



Differentiation of COVID-19 conditions in planar chest radiographs using optimized convolutional neural networks

Satyavratn Govindarajan¹ · Ramakrishnan Swaminathan¹

Accepted: 11 September 2020 / Published online: 6 November 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020, corrected publication 2021

Abstract

In this study, an attempt has been made to differentiate Novel Coronavirus-2019 (COVID-19) conditions from healthy subjects in Chest radiographs using a simplified end-to-end Convolutional Neural Network (CNN) model and occlusion sensitivity maps. Early detection and faster automated screening of the COVID-19 patients is essential. For this, the images are considered from publicly available datasets. Significant biomarkers representing critical image features are extracted from CNN by experimentally investigating on cross-validation methods and hyperparameter settings. The performance of the network is evaluated using standard metrics. Perturbation based occlusion sensitivity maps are employed on the features obtained from the classification model to visualise the localization of abnormal areas. Results demonstrate that the simplified CNN model with optimised parameters is able to extract significant features with a sensitivity of 97.35% and F-measure of 96.71% to detect COVID-19 images. The algorithm achieves an Area Under the Curve-Receiver Operating Characteristic score of 99.4% with Matthews correlation coefficient of 0.93. High value of Diagnostic odds ratio is also obtained. Occlusion sensitivity maps provide precise localization of abnormal regions by identifying COVID-19 conditions. As early detection through chest radiographic images are useful for automated screening of the disease, this method appears to be clinically relevant in providing a visual diagnostic solution using a simplified and efficient model.

Keywords Chest radiograph · COVID-19 · Convolutional neural network · Occlusion sensitivity · Visualisation

1 Introduction

Novel Coronavirus-2019 (COVID-19) is a pandemic affecting 212 countries and territories worldwide. As on May 82,020 report by the World Health Organisation, 37,59,967 confirmed cases and 2,59,474 deaths are reported [1]. The infections are seen to rise exponentially and rapidly, where an infected person transmits the disease to 406 individuals within 30 days. There is an urgent need for automated screening and early diagnosis of the disease to contain its prevalence [2].

Currently, Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is the gold standard diagnostic test for confirming COVID-19 patients. However, this method has been reported to suffer from high false negative rates and is

time consuming [3]. Evidences of imaging manifestations show promising directions to improve sensitivity in the detection. The imaging characteristics of COVID-19 are subtly different from other types of viral pneumonia [4]. Typical radiological manifestations of COVID-19 include patchy to confluent ground glass opacities distributed peripherally with or without consolidation appearing in bilateral lower lung zones [5, 6].

Chest Radiographic (CXR) imaging is considered as the initial diagnostic modality and primary screening tool to detect several respiratory pathologies such as Tuberculosis (TB) and Pneumonia [7, 8]. This modality is reliable and portable by avoiding risk of any disease exposure especially in remote settings [8]. Recent guidance by the World Health Organisation also suggests the prospective uses of chest imaging for COVID-19 patient diagnostics and therapy.

Computer Aided Diagnostic (CAD) systems driven by Artificial Intelligence (AI) have proven to be effective for disease screening [9]. The introduction of fully automated models in a CAD system has significantly reduced inter and intra observer variability in radiological examinations and improve diagnostic relevance through rapid screening rates.

✉ Satyavratn Govindarajan
s.vratan@gmail.com

¹ Non-Invasive Imaging and Diagnostics Laboratory, Biomedical Engineering Group, Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai, India

However, false negative rates produced are still high enabling continuous research potential in this field. Convolutional Neural Networks (CNN) have been the core technique to detect and localise disease manifestations and differentiate anomalies. These methods have been successfully employed for several radiographic applications [10–12]. These models utilise larger training samples to achieve better performance. Studies show that better performance can be achieved with tailor-made CNN trained from scratch [12, 13]. Simplified CNN models are also employed with shallow architecture for limited training datasets [14–16].

There have been significant efforts in the literature to detect and distinguish COVID-19 patients from normal using CXRs and CNNs [8]. Custom CNN models have been developed to classify CXR images with better accuracy [17]. In another study by Rahimzadeh [18], a concatenation of Xception and ResNet50V2 networks are performed on a combination of large dataset with a smaller number of COVID-19 X-ray images and achieved high performance. A deep stacking network is reported to yield better performance with the usage of transfer learning to detect COVID-19 from larger and smaller datasets [13]. These models employ image pretraining, data augmentation and transfer learning techniques by creating large datasets to achieve better performance [8, 17]. However, these networks require more computations for training and inference, and may not suit to medical tasks consisting of a limited amount of data [9].

Visualising and interpreting the network behaviour helps to validate the model's localization performance in assessing subtle characteristics and suspicious biomarkers associated with abnormal findings in medical images [10]. Besides quantitative validation of CNN model using performance measures, Explainable AI methods are essential to provide a visual diagnostic approach [19]. There have been attempts to visualise critical regions of COVID-19 using deep networks using Class Activation Map (CAM) methods based on gradients such as Grad-CAM, GradCAM++ to provide class-discriminating regions [8, 17, 20]. However, they tend to scatter from the areas of COVID-19 providing imprecise localizations [13, 17, 20]. Hence, sensitive methods that capture finer and subtle abnormal findings based on CNN's features are necessary. Occlusion sensitivity maps would provide appropriate visual diagnosis by delivering finer details of abnormality localization using sensitivity of softmax scores [10, 21].

The objective of this framework is to perform a pilot study for automated differentiation of COVID-19 patients from healthy subjects using chest X-ray images. For this, an efficient and simplified end-to-end CNN model for a limited training dataset is developed by extracting critical image features representing significant CXR biomarkers to identify COVID-19. The model is optimized by exploring its feasibility and efficiency in terms of its cross-validation analysis and hyperparameter setting. Besides quantitative validation using

performance measures, “Explainable AI” method is used to qualitatively validate the features obtained from CNN. This is performed using occlusion sensitivity maps which provides a visual diagnostic approach based on precise localization of abnormal regions within the lung fields.

Rest of the paper is organized as follows: Methods is described in section II, results are illustrated in section III, discussion is presented in section IV, followed by conclusions in section V and references.

2 Methods

This section describes about the image datasets, proposed CNN model, performance evaluation and occlusion sensitivity maps.

2.1 Image database

In this study, two publicly available datasets are considered for healthy and COVID-19 CXR images respectively.

