1-1-2020

Two-stage deep learning architecture for pneumonia detection and its diagnosis in chest radiographs

Barath Narayanan
University of Dayton

Venkata Salini Priyamvada Davuluru
University of Dayton

Russell C. Hardie
University of Dayton, rhardie1@udayton.edu

eCommons Citation

Barath Narayanan Narayanan, Venkata Salini Priyamvada Davuluru, Russell C. Hardie, "Two-stage deep learning architecture for pneumonia detection and its diagnosis in chest radiographs," Proc. SPIE 11318, Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications, 113180G (2 March 2020); https://doi.org/10.1117/12.2547635

Two-Stage Deep Learning Architecture for Pneumonia Detection and its Diagnosis in Chest Radiographs

Barath Narayanan Narayanan[a,b*], Venkata Salini Priyamvada Davuluru[a], Russell C. Hardie[a]
[a]Department of Electrical and Computer Engineering, University of Dayton, 300 College Park, Dayton, Ohio, USA 45469
[b]Sensors and Software Systems Division, University of Dayton Research Institute, 300 College Park, Dayton, Ohio, USA 45469

ABSTRACT

Approximately two million pediatric deaths occur every year due to Pneumonia. Detection and diagnosis of Pneumonia plays an important role in reducing these deaths. Chest radiography is one of the most commonly used modalities to detect pneumonia. In this paper, we propose a novel two-stage deep learning architecture to detect pneumonia and classify its type in chest radiographs. This architecture contains one network to classify images as either normal or pneumonic, and another deep learning network to classify the type as either bacterial or viral. In this paper, we study and compare the performance of various stage one networks such as AlexNet, ResNet, VGG16 and Inception-v3 for detection of pneumonia. For these networks, we employ transfer learning to exploit the wealth of information available from prior training. For the second stage, we find that transfer learning with these same networks tends to overfit the data. For this reason we propose a simpler CNN architecture for classification of pneumonic chest radiographs and show that it overcomes the overfitting problem. We further enhance the performance of our system in a novel way by incorporating lung segmentation using a U-Net architecture. We make use of a publicly available dataset comprising 5856 images (1583 – Normal, 4273 – Pneumonic). Among the pneumonia patients, 2780 patients are identified as bacteria type and the rest belongs to virus category. We test our proposed algorithm(s) on a set of 624 images and we achieve an area under the receiver operating characteristic curve of 0.996 for pneumonia detection. We also achieve an accuracy of 97.8% for classification of pneumonic chest radiographs thereby setting a new benchmark for both detection and diagnosis. We believe the proposed two-stage classification of chest radiographs for pneumonia detection and its diagnosis would enhance the workflow of radiologists.

Keywords: Pneumonia Detection, Computer Aided Detection, Computer Aided Diagnosis, Convolutional Neural Networks, Lung Segmentation

1. INTRODUCTION

Computer Aided Detection and Diagnosis (CADD) has been a research area attracting great interest in medical imaging over the past decade. CADD tools would assist the doctors and offer a valuable second opinion to the doctors. Pneumonia affects 7% of the global population and results in 2 million deaths every year[1]. Pneumonia is one of the most fatal diseases among children. Chest Radiographs (CRs) are one of the most commonly utilized imaging modalities by radiologists in order to detect and diagnose pneumonia[1]. In this research, we present a novel two-stage deep learning architecture to detect and diagnose pneumonia on CRs thereby assisting radiologists with the decision-making process.

Several machine learning and deep learning methods have been published in the literature for CADD in CRs. A Computer aided detection system to detect lung nodules in CRs is presented[2]. Lungs in CRs are segmented using active shape model and later the potential nodule candidates are determined using multi-scale weighted convergence index filter[2]. Later, the potential candidates are classified as nodule or non-nodule using 114 hand-crafted features and a Fisher linear discriminant classifier[2]. A novel 'N-Quoit' filter is presented for computer aided detection in CRs[3]. Feature selection based cluster and classification approach is presented in order to detect and identify lung nodules in CRs[4]. Comparative study of various traditional classification approaches for lung cancer detection in CRs is presented[5]. A set of 503 hand-crafted features are computed for lung cancer detection using conventional classification approaches[6]. A U-

Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications,
edited by Po-Hao Chen, Thomas M. Deserno, Proc. of SPIE Vol. 11318, 113180G
 · doi: 10.1117/12.2547635

Net architecture for segmentation of lungs in CRs[7]. CADD tools for pneumonia are presented[8-15]. A set of hand-crafted features are computed in order to detect pneumonia in CRs[8-11]. Gradient-based visualization method to localize the region of interest for pneumonia detection is presented[12]. An attention guided mask algorithm to locate the region of interest for pneumonia detection is presented[13]. However, very few research papers have emphasized on applying latest deep learning architectures for CADD of pneumonia using CRs[14, 15].

