Early Diagnosis Of Parkinson’s Disease In Elderly People Using Effective Expectation Maximization Method

M. Anitha, P.Tamije Selvy

Abstract: Nowadays, data mining techniques are applied in the healthcare sector which aids the medical experts in early diagnosis of diseases. This study aims to provide the effective machine learning and optimization techniques for diagnosing Parkinson’s disease (PD). As Parkinson’s disease affects brain’s neuromelanin of millions and millions of people, various soft computing techniques have emerged to help in early diagnosis of this deadly disease. This paper includes implementation of following steps: 1. Dataset Collection, 2. Preprocessing (Noise removal using Median Filter), 3. Segmentation (Expectation and maximization) and 4. Clustering (Fuzzy c-means clustering). Performance analysis is done based on sensitivity, specificity and accuracy. The experimental results resulted in significant improvement in diagnosing PD in early stage.

Index Terms: Brainstem, Fuzzy c-Means Clustering, Median filter, Neuromelanin, Parkinson’s disease, Particle Swarm Optimization

1. INTRODUCTION

The deepest part of the brain consists of basal ganglia, which constitutes caudate nucleus, putamen, globus pallidus, substantia niagra and subthalamic nucleus. Neuromelanin is dark-brown pigment which gives distinct color the brain cells of substantia niagra. The substantia niagra is part of basal ganglia which produces dopamine. Dopamine plays the major role in movement control by sending the signals to striatum. It has been proved that the symptoms of PD develop gradually. If the neurons get degenerated less dopamine is produced which does not provide movement precision and it impair function in brain regions. In general diagnosis of PD is done 30% of dopamine neurons of Substantia nigra have died already Fig 1(a) and 1(b) shows dopamine concentration in substantia nigra of brain. Positron Emission Tomography (PET) scan works by using radioactive tracers which functions by travelling into the brain and other parts of the body. These radioactive tracers are used to detect amyloid plaques. Plaques are nothing but abnormal cell clumps. PET scan is highly suggested for early diagnosis, identifying anatomical changes accurately in human brain. Fig 2(a) and 2(b) depicts the neuromelanin changes in PET scan. In addition it provides higher sensitivity compared to other brain imaging techniques.

Fig. 1. (a) Normal brain control indicating dopamine in substantia nigra and (b) Parkinsons disease affected brain

2 LITERATURE SURVEY

Christodoulou et al. proposed neuro fuzzy model to classify MRI brain images. Drawbacks of fuzzy system are overcome by neuro fuzzy model. As it incorporates both Artificial Neural Network (ANN) and fuzzy this neuro fuzzy model overcomes the fuzzy logic technique in terms of accuracy and convergence rate [1]. Ene M developed a probabilistic neural network model for the purpose of classification on Parkinson’s Disease. PNN consists of three search modes [2]. They are 1. Incremental Search 2. Monte Carlo Search and 3. Hybrid search. He finally proved that Hybrid search provided an accuracy of 81% compared to other search techniques. Rouzabahani et al suggested various brain imaging techniques to find the early symptoms of PD. Earlier, Parkinson’s causes are unknown. Recent researches prove that there are multiple factors that led to the cause of PD. In general, doctors rule out various causes when symptoms are elaborated briefly. [3] But they usually find it difficult to come to a conclusion. It has been suggested that PET brain imaging technique clearly shows the brain deformities compared to other MRI or CT. Indira Rustempasic [4] proposed FCM and pattern recognition method for early and accurate diagnosis of Parkinson’s disease using speech signals. Performance analysis is done in terms of sensitivity, specificity and accuracy. 80.88% is true positive predictive value achieved by FCM and pattern recognition. Sweety et al proposed earlier and accurate detection of Alzheimer’s Disease in three phases. Three phases include 1. Feature extraction, 2. Feature reduction and 3. Classification. Eight features are extracted which are further reduced using Particle Swarm Optimization. Then Decision

MR images with less or no dopamine.

(a) Healthy brain and (b) Parkinson’s brain

Fig. 2. (a) Healthy brain and (b) Parkinson’s brain
tree is used for classification purpose [5]. Performance analysis of classification is done using three parameters namely, sensitivity, specificity and accuracy.

3 PROPOSED METHODOLOGY
This section comprises four subsections. First section will present the preprocessing method, second section will explain segmentation procedure and third section describes the clustering method. Fig 3 demonstrates the framework of proposed method for Parkinson’s disease detection.

