A Novel Method for Diagnostic and Prognostic Detection of Alzheimer’s Disease

Antony Jasmine Dinu¹, Ria Manju²

¹Research Scholar, Department of Electronics and Communication Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Thuckalay, Tamil Nadu, India; ²Professor, Department of Electronics and Instrumentation Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Thuckalay, Tamil Nadu, India.

INTRODUCTION

The main aim of the research work is to develop a technique that could diagnose Alzheimer’s disease during its early stage itself from the brain MRI scans of an individual. In our new machine learning algorithm, this is done by extracting and analyzing the various corner features of the MRI scan images, so that the presence of AD can be identified during the initial stage of the disease. Thus its advancement can be controlled before it gets worsened by prescribing appropriate medications. The accurate diagnosis and the treatment strategy of AD are determined by a patient’s current AD stage. AD specialists use the level differences between various biomarkers based on the inflammatory responses and actions in the central nervous system to determine the presence of AD. This is the most acceptable technique to diagnose Alzheimer’s due to its high specificity and sensitivity. But the biomarkers are not always successful predictors especially for the diagnosis of AD during the early stages of the disease. Moreover, the medical team should take utmost care during the collection of cerebrospinal fluid without causing any impairment to the brain tissues and the spinal cord.

Neuroimaging researchers found that patients with AD have wide variations in the white matter when compared to the healthy control groups.¹ Therefore the total volume of the different regions like white matter, grey matter, CSF should be calculated and analyzed by a neurologist to conclude the

ABSTRACT

Introduction: Alzheimer’s is a degenerative neurological disease and is difficult to diagnose with early symptoms. Nowadays, the number of people diagnosed with Alzheimer’s disease is increasing rapidly due to lifestyle changes. Alzheimer’s disease can be diagnosed from MRI brain images. It requires medical expertise and the failure to identify the disease at an early stage will result in permanent disabilities. The automatic and accurate identification of Alzheimer’s disease from MRI images helps to eliminate the above issues and dispense better results.

Objectives: This work intends to develop an algorithm that could detect the presence of Alzheimer’s disease at an early stage by extracting the brain features of MRI images using a point-based feature extraction method.

Methods: Here, a new algorithm is proposed using combined point detection based feature extraction techniques like SURF, FAST, BRISK, Harris, Min Eigen and HOG methods and feature selection using Principal Component Analysis for early prediction of various stages of Alzheimer’s disease. An analysis of the proposed method is done by combining it with different classifiers and the performance parameters are evaluated. The performance of the proposed method is evaluated and analyzed using parameters such as classification accuracy, sensitivity, specificity and F1 score.

Results: From the analysis of the experimental results, the proposed algorithm was found to have a high accuracy rate of 98.62% for the detection and classification of Alzheimer’s disease.

Conclusion: The proposed method was found to be superior to the methods which use single feature extraction which is developed for the prediction and classification of Alzheimer’s disease.

Key Words: Alzheimer’s disease, Mild Cognitive Impairment, Feature Extraction, Feature Selection, Naïve Bayes, Principal Component Analysis.
stage of the disease. In this work, a new machine learning method is developed to find the presence of Alzheimer’s disease at a very early stage using combined point detection based feature extraction methods using SURF, FAST, BRISK, Harris and Min Eigen methods. Here Principal Component Analysis method feature selection method is used for high dimensionality reduction.

**FEATURE EXTRACTION AND MATCHING TECHNIQUES USED**

Feature extraction is the process of reducing the amount of data that is used to represent an image. In the proposed work point-based feature extraction method is used to extract the shape and texture features. The different feature extraction methods used in the proposed method are as follows:

**SURF (Speeded up robust features)**
SURF (Speeded Up Robust Features) is a feature detector that works based on a unique point matching technique that is used to identify and extract the key interest points. The SURF feature detector is faster and its similarity invariant representation makes it more robust to point matching especially for complex MRI images.

**FAST (Features from accelerated segment test)**
FAST (Features from Accelerated Segment Test) corner detector is well known for its high speed, reduced computational time and high efficiency. This method primarily works by reading the first frame followed by the grayscale conversion and afterwards it retrieves a vector of corner locations. This vector is used to draw bright green dots over the corner pixels in the output frame to display the corner locations.

