A Survey on Detection of Pneumonia from chest X-ray Images using Machine learning techniques

Bhumika Sharma

Abstract

One of the hottest topics in medical imaging research is the automatic detection of chest disease using chest radiography. The study conducts a thorough survey on machine learning methods for pneumonia detection based on clinical applications, with a special emphasis on artificial intelligence technology used in chest radiography. The paper presents several common chest X-ray datasets and briefly discusses general machine learning processes used in chest radiography, such as CNN, Deep learning, Supervised learning, and Transfer learning.

Finally, all these techniques are compared and analyzed to find out the best way for the detection of pneumonia disease.

Keywords - CNN, Machine learning, Supervised learning, Densenet, Deep learning, Transfer learning.

I. INTRODUCTION

In either or both lungs, pneumonia is an illness. It is brought on by bacteria, viruses, and fungi. Within the air sacs of your lungs, which are called alveoli, the infection triggers inflammation. The alveoli are filled with fluid or pus, making breathing difficult. Pneumonia-causing germs are infectious. This means that they will spread from individual to individual. Through inhaling airborne droplets from a sneeze or cough, both viral and bacterial pneumonia will spread to others.You can also get these kinds of pneumonia by coming into contact with surfaces or objects that are infected by bacteria or viruses that cause pneumonia. You'll contract environmental fungal pneumonia. It does not, however, spread from individual to person. Usually, symptoms include a combination of productive or dry cough, chest pain, fever, and breathing difficulties. The condition's severity is variable. Pneumonia is typically caused by infection with viruses or bacteria, and by other microorganisms less commonly.[a] It may be difficult to determine the responsible pathogen. Sometimes, diagnosis is based on symptoms and physical examinations. X-rays of the lung, blood tests, and sputum culture may help validate the diagnosis.

The condition can be defined by where it was acquired, such as neighborhood- or hospital-acquired or pneumonia associated with healthcare. Cystic fibrosis, chronic obstructive pulmonary disease (COPD), sickle cell disease, asthma, diabetes, heart failure, a history of smoking, poor cough capacity (such as after a stroke), and a weak immune system are risk factors for pneumonia.

Vaccines are available to avoid some forms of pneumonia (such as those caused by bacteria infected with Streptococcus pneumonia or influenza). Other methods of prevention include washing hands and not smoking to prevent infection. Treatment depends on the underlying cause. Pneumonia is treated with antibiotics that are thought to be due to bacteria. The infected person is usually hospitalised if the pneumonia is serious. If oxygen levels are low, oxygen therapy can be used.Pneumonia affects about 450 million people worldwide per year (7 percent of the population) and results in about 4 million deaths. Survival has improved dramatically with the advent of antibiotics and vaccines in the 20th century. However, pneumonia remains a leading cause of death in developed countries, as well as among the very old, the very young and the chronically ill. Pneumonia also shortens the suffering time of those who are already near to death and has thus been named "the old man's friend" .

Types of Pneumonia : The main types of pneumonia are:

1. **Bacterial pneumonia.** Different bacteria cause this form. Streptococcus pneumonia is the most common. It typically happens when the body, such as sickness, inadequate nutrition, old age, or compromised immunity, is damaged in some way, and the bacteria are capable of making their way into the lungs. Bacterial pneumonia can affect all ages, but if you drink alcohol, smoke cigarettes, are weakened, have recently undergone surgery, have a respiratory or viral infection, or have a weakened immune system, you are at higher risk.

2. **Viral pneumonia.** This type is caused by various viruses, including influenza (flu), and accounts for around one-third of all cases of pneumonia. If you have viral pneumonia, you might be more likely to get bacterial pneumonia.

3. **Mycoplasma pneumonia.** This form has symptoms and physical signs that are very different and is referred to as atypical pneumonia. It is caused by the Mycoplasma pneumonia bacterium. It normally triggers a moderate, widespread lung disease that affects all age groups.

4. **Community-acquired pneumonia.** This is a fancy way to suggest that you have been compromised somewhere other than a long-term care facility or hospital. Bacteria, viruses, and fungi may cause community-acquired pneumonia.Vaccines can help protect against some bacteria that can also cause pneumonia and against the flu virus. Aspiration pneumonia, which occurs when you breathe food, fluid, or vomit into your lungs, also involves community-acquired pneumonia. If you have trouble with swallowing or coughing, it's more likely. Bacteria will multiply in your lungs if you can not cough up the substance you have taken in.

