JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

An Enhanced Fuzzy Based KNN Classification Method for Alzheimer's Disease Identification from SMRI Images

Mathews Emmanuel
Research Scholar, Department of Computer Science, Sathyabhama Institute of Science and Technology, Chennai.

Dr. J. Jabez
Associate Professor, Department of Information Technology, Sathyabhama Institute of Science and Technology, Chennai.

ABSTRACT

The term "Alzheimer's Disease (AD)" relates to a brain disorder had been affecting a large number of individuals throughout the world annually. It may be dangerous to the individual if left undiagnosed or untreated, also this disorder has no successful therapy. There is a lot of interest in developing novel ways to detect AD more often. In the classification of AD using brain "Magnetic Resonance Imaging (MRI)" images, the "K-Nearest Neighbor (KNN)" classification was among the most successful technique. This algorithm compares the similarity between the training data and new instances before classifying them. However, its accuracy gets lacking when dealing with a complex dataset. To enhance its accuracy level in large datasets in this research we had enhanced the traditional KNN with FuzzyLogic by proposing "Enhanced Fuzzy-KNN (EFKNN)". A fuzzy degree of membership in the problem classes was calculated using the EFKNN. As a consequence, the boundaries between classes are smoother. The conventional KNN technique is unable to deal with large datasets because it lacks a fuzzy variation. The class-membership calculations, however, entail an additional computational burden, making them ineffective for dealing with huge datasets due to large storage requirements and higher running time. The primary goal of this work is to use "structural-MRI (sMRI)" images for obtaining the hippocampus volume area to automatically learn and categorize Alzheimer's disease. Here the proposed model consists of various stages namely "Pre-Processing", "Segmentation", "Feature Extraction", and "Classification". The EFKNN has been shown that it has not only a lower error in the classification of subjects but also more faith in the classification taking advantage of the FuzzyLogic principle. In this research, an EFKNN classifier is implemented for evaluating the subject of sMRI brain images as "Cognitive Normal (CN)", "Mild Cognitive Impairment (MCI)" or "Pure Alzheimer's Disease (AD)" classes during classification. The proposed EFKNN approach proves its efficiency in terms of detection and classification with its accuracy more than the existing KNN approach.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, Magnetic Resonance Imaging, K-Nearest Neighbor, Fuzzy Logic

1. INTRODUCTION

People throughout the globe are living longer due to medical advancements. Meanwhile, the proportion of senior people with AD also is on the rise. Loss of memory, inappropriate brain functions, difficulty in everyday life, and problems in expression and communication are all indications of this particular dementia [1].

As being the most prevalent dementia, AD seems to be a complicated illness that causes memory loss and many other cognitive impairments. AD is perhaps the most frequent dementia. For the time being, no recognized therapy will alleviate this illness, and the typical life expectancy following diagnosis is between 3 and 9 years [2]. About 50 million individuals throughout the globe have Alzheimer's in 2018, as per the "2018 World Alzheimer Report". By the year 2050, its proportion will have tripled to 152 million. A fresh case was reported every 3-seconds throughout the globe [3].

In an attempt to accurately identify AD, a thorough examination of a patient's clinical background, behavior evaluation, cognition assessment, neuroimaging, and also blood collection is required. The patient's clinical background and behavioral evaluation are two of the most important diagnosing findings. Still, several factors need to be taken into

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

consideration, including the patient's clinical background and behavioral testing results for accurate findings. As a result, developing an effective technique for automated diagnosis is a necessity.

Significant healthcare imaging techniques, such as MRI and "Positron Emission Tomography (PET)", have started developing professionally and may give more persuasive evidence for the diagnosis of AD. There are three stages of AD: "Alzheimer's disease (AD)", "Mild Cognitive Impairment (MCI)", and "Cognitive Normal (CN)". AD is the most severe of these. Figure 1 shows some examples of CN, MCI, and AD patients' brain sMRI scans.

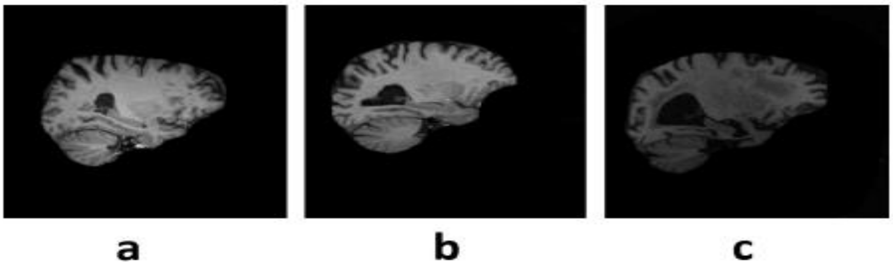

Figure 1. Patient's sMRI brain scans (a) CN (b) MCI (c) AD

Figure 1 shows that the brain's total gray matter size quickly varies from "CN to MCI, and from MCI to AD". For patients with MCI and AD, the hippocampus appears substantially lower than in the CN individual [4].

The MRI is the most advanced technology for evaluating brain tissue by providing detailed data in small, fine-grained sections. The use of MRI in the diagnosis of many disorders has proved effective [5]. Patients with CN, MCI, and AD may have their brain tissues differentiated using advanced image processing techniques. Using brain scans to classify AD takes lesser time and fewer resources. A person's predisposition to AD may be predicted using high-quality brain imaging and other biomarkers that are extracted from it.

Researchers have also employed machine intelligence to create the diagnostic imagery that correlates to the developed approach. Comparatively, an MRI scan may be obtained in a shortened time frame and is least economical for the patient than PET imaging [6]. Machine learning, on the other hand, has advanced considerably in the last several years. Researchers have devised a slew of imaging algorithms for diagnosing Alzheimer's disease using various forms of medical imagery.

As a result, brain MRI scans are being used to classify AD in a better way. Conventional classification methods can't differentiate AD from other diseases due to the obvious complexity of its architecture and the pixel content [7].

