

Automatika

Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/taut20>

Pneumonia detection in chest X-ray images using compound scaled deep learning model

Mohammad Farukh Hashmi, Satyarth Katiyar, Abdul Wahab Hashmi & Avinash G. Keskar

To cite this article: Mohammad Farukh Hashmi, Satyarth Katiyar, Abdul Wahab Hashmi & Avinash G. Keskar (2021) Pneumonia detection in chest X-ray images using compound scaled deep learning model, *Automatika*, 62:3-4, 397-406, DOI: [10.1080/00051144.2021.1973297](https://doi.org/10.1080/00051144.2021.1973297)

To link to this article: <https://doi.org/10.1080/00051144.2021.1973297>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 16 Sep 2021.



Submit your article to this journal [↗](#)



Article views: 601



View related articles [↗](#)



View Crossmark data [↗](#)



Pneumonia detection in chest X-ray images using compound scaled deep learning model

Mohammad Farukh Hashmi^a, Satyarth Katiyar^b, Abdul Wahab Hashmi^c and Avinash G. Keskar^d

^aDepartment of Electronics and Communication Engineering, National Institute of Technology Warangal, India; ^bDepartment of Electronics and Communication Engineering, Harcourt Butler Technical University Kanpur, India; ^cDepartment of Mechanical Engineering, Malaviya National Institute of Technology Jaipur, India; ^dDepartment of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology Nagpur, India

ABSTRACT

Pneumonia is the leading cause of death worldwide for children under 5 years of age. For pneumonia diagnosis, chest X-rays are examined by trained radiologists. However, this process is tedious and time-consuming. Biomedical image diagnosis techniques show great potential in medical image examination. A model for the identification of pneumonia, trained on chest X-ray images, has been proposed in this paper. The compound scaled ResNet50, which is the upscaled version of ResNet50, has been used in this paper. ResNet50 is a multilayer layer convolution neural network having residual blocks. As it was very difficult to obtain a sufficiently large dataset for detection tasks, data augmentation techniques were used to increase the training dataset. Transfer learning is also used while training the models. The proposed model could help in detecting the disease and can assist the radiologists in their clinical decision-making process. The model was evaluated and statistically validated to overfitting and generalization errors. Different scores, such as testing accuracy, F1, recall, precision and AUC score, were computed to check the efficacy of the proposed model. The proposed model attained a test accuracy of 98.14% and an AUC score of 99.71 on the test data from the Guangzhou Women and Children's Medical Center pneumonia dataset.

ARTICLE HISTORY

Received 12 February 2020
Accepted 23 August 2021

KEYWORDS

Pneumonia; chest X-ray images; Convolution Neural Network (CNN); ResNet50; transfer learning

1. Introduction

Pneumonia is an inflammatory disease of the lungs and is mainly caused due to pathogens like bacteria and viruses. It is one of the primary causes of death in countries such as the United States [1] and India [2]. In older patients, it goes unnoticed and untreated until it has reached a life-threatening point. Pneumonia is the primary cause of death in children below the age of 5 [3]. According to WHO

Every year, it kills an estimated 1.4 million children under the age of five years, accounting for 18% of all deaths of children under five years old worldwide. Pneumonia affects children and families everywhere but is most prevalent in South Asia and sub-Saharan Africa. Children can be protected from pneumonia, it can be prevented with simple interventions, and treated with low-cost, low-tech medication and care [2].

It is a very common disease all across the globe. Due to poverty, people refrain from having access to trained radiologists. As far as diseases like pneumonia are concerned, the level of accuracy in the diagnosis should be good enough to assure proper treatment of this fatal disease [4]. Hence, to reduce the mortality of pneumonia, there is a need for research in the field

of computer-aided diagnosis. There are many tests for pneumonia diagnosis, such as the chest ultrasound, chest MRI, chest X-ray, computed tomography of the lungs and needle biopsy of the lung [5]. X-rays are the most widely available diagnostics imaging technique [6]. The examination of chest X-rays is a difficult task for radiotherapists. Sometimes the appearance of pneumonia in the X-ray of the patient is very unclear. This leads to difficulty in the prediction of the disease as it is difficult to identify the features that describe the presence of the disease. This is the primary reason behind the misclassification of the X-ray images in the dataset. Several CAD systems have already been demonstrated to aid the medical area primarily in the detection of lung nodules, breast cancer, etc. Deep learning remains the most promising and widely used Machine Learning (ML) technique for radiology in particular and disease detection in general. Deep learning techniques have already proved their effectiveness and their disease prediction accuracy matches that of an average radiologist [7]. Presently, deep learning-based methods cannot be used as a replacement for trained clinicians in medical examination. Hence, computer-aided diagnosis techniques based on deep learning can be used as a supplement in the clinical decision-making process.

The rest of the paper is structured as follows: Section 2 contains the discussion of the related works. Methods used in this paper are discussed in Section 3, while the dataset used has been introduced in Section 3. The proposed methodology and the results are discussed in Sections 4 and 5, respectively. Finally, Section 6 has the conclusion of the paper.

2. Related work

Several biomedical image detection techniques have already been proposed by different authors. Authors in [8] discussed the identification of pneumonia. Razaak et al. addressed the challenges of methods for medical imaging [9]. Many authors proposed various methods for the detection of numerous diseases [10]. For example, Andre in his paper presented a deep CNN-based on Inception v3 architecture for skin cancer classification [11], Milletari also worked on a method for detecting the prostate, using CNN, in MRI volumes [12], Grewal used deep learning for detecting the brain haemorrhage in computed tomography scans [13], Varun worked on a method for the detection of diabetic retinopathy [14]. Lakhani proposed an algorithm based on conventional neural networks (CNNs) for the automated classification of pulmonary tuberculosis [15]. Bar also published a paper in which he used deep neural networks (DNNs) for chest pathology detection [16]. CNN is a much better advancement over DNN as it can very easily operate on 2-D and 3-D images and also extract the features responsible to classify the disease. This is possible in CNN because the max-pooling layer is very efficient and they are also attached to some weights. CNNs also deal with the grave situation of diminishing gradient as they involve gradient-based learning, while it is being trained. CNNs work on the principle of the features extracted by the various layers to identify the various chest X-rays and classify them accurately. Some algorithms were devised earlier in the papers of Shin for data mining [17,18] and the paper of Boussaid also proposed that labels are predicted by the extraction of features and application of segmentation techniques from radiology images of chest X-ray [19]. Methods regarding the diagnosis of disease using chest X-ray were also worked upon earlier by the authors such as Avni, Melendez and Jaeger who performed various examination techniques [20–22]. After the examination is performed classification algorithm on a chest X-ray view is performed which was discussed by Xue [23]. Roth et al. [24] proposed an algorithm for lymph node detection in computed tomography scans using 2.5D resampling approach and a deep CNN. The lung area is very well segregated from the whole chest X-ray where the network has to perform various operations looking for the features for the identification of the disease. Rajaraman et al. [25] presented a visualization strategy for localizing the region of interests in the chest

