PNEUMONIA CLASSIFICATION USING HYBRID METHODS AND IMAGE PROCESSING TECHNIQUES ON CHEST X-RAY IMAGES

D. NAPOLEON
Assistant Professor
Department of Computer Science
Bharathiar University
Coimbatore – 641046
mekaranapoleon@yahoo.co.in

I. KALAIARASI
Research Scholar
Department of Computer Science
Bharathiar University
Coimbatore – 641046
kalaiarasii3110@gmail.com

***Abstract*— lung diseases also known as respiratory diseases. The diseases in lung airways or other structure of the lungs called lung diseases. The diseases are many types especially the pneumonia is most dangerous lung disease due to its causing rate in children below the five years and adults. The pneumonia is an infection which is caused in lung air sacs filled with pus or fluid. The death rate of pneumonia causes is higher due to the late diagnosis. Therefore, the early detection of pneumonia is needed for increasing the survival rate of the people. Many CAD systems are employed, but still there is some space for improving the accuracy to detect the pneumonia disease. In this article the Hybrid Support Vector Machine and Artificial Neural Network of Back-Propagation Neural Network (HSA-BPNN) has been proposed for classifying pneumonia. The image has been pre-processed with Weiner filter, segmentation and feature extraction has been performed by Active Contour Integration (ACI) and Grey Level Co-occurrence Matrix (GLCM) method. In segmentation process, the ACI algorithm has separating or dividing the lung portion into left and right side of the lung. The GLCM has extract the thirteen number of features from segmented image. Finally, the proposed algorithm HSA-BPNN has been used for classifying the pneumonia disease which is based on features. Furthermore, the performance of the classifier model has been calculated based on three parametrics which are accuracy, specificity, and sensitivity which are compared with an existing algorithm. Hence, The proposed classifier HAS-BPNN shown the better classification accuracy compared than the existing methods.**

Keywords — *Chest X-Ray image, Pneumonia, Weiner filter, Active Contour Integration, GLCM, HSA-BPNN.*

I. INTRODUCTION

Lung disease is common world-wide. The pneumonia is an infectious disease which is caused by fungi, virus and bacterial. The infection affects the air sacs of human lung with full of fluid or pus. According from the recent survey which was given by WHO, totally 808 694 children were killed in the year 2017.The pneumonia accounts 15% of deaths in children that below the five years old [1]. The several medical modalities are used for detecting lung diseases such as the Skin test, Sputum test, Computed Tomography, Magnetic Resonance Imaging, Chest X-Ray examination. In many situations, the doctors are advised to consider Chest X-Ray image and CT scan for analyzing the lung diseases. The Chest X-Ray methods are cost effective and non-invasive properties comparing than CT screening. Therefore, the Chest X–Ray is a common method to detect lung diseases and also other diseases and its presents the affected area of lung accurately Jaeger et al.[2]. For better diagnosis, the skilled radiologist or physicians are needs for analyzing the diseases. In many hospitals, it is a main drawback for disease detection so that we can go for automated CAD system. Recently, the Artificial Intelligence having more space in disease prediction especially machine learning methods. The machine learning is a subset of AI and it's very efficient for disease prediction problem. The different types of machine learning techniques are there but the neural network and support vector machine has much space and gives effective results than other techniques. Therefore, the lung disease of pneumonia detection with machine learning techniques and image processing methods gives better accuracy which is increasing the survival rate is increased.

Many researchers contribute their work for lung disease detection using machine learning methods. P. Huang et.al [3] proposed automated system using chest radiographs for nodule detection and tuberculosis classification [4]. Rajpurkar et al [5] employed classical deep learning called DenseNet-121 used to diagnosis of pneumonia. Li et.al [6] proposed squeeze and excitation network (SENet) for identifying pneumonia in an image area. Nahid et.al [7] presents the CNN architecture with CLAHE method to identifying the patients suffered by pneumonia. Ayan et.al [8] used transfer learning and fine tuning of VGG16-Net, Xception-Net for classifying pneumonia.