5. **Walking pneumonia**. It is a type of bacterial pneumonia that is less severe. Doctors also call it 'atypical' pneumonia. There could be signs that are so slight that you don't know you have them. You can feel well enough that you can go about your daily activities, from which the "walking" in the name comes.

6. **Fungal Pneumonia.** A less common cause of pneumonia is fungi. When you're well, you're not likely to get fungal pneumonia. But if the immune system is compromised by an organ transplant, chemotherapy for cancer, medications to cure an autoimmune condition such as rheumatoid arthritis, HIV, you have a better risk of catching it. By breathing in tiny particles called fungal spores, you get fungal pneumonia.

Typically, the diagnosis is based on the recent health history (such as surgery, cold, or exposure to travel) and the seriousness of the disease. Your healthcare provider can diagnose pneumonia simply on a detailed history and physical exam based on these variables. The following tests may be used to confirm the diagnosis:

A. **Chest X-ray.** This test takes photographs, like the lungs, of internal tissues, bones, and organs.
B. **Blood tests.** The procedure will be used to see if there is an infection and whether the infection has spread to the bloodstream (blood cultures). The amount of oxygen in your bloodstream is tested by arterial blood gas monitoring.

C. **Sputum culture.** This procedure is carried out on the material from the lungs and into the mouth that is coughed up. It's also used to see if the lungs have an infection.

D. **Pulse oximetry.** A small machine which measures the amount of oxygen in the blood is an oximeter. It is taped or clipped onto a finger by a tiny sensor. When the computer is on, the sensor can see a tiny red light. The test is painless and it does not get hot in the red light.

E. **Chest CT scan.** In order to generate sharp, accurate horizontal or axial images (often called slices) of the body, this imaging technique uses a combination of X-rays and computer technology. A CT scan shows clear pictures, including bones, muscles, fat, and organs, of every part of the body. There are more detailed CT scans than normal X-rays.

F. **Bronchoscopy.** This is a direct inspection of the bronchi (the lungs' main airways) using a flexible tube (called a bronchoscope). It helps to determine and diagnose lung disorders, to evaluate blockages and to take tissue and/or fluid samples for examination.

G. **Pleural fluid culture.** A sample of a fluid sample is taken from the pleural space in this examination. This is the space between the chest wall and the lungs. A long, thin needle between the ribs and into the pleural space is inserted through the skin. Fluid is drawn into a needle-attached syringe. It is sent to the laboratory where it is examined to find out which pneumonia causes the bacteria.

II. LITERATURE SURVEY

Pneumonia is a life-threatening infectious disease affecting one or both lungs in humans commonly caused by bacteria called Streptococcus pneumoniae. Chest X-Rays which are used to diagnose pneumonia need expert radiotherapists for evaluation. Thus, developing an automatic system for detecting pneumonia would be beneficial for treating the disease without any delay particularly in remote areas. Due to the success of deep learning algorithms in analyzing medical images, Convolutional Neural Networks have gained much attention for disease classification. This paper primarily aims to improve the medical adeptness in areas where the availability of radiotherapists is still limited.Our study facilitates the early diagnosis of Pneumonia to prevent adverse consequences (including death) in such remote areas[1] Pneumonia is one of the most fatal diseases caused in the lungs. The diagnosis involves a chest x-ray which is interpreted by a radiologist. Human assisted diagnosis has its own limitations like the availability of an expert, cost, etc and hence an automated method for the detection of pneumonia from x-rays is a necessity. In this research, neural network models were developed to detect pneumonia from the chest x-ray images. Four models namely a basic

convolutional neural network [11] [16], VGG16, VGG19, Inception [18] V3 were constructed using CNN [11] and transfer learning [7] methodologies. The models were then trained on a pediatric pneumonia dataset which consisted of 2992 pneumonia and 2972 normal chest x rays. The results were then tested using 854 pneumonia and 849 normal images, and an accuracy of over 97 percent was obtained from all models[2].