Convolutional Neural Networks (CNNs) are effective for pneumonia detection[14, 15] and have provided good performance for various imaging applications which includes text, handwriting, and natural images. CNNs provide state-of-the-art performances for various visualization tasks[16-20]. We adopt the same in this paper. However, we present a two-stage deep learning architecture in order to detect pneumonia and classify its type as either bacterial or viral rather than a single stage architecture. We believe this type of architecture would be more helpful for radiologists and also help the CNNs to tune its weights/parameters accordingly for both detection and diagnosis stages.

For stage one, we present transfer learning based approaches for classification of a CR as either normal or pneumonic. Transfer learning approaches are implemented using well-established networks such as AlexNet[17], ResNet[18], VGG-16[19] and Inception-v3[20]. This type of architecture would help us exploit the wealth of information available from prior training for distinguishing wide range of classes and have proven to be highly effective for certain medical imaging applications[21]. Later, the overall probability of a CR being pneumonic is determined by averaging probabilities provided by all the transfer learning architectures implemented in this paper.

In stage two, we further subcategorize the pneumonic chest radiographs as either bacterial or viral for further diagnosis. We find that the transfer learning approaches tend to overfit the data. This could be attributed to lack of training data. For this reason, we present a computationally efficient and a simpler CNN architecture for classification of pneumonic chest radiographs as bacterial or viral. Based on the markings by radiologists, we realize that shape and size of the lung plays a huge role in determining the bacterial or viral category. Thus, we implement a lung segmentation algorithm using U-Net architecture[7]. In order to train this U-Net architecture, we utilize the true lung segmentation masks provided for a different database by radiologists. We present these results for a publicly available dataset thereby setting a new benchmark. Our proposed algorithm provides an area under the Receiver Operating Characteristic (ROC) curve value of 0.996 for pneumonia detection on a set of 624 test images. We also achieve an overall accuracy of 97.9% for classification of pneumonic chest radiographs as bacterial or viral thereby setting a new benchmark for both detection and diagnosis. We believe this type of two-stage architecture would help in enhancing the workflow of radiologists.

The remainder of the paper is organized as follows. Section 2 describes the database utilized for this research. Section 3 presents the computer aided detection architecture for pneumonia. Section 4 presents the CNN architecture implemented for classifying pneumonic chest radiographs as bacterial or viral. Experimental results obtained using the proposed approaches are presented in Section 5. Finally, conclusions are offered in Section 6.

2. MATERIALS

For this research, we utilize a publicly available dataset[15] containing CRs of children from 1 to 5 years of age. This data is collected from Guangzhou Women and Children's Medical Center in Guangzhou, China. Dataset comprises 5,856 images manually annotated by an expert reader as normal or bacterial or viral. Distribution of the dataset is as provided in Table 1. Figures 1-3 represent certain sample images marked as normal, bacterial or viral by an expert reader.

Table 1. Dataset Distribution

Type	Number of Chest Radiographs
Normal	1583
Pneumonia (Bacterial)	2780
Pneumonia (Viral)	1493

For the pneumonia detection stage, we combine the cases marked by reader as either bacterial or viral into a single pneumonia category thereby converting it into a binary classification scenario as either normal or pneumonic. We utilize the same set of training and testing cases[15]. We further divide training dataset into groups of 90% and 10% for training and validation purposes respectively in order to fine-tune the hyperparameters. Table 2 presents the distribution of

samples belonging to each type of dataset. We utilize the same set of cases for all the architectures presented for this stage.

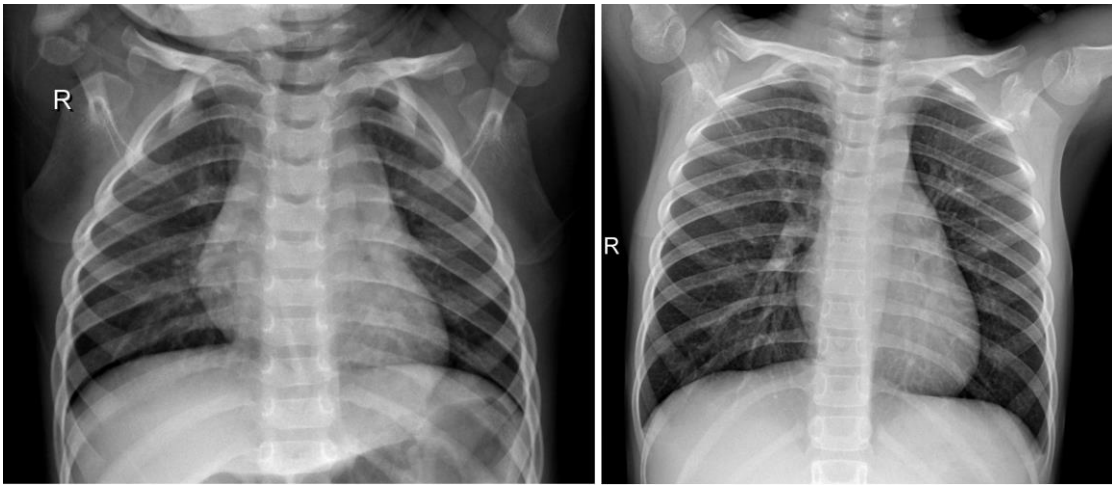

Figure 1. Sample images annotated as ‘Normal’ by Expert Readers.