- COVID-19 image collection:** The CXR images of COVID-19 patients are considered from a publicly accessible dataset [22]. This dataset contains CXR and CT images of various bacterial and viral pneumonias especially COVID-19 [20]. The images are being collected from various public sources, hospitals and physicians. It is part of a project approved by the University of Montreal Ethics Committee, which aims at improving prognostic predictions to triage and manage patient care. The objective behind the project is to develop AI based techniques for the diagnosis and prediction of infectious disease. As of April 28, 2020, an updated version of this dataset consists of 310 chest X-ray images in three views belonging to viral and bacterial pneumonia conditions with varied demographics. The dataset also contains patient metadata with clinical notes and patient details such as age, sex, survival information, date of image acquisition and name of the admitted hospital for most of the images. Bounding box denoting regions are provided for the detection of problematic regions in images. In this study, 151 COVID-19 CXR images in Posterior-Anterior (PA) view are used (refer Table 1).
- The Shenzhen set - Chest X-ray Database:** The CXR images of healthy subjects are considered from a public dataset [23]. This dataset has been created and published by the National Library of Medicine, Maryland, USA, in an effort to provide healthy and TB CXR images for training machine learning algorithms [24]. The dataset was obtained from Shenzhen No.3 People's Hospital, Guangdong Medical College, Shenzhen, China. The

Table 1 Dataset demographics

Category	COVID-19 image collection (As on April 28, 2020)	Shenzhen set
Number of images present	354 (CXR + CT)	662
Total Number of CXR images	310	662
Number of normal CXR images	–	326
Number of abnormal CXR images	310 (No Finding - 3)	336
Findings	Viral and bacterial pneumonias such as ARDS, Chlamydothila, COVID-19, E.Coli, Klebseilla, Legionella, Pneumocystis, Streptococcus, SARS	TB abnormalities
COVID-19 CXR images	247	–
CXR Image view	Frontal view (PA, AP, Lateral)	Frontal View (PA, AP)
Image type	RGB/Gray scale	RGB/Gray scale
Image format	PNG/JPEG/JPG	PNG
Dot Per Inch (DPI)	Variable	72
Bit depth	Variable	8
Image resolution	Variable	3 K × 3 K approx.

CXR images belong to the hospital outpatient clinics, and are acquired within a 1-month duration in September 2012 using a DR digital diagnose system from Philips [9, 23]. The patient metadata such as age, gender and TB abnormality are provided as separate files. The database has been exempted from IRB review for public use. In this study, CXR images in PA view pertaining to 150 healthy subjects are considered from this dataset.

Table 1 shows the demographic details of the two datasets. In this study, all the images corresponding to both the datasets are converted into grayscale format and down-sampled to be fixed at a resolution of 512×512 . Figure 1 shows Representative CXR original images of (a and b) healthy subjects and (c and d) COVID-19 patients considered from the datasets.

2.2 Proposed CNN model

To extract critical features from specific CXR biomarkers, a simple CNN deep network is employed to differentiate healthy and COVID-19 CXR images. CNN is the most popular deep learning method which provides the optimal architecture for image recognition and classification tasks by eliminating the use of hand-crafted features [25, 26]. The pipeline of the proposed approach with simplified and efficient CNN architecture is shown in Fig. 2.

The end-to-end CNN model considered in this study comprises convolutional sections (including batch normalization), max pooling layers, followed by a fully connected layer, a softmax layer and an output classification layer.

- **Input Image Layer:** This layer comprises of input CXR images with image sizes corresponding to the height, width, and the grayscale channel size.
- **Convolutional Section:** In the convolutional layers, filter window size and number of filters representing the number of feature maps are mentioned. It contains batch normalization to perform inbuilt optimization to accelerate network training and reduce overfitting. It normalizes each channel across a mini-batch. Then, Rectified Linear Unit (ReLU) non-linear activation layers are followed [25].
- **Max Pooling Layer:** These layers perform down-sampling to reduce the spatial size of the feature layer and eliminates redundant information.
- **Fully Connected Layer:** A fully connected layer connects all neurons of the network in the preceding layer by combining all the features extracted and trained by the preceding layers across the image to identify the biomarker patterns [27].
- **Softmax Layer and Classification layer:** The non-linear softmax activation normalizes the output of the previous layer. The output of this layer can then be used as probability scores by the classification layer to compute the loss.

2.2.1 Model construction

Firstly, the input images of size $512 \times 512 \times 1$ are subjected to CNN network. Each convolutional layer is made of varied number of filters as 8, 16, 32, 64, 128, 256 and 512 with varying filter sizes 3×3 and 5×5 [28]. The convolutions are all zero-padded in order to preserve the input resolution [9]. Further, each convolutional 2d-layer is followed by batch normalization layer and ReLU activation layers. An epsilon value of $1e-5$ is chosen in batch normalization.

Fig. 1 Representative CXR original images of (a and b) healthy subjects and (c and d) COVID-19 patients

