Detection of Pneumonia by Analyzing Chest X-Ray in Human using CNN

Dhanraj Babasaheb Bhosale[1], Nunna Lakshmi Saranya[2], Aamir Feroze Siddiqui[3], Tanya Bisht[4], Ashish Chandrashekar[5]

---***---

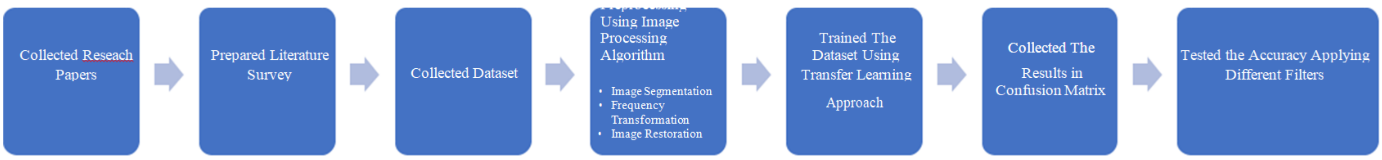
ABSTRACT

Pneumonia is one of the most common infection in the world today generally caused by bacteria, fungi or viruses. Pneumonia is swelling of the tissue in one or both lungs. It's usually caused by a bacterial infection. Few of the most common symptoms of pneumonia include cough, difficult breathing, rapid heartbeat, high temperature, sweating and shivering, loss of appetite, chest pain etc. It is responsible for more than one million hospitalizations and fifty thousand deaths per year in United States alone. Most common practice to identify the symptoms of pneumonia by performing chest X - ray. Two of the most common types of pneumonia found in human lungs are viral and bacterial. Doctors can distinguish between viral and bacterial pneumonia by examining the chest X – ray.

OBJECTIVE

In this project we are going to visualize and develop a Convolution Neural Network (CNN) to detect pneumonia symptoms in the patient and further differentiate between bacterial pneumonia and viral pneumonia and we are also going to preprocess the image dataset using various the image processing algorithms.

PROJECT WORK FLOW



DATASET USED

https://kaggle.com/paultimothymooney/chest-xray-pneumonia

Digital Image Processing means processing digital image by means of a digital computer. It includes the use of computer algorithms, in order to get enhanced image either to extract some useful information. Image processing consists of the various steps including importing the image called Image Acquisition. After that analysing and manipulating the image using various algorithms and then obtaining the output which can be the altered image or a report which is based on the analysis of the input image.

We already collected the image dataset in first review, now we are going to perform manipulation of image using below mentioned techniques.

LITERATURE REVIEW

JC Souza et al. [1] proposed that one of the most used imaging techniques for detection of pneumonia is Chest X-ray (CXR). Researchers are exploring the application of various image processing and machine learning algorithms to develop computer aided detection (CAD) systems which provides support in the challenging task. This paper proposed a fully automatic method using a CNN model for lung field segmentation in CXR.

HJ Zar et al. [2] proposed that one of the major causes of morbidity in children has turned out to be pneumonia. A study done by Darkenstein Child Health suggests pneumonia is spreading highly despite excellent immunization coverage. According to the study, new interventionsare required for preventing early life LRTI to ensure long term health.

JD Dear [3] proposed that dogs are mostly diagnosed with bacterial pneumonia while in the case of cat inflammatory bronchial disease is more frequent than bacterial pneumonia. A study tells that in young dogs, viral infection is followed by bacterial attack. Hence, the root cause of pneumonia has to be identified for the treatment followed by antibiotic therapy and prevention from air secretions.

D Wilmes et al. [4] proposed that people undergoing immunosuppressive therapy after kidney transplant are more prone to bacterial pneumonia. The author has described the most common pneumonia causing pathogen after kidney transplant, their clinical presentation, radiological features and treatments. This article suggests that recognizing the involved pathogen is most important to initiate any antibiotic therapy and to calculate duration of treatment.

O Henig et al. [5] proposed that older adults are more prone to pneumonia due to resistant organisms including Gram-negative bacilli. Various associated symptoms include confusion, clinical deterioration, new onset of recurrent fall etc. The author suggests treatment of pneumonia to be based on risk factors and multidrug resistant organisms common in older adults.

JL Wong et al. [6] proposed that patients with cancer have disproportionate morbidity and mortality. Cancer related disfunctions of lungs, mucositis all leads to bacterial pneumonia risks. Patients suspected of bacterial pneumonia are suggested to take early antibiotic therapy to cover pathogens encountered in health care setting.

D Das et al. [7] proposed that chest X-ray manifestations of pneumonia are important as in they guide the patients guide the patients to undergo appropriate therapy and anticipate for complications. CXR patterns can help the readers to find out different causes. Frequent CXRs should be taken during the treatment as the persistence or progression of the initial CXR can indicate the progress or failure of treatment.

AR Falsey et al. [8] proposed that most frequent victims of viral pneumonia are extreme age groups i.e., children and elders. In most of the cases the initial attack by the respiratory virus results in pneumonia in children who don't have immunity and the airways are not developed. Influenza virus is found to be main cause of morbidity in older adults. Respiratory viruses show different clinical syndromes within themselves and other bacterial pathogens.

S Nayak et al. [9] proposed that broad spectrum antibiotics are used to treat pneumonia because it is difficult to diagnose the pneumonia causing organism. The solution is to target the drug to a protein that is common to bacteria species. It can be used before identification of pneumonia causing organism as soon as symptoms are identified. The author discussed computational subtractive genomics approach which facilitates identification of putative drug targets.

SB Erdem et al. [10] proposed that atopy presence has an effect on young children within age 1 to 6 years suffering from viral pneumonia. Atopy is referred as positive skin prick tests to one or more allergens. The results of this study revealed that moderate pneumonia patients were compared with severe pneumonia patients, the frequency of atopy was not different. Hence, there was no effect of atopy on severity of pneumonia on children.

JP Amorin et al. [11] proposed that patients undergoing liver transplantation can experience morbidity and mortality due to postoperative pulmonary complication such as bacterial pneumonia. The author has conducted a study aiming to find the incidence, involved microbes and factors of bacterial pneumonia within a duration of 100 days after the liver transplant.

KA Radigan et al. [12] proposed that most of the community acquired pneumonias are caused by viruses. Infection in the respiratory tract are causing more than any number of symptomatic diseases and leave from work. This is resulted due to lack of antiviral treatment and poor diagnostic tests. With the increase in population of low immunity elderly population viruses contribute majorly in lower respiratory tract infections like bronchitis and viral pneumonia.

S Jain [13] proposed that respiratory viruses have more substantial contribution to pneumonia in elders and children even more than bacteria. Adenoviruses, coronaviruses, human rhinoviruses, influenza viruses are of the most commonly detected ones. While it is difficult to differentiate between viruses on their clinical effects, there is an increase in evidence of their role in pneumonia.

SN Grief et al. [14] proposed that the common symptoms considered for pneumonia include acute onset fever and cough, fatigue, anorexia and chest pain. Treatment of pneumonia should be aided by local resistance patterns. For previously healthy patients first-line treatment include macrolide antibiotic as azithromycin. Pediatric pneumonia is also quite common and first-line treatment include amoxicillin followed by macrolides.

C Feldman et al. [15] proposed that HIV patients are mostly affected by community acquired bacterial pneumonia (CAP). Changes in immunity system due to HIV led to increase in risk associated with pneumonia. When non-HIV infected patients of CAP are compared with infected ones there was no difference in clinical features. Prevention of CAP requires cigarette smoking cessation strategies, ART adherence etc.

M Kim et al. [16] investigated that whether prone Computer Tomography (CT) adds any value to the process of identifying honeycombing and classifying Usual Interstitial Pneumonia (UIP) patterns in terms of interobserver agreement and accuracy while taking pathological results as standard results. Different observers reviewed supine only and supine and prone combined tests of different patients to determine the presence of honeycombing and UIP classification. It was found that reviewing prone with supine improves the interobserver agreement of UIP classification for less experienced radiologists. There was no difference in the results for honeycombing.

Z Yunt et al. [17] proposed that interstitial pneumonia is a common manifestation of rheumatoid arthritis. It is associated with significant mortality. There is mot much existing data for the prediction of mortality. The authors seek to examine the prognostic value of computed tomography in rheumatoid arthritis patients. They took 158 RA-ILD patients with HRCT patterns of pneumonia. 100 of the subjects had definite usual interstitial pneumonia (UIP), 23 with possible UIP and 35 with NSIP. The results indicated that patients who were either definitely or possibly diagnosed with usual interstitial pneumonia have survival as poor as that of a patient with nonspecific interstitial pneumonia (NSIP).

I Sirazitdinov et al. [18] investigates the application of machine learning solutions to automatically detect the presence and positioning of pneumonia on chest x ray images. Here RetinaNet and Mask R-CNN are combined. Pneumonia is manifested on a relatively small region of a chest x-ray, which represents a challenge for modern object detectors. To tackle this problem, we used the FPN principle in the backbone of both models, since FPN generates multi-sczle feature maps with better quality information. The authors identified several future approaches. First that comprehensive diagnoses should be based on manifestation on specific clinical symptoms. Second, lateral and frontal chest x-ray should be augmented to obtain a detailed image. Third that meta information can be important in future diagnoses.

DS Kermany et al. [19] proposed a deep-learning framework for the screening of patients with blinding retinal diseases as clinical decision-making algorithm faces reliability challenges. The framework trains a neural network with fraction of data. On applying this approach to tomography images, the performance comparable to human experts can be achieved in classifying pneumonia. The regions identified by the neural network are also highlighted by the proposed tool. It is claimed to be useful in expediting the diagnosis and referral of these diseases. This tool can be useful in detecting pneumonia as well as distinguishing between bacterial and viral pneumonia.

EJ Yates et al. [20] aimed to develop a machine learning based model for the classification between chest radiography abnormalities as a tool to report prioritization. The model is based on Inception version 3 CNN and was trained on almost 50000 anonymous chest radiographs. The final layer of the deep neural network was developed using Tensorflow in order to classify between two anonymized image datasets. This study demonstrates the application of a machine learning-based approach to classify chest radiographs as normal or abnormal. Its applicationto real-world datasets may be warranted in optimizing clinician workload.