We develop an algorithm that can detect pneumonia from chest X-rays at a level ex- ceeding practicing radiologists. Our algorithm, CheXNet, is a 121-layer convolutional neural network trained on Chest X-ray 14, currently the largest publicly available chest X- ray dataset, containing over 100,000 frontal- view X-ray images with 14 diseases. Four practicing academic radiologists annotate a test set, on which we compare the performance of CheXNet to that of radiologists. We find that CheXNet exceeds average radiologist performance on the F1 metric. We extend CheXNet to detect all 14 diseases in ChestX-ray14 and achieve state of the art re- sults on all 14 diseases. [3]. Chest X-rays are often used to diagnose pneumonia, which kills around 50,000 people each year. Physicians can more effectively and quickly diagnose pneumonia using computer-assisted diagnosis. We hope to train a model using the dataset listed below to assist physicians in making pneumonia diagnoses in chest X-rays as part of this project.
The NIH dataset's images are 1024x1024. To begin, we used an anti-aliasing filter to resize each image. Our Logistic Regression baseline uses 32x32 resolution. The resolution of our Deep Learning model is 224x224. We also standardize the data to ensure that each function (pixel) has a zero mean and roughly unit variance.[4]

This research proposes a convolutional neural network model that can identify and detect the presence of pneumonia from a series of chest X-ray image samples, and it was trained from scratch. We built a convolutional neural network model from scratch to extract features from a given chest X-ray image and classify it to determine whether an individual is infected with pneumonia, unlike other methods that depend solely on transfer learning [2] approaches or conventional handcrafted techniques to achieve a remarkable classification output. The model may be able to help with the reliability and interpretability issues that come up often when working with medical images. Unlike other deep learning classification tasks with a large image repository, obtaining a large amount of pneumonia dataset for this classification task is difficult; as a result, we used multiple data augmentation algorithms to boost the CNN model's validation and classification accuracy, and we achieved remarkable validation accuracy [5]. Deep learning-based [16] computer assisted diagnosis systems and medical imaging are becoming highly popular research topics.

Currently, classification results are produced by hierarchically abstracting the original image using a classical convolutional neural network. These abstract features are less sensitive to the object's location and orientation, and the lack of spatial information restricts image classification accuracy development. As a result, researchers must continue to investigate how to build a suitable neural network architecture and training strategy in realistic clinical applications to avoid this issue. The novel [17] framework relies on the use of advanced classification to conduct direct lesion characterization and has a high level of accuracy in the classification task of children's pneumonia.[6]

III. METHODOLOGY

3.1 CNN

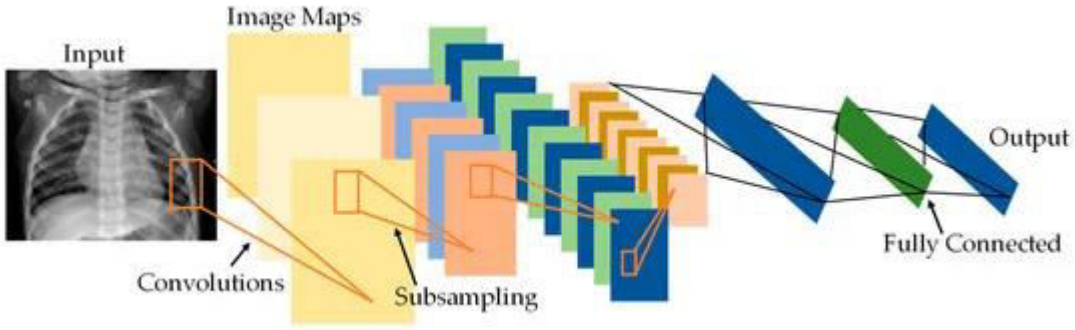

Fig 1: CNN model [7]

CNNs are simply feedforward artificial neural networks (ANNs) with two restrictions: neurons in the same filter are only bound to local patches of the image to maintain spatial structure, and their weights are shared to minimize the number of parameters in the model.

A CNN is made up of three layers: I a convolution layer for learning features, (ii) a max-pooling (subsampling) layer for downsampling the picture and reducing dimensionality, and (iii) a completely linked layer for equipping the network with classification capabilities [12]. Figure 1 provides a high-level overview of CNN's architecture.