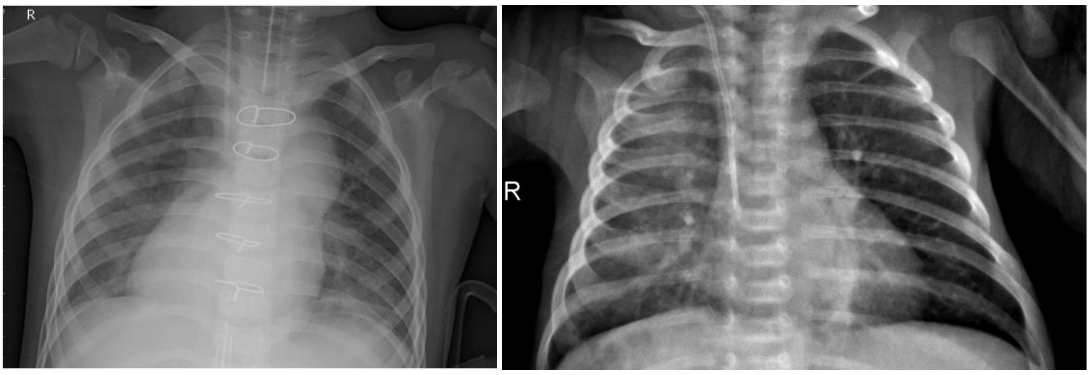

Figure 2. Sample images annotated as ‘Bacterial’ by Expert Readers.

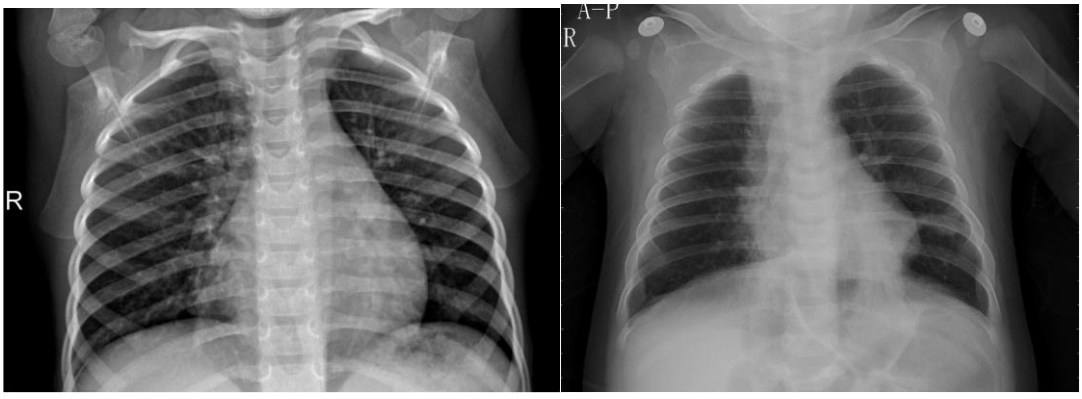

Figure 3. Sample images annotated as ‘Viral’ by Expert Readers.

Table 2. Training, Validation and Testing Dataset for Pneumonia Detection

Type of Dataset	Normal	Pneumonia
Training	1214	3494
Validation	135	389
Testing	234	390

For the diagnosis stage of pneumonia, we utilize the same set of cases[15] for training and testing purposes. We further subdivide training dataset into 10% for validation purposes in order to fine-tune the hyperparameters as implemented in the detection stage. The distribution is presented in Table 3.

Table 3. Training, Validation and Testing Dataset for Pneumonia Diagnosis

Type of Dataset	Bacteria	Virus
Training	2284	1210
Validation	254	135
Testing	242	148

3. PNEUMONIA DETECTION ARCHITECTURE

In this section, we present the transfer learning based CNN architecture adopted for classification of images as either normal or pneumonic. At first, we spatially re-sample all the images to 224 × 224 or 227 × 227 depending on the architecture implemented. Aforementioned, we utilize well-established networks such as AlexNet, ResNet, VGG16 and Inception-V3 for transfer learning purposes. Utilizing such well-established networks helps us retain wealth of information from prior training for classifying different objects. We retain all the weights and layers until the last fully connected layer. We replace the last layers with our own fully connected and softmax layers specific to this application. These approaches have proven to be highly effective especially for image-based classification problems[17-21]. Figure 4 presents the top-level block diagram of the transfer learning approach implemented in this study[21]. Moreover, study of such transfer learning approaches could provide us with insights about the x-ray images and would help us understand the effectiveness of such methodologies for medical imaging-based applications.