Orhan Er et al. [21] proposed that chest disease is very common in the today's world. A comparative study was performed on pneumonia disease diagnosis was done with neural network. According to them neural network is the structure of multilayer, probabilistic and learning vector quantization. Dataset was created using hospital's database using patient's epicrisis reports. They use the momentum training algorithm. In their study they got the better accuracy from the previous studies.

Gregory F. Cooper et al. [22] proposed that eight methods to derive computer models for predicting condition of patients with pneumonia. The primary metric was the error in the prediction to survival of the patient. The model is useful to decide whether the patient is required to admit in the hospital or can be treated from home. They studied in the population in which approximately 11% of the patients died of pneumonia, as a result shows that when 10% of the patient population is predicted to survive.

Hans-Ulnich Kauczor et al. [23] proposed a comparative study of multiple neural network with a density mask for the automatic detection and quantification of ground – glass opacities on high – resolution CT under clinical conditions. The density mask identified another 17.3% of the total lung area to be ground-glass opacities that were not detected by the neural network. Opacities on CT by a NN are accurate to be implemented for the interpretation of images in a clinical environment.

Orhan Er. et al. [24] proposed that chest disease are very important health problems in the world. In this paper a comparative diagnosis was calculated by using multilayer, probabilistic, learning vector quantization, and generalized regression neural networks. The chest diseases dataset was prepared by using patient's epicrisis reports from a chest diseases hospital's database. Due to different datasets used by the studies, the overall comparison of the results is impossible and irregular. LM training algorithm is better than BP with momentum training algorithm for MLNN structures for the diagnosis of chest disease.

Arati Gurung et al. [25] proposed that we sought to perform a systematic review and meta- analysis of studies implementing lung sound analysis (CLSA) to aid in the detection of abnormal lung sounds for specific respiratory disorders. They conducted analysis to calculate the sensitivity and specificity of CLSA for the detection of abnormal lung sounds. Most studies employed either electret microphones or piezoelectric sensors for auscultation and Fourier Transform and Neural Network algorithms for analysis. As the conclusion further research and product development could promote the value of CLSA in research studies or its diagnostic utility in clinical settings.

Amit Kumar Jaiswal et al. [26] proposed that pneumonia is one of the leading causes of death among children and old age people around the world and X-ray image analysis is considered as tedious and crucial tasks for radiology experts. In the paper they proposed their deep learning- based approach for the identification and localization of pneumonia in Chest X-rays (CXRs) images. Several factors such as positioning of the patient and depth of inspiration can change the appearance of the chest X-ray, complicating interpretation further. Their approach gives the outcomes through critical modifications of the training process and a step which merges bounding boxes from multiple models.

Feng Li et al. [27] proposed to evaluate radiologists' ability to detect focal pneumonia by use of chest radiographs and compared with plus bone-suppressed chest radiographs. A bone suppression image processing system was applied to the 56 radiographs to create corresponding bone suppression images. Receiver operating characteristic (ROC) analysis was used to evaluate the observers' performance. Usage of bone suppression images improved radiologists' performance for detection of focal pneumonia on chest radiographs.

Paul S. Heckerling et al. [28] proposed that Artificial neural networks (ANN) is used in the prediction of most of medical treatments but is not used to predict pneumonia. They trained the model on 1044 patients from the University of Illionis and was tested on 116 patients from the University of Nebrasaka. The method used by the authors was Feed-forward back propagation. ANN is trained using other strategies discriminated equally in the 2 cohorts but no better than did logistic regression. ANN improved on the accuracy of logistic regression.

Diana Carolina Moncada et al. [29] proposed that pneumonia is very common infection and a frequent cause of medical treatment, hospitalization and death around the world. Clinical treatment is basically based on the infiltrates in the chest X-ray. In the study authors targeted to find the interobserver and intraobserver agreement using two reading formats in patients with community-acquired pneumonia, and to explore their association with etiology and clinical outcomes. As the result, the observers made a joint reading to explore its prognostic usefulness via multivariate analysis and in patients with pneumonia, the interpretation of the chest X-ray, especially the smallest of details, depends solely on the reader.

Stacey J. Ackerman et al. [30] proposed that increasing pressures to improve the quality of care and enhance revenue while containing health care costs have prompted radiology departments to explore promising new technologies and services. The most important clinical acceptability for this technology is the ability to achieve a good accuracy when interpreting radiologic examinations from a digital-display workstation. Eight radiologists participated as image interpreter. As the result, variations in an observer's confidence threshold will result in movement along the same ROC curve, whereas a change in performance will result in a different ROC curve.

P.S. Heckerling et al. [31] proposed that Artificial neural networks (ANN) are those computer algorithms that can extract patterns from input variables, and use them to predict results. They used the genetic algorithm with selection based on maximizing network accuracy and minimizing network input-layer cardinality. The genetic algorithm evolved best 10-variable sets that discriminated pneumonia in the training cohort. Variable sets derived using a genetic algorithm for neural networks accurately discriminated pneumonia from other respiratory conditions, and did so with greater accuracy than variables derived using stepwise neural networks or logistic regression in some cases.

Wolfgang G. Kunz et al. [32] proposed that basal lung opacities are frequently observed on supine chest x – ray (SXCR) of intensive care patients, causing insecurity among clinicians and radiologists. We sought to determine the diagnostic accuracy of SCXR for basal pneumonia.

Sensitivity, specificity, and positive and negative predictive values (PPV/NPV) were calculated once pooling 0 and 1 as negative and once pooling 1 and 2 as positive finding. Interpreting only highly suspicious basal opacities as pneumonia considerably increases the PPV with almost constant NPV. Clinicians and radiologists should be aware of the limitations of SCXR regarding basal pneumonia.

Paul S. Heckerling et al. [34] proposed that optimized problems of artificial neural network (ANN) could be solved by Genetic Algorithms. We used genetic algorithms to search for optimal hidden layer architectures, connectivity, and training parameters for ANN for predicting community acquired pneumonia among patients with respiratory complaints. Predictive accuracy was measured as the area under a receiver operating characteristic (ROC) curve. Algorithms based on cross-generational selection, Gray coding of genes prior to mutation, and crossover recombination at different genetic levels, evolved optimized ANN identical to the baseline genetic strategy. Algorithms based on other strategies, including elite selection within generations and inversions of genetic material during recombination evolved less accurate ANN.

John R. Zech et al. [35] proposed that CNNs are widely use in today's world to analyze medical imaging to provide computer-aided (CAD). In 3 out of 5 natural comparisons, performance on chest X-rays from outside hospitals was significantly lower than on held-out X-rays from the original hospital system. Estimates of CNN performance based on test data from hospital systems used for model training may overstate their likely real-world performance.

LITERATURE SURVEY

Sl. No.	AUTHOR	TECHNIQUE	DATA SET	METRICS
1.	• JC Souza • JOB Diniz • JL Ferreira • GLF da Silva • AC Silva • AC de Paiva	• AlexNet based CNN • ResNet-18-based CNN • Chest X-ray • Image segmentation • Image Reconstruction • Keras Algorithm • Tensorflow	Montgomery county dataset is used which is a public chest x-ray dataset from department of health and human services of Montgomery County, Maryland. 138 CXRs out of which 80 are normal and 58 infected.	• Segmentation • Sensitivity • Specificity • Accuracy • Dice • Jaccard
2.	• HJ Zar	• Lung Ultrasound • Chest X-ray • PCR Analysis	The dataset consists of 1140 mother-child pairs with completion of all 1year follow-up visits and high cohort retention.	• Environmental factors • Infectious factors • Nutritional factors • Genetic factors • Psychological factors • Maternal factors • Immunological factors
3.	• JD Dear	• Thoracic Radiography • CT Scan • Hematology • Genetic factors	The research was conducted on an 8-year-old female spayed Chihuahua mix presented for wet cough and a 5-year-old male catered domestic medium hair cat presented for evaluation of acute respiratory distress.	• Temperature • Pulse • Respiratory rate • White blood cell count • Cellularity
4.	• D Wilmes • E Coche • H Rodriguez-Villalobos • N Kanaan	• Chest X-ray • Lung CT-scan • Spectrometry	Chest X-ray of a 75-year-old man admitted for 17 years after kidney transplantation and a 54-year-old man transplanted with a kidney one year ago presented with cough, fatigue and anorexia.	• Morbidity • Mortality • Bacterial behavior
5.	• O Henig • KS Kaye	• Management of Community-acquired Pneumonia • Outcomes of Pneumonia Among Residents of Long-term Care Facilities	Epidemiology of bacterial pathogens causing CAP among elderly patients and residents of LTCF, and risk factors for each organism.	• Range of Prevalence

6.	• JL Wong • SE Evans	• Radiographic images • Computed tomography • Microbiologic techniques • Molecular Diagnostics • Antibiotic Therapy • Non-bronchoscopic Diagnostics	Radiographic presentations of bacterial pneumonia in patients with cancer. Computed tomography images of patients with cancer with documented bacterial pneumonias	Presence of • Gram-positive bacteria • Gram-negative bacteria • Atypical Bacteria

7.	• D Das • DC Howlett	• Chest X-ray	Chest X-ray patterns, including lobar pneumonia, bronchopneumonia, nodular consolidation, interstitial consolidation, atypical pneumonia, and lung abscesses	• Lung abscess • Nodular consolidation • Nodular opacification
8.	• AR Falsey	• Chest Radiography • Reverse transcription polymerase chain reaction (RT-PCR)	• Chest radiograph of a 63-year-old woman who had mitral stenosis • Chest radiograph of a 75-year-old with fever, cough, and dyspnea • Chest radiograph of a 68-year-old man with history of emphysema.	• Opacity in chest radiography images • White blood cell count
9.	• S Nayak • D Pradhan • H Singh • MS Reddy	• Computational screening of potential drug targets • Comparative mining of putative drug targets • Chokepoint enzyme • PPI network analysis and Identification of hub genes	Representation of steps involved in comparative target identification of 13 bacterial pathogens. Identified targets of *Streptococcus pneumoniae TIGR4* were used to develop drugs or vaccines.	• Drug Targets Metabolic pathways • Protein-Protein Interaction • Chokepoint enzymes
10.	• SB Erdem • D Can • S Girit • F Catal	• Virus detection and identification • Skin Prick test • Statistical	A total of 280 patients from nine centers were included in the study. Of these patients, 163	• Atopy • Age • Body mass index • Hospitalization Period

