Detection and Classification of Lung Disease – Pneumonia and Lung Cancer in Chest Radiology Using Artificial Neural Network

Pavithra R[*] S.Y. Pattar[**]

[*] Department of Medical Electronics, BMS College of Engineering
[**] Department of Medical Electronics, BMS College of Engineering

Abstract- Chest radiology is the most common method used for diagnosis of lung diseases, the term lung disease refers to the abnormalities that effect the lung organ, diseases are such as asthma, COPD, lung cancer, pneumonia and many other breathing problems, in this paper, we develop a system that defects and classify the lung diseases as either pneumonia or lung cancer, this is accomplished by two stages they are feature extraction and classification, feature extraction is done through the use of Gabor filter, classification is through the use of neural network's like feed forward neural network(FFNN), Multi-layer perceptron neural network(MLPNN), Radial Basis Function(RBF).

Index Terms- feature extraction, classification, Neural Networks

I. INTRODUCTION

Lungs are the organs which are contained in the thoracic cavity, there are enveloped in two pleural membranes. There are two types of lung cancer, that is, small cell lung cancer and non-small cell lung cancer. The lungs are subdivided into lobes, compartments and fed by different parts of bronchial and vessel trees.

Lung cancer is a cancer that starts in the lungs. Lung cancer can start in the cells lining the bronchi and parts of the lungs such as the bronchioles or alveoli, changes in the genes (DNA) inside the lung cells may cause the cells to grow faster, they form a tumor.

Pneumonia is a other type of lung disease, which can be said as acute inflammation of lung parenchyma, inflammatory infiltrate in alveoli.

In this paper lung disease detection system is developed, this study classifies lung disease images as either lung cancer or pneumonia, this is accomplished by two stages of system, feature extraction and classification. Feature extraction is done through the use of Gabor filter, Gabor filters extract certain important features from the images Mean, Variance, Standard Deviation, Homogeneity, Energy, Contrast, Correlation. This set of extracted features are called as Gabor Feature set. Classification is done through the use of various types of Artificial Neural Networks. They are Radial Basis Function, Multi-Layer Perceptron and Feed Forward Neural Network.

II. PROPOSED WORK

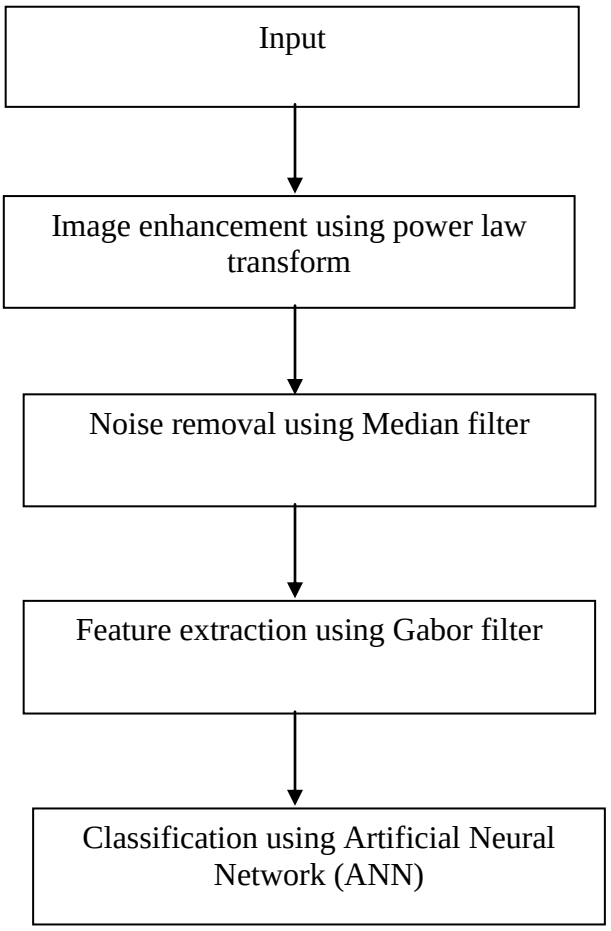


Figure: Flow chart of proposed work

The flow chart represents the developed system, which is used to detect and classify lung cancer and pneumonia lung disease from a collection of abnormal images, there are several stages in this system such as Enhancement of the image to improve the quality, removal of noise using Median filter, feature extraction using Gabor filter and finally classification using Artificial the flow chart of proposed system in detail.

The very first stage is to collect abnormal lung X ray images, the data set containing 116 images comprising lung cancer(58) and pneumonia(58) is been utilized.