Research Problem statement: The lack of high-quality open-source databases in the medical field is a major concern for researchers [8]. For example, the databases are unavailable for use in research since they don't contain the same amount of samples for every type [9]. This creates an issue with over-fitting and under-fitting. Currently, there's also no appropriate treatment for AD. Early detection and prevention have been the only established remedy for this disease. However, current models and algorithms are unable to effectively forecast MCI, resulting in a significant number of older persons worldwide suffering from this dangerous illness. The AD and other forms of dementia may be caught early enough that they can be treated before they progress to a terminal state.

Between 2000 and 2015, 32 million people were diagnosed with dementia, according to estimates. Over 152 million people worldwide are expected to suffer from dementia around the year 2050. In particular, the associated costs are too high for both individuals and healthcare providers. An estimated $7.9 trillion may be saved worldwide if this condition can be detected and treated early. As a result, MCI must be detected and treated as soon as possible.

Main Contribution: Since of its increased imaging adaptability, superior tissue contrasting and absence of radiation, we feel that the MRI modality is an appropriate choice for this research because it provides important knowledge on human brain architecture. The development of a better "Computer Aided Diagnostic (CAD)" model for interpreting sMRI scans and determining whether patients are normal or advancing to AD is regarded as fundamental. In an attempt to correctly diagnose the various phases of AD, the model was formulated from scratch.

Using sMRI brain scans, this proposed EFKNN algorithm is used to determine whether the person has CN, MCI, or Pure AD during the classification process. To test the models, it was trained using the ADNI MRI dataset [10]. The substantial number of subjects respective to MCI, AD, and CN are all included in the dataset. According to the findings, the proposed approach with a decreased number of parameters surpasses the current KNN approach.

The advantages of this proposed work are as follows:

- The EFKNN has been shown that it has not only a lower error in the classification of subjects but also more faith in the classification taking advantage of the FuzzyLogic principle.
- The EFKNN gives a more practical vector for the object's membership and thus provides for the object's class membership.

Paper Organization: Section 2 covers the recent publications in AD classification, Section 3 briefs about the methodologies involved in proposed and existing models module by module, Section 4 discusses the result obtained with a comparison of both existing and proposed models, and Section 5 concludes this research article with future scope.

2. RELATED WORKS

An innovative "Multimodal Deep Neural Network" using a multiphase approach was emphasized by the researchers in [11]. Using this strategy, 82.4 percent of individuals with MCI and those who are subsequently diagnosed with AD are being accurately predicted in 3 years. Accuracy is 86.3 percent for the AD category, and 94.23 percent for the CN category, according to this model.

The ADNI and "National Research Center for Dementia (NRCD)" datasets have been used by the researchers in [12] to suggest a diagnostic technique for the categorization of AD. A combination of characteristics from the subcortical, cortical, and hippocampus regions using MRI scans provides an accuracy of 96.42 percent for the categorization of AD from CN.

An aggregate CNN approach for extracting features and "SoftMax" classification was developed by the researchers in [13]. To avoid over-fitting, this method makes use of the right and left hippocampal regions in MRI scans.

For the class imbalance framework, the researchers in [14] used a "Pre-trained Alexnet" framework to identify the phases of AD. Features are extracted using "Support Vector Machine (SVM)", KNN, and "Random Forest (RF)" with an overall precision of 99.21 percent, by using a pre-trained network.

In [15], the researchers design a technique for automatically locating the desired locations inside huge MRI volumes. Accuracy rates of 94.82 percent and 94.02 percent are achieved by using information from the right and left hippocampus.

3. METHODOLOGIES

The CAD provides a broad range of capabilities to several disciplines of neuroimaging. A multitude of neuroimaging techniques was used to create diagnostic tools for a wide range of nonbrain and brain illnesses.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

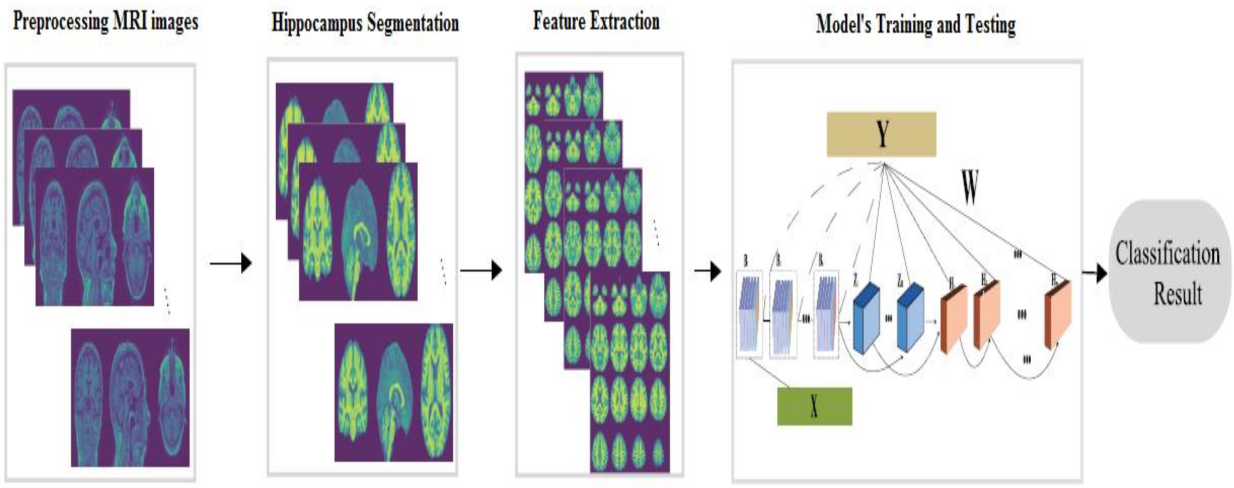

Figure 2. Proposed Methodology for AD Classification

In an attempt to identify different brain illnesses, such as Alzheimer's, seizures, etc, the sMRI has been one of the imaging modalities that gives higher resolution anatomical knowledge about the human brain. As a key indicator for AD diagnosis, sMRI may reveal shrinkage in the hippocampus structure. Also, Hippocampal volume measurement is a challenging task. In this research, we had undergone some preliminary stages before classifying the AD [16]. Figure 2 presents the overall flow of the proposed CAD model for AD diagnosis based on brain sMRI images.