II. METHODOLOGY

The proposed methodology HSA-BPNN has four steps for classifying pneumonia or normal. The work flow of proposed system described below.

A. DATA PRE-PROCESSING

The input image of Chest X-Ray image has been pre-processed by using wiener filter. The wiener filter which is used for removing noise in an input image. Here, the weiner filter is used for image restoration which is consider the image as Additive White Gaussian Noise (AWGN) when the input image contains some noise. The main requirement of the weiner filter is when the input image having nothing about spectral density function and particular noise. The output results for the proposed Weiner filter image have been given below in figure1 and figure 2.

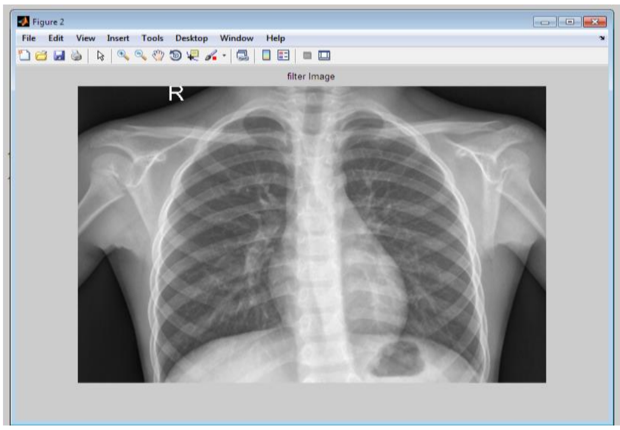

Fig. 1. Pre-processing Image for Normal

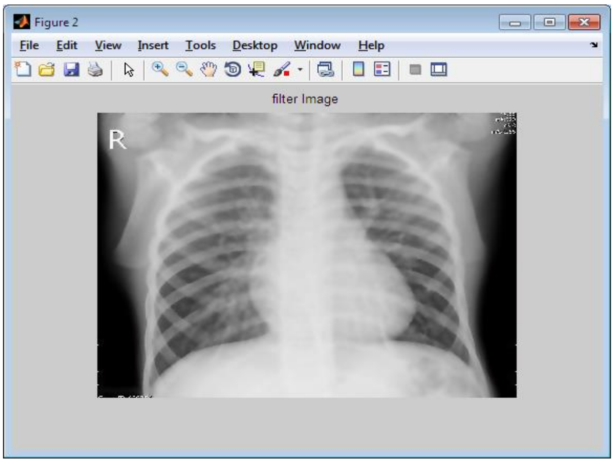

Fig. 2. Pre-processing Image for Pneumonia

B. SEGMENTATION

Segmentation means dividing the Chest X-Ray image into multiple segments. Instead of using entire image, only the segmented part has been used for further processing. In this system the Active Contour Integration has been used as a segmentation method for partitioning the chest X-Ray image. This proposed algorithm segment the Chest X-Ray image by using energy forces and constraints. The cotour is describing the curve which is in the form of splines, linear and polynomial. Here, by using energy forces the active contour segregating the region of interest pixel from the input image. The region of interest has been segmented from the chest X-Ray image which is used for further analysis of an image. The output for ACI segmentation is given below in figure 3 and figure 4.

C. FEATURE EXTRACTION

The features have been extracted from segmented region of interest image by using Grey-Level Co-occurrence Matrix (GLCM). The GLCM mainly used for extracting the texture feature of segmented image which is used for classification purpose. In this proposed work totally there are 13 features has been extracted which are entropy, energy, contrast, correlation, variance, skewness, smoothness, RMS, mean, standard deviation, average grey level, kurtosis. Hence, based on these thirteen features the Chest X-Ray image has been identified as pneumonia or normal lung. The test image features has been compared with this trained features, if the features are matched with testing image which is considered as pneumonia or it has to be normal. Here, the Table1 & Table2 shows the extracted features for normal and pneumonia from region of interest image.

