



# Deep Transfer Learning Based Unified Framework for COVID19 Classification and Infection Detection from Chest X-Ray Images

Sankar Ganesh Sundaram<sup>1</sup> · Saleh Abdullah Aloyuni<sup>2</sup> · Raed Abdullah Alharbi<sup>2</sup> · Tariq Alqahtani<sup>3</sup> · Mohamed Yacin Sikkandar<sup>3</sup> · Chidambaram Subbiah<sup>4</sup>

Received: 4 March 2021 / Accepted: 5 July 2021 / Published online: 11 August 2021  
© King Fahd University of Petroleum & Minerals 2021

## Abstract

The presentation of the COVID19 has endangered several million lives worldwide causing thousands of deaths every day. Evolution of COVID19 as a pandemic calls for automated solutions for initial screening and treatment management. In addition to the thermal scanning mechanisms, findings from chest X-ray imaging examinations are reliable predictors in COVID19 detection, long-term monitoring and severity evaluation. This paper presents a novel deep transfer learning based framework for COVID19 detection and segmentation of infections from chest X-ray images. It is realized as a two-stage cascaded framework with classifier and segmentation subnetwork models. The classifier is modeled as a fine-tuned residual SqueezeNet network, and the segmentation network is implemented as a fine-tuned SegNet semantic segmentation network. The segmentation task is enhanced with a bioinspired Gaussian Mixture Model-based super pixel segmentation. This framework is trained and tested with two public datasets for binary and multiclass classifications and infection segmentation. It achieves accuracies of 99.69% and 99.48% for binary and three class classifications, and a mean accuracy of 83.437% for segmentation. Experimental results and comparative evaluations demonstrate the superiority of this unified model and signify potential extensions for biomarker definition and severity quantization.

**Keywords** COVID19 · Classification · Segmentation · Chest X-ray · Transfer learning · Residual SqueezeNet · SegNet

✉ Sankar Ganesh Sundaram  
ssganesa@gmail.com

Saleh Abdullah Aloyuni  
s.aloyuni@mu.edu.sa

Raed Abdullah Alharbi  
r.abdullah@mu.edu.sa

Tariq Alqahtani  
talqahtani@mu.edu.sa

Mohamed Yacin Sikkandar  
m.sikkandar@mu.edu.sa

Chidambaram Subbiah  
chidambaramraj1@gmail.com

- <sup>1</sup> Department of Artificial Intelligence and Data Science, KPR Institute of Engineering and Technology, Coimbatore 641407, Tamil Nadu, India
- <sup>2</sup> Department of Public Health, College of Applied Medical Sciences, Majmaah University, 11952, Al Majmaah, Saudi Arabia
- <sup>3</sup> Department of Medical Equipment Technology, College of Applied Medical Sciences, Majmaah University, 11952, Al Majmaah, Saudi Arabia
- <sup>4</sup> Department of Information Technology, National Engineering College, Kovilpatti 628503, Tamil Nadu, India

## 1 Introduction

COVID-19 has evolved into a global pandemic since its outbreak in December 2019. Essentially, Coronavirus diagnosis depends on polymerase chain response (PCR) tests, while the use of clinical imaging as a symptomatic test for Coronavirus was dubious. Chest X-rays [1] (CXR) are preferred for COVID-19 screening owing to the ease of acquisition, cheaper cost, portability of the imaging mechanism and minimized risks of virus spreading. Several studies of CXR images [2, 3] divulged a low sensitivity for COVID19 detection from 25 to 69%. However, the ability to detect the disease correctly, i.e., the specificity, is achieved around 90% with these images.

Radiologists also recommend CXR image examinations for lung related pathologies in conjunction with COVID19 infection. CXR images are found appropriate for COVID19 patients to assess [4] the course of the disease and evaluate the severity. These imaging examinations reveal opacities such as consolidations, Ground Glass Opacity (GGO) and nodular shadowing in the peripheral and lower lung regions.



High penetration of COVID19 and characteristic findings from the chest images call for automated detection of the infections and localization of the infections. Deep learning models capable of intrinsic data representation, deep feature extraction and learning from priors are deployed in COVID19 detection and management. They are built on classical models pretrained on large datasets which are fine-tuned to specific applications such as COVID19 detection and infection segmentation.

In [1], three pretrained Convolutional Neural Network (CNN) models based on ResNet50, InceptionV3 and Inception ResNet V2 networks are deployed for COVID19 detection from CXR images, achieving the best classification accuracy of 98% with the ResNet50 network. Further, segmentation networks based on CNNs are also used in the COVID19 triage along with classifier models. A lung segmentation model proposed in [5] based on UNet for detection of nodules and assessment of infection staging and quantifications from CXR images reports a Dice score of 97.5%. However, there is a lack of integral models for COVID19 detection and infection segmentation from CXR images.

Prior investigations have shown that CXR images tend to be broad-spectrum in nature, with symptoms matching other types of viral infections. Nevertheless, GGOs common in COVID19 patients manifest as patches in the bilateral, peripheral and subpleural lung regions which modify the density and size of the lung lesions in various stages of the infection. Further, gap consolidations are also found to co-occur with opacities in viral pneumonia and COVID19 cases. Though CXR image findings are insightful in COVID19 detection, they differ with severity of the infections and co-morbidities. This renders the detection of COVID19 difficult and demands for segmentation of infections after classification, for better clinical decisions. Hence, the proposed framework is realized as a two-stage classification-segmentation model facilitating COVID19 detection and localization of infections.

Further, compared to other diseases, design of COVID19 detection and segmentation models poses the following challenges.

1. Overlap of symptoms with other virus infections
2. Risk of misclassifications which may result in adverse outcomes or false alarms
3. Need for large public datasets for testing the models
4. Lack of benchmarks for comparative analysis

Inspired by the performance of discrete deep learning models for classification and segmentation, this paper proposes a unified framework for COVID19 detection and infection segmentation from CXR images. This framework is deployed with a novel classification-segmentation

pipeline harnessing the potential of pretrained classification and segmentation networks and unsupervised segmentation driven by meta-heuristic optimization. The proposed framework is unique of its kind in COVID19 management for infection detection and prognosis.

This is realized as a classification-segmentation model which classifies a candidate image into one of the target classes viz. COVID19, viral pneumonia and Normal, and segments the infections from the COVID19 positive images. In order to improve the segmentation accuracy, the Classification Activation Maps (CAM) of the classifier are subjected to a Gaussian Mixture Model [6] super pixel segmentation and Region of Interest (RoI) construction based on a Cuckoo Search Optimization (CSO) [7] algorithm.

The classifier and segmentation subnetworks are built on the pretrained Residual SqueezeNet [8] and the SegNet [9] models, respectively, and this framework is called the RSqz-SegNet after the underlying models.

The main contributions of this paper are as follows.

1. An integral framework for COVID19 detection and infection segmentation is proposed
2. This framework detects diverse infections including opacities, granulomas and subtle infections from COVID19 positive images
3. The potential of CAMs in classification and infection detection is established
4. The subnetworks can be isolated and employed as classifier and segmentation models
5. The proposed framework demonstrates superior classification and segmentation performances, featuring potential extension for severity grading.

The experimental results achieved by this framework on two public datasets, interpretation of results and statistical analyses demonstrate the superiority of this model compared to the state-of-the-art approaches.

Rest of this paper is organized as follows. In Sect. 2, a background of the deep learning and deep transfer learning approaches pertaining to this research is given. A comprehensive review is presented on the prior work in the context of this research in Sect. 3. The datasets, architectures of the baseline models and other approaches employed in realizing the framework and details of implementation are described in Sect. 4. In Sect. 5, the architecture of the proposed framework is described with a schematic diagram followed by the functions of the subnetworks. The experimental results, performance analyses and comparisons are given in Sect. 6. An analysis of the results and ablation study are presented in Sect. 7. The paper is concluded in Sect. 8.



## 2 Background

In the recent years, deep learning models are widely used in the screening, diagnosis and treatment management of several pathologies by extracting discerning features from medical images. Contrary to the machine learning models which are trained on hand-crafted features, deep learning models are capable of learning intrinsic features from the images for clinical diagnosis. With the COVID19 upsurge, numerous automated models have been deployed for detection and classification of pathologies, segmentation of infected regions, severity analysis etc. Deep learning models deployed with state-of-the-art algorithms, trained on large and diverse datasets, run on high-end computing devices demonstrate superior prediction and classification abilities. This section presents a concise view of the deep learning and deep transfer learning approaches employed in this research.

### 2.1 Deep Learning

Deep learning is a class of machine learning which employs multiple layers to transform the raw input to gradually extract higher-level features at each layer. The word deep refers to the number of layers in a model through which the data is progressively transformed. Deep learning models are generally built with algorithms including Deep Neural Network (DNN), Deep Convolutional Neural Networks (Deep CNN) and Deep Belief Networks (DBN). CNNs [10] are superior to other algorithms in extracting low, mid and high level image features, learning from image content and have shown remarkable performances in image classification, detection, segmentation, and retrieval tasks. A CNN comprises an input layer, several hidden layers and an output layer. Basically, hidden layers consist of convolutional, pooling and Fully Connected (FC) layers. A convolutional layer is modeled with the following attributes.