X-rays. They used a customized VGG 16 model in their work. Togacar et al. [26] combined deep features from several deep models to form an efficient deep feature set and then used algorithms, such as k-nearest neighbours, decision tree, etc. for classification of the deep feature set. Then they used algorithms such as k-nearest neighbours, linear discriminant analysis, for pneumonia detection. Lakhani et al. [15] used GoogLeNet and AlexNet CNNs and reported an area under the curve (AUC) of 0.94–0.95. Ayan et al. [27] and Rahman et al. [28] used CNNs such as Vgg16, Xception, DenseNet201 and SqueezeNet network for pneumonia classification. Vikash et al. [29] proposed a model which was based on an ensemble architecture that combined outputs from several pre-trained models for pneumonia diagnosis. Jaiswal et al. [30] used a deep learning model based on mask-RCNN, which fuses local and global features for pixel-wise segmentation for pneumonia classification. Pezeshk et al. [31] proposed the use of 3-D CNNs for the fast screening of pulmonary nodule detection in chest CT. Our previous work [32] used an ensemble-like model that integrated five deep learning models (ResNet18, Xception, InceptionV3, DenseNet121 and MobileNetV3).

The main contribution of this paper is that a compound scaled version of ResNet50 for pneumonia detection, which required fewer computational resources compared to our previous work [32], is proposed. First, the ResNet50 architecture was used for binary pneumonia classification. Then the ResNet50 architecture was scaled up using the compound scaling method introduced by Tan et al. [33] and was used for pneumonia detection. The comparative analysis of both the architectures for pneumonia detection was done. Towards the end, the class activation maps were plotted to explain the results of the deep learning models.

3. Methods

3.1. Convolutional neural network

CNNs are arguably the most established deep learning architecture. LeCun et al. [34] first proposed the use of CNN for the recognition of handwritten Zip Code. They come in the category of feed-forward networks. The primary advantage of CNN is that it could consider the locality of features. The central part of the CNN is the convolutional layer that gives the network its name. The convolutional layer extracts the features from an input image. The pooling layer reduces the dimensionality of each input map but retains important information. The network architecture consists of several convolutional and pooling layers in series, as shown in Figure 1. At the end of the network, the softmax layer is used for image classification. The main drawback of deeper CNNs is of vanishing gradients.

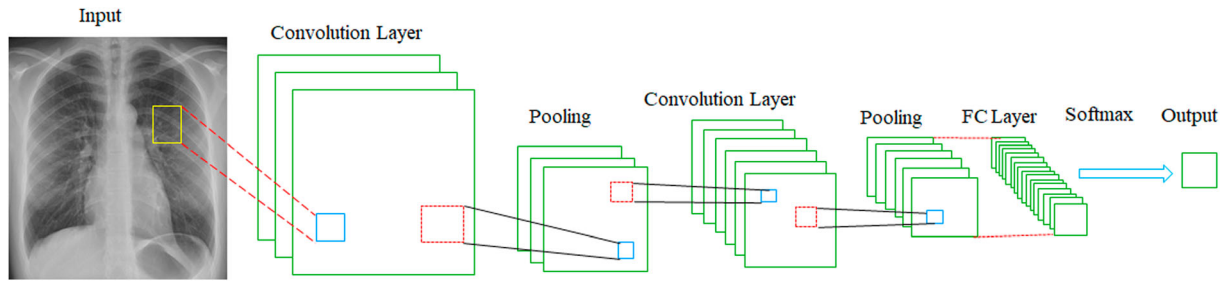


Figure 1. Architecture of the CNN.

3.2. Residual networks

Kaiming He et al. of Microsoft Research Institute [35] proposed the Residual Networks to solve the problem of vanishing gradients. The central idea of the residual block was to take into account both the incoming eigenvalues computed by the network layers and the original values while computing the result. This residual block can be seen in Figure 2.

$F(x)+x$ can be executed by feed-forward neural networks with “bypass connections”. One or more intermediate layers can be skipped using the bypass connections. The outputs of both the bypass connections and of the stacked layers are summed together. The training of the network can be done using SGD with backpropagation.

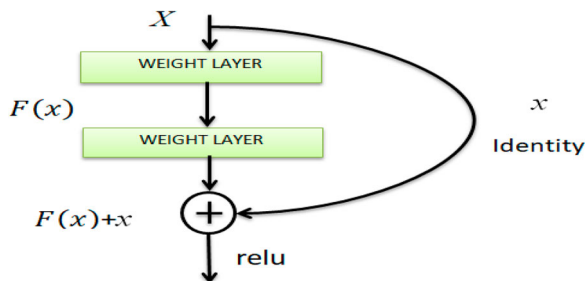


Figure 2. Residual block [35].

3.3. Transfer learning

CNNs generally require larger datasets to train on. When CNNs are trained on smaller datasets they generalize poorly. Transfer learning can be used in such cases. Figure 3 shows the process of transfer learning where the information gained by the model, while solving a particular problem is used for solving another problem.

In medical image processing-related problems, sufficient dataset is not available. Therefore, training a model without proper initialization of weights is expensive in terms of the computational resources involved and the achieved results are also poor. Transfer learning was used with both ResNet50 and the compound scaled version of ResNet50. The initialization of weights using transfer learning is effective as the model is already trained to detect generic features like edges in an image.

