3-25-2021

Pneumonia Radiograph Diagnosis Utilizing Deep Learning Network

Wesley O'Quinn
Georgia Southern University

Pneumonia Radiograph Diagnosis Utilizing Deep Learning Network

An Honors Thesis submitted in partial fulfillment of the requirements for Honors in Electrical Engineering

By
Wesley O'Quinn

Under the mentorship of Dr. Rami Haddad

ABSTRACT

Abstract—Pneumonia is a life-threatening respiratory disease caused by bacterial infection. The goal of this study is to develop an algorithm using Convolutional Neural Networks (CNNs) to detect visual signals for pneumonia in medical images and make a diagnosis. Although Pneumonia is prevalent, detection and diagnosis are challenging. The deep learning network AlexNet was utilized through transfer learning. A dataset consisting of 11,318 images was used for training, and a preliminary diagnosis accuracy of 72% was achieved.

Thesis Mentor:_________________________

Dr. Rami Haddad

Honors Director:________________________

Dr. Steven Engel

April 2019
Electrical Engineering
University Honors Program
Georgia Southern University

I. ACKNOWLEDGEMENTS

We want to acknowledge the Radiological Society of North America (RSNA) and the Pneumonia Detection Challenge for providing a rich dataset to forward machine learning in this field. I would also like to thank my mentor Dr. Rami Haddad without whose aid, this project would not have taken place. I would also like to thank David Moore for his tireless efforts to simplify the code associated with this project. Finally, I would like to thank the Georgia Southern Undergraduate Research Fund for supplying the monetary needs for me to travel and present this work at ICEICT 2019 in Harbin, China.

II. INTRODUCTION

Exploring the automated diagnosis of pneumonia through deep learning is an extremely important issue to examine due to a variety of reasons. Pneumococcal pneumonia is a form of the pneumococcal disease. This particular variety is most common in adults [1]. Even though pneumonia is common in adults, it remains the leading infectious cause of death among children under the age of five according to UNICEF. Statistically, 2,400 children die from this disease every day. In 2016, around 880,000 children died with most of the victims being less than two years of age [2]. In the United States alone, pneumonia accounted for 50,000 deaths in 2015; it continues to be listed in the top 10 causes of death in the country [3]. In addition, this disease is responsible for over 500,000 visits to emergency departments every year [4]. Though prevalent, pneumonia remains difficult to diagnose. According to the American Lung Association, pneumonia can be diagnosed in a variety of ways including using a blood test to check white blood cell count, Arterial blood gases test, Sputum tests, pleural fluid culture, pulse oximetry, and bronchoscopy [5]. Despite there being many methods available, a chest radiograph remains a primary method used for diagnosis. Although this technique is commonly used, it is challenging to diagnose based upon these images. Highly trained specialists are needed to review the chest radiographs creating a large amount of work for them. To further complicate the matter, reading these images is problematic because pneumonia is usually manifested through an area or areas of increased opacity [6]. This opacity is the result of a decrease in the ratio of gas to soft tissue (blood, lung parenchyma, and stroma) in the lung. When reviewing an area of increased attenuation (opacification) on a chest radiograph or CT, it is vital to determine where the opacification occurs [7]. The diagnosis is complicated because of other conditions such as fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, post-radiation changes, or surgical changes. Other factors such as positioning of the patient and depth of inspiration can alter the appearance of the radiograph. All these factors contribute to making the images very challenging to interpret. A physical representation of these challenges can be seen in figures provided. Figures 1 and 2 provides an example of a healthy patient radiograph and a pneumonia positive radiograph, respectively. These two images show the lack of visual differences between the two classes.