- i. A set of convolutional filters
- ii. Number of input and output channels
- iii. Padding, stride and dilation operators

The convolutional layer extracts features at various levels from the input images in the form of feature maps which are passed as input to the subsequent layers. The pooling layers operate independently on the feature maps, reducing their and computational costs in turn. The FC layers form the last few layers of the CNN which perform classification with a flattened vector constructed from the output of the previous layers. The schematic of a simple binary CNN classifier is shown in Fig. 1.

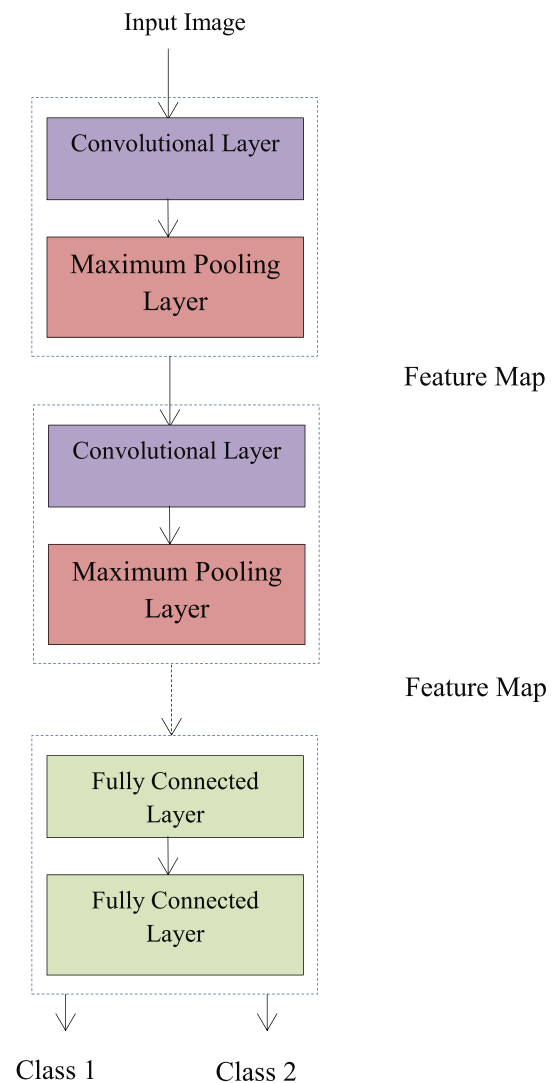


Fig. 1 CNN classifier architecture

A comprehensive survey by Sufian et al. [11] reports numerous studies exploring the potential application of deep learning models in COVID19 management.

CXR and CT images are widely used in the COVID19 clinical imaging examinations. Deep learning models trained with these images have simplified lung detection, infection localization, disease detection and classification, improving diagnostic values of these images. Mei et al. [12] employed two deep learning models for rapid COVID19 diagnosis. The first model was implemented with two CNN subnetworks, viz. Inception ResNet v2 and ResNet18 for identification of abnormal slices from CT volumes and disease diagnosis. The second model was implemented as a joint model combining the CNN and Multilayer Perceptron (MLP). Combined feature vectors comprising 512 dimensional features extracted from the CT image by the CNN and 12 clinical features were used to train the MLP for COVID19 classification.

This model reportedly achieves better classification accuracies compared to human experts. Similarly, Zhang et al. [13] proposed a two-stage deep learning framework for COVID19 management with CT lung images. It comprises a lung lesion segmentation network and a detection network which is trained with the lesion map. Further, this framework also supports prognosis analysis with clinical metadata and lung lesion features extracted from the lesion maps.

Though CT imaging examinations are highly recommended in COVID19 diagnosis, several deep learning models based on CXR images have been reported recently, due to their fast acquisition, low cost and simple imaging mechanism. Considering the significance of CXR images in COVID19, a public dataset called COVIDx containing 13,975 CXR images of 13,870 patients and a deep learning network called COVID-Net are presented by Wang and Wong [14]. This network reportedly achieves a classification accuracy of 93.3%.

## 2.2 Deep Transfer Learning

Transfer learning is a deep learning approach which makes use of the learning ability of a model for a particular problem to solve a similar problem. When the training dataset is smaller for a particular classification problem, transfer learning is applied to fine-tune the pretrained models trained with large volumes of data for a prior classification problem. Generally, transfer learning follows two approaches. In the first approach, the pretrained model is employed as a feature extractor and the learnt features are transferred to the new classifier which is trained with a smaller dataset specific to a classification problem. The second approach modifies the architecture of the pretrained models for improved classification performances. Generally, the convolutional layers are preserved for feature extraction at various levels and the Fully Connected (FC) layers of the pretrained models are replaced with new FCs with randomly initialized weights, which learn new discriminating patterns from the features extracted by the convolutional filters.

A pretrained VGG-16 model fine-tuned with a small CXR dataset of 1428 images is proposed in [15], achieving classification accuracies of 96% and 92.5% for binary and three class COVID19 classifications. This model is constructed by modifying the FC head of the VGG-16 with three FC layers with 512, 64 and 2 nodes for binary classification and 512, 64 and 3 nodes for three class classifications.

A most recent work in this context proposed in [16] presents a small dataset of chest CT and CXR images and two CNN architectures for COVID19 detection. The first CNN is implemented with a single convolution layer with 16 filters, followed by the batch normalization layer, a Rectified Linear Unit (ReLU), two FC layers and a softmax layer. This classifier achieves classification accuracies of 94% and 94.1% for

COVID19 detection with CT and CXR images, respectively. The second classifier is realized by modifying the last layers of the AlexNet [17] trained with millions of images. The weights and bias of the pretrained AlexNet are transferred to the new classifier which is fine-tuned with the smaller dataset. This classifier achieves classification accuracies of 82% and 98% for COVID19 detection with CT and CXR images, respectively. A recent investigation [18] on the application of machine and deep learning models for COVID19 triage emphasizes the usage of data driven deep learning models and transfer learning approaches in imaging examinations for COVID19 detection.

## 3 Related Work

As discussed in Sect. 2, deep learning models are successfully employed in COVID19 screening, diagnosis and treatment protocols with Chest CT and CXR images. However, a closer observation of the exiting deep learning models for COVID19 detection reveals that CXR images are best suited compared to CT images. In line with this observation, this section reviews excerpt literature and presents the research gaps from an inclusive study of the deep learning models for COVID19 detection and segmentation from CXR images.

Clinical findings in [19] show that peripheral distribution, GGO and bilateral involvement are widely observed in CXR images. COVID19 detection and segmentation models are trained to learn deep features from the CXR images to classify and segment them.

An investigation in [20] shows that high specificity and accuracy are achieved by MobileNetv2 and VGG19 models in comparison with the Xception, Inception and Inception ResNet V2 models, in COVID19 detection from two datasets with CXR images of normal, pneumonia and COVID19 cases. This study shows that VGG19 achieves highest accuracies of 98.75% and 93.48% for binary and three class classifications, while the best specificity is achieved with the MobileNetv2.

The authors extended the MobileNetv2 in [21] for the detection of six pulmonary disease classes such as COVID19, edema, effusion, emphysema, fibrosis and pneumonia which reportedly achieves an accuracy of 99.18% in COVID-19 detection. With the highest classification accuracy demonstrated by this model, the authors proceeded to standardize the deep features extracted from the CXR images as potential biomarkers of COVID19. The Covid-Resnet [22] constructed by adaptively fine-tuning the Resnet50 with decaying learning rates in three stages demonstrates stability with arbitrary test data achieving 96.23% overall accuracy.

COVIDX-Net [23], an integral framework of seven binary classifiers including the DenseNet201, InceptionV3, InceptionResNetV2, MobileNetV2, ResNetV2, VGG19 and Xception networks trained on a small dataset of CXR images, achieves the highest classification accuracy of 90%



with the VGG19 and Densenet classifiers. Similarly, few studies have been performed evaluating the performance of several deep learning models for COVID19 detection.

Narin et al. [24] have tested the pretrained InceptionV3, Inception ResNet V2 and ResNet50 networks on a small dataset with 50 COVID19 and 50 normal CXR images for COVID19 detection and achieved the highest classification accuracy of 98% with ResNet50. In [25], the authors have evaluated eleven CNNs on a small dataset and reported the highest classification accuracy of 95.38% with a combination of the Resnet50 for feature extraction and SVM classifier.

Recently capsule [26] networks which encode spatial information and probability of existence of an object in an image are employed in imaging problems. A capsule-based binary classifier, called the COVID-CAPS [27], reportedly achieves a classification accuracy of 95.7% without pretraining. This model exhibits an accuracy of 98.3%, pretrained on a small CXR dataset.

COVIDiagnosis-Net [28] is based on the lightweight SqueezeNet CNN, leveraging its small size, few parameters and demonstrated better accuracy compared to conventional CNNs. The hyper parameters of this classifier are tuned by and Bayesian optimization techniques to build an ideal classifier model. This model achieves an accuracy of 98.3% tested on the COVIDx dataset [14]. The COVID-Xpert [29] model based on DenseNet achieves a classification accuracy of 88.8% in three class COVID19 classification with CXR images.

Chowdhury et al. [30] employed three shallow and five deep learning models for COVID19, normal and viral infection detections. Binary and three class classifications are performed on a CXR image dataset constructed from six subsets with and without image augmentation. Performance evaluation of these models without augmentation shows that best classification accuracy of 99.41% is achieved by ResNet18 and CheXNet for two class, and 97.74% is achieved by CheXNet for three class classifications. Best accuracies of 99.70% and 97.94% are achieved for binary and three class classifications by the DenseNet201 under image augmentation. These results show that the deep CheXNet pretrained on a large CXR dataset demonstrates superior performance compared to other shallow and deep networks.

