next day.

The test of Covid-19 involves studying of RNA. The procedure is lengthy and takes at least 24 hours. Rapid detection kits failed to serve the purpose [2]. Asymptomatic individuals are another threat. Such patients do not show any symptom of Covid-19. Reverse Transcription Polymerase Chain Reaction (RT-PCR) depends on swabs from the nose and mouth. These swabs have the chances of getting an infection while collecting the samples. A big issue with RT-PCR testing is that it depends on patented products. Many manufacturers can't quickly develop these to meet the shortfall. Any glitch would further reduce RT-PCR 's reliability, which is already plagued with precision problems. An overview [2] of the available RT-PCR kits for COVID-19 reveals that some are only 90% sensitive and 96% accurate. This could be just 66 to 80 per cent in real-world conditions, which means one in every three would be falsely tested as negative.

Considering the challenges faced by the Doctors and scientists the authors of the paper are proposing a Deep\_Saliency approach to detect the corona positive cases. The aim is to detect the patient at a very early stage. A simple chest x-ray can immediately detect a Covid-19 suspect and the doctors can isolate the patient. An immediate isolation would help in stopping the community spread of the virus.

For the research work the dataset is taken from the x-ray dataset available in public domain. The x-rays are of Normal, Pneumonia (virus, and Bacteria), and Covid-19 positive and negative patients. For better results, chest X-ray has been taken from benchmark datasets, references and details given in experimental setup section of the paper. The objectives of the paper are to extract the salient features from the images and predict a Covid-19 positive case. To extract the features visual saliency, disparity, and motion saliency is calculated. A fusion map is generated using these features to get a Feature Group Map (FGM) for Long-Short Term Memory (LSTM) model. Before proceeding further, first understand what saliency is and why it was chosen along with motion and disparity. Saliency is the component which pops first to the human eye in an image.

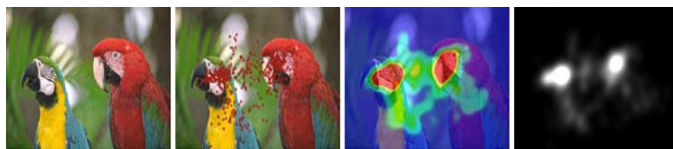


Figure 1 .1 the two parrots first pop up in our eyes and not in the background.

Figure 1.1. From left to right: (a) original, (b) fixation map with red fixation points, (c) heat map (red spots represent the most visually salient areas of the picture), and (d) saliency map Meur, le. O. et al (2012) [3]

There are two types of visual saliency detection techniques, top-down approach and bottom-up approach. Top-down approach is used to identify important objects like faces whereas bottom-up approach is used to analyze the pixel values and compute saliency values for each pixel like high pixel contrast and group highly salient pixels into regions. Based on depth, texture, luminance and contrast of color in [4] a novel stereoscopic saliency detection framework based on new fusion method was proposed, named Graph-based saliency detection algorithm. The algorithm in [5] was designed to represent an image as a fully connected graph. This model achieved 98% of the ROC. It was further [6] elaborated that how mathematics and biology could play an important role as image saliency detector. The mathematical model and understanding of a scene quickly break a complex problem into a simpler version which can be analyzed easily. As saliency was aimed at stereoscopic images a motion estimation method was developed in [7] which aimed at providing the motion vectors. It also gave a demonstration of full search algorithm with significantly less amount of time. Saliency considers left and right images when it comes to stereoscopic (or 3D) images. A method was proposed in [8] to detect saliency by the amount it differed from the object's neighborhood. The approach was easy to implement and could be used in many image and video content analysis applications. For improving visual attention in machine, an element in brain similar to neuron was designed in [9]. Koch used bottom up saliency model for this work. Various methods in [10-20] were proposed for saliency detection, but when it came to multiple objects the performance was not as expected. A new method in [21] came up which combined local feature, regional feature, edge and multiple scales. It did help to overcome the issue of multiple object detection, but the cost was too high and also it did not perform well on videos. To overcome this, a 2D model on 2D images in [22] was proposed but it only exploited depth contrast. Further, an automatic

detection system in [23] introduced a method that automatically detected salient image region in stereoscopic videos. The method was improved in [24] and tested on two publicly available stereoscopic datasets. The method was further improved in [25].

Though saliency was aimed at stereoscopic images or videos, but it has not limited the researcher. It has found its application in medical field as well. To detect breast cancer, a derivative-based feature saliency model in [28] was designed. It could highlight suspicious places in mammograms. For leakage detection in diabetic and malaria retinopathy, saliency map was used in [29]. Wireless capsule endoscopy (WCE) with saliency was used to detect ulcers in [30], it was further used with k-mean clustering to classify abnormalities in [31]. Feature extraction is a must for any classification algorithm to run successfully. Saliency with Convolutional Neural Network (CNN) is used to segment Ultrasound heart images in [32]. Exploiting saliency capabilities to detect boundaries and connectivity, optic disc segmentation is proposed in [33].

The analysis of previous work on saliency indicates that it is a good choice for medical images. Coming back to Covid-19 and the proposed model. As mentioned above RT-PCR in a real-world problem may give just 66% to 80% accuracy. Covid-19 is a chest related disease, improving the chest x-ray diagnosis would result in higher accuracy. Deep learning approaches have been studied and explored for diagnosis of Covid-19 [34-40]. Covid-Net proposed in [34] claims to achieve an accuracy of 80%. Inspired Covid-Net, a patch based convolutional neural network approach is proposed in [41]. CovidGan proposed in [42] had similar problem of small dataset for Convolutional Neural Network (CNN). They produced synthetic Covid X-ray images using their approach. The accuracy claimed is 95% but cannot be trusted as the dataset used is synthetic or manufactured to be more precise. The proposed models up until now have not been able to achieve more than 80% accuracy, due to limited dataset.