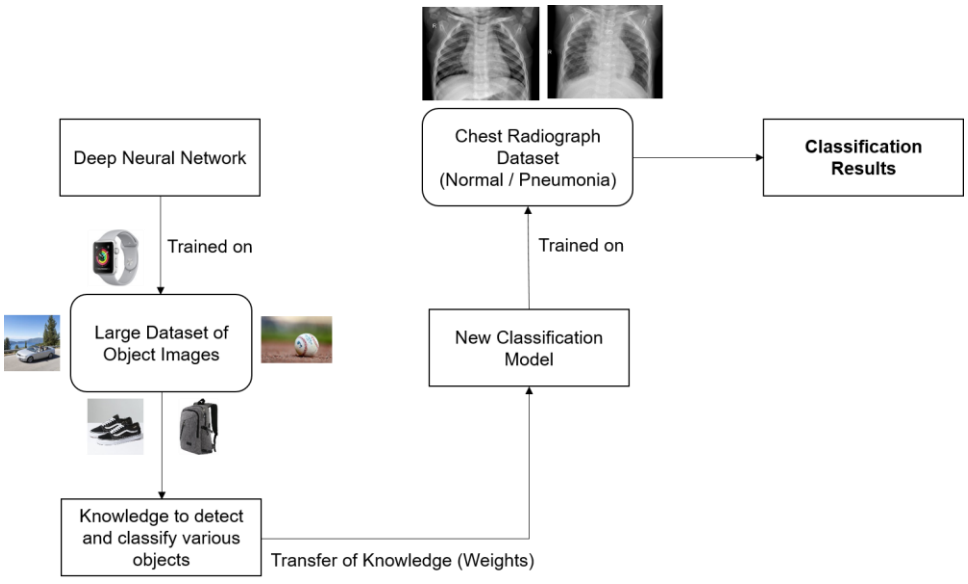

Figure 4. Top Level Block Diagram of Transfer Learning Approach.

4. PNEUMONIA DIAGNOSIS ARCHITECTURE

In this section, we propose and study various methods for classification of pneumonia chest radiographs as either bacteria or virus. We utilize the dataset distribution presented in Table 3 for this study. At first, we study the performance of transfer learning approaches based on the same set of networks utilized for detection stage.

In addition, we propose our own CNN architecture specific to this application and study its performance. Figure 5 presents our proposed CNN architecture. We resample all cases to a size of 256 × 256. Each convolutional layer present in the architecture is comprised of convolution operation (3 × 3), batch normalization, Rectified Linear Unit (ReLU) and a maximum pooling layer of window size 2 × 2 with a stride of 2 as shown in Figure 6. However, the number of convolution filters present in each layer differ and is as shown in Table 4. Hyperparameters for the proposed CNN architecture are determined solely based on the validation dataset. We choose 'adam' optimization technique with a mini batch size of 64 and an initial learning rate of 0.0001.

In addition to the CNN architecture, we also preprocess the images by segmenting the lungs using our previously proposed U-Net architecture[7]. U-Net architecture is solely trained based on the true lung masks provided for Shenzhen Dataset[22, 23] comprising 566 cases. Figure 7 presents the U-Net architecture adopted in this paper for lung segmentaition[7]. Proposed U-Net comprises of 3 stages of encoding and decoding[7]. All the convolution operations and pooling operations are performed using 3 x 3 and 2 x 2 filter. Deconvolution operation and up-pooling operations are performed using bilinear interpolation[24]. Lung segmentation is performed using MATLAB's built-in function unetLayers[25]. Based on the training dataset, we determine that shape of the lung plays a significant role in determining bacterial or viral type of pneumonia. Hence, we believe that segmenting the lungs would help the classifier in distinguishing bacterial and viral type of pneumonia.

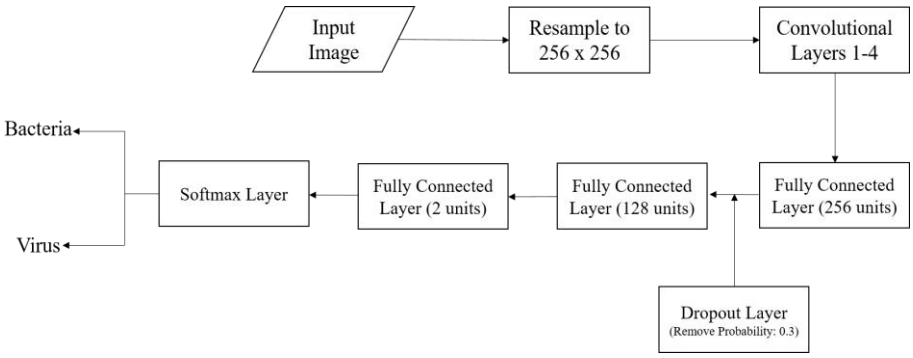

Figure 5. Proposed CNN architecture.