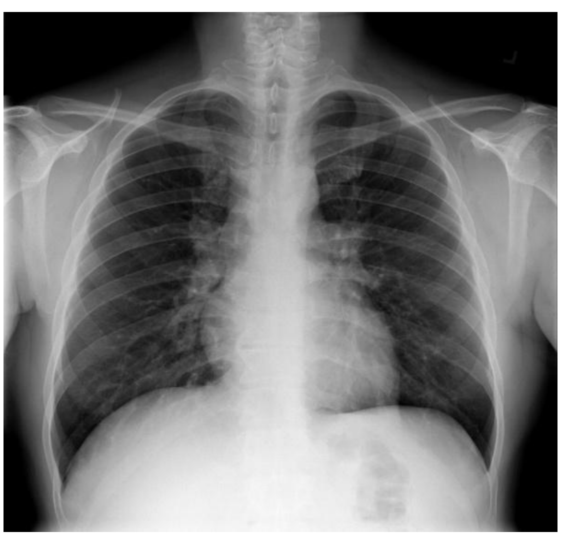

Figure 1: Pneumonia Negative

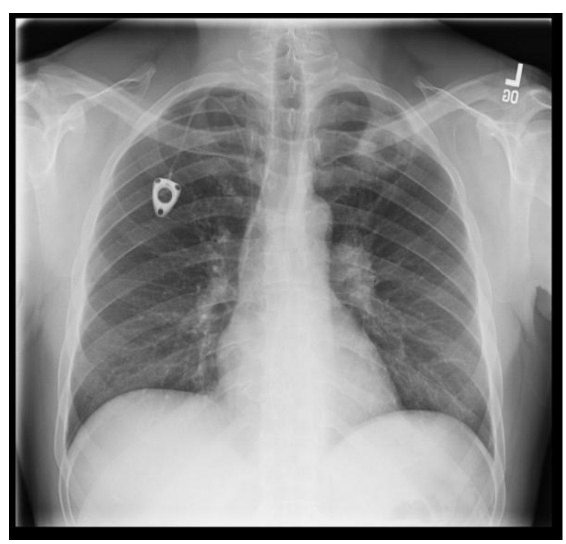

Figure 2: Pneumonia Positive

III. RELATED WORK

Though the fields of Pneumonia research and machine learning are highly developed individually, a much more limited amount of work is available on the application of machine learning to Pneumonia diagnosis. Deep learning algorithm research has progressed rapidly. In congruence, the diagnosis of pneumonia in the medical field is a widely known discipline. The combination of these two areas is novel. Research has been done in applying image processing techniques for identification. Sharma et al. investigated an approach that would assist in automating diagnostics [8]. Indigenous algorithms were developed for cropping and extracting the lung region from the images. Otsu thresholding was used to segregate the healthy part of the lung from the pneumonia infected cloudy region. In [9], a system detection method was proposed utilizing cellular neural networks. The simulation results show exceptional performance based on the difference grayscale color and segmentation between the normal area and lung region area.

IV. THEORY

Image classification is an important part of machine learning. This has led to a variety of developments in networks designed for image classification. These networks mostly utilize the backpropagation algorithm which was first proposed by Bryson and Yu-Chi [13]. In this research work, a subset of machine learning called deep learning networks were used. Deep learning mainly involves huge neural networks. There are a few commonly used deep learning architectures such as AlexNet, GoogleNet, and ResNet. AlexNet is the deep learning network chosen to be used which is a direct product of [12]. AlexNet was created as a deep convolutional neural network to classify over 1000 different classes. This network consists of 650,000 neurons and 60 million parameters which represent five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax [12]. This network proved to be state-of-the-art and has been used widely since its inception. The structure of AlexNet is shown in Figure 3.