4. Materials

4.1. Dataset

There were a total of 5836 images in the original dataset from the Guangzhou Women and Children’s Medical Center [36]. It included the images of both healthy and pneumonia cases. There were 1583 healthy chest X-ray images and 4273 pneumonia-infected chest X-ray

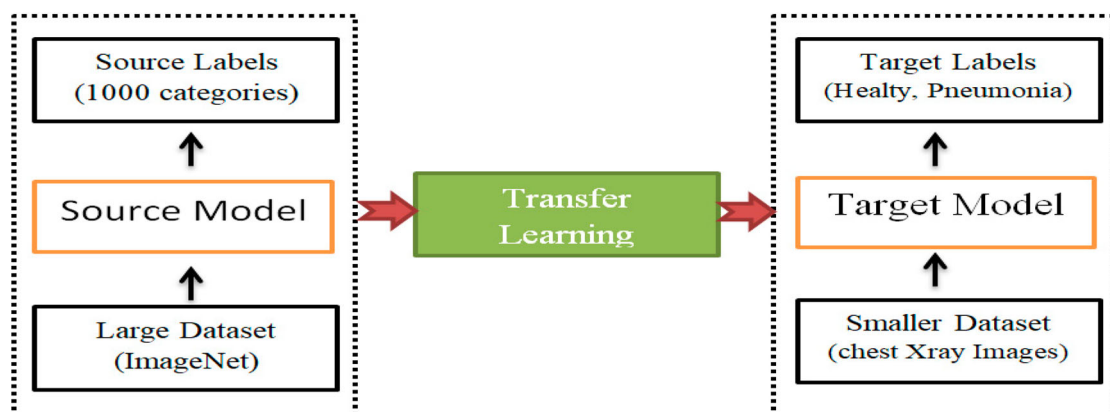


Figure 3. The mechanism of transfer learning.

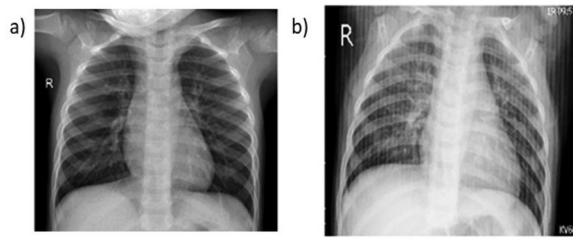


Figure 4. Chest X-ray of a healthy person (a) and a person suffering from pneumonia (b).

images. The entire dataset was rearranged into training and test set with 700 images in the test set and 5136 images in the training set. Two sample images of the dataset have been shown in Figure 4.

5. Proposed methodology

An optimal algorithm for pneumonia detection from Chest X-rays is proposed in this paper. Data augmentation techniques were deployed to increase the size

of the limited dataset. Pre-trained ResNet50 architecture is fine-tuned for the pneumonia classification task. Then, the scaling up of ResNet50 architecture is done using compound scaling, as introduced in [33]. Figure 5 shows the block diagram of the proposed methodology. Various techniques used in the proposed methodology are explained in the subsequent sections.

5.1. Pre-processing and data augmentation

First, the input images were resized to 224*224 and were used to train the model. Effective training of the neural net requires a large dataset. When the deep networks are trained on a smaller dataset, then the networks cannot generalize and hence poor testing accuracy. Data augmentation is one of the solutions to this problem and utilizes the existing dataset in an efficient manner and expands it. In the training dataset used in this paper, there were a total of 1283 healthy chest X-ray images and 3873 pneumonia-infected chest X-ray case images. The dataset had enough images

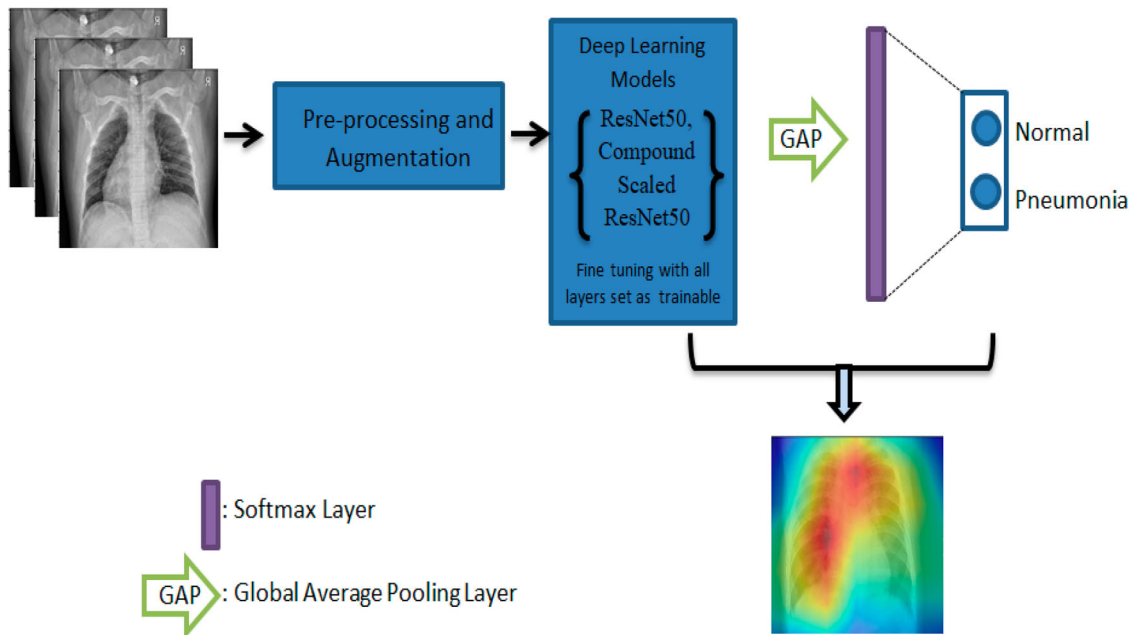


Figure 5. Block diagram representation of the proposed methodology.