To improve the accuracy the paper proposes to use Fusion Map (Saliency+Disparity+Motion) to generate a Group Feature MAP(GFM). GFM would be the input for the deep Learning Model. CNN has been the first choice of the researcher when it comes to classification. But, as seen in [41 and 42] the model is not able to achieve more than 80% accuracy. Keeping this in mind a CNN + LSTM model is designed for classification and prediction. LSTM's forget gate is expected to retain remember the features for improved classification.

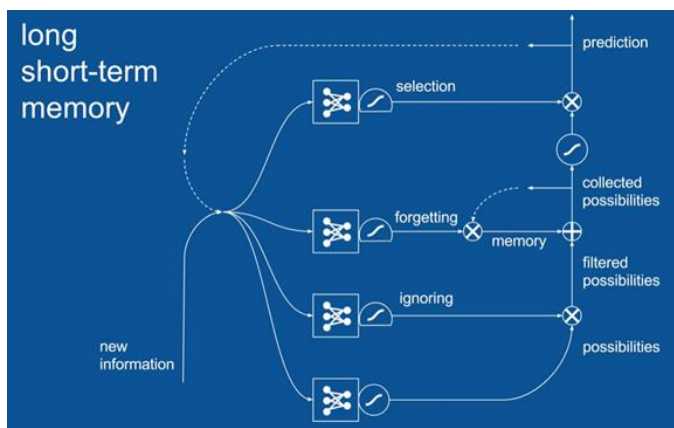


Figure 1.2: Taken from <https://towardsdatascience.com/machine-learning-recurrent-neural-networks-and-long-short-term-memory-lstm-python-keras-example-86001ceaaebc>

Figure 1.2 shows a general working model of LSTM.

The highlights of the study are:

1. The Saliency map gives the salient features but for image classification more detailed features are required for better accuracy. Disparity map improves the luminosity of the features. Motion map is drawn between original and Saliency Map, this enables to capture the minutest difference between the neighbouring pixels. The Fusion map generated has more visible features.
2. Covid-19 is found to mutate, it changed more than 20 times as of now. The virus thus is difficult to detect. The fusion map is able to retrieve the salient features more precisely. Classifiers work perfectly on benchmark datasets but in the case of Covid-19 there is limited availability of data and the data is in raw form as well. An image is normalized or pre-processed before it features are passed on to the classifier. Fusion map over comes the issue of visibility of salient features. The results obtained supports the claim.
3. The dataset has data from day 1 to day 30 of few patients, this enables the Deep Learning network to train itself better. The model is CNN+LSTM. LSTM work fine on large datasets but CNN works better on smaller datasets. The forget gate of LSTM enables the network to retain the features for a period of time. The forget gate will help to detect the Covid-19 positive case at an early stage.
4. The proposed model is expected to help the doctors and scientists to (a) detect a Covid-19 positive patient through the Chest X-rays at a very early stage, and (b) it would enable them to verify the authenticity of RT-PCR report.

## II Methodology

The aim is to reduce the objective function in fn (1).

$$\left[ pf \left( \sum_{\forall I} Fm \right) \right] \dots\dots\dots fn(1)$$

Here,

*pf* is Predictive function

$\sum_{\forall I} Fm$  is Fusion map of all the images,

The objective function is expanded and explained in the following section followed by brief steps and algorithms.

### 1. Objective functions explained

Fusion map is designed by the authors merging visual saliency map, disparity, and motion saliency.

## 1.1 Saliency map

$$S_{map} = \sum_{\forall I} GBVS(I) \dots \dots \dots \text{fn}(2)$$

Here,

$S_{map}$  is visual saliency

$\forall I$  is all the elements of Image

$GBVS(I)$  is saliency function

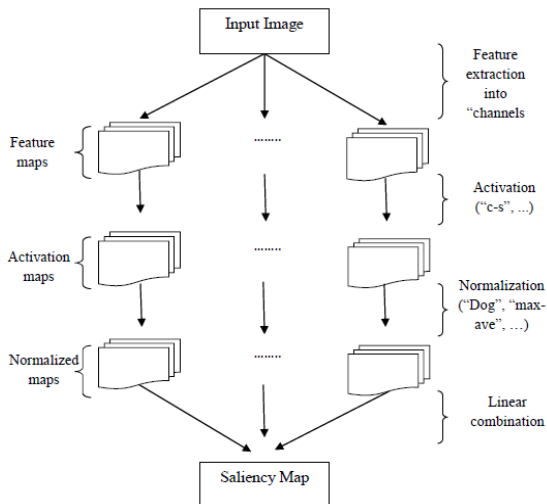


Figure 2: Graph based visual saliency, adopted from [32]

Figure 2 is a general working model of Graph based Video saliency(GBVS).

### 1.1.1 Step involved

- i. Compute feature maps from an x-ray image. Low-level features are derived from the color, intensity and orientation channels. The features extracted using GBVS are inspired by biologically based filters proposed in[4]. In order to create a set of scale-space images, each input image is first filtered by a series of Gaussian filters with different scales. A cross-scale image collection of differences is then specified by interpolation and subtraction between the variations of the chosen image size. Intensity and color characteristics are derived depending on contrast from each differential image diagram. The information regarding local orientation is also extracted at each scale by oriented Gabor pyramids. In all, 42 feature maps are calculated: 6 for strength, 12 for color and 24 for orientation.
- ii. Activation maps are computed using center-Surround Subtraction approach from feature maps. A graph map is identified as nodes by treating all the positions of pixels within the features detected in the last step. This determines and measures the dissimilarity between nodes and the weight of an edge of the graph map. To evaluate the dissimilarity between the node and its neighbouring nodes, the equilibrium distribution is computed.
- iii. Activation maps are normalized using global-mean maximum scheme. Saliency regions are clustered, dependent on distribution, and used as the final saliency map.