3.2 DEEP TRANSFER LEARNING

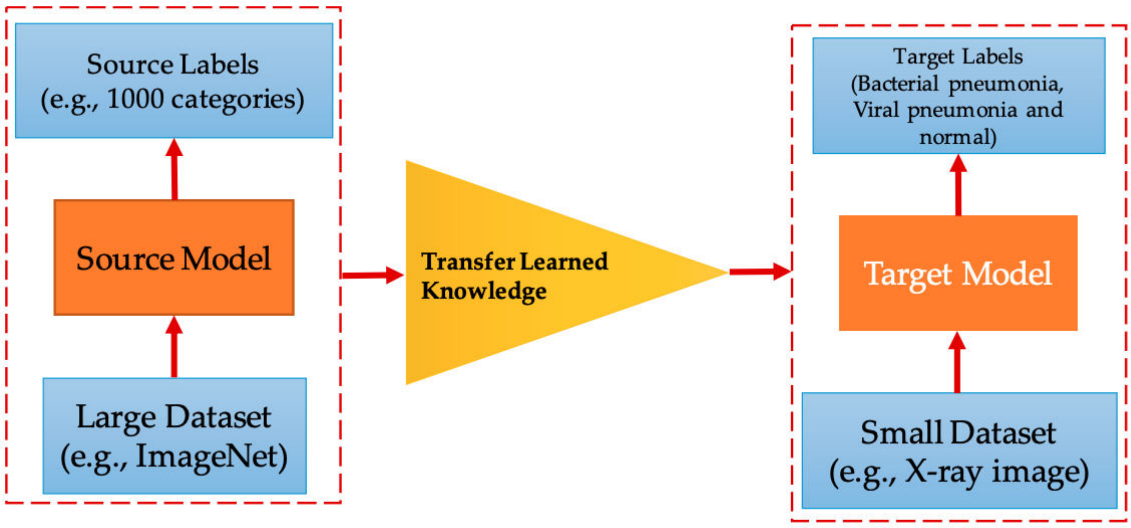

Fig 2: Transfer learning [7]

In most instances, a larger dataset outperforms a smaller one. In CNN applications where the dataset is not huge, transfer learning can be useful. The definition of transfer learning is illustrated in Figure 2, in which a trained model from a large dataset, such as ImageNet [13][14], can be applied to a smaller dataset.

3.3 DEEP NEURAL NETWORK

Deep neural network models have traditionally been developed, and human experts have conducted experiments on them in a trial-and-error method. This procedure necessitates a significant amount of time, expertise, and resources. To address this issue, a novel but simple model is presented that uses deep neural network architecture to automatically perform optimal classification tasks. The neural network architecture was developed specifically for the classification of pneumonia images. The suggested method is based on the convolutional neural network algorithm, which uses a set of neurons to convolve and extract relevant features from a given picture. The proposed method's effectiveness was demonstrated using the minimization of computational cost as the focal point, and it was compared to existing state-of-the-art pneumonia classification networks. [21]