	• V Sen • S Pekcan • R Ersu	Analysis • Ethical Approval	(58.2%) were male. RSV (29.7%), Influenza (20.5%), and many more were isolated from respiratory samples.	• Fever • SaO2
11.	• JP Amorin • M Lopez • K Rando • J Castelli • JM Presentado	• Multi-resistance profiling • Extreme resistance profiling	• Overview of the Population Having Received a Liver Transplant (n = 106)	• Multi-resistance profile • Extreme resistance profile

			• Form of Presentation, Diagnosis, and Microorganism Isolated (n = 9)	• Morbidity • Mortality
12.	• JC Souza • JOB Diniz • JL Ferreira • GLF da Silva • AC Silva • AC de Paiva	• AlexNet based CNN • ResNet-18-based CNN • Chest X-ray • Keras Algorithm • Tensorflow	Montgomery county dataset is used which is a public chest x-ray dataset from department of health and human services of Montgomery County, Maryland. 138 CXRs out of which 80 are normal and 58 infected.	• Segmentation • Sensitivity • Specificity • Accuracy • Dice • Jaccard
13.	• S Jain	• Chest Radiography • RT-PCR • Chest X-ray	The proportions of viral, viral-viral, bacterial-viral, bacterial, fungal or mycobacterial pathogens detected, and no pathogen detected.	• Influenza A • Rhinovirus • Adenovirus • Human metapneumovirus • Parainfluenza
14.	• SN Grief • JK Loza	• PCR ((polymerase chain reaction) • X-ray • Radiography	A 2015 prospective, multi-center study by the Centers for Disease Control and Prevention identified a responsible pathogen in only 38% of cases of community-acquired pneumonia in adults requiring hospitalization.	Risk Factors • Agricultural animals • AIDS • Alcoholism • Intravenous Drug use • Pulmonary abscess
15.	• C Feldman • R Anderson	• Clinical Presentation • Laboratory	An etiologic diagnosis is obtained in some 35% to 75% of HIV-	• CD4 cell count • Community-acquired bacterial pneumonia

		diagnosis of bacterial pneumonia • Treatment of bacterial CAP in HIV-infected persons	infected patients with bacterial CAP	• ART • HIV infection • Mortality • Smoking • Vaccination
16.	• M Kim • SM Lee • JW Song • KH Do • HJ Lee • S Lim • JB Seo	• Computed Tomography	HRCTs of 86 patients with pathologically proven UIP, NSIP and chronic HP between January 2011 and April 2015 evaluated by 8 observers .	• Honeycombing Identification • Interobserver agreement • Accuracy of UIP pattern

17.	• ZX Yunt • JH Chung • S Hobbs • ER Fernandez-Perez • AL Olson • TJ Huie • KK Brown	• High Resolution Computed Tomography	A total 158 rheumatoid arthritis with HRCT patterns of usual interstitial pneumonia or nonspecific interstitial were identified among individuals at National Jewish Health.	• Percent patients with definite UIP • Percent patients possible UIP • Percent patients with NSIP
18.	• I Sirazitdinov • M Kholiavchenko • T Mustafaev • Y Yixuan • R Kuleev • B Ibragimov	• Chest X ray • RetinaNet • Mask R-CNN	The solution was validated with a dataset of 26,684 images from Kaggle Pneumonia detection challenge.	• Precision • Recall • F1-score
19.	• DS Kermany • M Goldbaum • W Cai • CC Valentim • H Liang • SL Baxter • J Dong	• Computed Tomography • Deep Neural Networks • Chest X-ray	5232 chest x ray images from children with 3883 depicting pneumonia (2538 bacterial and 1345 viral)	• Accuracy • Specificity • Cross-Entropy Loss • Sensitivity
20.	• EJ Yates • LC Yates • H Harvey	• Chest Radiography • Convolutional Neural Network • Confusion Matrix Analysis	Re-training was performed on 47,644 images using commodity hardware, with validation testing on 5,505 previously unseen radiographs.	• Positive Predictive Value • Sensitivity • Specificity • Area under the curve
21.	• Orhan Er • Cengiz Sertkaya • Feyzullah Temurtas • A. Cetin Tanrikulu	• Radiography shape features for diagnosis. • BP is used is training algorithm.	Dataset prepared using epicrisis reports of 201 patients of a hospital of Turkey contains three classes	• MLNN with BP (one hidden layer) • MLNN with BPwM (two hidden layers)

22.	• Gregory F. Coopera • Constantin F. Aliferis • Richard and many more.	• Logistic Regression. • K – nearest method.	Training dataset of 9847 cases was prepared from MedisGroups Comparative Hospital Database (MCHD)	• Predictive Accuracy of several hybrid models
23.	• Hans-Ulrich • Kjell Heitmann • Claus Peter • Dirk Marwede • Uthmann • Manfred Thelen	• Contrast richness was detected as a more local and a more mid-range feature to segment. • The lungs were segmented and lung area affected was calculated.	Dataset of 84 patients - 54 men and 30 women; total of 99 consecutive CT scans was used.	• Image segmentation • Image development • Testing of the neural network

24.	• Orhan Er • Nejat Yumusak • Feyzullah Temurtas	• LM training algorithm. • BP with momentum training algorithm.	The dataset consists of the chest disease measurements contains six classes with 357 samples.	• Multilayer, probabilistic • Learning vector quantization • Regression neural network.
25.	• Arati Gurung • Carolyn G. Scrafford • James M. Tielsch • Orin S. Levine • William Checkley	• Fourier Transform Concept of Differential and Difference Equation. • Neural Network algorithms for analysis.	The dataset used in the study was actually the eight different studies of the previous authors from different parts of the world with more than 462 patients.	• Acute lower respiratory infections • Lung auscultation • Electronic auscultation • Acoustic signal Processing • Computerized lung sound Analysis • Crackle detection
26.	• Amit Jaiswal • Prayag Tiwari • Sachin Kumar • Deepak Gupta • Ashish Khanna • Joel J.P.C. Rodrigues	• Mask RCNN algorithm • ResNet101 algorithm • ROIAlign algorithm	Chest radiographs dataset from RSNA7 which annotated 30,000 exams from the original 112,000 chest X-ray dataset to identify instances of potential pneumonia as a training set.	• Mask RCNN • Pixel-wise Segmentation • RatinaNet Model
27.	• Feng Li • Roger Engelmann • Lorenzo Pesce • Samuel • Armato III • Heber MacMahon	• Computed Tomography (CT) • Bone Suppression Imaging (BSI) processing • Computed Radiography System	Dataset included 146 patients with infectious lung diseases on CXRs that were obtained from a radiology department of University of Chicago Medical Center	• Standard chest radiographs. • Focal airspace opacities. • ROC curve • Image processing

28.	● Paul S. Heckerling ● MD, Ben S. ● Gerber, MD ● Thomas G ● Robert S. Wigton	● Feed-forward back-propagation. ● Logistic regression. ● Stopped-training algorithm to prevent network overfitting was implemented.	Dataset is of 1044 patients from the University of Illinois (the training cohort) and were applied to 116 patients from the University of Nebraska	● Sociodemographic ● Symptom ● Sign ● Comorbidity ● Radiographic
29.	● Diana Carolina Moncada ● Zulma Vanessa Rueda ● Antonio Macías ● Tatiana Suárez ● Héctor Ortega ● Lázaro Agustín Vélez	● Kappa statistic and intraclass correlation coefficient). ● Quantitative analysis – pulmonary damage	A total of 211 patients with CAP who had chest X-rays available in a digital format were evaluated; 74 of those were excluded because the readers considered the chest X-rays of	● Interobserver Agreement ● Intraobserver Agreement ● Radiography ● Statistical Analysis

		categorized from 0 to 10.	poor quality for this study.	
30.	● Stacey J. Ackerman ● Joseph N. Gitlin ● Robert W. Gayler ● Charles D. Flagle ● R. Nicholas Bryan	● Null Hypothesis ● Consensus Panel ● Image Acquisition and Display ● Data Collection and Management ● ROC Analysis ● Reporting Protocol ● Accuracy Sensitivity	Dataset contained 160 examinations which were transmitted 5 miles over a dedicated telephone line at a T1 rate from Francis Scott Key Hospital to Johns Hopkins Hospital.	● Acceptable accuracy ● Specificity Measurements ● Variations in an observer's confidence threshold ● Analog and digital reading conditions.
31.	● P.S.Heckerling ● B.S. Gerber ● T.G.Tape ● R.S. Wigton	● Data Collection ● Genetic Algorithm ● Neural Networks ● Stepwise Neural Networks Logistic Regression ● Statistical Analysis	Dataset contained total of 1325 patients (1131 from the University of Illinois and 194 from the University of Nebraska) were enrolled in the original data collection.	● Partial Retraining ● Weight Elimination ● Linear Substitution ● Data Permutation
32.	● Wolfgang G. Kunz ● Maximilian Patzig ● Alexander ● Robert Stahl ● Maximilian F.	● SCXR examination ● Supine chest x-ray readings	Dataset contained 172 patients were identified and included in the study. "Possible pneumonia" was present in 57 cases on the right side and in 66	● Patient Population ● Image Analysis ● Statistical Analysis

