VOL. 17, NO. 5, MARCH 2022 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DETECTION OF ALZHEIMER'S DISEASE USING MRI IMAGES BASED ON SVM CLASSIFIER

P. M. K. Prasad[1] and Y. Raghavender Rao[2]
[1]Department of Electronics and Communication Engineering, GVP College of Engineering for Women, Visakhapatnam, Andhra Pradesh, India
[2]Department of Electronics and Communication Engineering, JNTUH College of Engineering Sultanpur, Sangareddy, Telangana, India
E-Mail: pmkp70@gmail.com

ABSTRACT

In the present scenario, most of the people are suffering from memory loss which leads to the cause of Alzheimer's disease. It slowly destroys the brain cells resulting in memory loss affects thinking, language skills and behavioral changes. As the age progresses, people are affected with this disease. So, it is required to detect at an initial stage, so that proper treatment can be given to the patient. There are various methods such as mini mental state examination, HOG and SURF, regional arthopy are used to detect Alzheimer's. But these methods are not reliable. So proposed method effectively detects Alzheimer's disease. This method consists of four steps that is, preprocessing, segmentation, feature extraction and classification. This method extracts various textual features using GLCM. SVM classifier is used to classify these features and it is more superior compared to other classifiers. The proposed SVM classifier based Alzheimer's detection is more superior when compared to the other methods.

Keywords: alzheimer's, feature extraction, GLCM, SVM classifier.

1. INTRODUCTION

Alzheimer's is a brain related disease that causes memory, decline in thinking and behavior. The main problem many people notice is forgetfulness and it affect their ability to work either at the office or at home. Due to this disease, people get confusion and misplace things. It usually stars slowly gets worse over period of time [1]. Presently, around fifty million people got affected due to Alzheimer's and due to population aging, it is expected that this number may increase up to 150 million by 2050 [2]. The diagnosis of Alzheimer's is very complicated due to various behavioral and cognitive symptoms [3]. Recent development and progress in the diagnosis and treatment of Alzhemiers disease has helped many patients. The accurate diagnosis leads to better treatment and it will reduce financial burden. Loss of memory is the important symptom of this disease. Memory impairments deteriorate and other symptoms may develop as the disease progresses. A family member may be likely to observe the symptoms.

Mini mental state examination technique is used to detect Alzheimer's. In this technique, pre-processing stage effects the final classification [4]. It can diagnose Alzheimer's early by filtering most of the noise in the image. But, there are still some limitations in this method. One of the limitation is that the pre-processed image has lost too many features, which could affect final classification results as there will be not be able to extract these features in the feature extraction stage. HOG and SURF descriptors can be used for detection of Alzheimer's due to good performance. This algorithm helps to find matching key points fastly. But here the classification was not accurate and efficient [5]. The regional Atrophy analyses of MRI images are also used for early Alzheimer's detection. The approach catches specific anatomical structures of the brain MRI images to detect the Alzheimer's. It also describes specific anatomically areas which are progressively affected by atrophy such as hippocampus, putamen, globuspallidus, thalamus, and caudate nucleus [6]. This method is not suitable for brain imaging, because it cannot analyze accurate volumetric of the atrophy and lesion.

The proposed method detects Alzheimer's disease accurately. It extracts various features effectively and it uses SVM classifier. It is a supervised learning machine model. SVM is mostly used to classify patterns [7]. The SVM classifier detects the Alzheimer's as Mild Decline stage that is; there are symptoms of Alzheimer's disease. People in this stage will have trouble in finding the exact word during discussions, remembering names.

2. PROPOSED METHOD

A new method for the detection of Alzheimer's disease is shown in Figure-1. It has four stages. Initially the approach starts with preprocessing of MRI images.

2.1 Preprocessing

Preprocessing step generally used to reduce noise and masking portions of images. The MRI image contains Salt and pepper noise. In this step, various filtering techniques are used for blurring and smoothing the images [8]

2.1.1 Median filter

The noise in the image is reduced using median filtering. It reduces the salt and pepper noise. It will be applied on each pixel and its value is replaced by the median. It preserves edges of the images under some conditions while removing the noise. In this filtering, the neighboring pixels are arranged as per the ascending order of the intensities and the centre pixel value is replaced by median value. This filter is very good in removing noise,

without the effects of smoothing that can happen with particular smoothing filters [9]

2.2 Segmentation

Segmentation extracts Region of Interest (ROI) from the image. Segmentation algorithms which are based on similarity and discontinuity properties of pixel values. Region growing works on the basis of similarity property. It is implemented based on region growing to calculate the segments corresponding to white matter and gray matter in order to extract the features, SVM classifier is used to classify the textural features.