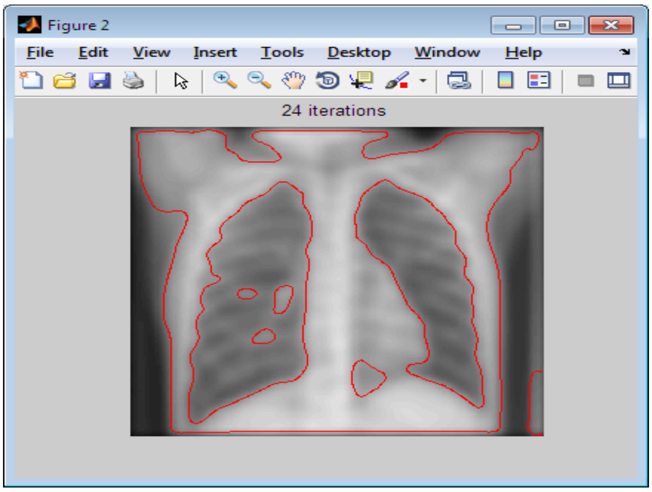

Fig. 3. Results for ACI Segmentation normal

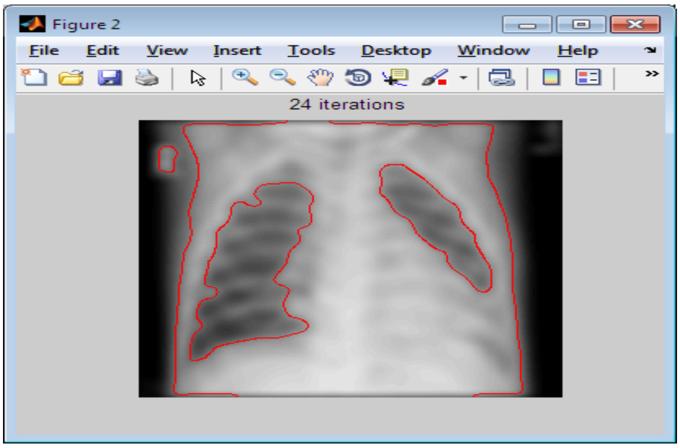

Fig. 4. Results for ACI Segmentation Pneumonia

Table 1. Features Extracted from GLCM Technique for Normal Lung

Variables - test_features
test_features
test_features <1x13 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0442	0.9935	0.1491	0.9825	130.9935	56.9452	7.5750	15.9468	2.0085e+03	1.0000	2.2999	-0.2821	255
2													

Table 2. Features Extracted from GLCM Technique for Pneumonia

Variables - test_features
test_features
test_features <1x13 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0442	0.9935	0.1491	0.9825	130.9935	56.9452	7.5750	15.9468	2.0085e+03	1.0000	2.2999	-0.2821	255
2													

D. CLASSIFICATION

In this proposed work the classification has been done by using hybrid supervised algorithms. The two supervised algorithms are used for classification purpose which are Support Vector Machine (SVM) and Artificial Neural Network of Back-Propagation Neural Network (ANN-BPNN). The combined feature of SVM and ANN has been named as HSA-BPNN which is used for classifying Chest X-Ray image as normal or Pneumonia. Based on statistical learning and the notion of structural risk minimization, SVM can perform pattern recognition, classification, and prediction. Vapnik and colleagues at AT&T Bell Laboratories invented the SVM to find the best hyperplane (w_0+x+b_0=0 that separates a set of training data ($x1$, $y1$), (xn, yn), Every xn is an actual pattern, and yn is either 1 or -1, with a maximum margin. When yn=1, the genuine pattern belongs to class 1, and when yn=1, it belongs to class 2.