	● Mike Notohamiprodjo,		cases on the left side. Others 92 were highly infected.	
33.	● Mark Cicero ● Alexander Bilbily ● Errol Colak ● Tim Dowdell ● Bruce Gray ● Kuhan Perampalada ● Joseph Barfett	● Python Programming Foundation) ● DICOM metadata for the acquisition technique. ● Inception architecture	Dataset included a total of 61,192 studies which were identified by the classification script and were exported from the PACS repository	● Study Population ● Neural Network Architecture ● Statistical Analysis ● Pulmonary edem ● Pleural effusion
34.	● Paul S. Heckerling ● Ben S. Gerber ● Thomas G. Tape ● Robert S. Wigton	● Genetic algorithms to search for optimal hiddenlayer architectures. ● Training parameters for ANN for predicting community acquired pneumonia among patients	A total of 1325 patients (1131 from the University of Illinois and 194 from the University of Nebraska) were enrolled in the original data collection.	● Data collection ● Neural networks ● ANN were feed-forward, back-propagation network.
35.	● John R. Zech ● Marcus A. ● Manway Liu ● Anthony B. ● Joseph J. Titano ● Eric Karl Oermann	● PyTorch 0.2.0 and torchvision were used for model training, ● All models were trained using a cross-entropy loss function with parameter update by stochastic gradient.	Three datasets were obtained from different hospital groups: NIH (112,120 radiographs from 1992 to 2015), Indiana University Network for Patient Care (IU; 7,470 radiographs, date range not available), and Mount Sinai Hospital (MSH; 48,915 radiographs from 2009 to 2016)	● Datasets ● Preprocessing: Frontal view filtering ● Preprocessing: Generating labels for pathology ● CNNs ● Internal and external performance testing ● Preprocessing: Identifying MSH portable scans ● Model training

Basic Architecture of Inception-v3

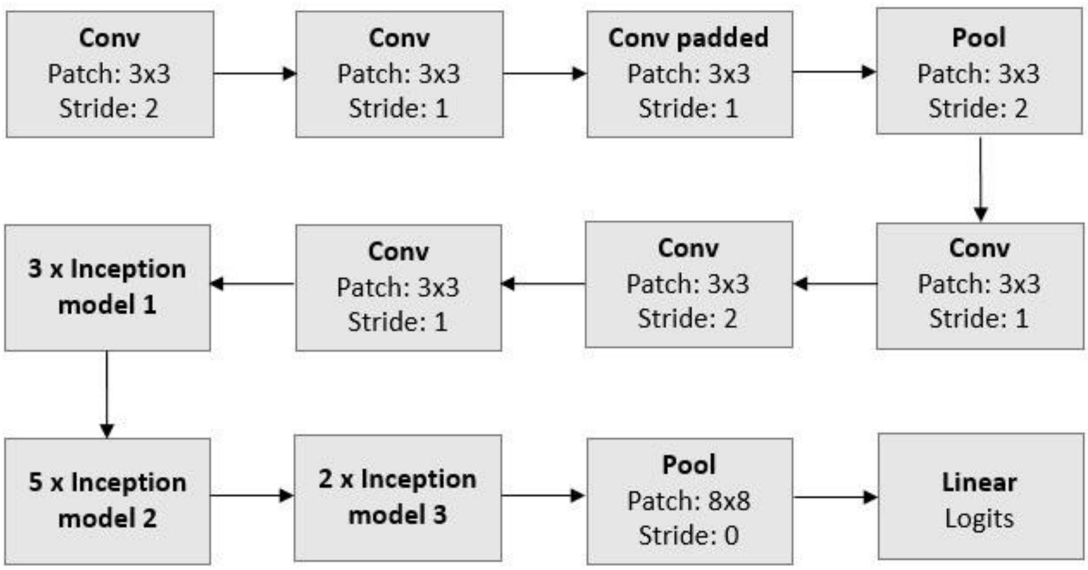
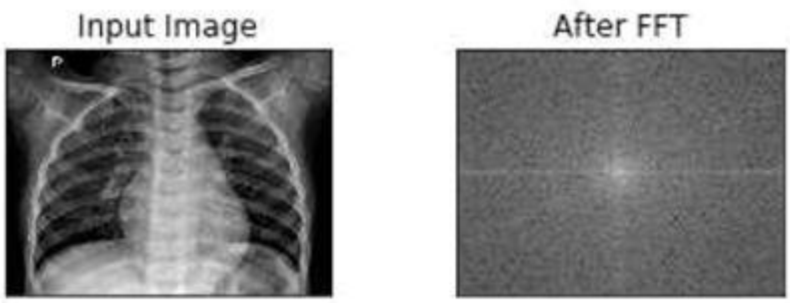


Image Enhancement in the Frequency Domain

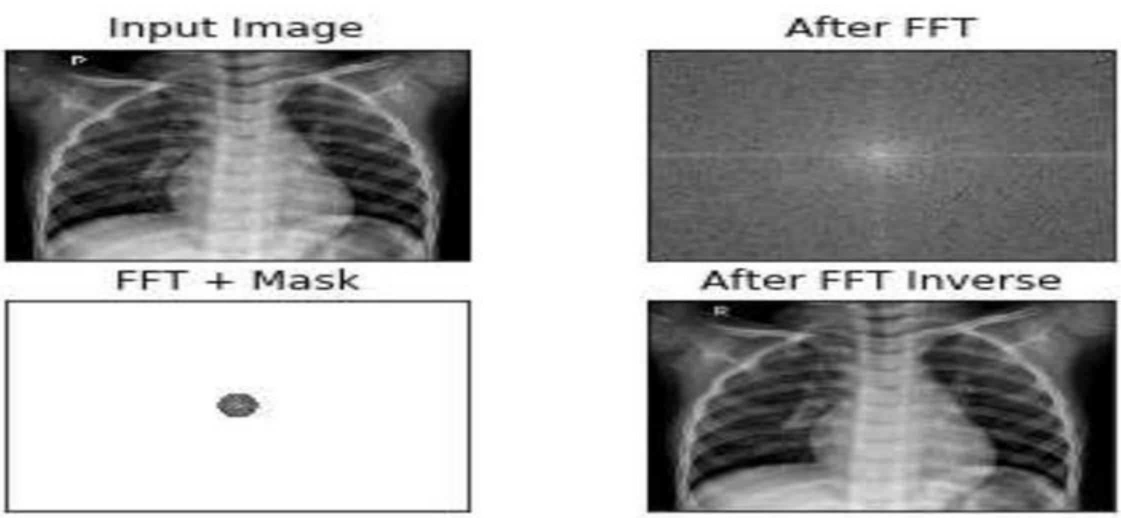
Fourier Transform

It converts the input image from spatial domain to frequency domain. It plots image as a set of high and low frequencies with low frequencies situated towards centre while high frequencies scattered around.

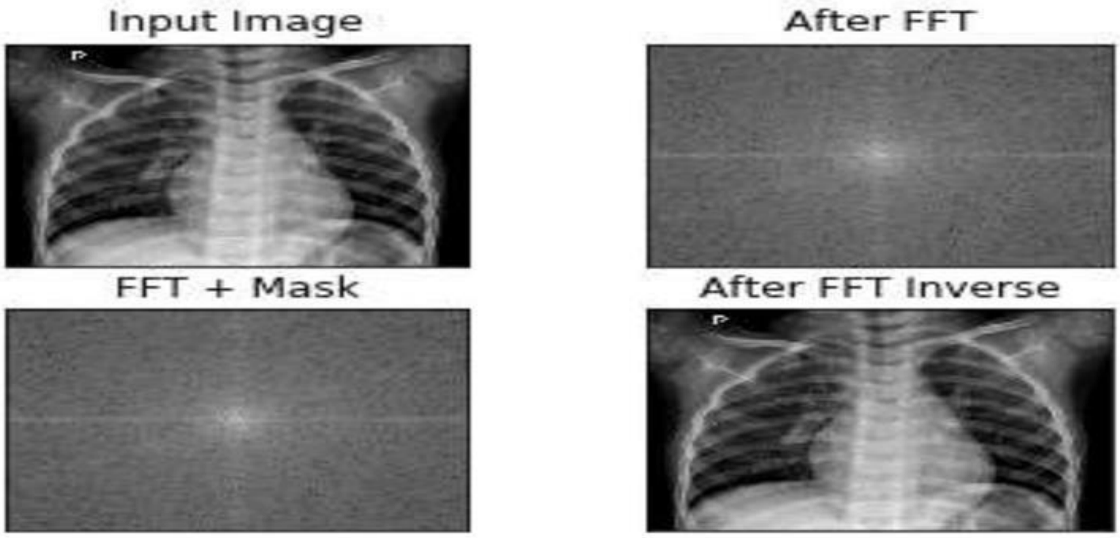


The Fourier transform of the image can be further processed for application like edge detection using high pass or band pass filter, noise reduction and blurring using a low pass filter.

Applying High pass filter



Applying Low pass filter



Applying Band pass filter

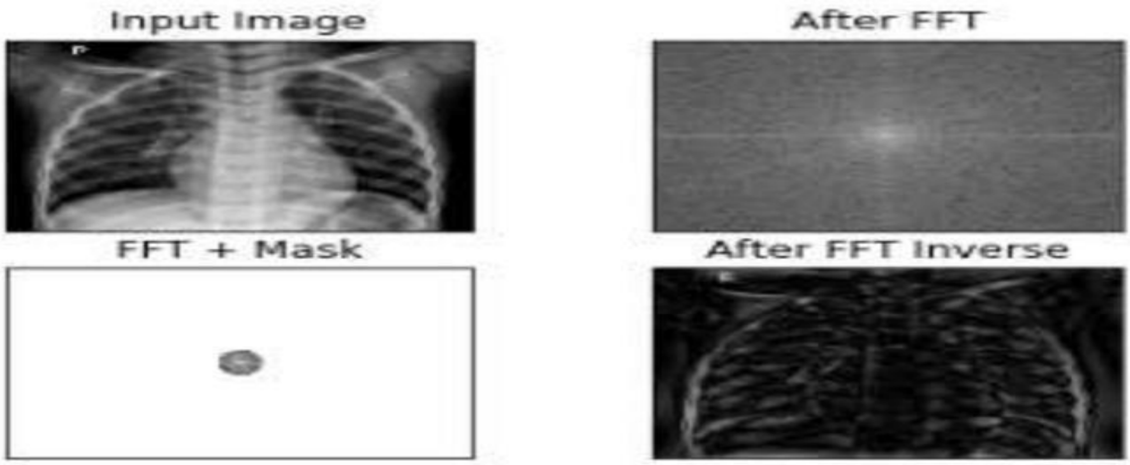
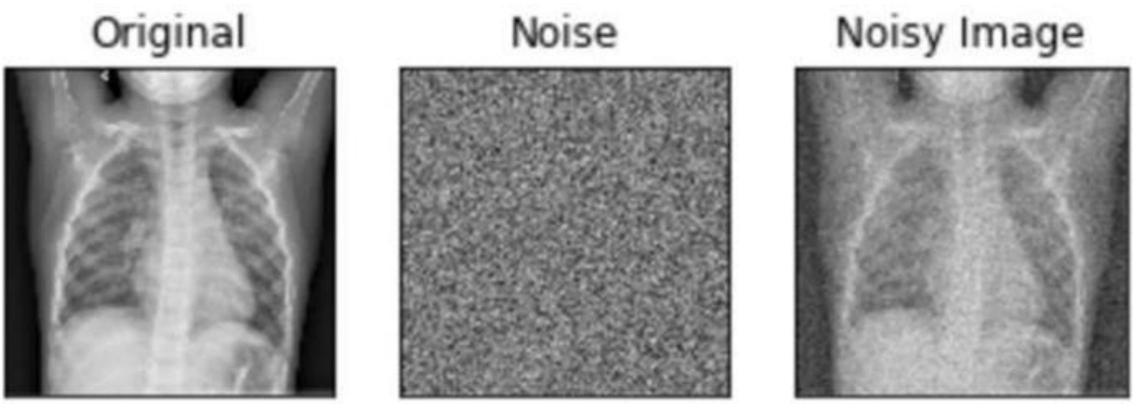