3.4 DENSENET-169

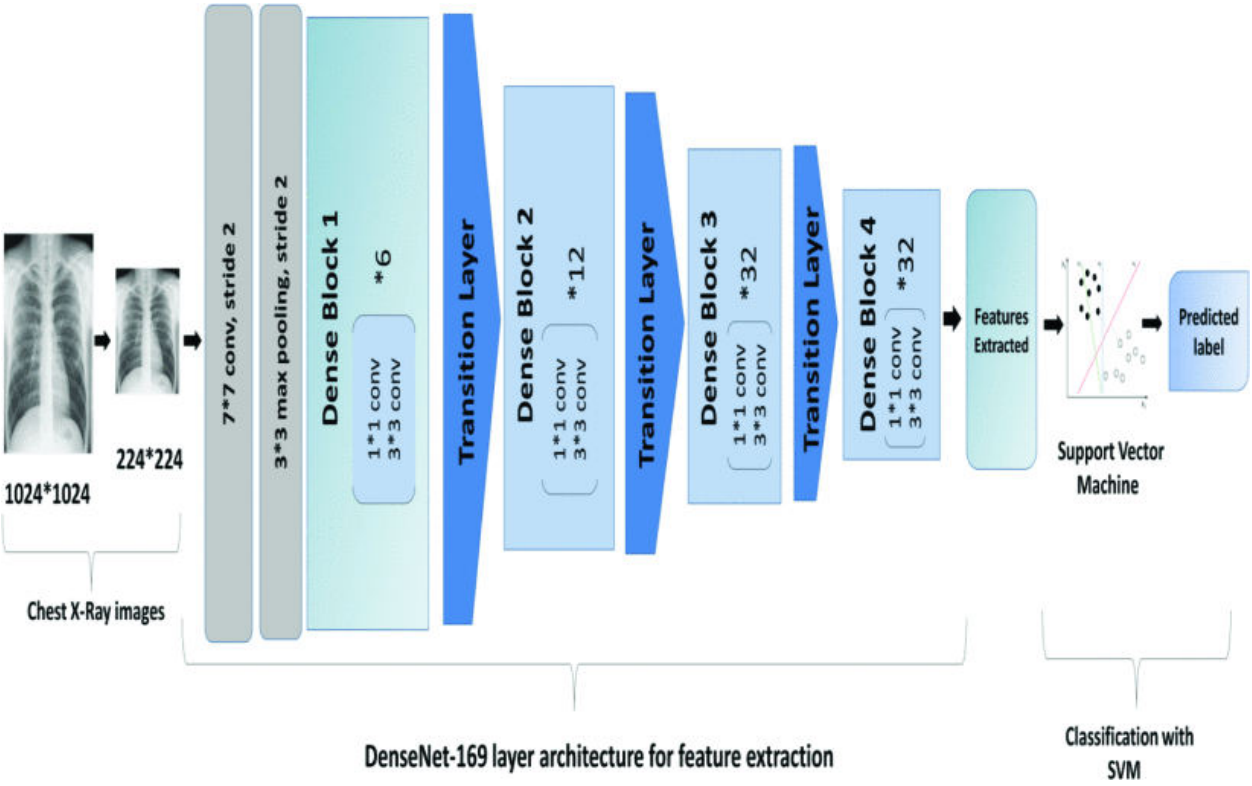

Fig 3: DenseNet-169 architecture [1]

One convolution and pooling layer, three transition layers, and four dense blocks make up the DenseNet-169 architecture. Following these layers is the classification layer, which is the final layer. With stride 2, the first convolutional layer performs 77 convolutions, followed by a maximum pooling of 33 with stride 2. The network then consists of a dense block, followed by three sets, each consisting of a transition layer and a dense block. DenseNets receives the dense connectivity suggested by Huang et al [22] by adding in direct connections from any layer to any other layer in the network. The network's lth layer receives the feature-maps of all the preceding layers, improving gradient flow across the entire network. This necessitates the concatenation of the feature-maps of the preceding layers, which can only be done if all of the feature-maps are the same size. However, because Convolutional Neural Networks are mainly designed to reduce feature-map size downsampling, the DenseNets architecture is split into several densely connected dense blocks.