3.1 MRI PREPROCESSING

The first stage in our suggested CAD model is data preprocessing. These steps are designed to decrease the amount of data needed to develop an accurate prediction model, and also to make it easier for the classification method to work with. Therefore, the 3-dimensional sMRI scans need a large layered structure, which raises the processing burdens. In addition, certain post-processing processes may necessitate increased computational time. Using the 3-dimensional scans, we've created a set of 2-dimensional slices. One of the most impacted parts of the brain in Alzheimer's disease is the hippocampus. Patients with MCI/AD show a steady decline in their hippocampus. As a result, we opted to utilize sagittal view imaging of the hippocampus for the classification model instead of coronal ones. We've looked at the 2-dimensional images and found the best slices for obtaining hippocampus tissue. A 256*256*1 sized image is generated as preprocessed output from all of the original sMRI sources. Figure 3 shows the preprocessed 2-dimensional sMRI brain image with "Coronal-View", "Sagital-View" and "Axial-View".

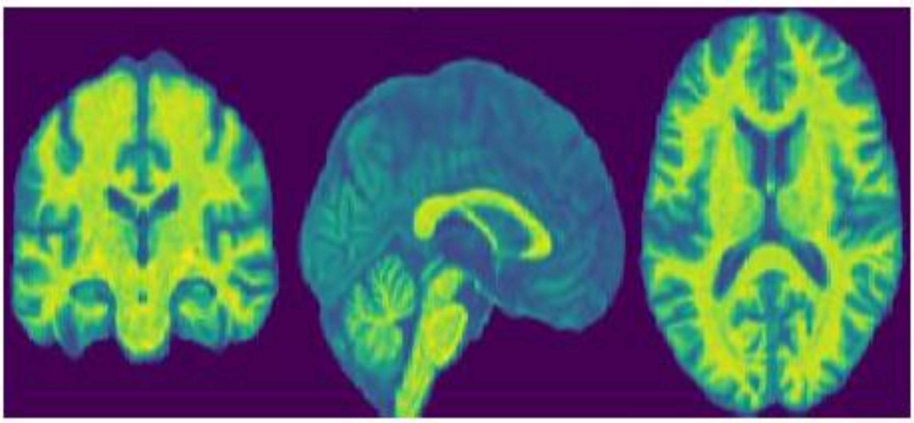

Figure 3. Preprocessed 2-dimensional Image

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

3.2 HIPPPOCAMPUS SEGMENTATION

Non-brain portions known as the skull may reside in the preprocessed brain sMRI images. The existence of a skull within the input data increases the dimensionality of the feature patterns, but it is not important for AD categorization in this research. As a result, we separated the brain from the skull portions. We employed the "Histogram Based Thresholding" strategy for segmenting the hippocampus area, which is one of the most widely employed image segmentation approaches. Figure 4 depicts one of the graphical consequences of skull removal and segmentation of the hippocampus area. Figure 4 (a) shows a preprocessed brain sMRI image, whereas Figure 4 (b) shows the visual result upon applying the segmentation.

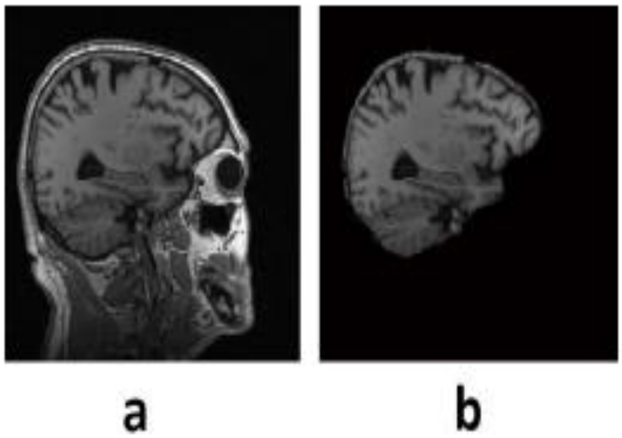

Figure 4. a) Preprocessed sMRI brain image, and b) Segmented image

3.3 FEATURE EXTRACTION

To obtain the category features, a "One-Hot Encoded" approach was employed to turn the data into a binary sequence. The "Z-score Normalization" was utilized to produce the typical scaling of quantitative features throughout the segmented hippocampus area. The "Pearson Correlation Coefficient (PCC)" has been used to filter out the correlated features from the segmented image by comparing the correlation between the "numerical-numerical" and "numerical-categorical" features. Pearson-Coefficient "correlation >0.99" features were tested through PCC. Using the PCC, we could see how the 2 predictors 'X' and 'Y' correlate linearly. A correlation coefficient of '0' implies no linear association among the 2 predictors, whereas a number close to '1' suggests the maximum association among the 2 predictors. Neither of the predictors has been eliminated since none were significantly "correlated (>0.99)".

3.4 CLASSIFICATION

3.4.1 KNN

The mechanism of this classification works in the order in which a single sample is allocated to a class based on information learned through training by the classifier. It is its job to give one of the several prespecified groups an input pattern represented by a variable. The KNN classifier is the existing classifier in this research to detect the subject as CN, MCI, or AD for the classification of the sMRI brain image. This classifier does a classification based on a non-parametric process. If a new training feature subset is applied to an established training collection, there is no need for previous information regarding the layout of the feature subsets gained in training sets (No retraining is required). The KNN technique performance can be used as an eternity of the input pattern of a given class. While 'k' rises, predicting confidence will increase.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

Work Flow of KNN for Classifying AD

- Creating a list of gathered subsets with optimal features.
- Using Euclidean for computing the distances among the saved subset of features (trained-set) with the subset of features that are unknown (tested-set) for classification.
- Identifying the K nearest by using the closest adjacent class-labels (CN, MCI, and AD) to identify the class-labels of the undefined record by voting has a higher majority.