In spite of several classification models, COVID19 detection from CXR images with subtle details and rare morphologies is highly challenging. It is essential that deep learning models are trained with diverse training images to learn the intrinsic features, generalizing well with arbitrary test data. COVID19 treatment regime can be enhanced with integral classification and segmentation models to localize and segment the infections for further clinical triages.

A unified model called the COVID\_MTNNet [31] performs COVID-19 detection and infected region segmentation from

CT and CXR images, implemented with two distinct networks, an Inception Recurrent Residual Convolutional Neural Network (IRRCNN) and a NABLA-N [32] network for model for COVID-19 detection and segmentation, respectively. This network shows an accuracy of 84.67% for CXR test images. A classification and detection pipeline proposed in [33] performs COVID19 classification, labeling lung manifestations and segmentation of manifestations from chest CT scans.

Concerned with segmentation of opacities from the CXR images, Tang et al. [34] have proposed a UNet-based semantic segmentation network with a ResNet18 backbone for segmentation of opacities from CXR images. This network achieves an average Intersection over Union (IoU) > 0.5 in the segmentation of opacity regions. A review of recent literature reveals that there exist several deep learning models for COVID19 detection from CXR images. Recently, the significance of infection segmentations from lung regions is well understood for COVID19 detection and explorations are performed in this context. It is evidenced that, despite the wealth of literature in this area, there are limited studies on localization of infestations in the CXR images. Segmentation of infections from CXR images is pertinent in COVID19 treatment pipeline to expedite clinical decisions. Overview of the existing works shows that prior works typically address either classification or segmentation and there is a strong need for unified models for classification and segmentation.

## 4 Materials and Methods

The proposed framework implemented as an integral model performing classification and segmentation attempts to solve COVID19 detection and infection segmentation, incorporating two subnetworks and segmentation and optimization algorithms. This framework is trained on augmented image subsets and tested on standard image subsets for performance evaluation and visual interpretations. This section presents an overview of the implementation of the proposed system and the underlying architectures and methodologies.

### 4.1 Datasets and Details of Implementation

In this research, we have employed two datasets, viz. the Kaggle CXR [35] public database for COVID19 detection and The Cancer Imaging Archive (TCIA) [34] for opacity segmentation. The Kaggle CXR dataset contains 1143 COVID19-positive images, 1345 viral pneumonia images and 1341 normal images of dimension  $1024 \times 1024$  in png format. The entire dataset is divided into training and testing datasets. The training dataset is further augmented with images generated by translation, rotation and scaling of the

training images. The TCIA contains 221 chest X-ray images and corresponding masks of the infections. This dataset is employed exclusively for testing the classifier model, and training and testing the segmentation subnetwork. Though this dataset contains 221 images, the segmentation network is trained with COVID19 images correctly classified by the classification subnetwork. Out of 221 images, 215 images are classified as COVID19 under three class classification and these images are used for training and testing the segmentation subnetwork. Out of the 215 images, 100 images are arbitrarily chosen and augmented with 200 more images for training and 115 image are employed in testing the segmentation subnetwork.

The summary of the training and testing data is presented Table 1. The proposed framework is implemented with Matlab 2020b software in an i7-7700 K processor with 32 GB DDR4 RAM equipped with a NVIDIA GeForce GTX1060 3 GB Graphics card.

## 4.2 GMM Super pixel Segmentation

GMM models based on the stochastic Gaussian process are leveraged with clustering algorithms in unsupervised learning problems. In this research, GMM super pixels are employed in the segmentation of CAMs into semantically identical regions.

In an image segmentation problem, a GMM models each image pixel as a mixture of Gaussian distributions, constructing a set of super pixels, each of which is associated one of the distributions. Each pixel is assigned a super pixel label based on Expectation Maximization (EM) and posterior probability as described below.

For a given image  $I$  of size  $M \times N$ , with the total number of pixels  $T = M \times N$ , each image pixel is assigned an index  $i$  in the range  $[0, T - 1]$ . The GMM super pixel algorithm clusters the image pixels based on  $d_x$  and  $d_y$ , which are the maximum window sizes of the super pixels along the horizontal and vertical directions such that  $M \bmod d_x$  and  $N \bmod d_y$  evaluate to 0. A set of  $L$  super pixels  $\{SP_1, SP_2, \dots, SP_L\}$  is generated, where  $L$ , the number of super pixels, is computed as in Eq. 1.

$$k_x = \left\lfloor \frac{M}{d_x} \right\rfloor, k_y = \left\lfloor \frac{N}{d_y} \right\rfloor; L = k_x \cdot k_y \quad (1)$$

Generally, for a random variable  $x$ , the probability density function (pdf) following a normal distribution is given in Eq. (2).

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right], -\infty < x < \infty \quad (2)$$

where  $\mu$  is the mean and  $\sigma$  is the standard deviation.

A super pixel  $l \in L$  is associated with a Gaussian distribution defined by the p.d.f  $p(z; \theta_l)$  as in Eq. (3) for the variable  $z$ .

$$p(z; \theta_l) = \frac{1}{(2\pi)^{D/2} \sqrt{\det(\Sigma_l)}} \exp \left[ -\frac{1}{2} (z - \mu_l)^T \sum_t^{-1} (z - \mu_l) \right] \quad (3)$$

where  $\theta_l = \{\mu_l, \Sigma_l\}$  and  $D$  is the number of components in  $z$ .

A pixel  $i$  belongs to a super pixel  $l$ , only if it exists in the pixels set  $I_l$  which is defined by Eqs. (4)–(8), where  $b$  and  $c$  are the lower and upper bounds of the super pixel  $l$ .

$$I_l = \{i | x_{l,b} < x_i < x_{l,c}, y_{l,b} < y_i < y_{l,c}, i \in I\} \quad (4)$$

where

$$x_{l,b} = \max(0, d_x(l_x - 1)) \quad (5)$$

$$x_{l,c} = \min(M, d_x(l_x + 2)) \quad (6)$$

$$y_{l,b} = \max(0, d_y(l_y - 1)) \quad (7)$$

$$y_{l,c} = \min(N, d_y(l_y + 2)) \quad (8)$$

For a super pixel  $l \in L$ ,  $l_x = l \bmod k_x$  and  $l_y = \left\lfloor \frac{l}{k_y} \right\rfloor$ , for any pixel  $i \in \{0, 1, \dots, T - 1\}$  the super pixel Label  $Lb_i$  is computed as in Eq. (9).

**Table 1** Training and testing datasets

Dataset	Infection	No. of images in database	No. of training images	No. of training images after image augmentation	No. of testing images	
					Binary classification	Multiclass classification
<i>Classification</i>						
Kaggle CXR	COVID19	1143	443	1772	700	700
Kaggle CXR	Viral pneumonia	1345	645	2580	1400	700
Kaggle CXR	Normal	1341	641	2564	1400	700
<i>Segmentation</i>						
TCIA	COVID19	221	100	300	115	115



$$Lb_i = \arg_l \max_{l \in L_i} \frac{p(z; \theta_l)}{\sum_{l \in L_i} p(z; \theta_l)} \quad (9)$$

From the above descriptions, it is understood that the GMM super pixels exhibit a good trade-off between accuracy and regularity. Further, the above computations can be parallelized increasing the speed of segmentation.

### 4.3 Cuckoo Search Optimization

Several image processing problems with nonlinear constraints are modeled as optimization problems solved by meta-heuristic algorithms. CSO is a kind of bio-inspired approach characterized by its simplicity and minimum number of tuning parameters. This algorithm has been successfully employed in complex medical image segmentation problems in segmenting the anatomical structures and RoI.

In this research, McCulloch's algorithm, encoded from the stable [36] random variable generation to model levy flights, is employed in levy flight generation and the Kapur's Entropy (KE) [37] is optimized for infection segmentations. The classical CSO algorithm is based on the following assumptions.

1. A cuckoo randomly chooses a nest at a time and lays an egg
2. The nests with best eggs will be carried over to the next generations
3. Always the total number of available host nests is fixed

A host can discover a foreign egg with a probability  $p_a \in [0, 1]$ , on which it may throw the egg or destroy the nest. This algorithm is extended to grouping the super pixels to construct the RoI, where each super pixel is analogous to a nest. A super pixel is either rejected or merged with other super pixels maximizing the Kapur's entropy, discriminating the infected regions from the normal regions. The CSO algorithm starts with a random solution and iteratively updates it by finding an alien egg until the solution converges. Similarly, in super-pixel-based RoI segmentation, a super pixel is treated to be a part of a RoI after evaluating its fitness and is either merged with the prospective super pixels or discarded.

### 4.4 Residual SqueezeNet Classifier Model

Compared to the conventional CNNs, the basic SqueezeNet classifier characterized by fewer hyper parameters is considerably smaller and highly accurate. In this research, a variant of this model called the Residual SqueezeNet which employs bypass connections to improve the learning accuracy is used in the classification of the CXR images. The architecture of the Residual SqueezeNet and its building blocks are depicted in Fig. 2. This network comprises a sequence of fire modules

which consists of a squeeze layer with  $1 \times 1$  filters in the first layer followed by an expansion layer with  $1 \times 1$  and  $3 \times 3$  filters. Finally, the output layer concatenates the outputs of the filters from the expansion layer. The skip connections enable the network to learn an identity function to learn the residual between the output and input, increasing the speed of learning. In this research, this architecture is employed in the construction of binary and three class classification models. Figure 2a shows the three class classifier based on the Residual SqueezeNet and construction of the CAM.