**BRISK (Binary Robust Invariant Scalable Key-points)**
The BRISK attains a scale of key points from the BRISK detector. The FAST score in the scale space pyramid is computed by the BRISK detector and then it will select the scale which has a high score.

**HOG (Histogram of Oriented Gradients)**
This feature extraction technique counts the number of occurrences of the orientation of gradients in the localized areas of an image.

**FEATURE SELECTION TECHNIQUE USED**
The feature selection is used to reduce high dimension feature points and to enhance the efficiency and performance of the classifiers by selecting the relevant features. PCA feature selection method is used in our proposed method to reduce high dimensionality.

**PCA (Principal Component Analysis)**
The PCA is an unsupervised machine learning technique. It uses a statistical method that uses an orthogonal transformation that will transform a list of correlated variables into a list of uncorrelated variables. PCA is a popular tool in exploratory data analysis and machine learning for predictive neural network models. It is used to analyse the interrelations between the variables. The PCA is used when the number of samples is low, and it focuses on determining the component axes that maximize the variance.

**CLASSIFICATION ALGORITHMS USED**
Different powerful machine learning algorithms that extract different features from MRI images for the detection of Alzheimer disease have been developed by Scientists. In this paper, various classification algorithms k Nearest Neighbor and Naive Bayes are incorporated along with the proposed method and each of their performance is compared and analysed.

**k Nearest Neighbor**
In the k-Nearest Neighbor classifier, the nearest neighbours of an instance are specified in terms of Euclidean distance in which all samples match the points in the n-dimensional space Rn. Euclidean distance between the samples

\[ d(X_i, X_j) = \sqrt{\sum_{n=1}^{n} (X_{in} - X_{jn})^2} \]

**Naive Bayes**
The naive Bayes classifier operates on the Bayes probability model. The Naive Bayes classifier integrates the Bayes probability model and a decision rule. A Bayes classifier delegate a class label \( y = C_k \) for some \( k \) having values from 1, 2,…K.
\[ y = \text{argmax} \ p(C_k)^n \prod_{i=1}^{n} p(x_i|C_k)^n \] (2)

**Random Forests**
Random Forests classifier is based on ensemble learning techniques. The works by training on different parts of the same training set followed by averaging multiple deep decision trees for reducing the variance. The predictions for unseen samples is given by

\[ \hat{f}(x') = \frac{1}{B} \sum_{b=1}^{B} f_b(x') \] (3)

**PROPOSED METHOD USING COMBINED FEATURE EXTRACTION AND FEATURE MATCHING TECHNIQUE**
In our work, a new method is developed for the identification and classification of early-stage Alzheimer’s disease using the combined point detection based feature extraction (CPDM) method.

In this method, at first, the noise removal in MRI brain images is done by using the Wiener filtering technique. To maintain appropriate ratios of distances connecting points on a straight line, we have used an affine transformation.

The MRI affected, and suspected brain images are pre-processed for enhancement of the images. The pre-processing is followed by different Feature Extraction like SURF, BRISK, FAST, Harris and MinEigen which extracts different corner points, shapes and Texture features. The match points are obtained by using these points based feature extraction methods. The features of these match points are extracted by using the HOG feature extraction method. The features of the trained images are also extracted by the above method for comparison.

The feature selection is used to extract the relevant features and to enhance the classification rate. The Principal Component Analysis method is used for Feature Selection. The combined features so obtained is finally, sent to different classifiers for classification of different stages of Alzheimer’s disease.

The Algorithm of our proposed method is as follows:
Step 1: Choose X number of training image samples and Y number of testing image samples
Step 2: For each Train and Test images, follow the below steps
   i. Detect the following key interest points from the test and train images
      - SURF points, Ps
      - Brisk Points, Pb
      - Fast Points, Pf
      - Harris Points, Ph
      - Min Eigen Points, Pm
   ii. Compare the detected points from the test and train images
   iii. Identify the matched key points from the test and train images
   iv. Create point pairs
   iv. Select strongest point pairs
   v. Extract HOG features
   v. Apply PCA Feature selection to find the optimal values
Step 3: Create the classifier Model
Step 4: Predict the classification outcome
Step 5: End

The MATLAB toolbox is used for design and programming. The MRI brain images which are used for the study is obtained from ADNI (Alzheimer’s disease Neuroimaging Initiative) Database. The entire workflow of the proposed method using combined feature extraction and feature matching methods is shown in Figure-1.