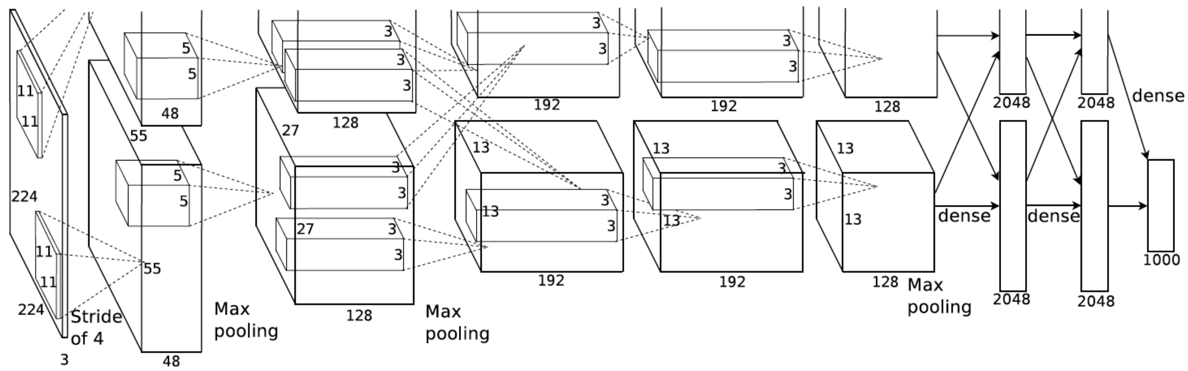

Figure 3: AlexNet Structure [12]

Structurally the first convolutional layer takes an input image of 227×227×3 with 96 kernels sized at 11×11×3 having a stride of 4 pixels. The output of this layer becomes the input of the second convolutional layer. This layer filters the data with 256 kernels each having a size of 5×5×48. This is transmitted to the third layer which utilizes 384 kernels sized at 3×3×256. The fourth layer has the same number of kernels as the third, but the size is 3×3×192. The final layer has 256 kernels at the same size as the fourth. Instead of using a standard model for the neurons output—represented by f(x) = tanh(x), Krizhevsky used a non-saturating nonlinearity represented as f(x) = max(0, x). Using this method, the neurons with this non-linearity are referred to as Rectified Linear Units (ReLUs). This system allowed the network to train several times faster than the equivalent models using tanh units. A graph from [12] is provided in Figure 4 to offer a visual representation of the training error rate vs. the number of epochs. The solid line represents a network using ReLU’s while the dashed line equates to a network using tanh neurons.

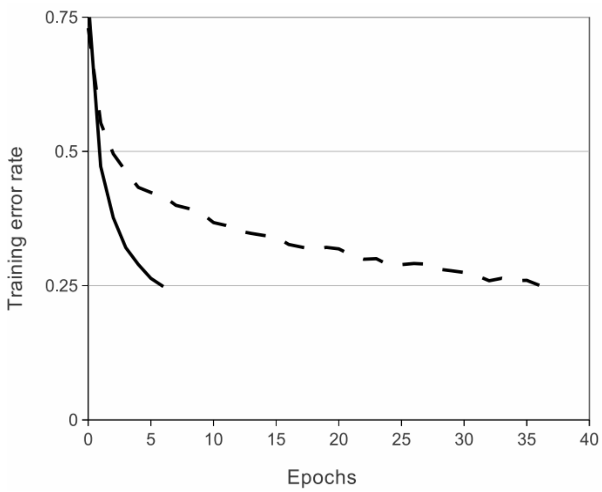

Figure 4: Training Error vs. Epochs

Another property of ReLUs is that they do not require input normalization to prevent them from becoming saturated.

V. PROPOSED METHOD

The proposed solution to the problem opened in the introduction was to train a deep learning network with the ability to recognize symptoms of pneumonia. Through this, pre-screening or even automated diagnosis will be made available. Deep learning was the computational model that was used. This model discovers “intricate structure in large datasets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer.” [10]. The specific platform used was a subset of Matlab. Recently, MatConvNet was developed as a toolbox for implementing

Convolutional Neural Networks (CNNs) for computer vision applications. As with most similar applications, this platform allows computations on a system's Graphics Processing Unit (GPU) for faster processing. A primary reason for choosing this application is that it is well supported in the Matlab environment. By using this powerful platform, we will be able to implement a pre-trained deep learning network to this field in a way that has not been attempted in prior research.

Training Data Utilized

The dataset utilized to train this network was obtained through the Radiological Society of North America (RSNA) Pneumonia Detection Challenge. This society worked in congruence with US National Institutes of Health, The Society of Thoracic Radiology, and MD.ai. These organizations were able to develop a large dataset of annotated images. The competition utilized a subset of chest radiographs released by the National Institutes of Health Clinical Center. In total, the training set available included over twenty-five thousand images. Each of these included a tag specifying either a positive or negative diagnosis for Pneumonia. The files in question were stored in Digital Imaging and Communications in Medicine (DICOM) format which is the standard image file format used by radiological hardware devices [11]. This type of image is utilized worldwide for storing and transmitting medical images since this is not a common file format.