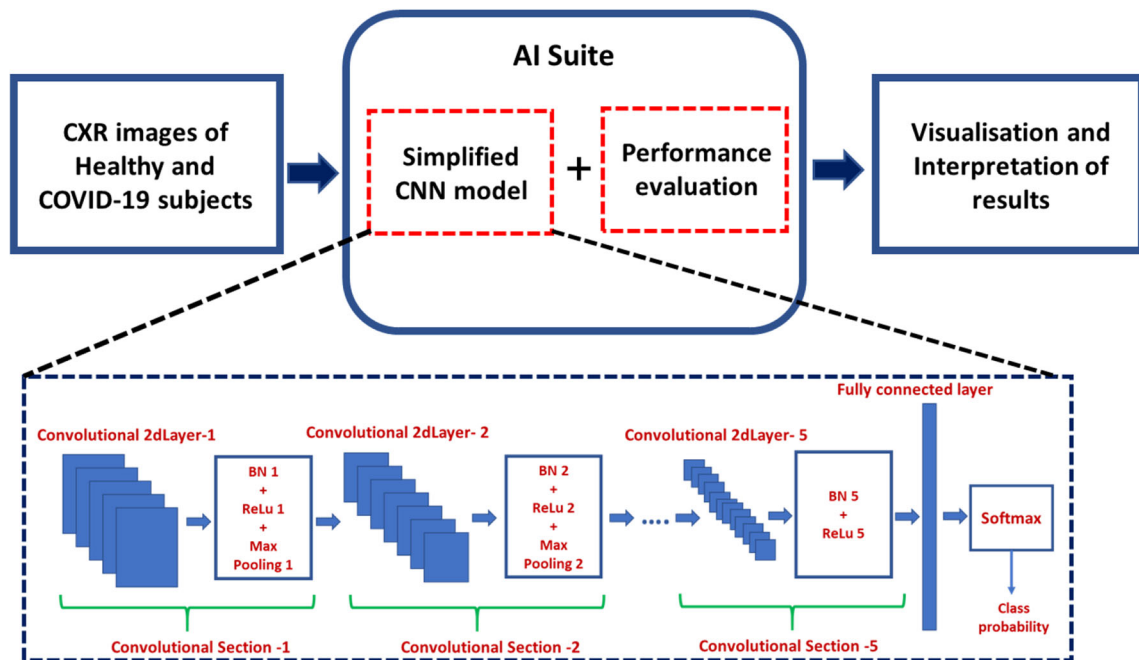
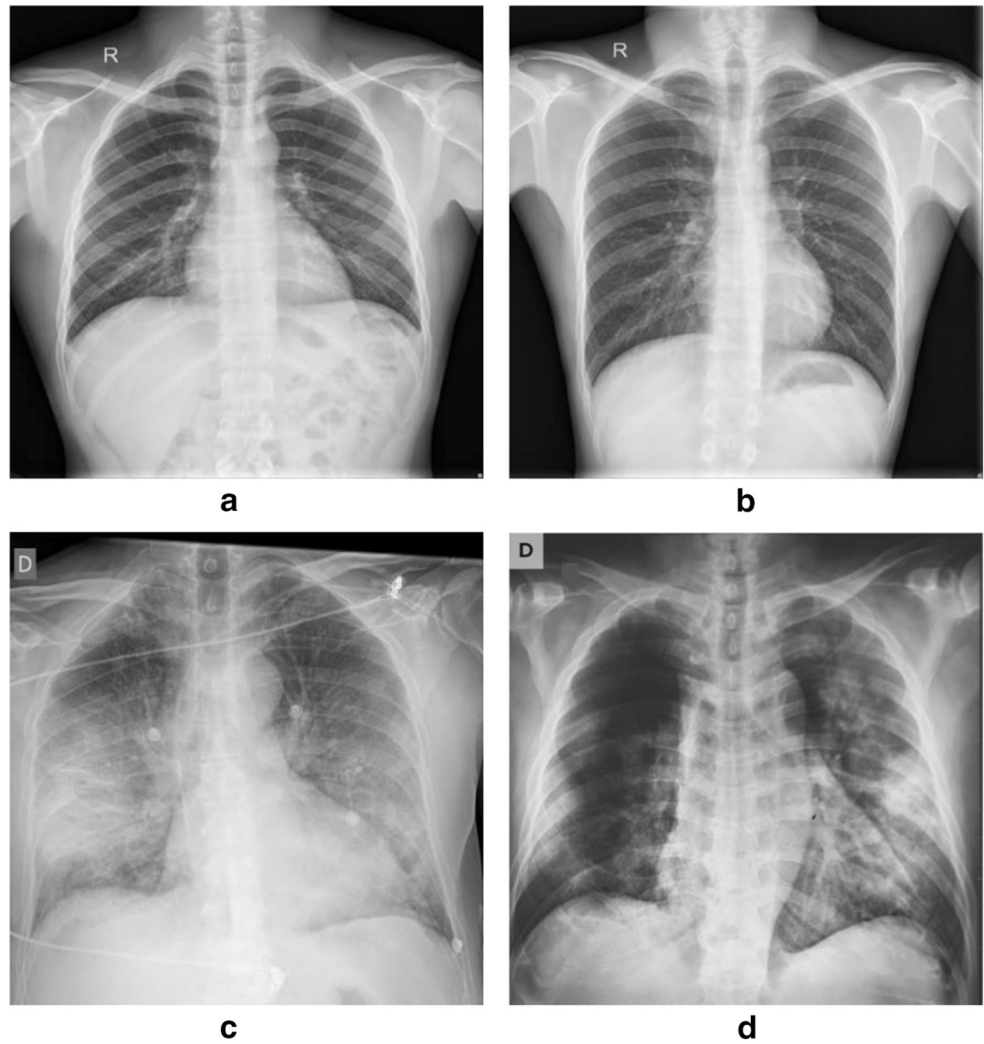


Fig. 2 Pipeline of the proposed methodology (BN – Batch normalisation, ReLu – Rectified linear unit)

Further, a max pooling layer is followed with a stride of [2] and pool size of 2 [25]. The feature maps obtained are fed to a fully connected layer, followed by the softmax layer and a classification layer. In this work, the output size of 2, corresponding to the two classes is fixed. The weights are initialized using the Glorot method on all layers where each weight is initialized from Gaussian distribution with a mean value of zero and finite variance. Finally, the classification scores are obtained to generate visualization-based occlusion sensitivity maps.

2.2.2 Model evaluation

Effect of comparisons in 5-fold and 10-fold cross-validation [9, 17] on classification accuracy is experimentally analyzed for different number of CNN layers. This is done to investigate on the cross-validation technique that yields the best performing feature test set [29]. This is in-line with literature stating that cross-validation analysis is important to construct a simplified network that provides high performance [30]. Number of CNN layers such as 3, 5 and 7 are selected for this purpose [12].

Network training is carried out by examining the effect of filter sizes on classification accuracy for different learning rates [0.001, 0.01] at different number of CNN layers [3, 5, 7]. Mini-batch samples and maximum epochs are selected empirically as 64 and 30 respectively with data shuffling at every epoch [9, 31]. In this work, varying CNN filter sizes of 3×3 and 5×5 are selected for experimental analysis [28]. Categorical cross-entropy is used as the error function and Adam optimizer is tested [9]. Parameters of Adam optimizer such as gradient decay factor, squared gradient decay factor and initial learn rate are fixed at 0.9, 0.999 and 0.001 respectively.

2.2.3 Model selection

An efficient architecture comprising of optimal number of layers with optimal hyperparameters for the corresponding model validation method that yields high performance in detection of COVID-19 is selected based on the obtained results. This provides a simplified and efficient CNN model optimized to achieve high classification performance, so as to help even unskilled technicians and practitioners of the medical community.

The whole system is trained and tested in about 20 min using Nvidia GeForce GTX 1050 Ti, Intel I7 processor, 16 GB RAM in parallel execution environment. The experiments are performed using MATLAB® 2019b software.

2.3 Performance evaluation

To evaluate the classification performance in differentiating COVID-19 and healthy images, various performance metrics namely, sensitivity, specificity, precision, F-measure, and Area Under the Curve-Receiver Operating Characteristic (AUC-ROC) are used [9, 32]. Average values of folds are reported in this study. In addition to these, to measure the effectiveness of detection and to compute the quality of binary classifier, Diagnostic odds ratio [33] and Matthews correlation coefficient (MCC) [34] metrics are calculated.

Diagnostic odds ratio: This ratio measures the efficacy of a diagnostic test and considered to be the single indicator of test performance. High value of Diagnostic odds ratio could be helpful [33] to medical practitioners in evaluating COVID-19 diagnoses. It is calculated by,

$$\text{Diagnostic odds ratio} = \frac{\text{sensitivity} * \text{specificity}}{(1 - \text{sensitivity}) * (1 - \text{specificity})}$$

2.4 Occlusion sensitivity maps

To identify whether the efficient CNN model is able to detect specific locations of significant healthy and COVID-19 biomarkers, occlusion sensitivity maps are used. Occlusion sensitivity maps provide localization information about certain specific areas in the image, when a grayscale mask is placed through the whole CXR image generating a probability map [10, 21]. It is used to analyse the network sensitivity to the occlusions of image regions [35]. These maps provide high spatial resolution and finer details.