3.5 ALEXNET ARCHITECTURE

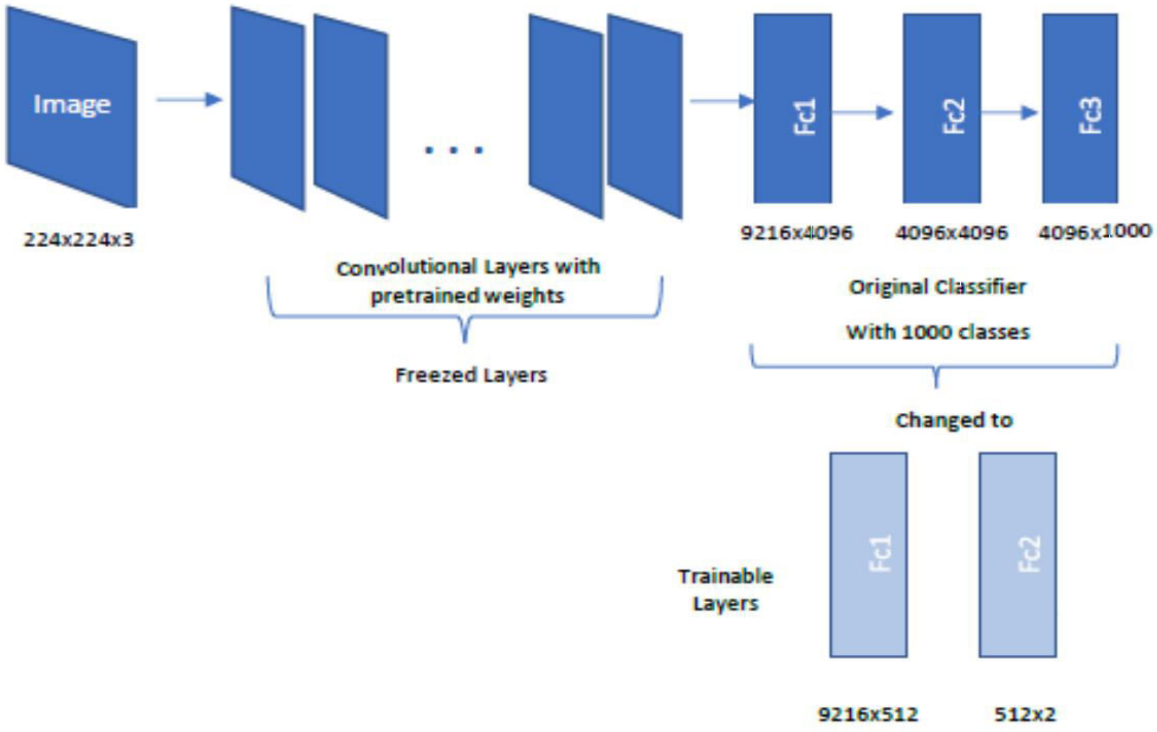

Fig 4: AlexNet with trainable and "frozen" layers [8]

AlexNet is a CNN similar to LeNet [15], but with more depth. To add non-linearity, this network replaced the tanh function with a Rectified Linear Unit (ReLU). To deal with overfitting, it used dropout layers rather than regularization. By freezing the convolutional layers and just training the classifier, overlapping pooling was also used to minimize.

3.6 RESNET-18

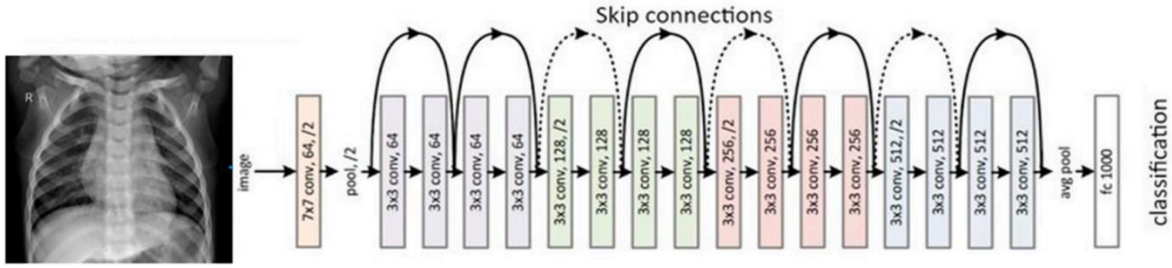

Fig 5: ResNet18 Structure

ResNet (short for Residual Network) was created to solve two problems: the vanishing gradient problem and the degradation problem [19]. Residual learning aims to address all of these issues. ResNet is divided into three types based on the number of layers in the residual network: ResNet18, ResNet50, and ResNet101. ResNet was successfully used for transfer learning in biomedical image classification [20]. We used ResNet18 for pneumonia detection in this study. During training, deep neural network layers typically learn low- or high-level features, while ResNet learns residuals instead of features.

IV. RESULT ANALYSIS

Table 1: Result analysis of various techniques

Reference	Method	Dataset	Result
1.	CNN	Chest X-RAY Image	0.80(AUC)
2.	CNN and Transfer Learning	Chest X-RAY Image	97-98%
3.	CNN with Deep learning	Chest X-RAY Image	95%
4.	Supervised Learning	Chest X-RAY Image	90%
5	Deep neural network	Chest X-RAY Image	93-95%
6	Deep learning	Chest X-RAY Image	96%
23	Machine learning	Chest X-RAY Image	80.3%

we can observe from table1 that on applying the method CNN the author gets 0.80 accuracy approx. The next method is CNN and Transfer learning where they get accuracy approx in between 97 to 98. Another method is CNN with Deep learning, here we get accuracy 95%

V. CONCLUSION

We have studied lots of papers on detection of pneumonia disease using machine learning techniques. And from Table 1 we can conclude that on applying the method CNN the author gets 0.80 accuracy approx. The next method is CNN and Transfer learning where they get accuracy approx in between 97 to 98. Another method is CNN with Deep learning, here we get accuracy 95% . Now move to another method which is supervised learning, the accuracy was 90%. Method is a Deep neural network where the author gets accuracy in between 93 to 95. Last method we studied was machine learning whose accuracy is 80.3%.