Image Restoration using morphological process

Introducing noise to image

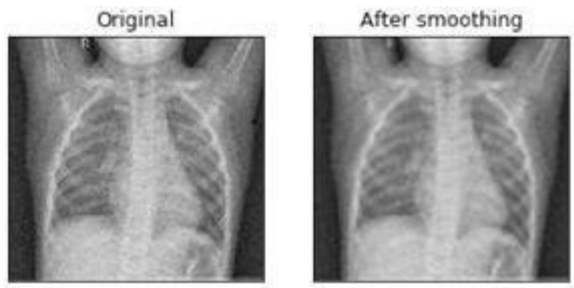
Here we are adding gaussian noise to the input image



To reduce the noise introduced in the image various filters are tried.

2D Convolution

Averaging Filter

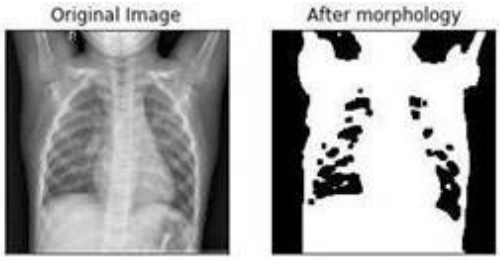


Median Filter

Image Segmentation

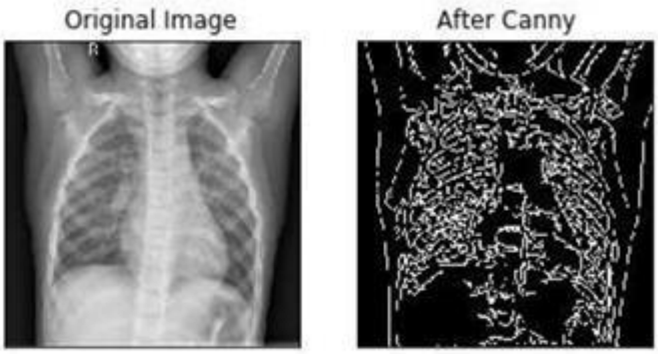
Image Thresholding

Morphological Process (for noisy and original)

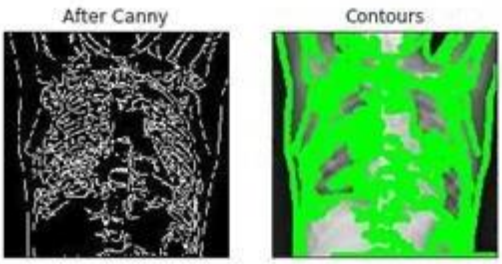


Canny algorithm

Here we have used canny algorithm of the OpenCV library in python for the detection of edges.



Contours



The next block of code is written to draw the image contours using drawContours function of the OpenCV library

Methodology

In this project we are going to visualize and develop a Convolution Neural Network (CNN) to detect pneumonia symptoms in the patient and further differentiate between bacterial pneumonia and viral pneumonia.

Recently, CNN-motivated deep learning algorithms have become the standard choice for medical image classifications although the state-of-the-art CNN-based classification techniques pose similar fixated network architectures of the trial-and-error system which have been their designing principle.

Here is a demonstration of how a doctor classifies between chest x-ray of normal patient with that of patient diagnosed with pneumonia. The left image depicts normal lungs with no areas of opacification in the image. The middle image depicts bacterial pneumonia exhibiting a focal lobar consolidation in the right upper lobe (shown by the arrows). The right image depicts viral pneumonia with a more diffuse pattern in both lungs.

How is pneumonia diagnosed and evaluated?

Your primary doctor will begin by asking you about your medical history and symptoms. You will also undergo a physical exam, so that your doctor can listen to your lungs. In checking for pneumonia, your doctor will listen for abnormal sounds like crackling, rumbling or wheezing. If your doctor thinks you may have pneumonia, an imaging test may be performed to confirm the diagnosis.

One or more of the following tests may be ordered to evaluate for pneumonia:

Chest x-ray: An x-ray exam will allow your doctor to see your lungs, heart and blood vessels to help determine if you have pneumonia. When interpreting the x-ray, the radiologist will look for white spots in the lungs (called infiltrates) that identify an infection. This exam will also help determine if you have any complications related to pneumonia such as abscesses or pleural effusions (fluid surrounding the lungs).

Dataset used

https://kaggle.com/paultimothymooney/chest-xray-pneumonia

This dataset consists of 5,863 jpeg images of chest X-ray divided into two categories – Pneumonia and normal. These images were selected from retrospective cohorts of pediatric patients of age one to five years from Guangzhou Women and Children's Medical Center, Guangzhou. All chest x-rays were imaged as a part of routine health care.

For the analysis of chest x-ray images, all chest radiographs were initially screened for quality control by removing all low quality or unreadable scans. The diagnoses for the images were then graded by two expert physicians before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert.

Chest X-Ray Dataset

Type	Test	Train
Normal	234	1341
Pneumonia	390	3875

Table 1. Number of images under test and train files

Expected Result

To find out how accurately the applied algorithm can distinguish between patients diagnosed with pneumonia or not on the basis of the chest x rays.

ALGORITHM USED

Transfer Learning

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task. It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

```
In [5]: print(os.listdir("chest_xray"))
        ['.DS_Store', 'test', 'train', 'val']

In [6]: print(os.listdir("chest_xray/train"))
```

Goals

1. To classify images without making any custom model but rather using keras built in models.
2. To get high accuracy for classifying almost any image.

Advantages of Transfer Learning

1. Super simple to incorporate.
2. Achieve same or even better (depending on the dataset) model performance quickly.
3. There's not as much labeled data required.
4. Versatile uses cases from transfer learning, prediction, and feature extraction.

Implementation

Importing libraries

```
In [4]: import pandas as pd
import cv2
import numpy as np
import os
from random import shuffle
from tqdm import tqdm
import scipy
import skimage
from skimage.transform import resize
```

Locating the dataset directories

A function to label the images in the dataset based upon whether the person is normal or diagnosed with pneumonia

Normal patient has been labeled 0

Pneumonia diagnosed patient has been labeled 1 Testing Accuracy with different Filters

```
In [8]: def get_label(Dir):
            for nextdir in os.listdir(Dir):
                if not nextdir.startswith('.'):
                    if nextdir in ['NORMAL']:
                        label = 0
                    elif nextdir in ['PNEUMONIA']:
                        label = 1
                    else:
                        label = 2
            return nextdir, label
```

A function which takes the dataset directory as an argument written to perform preprocessing on the image dataset which includes

- Image grey-scaling
- Image resizing
- Storing the greyscale values and labels in x and y variables

Combining the above two functions to write the function get_data with directory as an argument returning the x and y variables containing arrays of image greyscales and labels respectively.

```
In [10]:
def get_data(Dir):
    X = []
    y = []
    for nextDir in os.listdir(Dir):
        if not nextDir.startswith('.'):
            if nextDir in ['NORMAL']:
                label = 0
            elif nextDir in ['PNEUMONIA']:
                label = 1
            else:
                label = 2

            temp = Dir + nextDir

            for file in tqdm(os.listdir(temp)):
                img = cv2.imread(temp + '/' + file)
                if img is not None:
                    img = skimage.transform.resize(img, (150, 150, 3))
                    #img_file = scipy.misc.imresize(arr=img_file, size=(299, 299, 3))
                    img = np.asarray(img)
                    X.append(img)
                    y.append(label)

    X = np.asarray(X)
    y = np.asarray(y)
    return X,y
```

Applying the get_data function to train and test directories to get (X_train, y_train) and (X_test, y_test) pairs respectively.

```
In [11]: X_train, y_train = get_data(TRAIN_DIR)

100%|██████████| 3876/3876 [07:54<00:00,  8.18it/s]
100%|██████████| 1342/1342 [09:48<00:00,  2.85it/s]
```

```
In [12]: X_test , y_test = get_data(TEST_DIR)

100%|██████████| 390/390 [00:36<00:00, 12.23it/s]
100%|██████████| 234/234 [01:41<00:00,  2.08it/s]
```

```
In [37]: print(X_train.shape,'\n',X_test.shape)

(5216, 3, 150, 150)
 (624, 3, 150, 150)
```

```
In [57]: print(y_train.shape,'\n',y_test.shape)

(5216, 2, 2)
 (624, 2, 2)
```