2.2.1 Region growing algorithm

It will find out homogeneous regions within the image. In this approach pixels in a region are labeled with a distinctive label. This algorithm is based on the selection of initial pixel or seed point and the similarity criteria [10] [11]. The seed is assigned a label. Now each of the neighbor pixels is labeled recursively using a similarity measure or a homogeneity property. When no more pixels can be labeled with the same label, the first region is found. In the step, another seed pixel is selected out of the yet to be label pixel. The same labeling is followed. The overall process is repeated until no more pixels remain unlabelled. For determining neighbors, the concept of 4-connected or 8-connected pixel windows may be considered.

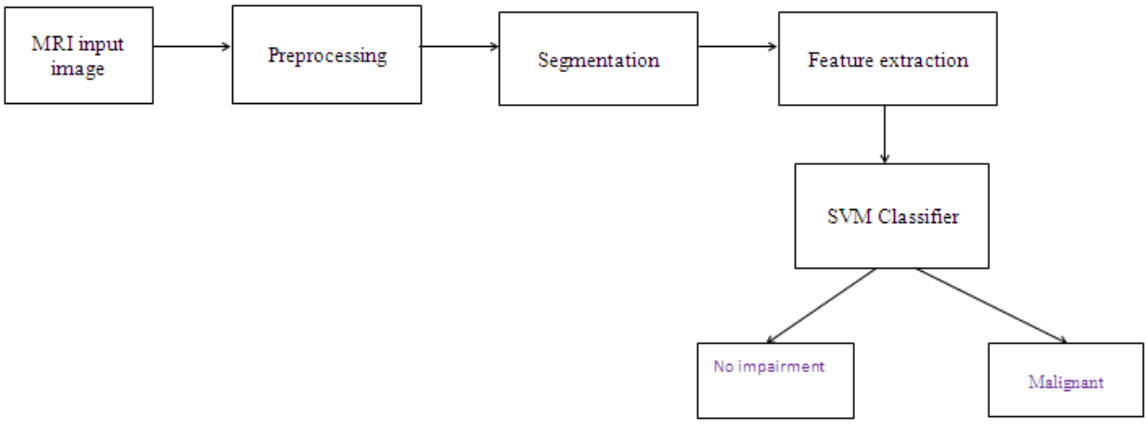

Figure-1. Block Diagram of Alzheimer's detection system.

2.3 Feature Extraction

It converts raw pixel values of an image in to useful data. This step extracts and quantifies some significant characteristics of the object. There are different types of features. But, texture is important feature which is useful for medical image retrieval applications [12]

2.3.1 Textural features based on GLCM

Gray level co-occurrence matrix (GLCM) extracts the textual features and estimates the image properties. The rows and columns of a GLCM are equivalent to grey levels of the image [13]. This method is used to extract information from MRI images [14]. In GLCM, (i,j) represents possible gray levels in the original image. It is obtained by identifying how many times that the ith pixel takes place in the selected spatial relationship to a jth pixel. GLCM can be specified in a matrix of P (i; j; d; θ) with two neighboring texture elements one with property i and the other with property j are separated by distance d at orientation θ. Here texture elements are pixels and properties are gray levels. There are various textural features such as contrast, Homogeneity, Entropy, Mean, Standard deviation, correlation, Energy, Variance, RMS value.