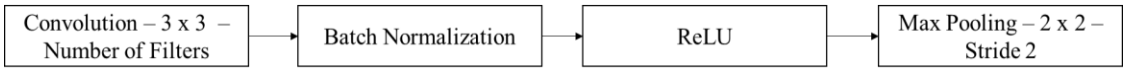

Figure 6. Convolutional Layer Structure.

Table 4. Number of convolution filters present in each convolutional layer.

Convolutional Layer#	Number of Filters
Layer #1	8
Layer #2	16
Layer #3	32
Layer #4	64

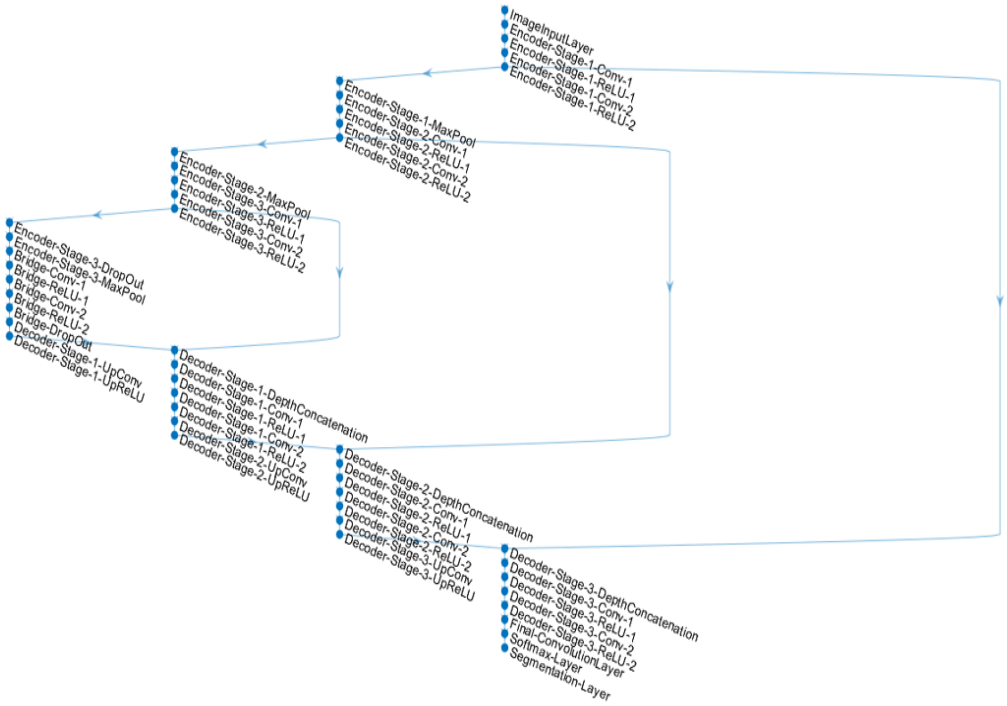

Figure 7. U-Net Architecture for Lung Segmentation in CRs[7]

Figure 8 presents the preprocessing results after the application of lung segmentation algorithm. We study the performance of our proposed approach with and without additional preprocessing. Note that, after preprocessing we solely present the segmented lung as the input image to the CNN architecture.

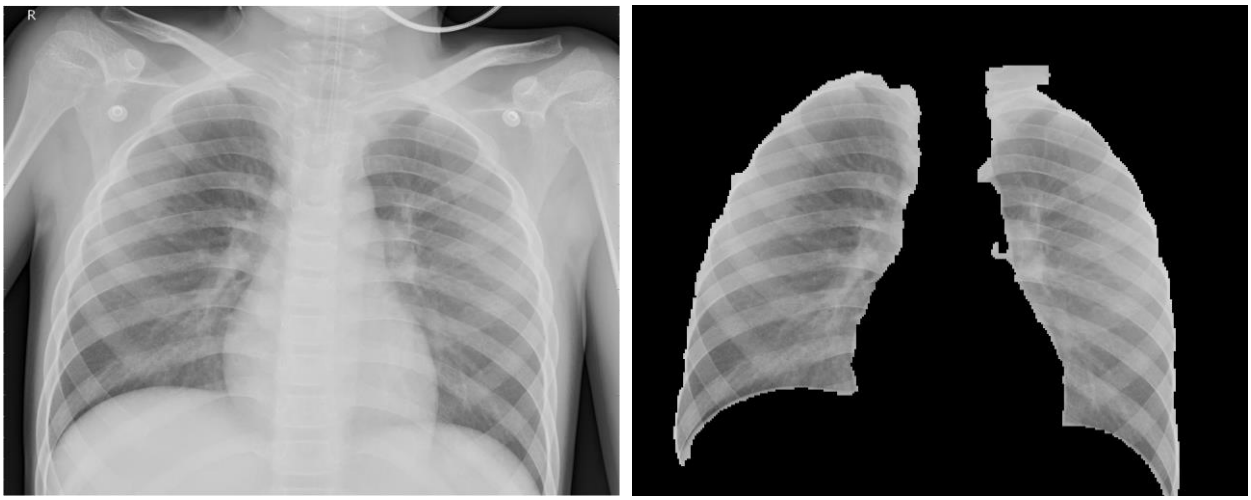

Figure 8. Lung Segmentation Results.