The real pattern is linearly separable if,.

$$\vec{w} \cdot \vec{x}_i + b \geq 1, \ \text{if } y_i = 1$$

$$\vec{w} \cdot \vec{x}_i + b \leq -1, \ \text{if } y_i = -1$$

The inequality can be represented as $(w \cdot xi + b) \geq 1, i = 1 \ldots n$. Minimizing |w| subject to $(w \cdot xi + b) \geq 1, i = 1 \ldots n$. yields the ideal hyperplane. As a result, the linear classifier for the ideal hyper plane is:

$$\vec{x} \rightarrow sign(w \cdot x_i + b)$$

SVM algorithm has been used for finding the best forecast model for classification. Here, the features from region of interest image has been obtained and feed them into SVM classification. The SVM can receive that all features and their labels for training purpose. First, Initializing the SVM for training purpose which contain thirteen numbers of features and their related labels. Initialize SVM trained data and their label. Repeat this for class 'Pneumonia' and 'Normal'. Append feature value and property value for training the SVM. The concatenation of feature data values and property label for both the classes of normal and pneumonia. Finally, start the training process for SVM with kernel type such as linear/polynomial. Here, the SVM property selection has been done and it transfer to back propagation neural network. ANN is based on human brain neuron networks. A layer is a cluster of neurons. An ANN has three layers: an input layer, a hidden layer, and an output layer. The input layer and the output layer are connected by an activation function in the hidden layer. In the input layer, each node Xi has a weight, Wi. The sum of each node's product with its weight is referred to as a net$\sum_i^n x_i w_i$. The net is sent into an activation function, which produces the following prediction or classification:

$$Y_i = f(\sum W_i X_i)$$

In this proposed work, the selected attributes or features of SVM and their related labels have been taken as input. The steps include for training the ANN include,

1. Initialize the back-propagation neural network

2. Setting the propagation iteration and number of hidden layer count.

3. Finally, setting the gradient value and training the back-propagation neural network with SVM selected features and their labels.

4. End the training process.

Therefore, the classification performed by HSA-BPNN, so the hybrid method of SVM and ANN has been given the classification result for pneumonia and normal lung. Thus, the classification output shown below in figure5 and figure6 for pneumonia and normal lung.

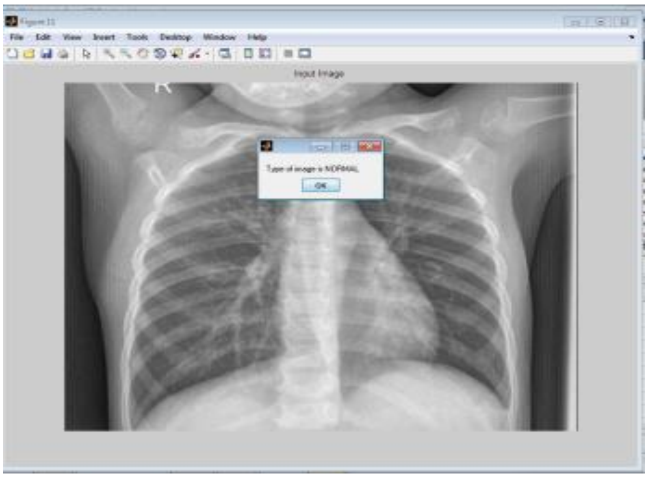

Fig. 5. Classification Results for Normal Lung

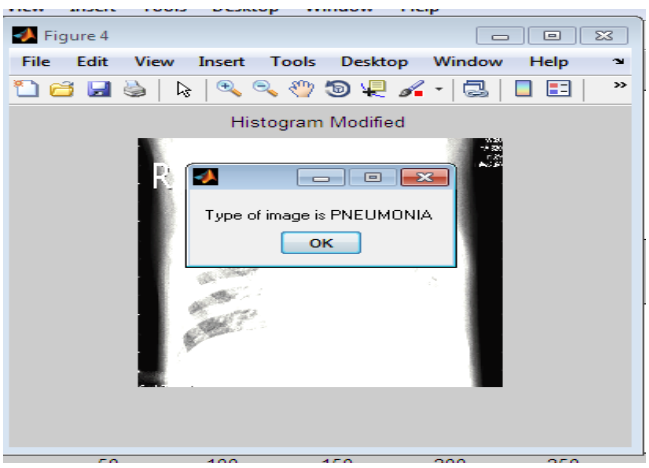

Fig. 6. Classification Results for Pneumonia

III. RESULTS AND DISCUSSION

The performance has been calculated for HSA-BPNN classification model by using three parameters which are accuracy, sensitivity and specificity. Hence, based on this parametric calculation, the proposed algorithm has been compared with an existing algorithm of Support Vector

