Feature selection and Disease identification in CT images based on particle swarm optimization

J. Nithiya 1, A. Saleema Banu 2, S. Yoga Lakshmi 3, E. Mallika 4

1, 2, 3 UG Scholar and 4 asst. professor CSE.
1, 2, 3, 4 P.S.R. Rengasamy college of engineering for women.

Abstract- The process of Modifying, enhancing or analyzing the properties of the image is called image processing and it is widely used in the computer vision process. For Disease identification selecting the best features reduces the time complexity and algorithm complexity of the process. In image features were are extracted based on B-HOG features, Wavelet features, LBP features and CVH features. Particle Swarm Optimization algorithm used for feature selection process. There are five different classification to test the efficiency of the selected features. The overall performance of the process were measured based on the performance metrics.

1. INTRODUCTION
Image is defined as a 2D function \( f(x,y) \) and giving out is application, pixel are small individual elements of a digital image, each and every pixel has a particular place and brightness or intensity value. Two types of image processing are one is analog image processing and another one is digital image processing. The important needs for DIP are to improve the pictorial information for human interpretation and to process image data for storage, transmission and representation.

Histogram is a graphical representation and plots the number of pixel.

Support vector machine is use the training data to create the optimal separating hyperplane connecting the classes. The optimal hyperplane maximizes the margin of the closest data points.

The process of modifying enhancing or analyzing the properties of the images is called image processing. It is widely used in the computer vision process. Image processing approaches we based on medical images called as Bio-medical images. In this medical images were taken as the input. From that medical image the disease in the images can be identified. The process of finding the location diseased portion is called as segmentation process.

2. PROBLEM STATEMENT
The ROI region in the lung CT images were obtained and the features were extracted from those regions. There are four features are used B-HOG feature, wavelet feature, LBP feature and CVH features used to extracted from images. To identify the diseases in the images by using the combination of spectral and the spatial features. The extracted features were very large in numbers and hence the best features were selected from the
The algorithm used in this system is Fisher criterion. It is used in the place of the objective functions in a generic optimization algorithm.

3.RELATED WORK

To provide useful insights for neural network application in medical imaging and computational intelligence, we structure the rest of this paper in six further sections, where section provides some basics about neural network to enable beginners to understand the structure, the connection and the neuron functionalities[1]. Non-rigid registration of medical images can also be used to register a patient’s data to an anatomical atlas[1]. Neural networks are notoriously hard to interpret and analyses, with in situations where it is desirable to simply and concisely define the process transforming inputs to output values it preserve be difficult to justify their use[1]. Tuberculosis (TB) is very dangerous and rapidly spread disease in the world. into the investigating cases for suspected tuberculosis (TB), chest radiography is not only the key techniques of diagnosis based on the medical imaging but also the diagnostic radiology[2]. Fuzzy-ART neural network integrate advantages of fuzzy logic operators and the basic characteristics of the implementations ART[2]. Geometric feature learning methods not only can be solved the recognition troubles other than also can be predicted the subsequent actions which analyze the sequential input sensory images set[2]. Identification and description of diffuse parenchyma lung disease (DPLD) patterns is very difficult so an automated scheme for volumetric quantification of interstitial pneumonia (IP) pattern, a subset of DPLD, is presented, utilizing a multi detector CT (MDCT) dataset[3]. It gives deep understanding of the Co-occurrence feature extraction[3]. The most important disadvantage of standard FCM algorithm is it deals with image as separate points because the fuzzy function does not consider the spatial dependence[3]. We propose a locality constrained linear coding based approach for classifying lung descriptions in a computed tomography (CT) image set. This method can be utilize to assess the tissue patterns in CT lung images, and thus assisting the diagnosis of pulmonary emphysema[4]. It can be used to assess the pathological extent of emphysema[4]. Unlike these methods, we combine texture features with feature representation which captures spatial distribution[4]. Into this document, we propose using texton signatures based on raw pixel representation along with a parallel multiple classifier system for the classifier of emphysema in computed tomography images of the lung[5]. Irrespective of the image used to describe local image information, the text on - based approach consists of learning and classification stages[5]. This explains why combining them do not improve the performance[5]. In this paper, the registration problem for

Infrared visible stereo pairs, the properties of infrared and visible images that make them mostly similar near boundaries, propose a technique to extract keypoints on the boundary and on the skeleton of region of interest (ROI)[6]. Key points are matched using our future metrics, and the matching pairs are used to calculate the homography[6]. Skeleton keypoints are lower into number, and unfortunately do
not give a small registration error as we would have wished[6].

R. K. Samanta#1 and Malay Mitra[7]. Intelligent automated decision support systems have been attempt with varying accuracies for early detection of breast disease. One of the important tools in this context be neural network. Various feature selection techniques have been deployed as prerequisite[7]. Validation set is a division of our data used to tune the network topology or network parameters other than weights[7]. We could not compare this result as there is no such data available from the literature of this kind[7].

3. PROPOSED WORK

The ultimate goal of this project is to select the optimal features from the different type of features extracted using different methods. The objective functions were initially defined and provided the initial position and velocity. The generated value for the objective function is the local best of the process and this process is repeated in iterations till the global best solution is obtained. The performance of the process were measured based on the performance metrics like Accuracy, sensitivity and specificity. The performance measure proves that the proposed feature extraction more efficient while applying the feature using SVM classifier. This paper used is mainly used to identify the diseases.

![Architecture Diagram](image)

**Fig.1.ARCHITECTURE DIAGRAM**

**PREPROCESSING**

Preprocessing images commonly involves removing low frequency background noise. To collect the CT images on different person in different descriptions, each images are collected the data set and choose the any one image to perform preprocessing.

**STATISTICAL FEATURE**

B-HOG Features are surface features extracted from the images. The images were divided into small regions. Histogram values were calculated for the divided regions and the gradient information’s were obtained from the histogram values. The code words were generated for the extracted values based on the clustering using GMM model and PDF estimation. Using B-HOG 36 features were extracted.
TEXTURE FEATURE

The wavelet based features were extracted by splitting images using wavelets in for levels. The mean, standard deviation and variance value of the each identified wavelet coefficients were calculated. The extracted values were then saved as features. Using wavelet features 18 features were extracted.

FEATURE SELECTION

The images were divided into small patches. The center pixel is identified in the patches. If the pixels were greater than the center pixel then value is assigned as 1 otherwise the value is assigned as 0. The histogram is calculated for the generated LBP codes and then the calculated histogram values were normalized. Using LBP Features 96 features were extracted.

CLASSIFICATION

The histogram of the images gives the intensity plot of the images. The values of the pixels in the images were identified as the intensity of the images. The histogram of the images was calculated and they are saved as the features. Using CVH features 40 features were extracted.

Nine disease is used
1. grand grass opacity (GGO).
2. lobulation.
3. cavity and vacuolous (CV)
4. spiculation.
5. pleural Indentation (PI).
6. Obstructive pneumonia (OP).
7. Calcification.
8. AirBronchogram (AB).

PERFORMANCE

The base paper of this project is “recognize common CT imaging signs of lung syndrome through a original feature selection method based on fisher criterion and general optimization”. The author used fisher criterion and genetic optimization to identify the disease but it is not accurate so that we use particle swarm optimization in that it identify the disease accurately using value.

CONCLUSION

The ultimate goal of the project to select the best feature from different type of feature extracted using different method. To compare the performance of different classifier based on the selected features. Optimize the data in an efficient manner based on particle group optimization. So this project used to identify the accurate disease.

REFERENCES

[5] Meng Yil, Tatyana Nuzhnaya 1, Vasileios Megalooikono, Longi