3.4.2 EFKNN

The methods for this classification are primarily employed towards some correlation calculation of a group of items dependent on such distance measurements. One of the existing template classification techniques without dimensionality reduction includes the KNN method. With Object-Membership functions, the decision rule for existing KNN assumes equivalent weight, ignoring numerous similarity patterns. The proposed EFKNN in this research has shown that it has not only a lower error in the classification of subjects but also more faith in the classification taking advantage of the Fuzzy-Logic principle. The EFKNN gives a more practical vector for the object's membership and thus provides for the object's class membership. A class with its nearest K-neighbors is allocated in this algorithm to the most common form. EFKNN assigns the sample's fuzzy membership and allows policymakers to make fuzzy choices. In this research for evaluating the subject of brain sMRI images as CN, MCI, or AD the EFKNN classification is proposed.

3.4.2.1 FUZZY LOGIC (FL)

The FL identified the role that ties the reality of the proposal to other proposals. A collection of an infinite series, an FL determines the actual [0, 1] number with the number between true and false. Another real meaning is contained in the FL describing the truth tables of varying adjective degrees. The principle presents as knowledge linguistically that offers a systemic calculus and linguistic labels indicate numerical estimation utilizing the membership function. A systemic calculus is given. For instance, take into account a classical "set A" with a narrow limit and contains a real number higher than 'K' which was described as per the following Equation (1):

$$A = \{x > K\}$$

Eq→1

It is observed that 'K' is simple and unequivocal. If 'x' is larger than 'K', 'x' will be "set A", otherwise, it is not "set A" as per Equation (1). In numerous configurations including certain science and engineering, the classical set is used and does not represent human existence, however, it appears to be complex and imprecise. The fuzzy-set will simplify it.

3.4.2.2 CLASSIFICATION OF ALZHEIMER'S DISEASE USING EFKNN

The proposed EFKNN is intended to divide the sub-set "X = {x1, x2, . . ., xn}< Rn" of vector samples into the cluster of sub-sets with fuzzy of "c (1<c<n)". In that case "(i = 1, 2,. ., c) & (j = 1, 2,. ., n)" the fuzzy matrix of membership is 'U', in which "U_{ij}" is the "x_j" fuzzy in class 'i'. The object "j^{th}" is allocated to the class "i^{th}", which has the highest "U_{ij}", relative to the membership of the fuzzy with other groups in a non-fuzzy variant of the method. Two restrictions are present in matrix "U" as given below:

$$\sum_{i=1}^{c} u_{ij} = 1,\ j$$

Eq→2

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

$$u_{ij}\hat{I}\ [0,1], \quad 0 < \sum_{j=1}^{n} u_{ij} < n.$$

Eq→3

The first-ever limitation as in Equation (2) guarantees that all membership object's grades are earned in all groups "i=1, 2, .,c", and all the membership grades are summed as one. Equation (3) notes that for all objects, the membership of fuzzy classes lies at or above zero and is equivalent to or below one. These 2 limitations illustrate that if an object is in the "U = 1" class, it certainly has no participation in the other classes. Furthermore, for all objects in a class the total of all the membership of fuzzy grades surpasses zero, or else the class does not exist and is consequently fewer than the total of objects 'n'. In the algorithm of FKNN each vector is given the degree of fuzzy membership, regarding the distances between vectors and their membership in KNN which was given below:

$$u_i(x) = \frac{\sum_{j=1}^{K}\left(\frac{u_{ij}}{||x-x_j||^{\frac{2}{(m-1)}}}\right)}{\sum_{j=1}^{K}\left(\frac{1}{||x-x_j||^{\frac{2}{(m-1)}}}\right)}$$

Eq→4

Here 'K' is the neighbors that are nearest with a predefined number, and 'm' is the parameter with a constant. In the measurement of membership of fuzzy value, the 'm' parameter defines the weight of each closest neighbor. In Equation (4) it's a central part in the calculation of the degree of membership of an "i^{th}" object class that was regulated by inverted distances between an 'x' object and its closest neighbors and memberships of a 'K' class.

The reverse association between membership level and distances, unlike most of the non-fuzzy version classification techniques, acts as the part that a weighting feature performs in rewarding/penalizing those with farther or closer distance from other class objects. Clearer is that since an object belongs to the 'A' class with a degree of "0.95" when it belongs only to the 'B' class with a degree of "0.05" then it will be rational that an object should belong to 'A'.

If therefore the object membership ratings were "0.55" and "0.45" respectively for classes 'A' and 'B', there could be some hesitations in either of the classes 'A' or 'B' until the object was allocated. Ultimately, the mission, which creates a greater degree of resemblance, specifies the components of either 'A' or 'B' of the object. The weighting function in Equation (4) illustrates all such circumstances dependent on the inverse distances of the objection set in the class.

The 'm' parameter shows the magnitude when the items are rewarded/ penalized with distances from some of the other objects. The near neighbors play a far more significant function in determining the membership standard of the subject to be listed, 'm' is often larger than one and the closer its worth. Through raising this metric, from the other side, neighbors are weighed more equally and are less likely to have relative distances from the categorized object. The distance of Euclidean among its "j^{th}" and 'x' is "$||x\text{-}x_j||$"

Whenever the number of input parameters will be less than the scale of the training set, in machine learning models, there is rather a slight chance of "Over-Fitting (OF)". "Cross-Validation (CV)" is also implemented for model precision without depending upon the data used during the training set in the calculation of the performance variables.

The CV will focus on saving predictive models from problems like OF and test the independence of the model in its data collection. The OF is rendered by storing the mapping feature from input to output variable when the forecast loses its true sense throughout the process of the testing. This is normal when the data number to be separated between the training and test dataset is not adequate without applying data damage.