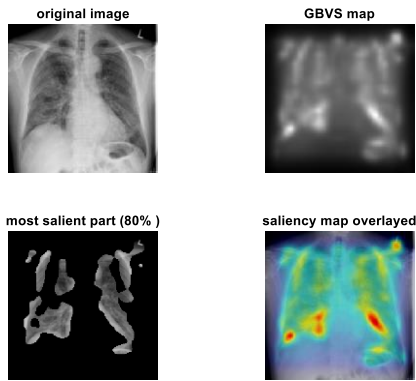


Figure 3: Different stages in saliency map

Figure (2) shows the original image. GBVS map drawn from the original image. Salient parts detected.

## 1.2 Disparity map

$$D_{map} = \sum_{\forall I} Disparity(S_{map}) \dots \dots \text{fn}(3)$$

Here,

$D_{map}$  is disparity map

$\forall I$  is all the elements of Image

$Disparity()$  is disparity function



Figure 4: Disparity Map

### 1.2.1 Steps

To get the disparity map, as shown in Fig. 3, a matching procedure is executed. Every pixel in the original image and corresponding pixel in the GBVS feature map image is taken and disparity map (distance between the pixel values) is computed. In the disparity map, brighter shades indicates more shifts and lesser distance as compared to darker shades, which gives lesser shifts and more distance from the GBVS feature map.

### 1.2.2 Mathematical Formulation

To calculate disparity local optimization technique is used. It calculates the disparity of each pixel independently, using the selected pixel's single matching cost. Local optimizations typically yield precise estimates of the disparity in textured regions. The "Winner-Take-All"[45] (WTA) is a basic local optimisation technique. The optimization of WTA consists of selecting the disparity  $d$  for each pixel  $(x, y)$  which is represented as:

$$d_x = \arg \min_{\tilde{d}} cf(x, x - \tilde{d}) \dots\dots \text{eq}(1)$$

here,

$cf(x, x - \tilde{d})$  is the cost function

$\tilde{d}$  is the maximum disparity value defined

The cost function corresponds to Sum of Absolut Difference (SAD). Cost function results are not shared with adjacent pixels. The minimal cost, therefore, yields the final value of variance. The optimization of eq(1) yields a disparity estimates in texture less regions. The algorithm explains the steps

*Algorithm 1: Disparity*

Require:  $d_{map}(x, y)$  a 2D array to hold disparity map

Initialize  $d_{map}(x, y)$ , min\_disparity=-16, max\_disparity=6

$I_1 \leftarrow$  read original x-ray image

$I_2 \leftarrow$  read saliency map

$[r,c,n] \leftarrow$  get size of  $I_1$

$I_2 \leftarrow I_2(\text{size}, I_1)$  // normalize size of  $I_2$  to match  $I_1$

for i in n do

    for j in r do

        for k in c

            // calculate disparity of Original ( $I_{d1}$ ) and Saliency map ( $I_{d2}$ )

$$I_{d1} = \arg \min_{\tilde{d}} cf(x(j, k, i), x(j, k, i) - \tilde{d})$$

$$I_{d2} = \arg \min_{\tilde{d}} cf(x(j, k, i) + \tilde{d}, x(j, k, i))$$

        If  $I_{d2}(x(j, k, i) - I_{d1}) = I_{d1}$  then

                the match is reliable

        else

                unreliable match

                pixels padded with disparity value of neighbouring pixels

        end if

    end for

end for

end for

**1.3 Motion Map**

$$M_{map} = \sum_{\forall} Motion(I_{gray}, S_{map}) \dots\dots\dots \text{fn}(4)$$

Here,

$M_{map}$  is motion map

$\forall$  is all the elements of the image

$I_{gray}$  is original grey scale image

$Motion()$  is motion estimation



Figure 5: Motion Map

### 1.3.1 Mathematical formulation

Motion Estimation is done using a combination of Block Matching algorithm and Optical Flow algorithm. For optical flow estimation, two consecutive consider frames  $f(x, y)$  and  $g(x, y)$ .

We have

$$G(x,y) = f(x + Dx, y + Dy)$$

$$\approx f(x, y) + \Delta x \frac{\partial}{\partial x} f(x, y) + \Delta y \frac{\partial}{\partial y} f(x, y)$$

which is the first order Taylor series approximation. The optical shift can be found by solving the minimization problem

$$\underset{\Delta x, \Delta y}{\text{minimize}} \Phi(\Delta x, \Delta y)$$

Where,

$$\Phi(\Delta x, \Delta y) = \sum_{x,y} (g(x, y) - f(x, y) - \Delta x \frac{\partial}{\partial x} f(x, y) - \Delta y \frac{\partial}{\partial y} f(x, y))^2$$

Since this is a linear least square problem, the optimal  $Dx$ , and  $Dy$  can be determined by setting the derivative of the objective function to zero. Thus we have,

$$\frac{\partial \Phi}{\partial \Delta x} = 0 \quad \text{and} \quad \frac{\partial \Phi}{\partial \Delta y} = 0$$

and consequently we can setup the following system of linear equations:

$$\begin{pmatrix} \sum_{x,y} \left(\frac{\partial f}{\partial x}\right)^2 & \sum_{x,y} \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \\ \sum_{x,y} \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} & \sum_{x,y} \left(\frac{\partial f}{\partial y}\right)^2 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} \sum_{x,y} (g - f) \frac{\partial f}{\partial x} \\ \sum_{x,y} (g - f) \frac{\partial f}{\partial y} \end{pmatrix}$$

Therefore, by solving this system of linear equations we can determine the optimal solution. Note that in order to make Taylor approximation valid, we implicitly assume that

$$|\Delta x| \ll 1 \quad \text{and} \quad |\Delta y| \ll 1$$

So the overall displacement can be determined as

$$Dx = \Delta x + dx \quad \text{and} \quad Dy = \Delta y + dy$$



The image formed in Fig. 4. show the motion vectors of x-ray image. To find the motion parts, the difference between the original image and GBV map (figure (2)) is drawn. By subtracting each frame from its neighbouring frame, successive frames are calculated. Sum of Absolute Difference (SAD) is calculated between the frames to get the motion map.