Importing to_categorical function from keras library which returns a binary matrix representation of the input vector supplied. Applying the to_categorical function on the y_train and y_test arrays (label arrays)

```
In [15]: from keras.utils.np_utils import to_categorical

Using TensorFlow backend.
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:516: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_qint8 = np.dtype([("qint8", np.int8, 1)])
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:517: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:518: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_qint16 = np.dtype([("qint16", np.int16, 1)])
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:519: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_quint16 = np.dtype([("quint16", np.uint16, 1)])
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:520: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_qint32 = np.dtype([("qint32", np.int32, 1)])
/anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/dtypes.py:525: FutureWarning: Passing (type, 1) or '1type' a
s a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  np_resource = np.dtype([("resource", np.ubyte, 1)])
/anaconda3/lib/python3.7/site-packages/tensorboard/compat/tensorflow_stub/dtypes.py:541: FutureWarning: Passing (type, 1) or '1
type' as a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_qint8 = np.dtype([("qint8", np.int8, 1)])
/anaconda3/lib/python3.7/site-packages/tensorboard/compat/tensorflow_stub/dtypes.py:542: FutureWarning: Passing (type, 1) or '1
type' as a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
/anaconda3/lib/python3.7/site-packages/tensorboard/compat/tensorflow_stub/dtypes.py:543: FutureWarning: Passing (type, 1) or '1
type' as a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
  _np_qint16 = np.dtype([("qint16", np.int16, 1)])
/anaconda3/lib/python3.7/site-packages/tensorboard/compat/tensorflow_stub/dtypes.py:544: FutureWarning: Passing (type, 1) or '1
type' as a synonym of type is deprecated; in a future version of numpy, it will be understood as (type, (1,)) / '(1,)type'.
```

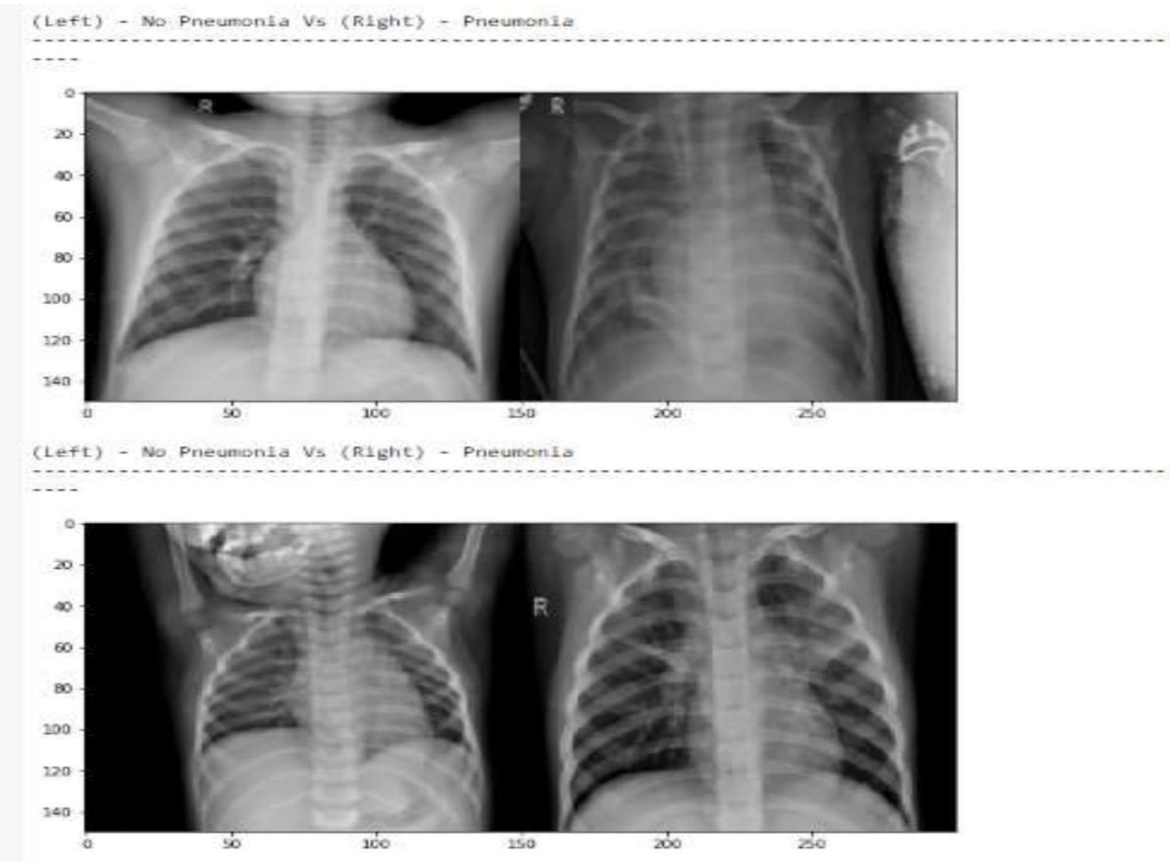
```
In [55]: y_train = to_categorical(y_train, 2)
         y_test = to_categorical(y_test, 2)
```

Using matplotlib to demonstrate pneumonia and no pneumonia images side by side.

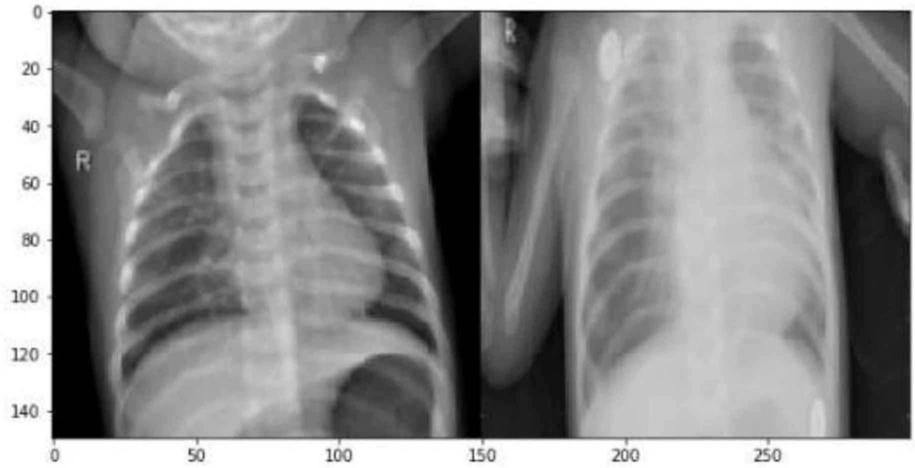
```
In [17]: Pimages = os.listdir(TRAIN_DIR + "PNEUMONIA")
         Nimages = os.listdir(TRAIN_DIR + "NORMAL")
```

```
In [20]: import matplotlib.pyplot as plt
         def plotter(i):
             imagep1 = cv2.imread(TRAIN_DIR+"PNEUMONIA/"+Pimages[i])
             imagep1 = skimage.transform.resize(imagep1, (150, 150, 3) , mode = 'reflect')
             imagen1 = cv2.imread(TRAIN_DIR+"NORMAL/"+Nimages[i])
             imagen1 = skimage.transform.resize(imagen1, (150, 150, 3))
             pair = np.concatenate((imagen1, imagep1), axis=1)
             print("(Left) - No Pneumonia Vs (Right) - Pneumonia")
             print("-----------------------------------------------------------------------------
             plt.figure(figsize=(10,5))
             plt.imshow(pair)
             plt.show()
         for i in range(5,10):
             plotter(i)
```

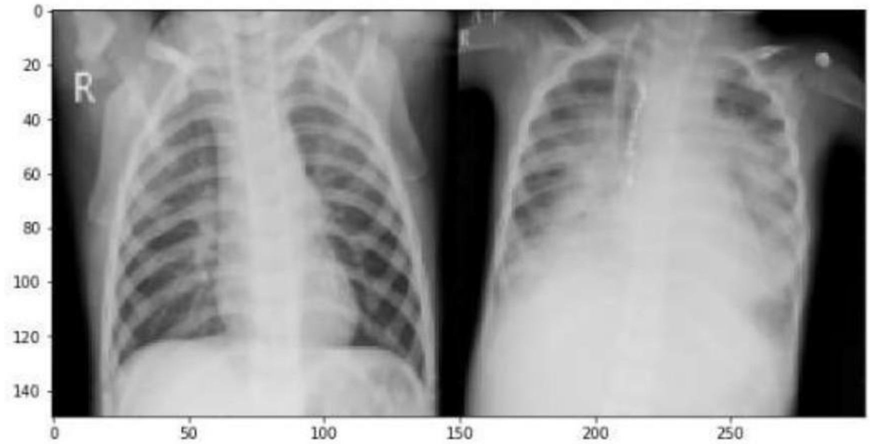
In the below images, it can be seen that the right-side images are chest x-rays of the person diagnosed with pneumonia. It is differentiated by observing the increased amount of cloudiness in the lungs.



```
(Left) - No Pneumonia Vs (Right) - Pneumonia
----------------------------------------------------------------------------------------------------
----
```




```
(Left) - No Pneumonia Vs (Right) - Pneumonia
----------------------------------------------------------------------------------------------------
----
```



Models often benefit from reducing the learning rate once learning stagnates. For this, we used ReduceLROnPlateau which monitors accuracy setting the factor by which to reduce learning rate as 0.1. Verbose is 1: update messages. P**atience**: number of epochs that produced the monitored quantity with no improvement after which training will be stopped.

```
In [21]: from keras.callbacks import ReduceLROnPlateau , ModelCheckpoint , LearningRateScheduler
         lr_reduce = ReduceLROnPlateau(monitor='val_acc', factor=0.1, epsilon=0.0001, patience=1, verbose=1)

/anaconda3/lib/python3.7/site-packages/keras/callbacks.py:1335: UserWarning: `epsilon` argument is deprecated and will be remov
ed, use `min_delta` instead.
  warnings.warn('`epsilon` argument is deprecated and '
```

Transfer learning

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

We specified the transfer learning weights filepath and called ModelCheckPoint to save model to the given filepath after every epoch.

```
In [22]: filepath="transferlearning_weights.hdf5"
         checkpoint = ModelCheckpoint(filepath, monitor='val_acc', verbose=1, save_best_only=True, mode='max')
```

Inception v3 is a widely-used image recognition model that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years.