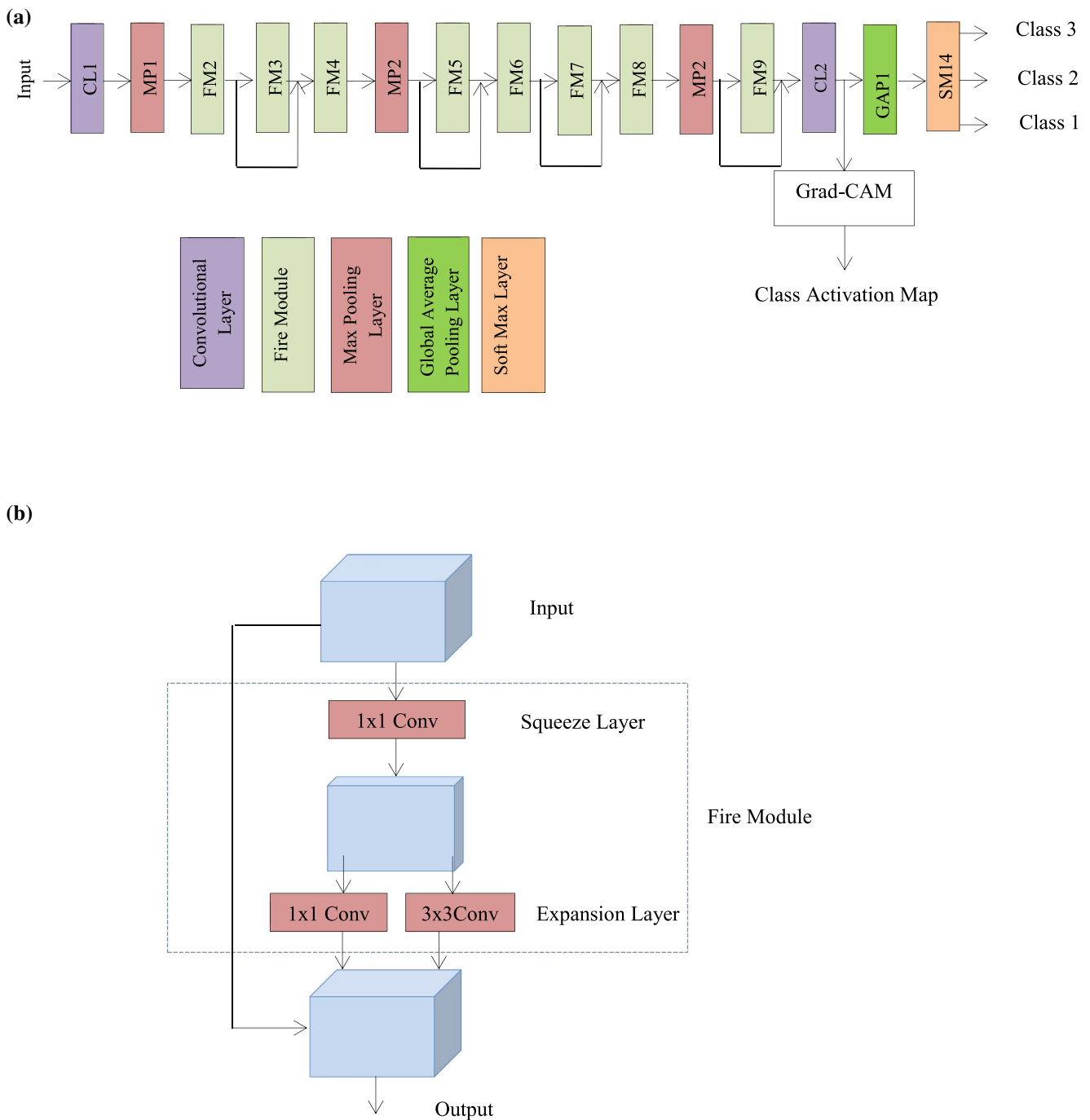
### 4.5 Semantic Segmentation Network

The SegNet meant for road scene segmentation which follows an encoder-decoder architecture is fine-tuned in this research to segment the opacity regions. The schematic of this network is shown in Fig. 3. The encoder is built with convolution layers from the VGG-16 network and maxpooling layers and the convolution and upsampling operations are performed at the decoder. The final layer is a softmax layer which classifies each pixel of the decoder output to construct the segmentation masks.

## 5 Proposed System

The proposed COVID19 detection and opacity segmentation system is realized as a two-stage framework with a classifier subnetwork cascaded to a semantic segmentation subnetwork. The schematic of this model is given in Fig. 4 and the four phases of implementation are described in the following subsections. This system harnesses the potential of deep transfer learning in COVID19 detection and infection localization. It follows the domain adaptation transfer learning approach when the source and target domains are different. By this approach, the deep learning models trained with a large dataset of a source domain are tweaked to adapt to a target domain with a comparatively smaller dataset. The residual SqueezeNet and the GMM super pixel driven SegNet employed in this framework are the novel and original contributions of this research. Basically, the conventional SqueezeNet trained with more than a million of images classifies input images into categories of objects such as mouse, key board, and animals. In this research, the pretrained SqueezeNet is strengthened with residual connections for faster learning and fine-tuned with the CXR images, preserving the original weights. This approach significantly improves the prospective of the classifier network to learn discerning features from CXR images for improved COVID19 classification.

Similarly, the original SegNet trained for indoor and road scene segmentation for object detection is fine-tuned with the RoIs extracted from the CXR images. While the



**Fig. 2** a. Residual SqueezeNet architecture, b. Residual SqueezeNet building blocks

learned weights of the basic VGG-16-based semantic segmentation network are preserved, the softmax layers of this network are modified as a 2-class classifier to classify each pixel of the segmented RoI as background or infection. The proposed framework offers significant advantages in terms of training speed, accurate classifications and localization of infections.

## 5.1 Image Preprocessing

The images in the kaggle CXR dataset are resized to  $227 \times 227$  to match the input size of the image input layer of the pretrained residual SqueezeNet. The TCIA dataset contains chest X-ray images and corresponding segmentation masks in JPG and PNG formats, respectively. These images are of varying large dimensions such as





Fig. 3 SegNet architecture

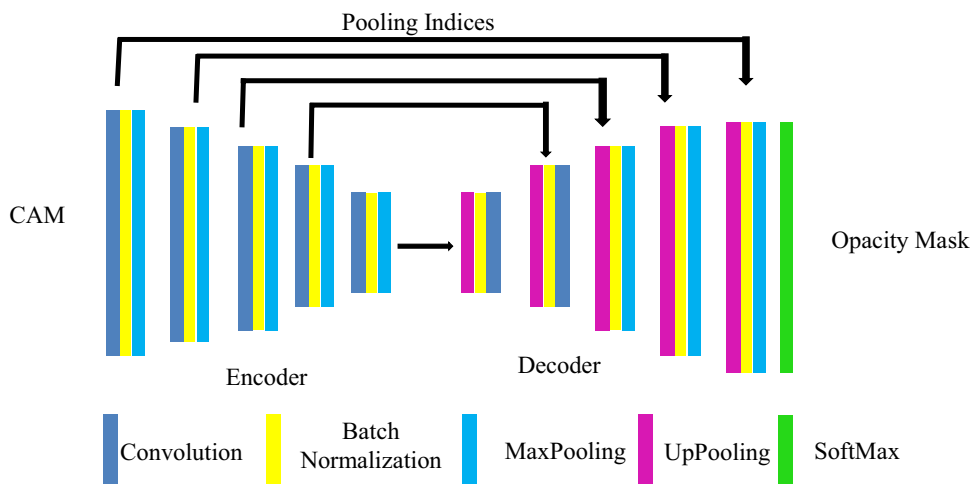
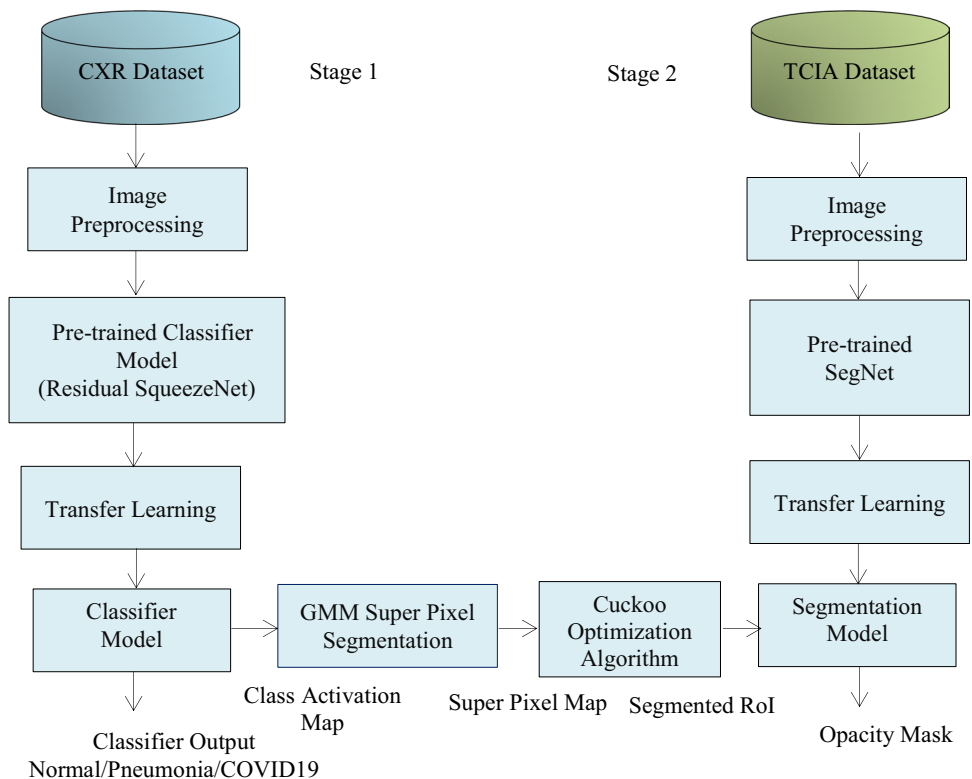


Fig. 4 RSqz-SegNet COVID19 detection and infection segmentation network



1760 × 1760, 2846 × 2330 etc. and they are uniformly resized to 256 × 256.