### 1.4 Fusion Map

$$F_m = [S_{map} \cup D_{map} \cup M_{map}] \dots \text{fn}(5)$$

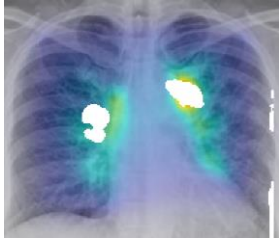


Figure 6: Saliency and Disparity Fused

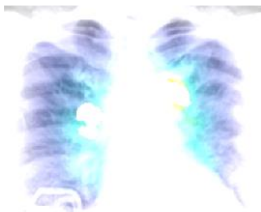


Figure 7: Saliency, Disparity, Motion fused

Fusion is a simple addition of the three maps.

#### Algorithm 2: fusion Map

Initialize  $F_{map}$ , and  $M_{image}$

Start

$D_{map} \leftarrow \text{resize}(\text{size}(S_{map}))$

$M_{map} \leftarrow \text{resize}(\text{size}(S_{map}))$

$M_{image} \leftarrow S_{map} + D_{map}$

$F_{map} \leftarrow M_{image} + M_{map}$

The resultant map is a fusion map of Saliency, Disparity and Motion. Figure 6 shows a fusion map. The source image is chest x-ray of a covid-19 positive patient.

### 1.5 Group Feature Map

$$GFM = \sum_{i=0}^n F_{map}(i) \dots \text{fn}(6)$$

here,

GFM is Group fusion map

$n$  is the number of images in the folder

$GFM$  is a collection of all the fusion maps created using  $\text{fn}(5)$  and saved in a separate folder.

## 1.6 Prepare Data

$TR_{fm}, TS_{fm} = split(GFM) \dots$  fn(7)

here,

$TR_{fm}$  is Training Set

$TS_{fm}$  is Testing Set

The Group feature map created using fn(6) is split into Training and Testing set for predictive analysis.

## 1.7 Predictive Function

$P_o = model.predict(lstm.eval(TR_{fm}, TS_{fm}), NF_{map}) \dots$  fn(8)

here,

$NF_{map}$  is the new Fusion map

$P_o$  is the predicted output

This is the final objective to predict Covid-19 using chest x-ray of a new patient. Algorithm, given in the following section, explains clearly how the entire model works.

## III Experimental Setup

There are two parts to any software based experimental setup. Hardware and Dataset. The model is tested using Matlab.

### 1. Hardware

Processor: Intel core i5, fifth generation

RAM: 16 GB

Graphic card: Nvidia

HDD: 1 TB

OS: Windows 10

### 2. Dataset

Pneumonia, and Normal chest x-ray dataset has been drawn from MIMIC-CXR dataset [43,44,45], NIH dataset [46] contains 1575 normal x-ray images, 4975 Pneumonia x-rays (virus and bacteria). Covid-19 x-rays from IEEE-8023[47], Covid-19 radiography [48], and Covid-19 X-ray dataset [49]. The Covid-19 dataset has images of 852 positive cases. Out of 852, 763 images were taken as rest of the images had labels or equipment impressions(as shown in figure(8)), making it difficult to get the salient features.



Figure(8): Covid-19 with instruments

### 3. Model and implementation

The training and testing set is given in fn(7). The Deep learning model based on LSTM requires an activation function.

The implementation model is summarised in the following algorithm

*Algorithm 3: implementation*

i. Setup

Initialize required variables

Resize all the images to 200X158

Start

Step 1. Generate GFM using function 1 to 7

Step 2. Tran, Test  $\leftarrow$  using function 8 get training and testing set

Step 3. Model  $\leftarrow$  create a function

ii. Design the network

Step 1. Input layer

GFM

Step 2. Middle Layer

1. Convolution2D 64 4x4 convolutions with stride [0 1] and padding [1 2 2 1]

2. SWISH

$$SWISH(x) = x \times sigmoid(x) \dots eq(2)$$

here,

$x$  is the input neuron

$$sigmoid(x) = (1 + e^{-x})^{-1}$$

3 Max Pooling 4x4 max pooling with stride [1 0] and padding [0 1 0 1]

4 Convolution2D 32 4x4 convolutions with stride [2 2] and padding [1 2 2 1]

5 SWISH (eq 2)

6 Max Pooling 4x4 max pooling with stride [2 2] and padding [0 1 1 0]

7 Convolution 64 4x4 convolutions with stride [2 2] and padding [1 1 1 1]

8 SWISH (eq 2)

9 Max Pooling 2x2 max pooling with stride [2 2] and padding [0 1 1 0]

Step 3. Final Layer

1. Fully Connected 128

2. SWISH (eq 2)

3. Fully connected 32

4. Softmax

5. Adam optimizer

6. Classification output crossentropy

iii evaluate the model

iv Predictive analysis

1.  $I_p \leftarrow$  read image to be tested for prediction

2.  $N_{map} \leftarrow$  create a new fusion map for  $I_p$

3.  $I_o \leftarrow$  predict (model,  $N_{map}$ )

iv. Evaluate the model

## IV Model Evaluation

For accuracy True Negative (TN), False Negative (FN), True Positive (TP), and True Negative (TN) are measured. These are further used to calculate Precision, Recall, Accuracy and F1 score. The mathematical formulation used is:

$$\text{Precision} = \frac{TP}{TP + FP} \dots \text{eq(3)}$$

$$\text{Recall} = \frac{TP}{TP + FN} \dots \text{eq(4)}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \dots \text{eq(5)}$$

$$F1 = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \dots \text{eq(6)}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \dots \text{eq(7)}$$

TP is the number of correctly detected ailment; FN is the number of ailments that have not been correctly identified or wrongly identified. FP is the number of ailments identified as one in the list of detection, but they are not. In simple terms they are falsely identified. TN is the number of not an ailment correctly identified. F1 score is taken for early stoppage of the classification.