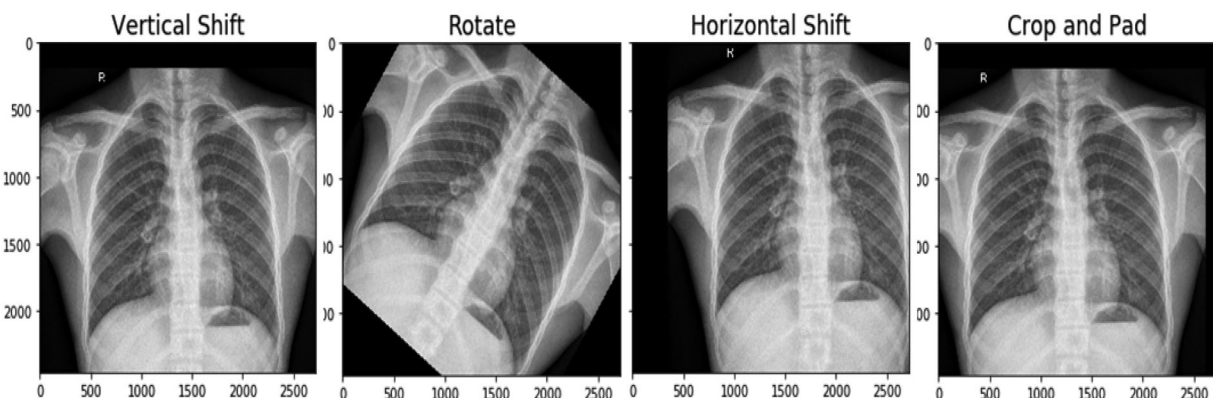


Figure 6. Resultant images after using augmentation techniques.

Table 1. Image augmentation settings.

Technique	Setting
Crop and Pad	0.25
Horizontal Shift	0.15
Rotation	35
Vertical Shift	0.2

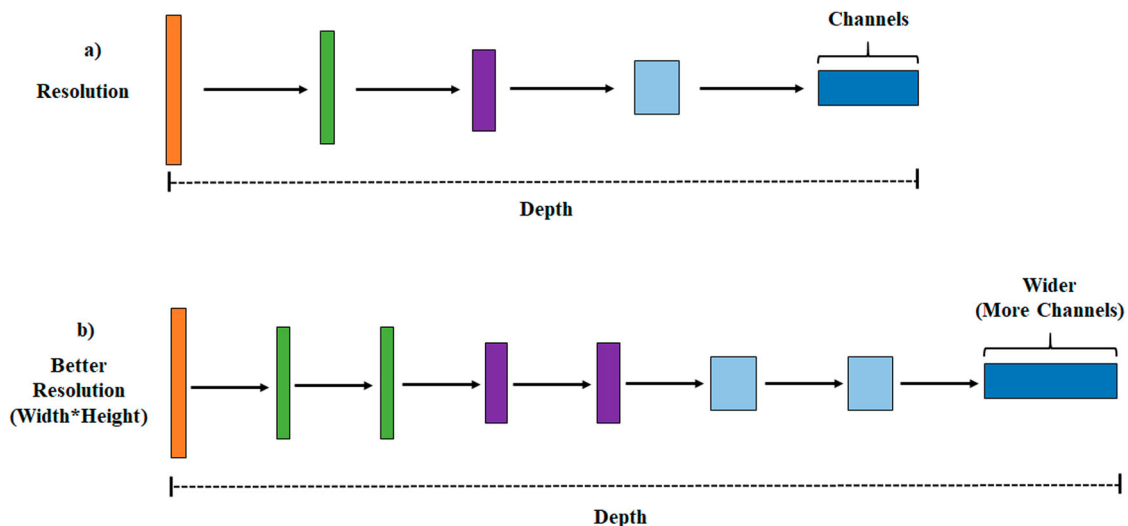
of the pneumonia case and hence only the images of the normal case were augmented twice. The settings used to create augmented images (Figure 6) are shown in Table 1. After augmentation, there were 3849 normal images and 3873 pneumonia images. The test set images were not augmented.

5.2. Compound scaling

Scaling of CNNs can be done across three dimensions: depth, width and resolution. The depth of a CNN is equivalent to the number of layers it has, width refers to the number of channels in the convolution layer and resolution refers to the resolution of the image passed to a CNN. Scaling can be used: to achieve greater performance for a particular task and to create more efficient models. Scaling can be done on width, height and resolution. Figure 7 explains what scaling means across different dimensions.

Scaling up any dimension of the network (width, depth or resolution) should improve accuracy, but the accuracy gain diminishes for bigger models. While scaling CNNs, it is critical to balance all dimensions of a network (width, depth and resolution) for getting improved accuracy and efficiency. Tan et al. [33] proposed compound scaling technique in their network: EfficientNet. A CNN can be represented mathematically as Equation (1):

$$N(d, w, r) = \bigodot_{i=1 \dots s} \hat{F}_i^{d, \hat{L}_i} (X_{(r, \hat{H}_i, r, \hat{W}_i, w, \hat{C}_i)}) \quad (1)$$

**Figure 7.** Compound scaling used in the paper. (a) Base model; (b) compound up scaled model.

In this equation, N is the network, i is the stage number, F_i denotes the convolution operation for the i -th stage, and L_i represents the number of times F_i is repeated in stage i . H_i , W_i and C_i simply denote the input tensor shape for stage i . Hence, N ; $\hat{F}_i \hat{F}_i$, $\hat{L}_i \hat{L}_i$, $\hat{H}_i \hat{H}_i$, $\hat{W}_i \hat{W}_i$, $\hat{C}_i \hat{C}_i$ are the pre-defined parameters in the baseline network. The coefficients for scaling network width, depth, and resolution are w , d and r , respectively. The compound scaling technique proposed Tan et al. [33], consisted of a compound coefficient Φ that uniformly scales the networks width, depth and resolution.

$$\text{depth : } d = \alpha^\Phi$$

$$\text{width : } w = \beta^\Phi$$

$$\text{resolution : } r = \gamma^\Phi$$

$$\text{Such that, } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \text{ and } \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \quad (2)$$

Φ is a user-specified coefficient which controls the number of resources available and α , β and γ specify how to designate these resources to network depth, width and resolution, respectively. In a CNN, Conv layers are the most computationally expensive part of the network. Also, FLOPS of a convolution operation is almost proportional to $d, w \triangleright 2, r \triangleright 2$. Scaling the network using Equation (2) will increase the total FLOPS by $(\alpha \cdot \beta^2 \cdot \gamma^2)^\Phi$. The constraint, $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$, was applied so that the total FLOPS do not exceed 2^Φ . Here, α , β , and γ can be determined using grid search and the compound coefficient Φ can be increased to get larger but more accurate models. In this work, ResNet 50 model was scaled up with $d = 1.4$, $w = 1.2$ and $r = 1.3$.