Even though the amount of data available is substantially larger than the proportion of input variables, there is a slight possibility that the OF is required to obtain an understanding of how quantitative models can be confident, in reality, the "Nash-Sutcliffe (NS)" process efficiency coefficient can be used. The coefficient of NS is as follows:

$$E_{NS} = 1 - \frac{\sum_{i=1}^{n} (h_i^* - h_i)^2}{\sum_{i=1}^{n} (h_i^* - H)^2}$$

Eq→5

The coefficient for NS ranges from "-∞ to 1", in this near to the value of one implies greater preciseness and an appropriately zero value indicates that the predictive model is a reasonable mean, and some negative values reflect that even the mean of the observed results is less reliable than the predictive model result.

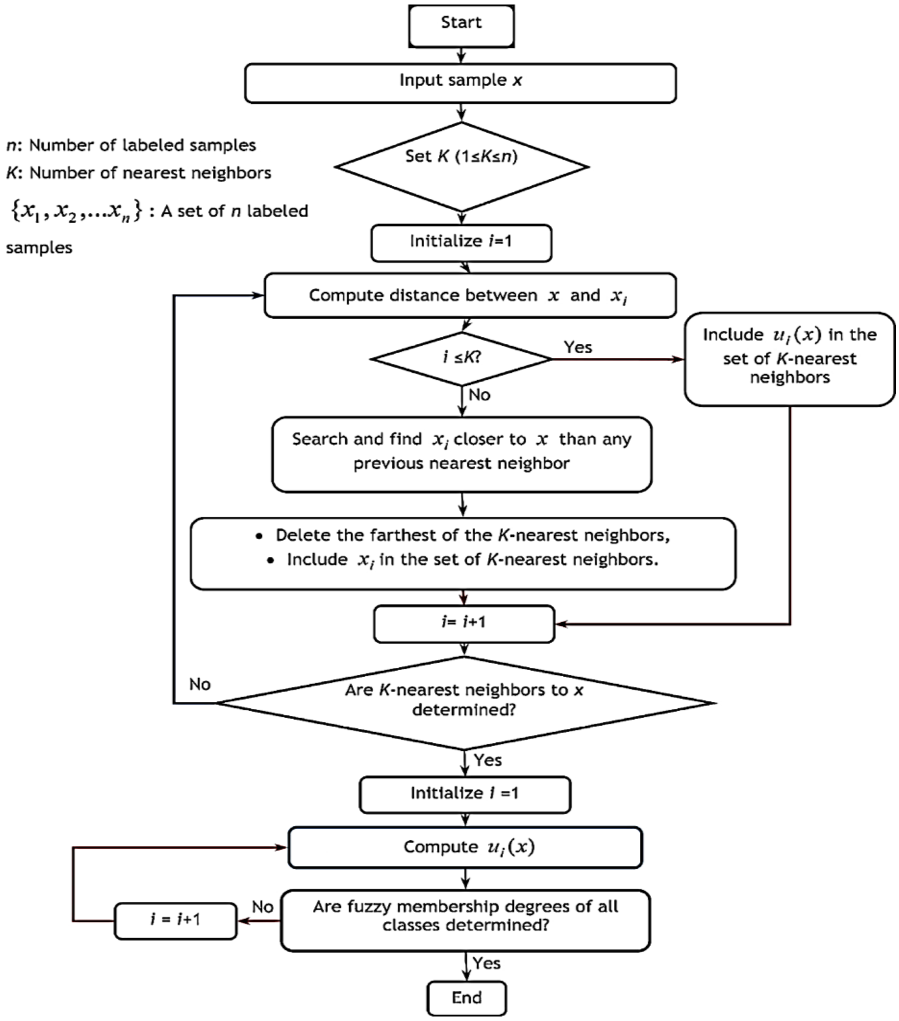

Figure 5. A flowchart of the EFKNN algorithm

Figure 5 shows the flowchart of the proposed FKNN classifier for classifying AD. The FKNN method operates with the number of the closest neighbors by giving an 'x' sample of input and neighbors that are nearest 'K'. With it the first class (i=1) the method is initiated then the distances between 'x' and 'x_i' are determined.

Moreover, unless the quantity of classes, 'i' is equivalent to or lower than 'K', the algorithm scans the "x_i" closest to 'X' than the nearest neighbor. In addition to this, "$U_i(x)$" is included in the 'K' neighboring set. The closest neighbors from 'K' are excluded from the collection if the second condition is valid and substituted by "x_i".

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

The 'i' is then incremented by one and the process is performed until all K-nearest neighbors to the 'x' are found, if the closest K neighbors to the 'x' are not determined. The membership of fuzzy values for both groups is then determined as in Equation (4).

(i) TRAINING PHASE

Here it has been considered the probability of distribution as "$Pd = (pd1, pd2, p3, \ldots, pdn) \in \Delta n$" where "$\Delta n = \{(pd1, pd2, p3, \ldots, pdn) | pdi \geq 0, i = 1,2, \ldots, n, n \geq 2, \Sigma\ pdi = 1\ ni = 1\ \}$" for 'n' with finite discrete sets. Subsets with the feature from the sMRI the entropy were calculated by following Equation (6):

$$TE(P) = -\sum_{i=1}^{n} pd_i log_2 pd_i$$

Eq→ 6

Here the 'i' indicates the essential stage in the sMRI dimensional sub-set "M*N". The normalized histogram is specified as 'P' for a sub-set with features. Then the normalized histogram into 'n' groups is partitioned by "N-1" thresholds (t). It was classified as CN, MCI, and AD into three classes. The entropy for each class is defined as follows:

$$TE_1(t) = -\sum_{i=0}^{t_1} \frac{pd_i}{Pd_1} In \frac{pd_i}{Pd_1}$$

Eq→ 7

$$TE_2(t) = -\sum_{i=t_1+1}^{t_2} \frac{pd_i}{Pd_1} In \frac{pd_i}{Pd_1}$$

Eq→ 8

$$TE_n(t) = -\sum_{i=t_{n-1}+1}^{L-1} \frac{pd_i}{Pd_1} In \frac{pd_i}{Pd_1}$$

Eq→9

Where

$$Pd_1(t) = \sum_{i=0}^{t_1} pd_i, Pd_2(t) = \sum_{i=t_1+1}^{t_2} pd_i, \ldots, Pd_n(t) = \sum_{i=t_{n-1}+1}^{L-1} pd_i$$