Contrast: High degree of color or gray scale variation is displayed by high contrast images. The contrast can be computed using equation 1.

$$\text{Contrast}= \sum_i \sum_j (i-j)^2 p_{i.j} \tag{1}$$

Homogeneity: It indicates the closeness of the distribution of elements in the GLCM to the GLCM diagonal [15]. The distribution parameter is high for best segmentation. The image is divided into a set of homogeneous texture regions, and then the texture features associated with the regions are indexed in the image data. The Homogeneity can be computed using equation 2.

$$\text{Homogeneity}= \sum_i \sum_j \frac{1}{1+(i-j)^2} p_{i,j} \tag{2}$$

Entropy: It is the significant characteristic used in segmentation method. The Mathematical representation of entropy is given in equation 3.

$$\text{Entropy}= -\sum_i \sum_j p_{i,j} \log_2 p_{i,j} \tag{3}$$

Mean: the contribution of individual pixel intensity of the entire image is indicated by mean. The mean can be computed using equation 4.

$$\text{Mean} = \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

Standard deviation: It describes probability distribution of an observed population. The standard deviation can be computed using equation 5.

$$\text{Standard Deviation: } \sqrt{\sum_{i=0}^{N-1}(i-\mu)^2 p(i)} \tag{5}$$

Correlation: It indicates the spatial dependencies between the pixels. Correlation can be computed using equation 6.
Correlation:

$$\frac{1}{\sigma_x \sigma_y}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i,j)p(i,j)^2 - \sigma_x\sigma_y \tag{6}$$

σx, σy is the standard deviation in the horizontal spatial domain. and vertical domain respectively.

Energy: It is quantifiable amount of the extent of pixel pair repetitions. The energy value is high for highly correlated pixels; the energy value is very large [16]. The energy can be computed using equation 7

$$\text{Energy: } \sum_i \sum_j p(i,j)^2 \tag{7}$$

Variance: It shows the variation in intensity between neighboring pixels. The variance can be computed using equation 8.

$$\text{Variance: } \sigma^2 = \sum_{i=0}^{N-1}(i-\mu)^2 p(i) \tag{8}$$

RMS value: The RMS value can be computed using equation. 9

$$\text{RMS value: } \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|x_i\right|^2} \tag{9}$$

In feature extraction step, feature vector represents various features [17]

2.4 Classification

The proposed method uses supervised classification as it is more superior than unsupervised classification. There are many techniques of supervised classification for assigning pixels to informational classes, Minimum Distance from Mean Parallelepiped, Maximum Likelihood, Support Vector Machine (SVM), Artificial Neural Networks. But, SVM is superior compared to other classifiers.

2.4.1 Support vector machine (SVM)

SVM is a machine learning classification algorithm and is used for solving binary classification problems [18]. It draws a decision boundary also known as a hyperplane [19]. The data in the original input space can be linear separable in a higher-dimensional feature space [20].SVM is superior to other supervised learning methods.

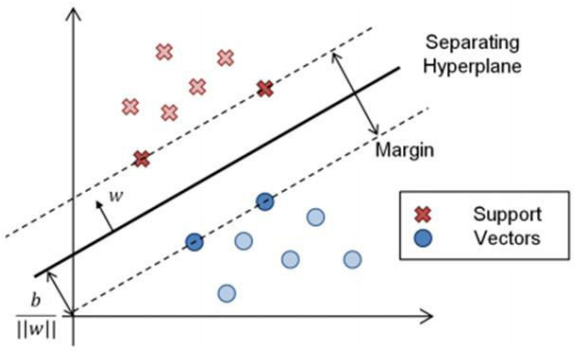

Figure-2. Separation of hyper plane in SVM.

In this method, training samples of one group are forced to be on one side of the hyper plane and the samples of the other group are forced to be on the opposite side of the hyperplane. This problem can be solved with the help of various non-linear programming techniques that results in many active constraints. These constraints are very close to the two groups. These are called "support vectors". The rest of the training samples do not contribute for the division of hyper plane. Due to this feature SVM is considered as an effective classifier. The hyper plane is determined only by a relatively small number of samples that are close to the opposite group. There is no influence on the results when the samples are far away. The samples that are far away have no influence on the results. The classifier focuses on the refinements of the morphological differences between the two groups and not on gross differences. Therefore, it is more effective.