Data preprocessing

Preprocessing is the process of improving or enhancing the quality of input image and make the feature extraction phase more reliable, main motive of preprocessing stages is to remove noise present in input image, here in the preprocessing stage median filter is used to remove noise from the input image and for image enhancement power law transform's is been used.

Median filter is another type of noise removal filter, it creates boundary of n*n in the noise input image, the n*n sub region is scanned over the entire image. The pixel values present in the n*n boundary is listed or arranged in the descending order, the middle value (median) is replaced with the centre pixel value of the n*n boundary. Usually n = 3 or 5 in median filter.

The above process is repeated for all the n*n sub regions of the input noisy image, such that the pixel value from which noise has to be removed will be the centre pixel value of each non boundary.

Feature extraction

Images have a huge number of features, it is important to recognize and extract such features from input images, feature extraction is the process done to reduce the complexity of processing, here for feature extraction we use Gabor filter, Gabor filter extract local pieces of information which are then combined to recognize an object or ROI 2D Gabor filter function.

$$\Psi(x, y) = f^2/\pi\Upsilon\eta e^{-(\frac{f^2}{\Upsilon^2}x'^2+f^2/\Upsilon^2y'^2)}e^{j2\pi fx'}$$

$$x' = x\cos\theta+y\sin\theta$$
$$y' = -x\sin\theta+y\cos\theta$$

The Gabor filter will extract the features like Mean, Variance, Standard Deviation, Contrast, Correlation, Homogeneity, and Energy.

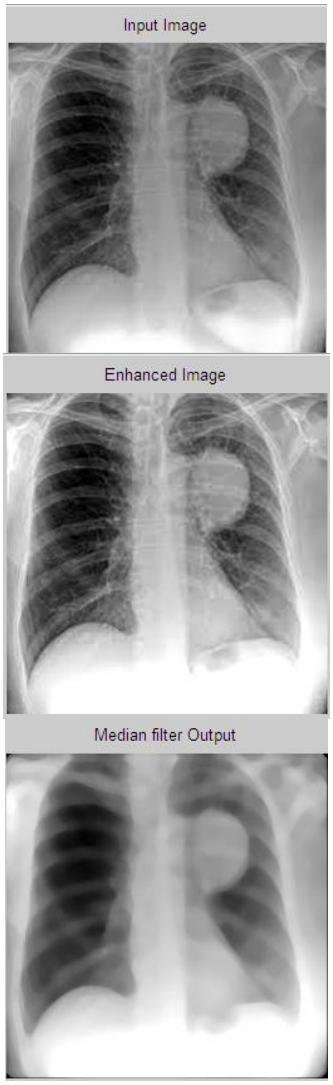
CLASSIFIER

The classifier is a mathematical function which is implemented using classification algorithm which maps input data to a particular category. There are various types of classifiers. One such is Artificial Neural Network(ANN) which is used in this paper. Artificial neural network is a network of simple processing elements called neurons, which operates on their local data and communicates with other elements, three types of ANN is used here namely Feed forward Network, Radial Basis function and Multilayer perceptron Network.

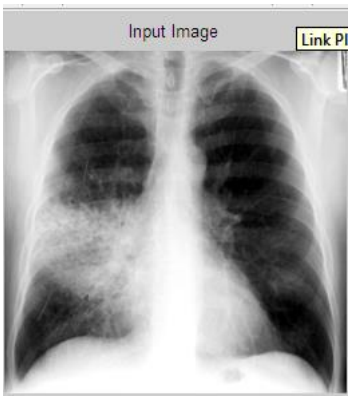
III. RESULTS AND DISCUSSION

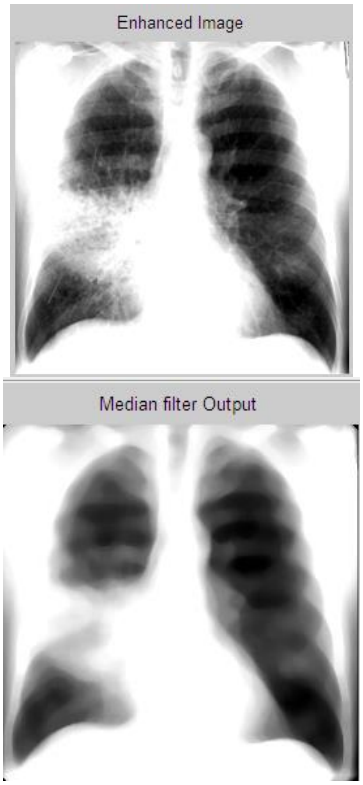
1. Preprocessing results for lung cancer detection