5.3. Fine-tuning the models

ResNet50, which was pre-trained on ImageNet dataset [37] is used as the base architecture. Table 2 and Figure 8

Table 2. ResNet50 Architecture.

Layer name	Convolutional layers	Output size
Conv1	7*7, 64 Stride 2	112*112
Conv2_x	3*3 Max Pool, Stride 2	56*56
	1 * 1, 64 1 * 1, 64	
	3 * 3, 64 3 * 3 64	
	1 * 1, 256 1 * 1, 256 * 3	
Conv3_x	1 * 1, 128 1 * 1, 128	28*28
	3 * 3, 128 1 * 1, 128	
	1 * 1, 512 1 * 1, 512 * 4	
Conv4_x	1 * 1, 256 1 * 1, 256	14*14
	3 * 3, 256 3 * 3, 256	
	1 * 1, 1024 1 * 1, 1024 * 6	
Conv5_x	$\begin{bmatrix} 1 * 1, & 512 \\ 3 * 3, & 512 \\ 1 * 1, & 2048 \end{bmatrix} \begin{bmatrix} 1 * 1, & 512 \\ 3 * 3, & 512 \\ 1 * 1, & 2048 \end{bmatrix} \cdot * 3$	7*7
Global Average Pool, 2-d fc, Softmax		

show the framework of the network and its details. The network has five stages. The first stage only has one convolutional layer. The second, third, fourth and fifth stages have 9, 12, 18 and 9 convolutional layers, respectively. The kernel size of each convolutional layer is 3*3. After adding a dense layer with two output classes to this imported architecture, every layer of the network was unfrozen. Stochastic Gradient Descent (SGD) is used as the optimizer and the model is trained for 50 epochs. Different learning rates were used, while training the model and different results were obtained. The most optimum performance was achieved while using 0.001 as the learning rate. Then the scaled version of ResNet 50 was created, as discussed in Section 3.2.

5.4. Class activation maps

One of the problems with DNNs is that they are black-box models and it is difficult to explain their final

predictions. Class activation maps (CAMs) are one of the solutions for this problem. The interpretation of the results of the model is necessary for clinical decision-making. This can also help in understanding the points of failure of the model. There are several ways to generate CAMs. In this paper, first, the global average pooling (GAP) layer was attached to the model. After GAP, the softmax layer was attached, which gives the final prediction. The network can be seen in Figure 5. CAMs were generated for the ResNet50 model and the compound scaled ResNet 50 model and the results are presented in Section 5.

6. Experiment results and discussion

The details of the experiments, conducted to test the proposed architecture, are presented in this section. Keras open-source framework with TensorFlow as the backend has been used to implement the deep learning networks. Computation was done on a system having 16 GB RAM and NVIDIA Quadro 4000 graphic card with 2 GB GDDR5 GPU Memory and Intel i7, 7th generation processor.

6.1. Testing accuracy and testing loss

Each experiment was conducted several times to examine the proposed model’s performance. Most optimum results were obtained with 0.001 as the learning rate with SGD as the optimizer. The testing accuracy achieved by ResNet50 architecture was 97.85 and the testing loss was 0.070. With 0.1 as the learning rate, the model was unable to learn anything and testing and training accuracy was stuck at 50%. Table 3 shows the

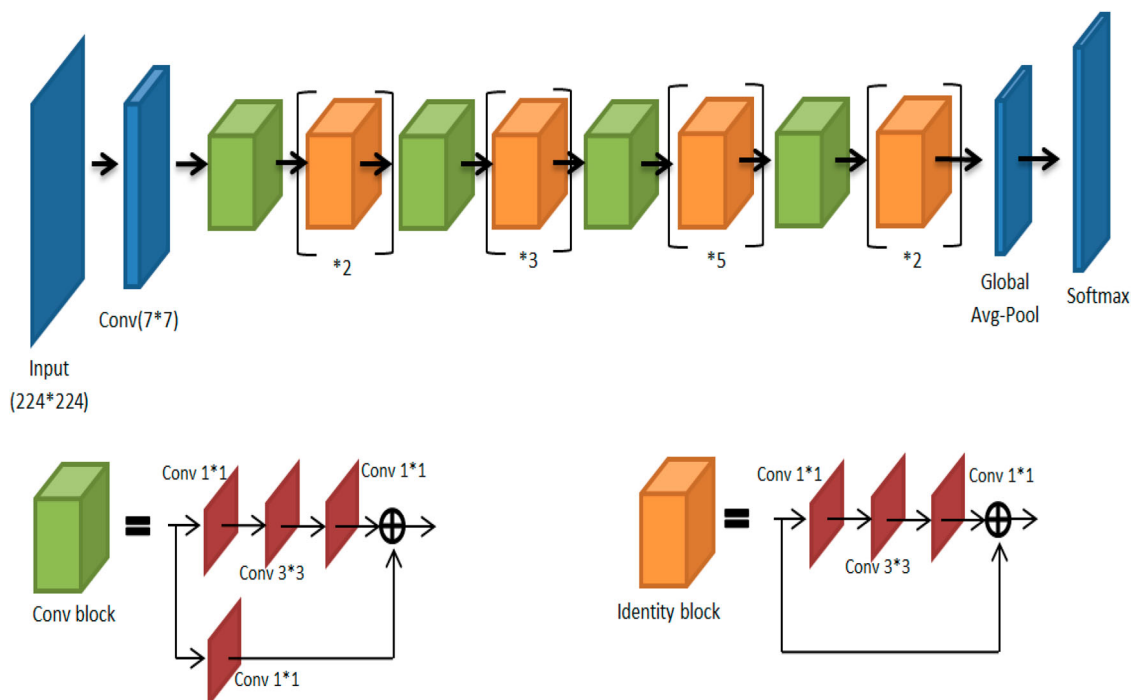


Figure 8. ResNet50 architecture [35] Mebbe a new diagram.