![Figure 1: Block Diagram of the proposed system.](image-url)
METHOD

In the proposed method, the features of SURF, FAST, BRISK, Harris, Min Eigen methods are extracted and piled up into a corresponding feature vector which consists of the metrics data and location information of the features of the images. The number of features of each method is different and the HOG features are then extracted from these vector points and compared to ensure that the points are matched correctly. The HOG features are extracted in a 1-by-N vector, Here N is the HOG feature-length. MSAC method is lastly used to remove the outliers and to eliminate the mismatches due to Gaussian noise of σ. The value of 1.96σ is used to remove around 5% of outliers (Table 1).

Table 1: Number of match points obtained by using individual and combined feature extraction techniques

<table>
<thead>
<tr>
<th>Match Points of Feature Extraction Techniques</th>
<th>Number of Match Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>AD</td>
</tr>
<tr>
<td>SURF</td>
<td>33</td>
</tr>
<tr>
<td>FAST</td>
<td>27</td>
</tr>
<tr>
<td>HARRIS</td>
<td>21</td>
</tr>
<tr>
<td>BRISK</td>
<td>24</td>
</tr>
<tr>
<td>MIN EIGEN</td>
<td>13</td>
</tr>
<tr>
<td>SURF+FAST</td>
<td>84</td>
</tr>
<tr>
<td>SURF+BRISK</td>
<td>76</td>
</tr>
<tr>
<td>SURF+HARRIS</td>
<td>71</td>
</tr>
<tr>
<td>SURF+MIN EIGEN</td>
<td>62</td>
</tr>
<tr>
<td>FAST+HARRIS</td>
<td>76</td>
</tr>
<tr>
<td>FAST+BRISK</td>
<td>69</td>
</tr>
<tr>
<td>FAST+MIN EIGEN</td>
<td>74</td>
</tr>
<tr>
<td>HARRIS+MIN EIGEN</td>
<td>59</td>
</tr>
<tr>
<td>HARRIS+BRISK</td>
<td>63</td>
</tr>
<tr>
<td>SURF+FAST+HARRIS</td>
<td>152</td>
</tr>
<tr>
<td>SURF+FAST+BRISK</td>
<td>184</td>
</tr>
<tr>
<td>FAST+HARRIS+MIN EIGEN</td>
<td>113</td>
</tr>
<tr>
<td>FAST+HARRIS+MIN EIGEN</td>
<td>171</td>
</tr>
<tr>
<td>HARRIS+MIN EIGEN+BRISK</td>
<td>164</td>
</tr>
<tr>
<td>SURF+FAST+MIN EIGEN</td>
<td>97</td>
</tr>
<tr>
<td>FAST+HARRIS+MIN EIGEN+BRISK</td>
<td>208</td>
</tr>
<tr>
<td>SURF+FAST+HARRIS+MIN EIGEN</td>
<td>244</td>
</tr>
<tr>
<td>SURF+FAST+HARRIS+MIN EIGEN+BRISK</td>
<td>276</td>
</tr>
</tbody>
</table>

Table-1 shows the number of match points obtained by using the single feature extraction technique and the number of match points obtained by using the multiple feature extraction technique. From the results, the number of match points retrieved is more when combined feature extraction techniques are used when compared to single feature extraction techniques. Here the feature extraction techniques such as SURF, FAST, HARRIS and Min Eigen are chosen in such a way that the match points obtained from these techniques are suitable to extract the information from the MRI images.

Let us consider the following number of features retrieved from using the different feature extraction technique

The number of features extracted using SURF is an \( \rightarrow a>0 \)

The number of features extracted using FAST is b \( \rightarrow b>0 \)

The number of features extracted using BRISK is c \( \rightarrow c>0 \)

The number of features extracted using HARRIS is d \( \rightarrow d>0 \)