Machine (SVM), Linear Discriminant Analysis (LDA), and Artificial Neural Network (ANN). The comparison result for accuracy, specificity and sensitivity of HSA-BPNN has been shown in Table3, Table4, Table5 and figure7, Figure8, Figure9.

$$\text{Accuracy} = \frac{\textit{Number of all correct predictions}}{\textit{Total number of dataset}}$$

$$\text{Specificity} = \frac{\textit{Correct negative predictions}}{\textit{Total number of negatives}}$$

$$\text{Sensitivity} = \frac{\textit{Correct positive predictions}}{\textit{Total Number of Positives}}$$

Table 3. Comparison of HAS-BPNN accuracy with existing methods

Methods	Accuracy
SVM	78.08%
LDA	84%
ANN	87.5%
HSA-BPNN	96.4

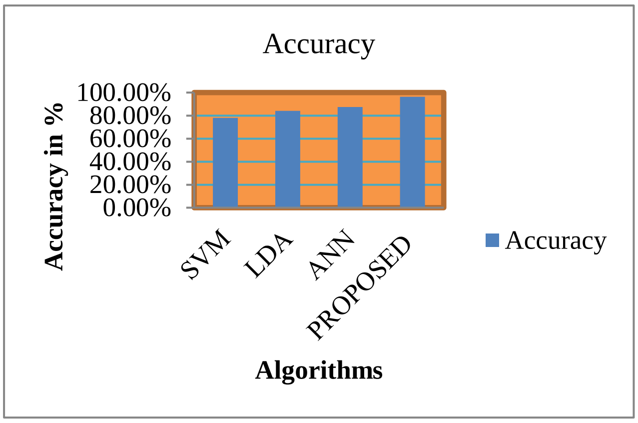

Fig. 7. Accuracy of HSA-BPNN

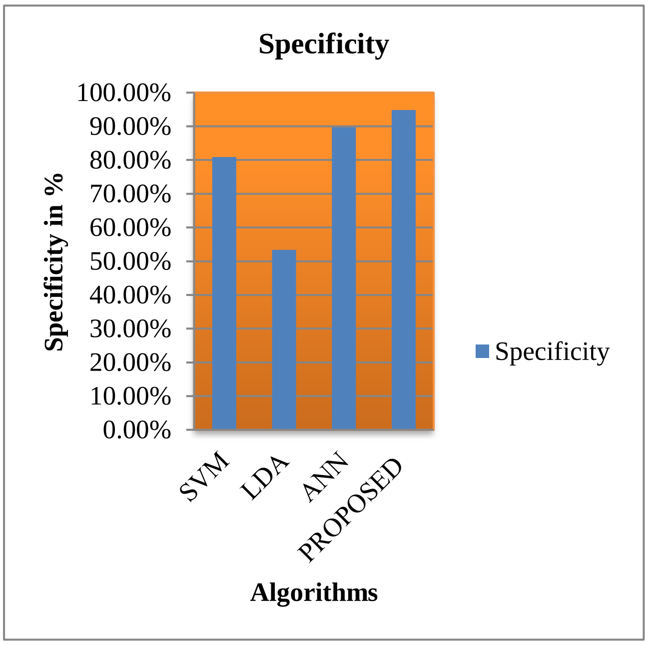

Fig. 8. Specificity of HSA-BPNN

Table 4. Comparison of HAS-BPNN specificity with existing methods

Methods	Specificity
SVM	**80.92%**
LDA	**53.33%**
ANN	**89.75%**
HSA-BPNN	**94.85%**