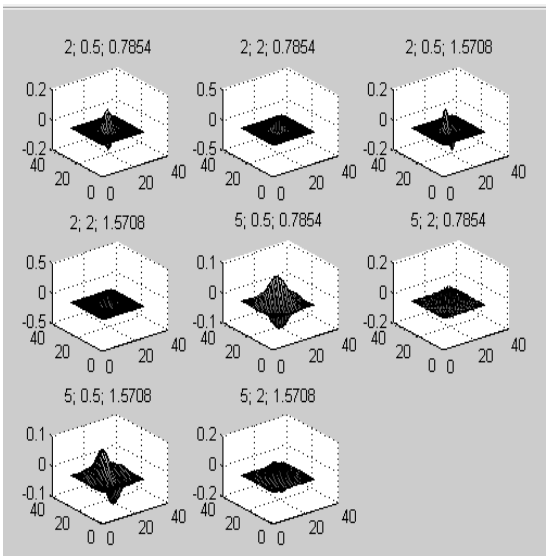
Case 1: lung cancer image



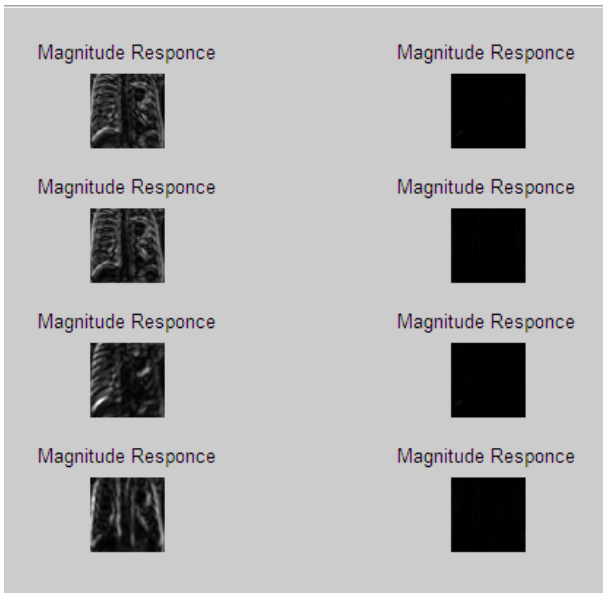
Case 2: Pneumonia image







Gabor filter response

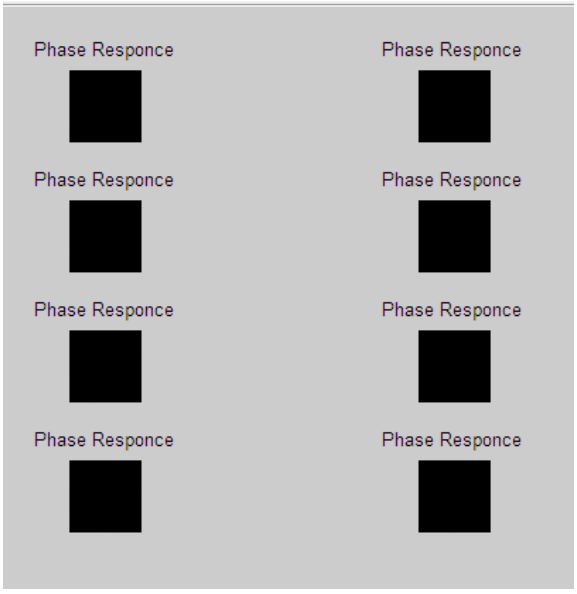


Magnitude response of chosen input abnormal image

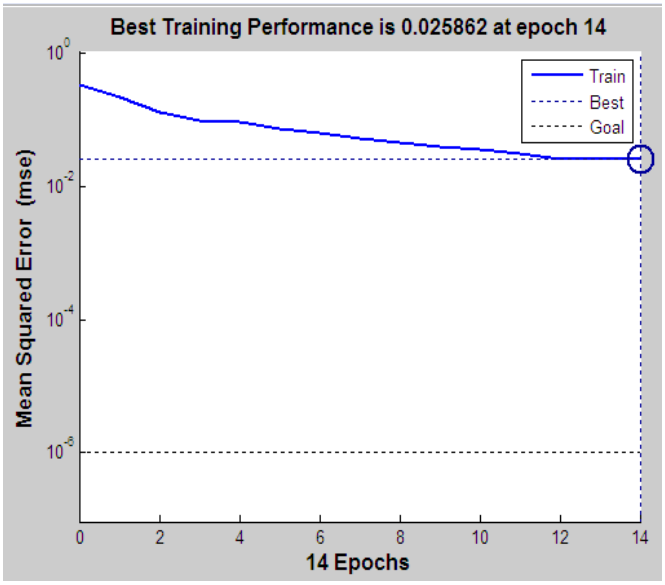
Feature extraction results:

	1	2	3	4	5	6	7	8	9
1	53.4401	59.9849	53.1754	60.7115	58.1471	62.4787	41.3473	76.0473	81.6538
2	2909	2496	2535	2451	2683	1697	1589	6706	4827
3	.0190e+03	8.7859e+03	1.1052e+04	9.6193e+03	1.0627e+04	8.6226e+03	1.0166e+04	9.9791e+03	8.6049e+03
4	0.0673	0.0114	-0.2268	-0.0181	-0.2224	-0.0108	-0.0016	0.0193	0.0825
5	.7508e-05	3.1583e-05	3.5281e-05	3.1150e-05	3.3444e-05	2.6874e-05	3.5816e-05	3.4688e-05	3.1642e-05
6	0.0380	0.0390	0.0298	0.0373	0.0315	0.0439	0.0400	0.0359	0.0388
7	53.9344	49.9569	50.3482	49.5097	51.7952	41.1940	39.8593	81.8929	69.4754
8	0.7631	0.7435	0.8099	0.7853	0.9548	1.1278	0.6654	0.8498	0.9874
9	2	1	1	1	2	3	1	1	3
10	.3408e+03	8.9562e+03	1.1107e+04	9.1991e+03	9.5824e+03	8.7699e+03	8.0846e+03	8.4411e+03	9.6470e+03
11	-0.0149	-0.0444	-0.3168	5.2684e-04	-0.1101	-0.0843	0.0441	0.0688	0.0551
12	.6227e-05	6.1008e-05	6.9900e-05	6.3841e-05	6.4348e-05	7.8858e-05	6.1201e-05	6.3584e-05	8.5408e-05
13	0.0317	0.0413	0.0291	0.0437	0.0380	0.0414	0.0435	0.0404	0.0422
14	1.2324	0.9825	1.1740	1.0395	1.3437	1.8631	0.7876	1.1608	1.6926
15	53.1765	61.2258	51.8611	63.6426	57.8600	68.1740	41.9395	72.4537	88.6724
16	3012	2653	2345	2707	2579	1900	1512	4947	5812

Feature Extraction Values



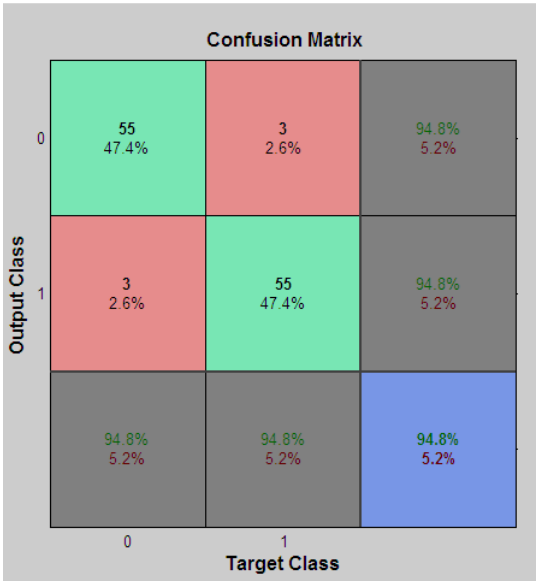
Phase response of chosen input abnormal image



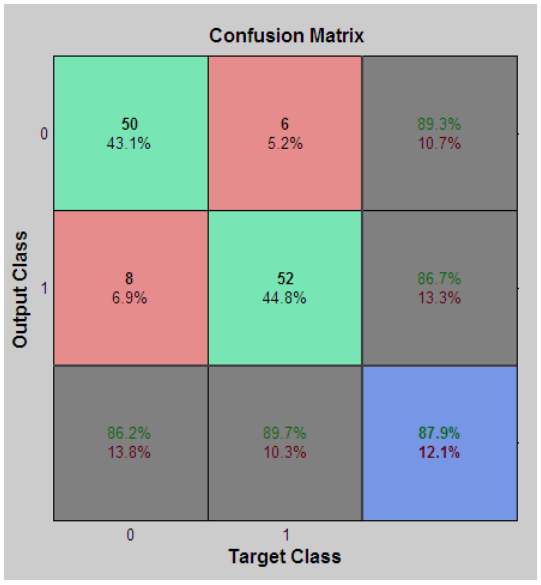
Graph of Mean Square Error for different iteration