The number of features extracted using MIN EIGEN is e \( \rightarrow e>0 \)

\[ a+b >a,b \quad \& \& \quad b+c>b,c \quad \& \& \quad c+d>c,d \quad \& \& \quad d+e>d,e \quad \ldots \] (4)

\[ a+b+c >a,b,c \quad \& \& \quad a+c+d >a,c,d \quad \& \& \quad a+d+e >a,d,e \quad \& \& \quad b+c+d >b,c,d \quad \ldots \] (5)

\[ a+b+c+d >a,b,c,d \quad \& \& \quad a+c+d+e >a,c,d,e \quad \& \& \quad d+c+a+e >d,c,a,e \quad \ldots \] (6)

\[ a+b+c+d+e >a,b,c,d,e \quad \ldots \] (7)

These are applicable for \( \forall a,b,c,d,e >0 \)

From the above equation the number of match points retrieved by using combined feature extraction technique is high when compared to the number of match points retrieved by using single feature extraction techniques. Due to this the accuracy of feature extraction is high and which in turn increases the performance and stability of the application. The combined feature extraction module is shown in Figure-2.

Figure 2: Combined feature extraction module.
EXPERIMENTAL RESULTS

In our work, 2000 MRI T1 weighted images are collected and grouped into Normal Cognitive (NC), Mild Cognitive Impairment (MCI) and Alzheimer’s disease (AD).

Number of MRI scans of AD → 1004
Number of MRI scans of MCI → 611
Number of MRI scans of NC → 385

The performance analysis of the proposed method is evaluated based on four parameters such as

1. True Positive (TP)
2. True Negative (TN)
3. False Positive (FP)
4. False Negative (FN)

These parameter values are obtained from the AD, MCI and NC group images present in the datasets. The Classification Accuracy, Sensitivity, Specificity, F1 score and Time Consumption is calculated using the values of TP, TN, FP and FN.

Classification Accuracy
The classification accuracy is the rate of correct predictions. It is given by

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)
\]

Sensitivity
Sensitivity is the proximity of patients with Alzheimer’s disease who test positive.

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (9)
\]

Specificity
Specificity is the proximity of patients without Alzheimer’s disease who test negative.

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (10)
\]

F1 Score
F1 score determines the balance between the precision and the recall.

\[
F_1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (11)
\]

Time Consumption
Time taken to obtain the analysis result is termed as time consumption and it is available from MATLAB results viewer and it is measured in Seconds.\(^\text{14}\)

Figure 3 shows the feature matching technique and Figure 4 shows the feature selection process to retrieve the corresponding strongest match points from brain MRI images. With the combined point-based feature extraction technique, superior performance, robustness and stability are achieved (Table 2 and 3).
Table 2: Average performance parameters values when single feature extraction and feature selection method is used along with different classifiers

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>Number of Samples</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>F1 Score</th>
<th>Execution Time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>k Nearest Neighbor</td>
<td>50</td>
<td>65.57</td>
<td>52.44</td>
<td>54.10</td>
<td>0.510</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>67.43</td>
<td>57.52</td>
<td>58.47</td>
<td>0.512</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>73.27</td>
<td>68.75</td>
<td>70.63</td>
<td>0.523</td>
<td>0.67</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>50</td>
<td>65.43</td>
<td>51.43</td>
<td>53.37</td>
<td>0.510</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>69.41</td>
<td>56.49</td>
<td>59.42</td>
<td>0.512</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>71.42</td>
<td>69.71</td>
<td>71.36</td>
<td>0.519</td>
<td>0.63</td>
</tr>
<tr>
<td>Random Forests</td>
<td>50</td>
<td>67.59</td>
<td>56.13</td>
<td>58.23</td>
<td>0.513</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>69.22</td>
<td>63.46</td>
<td>65.37</td>
<td>0.536</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>76.11</td>
<td>70.37</td>
<td>71.71</td>
<td>0.591</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Table 3: Performance parameters values when proposed combined feature extraction and feature selection method is used along with different classifiers

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>Number of Samples</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>F1 Score</th>
<th>Execution Time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-Nearest Neighbor</td>
<td>50</td>
<td>95.64</td>
<td>66.14</td>
<td>67.44</td>
<td>0.503</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>97.43</td>
<td>73.22</td>
<td>74.27</td>
<td>0.511</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>98.40</td>
<td>86.35</td>
<td>8.78</td>
<td>0.525</td>
<td>0.76</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>50</td>
<td>95.52</td>
<td>63.52</td>
<td>64.57</td>
<td>0.504</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>96.31</td>
<td>69.17</td>
<td>69.37</td>
<td>0.516</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>98.21</td>
<td>85.26</td>
<td>86.21</td>
<td>0.521</td>
<td>0.83</td>
</tr>
<tr>
<td>Random Forest</td>
<td>50</td>
<td>95.1</td>
<td>68.56</td>
<td>69.18</td>
<td>0.570</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>96.18</td>
<td>71.34</td>
<td>72.11</td>
<td>0.590</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>98.62</td>
<td>86.13</td>
<td>87.22</td>
<td>0.613</td>
<td>0.90</td>
</tr>
</tbody>
</table>