5. EXPERIMENTAL RESULTS

In this section, we present the results obtained for approaches described in Sections 3 and 4. We study the performance of pneumonia detection in terms of ROC curve thereby providing an option to radiologists to choose his/her choice of operating point. As mentioned earlier, we study the performance of the pneumonia detection based on the dataset

distribution provided in Table 2. Table 5 presents the training summary for all the transfer learning based approaches in terms of overall accuracy. Table 5 clearly indicates that transfer learning approaches provide good results for validation dataset and the systems are well trained without any additional preprocessing.

Table 5. Training and Validation Accuracy for Pneumonia Detection.

Network Adopted for Transfer Learning	Training Accuracy	Validation Accuracy
AlexNet	97.2	96.5
VGG16	95.1	94.8
ResNet	98.5	98.4
Inception-v3	**99**	**98.9**

We present the results obtained for pneumonia detection in terms of ROC curve for test cases. Figure 9 presents the results obtained using transfer learning approaches for the test dataset. Results clearly indicate that transfer learning approaches using well established networks are highly effective for classification of normal and pneumonic chest radiographs. In addition, we also average all the posterior probabilities provided by all the approaches and study its results. We term this method as 'Average of All'. We study the results in terms of Area under the ROC curve (AUC) and they are summarized for the test dataset in Table 6.

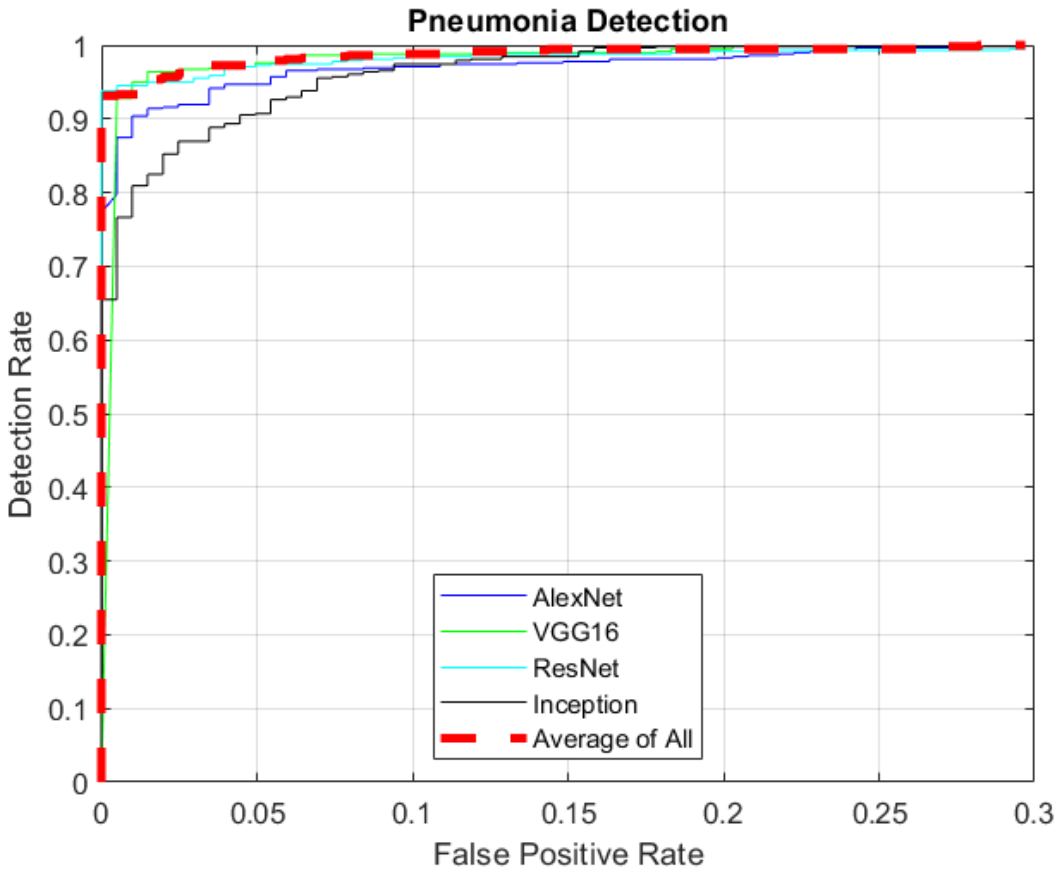

Figure 9. ROC curve obtained for various classification algorithms for detection of Pneumonia.

Table 6. AUC for Detection of Pneumonia on Testing Set.