### 5.2 COVID19 Detection with Residual SqueezeNet

As shown in Fig. 4, the COVID19 detection subnetwork is modeled as a multiclass classifier which is built from the Residual SqueezeNet pretrained with the ImageNet [38] dataset. Initially, each CXR image is convolved with the convolution filters and maxpooling is applied. This is

followed by a sequence of fire modules with skip connections and a max pooling layer. The output from the final fire module is convolved with the final convolutional layer. In CNN classifier models, Generalized Gradient-weighted Class Activation Mapping (Grad-CAM) [39] is employed in the construction of class localization maps from the gradients of the target classes to analyze the model behavior. In the proposed framework, the Grad-CAM is applied on the CAMs output from the final convolutional layer. The global average pooling layer converts these CAMs into a single

**Table 2** Training parameters of residual SqueezeNet classifier

Parameter	Value
Maximum epochs	100
MiniBatchSize	128
Momentum	0.9000
Learning rate	0.001
Optimization	SGDM
L2 regularization parameter	1.0000e−04

value, based on which the softmax layer provides probability distributions of the target classes. The CAMs constructed from the final convolution layer provide a visual interpretation of the regions which attribute the classifier decision. The training parameters of this classifier subnetwork are given in Table 2.

### 5.3 RoI Segmentation with CSO

Segmentation of the RoI from the CAMs can considerably improve the segmentation accuracy of the opacity regions from the chest X-ray images. In this research, RoI segmentation is performed by clustering the GMM super pixels of the CAMs following the CSO approach as in Algorithm 1.

**Table 3** Training parameters for RoI segmentation

Parameter	Value
No. of super pixels	100
No. of nests	100
Maximum iteration	100
Number of dimensions	4
Levy exponent $\beta$	1.5
Discovery rate of alien solutions $p_a$	0.25
Lower bound	0.001
Upper bound	SGDM

In Algorithm 1, the segmentation problem is conceptualized as a maximization problem to maximize KE in the super pixels. This process is characterized as classifying the super pixels as background or RoI based on the entropy. The random numbers for Levy flight generation are obtained with the McCulloh's algorithm to select the GMM super pixels in each iteration. The super pixels with high entropy are added to the RoI and the others are discarded. For given set of  $L$  super pixels, RoI segmentation is performed applying multilevel thresholding on the GMM super pixels as described below.

#### Algorithm 1-CAM-RoI Segmentation

**Input:** Classification Activation Map  $M$

**Output:** Segmented RoI  $CAMSeg$

**Steps:**

1. Initialize the index variable  $i$ , number of iterations  $n$  and the RoI  $CAMSeg$ 
  - a.  $i \leftarrow 1$
  - b.  $CAMSeg \leftarrow \{NULL\}$
  - c.  $n \leftarrow \text{number of iterations}$
2. Construct GMM super pixels from  $M$  and assign a label to each super pixel  $SP_i$ 
  - a.  $SP_i \leftarrow \{1, 2, \dots, L\}$
3. Assign each super pixel to a nest  $N_i$ 
  - a.  $N_i \leftarrow SP_i \ i \in \{1, 2, \dots, L\}$
4. Define the Objective function  $f(x)$ ,  $x = \{SP_i, i = 1, 2, \dots, L\}$
5. Repeat for  $j \leftarrow 1$  to  $n$ 
  - a. Start a random walk defined by Levy flight
  - b. Compute the fitness function  $f_j(x)$  to get the best nest  $N_b$   
 $N_b \leftarrow \{SP_i, i = 1, 2, \dots, L\}$  such that  $\arg \max(KE(SP_i))$
  - c. Generate new population retaining the best nest  $N_b$
  - d. Reassign the Levy step size
6. Continue
7. Repeat for each super pixel  $SP_i$  in  $N_b$ 
  - a.  $CAMSeg \leftarrow CAMSeg + SP_i$
8. Continue
9. Display  $CAMSeg$



For each super pixel  $SP_i$ , the mean intensity value ranges from 0 to 255. Let the number of distinct intensities of  $N$  nests be  $D$ . The entropy values of  $L$  super pixels are as given in Eqs. (10)–(12).

$$E_1 = - \sum_{i=0}^{t1-1} \frac{p_i}{w_1} \log_2 \frac{p_i}{w_1}, w_1 = \sum_{i=0}^{t1-1} p_i \tag{10}$$

$$E_2 = - \sum_{i=t1}^{t2-1} \frac{p_i}{w_2} \log_2 \frac{p_i}{w_2}, w_2 = \sum_{i=t1}^{t2-1} p_i \tag{11}$$

$$E_L = - \sum_{i=tL}^D \frac{p_i}{w_L} \log_2 \frac{p_i}{w_L}, w_L = \sum_{i=tL}^D p_i \tag{12}$$

where  $t1, t2 \dots tL$  are the maximum intensities of super pixels  $SP_1, SP_2 \dots SP_L$ . For each super pixel  $SP_i$ , the entropy computed and the sum of the entropies of these super pixels are maximized to separate the RoI. This algorithm results in the separation of the RoI of the CAMs contributing to the classifier output. The tuning parameters of the CSO algorithm for RoI segmentation from CAM are given in Table 3.

### 5.4 Opacity Segmentation with SegNet

Opacity regions in COVID19-positive CXR images feature lung consolidations, GGO and haziness. Segmentation of these infections from the chest X-ray images is challenging due to the diversities of these patterns. In this research, the pretrained SegNet semantic segmentation network is trained to segment the opacities from the RoIs segmented from the CAMs. As shown in Fig. 4, the segmentation sub-network is a pretrained encoder-decoder model, with symmetric layer to layer propagation. The encoder–decoder pairs capture feature maps of the CAMs for pixel-wise classification at different resolutions. Boundary information invariant over spatial shifts is very important for localization of the opacities. The maxpooling operation which provides translation invariance combined with

subsampling results in a large spatial context governed by each pixel.

However, this results in smaller feature maps with blurred boundaries which are not desirable for semantic segmentation. In order to match the resolution of the segmented region with that of the input CAM, upsampling is performed by decoder. The segNet stores the maxpooling indices of the feature maps generated by the encoder, to capture the boundary information which is passed to the decoder which generates sparse feature maps at high resolutions. Dense feature maps are generated with the trainable filter banks at the decoder. The final stage of the decoder is connected to the softmax classifier which segments the opacity region by classification of the pixels in the dense feature maps. The training parameters of this segmentation subnetwork are given in Table 4.

## 6 Experimental Results and Comparative Analyses

The classification and segmentation results obtained with the proposed framework are presented in this section.

### 6.1 COVID19 Detection with Residual SqueezeNet

The residual SqueezeNet classifier is initially fine-tuned with the augmented training dataset and tested with the test dataset as in Table 1. The classifier is trained and tested for binary and three class classifications. During training, the training dataset is divided into training and validation subsets in the ratio 70:30 for binary and three class classifications. The Receiver Operating Characteristics (ROC) curves for the classifiers are shown in Fig. 5 which testify their learning accuracy.

These classifiers are tested with the test datasets, and the resulting confusion matrices are shown in Fig. 6. The accuracy, sensitivity, specificity, precision, F1-score and Mathews Correlation Coefficient (MCC) metrics evaluated with Eqs. (13–18) are shown in Table 5.