DISCUSSIONS

The performance metrics such as classification accuracy, sensitivity, specificity, F1 score and time consumption is calculated using the True Positive, True Negative, False Positive, and False Negative values. Table-2 shows the measurements of the parameters from the output of different classifiers using a single point-based feature extraction method followed by the HOG feature extraction method and PCA feature selection method and Table-3 shows the measurements of the parameters from the output of different classifiers by the newly proposed method using combined feature extraction followed by the feature selection method. It was seen from the investigational results that; the accuracy of classification is found to be 98.62% which is very high when newly proposed combined feature extraction and feature selection method when compared to classifiers which use single feature extraction method. The sensitivity rate obtained by the new method is 86.13% and the specificity rate obtained by the new method is 87.22%. The sensitivity and specificity rates are also very high in our new proposed method. The F1 score is also found to be high in our proposed method. The execution time is found to be high when a newly proposed method is used, but it is more accurate in the diagnosis of AD. Here accuracy of disease prediction is more important than time consumption.

The new method has the advantages of combined multiple points based feature extraction methods and the advantages of combining multiple feature selection methods. Hence the accuracy rate is very high and so it can be used with any classifiers to yield high accuracy rates for classification purposes. The proposed method is found to be superior due to its high accuracy rate and high sensitivity.

The ROC curves are used to evaluate the performance of the classifier model developed over its complete operating range. The AUC (Area Under Curve) of the classifier model is used to evaluate the performance of the model developed. Here we have developed models using different classifiers such as k Nearest Neighbor, Naive Bayes and Random Forest classifiers after integrating with the proposed feature extraction method. The Receiver Operating Characteristic curves (ROC) are obtained for the proposed combined feature extraction method and combined feature selection methods.
Dinu et al: Early detection and classification of Alzheimer's disease

In our proposed model, the AUC of the ROC of the newly developed classifier model is high when the new proposed algorithm is used with different classifiers. The analysis of ROC Curve’s AUC implies that the new proposed model works with high efficiency in the early prediction and classification of various stages of Alzheimer’s disease. The ROC Curves obtained when single feature extraction and feature selection is used and ROC curves when the proposed method is used are shown in Figure-5 and the classification outcome of Alzheimer’s disease is shown in Figure-6.

**CONCLUSION**

In the study, a new method is proposed for the early prediction of Alzheimer’s disease using combined point detection feature extraction methods and combined feature selection methods along with optimization techniques. The different feature extraction methods such as SURF, FAST, BRISK, Harris and Min Eigen methods are combined, and corresponding match points are obtained from Test and Train images. This is followed by the feature extraction of HOG features from the match points. The feature selection is carried out using the Principal Component Analysis method and Linear Discriminant Analysis method. The proposed algorithm is combined with different neural network classifiers like k Nearest Neighbor, Naïve Bayes and Random Forests and the efficiency and performance of each classifier is evaluated. The performance metrics and ROC curves are obtained for each method. The implementation results have been analyzed and compared in terms of classification accuracy, sensitivity, specificity and time consumption. The Area Under the Curve of all the classifier’s ROC’s is high when the new proposed feature algorithm is incorporated in different classifiers. The experimental results indicate that the new proposed approach is most efficient and will deliver promising outcomes for clinical applications.

**ACKNOWLEDGEMENT**

Authors acknowledge the immense help received from the scholars whose articles are cited and included in references of this manuscript. The authors are also grateful to authors/editors/publishers of all those articles, journals and books from where the literature for this article has been reviewed and discussed.

**Source of Funding:** Nil

**Conflicts of Interest:** Nil

**Author Contributions**

Dinu A J: Conceptualization, Methodology, Writing—Original Draft Preparation, Dr. Manju R: Reviewing & Supervision

**REFERENCES**