Network Adopted for Transfer Learning	AUC
Kermany et al.[15]	0.968
AlexNet	0.993
VGG16	0.990
ResNet	0.993
Inception-v3	0.987
Average of All	**0.996**

For diagnosis of pneumonic chest radiographs as either bacterial or viral, Table 7 presents the training summary for all the approaches presented in Section 4. Table 7 clearly indicates that transfer learning based approaches tend to overfit the dataset and our proposed CNN overcame the overfitting issue due to its simplistic architecture. Table 7 also indicates that segmentation of lungs using U-Net architecture increased the validation accuracy by 2.3%.

Table 7. Training and Validation Accuracy for Pneumonia Diagnosis (Bacterial vs. Viral).

Network	Training Accuracy	Validation Accuracy
AlexNet	96	68
VGG16	97	75
ResNet	93	73
Inception-v3	98	72.2
Proposed CNN	**96.7**	**96.0**
Proposed CNN with Lung Segmentation	**98.5**	**98.3**

We measure the performance of our proposed CNN architecture for classification of pneumonic chest radiographs in terms of confusion matrix for our test dataset. Figure 10 presents the confusion matrix for classification of pneumonic chest radiographs without and with lung segmentation using our proposed CNN architecture. Table 8 presents the AUC value for viral detection on test cases for pneumonia diagnosis.

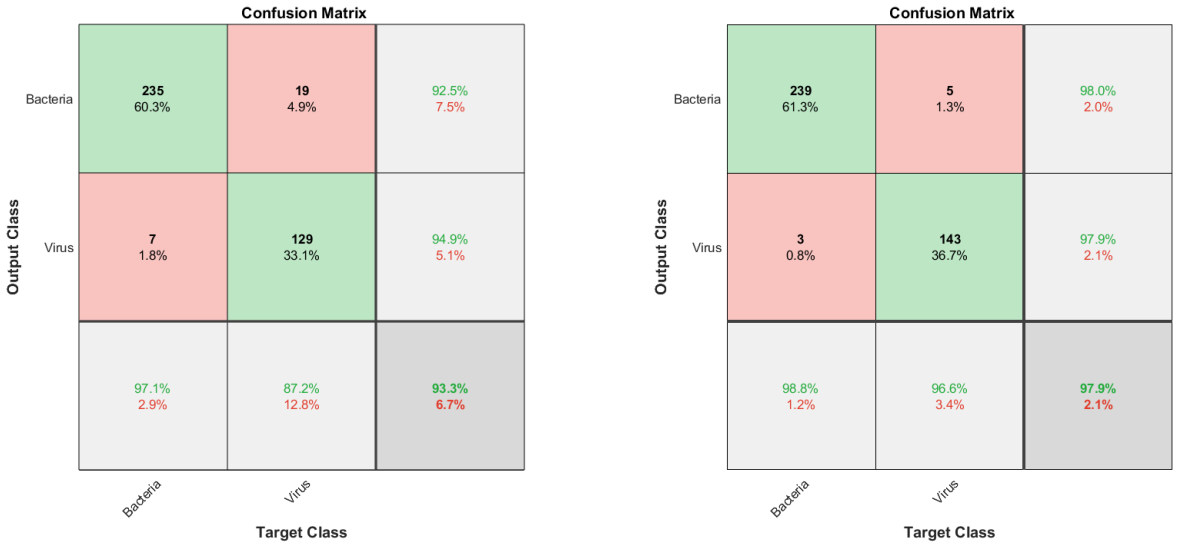

Figure 10. Confusion matrix on test dataset using our proposed CNN architecture for classification of pneumonic chest radiographs (a) without any additional preprocessing, (b) after the application of lung segmentation

Table 8. AUC for Pneumonia Diagnosis (Viral Detection) on Testing Set.

Network	AUC
Kermany et al.[15]	0.94
Proposed CNN	0.96
Proposed CNN with Lung Segmentation	0.99

6. CONCLUSIONS

In this research, we have presented a two-stage deep learning architecture for detection and diagnosis of pneumonia. All the classification algorithms presented in this paper performed relatively well. We observe that the transfer learning approaches helped us achieve an AUC value of 0.996 for pneumonia detection thereby setting a high benchmark. Results indicated that transfer learning approaches using well-established deep learning networks led to overfitting problem for classification of pneumonia images as bacteria and virus. However, we overcame this problem with the help of a simple CNN architecture specific to this application and we achieved an accuracy of 93.3%. We enhanced the performance further to 97.9% by segmenting lungs using our U-Net architecture. Aforementioned, shape and structure of the lung plays an important role in determining the type of pneumonia category. In addition, segmenting lungs also aids in identifying region of interest for the classifier and the CNN architecture is tuned accordingly. We believe this type of two-stage architecture would help in determining parameters/weights particular to that stage. Having separate network architectures for detection and diagnosis would also help in re-training with new labeled data based on the application. This type of two-stage architecture would be highly essential for pneumonia detection and its diagnosis and would enhance the workflow of radiologists.