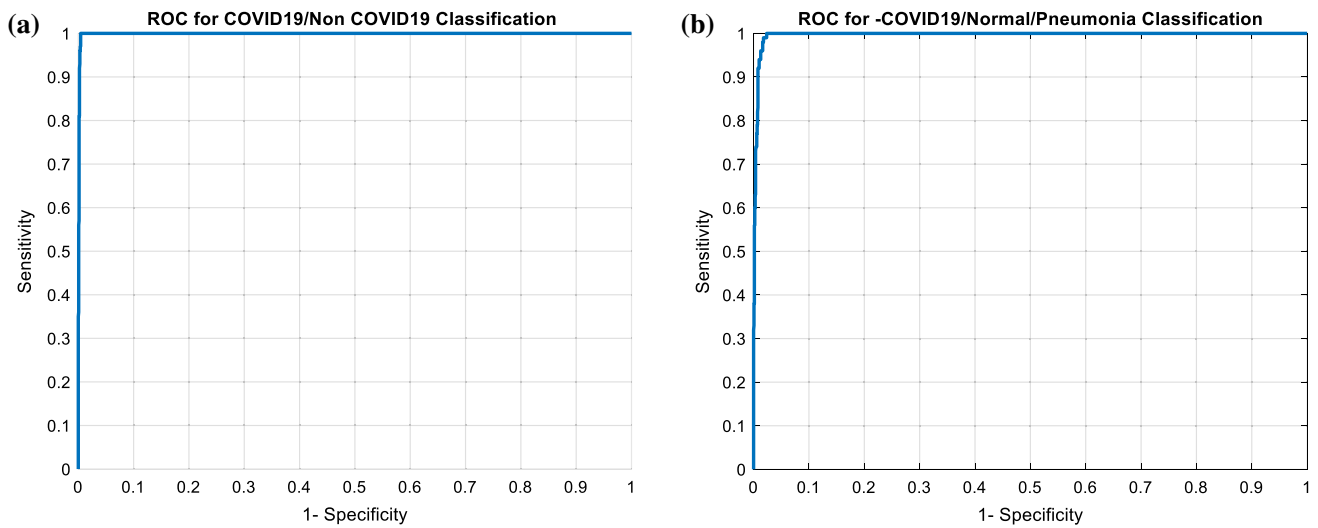
$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{13}$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \tag{14}$$

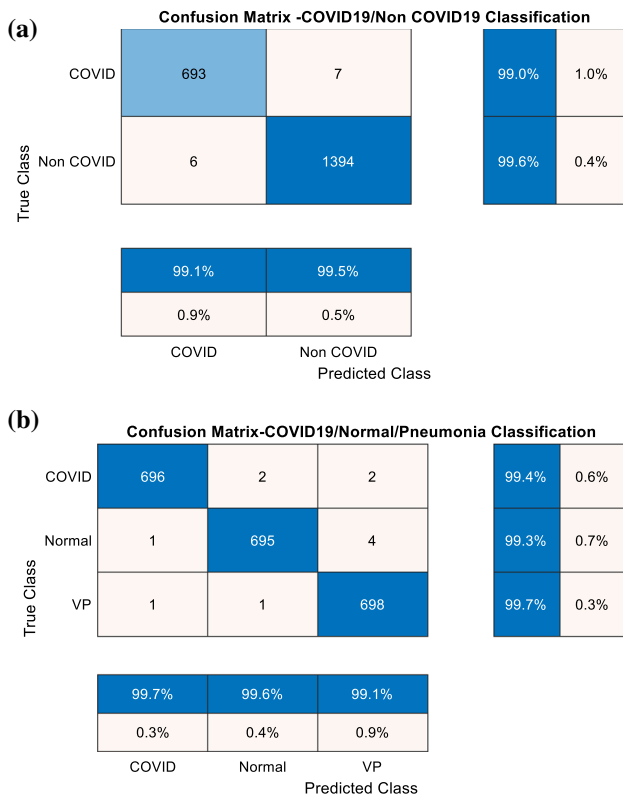
$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{15}$$

**Table 4** Training parameters of SegNet

Parameter	Value
Maximum Epochs	100
MiniBatchSize	128
Momentum	0.9000
Learning rate	0.001
Optimization	SGDM
L2 regularization parameter	1.0000e−04



**Fig. 5** ROC curves **a** Binary **b** Three class classifications



**Fig. 6** **a.** Confusion matrix binary classification—Kaggle CXR dataset, **b** Confusion matrix three class classification—Kaggle CXR dataset

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{16}$$

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \tag{17}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{18}$$

where TP—True Positive, TN—True Negative, FP—False Positive, FN—False Negative.

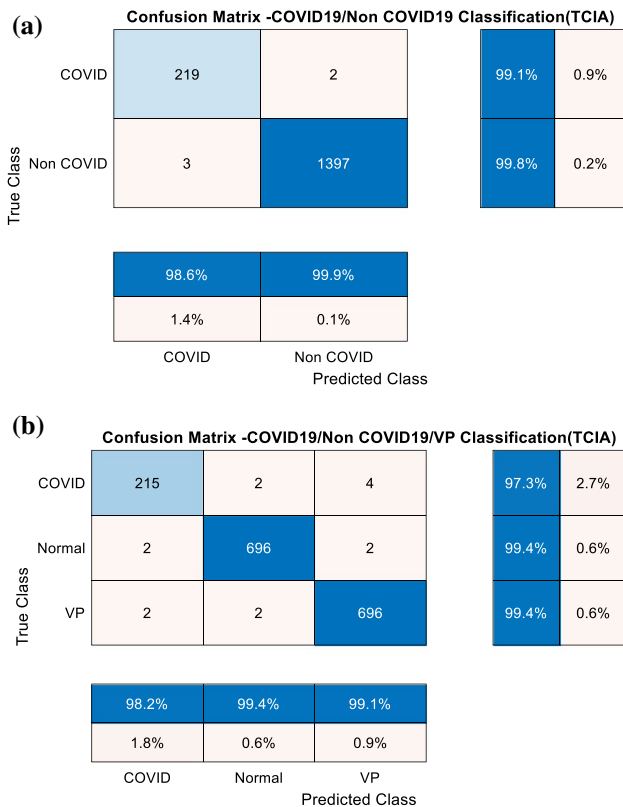
It is seen from Table 5 that best classification accuracies are achieved for three class and binary classifications with the Kaggle CXR and TCIA datasets, respectively. It is evident from Table 1 that the TCIA dataset has only COVID19-positive images. For classification of these images, the training and test datasets are constructed with viral pneumonia and normal images from the Kaggle CXR dataset as Non COVID images. The confusion matrices depicting the performance of the classifiers are shown in Figs. 6 and 7 for the Kaggle CXR and TCIA datasets, respectively, for visual interpretation of the results.

The performance of the proposed model is compared with that of state-of-the-art deep learning models for COVID19 detection from CXR images in Table 6.

The performance metrics show that the classification accuracy of the proposed RSqz-SegNet is superior to other state-of-the-art models. It is also observed that three class classification with TCIA model results in lowest accuracy. However, it is also seen that precision metric is higher for this classification compared to that of binary classification with this dataset. This disparity is attributed to the small number of COVID19 images in this training and testing dataset and the diversities of two datasets. In this research, along with the conventional performance metrics, the

**Table 5** Performance metrics

Dataset	Classifier type	Accuracy	Metrics					
			Sensitivity	Specificity	Precision	FP	F1-score	MCC
Kaggle CXR	Binary	0.9938	0.9900	0.9957	0.9914	0.0043	0.9907	0.9861
	Three class	0.9948	0.9948	0.9974	0.9948	0.0024	0.9952	0.9921
TCIA	Binary	0.9969	0.9910	0.9979	0.9865	0.0021	0.9887	0.9869
	Three class	0.9914	0.9871	0.9954	0.9892	0.0046	0.9881	0.9837



**Fig. 7** a Confusion matrix binary classification—TCIA dataset, b Confusion matrix three class classification—TCIA dataset

balanced measures F1 score and MCC are used to evaluate the models under imbalanced training and testing data. The MCC ranging from  $-1$  to  $+1$  evaluates near to  $+1$  only if the balanced ratio of the TP, FP, TN and FN values is high. A higher value of MCC signifies perfect classification. F1 score is the weighted average of the precision and sensitivity measures where precision is the ratio of the number of correct positive predictions out of the total number of positive predictions and sensitivity is the ratio of the number of correct positive predictions to the total number of positive samples. F1 in the range  $[0, 1]$  evaluates to the ideal value 1 on perfect classification. These measures reported in Table 5 demonstrate the classification efficacy of the proposed RSqz-SegNet.

### 6.2 Opacity Segmentation

After classification, the CAM is constructed from the gradients of each image correctly classified as COVID19 by the residual SqueezeNet classifier subnetwork. This classifier identifies 215 out of 221 images of the TCIA dataset as COVID positive under three class classification as seen in Fig. 7b. The RoI is constructed from these CAMs by super pixel segmentation followed by CSO. The segmented RoI is then fused with the CXR image to fine-tune the pretrained SegNet subnetwork. Basically, this network is trained to predict pixel-wise class labels of the images. The 215 images are divided into training and testing subsets comprising 100 and 115 images, respectively. The training dataset is augmented with 200 more images by geometric transformation of the original hundred images. The SegNet is trained with these images superimposed with the segmented CAMs and the corresponding segmentation masks as labels. This network is then tested with 115 test images fused with the CAMs and the segmentation masks are generated from the pixel-wise class labels predicted by this network. The only work reported in literature, employing the TCIA dataset in [34] performs opacity segmentation with a UNet based on a ResNet backbone and evaluates the segmentation performance with the IoU metric. This metric quantifies the overlap between the ground truth and the predicted segmentation masks. The performance metric reported in [34] and those obtained with the RSqz-SegNet are given in Table 7. In this paper, along with the IoU, the global and mean accuracies of segmentation and Boundary F1 (BF) score are evaluated. The global accuracy is the ratio of correctly classified pixels out of the total number of pixels, irrespective of the target class. Accuracy measures the number of correctly classified pixels in each class. The BF score in the range  $[0, 1]$  measures the similarity of the boundaries of the target and predicted segmentation masks. It is seen that the fine-trained segNet performs better than the UNet.