Consider a linearly separable problem where the training set D is formed by K instances, each one belonging to one of two classes

$$D = \{(x_1, y_2), \ldots, (x_k, y_k)\} x_k \in R^n, y_k \in \{-1,1\} \tag{11}$$

Where $x_k$one instance of the N-dimensional feature vector is built by one of the selection methods. It should be noticed that the iterate k for the training samples is here given in subscript, instead of in superscript for presentation purposes. A given hyper plane, which can be parameterized by its normal vector $w$and a constant b

$$w \,.\, x + b = 0 \tag{12}$$

is said to be a separating hyper planeif and only if the decision function given by

$$f(x) = sign(w\,.\,x + b) \tag{13}$$

correctly classifies all training instances, i.e., if $f(x_k)_= y_k$for all instances k, which can also be compactly rewritten in the following way

$$\mathrm{y_k} = (\mathrm{w}\,.\,\mathrm{x_k} + \mathrm{b}) \geq \ 1 \quad \forall \mathrm{k} \tag{14}$$

Two details should be noticed in the previous inequality. First, the constant on the right-hand side could actually be any strictly positive number due to the fact that any hyper plane represented by (w, b) can also be represented by any scaled pair $(\lambda\,\mathrm{w}, \lambda b)$ with $\lambda \in \mathrm{R}^+$. The second and most important detail is that any separating hyper plane can be represented in such a way that equation 14 is met as equality for the nearest training samples by changing the scaling factor λ. Now, in order to select the best hyper plane from the infinite set of separating ones, its margin should be maximized. Bearing in mind that the distance between the hyper plane and the nearest vectors is given by:

$$\mathrm{d}\big((\mathrm{w}, \mathrm{b}), \mathrm{x_k}\big) = \frac{\mathrm{y_k}(\mathrm{w}\,.\,\mathrm{x_k}+\mathrm{b})}{\|\mathrm{w}\|} \tag{15}$$

$$= \frac{1}{\|\mathrm{w}\|} \tag{16}$$

one can conclude that the optimal hyper plane can be obtained by minimizing $\|w\|$ or equivalently $\frac{1}{2}\|w\|^2 = \frac{1}{2}w^T w$, under the constraints eq.14, i.e., by minimizing the following convex quadratic programming problem:

$$\underset{\mathrm{w,b}}{\text{minimise}}\,\frac{1}{2}\mathrm{w^T w}$$

$$\text{subject to} \quad y_k(w\,.\,x_k + b) \geq 1 \quad \forall k \tag{17}$$

The main method to solve this problem is based on its Lagrangian dual formulation which can be obtained as follows. First, consider the Lagrangian function associated with problem

$$\mathrm{L(w, b, \Delta)} = \frac{1}{2}\mathrm{w^T w} - [\mathrm{y_k}(\mathrm{w^T}\,.\,\mathrm{x_k} + \mathrm{b}) - 1] \tag{18}$$

where$\Delta = (\alpha_{1,\ldots,}\alpha_k)$ is the vector of non-negative Lagrangian multipliers associated with constraints eq.14. Then, the infimum of $L(w, b, \Delta)$ with respect to w and b, quantity that the dual problem maximizes, can be determined using $\nabla_{w,b}L(w, b, \Delta)$ and plugging the results in to equation 18. Specifically, the conditions$\frac{\delta L(w,b,\Delta)}{\delta w} = 0$ and $\frac{\delta L(w,b,\Delta)}{\delta b} = 0$yield:

$$w = \sum_{k=1}^{k} \alpha_k y_k x_k \tag{19}$$

$$\sum_{k=1}^{k} \alpha_k y_k = 0 \tag{20}$$

respectively, and after some manipulation, the infimum of L can be given by:

$$\inf_{\mathrm{w,b}}\{\mathrm{L(w, b, \Delta)}\} = \sum_{\mathrm{k=1}}^{\mathrm{k}} \alpha_{\mathrm{k}} - \frac{1}{2}\sum_{\mathrm{k=1}}^{\mathrm{k}}\sum_{\mathrm{k=1}}^{\mathrm{k}} \alpha_{\mathrm{k}}\,\alpha_1 \mathrm{y_k y_1}(\mathrm{x_k^T x_l}) \tag{21}$$

Dual maximization problem for the Lagrangian coefficients, which can be stated, using vector notation, in the following form:

$$\begin{aligned} &\underset{\Delta}{\text{maximize}}\ \Delta^T 1 - \frac{1}{2}\Delta^T D \Delta \\ &\text{Subject to} \quad \Delta \geq 0 \\ &\Delta^T y = 0 \end{aligned} \tag{22}$$

Where 1= (1,…,1) is a K-dimensional unit vector $y = (y_1, \ldots., y_k)$ is the vector of labels and D is the scalar matrix $D_{kl} = y_k y_l x_k^T x_k$ and $l \in \{1, \ldots, k\}$ the dual problem is still a quadratic problem but instead of scaling with the number of dimensions in the feature space, it scales with the number of scaling instances. In addition one can notice that each input vector x_k always appears in a dot product with some other vectors x_l a property whose usefulness become clear later.