Table 3. Maximum testing accuracy achieved by ResNet50 with different learning rate with SGD as the optimizer.

Learning rate	Testing accuracy achieved
0.1	Nil (accuracy stuck at 50%)
0.01	88.19%
0.001	97.85%
0.0001	95.14%

testing accuracy achieved by ResNet50 with different learning rates.

Finally, with the compound scaled ResNet50 architecture, and with the same hyperparameters, the testing accuracy increased to 98.14 and the testing loss decreased to 0.061. Training accuracy and loss graphs are shown in Figure 9. The models were trained for 50 epochs in every experiment. The average inference time was 0.042 s for both the models.

6.2. Performance analysis

Accuracy, recall and precision score for the two models, ResNet50 and the compound scaled ResNet50 were calculated. Accuracy can be defined as the ratio of the right predictions to the total number of predictions (Equation (3)). Precision is how precise the model is

in predicting the true labels (Equation (4)). Recall is the ratio of the true positives to the summation of true positives and false negatives (Equation (5)).

$$\text{Accuracy} = (TP + TN)/(TP + TF + FP + FN) \quad (3)$$

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

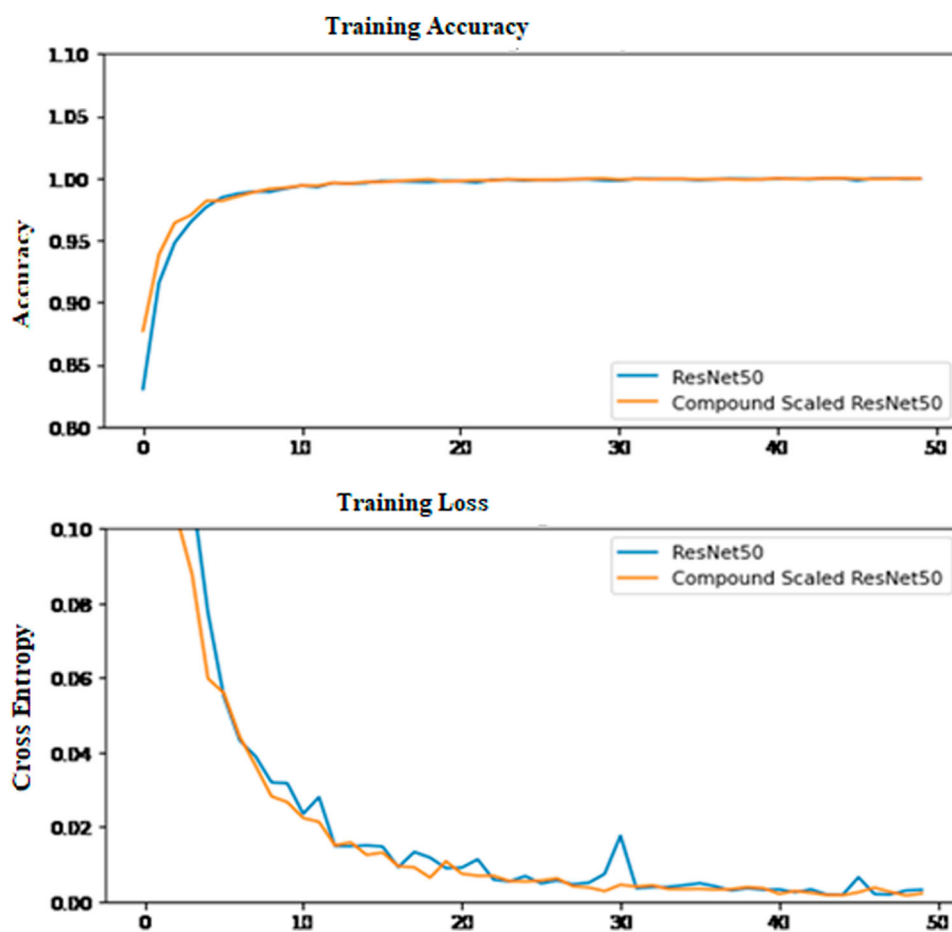
$$\text{Recall} = TP/(TP + FN) \quad (5)$$

To calculate these scores, the confusion matrix for both ResNet50 and compound scaled ResNet50 was obtained (Figures 10 and 11, respectively). The confusion matrix can help in obtaining false positives, true negatives, true positives and false negatives and further these can help in analysing the performance of the model and calculating the above-discussed scores.

As recall is increased, the precision decreases and vice versa. F1 score can be calculated to find the optimal blend of recall and precision.

$$F1 = (2 * \text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall}) \quad (6)$$

The F1 score of the ResNet 50 model was 99.50 and the compound scaled ResNet50 model was 99.71. Furthermore, the AUC curves were plotted and can be seen in Figure 12. Compound scaled ResNet50 architecture achieved a higher AUC score. All the results

**Figure 9.** Training accuracy and training loss curves with respect to number of epochs (50 Epochs).

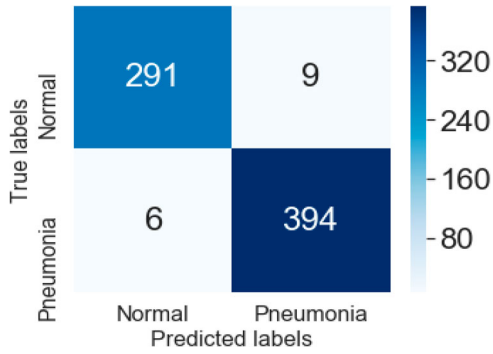


Figure 10. Confusion matrix for ResNet50 over the testing dataset.