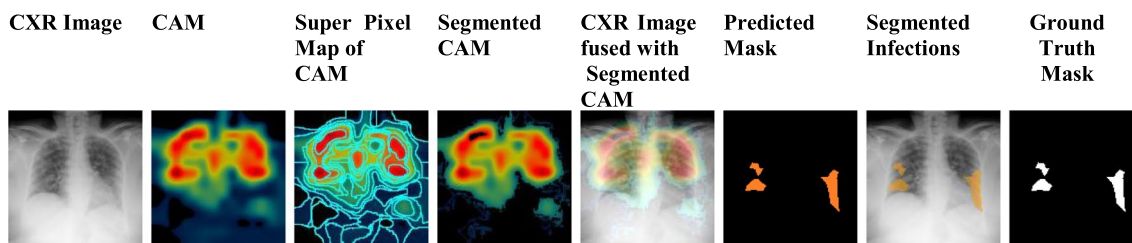
For a given CXR image, the output images generated in the pipeline of the proposed framework are shown in Fig. 8. The input CXR image is characterized by bilateral peripheral patchy infiltrates. The CAM captured from the final convolutional layer of the Residual SqueezeNet highlights the image features influencing the classifier

**Table 6** Classification performance comparison of deep learning models

Reference	Classifier model	Test images	Classification accuracy %
Proposed	RSqz-SegNet	Kaggle CXR dataset	99.38
		700 COVID19	99.48
		1400 Non-COVID19	99.69
		700 COVID	99.14
		700 Viral Pneumonia	
		700 Normal	
		TCIA Dataset	
		221 COVID19	
		1400 Non-COVID19	
		221 COVID	
700 Viral Pneumonia			
700 Normal			
Apostolopoulos et al. [20]	VGG-19	224 COVID19 700 Viral Pneumonia 504 Normal	93.48
Sethy and Behra [25]	ResNet50 + SVM	25 COVID-19(+) 25 COVID-19(-)	95.38
Narin et al. [24]	Deep CNN ResNet-50	50 COVID-19(+) 50 COVID-19(-)	98.0
Chowdhury et al. [30]	SqueezeNet	60 COVID-19	93.3
		60 Viral Pneumonia	96.6
		60 Normal	
Wang and Wong [14]	COVID-Net	60 COVID-19	
		60 Viral Pneumonia	
		60 Normal	
Wang and Wong [14]	COVID-Net	53 COVID-19(+) 5526 COVID-19(-) 8066 Normal	93.3
Hemdan et al. [23]	COVIDX-Net	25 COVID-19(+)	90.0
		25 Normal	

**Table 7** Comparison of segmentation metrics

Reference	Method	No. of test images	Global accuracy	Accuracy	IoU	BF score
Proposed Work	RSqz-SegNet	115	82.95	83.437	89.93	90.34
Tang et al. [34]	UNet	67	–	–	47.55	–



**Fig. 8** Images generated in the RSqz-SegNet pipeline

decision. This CAM is segmented with GMM super pixels from which the RoI is constructed employing the CSO. It is seen that the segmented CAM contains only the RoI. This RoI is fused with the original CXR for segmentation to train the fine-tuned SegNet. Also, the segmented mask

constructed from the SegNet output closely matches the ground truth mask and it is overlaid on the CXR image localizing the infections.

In continuation, the segmentation results obtained with this subnetwork for a few test image samples with and

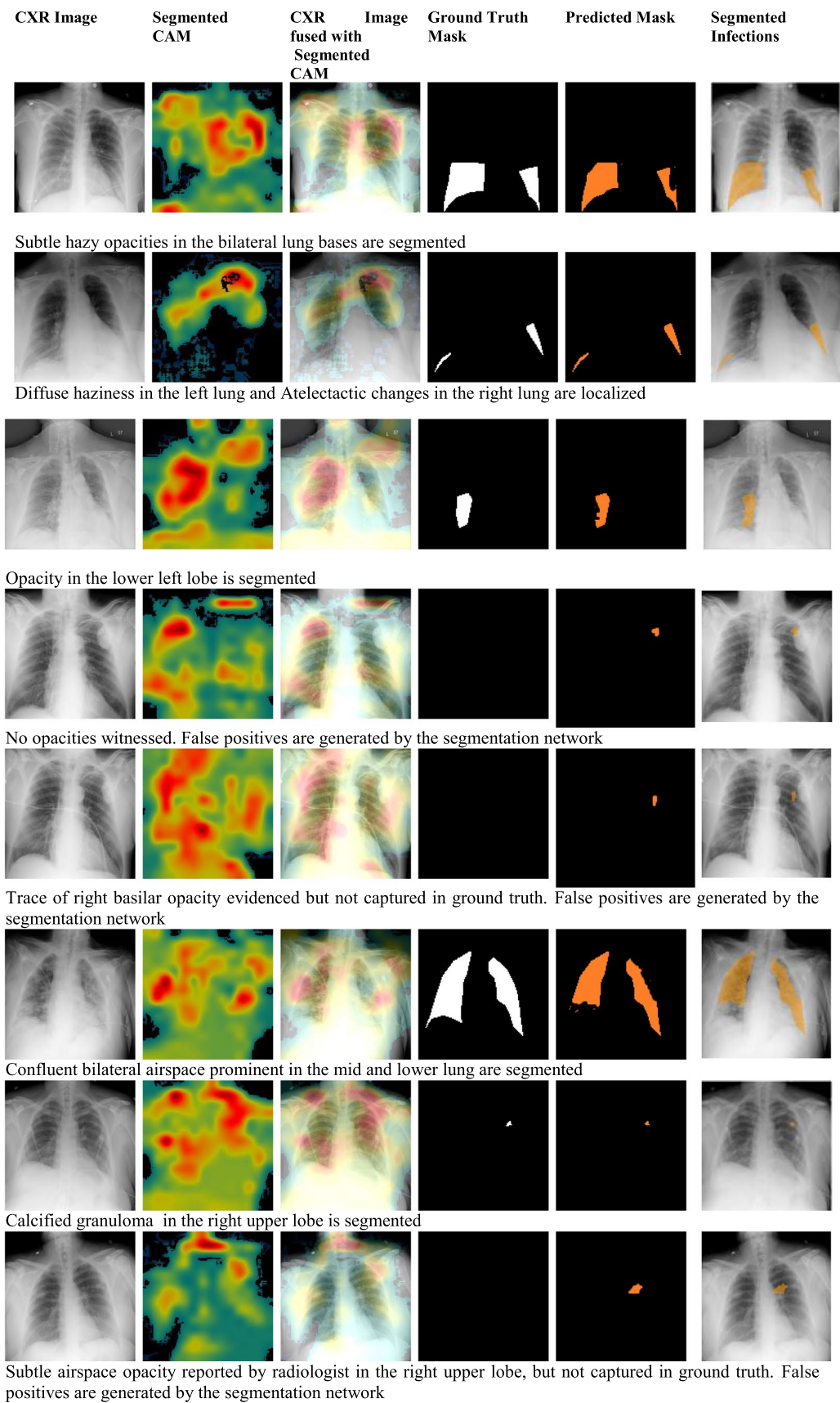


Fig. 9 Segmentation results

**Table 8** Classification performance metrics for ablation study

Dataset	Classifier type	Metrics						
		Accuracy	Sensitivity	Specificity	Precision	FPR	F1-Score	MCC
Kaggle CXR	Binary	0.9839	0.9801	0.9857	0.9815	0.0042	0.9808	0.9762
	Three class	0.9849	0.9845	0.9874	0.9849	0.0023	0.9852	0.9822
TCIA	Binary	0.9869	0.9811	0.9879	0.9766	0.0020	0.9788	0.9770
	Three class	0.9815	0.9772	0.9854	0.9793	0.0045	0.9782	0.9739

without opacities are shown in Fig. 9 with annotations. This figure shows that the proposed SegNet model captures the infections including haziness, opacities and granuloma. Also, the FPs generated by this model are reported in the figure.

The segmentation results show that different kinds of opacities are captured by the SegNet including the hazy and subtle manifestations. The FPs generated are worth further morphological analysis for understanding the infection patterns.

## 7 Discussion

This paper proposes a two-stage model for COVID19 detection and infection segmentation based on deep transfer learning and bioinspired segmentation. Several studies show that CXR images of COVID19 patients are non-specific, variable and overlap with the symptoms of Human parainfluenza, respiratory syncytial viruses, H7N, H7N9 and Middle East respiratory syndrome coronavirus (MERS-COV), which render the detection and segmentation tasks difficult. Generally, GGOs manifest in the COVID19 positive CXR images as basal, medial, peripheral, bilateral, posterior and subpleural regions. However, GGOs are also evidenced unilaterally in CXR images of viral pneumonia cases. Further, airspace consolidations are conjoint with GGOs in COVID19 and viral pneumonia cases. In spite of these inconsistencies, the proposed RSqz-SegNet framework demonstrates exemplary classification performances with a few misclassifications in binary and multiclass classifications in COVID19 detection. It is attributed to the ability of the residual classifier to learn the discriminating features.

Though the classification results are more than 99% for the two datasets for the two classifications categories, the performance of the SegNet model demonstrating a mean accuracy of 63.437% and IoU of 69.93% is not appreciable, concerning the severity of the infections. However, the segmentation results are comparatively better than the most recent representative work reported in [34]. It is evident that segmentation of the CAMs facilitates the localization of GGOs with irregular, hazy and patchy manifestations. Follow-up CXR imaging shows progressive transformations of

the infections resulting in temporal changes of lungs featuring varied severity scores. This framework can be extended for severity assessment in future.