Finally, from optimization theory, more specifically from the complementary slackness condition of the Karush-Kuhn-Tucker theorem, one can conclude that, when the solution to the problem eq.22 is met, one of two possible situations holds true. If a given instance $x_k$is a support vector, then the associated Lagrangian multiplier $\alpha_k$is non-negative, otherwise, α_k is zero. Consequently, the optimal hyper plane can be constructed as a linear combination of the support vectors using equation. 19. In addition, the bias b can be found from the constraints. 14 for the support vectors, since they are met as equalities for such training instances.

Following the historical evolution, the first extension of this concept was the introduction of *kernels*. As mentioned before, this extension is motivated by the fact that even if one dataset is not linearly separable, it can be separated by a nonlinear separation surface. The introduction of kernels will solve this problem by using a mapping $z = \varphi(x)$ that transforms the original N - dimensional input space into a new N - dimensional feature space, where an hyper plane will try to separate the new transformed data $\{(\varphi(x_k), y_k)\}$.

Two complications can be identified at this stage. The first is computational and is associated with the dimension of the feature space, which can even be infinite for some types of kernels, precluding a naive straightforward approach. However, after replacing all occurrences of $x$by φ, one can see that each time a given $\varphi(x_k)$ appears, is in a dot product with another $\varphi(x_l)$. As a consequence, one only needs to define that the inner product in the feature space, without having to explicitly compute the solution is then obtained by solving the mapping of the input vectors. Specifically, the elements of

matrix D of the dual problem equation 22 become $D_{kl} = y_k y_l(\varphi(x_k).\varphi(x_l))$

$$w = \sum_{k=1}^{k} \alpha_k y_k \varphi(x_k) \quad (23)$$

and then replacing into w yeilds to the termas decision function

$$f(x) = sign \sum_{k=1}^{k} \alpha_k y_k(\varphi(x_k).\varphi(x)) + b \quad (24)$$

The inner products are given by the so called kernel function:

$$k(x_k, x_l) = (\varphi(x_k).\varphi(x_l)) \quad (25)$$

The second problem arises with the choice of the kernel. In low-dimensional spaces, one might be able to imagine a mapping function that would separate the training set in the feature space, but this task becomes impossible when the dimension of the input space starts to grow. Fortunately, several kinds of kernels with good generalization ability were already identified. The RBF kernel, defined in equation. 24, is one of them and it has been extensively used for the CAD of Alzhemier's Detection.

$$k(x_k, x_l) = exp\{-\gamma\|x_k - x_l\|^2\} \quad (26)$$

So far, the training data was assumed to be separable either on the input space or in the feature space. However, if the data cannot be fully separated without committing a small number of errors, the dual problem becomes unbounded and no solution can be found.

In addition, the use of complex kernels to separate the data often leads to poorer classification performance. To handle this problem Cortes and Vapnik relaxed the constraints in equation. 14, introducing a positive slack variable $\varepsilon_k$in each one:

$$y_k(w.x_k) \geq 1 - \varepsilon_k \quad (27)$$

Those slack variables, which quantify the errors committed by each vector, were then weighted in the cost function in order to keep them as small as possible. The new primal problem can therefore be stated as follows:

$$\underset{w,b}{\text{minimise}} \frac{1}{2} w^T w + c \sum_{k=1}^{k} \varepsilon_k$$
$$\text{subject to} y_k(w.x_k) \geq 1 - \varepsilon_k \; \forall k \quad (28)$$

where C is a tuning parameter that controls the cost of misclassification. This optimization problem, which is still convex and quadratic as the original one, is often solved by exploiting its dual representation, which can be obtained following reasoning similar to the separable case. At the end, the dual problem takes the form:

$$\underset{\Delta}{\text{maximise}} \, \Delta^T 1 - \frac{1}{2} \Delta^T D \Delta$$
$$\text{subject to } 0 \leq \Delta \leq C \quad (29)$$
$$\Delta^T y = 0$$

Therefore the mathematics for the implementation of the hard margin SVM has been implemented by the separation of hyper plane in the classification technique.