Deep Learning Model

Two options were considered to apply deep learning to this project. The first option was to go through the training process of a new deep neural network. Although this method could be pursued, large amounts of data and training time were required. The second option was to fine-tune a pre-trained model using transfer learning. In this study, transfer learning was used. Therefore, a pre-trained deep convolutional neural network was employed. The CNN utilized was AlexNet. As training is computation-intensive, a graphics card was used to reduce the training time. This specific network was initially trained to identify one thousand different classes. This being the case, the network was manipulated by performing net surgery. The last three layers were removed and replaced by a new fully-connected layer specifically for classifying two categories. A Softmax layer also added thereby applying the Softmax function to the input. Finally, the remaining layers were standard classification layers. The layers graphed in can be seen in Figure 5.

21	'fc8_2'	Fully Connected	2 fully connected layer
22	''	Softmax	softmax
23	''	Classification Output	crossentropyex

Figure 5: Last Three Layers after Net Surgery

Next, the learning rates of the newly added layers were altered. These learning rates were boosted while keeping the original layers nearly the same. The structure of AlexNet prior to net surgery is shown in Figure 3.

Input Data

Since the images were made available in DICOM format, it was decided that we would continue to process the data in this format. This was done by creating a customized datastore function. For initial training purposes, each positive/negative label was counted to ascertain which group was the smallest. This group happened to be pneumonia positive. This class was tabulated to contain 5659 images. The negative pneumonia set was trimmed to obtain an equal number of each. Once this was accomplished, the entire set of images was divided into two separate groups. One group was the training dataset utilizing 70% of the images, and the remaining 30% of images were stored as validation data.

Executing Program

Once all the information was loaded into the program and the data manipulated appropriately the system could be run. The batch size was set to 128 images. The number of epochs was set to twenty, an epoch being the entire dataset being passed forward/backward through the network once. Since the total number of training images was 7,922, it took approximately 62 iterations to complete one epoch. Different networks were trained with variations in the amount of time needed to produce a working network. On average, the training took slightly over two hours. Provided in Figure 6 is an example of the training process.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy
1	1	00:00:08	44.53%
1	50	00:06:37	71.88%
2	100	00:12:23	71.88%
3	150	00:18:06	75.00%
4	200	00:24:02	78.13%
5	250	00:29:54	80.47%
5	300	00:35:00	81.25%
6	350	00:40:14	85.16%
7	400	00:45:03	83.59%
8	450	00:49:40	82.03%
9	500	00:54:47	84.38%
10	550	01:00:10	85.94%
10	600	01:05:02	87.50%
11	650	01:09:42	85.94%
12	700	01:14:27	89.06%
13	750	01:19:22	89.84%
14	800	01:24:05	91.41%
14	850	01:28:49	94.53%
15	900	01:33:41	96.09%
16	950	01:38:47	96.88%
17	1000	01:43:33	98.44%
18	1050	01:48:06	99.22%
19	1100	01:53:05	96.88%
19	1150	01:57:52	95.31%
20	1200	02:02:25	95.31%
20	1220	02:04:12	95.31%

Figure 6: Net Training Process

VI. RESULTS

Different training sessions were run with mixed results. Using a separate validation set the network was able to achieve an initial accuracy of 72%. This is significant because the images used for validation were different from the training images. This gives an accurate representation of how the network would function in real-world applications. Although this was lower than expected, it provides a good baseline for future research. Provided in Figure 7 is the confusion matrix for the network run on randomly selected images from the training set.

```
confMat =

    0.6908    0.3092
    0.0389    0.9611
```

Figure 7: Confusion Matrix

Identification Insight

After receiving the results, a script was developed to run through all the images in the validation set. This code would then identify which images were misclassified and display them in a visual format. From this, we were able to gain significant insight into how the network classified images. The misclassified images consisted of several outliers that revealed three common causes of misidentification. The images usually were lower quality with regard to clarity, were taken at a wrong angle, or contained a medical device such as a pacemaker. Figures 8, 9 and 10 provide examples of these respectively.