REFERENCES

1. Dimpy Varshni ,Rahul Nijhawan,Kartik Thakral ,Lucky Agarwal , Ankush Mittal. 2019. Pneumonia Detection Using CNN based Feature Extraction (2019), IEEE 2019
2. Gaurav Labhane; Rutuja Pansare; Saumil Maheshwari; Ritu Tiwari; Anupam Shukla.2020. Detection of Pediatric Pneumonia from Chest X-Ray Images using CNN and Transfer Learning(2020) IEEE 14 May 2020
3. Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, and others. 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)
4. Benjamin Antin , Joshua Kravitz , and Emil Martayan. Detecting Pneumonia in Chest X-Rays with Supervised Learning
5. Okele stephen et.al " An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare ". Journals of healthcare engineering 2019(2019)
6. Liang gaobo and Lixin Zheng. "A transfer learning method with deep residual network for pediatric pneumonia diagnosis." computer method and programs in Biomedicine (2019)
7. Rahman, Tawsifur, et.al. "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray." Applied Sciences 10.9 (2020): 3233.
8. Chouhan, Vikash, et al. "A novel transfer learning based approach for pneumonia detection in chest X-ray images." Applied Sciences 10.2 (2020): 559.
9. https://www.webmd.com/lung/pneumonia-types, Reviewed by Brunilda Nazario, MD on April 08, 2020
10. https://www.hopkinsmedicine.org/health/conditions-and-diseases/pneumonia

11. Varshni, Dimpy, et al. "Pneumonia detection using CNN based feature extraction." *2019* IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2019.
12. Bailer, C.; Habtegebrial, T.; Varanasi, K.; Stricker, D. Fast Feature Extraction with CNNs with Pooling Layers. arXiv 2018, arXiv:1805.03096
13. Wang, S.-H.; Xie, S.; Chen, X.; Guttery, D.; Tang, C.; Sun, J.; Zhang, Y.-D. Alcoholism Identification Based on an AlexNet Transfer Learning Model. Front. Psychol. 2019, 10, 205. [CrossRef]
14. Christodoulidis, S.; Anthimopoulos, M.; Ebner, L.; Christe, A.; Mougiakakou, S.G. Multisource Transfer Learning With Convolutional Neural Networks for Lung Pattern Analysis. IEEE J. Biomed. Health Inform. 2016, 21, 76–84. [CrossRef]
15. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. Proc. IEEE 1998, 86, 2278–2324. [CrossRef]
16. Hashmi, Mohammad Farukh, et.al. "Efficient pneumonia detection in chest x ray images using deep transfer learning." Diagnostics 10.6 (2020): 417.
17. Erdem, Ebru, and Tolga AYDİN. "Detection of Pneumonia with a Novel CNN-based Approach." Sakarya University Journal of Computer and Information Sciences 4.1 (2021): 26-34.
18. Harford, Carl G., and Mary Hara. "Pulmonary edema in influenzal pneumonia of the mouse and the relation of fluid in the lung to the inception of pneumococcal pneumonia." The Journal of experimental medicine 91.3 (1950): 245.
19. ResNet, AlexNet, VGGNet, Inception: Understanding various architectures of Convolutional Networks. Available online: https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/ (accessed on 23 December 2019)
20. LeCun, Y.; Kavukcuoglu, K.; Farabet, C. Convolutional networks and applications in vision. In Proceedings of the 2010 IEEE International Symposium on Circuits and Systems, Paris, France, 30 May–2 June 2010; pp. 253–256.
21. Hashmi, Mohammad Farukh, et al. "Efficient pneumonia detection in chest xray images using deep transfer learning." *Diagnostics* 10.6 (2020): 417.
22. Gao Huang, Zhuang Liu, Laurens Van Der Maaten and Kilian Q Weinberger, "Densely Connected Convolutional Networks", *CVPR*, vol. 1, pp. 3, 2017.

23. Tajima, Takeshi, et al. "Etiology and clinical study of community-acquired pneumonia in 157 hospitalized children." Journal of infection and chemotherapy 12.6 (2006): 372-379.