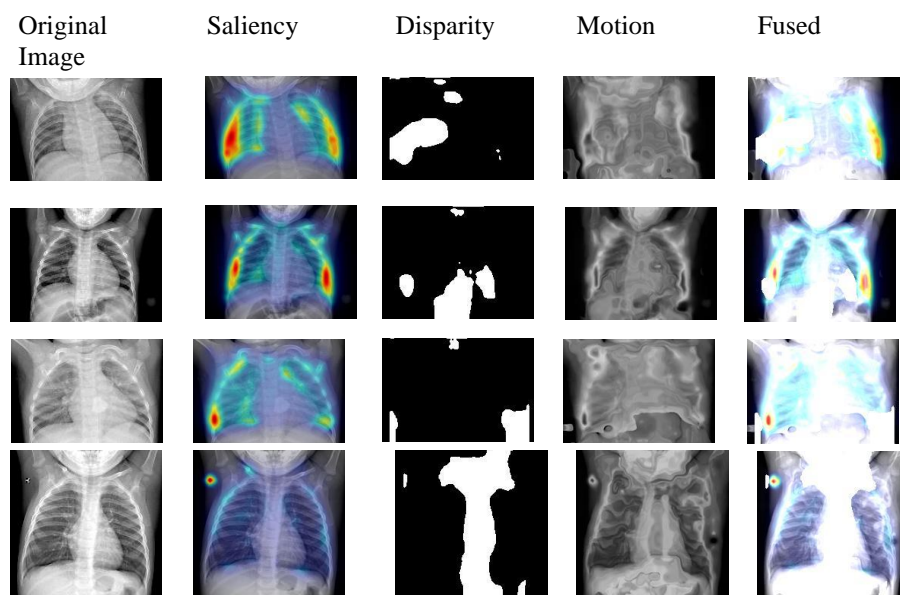
## V Results

The quantitative comparison results of the fusion map are given in Table 1. In this table, Model 1 [ITTI] represents the saliency detection model by fusion method of linear combination from 2D saliency detection model in (Itti et al., 1998) and depth model in (Wang et al., 2013); Model 2 [AIM] represents the saliency detection model by fusion method of linear combination from 2D saliency detection model in (Bruce and Tsotsos et al., 2006 [26]) and depth model in (Wang et al., 2013 [22]); Model 3 [HOU'S] represents the saliency detection model by fusion method of linear combination from 2D saliency detection model in (Hou and Zhang et al., 2007) [27] and depth model in (Wang et al., 2013). From this table, we can see that the proposed model performs better than the existing fusion models.

Table 1: The AUC and CC comparison.

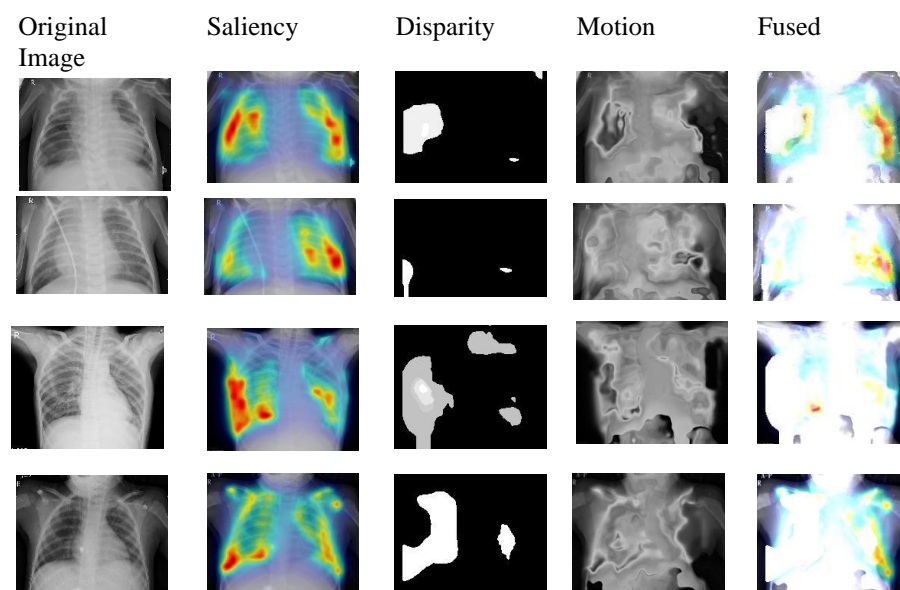
Model	CC (Correlation Coefficient)	AUC
IT+DSM	0.3752	0.8490
GBVS+DSM	0.3903	0.8509
AIM+DSM	0.3419	0.8495
FT+DSM	0.3148	0.7971
Fang et al's Method	0.5453	0.8650
<b><math>F_{map}</math></b>	<b>0.7904</b>	<b>0.9488</b>

The results of Deep-Saliency are shown on four sample images. The actual results contain separate folders for Saliency, Disparity, Motion, and Fusion.