Eq→10

Here the thresholds with a dummy of two for classifying the MRI as "$tn = L - 1$" and "$t0 = 0$" were initiated by "$t0 < t1 < \cdots < tn-1 < tn$". To find out the optimal value for the threshold Equation (11) is used.

$$\varphi(t_1, t_2, \ldots, t_n) = Arg\ max([TE_1(t) + TE_2(t) + \cdots + TE_n(t)])$$

Eq→11

The EFKNN classifier's training process is quick and the EFKNN classifier's training samples are stored in a local sub-region that is required mostly during the EFKNN classifier test stage. EFKNN preserves all trained samples that contribute to a lack of generalization.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

(ii) TESTING PHASE

The 'A' set is a classic unit, which is described by an element array that may or may not form part of the 'A' set. Centered on the fuzzy-set it is a classic-set generalization, an aspect is partly origin from the 'A' set. The 'A' set is determined by the given Equation (12).

$$A = \{(y, \mu_A(y)) | y \epsilon Y\}$$

Eq→12

Here the "μA" is the function with membership, it calculates 'y' that is nearer to 'A'. The Trapezoidal Membership Features operate for the sMRI Feature Subset, "$\mu_1, \mu_2, ..., \mu_n$" by 2*(n − 1) as a non-recognized area with fuzzy parameters $r_1, s_1, ..., r_{n-1}, s_{n-1}$ in which "$0 \leq r1 \leq s1 \leq \cdots \leq rn{-}1 \leq sn{-}1 \leq L - 1$". In this proposed system the trapezoidal membership is used for this computation. The threshold with level 'n' is extracted by the next membership function as per the following Equations (13, 14, 15).

$$\mu_1(k) = \begin{cases} 1 & k \leq r_1 \\ \frac{k-s_1}{r_1-s_1} & r_1 \leq k \leq s_1 \\ 0 & k > s_1 \end{cases}$$

Eq→13

$$\mu_{n-1}(k) = \begin{cases} 0\ k \leq r_n - 2 \\ \frac{k-r_n-2}{s_n-2-r_n-2}\ r_{n-2} < k < s_{n-2} \\ 1\ s_{n-2} < k < r_{n-2} \\ \frac{k-s_n-1}{r_{n-1}-s_{n-1}}\ r_{n-1} < k < s_{n-1} \\ 0\ k > c_{n-1} \end{cases}$$

Eq→14

$$\mu_n(k) = \begin{cases} 1 & k \\ \frac{k-r_n}{s_n-r_n} & r_{n-1} < k < s_{n-1} \\ 1 & k > s_{n-1} \end{cases}$$

Eq→15

For the feature subset of sMRI, the MaximumFuzzyEntropy (MFE) for classification at the n level was performed by following Equations (16,17, 18).

$$MFE_1 = -\sum_{i=0}^{L-1} \frac{pd_i * \mu_1(i)}{Pd_1} * In\left(\frac{pd_i * \mu_1(i)}{Pd_1}\right)$$

Eq→16

$$MFE_1 = -\sum_{i=0}^{L-1} \frac{pd_i * \mu_2(i)}{Pd_2} * In\left(\frac{pd_i * \mu_2(i)}{Pd_2}\right)$$

Eq→17

$$MFE_n = -\sum_{i=0}^{L-1} \frac{pd_i * \mu_n(i)}{Pd_n} * In\left(\frac{pd_i * \mu_n(i)}{Pd_n}\right)$$

Eq→18

For obtaining the parameter's optimal value it needs to maximize the total entropy as given in Equation (19):

$$\varphi(r_1, s_1, \ldots, r_{n-1}, s_{n-1}) = Argmax([MFE_1(t) + MFE_2(t) + .. + MFE_n(t)])$$

Eq→19

As per Equation (19) is essential to maximize the measurement time of the proposed MRI AD classification by employing a global optimization technique. Using the fuzzy-parameter (n-1) the following is given for several threshold values:

$$t_1 = \frac{(r_1 + s_1)}{2}$$

Eq→20

$$t_2 = \frac{(r_2 + s_2)}{2}$$

Eq--→21

$$t_{n-1} = \frac{(r_{n-1} + c_{n-1})}{2}$$

Eq--→22

By utilizing these decision rules this EFKNN classifier was able to classify the MRI images as ‘CN’, 'MCI' or ‘AD' with less computational time and improved accuracy.

4. RESULTS AND DISCUSSIONS:

DATASET

An "Original volumetric T1-weighted, Magnetization Prepared Rapid Gradient Echo (MPRAGE) sMRI" images were collected from the public dataset "Alzheimer's Disease Neuroimaging Initiative (ADNI)" for this research. Images of "210 (Male:105, Female:105) distinct subjects (CN: 70, MCI 70, AD: 70)" are taken from more than 2000 data.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

TOOLS

The mathematical programming language Matlab is widely utilized in a multitude of medical imaging applications. Matlab's user-friendly interactions make it quicker than many of the other toolboxes in implementing new ideas. To evaluate all of the model designs, we have utilized the Matlab software toolbox. We've employed data generator methods including rotational, contrast adjustment, inverting, and more to boost generalization ability. The predicted class is shown with corresponding labeled classes from the four separate categories "True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN)" in the "Confusion-Matrix (CM)". Over the training dataset, the CM delivers the model's performance.

ACCURACY

The accuracy of classification is one of the most general methods for estimating system performance. It's being used to evaluate the ability of the classification system as the important standard parameter. The higher the accuracy of the categorization, the better the efficiency is. The benefits of this measure are simple and are calculated using the CM as given in Equation (23).