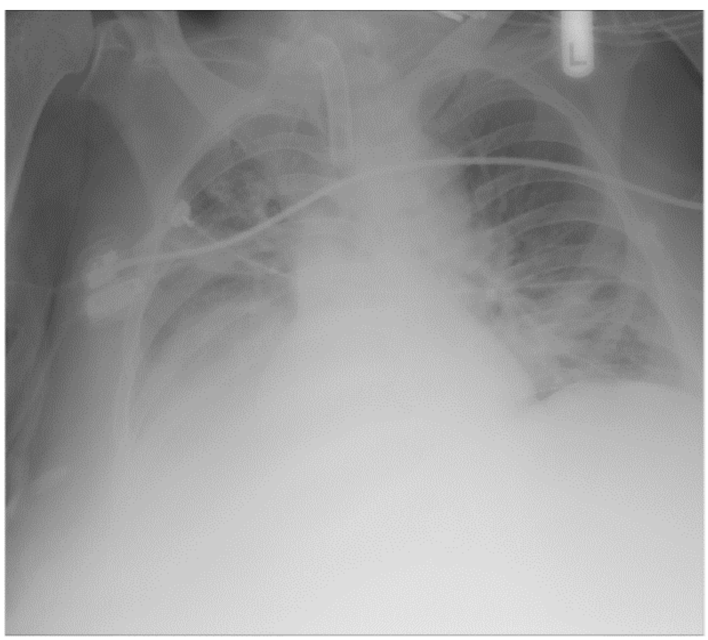

Figure 8: Misidentified Image (Low Quality)

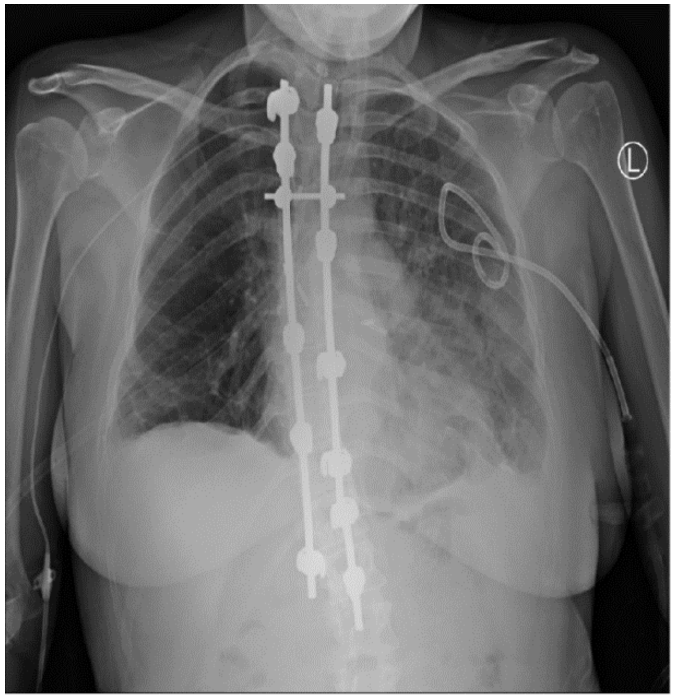

Figure 9: Misidentified Image (Irregular Device)