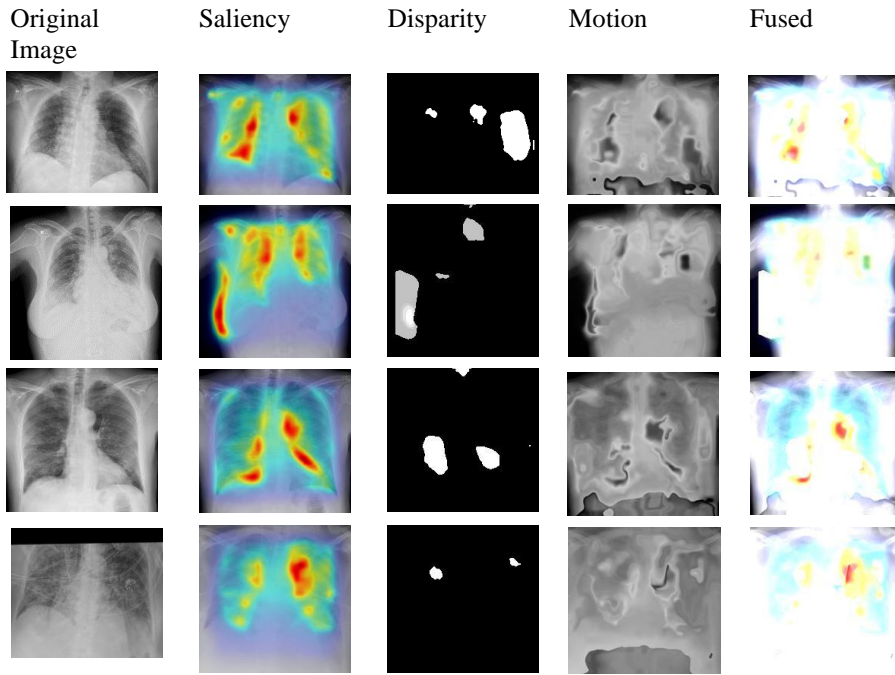
Figure(9): Normal x-ray images

Figure(9) shows the saliency, disparity, motion and fusion map. Four sample images are shown in the table to give an idea of these maps on Normal x-rays.



Figure(10): Pneumonia x-ray images (bacteria and Virus)

Figure(10) shows the different maps on Pneumonia -ray images. The two cases of pneumonia, virus or bacteria have been considered.



Figure(11): Covid-19

Figure(11) shows the different maps on Covid-19 x-ray images. The images are from day 1 to day 30 in cases of acute infection (SARS, pneumonia). The two cases of pneumonia, virus or bacteria have been considered.

The fused map in Figures 9, 10 and 11 shows that the most outstanding component is illuminated, and the luminosity of the picture is increased. The images are maps of features so the original image will only be available for Saliency. Before choosing LSTM+CNN model with SWASH, various classification algorithms were tried out.

Table 2: Different Classification algorithms using NIH dataset

Model	Precision	Recall	Accuracy	F1-score
Kmeans	0.700934579	0.6355932	0.7866667	0.6666667
Decision Stumps	0.70754717	0.6302521	0.7933333	0.6666667
Naïve Bayes	0.700934579	0.6355932	0.7866667	0.6666667
LSTM	0.72815534	0.6147541	0.8133333	0.6666667
Deep-Saliency	0.9375	0.5172414	0.9666667	0.6666667

Table 2 helped in choosing the LSTM-CNN model for the proposed Deep-Saliency. The dataset does not contain any Covid-19 images. To select the correct optimizer and the activation function the code was again run on NIH dataset using different activation functions and optimizers.

Table 3: Activation functions tested on NIH dataset

Optimizer	Activation Function		
	Relu	Sigmoid	SWASH
Adam	90.18	90.01	93.33
Sgd	88.2	69	91.2
Rmsprop	89.23	89.5	91.5
Adagrad	88.2	85.2	92.13

From table 3 it is concluded that SWASH activation function with Adam performs better than the other functions tested.

Table 4: Classification Results

Model	Precision	Recall	Accuracy	F1-score	Specificity
Local Approach [42]	60.6	60.1	70.7	59.3	89.7
Deep-Saliency	93.75	51.72414	96.66667	66.66667	99.89

Table 5: Sensitivity Calculation

Methods	Sensitivity			
	Normal	Bacteria	Virus	Covid-19
[42]	89.5	91	91	92.5
Deep-Saliency	91.5	92.8	93.2	94.5

Table 4 shows the classification results using the local approach [42] and the proposed model. An accuracy of 96.66 is good considering the limited dataset of Covid-19 dataset.

Table 6: Comparative analysis on the basis of sensitivity

Methods	Sensitivity			
	Normal	Pneumonia (Virus)	Bacteria	Covid-19
[41]	95	91	-	80
[42]	90	93	-	100
Deep-Saliency	99	99	98	100

Table 7: Comparative analysis on the basis of Precision

Methods	Precision			
	Normal	Pneumonia (Virus)	Bacteria	Covid-19
[41]	91.3	93.8	-	88.9
[42]	95.7	90.3	-	76.9
Deep-Saliency	98.4	97.2	97.4	99.98

Table 6 and 7 are comparative analysis of the proposed model with the existing models proposed for Covid-19. The results give an accuracy of 96.66% in case of Covid-19. As LSTM has an issue of taking more time a comparison is drawn between different classifiers.

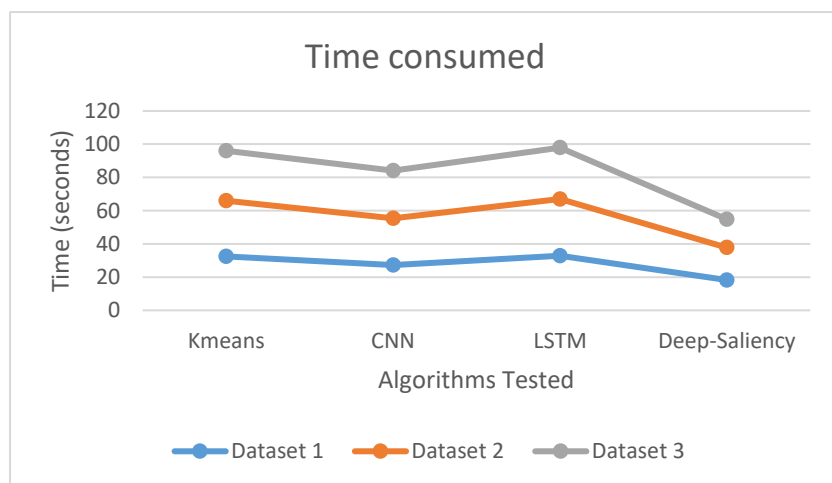


Figure (12): Comparative analysis of time consumed by different algorithms

As there is always a challenge of LSTM taking longer time, a comparison of time taken by the proposed model on the three Covid-19 datasets has been drawn. The comparison is between Kmeans, CNN, LSTM, and Deep-Saliency. The time taken by the proposed model is lesser than other algorithms.