Multilayer perceptron

Classifier results:
Feed forward neural network

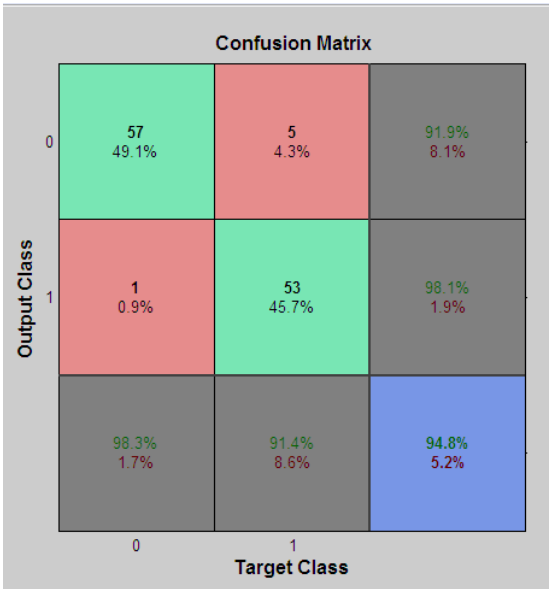


Confusion matrix

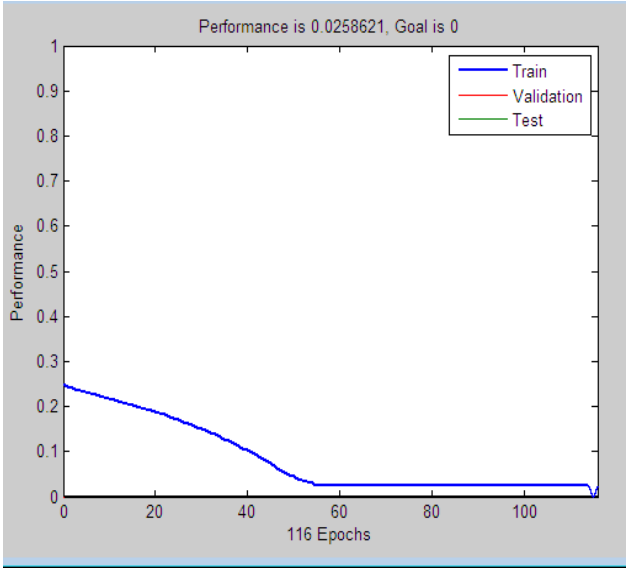


Confusion matrix

Radial Basis Function Neural Network



Confusion matrix



Graph of Mean Square Error for Different Iteration

IV. Conclusion

In this project work, lung disease detection system has been developed, abnormal images are considered as input, on which preprocessing techniques are applied using power law transform and median filter to remove noise, features extraction is done by Gabor filter and finally classification of images as either lung cancer or pneumonia is done by Artificial neural network, Radial Basis Function Neural Network, Multi Layer Perceptron Neural Network, Feed Forward Neural Network. Here three types of ANN classifiers have been used in this project out of which we have got best results for Feed Forward and Radial Basis Function Network than compared to all other existing methods of classifiers, percentage of accuracy been obtained is about 94.8% for Feed Forward and 94.82% for Radial Basis Function Network.

References

[1] Yetisgen-Yildiz M, Glavan BJ, Xia F, Vanderwende L, Wurfel MM. Identifying patients with pneumonia from free-text intensive care unit reports. Proceedings of Learning from Unstructured Clinical Text Workshop of the International Conference on Machine Learning 2011.

[2] Bejan CA, Xia F, Vanderwende L, Wurfel MM, Yetisgen-Yildiz M. Pneumonia identification using statistical feature selection. In submission.

[3] K.A.G. Udeshani, R.G.N. Meegama, T.G.I. Fernando, "Statistical Feature-based Neural Network Approach for the Detection of Lung Cancer in Chest X-Ray Images," International Journal of Image Processing (IJIP), Volume (5), Issue (4) , 2011.

[4] V.Krishnaiah, Dr.G.Narsimha, Dr.N.Subhash Chandra. 2013, "Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques," International Journal of Computer Science and Information Technologies, Vol. 4 (1), 2013, 39 – 45.

[5] S. A. Patil and V. R. Udupi, Chest X-rays features extraction for lung cancer classification, Journal of Scientific and Industrial Research, vol. 69, pp. 271-277, 2010.

Authors

First Author – Pavithra R, Bachelor of Engineering (M.Tech), B M S College of Engineering, pavi2710@gmail.com
Second Author – S.Y. Pattar, M.Tech (PhD), B M S College of Engineering, sypmudgal@yahoo.com