Table 5. Comparison of HAS-BPNN sensitivity with existing methods

Methods	Sensitivity
SVM	**84.93%**
LDA	**97.14%**
ANN	**97.75%**
HSA-BPNN	**98.40%**

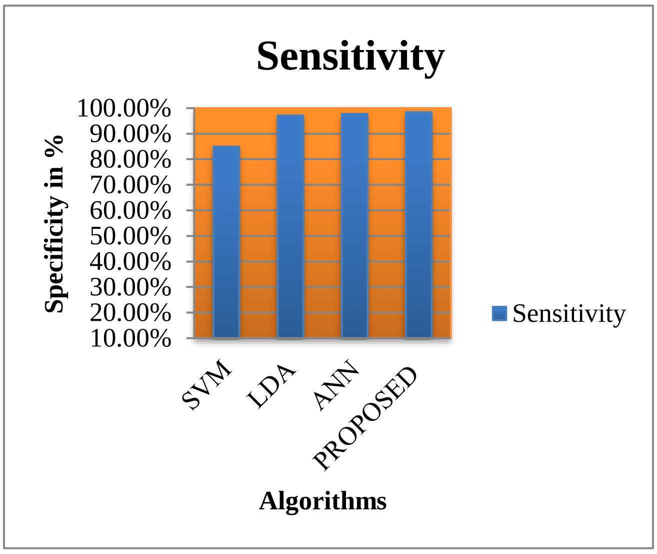

Fig. 9. Sensitivity of HSA-BPNN

Therefore, based on this parametric calculation, the HSA-BPNN has efficient results than the existing methods.

IV. CONCLUSION

The hybrid method of HSA-BPNN model has been developed for classifying the pneumonia disease. For this classification, first the pre-processing has been done by using Weiner filter on Chest X-Ray images. Then, the segmentation has been done by using Active Contour Integration on pre-processed image. This segmentation process has segregated the region of interest from the pre-processed image which is used for analyzing the image or finding the features for further processing. The feature extraction has been performed with GLCM technique and extracts 13 features from the segmented result. Finally, the classification of pneumonia and normal stage of the lung has been done by using hybrid support vector machine and artificial neural network of back-propagation neural network. The hybrid method has given the classification result for Accuracy is 96.4, Specificity is 94.85% and Sensitivity is 98.40%. Therefore, the hybrid model of HSA-BPNN is effective method for early detection of pneumonia which is helpful to increasing the patient's life time.

References

[1] https://www.who.int/news-room/fact-sheets/detail/pneumonia

[2] Jaeger S, Candemir S, Antani S, Wng YXJ, Lu PX, Thoma G, "Two public chest x-ray datasets for computer-aided screening of pulmonary diseases," Quant Imaging Med Surg 2014:4. 2014.

[3] P. Huang, S. Park, R. Yan et al., "Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study," Radiology, vol. 286, no. 1, pp. 286–295, 2017.

[4] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks," Radiology, vol. 284, no. 2, pp. 574–582, 2017.

[5] Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al.,"Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," arXiv , arXiv:1711.05225, 2017.

[6] Li, B.; Kang, G.; Cheng, K.; Zhang, N., "Attention-Guided Convolutional Neural Network for Detecting Pneumonia on Chest X-Rays," In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27; pp. 4851–4854, July 2019.

[7] Nahid, A.-A.; Sikder, N.; Bairagi, A.K.; Razzaque, A.; Masud, M.; Kouzani, A.Z.; Mahmud, M.A.P.,"A Novel Method to Identify Pneumonia through Analyzing Chest Radiographs Employing a Multichannel Convolutional Neural Network," Sensors 2020, 20, 3482. [CrossRef] [PubMed], 2020.

[8] Ayan, E.; Unver, H.M.,"Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning," In Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 24–26,; pp. 1–5, April 2019.