3. RESULTS AND ANALYSIS

Step 1: The brain MRI noisy image is considered as an input image for the detection of Alzheimer's image is shown in Figure-3.

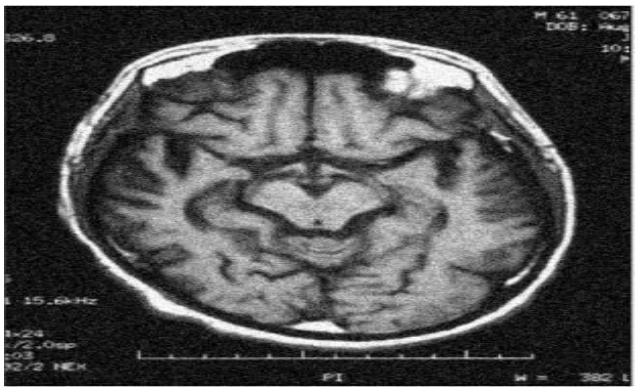

Figure-3. Input image 1.

Step 2: The noise in the MRI image is reduced using Median filtering which is shown in Figure-4.

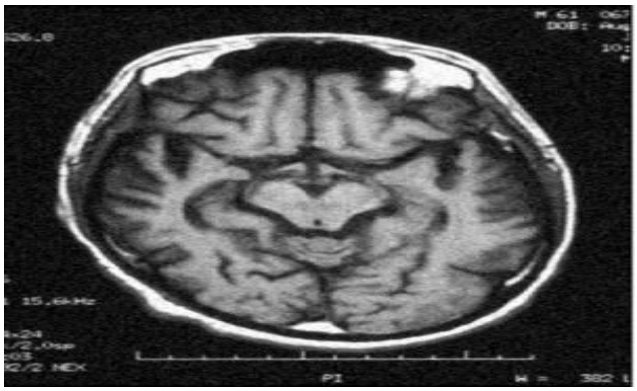

Figure-4. Median filter output image.

Step 3: The noise free image is segmented using region growing into different regions in the image. This method involves the selection of initial seed point. Segmented image using Region growing is shown in Figure-5.

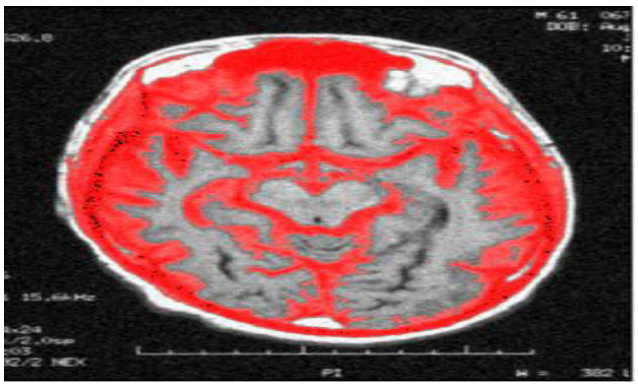

Figure-5. Segmented image.

Step 4: Various textual features are extracted from the segmentation output by using GLCM which determines the spatial dependence of gray level in an image. Table-1 shows feature values for input image 1.

Table-1. Feature values for input image1.

S. No	Feature	Value
1	Contrast	0.0141
2	Correlation	0.0510
3	Energy	0.9734
4	Homogeneity	0.9934
5	Mean	2.29966e-04
6	Standard deviation	0.0370
7	Entropy	2.8273
8	RMS	0.0370
9	Variance	0.0014

Step 5: These extracted features are given to the SVM classifier. SVM Classifier's job is to tell the final decision whether the patient suffers from the Alzheimer's or not.

Output: no impairment: a symptom of Alzheimer's.