REFERENCES

[1] Le Roux, D. M., Myer, L., Nicol, M. P., & Zar, H. J., "Incidence and severity of childhood pneumonia in the first year of life in a South African birth cohort: the Drakenstein Child Health Study", The Lancet Global Health, 3(2), e95-e103 (2015).

[2] Hardie, R. C., Rogers, S. K., Wilson, T., & Rogers, A., "Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs", Medical Image Analysis, 12(3), 240-258 (2008).

[3] Okumura, T., "Variable N-Quoit filter applied for automatic detection of lung cancer by X-ray CT", Computer Assisted Radiology and Surgery (CAR 1998), 242-247 (1998).

[4] Narayanan, B. N., Hardie, R. C., Kebede, T. M., & Sprague, M. J., "Optimized feature selection-based clustering approach for computer-aided detection of lung nodules in different modalities", Pattern Analysis and Applications, 22(2), 559-571 (2019).

[5] Narayanan, B. N., Hardie, R. C., & Kebede, T. M., "Analysis of various classification techniques for computer aided detection system of pulmonary nodules in CT", Proc. IEEE National Aerospace and Electronics Conference and Ohio Innovation Summit, 88-93 (2016).

[6] Narayanan, B. N., Hardie, R. C., & Kebede, T. M., "Performance analysis of a computer-aided detection system for lung nodules in CT at different slice thicknesses", Journal of Medical Imaging, 5(1), 014504 (2018).

[7] Narayanan, B. N., & Hardie, R. C., "A Computationally Efficient U-Net Architecture for Lung Segmentation in Chest Radiographs", Proc. IEEE National Aerospace and Electronics Conference and Ohio Innovation Summit (2019).

[8] Oliveira, L. L. G., e Silva, S. A., Ribeiro, L. H. V., de Oliveira, R. M., Coelho, C. J., & Andrade, A. L. S., "Computer-aided diagnosis in chest radiography for detection of childhood pneumonia", International journal of medical informatics, 77(8), 555-564 (2008).

[9] Abe, H., MacMahon, H., Shiraishi, J., Li, Q., Engelmann, R., & Doi, K., "Computer-aided diagnosis in chest radiology", Seminars in Ultrasound, CT and MRI, 25(5), 432-437 (2004).

[10] Giger, M., & MacMahon, H., "Image processing and computer-aided diagnosis", Radiologic Clinics of North America, 34(3), 565-596 (1996).

[11] Monnier-Cholley, L., MacMahon, H., Katsuragawa, S., Morishita, J., Ishida, T., & Doi, K., "Computer-aided diagnosis for detection of interstitial opacities on chest radiographs", American journal of roentgenology, 171(6), 1651-1656 (1998).
[12] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., & Lungren, M. P., "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning", *arXiv preprint arXiv: 1711.05225* (2018). Available online: https://arxiv.org/abs/1711.05225 (accessed on August 18, 2019).
[13] Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y., "Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification", arXiv preprint arXiv: 1801.09927. Available online: https://arxiv.org/abs/1801.09927v1 (accessed on August 18, 2019).
[14] Rajaraman, S., Candemir, S., Kim, I., Thoma, G., & Antani, S., "Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs", Applied Sciences, 8(10), 1715 (2018).
[15] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning", Cell, 172(5), 1122-1131 (2018).
[16] LeCun, Y., & Bengio, Y., "Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks", 3361(10), (1995).
[17] Krizhevsky, A., Sutskever, I., & Hinton, G. E., "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, 1097-1105 (2012).
[18] He, K., Zhang, X., Ren, S., & Sun, J., "Deep residual learning for image recognition", Proc. IEEE conference on computer vision and pattern recognition, 770-778 (2016).
[19] Simonyan, K., & Zisserman, A., "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556 (2014).
[20] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z., "Rethinking the inception architecture for computer vision", Proc. of the IEEE conference on computer vision and pattern recognition, 2818-2826 (2016).
[21] Narayanan, B. N., Ali R., & Hardie, R. C., "Performance analysis of machine learning and deep learning architectures for malaria detection on cell images", Proc. SPIE 11139, Applications of Machine Learning, 111390W (2019). https://doi.org/10.1117/12.2524681 (accessed on September 11, 2019)
[22] Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Xue, Z., Palaniappan, K., Singh, R.K., Antani, S. and Thoma, G., "Automatic tuberculosis screening using chest radiographs", IEEE transactions on medical imaging, 33(2), 233-245 (2013).
[23] Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Xue, Z., Karargyris, A., Antani, S., Thoma, G. and McDonald, C.J., "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration", IEEE transactions on medical imaging, 33(2), 577-590 (2013).
[24] Noh, Hyeonwoo, Seunghoon Hong, and Bohyung Han. "Learning deconvolution network for semantic segmentation." Proceedings of the IEEE international conference on computer vision. 2015.
[25] U-Net Architecture Implementation in MATLAB, https://www.mathworks.com/help/vision/ref/unetlayers.html Accessed September 26, 2019.