It works when the occluded images are passed to CNN network and a Euclidean distance is calculated for each iteration. The difference between the occluded distance and non-occluded distance is computed. This difference increases once a patch occludes an area in the image relevant to the network, thus creating a heat map. The pathological regions correspond to higher probability and a drop in the value indicates that the pathological locations have been occluded. In this study, a mask size of 15 and stride value of 10 pixels are heuristically chosen. Thus, COVID-19 specific CXR biomarkers can be localised providing an approximate visual diagnosis.

3 Results

The following section contains image description, CNN model selection, Performance evaluation and diagnostic visualisation of CXR images.

3.1 Image description on healthy and COVID-19 chest radiographs

A representative set of CXRs of healthy and COVID-19 subjects in PA view are shown in Fig. 1. In Fig. 1(a and b), it is visually observed that the lung fields have better image contrast with respect to background structures. It might be due to lung sacs that are filled with sufficient amount of air without forming any white obscurity/opacity. However, the images suffer from intensity inhomogeneity across lung fields and along the radiolucent ribs due to overlay of sub-structures. The lung fields display structural dissimilarity across subjects.

In Fig. 1(c and d), the lung fields exhibit poor contrast with respect to surrounding regions. This might be due to the radiological manifestations that restrict the flow of air inside the lungs and its passages. Biomarkers such as patchy and segmental white opacities distributed in central, perihilar and lung periphery are observed in the bilateral lungs which depict intensity variations locally. In Fig. 1(c and d), these markers appear as homogenous in intensity with surrounding opaque regions such as rib contours, bony areas, mediastinal reflections and heart vessels. These opaque markers co-exist in multiple lung areas showing as indefinite image attributes demonstrating a major challenge to detect and differentiate COVID-19 CXRs from healthy images.

3.2 CNN model selection

A comparison of 5-fold and 10-fold cross-validation methods are performed on classification accuracy for different hidden layers of CNN (refer Table 2). It can be seen that high accuracy values are obtained for 5-fold as compared to 10-fold cross-validation technique. It is also observed that there is an increase in accuracy values from 3-layer to 5-layer and 7-layer networks. However, it can be seen that the accuracies are similar with a higher value of 96.6% in 5-layer and 7-layer networks against the 3-layer model. This might be due to the fact that the CNN could have extracted most of the relevant features in 5-layer architecture for the differentiation of classes.

The experimental analysis of the effect of filter sizes on classification accuracy for 3, 5 and 7 layers of CNN at two different learning rates are shown in Fig. 3. From Fig. 3(a), it is observed that a maximum accuracy value of 96.6% is obtained at 5-layer CNN with filter size 3×3 as compared to the

Table 2 Comparison of cross-validation methods on classification accuracy (in %) for different layers of CNN

Layers	K = 5	K = 10
3	95.3	94.3
5	96.6	95.3
7	96.6	95.3

accuracy of 96% with filter size 5×5 at learning rate of 0.001. Similar is the case in Fig. 3(b) with 5-layer CNN reporting a maximum accuracy using filter size 3×3 in comparison with filter size 5×5 and other layers. Among Fig. 3(a) and (b), it is seen that filter size 3×3 at a learning rate of 0.001 yields the highest accuracy with 5-layers as against the learning rate 0.01.

From the above results, it is seen that a depth of 5-layer convolutional blocks is chosen. Network training and testing using 5-fold cross-validation is found to yield better results. The optimal set of parameter values which provided maximum accuracy are obtained to be filter size of 3×3 , a learning rate of 0.001, batch size of 64 and maximum epochs of 30 (refer Table 3).

3.2.1 Performance evaluation

The classification performance of the proposed CNN model with 5-layers is quantitatively evaluated using performance measures (refer Table 4). A maximum sensitivity value of 97.35% is obtained in the detection of COVID-19 images. The percentage specificity value is also found to be better in identifying healthy subjects not being detected as diseased. Similarly, precision and F-measure values are found to be high (> 96%). It can be noted that equivalent values of F-measure for healthy and COVID-19 classes denote better classification performance by 5-layered CNN model. Diagnostic odds ratio is obtained with a high value (882) indicating a better test performance of the diagnostic CNN model. Quality of this binary classifier is also found to be better with perfect prediction of two classes using MCC measure.

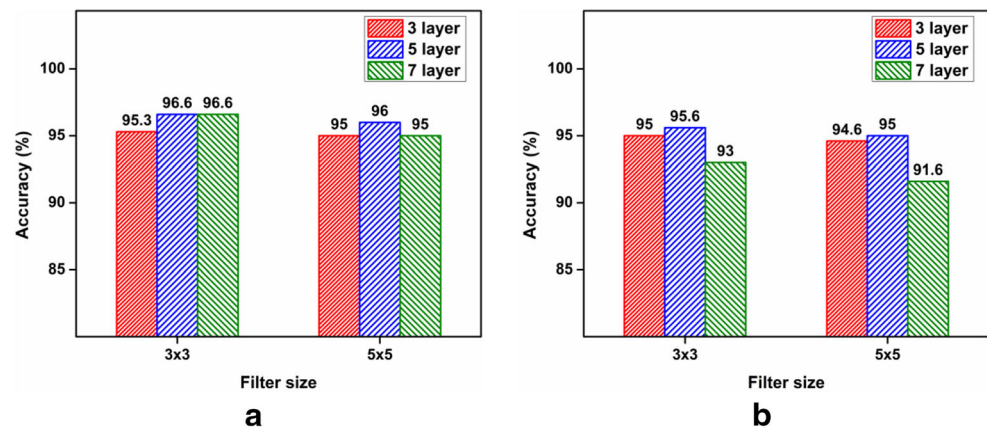
Further, a confusion matrix is generated to analyse the proportion of images being classified and misclassified. In Fig. 4(a), the number of images belonging to true positive (TP), false positives (FP), true negatives (TN) and false negative (FN) classes for healthy and COVID-19 subjects are shown.

It is observed that the greater number of images are identified as TP. This states that the CNN model is able to detect a greater number of COVID-19 images correctly with low FN. Similarly, TNs are also detected more precisely by the model with 145 healthy images being correctly classified out of 150 with low false negative images. From Fig. 4(b), ROC analysis shows that a high AUC value of 0.994 is obtained indicating better classification predictability with 99.4% chance that the 5-layer CNN model distinguishes the two classes. This is evident from the curve showing increased threshold points at higher sensitivity close to 1.0.