An ablation study is performed with this framework to understand the significance of RoI segmentation from the GMM super pixels of the CAM. The RSqz-SegNet model is modified by removing the RoI segmentation process, and the CXR image fused with CAM is directly given as input to the SegNet. The performance metrics given in Table 8 for this study show a noticeable degradation in the performance of the model when the SegNet is trained with the CXR image overlaid with the CAM.

A few limitations of this research are worth discussion, providing insights into further research. The potential limitations are the lack of standard datasets for infection segmentation from CXR images and integral models for classification and localization. These drawbacks limit the evaluation of the proposed framework with benchmark models. The comparative analyses of the classification and segmentation performances presented in Sect. 6 are based only on the results reported in existing literature. However, exercising the benchmarked models with public datasets can provide better perceptions of the model.

## 8 Conclusion

This paper presents a deep transfer learning based novel two-stage framework called the RSqz-SegNet for COVID19 detection and infection segmentation from CXR images. This framework comprises fine-tuned residual SqueezeNet and SegNet subnetworks for classification and segmentation, respectively. This classifier model is trained and tested with two standard public datasets for binary and three class classifications of COVID19, Viral Pneumonia and Normal cases, which achieves 99.69% and 99.48% classification accuracies for binary and three class classifications, respectively. Similarly, the segmentation subnetwork also achieves a global and mean accuracy of 82.95% and 3.43%, respectively. The performance of the segmentation model is enhanced with RoI segmentation of the CAMs employing GMM super pixels and the meta-heuristic CSO. This research paves a new approach, integrating the classification and segmentation



tasks for COVID19 management. This framework can be further enhanced by improving the segmentation subnetwork towards accurate infection segmentations for defining the potential biomarkers for COVID19 detection and severity evaluation in prognosis management.

## References

- Pereira, R.M.; Bertolini, D.; Teixeira, L.O.; Silla Jr, C.N.; Costa, Y.M.: COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comput. Methods Programs Biomed.* **194**, 105532 (2020)
- Wong, H.Y.F.; Lam, H.Y.S.; Fong, A.H.T.; Leung, S.T.; Chin, T.W.Y.; Lo, C.S.Y.; Lui, M.M.S.; Lee, J.C.Y.; Chiu, K.W.H.; Chung, T.; Lee, E.Y.P.: Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology* **296**, 201160 (2020)
- Choi, H.; Qi, X.; Yoon, S.H.; Park, S.J.; Lee, K.H.; Kim, J.Y., et al.: Extension of coronavirus disease 2019 (COVID-19) on chest CT and implications for chest radiograph interpretation. *Radiol. Cardiothorac. Imaging* **2**(2), e200107 (2020)
- Rubin, G.D.; Ryerson, C.J.; Haramati, L.B.; Sverzellati, N.; Kanne, J.P.; Raoof, S.: The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner society. *Radiology* **296**, 172–180 (2020)
- Gaál, G.; Maga, B.; Lukács, A.: Attention U-net based adversarial architectures for chest X-ray lung segmentation (2020). [arXiv:2003.10304](https://arxiv.org/abs/2003.10304)
- Ban, Z.; Liu, J.; Cao, L.: Superpixel segmentation using Gaussian mixture model. *IEEE Trans. Image Process.* **27**(8), 4105–4117 (2018)
- Yang, X.S.; Deb, S.: Search via Lévy flights. In: 2009 World Congress on Nature and Biologically Inspired Computing (NaBIC), pp. 210–214. IEEE (2009)
- Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K.: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size (2016). [arXiv:1602.07360](https://arxiv.org/abs/1602.07360)
- Badrinarayanan, V.; Kendall, A.; Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017)
- Liu, X.; Deng, Z.; Yang, Y.: Recent progress in semantic image segmentation. *Artif. Intell. Rev.* **52**(2), 1089–1106 (2019)
- Sufian, A.; Ghosh, A.; Sadiq, A.S.; Smarandache, F.: A survey on deep transfer learning to edge computing for mitigating the COVID-19 pandemic. *J. Syst. Archit.* **108**, 101830 (2020)
- Mei, X.; Lee, H.C.; Diao, K.Y.; Huang, M.; Lin, B.; Liu, C., et al.: Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat. Med.* **26**(8), 1224–1228 (2020)
- Zhang, K.; Liu, X.; Shen, J.; Li, Z.; Sang, Y.; Wu, X., et al.: Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* **181**(6), 1423–1433 (2020)
- Wang, L.; Lin, Z.Q.; Wong, A.: COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. *Sci. Rep.* **10**(1), 1–12 (2020)
- Pandit, M.K.; Banday, S.A.: SARS n-CoV2–19 detection from chest x-ray images using deep neural networks. *Int. J. Pervasive Comput. Commun.* **16**, 419–427 (2020)
- Maghdid, H.S.; Asaad, A.T.; Ghafoor, K.Z.; Sadiq, A.S.; Mirjalili, S.; Khan, M.K.: Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms. In: *Multimodal Image Exploitation and Learning 2021*, vol. 11734, pp. 117340E. International Society for Optics and Photonics (2021)
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Adv. Neural. Inf. Process. Syst.* **25**, 1097–1105 (2012)
- Alafif, T.; Tehame, A.M.; Bajaba, S.; Barnawi, A.; Zia, S.: Machine and deep learning towards COVID-19 diagnosis and treatment: survey, challenges, and future directions. *Int. J. Environ. Res. Public Health* **18**(3), 1117 (2021)
- Salehi, S.; Abedi, A.; Balakrishnan, S.; Gholamrezanezhad, A.: Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *Am. J. Roentgenol.* **215**, 1–7 (2020)
- Apostolopoulos, I.D.; Mpesiana, T.A.: Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **43**, 1 (2020)
- Apostolopoulos, I.; Aznaouridis, S.; Tzani, M.: 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](https://arxiv.org/abs/2004.00338)
- Farooq, M.; Hafeez, A.: Covid-resnet: a deep learning framework for screening of COVID19 from radiographs (2020). [arXiv:2003.14395](https://arxiv.org/abs/2003.14395)
- Hemdan, E.E.D.; Shouman, M.A.; Karar, M.E.: Covidx-net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images (2020). [arXiv:2003.11055](https://arxiv.org/abs/2003.11055)
- Narin, A.; Kaya, C.; Pamuk, Z.: Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* **24**, 1207–1220 (2021). <https://doi.org/10.1007/s10044-021-00984-y>
- Sethy, P.K.; Behera, S.K.: Detection of coronavirus disease (Covid-19) based on deep features (2020). <https://doi.org/10.20944/preprints202003.0300.v1>
- Hinton, G.E.; Sabour, S.; Frosst, N.: Matrix capsules with EM routing. In: *International Conference on Learning Representations* (2018)
- Afshar, P.; Heidarian, S.; Naderkhani, F.; Oikonomou, A.; Plataniotis, K.N.; Mohammadi, A.: Covid-caps: a capsule network-based framework for identification of COVID-19 cases from x-ray images. *Pattern Recogn. Lett.* **138**, 638–643 (2020)
- Ucar, F.; Korkmaz, D.: COVIDiagnosis-Net: deep Bayes-SqueezeNet based diagnostic of the coronavirus disease 2019 (COVID-19) from X-ray images. *Medical Hypotheses*, 109761 (2020)
- Li, X.; Zhu, D.: Covid-xpert: an AI powered population screening of COVID-19 cases using chest radiography images (2020). [arXiv:2004.03042](https://arxiv.org/abs/2004.03042)
- Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Mahbub, Z.B.; et al.: Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access* **8**, 132665–132676 (2020)
- Alom, M.Z.; Rahman, M.M.; Nasrin, M.S.; Taha, T.M.; Asari, V.K.: COVID\_MTNNet: COVID-19 detection with multi-task deep learning approaches (2020). [arXiv:2004.03747](https://arxiv.org/abs/2004.03747)
- Alom, M.Z.; Aspas, T.; Taha, T.M.; Asari, V.K.: Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks (2019). [arXiv:1904.11126](https://arxiv.org/abs/1904.11126)
- El-bana, S.; Al-Kabbany, A.; Sharkas, M.: A multi-task pipeline with specialized streams for classification and segmentation of infection manifestations in covid-19 scans. *PeerJ Comput. Sci.* **6**, e303 (2020)



34. Tang, H.; Sun, N.; Li, Y.: Segmentation model of the opacity regions in the chest X-rays of the Covid-19 patients in the us rural areas and the application to the disease severity. medRxiv (2020)
35. Muhammad, T.R.; Chowdhury, E.H.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Mahbub, Z.B.; Islam, K.R.; Khan, M.S.; Iqbal, A.; Al-Emadi, N.; Ibne Reaz, M.B.: COVID-19 Radiography Database (2020). <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>
36. Chambers, J.M.; Mallows, C.L.; Stuck, B.W.: A method for simulating stable random variables. *J. Am. Stat. Assoc.* **71**(354), 340–344 (1976)
37. Kapur, J.N.; Sahoo, P.K.; Wong, A.K.: A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Vis. Graph. Image Process.* **29**(3), 273–285 (1985)
38. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
39. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D.: Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626 (2017)