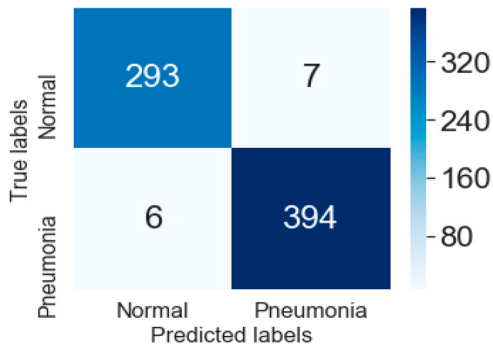


Figure 11. Confusion matrix for the proposed compound scaled ResNet50 architecture over the testing dataset.

have been summarized in Table 4. The achieved results also show that the generic image features, such as the edges and the curves, learned by the deep neural networks while getting trained on the ImageNet dataset, worked as a good initialization of the weights. Despite the domain difference, transfer learning was effective. Hence, initialization of weights by models trained on any classification dataset is better than random initialization of weights. It can also be seen that the compound

Table 4. Accuracy, precision and recall of the proposed methodology.

Architecture	Precision	Recall	F1 Score	Accuracy	AUC Score
ResNet50	97.76	98.5	98.13	97.85	99.50
Compound Scaled ResNet 50	98.25	98.14	98.3	98.14	99.71

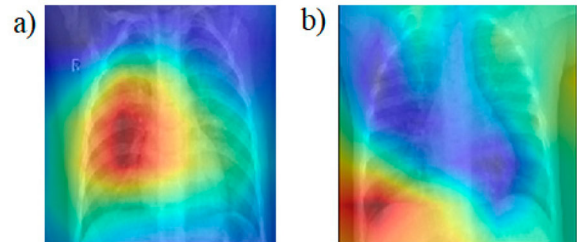


Figure 13. Class activation map for (a) Pneumonia-infected Chest X-ray and (b) healthy chest X-ray.

scaling used in the paper was effective in scaling up the model and increasing the overall performance of the model.

Furthermore, the CAM was plotted to explain the predictions of the deep model. In the Figure 13, it can be seen that the proposed model was able to differentiate between a normal and pneumonia-infected chest X-ray. The deep learning model identified abnormal lungs to detect pneumonia in the chest X-ray. The image parts towards which the model focussed while giving the output were highlighted in the CAM.

6.3. Comparative analysis with respect to various existing methods

Furthermore, a comparative study is presented in Table 5. The accuracy of various existing methods and the proposed methodology is compared. Rajaraman

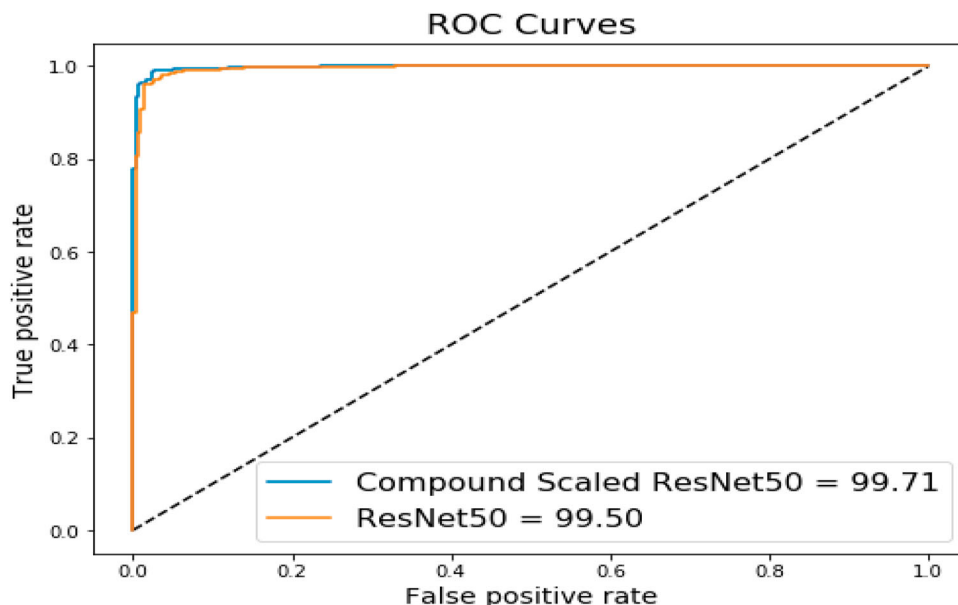


Figure 12. ROC AUC curve for ResNet50 and the proposed compound scaled ResNet50.

Table 5. Comparison of the proposed methodology with respect to various existing methods.

Model	No. of images	Accuracy	Precision	Recall	AUC
Rajaraman et al. [25]	5856	96.2	97.0	99.5	99.0
Rahman et al. [28]	5247	98.0	97.0	99.0	98.0
Ayan et al. [27]	5856	84.5	91.3	89.1	87.0
M.Togacar et al. [26]	5849	96.84	96.88	96.83	96.80
Vikash et al. [29]	5232	96.39	93.28	99.6	99.34
Hashmi et al. [32]	5856	98.26	99.00	98.43	99.76
Proposed Methodology	5856	98.3	98.25	98.14	99.71

et al. [25] used customized CNNs and were able to achieve a testing accuracy of 96.2% and an AUC score of 99.0. M. Togacar et al. [26] reduced the number of deep features and reported the testing accuracy of 96.84%. Ayan et al. [27] and Rahman et al. [28] used CNNs such as Vgg16, Xception, DenseNet201 and SqueezeNet and achieved a testing accuracy of 84.5% and 98.0%, respectively. Vikash et al. [29] proposed an ensemble model to combine predictions from multiple deep learning models and reported a testing accuracy of 96.36%. The proposed methodology achieved a testing accuracy of 98.14% with an AUC of 99.71. Our previous work [32] achieved a higher testing accuracy (98.43%) and a higher AUC score (99.71%), but it required more computational resources as five deep learning models were integrated. Five models (ResNet18, Xception, InceptionV3, DenseNet121 and MobileNetV3) had to be trained and the predictions from all the five models had to be obtained to give the final output. The average inference time for this approach was 0.042 s, while for the previous approach it was 0.203 s. Hence, even though the present approach achieved lesser accuracy compared to the previous approach, but its application on a real-world device would be easier.