3.2.2 Diagnostic visualisation: Correctly classified

To understand the diagnostic information obtained from CNN model, it is essential to visualise the image regions responsible

Fig. 3 Effect of filter sizes on classification accuracy for different number of layers with (a) Learning rate = 0.001 and (b) Learning rate = 0.01



for abnormality detection through feature activations of the network. Figure 5 shows occlusion sensitivity maps and their extracted lung field regions for representative healthy and COVID-19 subjects. The localization is mainly obtained in the lung field regions. The first and second rows of Fig. 5(a) show a representative chest radiograph truly classified as healthy. The occlusion sensitivity maps shown in Fig. 5(b) (Row 1 and Row 2) have blue areas highlighted in entire lung field regions and red areas as surrounding regions. These findings are correctly identified as healthy by the CNN model and in line with the database annotations.

As lungs are the region of interest, this can be visualised in the extracted lung field maps shown in Fig. 5(c) (Row 1 and Row 2). Here, small spots of red markings are visible inside lung fields. However, there are no increased signal variations in any of the regions. The first and second rows of Fig. 5(d) show a representative chest radiograph truly classified as COVID-19. The first row of Fig. 5(d) shows a hazy patchy and streaky opacities in the bilateral lung bases covering a large portion in the lower zones. This might indicate a high severity of disease.

Similarly, the second row of Fig. 5(d) shows a striated hazy opacity near the middle lower lobes evolving from the lung borders. This might also represent a highly severe condition. These findings are consistent with COVID-19. In the occlusion sensitivity maps shown in Fig. 5(e), the areas inside lung fields are highlighted in red. It can also be seen that the areas adjacent to red regions inside lung fields denote colour changes between greenish yellow to brownish red. This correlates well with the pathological changes seen in the original image and inline with the database annotations indicating an increased density of abnormality. These visualizations represent

positive contributions to classification and are better represented through extracted lung field maps shown in Fig. 5(f).

3.2.3 Diagnostic visualisation: Misclassified

The first row of Fig. 6 shows a healthy subject which is wrongly classified as positive for COVID-19. Figure 6(a and d) represents original healthy and COVID-19 images respectively. The occlusion sensitivity map in Fig. 6(b and c) shows brighter red areas along the lung borders and greenish yellow to brownish red areas inside lung fields suspicious for COVID-19. This area attributes to much decreased radiolucency, as this may be the case for a pediatric with reduced air inspiration. This is inline with the patient history of a 3-year-old subject. The second row corresponds to COVID-19 subjects misclassified as negative. In the second row of Fig. 6, the dataset annotations describe that patchy ill-defined subpleural opacities are found in the middle zone of the right lung. However, the occlusion sensitivity maps and extracted lung field maps shown in Fig. 6(e and f) couldn't reveal the associated pathology as whole of the lung fields represent blue areas corresponding to negative contribution to classification.

4 Discussion

This work focusses on the differentiation of healthy subjects from COVID-19 CXR images using a simplified and efficient end-to-end CNN model and occlusion sensitivity maps. In this work, the images are obtained from public datasets and are trained and tested using a simplified end-to-end CNN model for classification of the two classes. Cross-validation analysis and hyperparameters settings are performed experimentally to

Table 3 Custom CNN model selection based on maximum classification performance

Parameters	Convolutional layers	Cross validation	Filter size	Learning rate	Mini-batch samples	Max. epochs
Optimal values	5	5-fold	3 × 3	0.001	64	30

Table 4 Performance measures obtained for healthy and COVID-19 images using CNN

Performance measures	Healthy	COVID-19
Sensitivity (%)	96.00	97.35
Specificity (%)	97.35	96.00
Precision (%)	97.30	96.08
F-measure (%)	96.64	96.71
Diagnostic odds ratio		882
Matthews correlation coefficient		0.93

construct an efficient model optimized for the considered problem. Further, perturbation-based visualization method is employed to capture finer details of localization of abnormal regions within lung fields, so as to provide approximate visual diagnostics helpful in the clinical setting.

Based on the obtained results, the proposed tailor-made model is able to extract significant features corresponding to healthy and COVID-19 regions from CXRs with better discriminating ability. From the results of cross-validation analysis, high network performance with 5-fold cross-validation could be due to large number of test images that are validated at every fold thus, providing the best performing feature test set than 10-fold method. Maximum accuracy obtained for filter 3×3 with a learning rate of 0.001 implies that the network is able to better capture local complex features in massive amount of information with better weight sharing, making it an efficient network. The model provides maximum performance using small learning rate which indicates that it is able to learn the set of weights in an optimal way with a small learning rate providing better input-output mapping.

From the experimental investigations, it is found that 5-layer CNN provides maximum classification performance with a simplified model optimised for the detection of COVID-19. Five convolutional blocks correspond to a

receptive window covering the whole input image, and this window size allows the network to access a large context for its decisions at each location [9]. This is also evident from the quantitative measures of sensitivity, specificity, precision, AUC-ROC, F-measure and Diagnostics odds ratio. The quality of the binary classifier is validated to be high using MCC.

Results of Occlusion sensitivity maps show that subtle and non-specific abnormal findings of COVID-19 could be identified with better visualization by providing finer localizations of regions. It is obtained that the visual information from the maps correlates well with the pathological changes seen in the original image and inline with the database annotations indicating evidence of abnormality.

A detailed discussion and comparison with existing studies that report the implementation of CNN models and visualization methods for the detection of COVID-19 using CXR images are presented (refer Table 5). Only peer-reviewed articles are considered for comparison. Ozturk et al. [20] performed DarkNet with 17 convolutional layers and Grad-CAM based visualization technique. Limitations of this work correspond to an imbalanced dataset, and imprecise localization of abnormal areas. A maximum sensitivity value of 90.65% is reported for the differentiation of healthy and COVID-19 images.

Brunese et al. [8] applied VGG-16 based CNN model and Grad-CAM visualization method for significant number of healthy images when compared to COVID-19 CXRs. A sensitivity value of 87% is reported for binary classification. The work proposed to investigate the use of verification techniques to obtain better results in the future.

Another study by Mahmud et al. [13] used a stacked multi resolution-based CNN model with balanced data set for the classification of healthy and COVID-19 pneumonia. Although high sensitivity and AUC have been obtained, a scattering in CNN feature localizations out of a region of interest has been reported. Rajaraman et al. [36] implemented pretrained models and a residual network on a very small collection of COVID-19 data. This work performed data