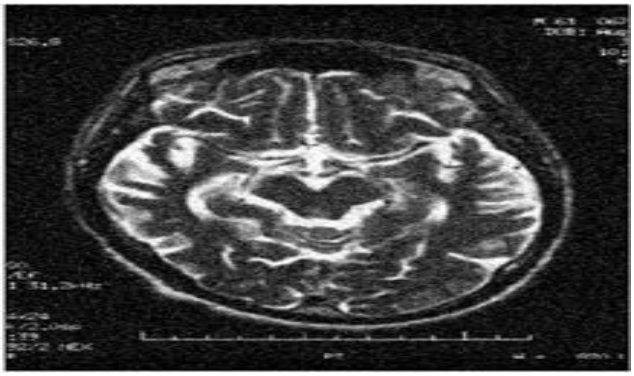

Figure-6. Input image 2.

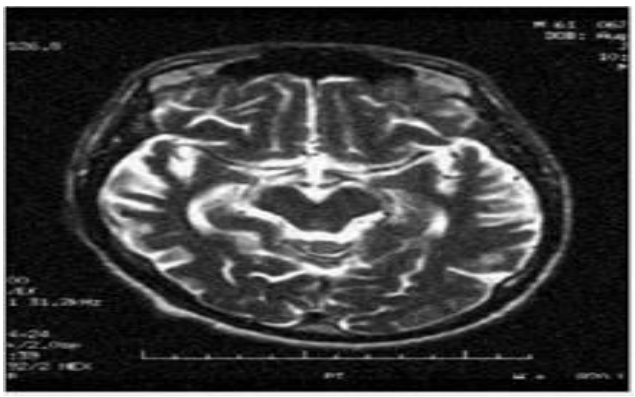

Figure-7. Median filtered image.

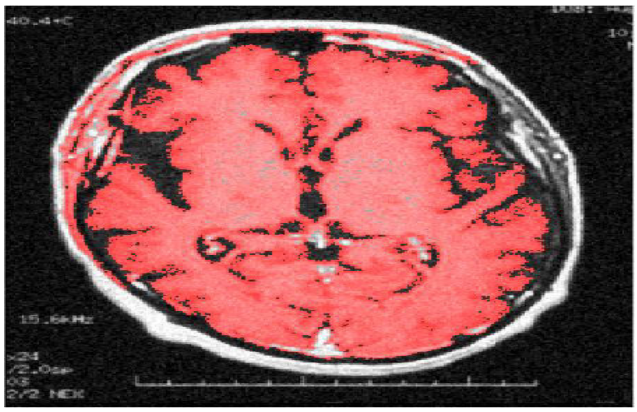

Figure-8. Image segmentation result.

Table-2. Feature values for input image2.

S. No	Feature	Value
1	Contrast	0.0161
2	Correlation	0.0282
3	Energy	0.9693
4	Homogeneity	0.9923
5	Mean	2.4184e-04
6	Standard deviation	0.0378
7	Entropy	2.8119
8	RMS	0.0378
9	Variance	0.0014

Output: Malignant: Mild decline

4. CONCLUSIONS

In this paper, the proposed method detects Alzheimer's disease using SVM classifier. In preprocessing step, noise present in MRI image is removed using median filtering. It is an effective method for reducing the noise. Then in segmentation step, the region growing is applied on denoised image to extract Region of interest. The proposed method extracted nine textual features using GLCM. It is a widely used method for extracting the texture based on features. These textual

features are classified using SVM. SVM classifier is more robust in classifying the features. This method detects the Alzheimer's in two stages i.e. No impairment means there is no symptom of Alzheimer's and malignant i.e. mild decline. This method is an effective method to detect Alzheimer's disease.

REFERENCES

[1] Shuai Mao; Changle Zhang; Na Gao; Yan Wang; Yanwu Yang; XinGuo, Ting Ma. 2017. A Study of Feature Extraction for Alzheimer's Diseaese Based on Resting. State 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) Jeju, Korea 10.1109/EMBC.2017.8036875.

[2] Martin Prince, Renata Bryce, Emiliano Albanese, AndersWimo, WagnerRibeiro, Cleusa P. Ferri. 2013. The global prevalence of dementia: A systematic review and Meta analysis. Alzheimer's & Dementia. 9(1): 63-75, https://doi.org/10.1016/j.jalz.2012.11.007.

[3] Lucia Billeci, Asia Badolato, Lorenzo Bachi and Alessandro Tonacci. 2020. Machine Learning for the Classification of Alzheimer's disease and Its Prodromal Stage Using Brain Diffusion Tensor Imaging Data: A Systematic Review. MDPI Processes 8, 1071doi:10.3390/pr8091071.