7. Conclusions and future scope

Pneumonia is a life-threatening infectious disease. For patients over 75 years, the mortality rate of pneumonia is 24.8% [38]. In this paper, an algorithm which can further support computer-aided diagnosis of pneumonia has been proposed. The deep residual network, proposed in the paper, has a more complex structure but fewer parameters and higher accuracy. Furthermore, this model was scaled up efficiently using the method of compound scaling. Data augmentation and transfer learning have also been used to tackle the obstacle of the insufficient training dataset. Different scores, such as recall, precision and accuracy, were computed to prove the robustness of the model. The proposed model attained an accuracy of 98.14%, a high AUC score of 99.71 and an F1 score of 98.3. The future works involve developing an algorithm which can localize the parts of the lung affected by pneumonia.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Mohammad Farukh Hashmi  <http://orcid.org/0000-0002-3808-9122>

References

- [1] https://healthcare.utah.edu/the-scope/shows.php?shows=0_r1w4wti7.
- [2] https://www.who.int/maternal_child_adolescent/news_events/news/2011/pneumonia/en/.
- [3] Rudan I, Tomaskovic L, Boschi-Pinto C, et al. Global estimate of the incidence of clinical pneumonia among children under five years of age. *Bull World Health Organ.* 2004;82:85–903.
- [4] Naicker S, Plange-Rhule J, Tutt RC, et al. “Shortage of healthcare workers in developing countries. *Afr Ethn Dis.* 2009;19:60.
- [5] <https://www.radiologyinfo.org/en/info.cfm?pg=pneumonia>.
- [6] Franquet T. Imaging of pneumonia: trends and algorithms. *Eur Respir J.* 2001;18:196–208.
- [7] Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18(8): 500–510. doi:10.1038/s41568-018-0016-5
- [8] Gonçalves-Pereira J, Conceição C, Póvoa P. Community-acquired pneumonia: identification and evaluation of nonresponders. *Ther Adv Infect Dis.* 2013;1(1):5–17.
- [9] Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and the future. 2017.
- [10] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annual review of biomedical engineering.* 2017;19:221–248. doi:10.1146/annurev-bioeng-071516-044442.
- [11] Andre E, Brett K, Roberto A, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115–118.
- [12] Milletari F, Navab N, Ahmadi S. V-net: fully convolutional neural networks for volumetric medical image segmentation. 2016 fourth International Conference on 3D Vision (3DV); Stanford, CA; 2016:565–571. doi:10.1109/3DV.2016.79.
- [13] Grewal M, Srivastava MM, Kumar P, et al. 2017. Radiologist level accuracy using deep learning for haemorrhage detection in CT scans. arXiv preprint arXiv:1710.04934.
- [14] Varun G, Lily P, Marc C, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2017;316(22):2402–2410.
- [15] Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology.* 2017;284(2):574–582.
- [16] Bar Y, Diamant I, Wolf L, et al. Chest pathology detection using deep learning with non-medical training. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI); New York (NY); 2015:294–297. doi:10.1109/ISBI.2015.7163871.

- [17] Shin HC, Lu L, Kim L, et al. Interleaved text/image deep mining on a very large- scale radiology database. Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR); Boston (MA); June 2015.
- [18] Shin HC, Lu L, Kim L, et al. Interleaved text/image deep mining on a large-scale radiology database for automated image interpretation. *J Mach Learn Res.* **2016**;17(40):1–31.
- [19] Boussaid H, Kokkinos I. Fast and exact: ADMM-based discriminative shape segmentation with loopy part models. Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR); Columbus (OH); June 2014.
- [20] Avni U, Greenspan H, Konen E, et al. X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words. *IEEE Trans Med Imaging.* **2011**;30(3):733–746. doi:10.1109/TMI.2010.2095026.
- [21] Melendez J, Van GB, Maduskar P, et al. A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-rays. *IEEE Trans Med Imaging.* **2015**;34(1):179–192.
- [22] Jaeger S, Karargyris A, Candemir S. Automatic tuberculosis screening using chest radio-graphs. *IEEE Trans Med Imaging.* **2014**;33(2):233–245.
- [23] Xue Z, You D, Candemir S, et al. Chest x-ray image view classification. Proceedings of the computer-based medical systems IEEE 28th International symposium; São Paulo, Brazil; June 2015.
- [24] Roth HR, et al. A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. International conference on medical image computing and computer-assisted intervention; Springer, Cham; **2014**.
- [25] Rajaraman S, Candemir S, Kim I, et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Appl Sci.* **2018**;8(10):1715.
- [26] Toğaçar M, et al. A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models. *IRBM.* **2019**;41(4):212–222. doi:10.1016/j.irbm.2019.10.006.
- [27] Ayan E, Ünver HM. Diagnosis of pneumonia from chest X-Ray images using deep learning. 2019 scientific meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT); IEEE; 2019.
- [28] Rahman T, Chowdhury MEH, Khandakar A, et al. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Appl Sci.* **2020**;10(9):3233.
- [29] Chouhan V, Singh SK, Khamparia A, et al. A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl Sci.* **2020**;10(2):559.
- [30] Jaiswal AK, Tiwari P, Kumar S, et al. Identifying pneumonia in chest X-rays: A deep learning approach. *Measurement.* **2019**;145:511–518.
- [31] Pezeshk A, Hamidian S, Petrick N, et al. 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. *IEEE J Biomed Health Inform.* **2019**;23(5):2080–2090.
- [32] Hashmi MF, Katiyar S, Keskar AG, et al. Efficient pneumonia detection in chest xray images using deep transfer learning. *Diagnostics.* **2020**;10(6):417.
- [33] Tan M, Le QV. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*; **2019**.
- [34] LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**;1(4):541–551.
- [35] He K, Zhang X, Ren S, et al. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016:770–778.
- [36] Kermany D, Zhang K, Goldbaum M. (2018). Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. Mendeley Data, v2 doi:10.17632/rscbjbr9sj.2.
- [37] Fei-Fei L. Imagenet: crowdsourcing, benchmarking & other cool things. *CMU VASC Seminar.* 2010;16:18–25.
- [38] Henig O, Kaye KS. Bacterial pneumonia in older adults. *Infectious Disease Clinics.* **2017**;31(4):689–713.