Importing and preparing the base InceptionV3 model

```
In [39]: from keras.applications.inception_v3 import InceptionV3
         # create the base pre-trained model
         base_model = InceptionV3(weights=None, include_top=False , input_shape=(150, 150,3))
```

```
In [40]: x = base_model.output
         x = Dropout(0.5)(x)
         x = GlobalAveragePooling2D()(x)
         x = Dense(128, activation='relu')(x)
         x = BatchNormalization()(x)
         predictions = Dense(2, activation='sigmoid')(x)
```

Configuring the base model weights and compiling the model set for accuracy prediction

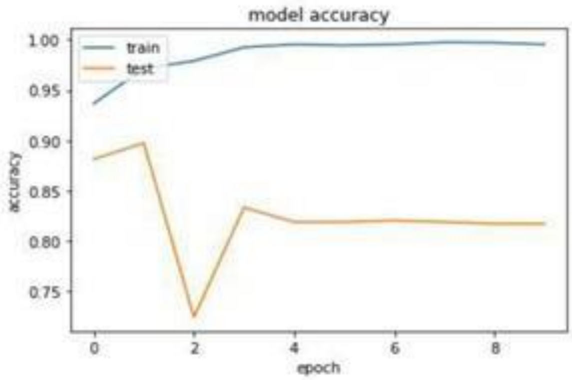
```
In [42]: base_model.load_weights("chest_xray/inception_v3_weights.h5")

In [43]: model = Model(inputs=base_model.input, outputs=predictions)

In [44]: model.compile(loss='categorical_crossentropy',
                       optimizer='adam',
                       metrics=['accuracy'])
```

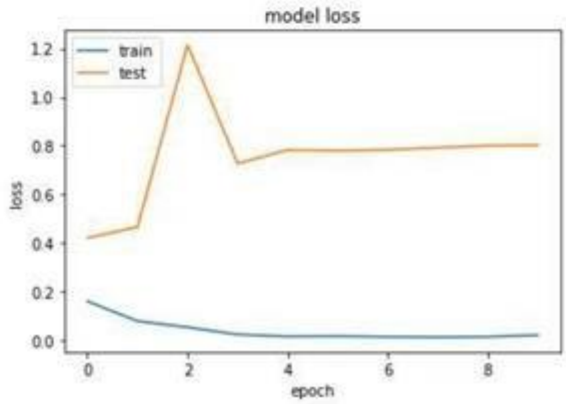
```
In [45]: print(model.summary())

Model: "model_1"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
input_1 (InputLayer)            (None, 150, 150, 3)  0
__________________________________________________________________________________________________
conv2d_1 (Conv2D)               (None, 74, 74, 32)   864         input_1[0][0]
__________________________________________________________________________________________________
batch_normalization_1 (BatchNor (None, 74, 74, 32)   96          conv2d_1[0][0]
__________________________________________________________________________________________________
activation_1 (Activation)       (None, 74, 74, 32)   0           batch_normalization_1[0][0]
__________________________________________________________________________________________________
conv2d_2 (Conv2D)               (None, 72, 72, 32)   9216        activation_1[0][0]
__________________________________________________________________________________________________
batch_normalization_2 (BatchNor (None, 72, 72, 32)   96          conv2d_2[0][0]
__________________________________________________________________________________________________
activation_2 (Activation)       (None, 72, 72, 32)   0           batch_normalization_2[0][0]
__________________________________________________________________________________________________
conv2d_3 (Conv2D)               (None, 72, 72, 64)   18432       activation_2[0][0]
__________________________________________________________________________________________________
batch_normalization_3 (BatchNor (None, 72, 72, 64)   192         conv2d_3[0][0]
__________________________________________________________________________________________________
activation_3 (Activation)       (None, 72, 72, 64)   0           batch_normalization_3[0][0]
__________________________________________________________________________________________________
max_pooling2d_1 (MaxPooling2D)  (None, 35, 35, 64)   0           activation_3[0][0]
__________________________________________________________________________________________________
```

Setting up batch size and number of epochs and calling model.fit to train the model and storing the output in history variable. The batch size has been set up to be 64 and number of epochs are 10.

```
In [46]: batch_size = 64
         epochs = 10

In [47]: history = model.fit(X_train, y_train, validation_data = (X_test , y_test) ,callbacks=[lr_reduce,checkpoint] ,
                              epochs=epochs)

WARNING:tensorflow:From /anaconda3/lib/python3.7/site-packages/tensorflow/python/ops/math_grad.py:1250: add_dispatch_support.
<locals>.wrapper (from tensorflow.python.ops.array_ops) is deprecated and will be removed in a future version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
Train on 5216 samples, validate on 624 samples
Epoch 1/10
5216/5216 [==============================] - 692s 133ms/step - loss: 0.1600 - acc: 0.9367 - val_loss: 0.4200 - val_acc: 0.881
4

Epoch 00001: val_acc improved from -inf to 0.88141, saving model to transferlearning_weights.hdf5
Epoch 2/10
5216/5216 [==============================] - 761s 146ms/step - loss: 0.0782 - acc: 0.9714 - val_loss: 0.4654 - val_acc: 0.897
4

Epoch 00002: val_acc improved from 0.88141 to 0.89744, saving model to transferlearning_weights.hdf5
Epoch 3/10
5216/5216 [==============================] - 793s 152ms/step - loss: 0.0527 - acc: 0.9789 - val_loss: 1.2146 - val_acc: 0.724
4
```

Using matplotlib to trace the model accuracy and model loss with the respect to the number of epochs





Representing the model accuracy by plotting the confusion matrix with x axis is the predicted result and y axis as the true result.

First quadrant represents that patient is not diagnosed with pneumonia and model predicted for the same.

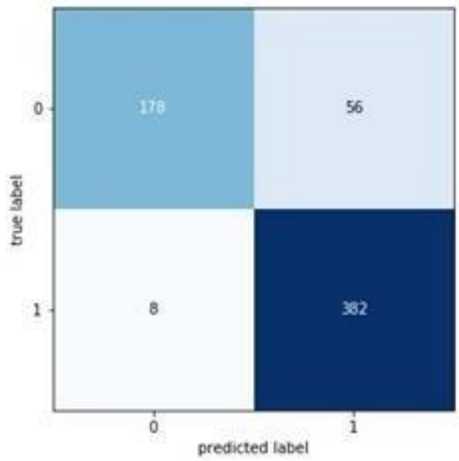
Second quadrant shows the number of cases where person is not diagnosed with pneumonia but model predicted that patient has pneumonia.

Third quadrant is most critical one as it shows the number of cases where the patient is diagnosed with pneumonia and model predicted opposite.

Second quadrant shows the number of cases where person is diagnosed with pneumonia and model predicts for the same.

```
In [50]: from sklearn.metrics import confusion_matrix
         pred = model.predict(X_test)
         pred = np.argmax(pred,axis = 1)
         y_true = np.argmax(y_test,axis = 1)
```

```
In [52]: CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()
```



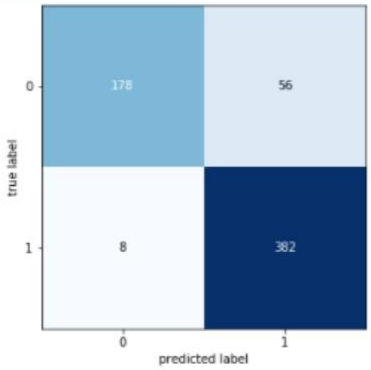
Accuracy of the model when the person is not actually suffering from pneumonia came out to be 76%

```
In [1]: 178/(178+56)
Out[1]: 0.7606837606837606
```

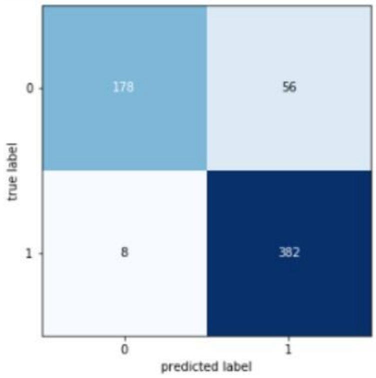
Accuracy of the model when the person is actually suffering from pneumonia came out to be 98%

```
In [54]: 382/(382+8)
Out[54]: 0.9794871794871794
```

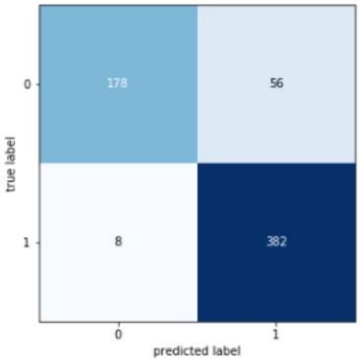
```
In [47]: fx_test , fy_test = get_data("Median/")
         fy_test = to_categorical(fy_test, 2)
         fx_test=fx_test.reshape(624,150,150,3)

         fpred = model.predict(fx_test)
         fpred = np.argmax(fpred,axis = 1)
         fy_true = np.argmax(fy_test,axis = 1)
         CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()

100%|██████████| 234/234 [00:10<00:00, 21.91it/s]
100%|██████████| 390/390 [00:16<00:00, 23.22it/s]
```



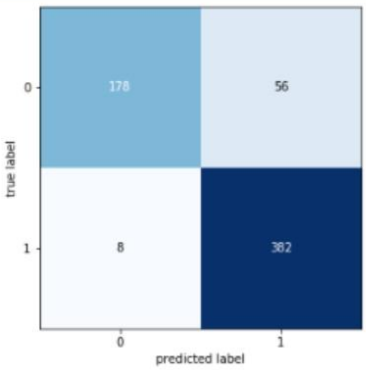

```
In [49]: fx_test , fy_test = get_data("Blurring/")
         fy_test = to_categorical(fy_test, 2)
         fx_test=fx_test.reshape(624,150,150,3)

         fpred = model.predict(fx_test)
         fpred = np.argmax(fpred,axis = 1)
         fy_true = np.argmax(fy_test,axis = 1)
         CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()

100%|██████████| 234/234 [00:09<00:00, 24.07it/s]
100%|██████████| 390/390 [00:16<00:00, 23.76it/s]
```