Fig. 4 Model performance shown as (a) Confusion matrix and (b) ROC analysis

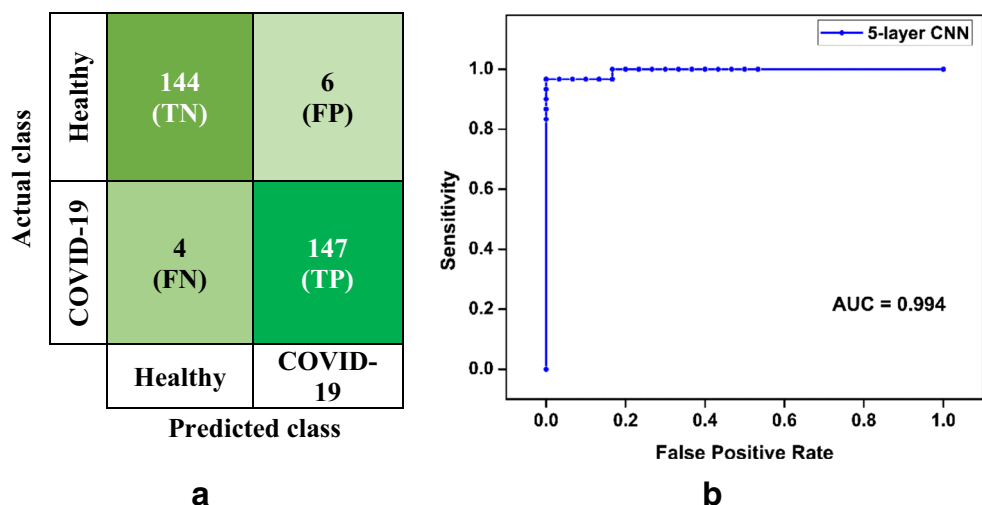
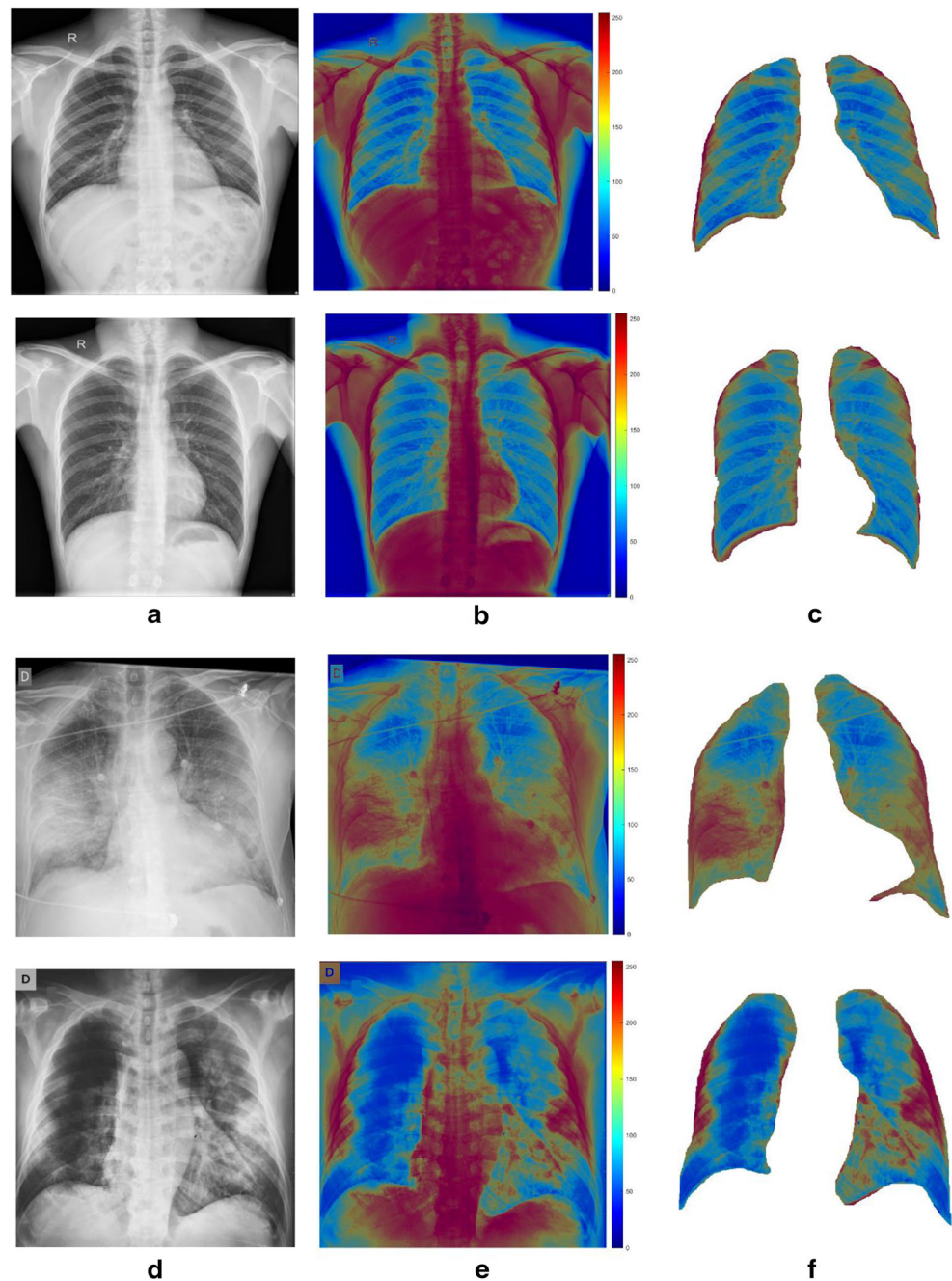


Fig. 5 Correctly classified images: Original images of healthy (a), COVID-19 subjects (d) and overlay of occlusion sensitivity maps (b and e). Extracted lung field maps (c and f)



augmentation technique. Large imbalance between the classes hinders the appropriate selection of augmented training images. A Truncated Inception Net proposed by Das et al. explored different sets of input X-ray images achieved high sensitivity and AUC values. However, poor localization of abnormal areas using activation maps could be observed.

From the discussion, it could be observed that most of the works utilized imbalanced dataset for binary classification of healthy and COVID-19 CXR images. CAM based visualization approaches are predominantly reported to validate the quality of features obtained from deep architectures. Such

approaches illustrated imprecise localization of COVID-19 regions.

To summarise, this study rules out the usage of such pretrained networks, which have been established for natural image classification tasks. The pretrained networks require more computations for training and inference and may not suit to medical tasks consisting of limited amount of data.