3.1 Preprocessing
PET images of brain are obtained for this work as per the protocol. The collected PET images belong to age group of 50 to 75. 208 images are collected which includes normal and PD affected human brain. Preprocessing is done to remove unwanted noise or outliers in medical images. Initially noise in brain images is removed using median filter. This filter is generally used to preserve the edges of the medical images. This effective filtering technique is easy to implement and it has been proved that median filter is less sensitive to outliers compared to mean filtering method [6]. This particular filter is used to remove salt and pepper noise. This impulse type of noise is added to the medical images by adding both random bright (255 pixel value) and random dark (0 pixel value) all over the brain image [7]. In order to increase robustness, the noisy medical image is pre-processed using median filter.

3.2 Segmentation
Medical image segmentation is a critical task [8]. In order to segment brain image automatically, Expectation- Maximization (EM) algorithm is proposed which is a widely used model. EM is an iterative mechanism to estimate the parameters in case of missing data. In EM, cluster representation is done by means of mean, variance and weight. It includes the joint distribution of parameters like color, position and shape features.

Algorithm:
S1 Considering a set of parameters in the presence of missing or incomplete data.
S2 Estimate the values of the missing data using data which are available (E-step)
S3 Parameters are updated once the complete data generated after the S2 step (E-step).
S4 Repeat steps 2 and 3 until convergence.

3.3 Clustering
Fuzzy c-means (FCM) Clustering algorithm which is considered as unsupervised method finds its applications in various fields such as agricultural engineering, image and analysis, medical diagnosis etc,. FCM works by calculating its objective function iteratively [9] as in (1). This can also be defined as least squared function where ‘n’ denotes the total number of datasets, ‘c’ defines the number of classes or partitions and ‘ ’ represents the squared distance, ‘ ‘ denotes a weigh measure of ith cluster.

\[ J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m \| \bar{a}_{ik} \|^2 \]  

Algorithm:
S1 Randomly select ‘c’ cluster centers.
S2 Calculate the fuzzy membership 'u_{ik}':
S3 Compute the fuzzy centers 'v_j':
\[
\nu_j = \left(\sum_{i=1}^{n} (\mu_{ij})^m x_i\right)/\left(\sum_{i=1}^{n} (\mu_{ij})^m\right), \forall j = 1, 2 \ldots c
\]

(3)

where ‘n’ is the number of data points.

S4 Repeat steps S2 and S3 until minimum ‘J’ value is achieved.

![Fig 6. (a) represents the clustered image using FCM and (b) denotes extracted ROI](image)

Brain contains four main lobes namely frontal, parietal, occipital and temporal. Parkinson’s disease involves the symptoms of lesion changes along with dopamine level reduction in cell of midbrain. As the dopamine level changes are accurately visible in upper two lobes the segmented images are clustered using FCM [10]. When FCM is applied false positive regions are removed.

4 PERFORMANCE METRICS

Performance metrics include classification accuracy, sensitivity, specificity, positive predictive value and negative predictive values. The confusion matrix is depicted in table 1 and other parameters are depicted with the following formulas as in (4), (5), (6), (7) and (8).

Accuracy (%) = \(\frac{TP+TN}{(TP+FN+TN+FP)}\times 100\) \hspace{1cm} (4)
Sensitivity (%) = \(\frac{TP}{(TP+FN)}\times 100\) \hspace{1cm} (5)
Specificity (%) = \(\frac{TN}{(FP+TN)}\times 100\) \hspace{1cm} (6)
Positive predictive value (%) = \(\frac{TP}{(TP+FP)}\times 100\) \hspace{1cm} (7)
Negative predictive value (%) = \(\frac{TN}{(TN+FN)}\times 100\) \hspace{1cm} (8)

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>PERFORMANCE METRICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual Value</strong></td>
<td><strong>Predicted Value</strong></td>
</tr>
<tr>
<td>Positive</td>
<td>TP – True Positive</td>
</tr>
<tr>
<td>Negative</td>
<td>FP – False Positive</td>
</tr>
</tbody>
</table>

5 CONCLUSION

Diagnosing neurodegenerative disease like Parkinson’s Disease (PD) by manual interpretation may lead to wrong prediction. To overcome this problem, certain image processing machine learning techniques are proposed. The proposed work included preprocessing, segmentation and effective clustering methods. Preprocessing involves filtering technique by median filter. Salt and pepper noise is removed in PET-brain image. This preprocessed image is segmented using Expectation and Maximization algorithm. Further FCM is applied to the segmented image and the image is clustered effectively which produces best result in terms of sensitivity, specificity and accuracy. Thus proposed algorithms clearly reveal that early and accurate diagnosis is possible.

REFERENCES