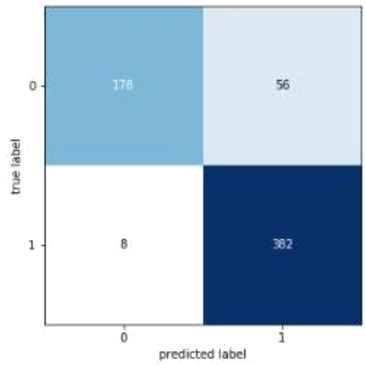
```
In [48]: fx_test , fy_test = get_data("Bilateral/")
         fy_test = to_categorical(fy_test, 2)
         fx_test=fx_test.reshape(624,150,150,3)

         fpred = model.predict(fx_test)
         fpred = np.argmax(fpred,axis = 1)
         fy_true = np.argmax(fy_test,axis = 1)
         CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()

100%|██████████████████████████████████████████████████████████████████████████████████
█████| 234/234 [00:11<00:00, 20.74it/s]
100%|██████████████████████████████████████████████████████████████████████████████████
█████| 390/390 [00:16<00:00, 24.35it/s]
```




```
In [46]: fx_test , fy_test = get_data("Affine/")
         fy_test = to_categorical(fy_test, 2)
         fx_test=fx_test.reshape(624,150,150,3)

         fpred = model.predict(fx_test)
         fpred = np.argmax(fpred,axis = 1)
         fy_true = np.argmax(fy_test,axis = 1)
         CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()

100%|██████████████████████████████████████████████████████████████████████████████████
█████| 234/234 [00:10<00:00, 21.34it/s]
100%|██████████████████████████████████████████████████████████████████████████████████
█████| 390/390 [00:17<00:00, 22.75it/s]
```



```
In [45]: fx_test , fy_test = get_data("2DConv/")
         fy_test = to_categorical(fy_test, 2)
         fx_test=fx_test.reshape(624,150,150,3)

         fpred = model.predict(fx_test)
         fpred = np.argmax(fpred,axis = 1)
         fy_true = np.argmax(fy_test,axis = 1)
         CM = confusion_matrix(y_true, pred)
         from mlxtend.plotting import plot_confusion_matrix
         fig, ax = plot_confusion_matrix(conf_mat=CM ,  figsize=(5, 5))
         plt.show()

100%|██████████████████████████████████████████████████| 234/234 [00:10<00:00, 23.28it/s]
100%|██████████████████████████████████████████████████| 390/390 [00:20<00:00, 19.27it/s]
```



Results

Hence, we achieve 97% accuracy of actually diagnosed with Pneumonia after applying Convolution Neural network to the give dataset and 76% accuracy of not diagnosed with Pneumonia.

Conclusion

This project is a small step toward the huge contribution of image processing and deep learning towards the medical field with a lot more to come in the future. The model we prepared was successful with 98% approximately when patient was diagnosed with pneumonia. Transfer Learning along with image processing algorithm are the main driving force behind this project. Transfer learning definitely going to be one of the key drivers for machine learning and deep learning success in mainstream adoption in the industry.

REFERENCES

[1] Souza, J. C., Diniz, J. O. B., Ferreira, J. L., da Silva, G. L. F., Silva, A. C., & de Paiva,

A. C. (2019). An automatic method for lung segmentation and reconstruction in chest X-Ray using deep neural networks. Computer Methods and Programs in Biomedicine.

[2] Zar, H. J. (2017). Bacterial and viral pneumonia: New insights from the Drakenstein Child Health Study. Paediatric respiratory reviews, 24, 8-10.

[3] Dear, J. D. (2014). Bacterial pneumonia in dogs and cats.

[4] Wilmes, D., Coche, E., Rodriguez-Villalobos, H., & Kanaan, N. (2018). Bacterial pneumonia in kidney transplant recipients. Respiratory medicine, 137, 89-94.

[5] Henig, O., & Kaye, K. S. (2017). Bacterial pneumonia in older adults. Infectious Disease Clinics, 31(4), 689-713.

[6] Wong, J. L., & Evans, S. E. (2017). Bacterial pneumonia in patients with cancer: novel risk factors and management. Clinics in chest medicine, 38(2), 263-277.

[7] Das, D., & Howlett, D. C. (2009). Chest X-ray manifestations of pneumonia. Surgery- Oxford International Edition, 27(10), 453-455.

[8] Falsey, A. R. (2007). Community-acquired viral pneumonia. Clinics in geriatric medicine, 23(3), 535-552.

[9] Nayak, S., Pradhan, D., Singh, H., & Reddy, M. S. (2019). Computational screening of potential drug targets for pathogens causing bacterial pneumonia. Microbial pathogenesis, 130, 271-282.

[10] Erdem, S. B., Can, D., Girit, S., Çatal, F., Şen, V., Pekcan, S., & Ersu, R. (2018). Does atopy affect the course of

viral pneumonia? Allergologia et immunopathologia, 46(2), 119- 126.

[11] Amorin, J. P., Lopez, M., Rando, K., Castelli, J., & Presentado, J. M. (2018, March). Early bacterial pneumonia after hepatic transplantation: epidemiologic profile. In Transplantation proceedings (Vol. 50, No. 2, pp. 503-508). Elsevier.

[12] Radigan, K. A., & Wunderink, R. G. (2011). Epidemic viral pneumonia and other emerging pathogens. Clinics in chest medicine, 32(3), 451-467.

[13] Jain, S. (2017). Epidemiology of viral pneumonia. Clinics in chest medicine, 38(1), 1-9.

[14] Grief, S. N., & Loza, J. K. (2018). Guidelines for the evaluation and treatment of pneumonia. Primary Care: Clinics in Office Practice, 45(3), 485-503.

[15] Feldman, C., & Anderson, R. (2013). HIV-associated bacterial pneumonia. Clinics in chest medicine, 34(2), 205-216.

[16] Kim, M., Lee, S. M., Song, J. W., Do, K. H., Lee, H. J., Lim, S., ... & Seo, J. B. (2017).

Added value of prone CT in the assessment of honeycombing and classification of usual interstitial pneumonia pattern. European journal of radiology, 91, 66-70.

[17] Yunt, Z. X., Chung, J. H., Hobbs, S., Fernandez-Perez, E. R., Olson, A. L., Huie, T. J., ... & Brown, K. K. (2017). High resolution computed tomography pattern of usual interstitial pneumonia in rheumatoid arthritis-associated interstitial lung disease: relationship to survival. Respiratory medicine, 126, 100-104.

[18] Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., & Ibragimov,

B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. Computers & Electrical Engineering, 78, 388-399.

[19] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., ... & Dong, J. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell, 172(5), 1122-1131.

[20] Yates, E. J., Yates, L. C., & Harvey, H. (2018). Machine learning “red dot”: open- source, cloud, deep convolutional neural networks in chest radiograph binary normality classification. Clinical radiology, 73(9), 827-831.

[21] Er, O., Sertkaya, C., Temurtas, F., & Tanrikulu, A. C. (2009). A comparative study on chronic obstructive pulmonary and pneumonia diseases diagnosis using neural networks and artificial immune system. Journal of medical systems, 33(6), 485-492.

[22] Coopera, G. F., Aliferis, C. F., Ambrosinoa, R., Aronisb, J., Buchananb, B. G., Caruana, R., ... & Meek, C. (1996). Artificial intelligence in medicine.

[23] Kauczor, H. U., Heitmann, K., Heussel, C. P., Marwede, D., Uthmann, T., & Thelen,

M. (2000). Automatic detection and quantification of ground-glass opacities on high- resolution CT using multiple neural networks: comparison with a density mask. American Journal of Roentgenology, 175(5), 1329-1334.

[24] Er, O., Yumusak, N., & Temurtas, F. (2010). Chest diseases diagnosis using artificial neural networks. Expert Systems with Applications, 37(12), 7648-7655.

[25] Gurung, A., Scrafford, C. G., Tielsch, J. M., Levine, O. S., & Checkley, W. (2011). Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis. Respiratory medicine, 105(9), 1396-1403.

[26] Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. (2019). Identifying Pneumonia in Chest X-Rays: A Deep Learning Approach. Measurement.

[27] Li, F., Engelmann, R., Pesce, L., Armato, S. G., & MacMahon, H. (2012). Improved detection of focal pneumonia by chest radiography with bone suppression imaging. European radiology, 22(12), 2729-2735.

[28] Heckerling, P. S., Gerber, B. S., Tape, T. G., & Wigton, R. S. (2003). Prediction of community-acquired pneumonia using artificial neural networks. Medical decision making, 23(2), 112-121.

[29] Moncada, D. C., Rueda, Z. V., Macías, A., Suárez, T., Ortega, H., & Vélez, L. A. (2011). Reading and interpretation of chest X-ray in adults with community-acquired pneumonia. The Brazilian Journal of Infectious Diseases, 15(6), 540-546.

[30] Ackerman, S. J., Gitlin, J. N., Gayler, R. W., Flagle, C. D., & Bryan, R. N. (1993). Receiver operating characteristic analysis of fracture and pneumonia detection: comparison of laser-digitized workstation images and conventional analog radiographs. Radiology, 186(1), 263-268.

[31] Heckerling, P. S., Gerber, B. S., Tape, T. G., & Wigton, R. S. (2003). Prediction of community-acquired pneumonia using artificial neural networks. Medical decision making, 23(2), 112-121.

[32] Kunz, W. G., Patzig, M., Crispin, A., Stahl, R., Reiser, M. F., & Notohamiprodjo, M. (2018). The Value of Supine Chest X-Ray in the Diagnosis of Pneumonia in the Basal Lung Zones. Academic radiology, 25(10), 1252-1256.

[33] Cicero, M., Bilbily, A., Colak, E., Dowdell, T., Gray, B., Perampaladas, K., & Barfett,

J. (2017). Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. Investigative radiology, 52(5), 281-287.

[34] Heckerling, P. S., Gerber, B. S., Tape, T. G., & Wigton, R. S. (2004). Use of genetic algorithms for neural networks to predict community-acquired pneumonia. Artificial Intelligence in Medicine, 30(1), 71-84.

[35] Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. PLoS medicine, 15(11), e1002683.