Secondly, previous studies have shown the visualization and interpretability task on a superficial level. A deeper understanding of radiological aspects provides deeper insights to

Fig. 6 Misclassified images: Original images of healthy (a), COVID-19 subjects (d) and their corresponding occlusion sensitivity maps (b and e). Extracted lung field maps (c and f)

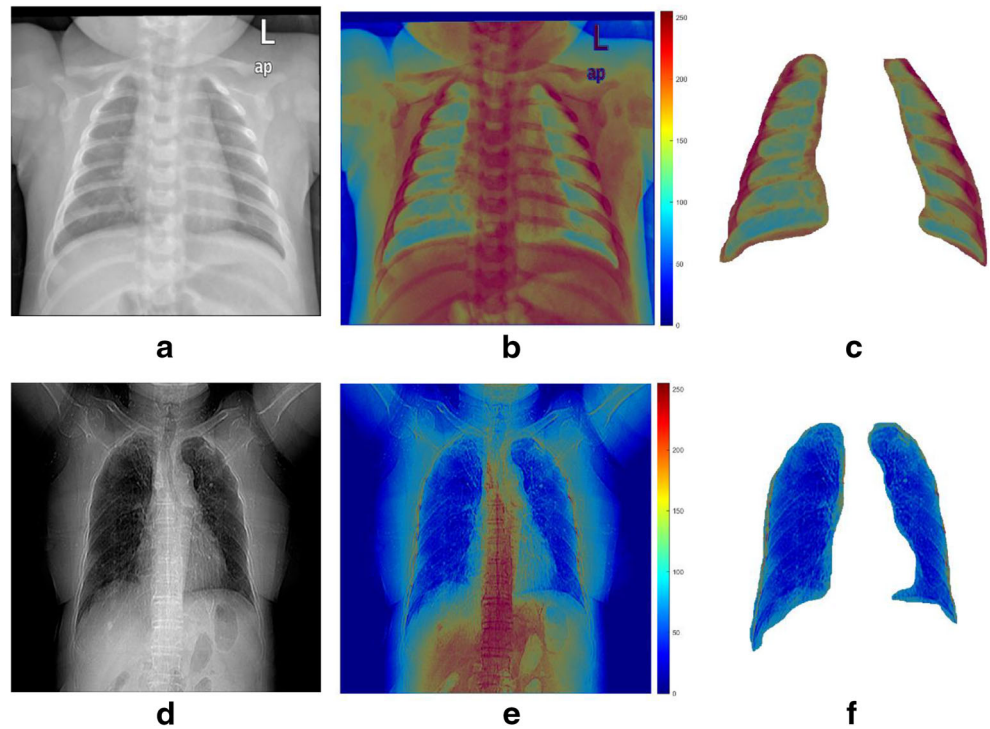


Table 5 Discussion and comparison with existing studies

Author	Images considered		Architecture used	Visualization method	Validation method	COVID-19 detection Results (%)	Remarks/Limitations
	Healthy	COVID-19					
Ozturk et al. [20]	1000	250	DarkNet-17	Grad-CAM	5-fold CV	Sensitivity: 90.65 AUC-ROC: –	Imprecise localization of areas on the chest region
Brunese et al. [8]	3520	250	VGG-16	Grad-CAM	CV	Sensitivity: 87 AUC-ROC: –	Proposed to investigate if formal verification techniques can be helpful to obtain better results
Mahmud et al. [13]	305	305	Stacked Multi-resolution CovXNet	Grad-CAM	5-fold CV	Sensitivity: 97.8 AUC-ROC: 96.9	Scattering in gradient based localizations out of the region of interest
Rajaraman et al. [36]	1583	314	Wide residual network and pretrained models	Grad-CAM	Random Split	Sensitivity: – AUC-ROC: –	Very small collection of COVID-19 data to select augmented training images, Imbalanced dataset and Imprecise localization of areas on the chest region belonging to COVID-19
Das et al. [17]	D1:1583 D2: 80	162 162	Truncated Inception Net	Activation map	10-fold CV	Sensitivity: 95 AUC-ROC: 99	Maximum values are reported for imbalanced dataset. Poor localization of areas of COVID-19
<i>Proposed Work</i>	150	151	<i>CNN - 5</i>	<i>Occlusion sensitivity</i>	<i>5-fold CV</i>	<i>Sensitivity: 97.35 AUC-ROC: 99.4</i>	<i>Simplified, efficient CNN network for limited dataset Perturbation based visualization method for precise localization</i>

gain trust, transparency and can be an additional opinion in the medical community. This work provides such visual diagnosis which can be performed on low-end performance computational systems. The results of visualization are found to be remarkable, even if absence of any annotation is considered. Thirdly, the study emphasizes the importance of Diagnostic odds ratio metric, which measures the effectiveness of a diagnostic test in a medical setting. High value of Diagnostic odds ratio could be helpful to medical practitioners in evaluating COVID-19 diagnoses.

5 Conclusion

The coronavirus pandemic has had a large impact on global health and well-being. Early detection of the disease is essential to reduce its impact for directing proper treatment decisions and strategies. Automated AI based diagnoses using chest radiographs provide a rapid and reliable solution for patient screening of COVID-19 abnormalities, especially in remote and resource-poor regions. In this work, an attempt has been made to develop a simplified and efficient end-to-end CNN model optimised for the differentiation of COVID-19 from healthy subjects based on the identification of significant image biomarkers at specific locations of chest radiographs. Experimental investigations on model validation methods and hyperparameters settings of the model are performed by extracting critical features.

Maximum sensitivity and AUC-ROC values demonstrate the discriminating ability of the considered CNN model in the detection of Novel coronavirus-19 images. The proposed network could offer faster chest X-ray screening by reducing the computational requirement significantly using less power-hungry hardware so as to help even unskilled technicians and practitioners of the medical community. Low false positives from classification results suggest that many healthy subjects need not undertake strenuous treatment process and quarantine procedures as COVID-19 patients. Perturbation based Occlusion sensitivity maps provide localization of abnormal areas by indicating subtle findings and non-specific alterations with inter and intra patient variability. This automated CAD system could be clinically useful to assist radiologists to take precise decisions through visual diagnosis in minimal amount of time in healthcare settings. When the resources are scanty, such visually assisted tools help medical practitioners to triage patients by highlighting the pathological areas despite the lack of fine-grained annotations. This helps in easy decision making and treatment strategy. Future work would involve full-scale research on implementing deeper architectures of CNN models using large image datasets and performing lung field segmentation for better prognosis of the disease.

Data availability The data used in the manuscript is publicly accessible dataset that has been constantly updated. Check Reference [22] for details.

Compliance with ethical standards

Conflicts of interest/competing interests The authors express no conflicts of interest by any means in the manuscript preparation.

Ethics approval Not Applicable.

Consent to participate Not Applicable.

Consent for publication The authors provide full consent for submission and publishing the manuscript in the journal.