[4] A. M. Fjell I. K. Amlien L. T. Westlye K. B. Walhovd. 2009. Mini-Mental State Examination is Sensitive to Brain Atrophy in Alzheimer's disease Dement Geriatr Cogn Disord. 28: 252-258.

[5] Detection Ameer Nisha. S, Shajun Nisha S., Dr. M. Mohamed Sathik. 2017. A Study on Surf & Hog Descriptors for Alzheimer's disease. International Research Journal of Engineering and Technology. 04(02).

[6] Rowayda A, Sadek. 2013. Regional Atrophy Analysis of MRI for Early Detection of Alzheimer's disease. International Journal of Signal Processing, Image Processing and Pattern Recognition. 6(1).

[7] Sambath Kumar Dr. M. Nandhini 2017. A Comprehensive Survey: Early Detection of Alzheimer's Disease Using Different Techniques and Approaches. International Journal of Computer Engineering & Technology, 8(4): 31-44, Article ID: IJCET_08_04_004.

[8] Bharathi. A, A. S. Arunachalam. 2021. Pre-Processing on Alzheimer MRI images. Annals of R.S.C.B., ISSN: 1583-6258, 25(4): 4433-4441.

[9] Nalini Bodaisingi, Balaji Narayanam. 2018. Techniques for De-noising of Bio-Medical Images. International Journal of Biology and Biomedical Engineering. Vol. 12.

[10] Erena Siyoum Biratu, Friedhelm Schwenker, Taye Girma Debelee, Samuel Rahimeto Kebede, Worku Gachena Negera, Hasset Tamirat Molla. 2021. Enhanced Region Growing for Brain Tumor MR Image Segmentation. MDPI, Journal of Imaging, 7(2): 22. https://doi.org/10.3390/jimaging7020022.

[11] Shanshan Sun a, Runtong Zhang. 2017. Region of Interest Extraction of Medical Image based on Improved Region Growing Algorithm. Advances in Engineering Research, vol. 125, Atalantis press.

[12] Wamidh K. Mutlag, Shaker K. Ali, Zahoor M. Aydam and Bahaa H. Taher. 2020. Feature Extraction Methods: A Review. Journal of Physics, doi: 10.1088/1742- 6596/1591/1/012028.

[13] Sumit Salunkhe, Mrinal Bachute, Shilpa Gite, Nishad Vyas, Saanil Khanna, Keta Modi, Chinmay Katpatal and Ketan Kotecha. 2021. Classification of Alzheimer's Disease Patients Using Texture Analysis and Machine Learning. Applied System Innovation MDPI, 4, 49. https://doi.org/10.3390/asi4030049

[14] Ramola A., Shakya A. K., Van Pham D. 2020. Study of statistical methods for texture analysis and their modern evolutions. Engineering Reports. Wiley 2:e12149.https://doi.org/10.1002/eng2.12149.

[15] Pradeep N., Girisha H., Sreepathi B., Karibasappa K. 2012. Feature Extraction of Mammograms. International Journal of Bioinformatics Research. 4(1): 241-244.

[16] Pawan Kumar Mall, Pradeep Kumar Singh, Divakar Yadav. 2019. GLCM Based Feature Extraction and Medical X- RAY Image Classification using Machine Learning Techniques. IEEE Conference on Information and Communication Technology (CICT), DOI: 10.1109/CICT48419.2019.9066263, 2019

[17] Gaurav Kumar, Pradeep Kumar Bhatia. 2014. A Detailed Review of Feature Extraction in Image Processing System. IEEE Fourth International Conference on Advanced Computing &

Communication Technologies, DOI: 10.1109/ACCT.2014.74

[18] Jair Cervantes, Farid Garcia-Lamont, Lisbeth Rodríguez - Mazahua, Asdrubal Lopez. 2020. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, Elsevier. 408: 189-215.

[19] Sameer Mahajan, Gayatri Bangar, Nahush Kulkarni. 2020. Machine Learning Algorithms for Classification of Various Stages of Alzheimer's Disease: A review. International Research Journal of Engineering and Technology. 07(08).

[20] Aditi, Sheetal Chhabra. 2017. Performance Analysis of SVM Classifier for Classification of MRI Image. International Research Journal of Engineering and Technology. 04(06).