## VI Conclusion

Since the pandemic outbreak the authors have been looking for the datasets, clinical datasets have been uploaded as of Feb 2020. The datasets were good for analysing the severity of the outbreak and understanding the virus too. The number of cases around the world was rising at

a high - speed, more testing was performed and complete lockdown in some countries was enforced. Europe saw deserted streets for the first time, everyone staying at home, but the cases continue to shoot up. The number is rising even today, and there's no cure for it either. Isolation is the only solution, until the vaccine is out. But who isolated? Do the tests give accurate results? Covid-19 has a quality to mutate itself, in such scenarios how accurate the tests can be?

The proposed model is an attempt by the authors to detect Covid-19 at an early stage using chest x-rays. The model has achieved an accuracy of 96.66%, it could have been better had there been more images available. The available datasets are raw datasets, some are even with equipment. The proposed work is able to identify any virus or bacteria in the lungs. Saliency, motion, disparity, and fusion maps would work as biomarkers for the doctors and researchers. The dataset generated would act as benchmark dataset and would be soon uploaded in the public domain. RT-PCR claims to be 96% accurate but in the real-world scenario the expected accuracy is just 80% as there are always chances of swab taken as sample could be infected by a foreign element.

## VII References

- [1] WHO Coronavirus Disease (COVID-19) Dashboard. <https://covid19.who.int>
- [2] How Is COVID-19 Diagnosed and Are the Tests Reliable? <https://www.downtoearth.org.in/news/science-technology/how-is-covid-19-diagnosed-and-are-the-tests-reliable--71759>.
- [3] Meur, I. O. and Baccino, T. 2012. "Methods for comparing scanpaths and saliency maps: strengths and weaknesses", ©Psychonomic Society, Inc., pp. 1-16.
- [4] Fang, Y.; Wang J.; Narwaria, M.; Callet, L. P. and Lin, W. 2014. "Saliency Detection for Stereoscopic Images". *IEEE Transaction on Image Processing*, vol 6, no. 23.
- [5] Harel, J.; Koch, C. and Perona, P. 2007. "Graph-based Visual Saliency". In: *Proc. Advance in Neural Information Processing Systems*, Cambridge, MA: MIT Press, pp. 545-552.
- [6] Itti, L.; Koch, C. and Niebur, E. 1998. "A model of saliency-based visual attention for rapid scene analysis". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259.
- [7] Chan, S. H.; Vo, D. T. and Nguyen, T. Q. 2010. "Sub-pixel motion estimation without interpolation". In: *Proceedings of International Conference on Acoustics, Speech*
- [8] He, S.; Han, J.; Hu, X.; Xu, M.; Guo, L. and Liu, T. 2011. "A biologically inspired computational model for images saliency detection". In: *Proceedings of the 19th ACM international conference on Multimedia*, New York, pp. 1465-1468.
- [9] Koch, C. and Ullman, S. 1985. "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry". *Human Neurobiology*, pp. 115-141.
- [10] Oliva, A.; Torralba, A.; Castelhana, M. S. and Henderson, J. M. 2003. "Top-down control of Visual Attention in Object Detection". In: *Proceedings of International Conference on Image Processing*, IEEE Conference, vol. 1, pp. 253-256.
- [11] Han, J.; Ngan, K. N.; Li, M. and Zhang, H. J. 2006. "Unsupervised Extraction of Visual Attention Objects in Color Images". *IEEE Transaction on Circuits and Systems for Video Technology*, vol 16, pp-141-145.
- [12] Walther, D. and Koch, C. 2006. "Modeling attention to salient proto-objects". *Neural Networks*, vol. 19, no. 9, pp. 1395-1407.
- [13] Torralba, A.; Oliva, A.; Castelhana, M. S. and Henderson, J. M. 2006. "Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search". *Psychological Review*, vol. 113, no. 4, pp. 766-786.
- [14] Zhang, L; Tong, M. H.; Marks, T. K.; Shan, H. and Cottrell, G. W. 2008. "Sun: A bayesian framework for saliency using natural statistics." *Journal of Vision*, vol. 8, no.7, pp. 1-20.
- [15] Achanta, R.; Hemami, S.; Estrada, F. and Susstrunk, S. 2009. "Frequency-tuned salient region detection". In: *Computer Vision and Pattern Recognition*, IEEE Conference, pp. 1597-1604.
- [16] Shi, L.; Wang, J.; Xu, L. and Lu, H. 2009. "Context Saliency based Image Summarization". In: *International Conference on Multimedia and Expo*, IEEE Conference, pp. 270-273.
- [17] Shafie, A. A.; Hafiz, F. and Ali, M. H. 2009. "Motion Detection Techniques using Optical Flow". *World Academy of Science, Engineering and Technology*, vol. 3, no. 8, pp. 503-505.
- [18] Gopalakrishnan, V.; Hu. Y. and Rajan, D. 2009. "Salient region detection by modeling distributions of color and orientation". *IEEE Transactions on Multimedia*, vol. 11, no. 5, pp. 892-905.