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} * 100$$

Eq→23

In this research, the Accuracy of KNN, and EFKNN for classifying the CN, MCI, and AD from the sMRI datasets are evaluated and compared in this section. Table 1 and Figure 6 show the accuracy performance for different sMRI images. Here for comparison, it had been taken 6 images. The accuracy level of the EFKNN is higher for all images when comparing it with KNN. Thus for AD subjects, it had proved that the EFKNN algorithm provides a better accuracy rate when comparing it with KNN classifier algorithms.

Table 1. Accuracy Comparison

SUBJECTS	KNN	EFKNN
AD-sMRI1	81	92
AD-sMRI2	82	93
MCI-sMRI1	79	90
MCI-sMRI2	80	91
CN-sMRI1	85	96
CN-sMRI2	84	95

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 89 - 103
https://publishoa.com
ISSN: 1309-3452

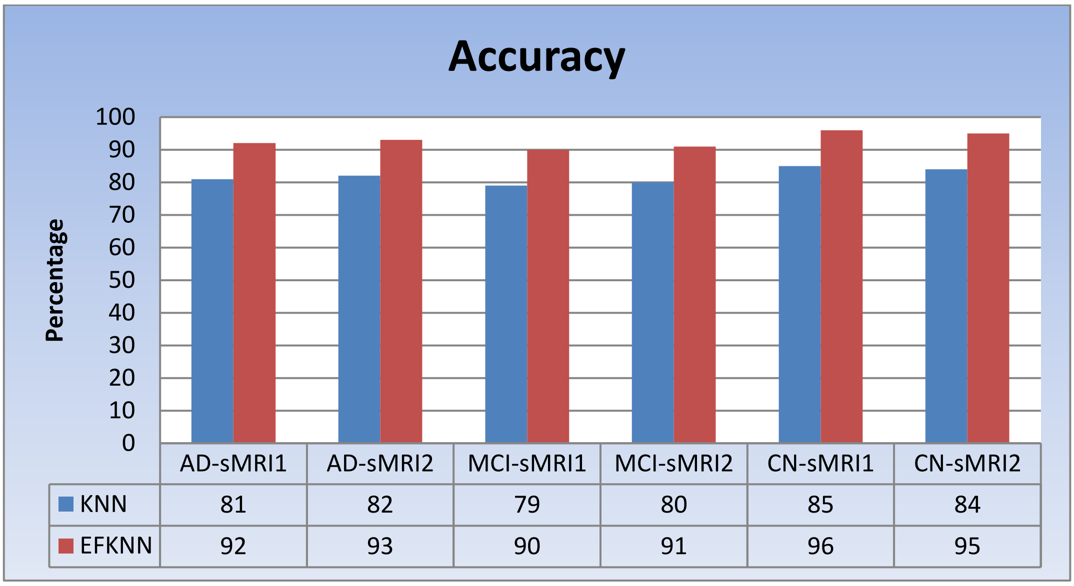

Figure 6. Accuracy Comparison Graph

SENSITIVITY / RECALL

This section illustrates the exact prediction of dementia and it was labeled as TP value. Similarly, the non-dementia was labeled as TN. The TP and TN indicate the accurate outcomes of medical tests and have been proven to be valid. Both in FN and FP, the actual circumstances are inverted. The examination specifies the precise effects of the TP which is commensurate with the exposure. This sensitivity analysis proves the recognition of the diseases and the percentage of FN and TP identification is effectively calculated as specified in Equation (24).

$$\text{Sensitivity}=\frac{TP}{TP+FN}$$

Eq→24

Table 2. Sensitivity Comparison

SUBJECTS	KNN	EFKNN
AD-sMRI1	77	88
AD-sMRI2	78	89
MCI-sMRI1	75	86
MCI-sMRI2	76	87
CN-sMRI1	81	92
CN-sMRI2	80	91

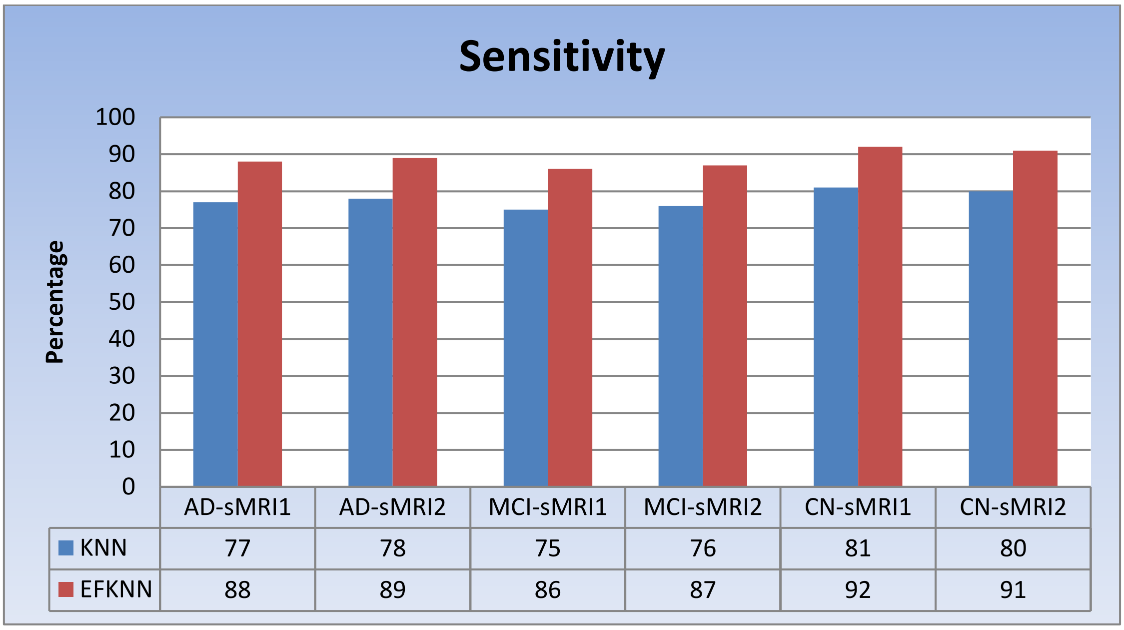

Figure 7. Sensitivity Comparison Graph

In this research, the Sensitivity of KNN, and EFKNN for classifying the CN, MCI, and AD from the sMRI datasets are evaluated and compared in this section. Table 2 and Figure 7 show the sensitivity performance for different sMRI images. Here for comparison, it had been taken 6 images. The sensitivity level of the EFKNN is higher for all images when comparing it with KNN. Thus for AD subjects, it had proved that the EFKNN algorithm provides a better sensitivity rate when comparing it with KNN classifier algorithms.