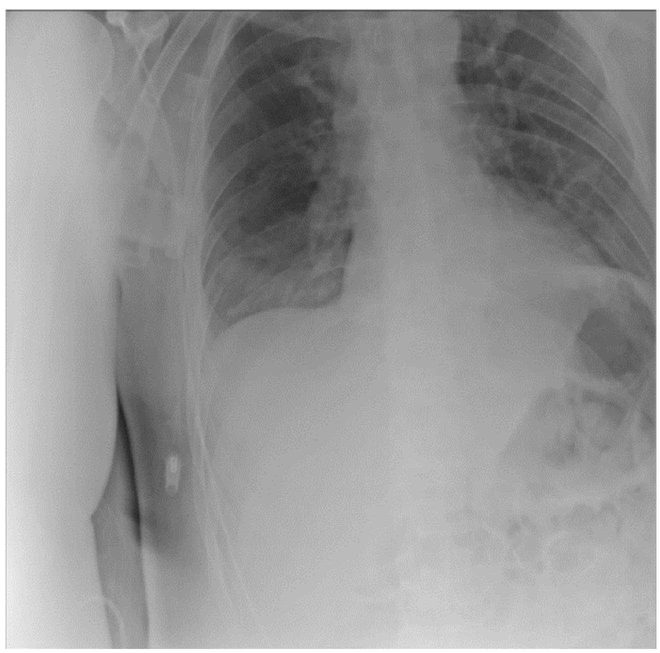

Figure 10: Misidentified Image (Bad Angle)

Proposed Solutions

One straightforward approach to increasing the accuracy will be to review and manually remove all images that are irregular or low quality. The reason we do not plan on utilizing this approach is that we desire the network to be utilized in the real world. This solution would only mask the issue. The other approach will be to increase our dataset to allow the network to train itself to identify irregularities and still be able to classify the image. It is also planned to apply some form of pre-processing to improve the quality of low-resolution scans.

VII. CONCLUSION

In conclusion, a model was developed that utilizes the deep learning functions of convolutional neural networks to detect visual signals for pneumonia in medical images and make a diagnosis. Potential pitfalls have been identified in the process that will allow for the streamlining of the process as the research moves forward. This research also provides insight into the testing difficulties associated with large datasets of radiograph images. Most research in this field has focused on small, controlled datasets. With these datasets, irregular features may not have been much of an issue. But as the size of the dataset increases, so does the number of irregular images. Although this is a potential difficulty, it is desired because a model utilized by the medical industry needs to be able to consider all forms of radiographs. Even if these types of images are imputed, the model still needs to function at a high level.

VIII. REFERENCES

[1] Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases. Hamborsky J, Kroger A, Wolfe S, eds. 13th ed., Washington DC: Public Health Foundation, 2015.

[2] "Pneumonia." UNICEF DATA, data.unicef.org/topic/child-health/pneumonia/.

[3] Deaths: Final Data for 2015. Supplemental Tables. Tables I-21, I-22. Available: www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr66_06_tables.pdf

[4] Rui P, Kang K. National Ambulatory Medical Care Survey: 2015 Emergency Department Summary Tables. Table 27. Available from: www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf

[5] Diagnosing and Treating Pneumonia. (n.d.). Retrieved from https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/pneumonia/diagnosing-and-treating.html

[6] Franquet T. Imaging of community-acquired pneumonia. J Thorac Imaging 2018 (Epub ahead of print). PMID 30036297

[7] Jones, J. (n.d.). Pulmonary opacification | Radiology Reference Article. Retrieved from https://radiopaedia.org/articles/pulmonary-opacification

[8] A. Sharma, D. Raju, and S. Ranjan, "Detection of pneumonia clouds in chest X-ray using image processing approach," 2017 Nirma University International Conference on Engineering (NUiCONE), Ahmedabad, 2017, pp. 1-4.

[9] A. A. Abdullah, Norafifah Md Posdzi and Y. Nishio, "Preliminary study of pneumonia symptoms detection method using Cellular Neural Network," International Conference on Electrical, Control and Computer Engineering 2011 (InECCE), Pahang, 2011, pp. 497-500.

[10] LeCun, Y., Bengio, Y., & Hinton, G. (2015, May 27). Deep learning. Retrieved from https://www.nature.com/articles/nature14539

[11] DICOM demystified: A review of digital file formats and their use in radiological practice Graham R.N.J., Perriss R.W., Scarsbrook A.F. (2005) Clinical Radiology, 60 (11), pp. 1133-1140.

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12), F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), Vol. 1. Curran Associates Inc., USA, 1097-1105.

[13] Bryson A E, Jr. & Ho Y C. Applied optimal control: optimization, estimation, and control. Waltham, MA: Blaisdell, 1969. 481 p. [Stanford Univ., Stanford, CA and Harvard Univ., Cambridge, MA]