- [19] Goferman, S.; Zelnik-Manor, L. and Tal, A. 2010. "Context-Aware Saliency Detection". In: *Computer Vision and Pattern Recognition (CVPR)*, IEEE Conference, pp. 2376-2383.
- [20] Cheng, M.-M.; Zhang, G.-X.; Mitra, N. J.; Huang, X. and Hu, S.-M. 2011. "Global contrast based salient region detection". In: *Computer Vision and Pattern Recognition*, IEEE Conference, pp. 409-416.
- [21] Guo, C and Zhang, L. 2010. "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression". *IEEE Transaction on Image Processing*, vol. 19, no. 1, pp. 185-198.
- [22] Wang, J.; Silva, M. P. D.; Callet, L. P. and Ricordel, V. 2013. "Computational Model of Stereoscopic 3D Visual Saliency". *IEEE Transaction on Image Processing*, vol 5, no. 5, pp. 1-11.
- [23] Kopf, S.; Dittrich, T.; Schaber, P.; Guthier, B. and Effelsberg, W. 2013. "Saliency Detection for Stereoscopic Detection". In: *Proc. of the 4th ACM Multimedia System Conference (MMSYS)*, pp. 12-23.
- [24] Cong. Runmin, Lei; Jianjun, Zhang; Changquig, Huang; Quingming, Cao; Xiaochun and Hou. Chunping, "Saliency Detection for Stereoscopic Images Based on Depth Confidence Analysis and Multiple Cues Fusion" *IEEE SIGNAL PROCESSING LETTERS*, VOL. 23, NO. 6, JUNE 2016
- [25] Du. Huan, Liu. Zhi, Song. Hangke, Mei. Lin and Xu. Zheng, "Improving RGBD Saliency Detection Using Progressive Region Classification and Saliency Fusion," *Special section on intelligent sensing on mobile and social media analytics*, Digital Object Identifier 10.1109/ACCESS.2016.2632724
- [26] Bruce, ND. and Tsotsos, JK. 2009 "Saliency, attention, and visual search: an information theoretic approach" *Journal of Vision March 2009, Vol.9, 5.* doi:10.1167/9.3.5
- [27] Hou and Zhang et al., 2007 "Saliency Detection: A Spectral Residual Approach Proceedings / CVPR", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- [28] W. E. Polakowski et al., "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," in *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 811-819, Dec. 1997, doi: 10.1109/42.650877.
- [29] Y. Zhao et al., "Intensity and Compactness Enabled Saliency Estimation for Leakage Detection in Diabetic and Malarial Retinopathy," in *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 51-63, Jan. 2017, doi: 10.1109/TMI.2016.2593725.
- [30] Y. Yuan, J. Wang, B. Li and M. Q. -. Meng, "Saliency Based Ulcer Detection for Wireless Capsule Endoscopy Diagnosis," in *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 2046-2057, Oct. 2015, doi: 10.1109/TMI.2015.2418534.
- [31] Y. Yuan, B. Li and M. Q. -. Meng, "WCE Abnormality Detection Based on Saliency and Adaptive Locality-Constrained Linear Coding," in *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 149-159, Jan. 2017, doi: 10.1109/TASE.2016.2610579.
- [32] H. Liu, W. Chu and H. Wang, "Automatic Segmentation Algorithm of Ultrasound Heart Image Based on Convolutional Neural Network and Image Saliency," in *IEEE Access*, vol. 8, pp. 104445-104457, 2020, doi: 10.1109/ACCESS.2020.2989819.
- [33] B. Zou, Q. Liu, K. Yue, Z. Chen, J. Chen and G. Zhao, "Saliency-Based Segmentation of Optic Disc in Retinal Images," in *Chinese Journal of Electronics*, vol. 28, no. 1, pp. 71-75, 1 2019, doi: 10.1049/cje.2017.12.007.
- [34] L. Wang and A. Wong, "COVID-net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," 2020, *arXiv:2003.09871*.
- [35] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, Apr. 2020.
- [36] E. El-Din Hemdan, M. A. Shouman, and M. Esmail Karar, "COVIDXnet: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, *arXiv:2003.11055*.
- [37] I. D. Apostolopoulos, S. Aznaouridis, and M. Tzani, "Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases," 2020, *arXiv:2004.00338*.
- [38] M. Farooq and A. Hafeez, "COVID-ResNet: A deep learning framework for screening of COVID19 from radiographs," 2020, *arXiv:2003.14395*.
- [39] 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," 2020, *arXiv:2004.02696*.
- [40] L. Zhong, L. Mu, J. Li, J. Wang, Z. Yin, and D. Liu, "Early prediction of the 2019 novel coronavirus outbreak in the mainland China based on simple mathematical model," *IEEE Access*, vol. 8, pp. 51761\_51769, 2020.
- [41] Y. Oh, S. Park and J. C. Ye, "Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2688-2700, Aug. 2020, doi: 10.1109/TMI.2020.2993291.
- [42] Multi-View Video Coding. <http://epixea.com/research/multi-view-coding-thesis3.html>.

- [43] Johnson, A., lungren, m., Peng, Y., Lu, Z., Mark, R., Berkowitz, S., & Horng, S. (2019). MIMIC-CXR-JPG - chest radiographs with structured labels (version 2.0.0). *PhysioNet*. <https://doi.org/10.13026/8360-t248>.
- [44] Johnson AE, Pollard TJ, Berkowitz S, Greenbaum NR, Lungren MP, Deng CY, Mark RG, Horng S. MIMIC-CXR: A large publicly available database of labeled chest radiographs. *arXiv preprint arXiv:1901.07042*. 2019 Jan 21.
- [45] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101 (23), pp. e215–e220.
- [46] Cohen, Joseph Paul. *Ieee8023/Covid-Chestxray-Dataset*. 2020. 2020. *GitHub*, <https://github.com/ieee8023/covid-chestxray-dataset>.
- [47] *IEEE Covid Chest X-Ray Dataset*. Accessed: August. 7, 2020. [Online]. Available: <https://github.com/ieee8023/covid-chestxray-dataset>.
- [48] *Covid19 Radiography Database*. Accessed: August. 10, 2020. [Online]. Available: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>
- [49] *COVID-19 Chest X-Ray Dataset Initiative*. Accessed: August. 11, 2020. [Online]. Available: <https://github.com/agchung/Figure1-COVIDchestxray-dataset>