5. CONCLUSION

AD is very difficult to treat, also the individuals are afflicted for the remaining part of their lives after being diagnosed. Early detection of AD is significant for delaying the illness's progression. There is a requirement for a large number of MRI scans in particular to investigate multi-class categorization in dementia in depth. In this research work, the fuzzy model combined with the enhanced KNN method for the detection of "AD, CN, and MCI" based on the hippocampus region from the sMRI images. The proposed EFKNN works as the principal concept to delegate the representatives for the potential groups as a mechanism that distinguishes the objects from their KNNs. It is a true member of one class and not a part of another. The membership is allocated with a medium distance between the class and its components. Here the enhanced KNN measures the first centroid then the fuzzy component is determined. The models are evaluated in the ADNI database and compared with the conventional KNN classifier approach and findings are shown. Experiments have shown that our proposed approach, called the EFKNN method, is superior to the KNN method. The proposed model's accuracy will be improved in the future using further augmentation approaches. Speeding up computations by reducing the number of features. Optimizing hyperparameters to identify better optimizers and features to enhance the premature dementia diagnosis is a continual effort.

REFERENCES

[1] J. Giorgio, S. M. Landau, W. J. Jagust, P. Tino, and Z. Kourtzi, "Modelling prognostic trajectories of cognitive decline due to Alzheimer's disease", NeuroImage, Clin., vol. 26, Jan. 2020, Art. no. 102199, DOI: 10.1016/j.nicl.2020.102199.

[2] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, "A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia", Neural Netw., vol. 123, pp. 176-190, Mar. 2020, DOI: 10.1016/j.neunet.2019.12.006.

[3] C. Patterson, "World Alzheimer report 2018 the state of the art of dementia research: New frontiers", Tech. Rep., Jun. 2018.

[4] H. Wang, Y. Shen, S. Wang, T. Xiao, L. Deng, X. Wang, and X. Zhao, "Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease", Neurocomputing, vol. 333, pp. 145-156, Mar. 2019, DOI: 10.1016/j.neucom.2018.12.018.

[5] C. Ge, Q. Qu, I. Y.-H. Gu, and A. Store Jakola, "Multiscale deep convolutional networks for characterization and detection of Alzheimer's disease using MR images", in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2019, pp. 789-793, DOI: 10.1109/ICIP.2019.8803731.

[6] B. Deepa, M. G. Sumithra, V. Chandran, and V. Gnanaprakash, "Fusion-based segmentation technique for improving the diagnosis of MRI brain tumor in CAD applications", in Proc. Int. Conf. ISMAC Comput. Vis. Bio-Eng., vol. 30. Cham, Switzerland: Springer, 2019, pp. 299-307.

[7] R. Guan, X. Wen, Y. Liang, D. Xu, B. He, and X. Feng, "Trends in Alzheimer's disease research based upon machine learning analysis of PubMed abstracts", Int. J. Biol. Sci., vol. 15, no. 10, p. 2065, 2019.

[8] T. Pan, J. Zhao, W.Wu, and J. Yang, "Learning imbalanced datasets based on SMOTE and Gaussian distribution", Inf. Sci., vol. 512, pp. 1214-1233, Feb. 2020, DOI: 10.1016/j.ins.2019.10.048.

[9] S. Afzal, M. Maqsood, F. Nazir, U. Khan, F. Aadil, K. M. Awan, I. Mehmood, and O.-Y. Song, "A data augmentation-based framework to handle class imbalance problem for Alzheimer's stage detection", IEEE Access, vol. 7, pp. 115528-115539, 2019, DOI: 10.1109/ACCESS.2019.2932786.

[10] ADNI. Alzheimer's Disease Neuroimaging Initiative: ADNI. Accessed: Jun. 21, 2021. [Online]. Available: http://adni.loni.usc.edu/data-samples/access-data

[11] D. Lu, A. Disease Neuroimaging Initiative, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images", Sci. Rep., vol. 8, no. 1, pp. 1-13, Dec. 2018, DOI: 10.1038/s41598-018-22871-z.

[12] Y. Gupta, K. H. Lee, K. Y. Choi, J. J. Lee, B. C. Kim, and G. R. Kwon, "Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images", PLoS ONE, vol. 14, no.10, 2019, Art. no. e0222446, DOI: 10.1371/journal.pone.0222446.

[13] S. Ahmed, K. Y. Choi, J. J. Lee, B. C. Kim, G.-R. Kwon, K. H. Lee, and H. Y. Jung, "Ensembles of patch-based classifiers for diagnosis of Alzheimer diseases",' IEEE Access, vol. 7, pp. 73373-73383, 2019, DOI: 10.1109/ACCESS.2019.2920011.

[14] H. Nawaz, M. Maqsood, S. Afzal, F. Aadil, I. Mehmood, and S. Rho, "A deep feature-based real-time system for alzheimer disease stage detection", Multimedia Tools Appl., pp. 1-19, Jun. 2020, DOI: 10.1007/s11042-020-09087-y.

[15] A. Basher, B. C. Kim, K. H. Lee, and H. Y. Jung, "Volumetric feature-based Alzheimer's disease diagnosis from sMRI data using a convolutional neural network and a deep neural network", IEEE Access, vol. 9, pp. 29870-29882, 2021, DOI: 10.1109/ACCESS.2021.3059658.

[16] Mathews Emmanuel, & Jabez, J. (2022). A brief survey on different methods employed in each module for diagnosing alzheimer disease in the early stage. International Journal of Health Sciences, 6(S1). https://doi.org/10.53730/ijhs.v6nS1.7631