References

1. World Health Organization (2020) Coronavirus disease 2019 (COVID-19): situation report-109 http://www.who.int/docs/default-source/coronaviruse/situation-reports/20200508covid-19-sitrep-109.pdf?sfvrsn=68f2c632_6. Accessed 8 May 2020
2. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med* 27:1–4. <https://doi.org/10.1093/jtm/taaa021>
3. Karim MR, Döhmen T, Rebholz-Schuhmann D et al (2020) DeepCOVIDExplainer: explainable COVID-19 predictions based on chest X-ray images. arXiv preprint arXiv:2004.04582
4. Butt C, Gill J, Chun D, Babu BA (2020) Deep learning system to screen coronavirus disease 2019 pneumonia. *Appl Intell* 24:1. <https://doi.org/10.1007/s10489-020-01714-3>
5. Bai HX, Hsieh B, Xiong Z et al (2020) Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology*: 200823. <https://doi.org/10.1148/radiol.2020200823>
6. Yoon SH, Lee KH, Kim JY, Lee YK, Ko H, Kim KH, Park CM, Kim YH (2020) Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea. *Korean J Radiol* 21:494–500. <https://doi.org/10.3348/kjr.2020.0132>
7. Govindarajan S, Swaminathan R (2019) Analysis of tuberculosis in chest radiographs for computerized diagnosis using bag of keypoint features. *J Med Syst* 43:87. <https://doi.org/10.1007/s10916-019-1222-8>
8. Brunese L, Mercaldo F, Reginelli A, Santone A (2020) Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comput Methods Prog Biomed* 196: 105608. <https://doi.org/10.1016/j.cmpb.2020.105608>
9. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D (2019) Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization. *Sci Rep* 9:6268. <https://doi.org/10.1038/s41598-019-42557-4>
10. Islam MT, Aowal MA, Minhaz AT et al (2017) Abnormality detection and localization in chest X-rays using deep convolutional neural networks. arXiv preprint arXiv:1705.09850
11. Varma M, Lu M, Gardner R, Dunnmon J, Khandwala N, Rajpurkar P, Long J, Beaulieu C, Shpanskaya K, Fei-Fei L, Lungren MP, Patel BN (2019) Automated abnormality detection in lower extremity radiographs using deep learning. *Nat Mach Intell* 1:578–583. <https://doi.org/10.1038/s42256-019-0126-0>
12. Mendoza J, Pedrini H (2020) Detection and classification of lung nodules in chest X-ray images using deep convolutional neural networks. *Comput Intell* 36:370–401. <https://doi.org/10.1111/coim.12241>

13. Mahmud T, Rahman MA, Fattah SA (2020) CovXNet: a multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput Biol Med* 122:103869. <https://doi.org/10.1016/j.combiomed.2020.103869>
14. Tajbakhsh N, Suzuki K (2017) Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recogn* 63:476–486. <https://doi.org/10.1016/j.patcog.2016.09.029>
15. Ganapathy N, Veeranki YR, Swaminathan R (2020) Convolutional neural network based emotion classification using Electrodermal activity signals and time-frequency features. *Expert Syst Appl* 159:113571. <https://doi.org/10.1016/j.eswa.2020.113571>
16. Xu S, Wu H, Bie R (2018) CXNet-m1: anomaly detection on chest X-rays with image-based deep learning. *IEEE Access* 7:4466–4477. <https://doi.org/10.1109/ACCESS.2018.2885997>
17. Das D, Santosh KC, Pal U (2020) Truncated inception net: COVID-19 outbreak screening using chest X-rays. *Phys Eng Sci Med*: 1–11. <https://doi.org/10.1007/s13246-020-00888-x>
18. Rahimzadeh M, Attar A (2020) A new modified deep convolutional neural network for detecting COVID-19 from X-ray images. *arXiv preprint arXiv:2004.08052*
19. Xie N, Ras G, van Gerven M, Doran D (2020) Explainable deep learning: a field guide for the uninitiated. *arXiv preprint arXiv:2004.14545*
20. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 121:103792. <https://doi.org/10.1016/j.combiomed.2020.103792>
21. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: Fleet D, Pajdla T, Schiele B et al (eds) *Computer vision – ECCV 2014*. Springer, Cham, pp 818–833
22. Cohen JP, Morrison P, Dao L (2020) COVID-19 image data collection. *arXiv preprint arXiv:2003.11597*
23. Jaeger S, Candemir S, Antani S, Wang YX, Lu PX, Thoma G (2014) Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imaging Med Surg* 4:475–477. <https://doi.org/10.3978/j.issn.2223-4292.2014.11.20>
24. Jaeger S, Karagyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Zhiyun Xue, Palaniappan K, Singh RK, Antani S, Thoma G, Yi-Xiang Wang, Pu-Xuan Lu, McDonald CJ (2014) Automatic tuberculosis screening using chest radiographs. *IEEE Trans Med Imaging* 33:233–245. <https://doi.org/10.1109/TMI.2013.2284099>
25. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9:611–629. <https://doi.org/10.1007/s13244-018-0639-9>
26. Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A (2019) Comparison of deep learning approaches for multi-label chest X-ray classification. *Sci Rep* 9:6381. <https://doi.org/10.1038/s41598-019-42294-8>
27. Wang SH, Muhammad K, Hong J, Sangaiah AK, Zhang YD (2020) Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. *Neural Comput Appl* 32:665–680. <https://doi.org/10.1007/s00521-018-3924-0>
28. Zhang L, Ren Z (2020) Comparison of CT and MRI images for the prediction of soft-tissue sarcoma grading and lung metastasis via a convolutional neural networks model. *Clinical Radiol* 75:64–69. <https://doi.org/10.1016/j.crad.2019.08.008>
29. Otoom AF, Abdallah EE, Hammad M (2015) Breast cancer classification: comparative performance analysis of image shape-based features and microarray gene expression data. *Int J Bio-Sci Bio-Technol* 7:37–46. <https://doi.org/10.14257/ijbsbt.2015.7.2.04>
30. Zhang B, Liu S, Shin YC (2019) In-process monitoring of porosity during laser additive manufacturing process. *Additive Manufacturing* 28:497–505. <https://doi.org/10.1016/j.addma.2019.05.030>
31. Toğaçar M, Ergen B, Cömert Z (2020) COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Comput Biol Med* 121:103805. <https://doi.org/10.1016/j.combiomed.2020.103805>
32. Onan A (2015) A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer. *Expert Syst Appl* 42:6844–6852. <https://doi.org/10.1016/j.eswa.2015.05.006>
33. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM (2003) The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 56:1129–1135. [https://doi.org/10.1016/S0895-4356\(03\)00177-X](https://doi.org/10.1016/S0895-4356(03)00177-X)
34. Chicco D, Jurman G (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21:6. <https://doi.org/10.1186/s12864-019-6413-7>
35. Li MD, Chang K, Bearce B, Chang CY, Huang AJ, Campbell JP, Brown JM, Singh P, Hoebel KV, Erdoğan D, Ioannidis S, Palmer WE, Chiang MF, Kalpathy-Cramer J (2020) Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging. *NPJ Digital Med* 3:1–9. <https://doi.org/10.1038/s41746-020-0255-1>
36. Rajaraman S, Antani S (2020) Weakly labeled data augmentation for deep learning: a study on COVID-19 detection in chest X-rays. *Diagnostics* 10:358. <https://doi.org/10.3390/diagnostics10060358>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Satyavratn Govindarajan is currently a PhD scholar in Biomedical Engineering Group, Department of Applied Mechanics, Indian Institute of Technology Madras, India.

His research interests include Medical image processing, Image informatics, Machine learning, Deep Learning, Control and instrumentation.