How to Cite:
Emmanuel, M., & 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), 10953–10970. https://doi.org/10.53730/ijhs.v6nS1.7631

A brief survey on different methods employed in each module for diagnosing alzheimer disease in the early stage

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*---**Numerous individuals are afflicted with disorders of the brain all around the globe. There is a need to research and discover an approach to these disorders is crucial currently. Dementia is perhaps an occurrence of brain disease. Injuries or typical elderly might lead to a decline in cognitive processes including thinking, remembering, and some other psychological attributes. Alzheimer's disease (AD) accounted for 60 to 80 percent of all cases of mental ailment and is among the most common forms of dementia. "Magnetic Resonance Imaging (MRI)" could be used to diagnose this condition at a preliminary phase and allow individuals to have a long and healthy life. An intellectual system widely recognized as "Machine Learning (ML)" utilizes many probabilities and optimizing approaches that provide PCs the ability to benefit from massive and complicated information. As a consequence, experimenters are increasingly relying on ML for the early detection of AD. The purpose of this research is to conduct a review of the "Automated Alzheimer's Disease Diagnosis (AADD)" framework. Commonly the AADD framework is composed of the following modules such as MRI-Preprocessing, Hippocampus-Segmentation, Feature-Extraction to extract brain structural features, Feature-Selection to select optimal features from the feature vector, and finally Classification to classify the individual for Cognitive-Normal (CN) or AD from MRI image datasets. Early identification of AD through ML approaches is reviewed, analyzed, and critically evaluated in this research. There were some intriguing approaches, but they were tested on a diversity of unverified large datasets from various neuroimaging modalities, making comparisons challenging. This survey article will be very useful for Alzheimer's researchers and also for ML researchers.

International Journal of Health Sciences ISSN 2550-6978 E-ISSN 2550-696X © 2022.
Manuscript submitted: 27 March 2022, Manuscript revised: 9 April 2022, Accepted for publication: 18 May 2022

***Keywords*---**Alzheimer’s disease, Magnetic Resonance Imaging, Machine learning, Classification.

Introduction

Dementia is a concept that encompasses several symptoms, including the loss of recollection, thought, vocabulary, or vision. Such signs are common variations used to classify multiple forms of the disease, such as Alzheimer's or Parkinson's. AD is typically an aging condition, although, in the past decades, the number of cases grew. In several nations, too, there has been a rise in the number of individuals suffering from early-stage dementia (persons under 65 years old). Alzheimer's is the most prevalent form of diseased patients and comprises 60 to 80 percent of cases [1]. However, as the condition develops, there will be other signs, including changes in disorientation, disposition, and behavior. There may be uncertainty about the activities, time, and location of residing. 1.37% of the world's estimated deaths in 2030 are estimated to be caused by AD and other dementias, according to the World Health Organization (WHO) as shown in Figure 1.

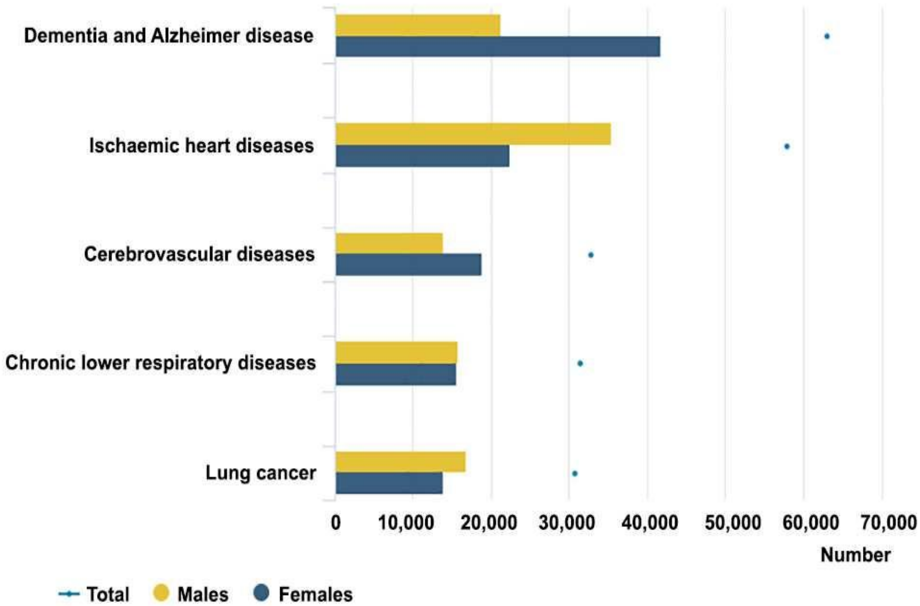

Figure 1. World Health Organization Statistics

Small dementia people have an estimated eight-year life expectancy. The synaptic loss is widely distributed to render the condition permanent as AD is diagnosed. The remaining neurons will not separate and replace the other neurons as neurons die, as do other cells, and therefore the damage cannot be reversed.
To order to reduce the decline speed it is also critical that dementia is identified early in the process. Another question to take into consideration is the expense of dementia. Global dementia costs "819 billion US dollars" and by 2023 will reach a trillion [2]. Most nations reported the expense of dementia to be 26.3 billion

pounds in 2018. The incidence of dementia was 4.3 billion in hospitals and about 85 million in diagnoses. Therefore, affordable diagnostic and support instruments must be developed to help reduce the rising costs of dementia. Several of the proposals suggested is centered on healthcare approaches to minimize expenses and render health services and technologies easily available [3].

(i) Symptoms of AD:

Alzheimer's symptoms typically begin with the difficulty of gaining new knowledge, leading to increasingly severe symptoms such as "disorientation, mood and behavior, deepening confusion over events, times and places; unfounded suspicion about family, friends, and caregivers; more severe memory loss and changes in behavior; and difficulty in the talk, swallowing symptoms may be more noticeable to relatives or associates in certain situations than to the person involved". In the hippocampus (in control of memory formation), a disruptive combination of plaques and tangles starts [4]. The first symptom of AD is typically a short-term memory failure. Proteins slowly infiltrate certain brain regions, triggering subtle modifications and signaling several disease levels. It destroys the ability to process logical thinking in front of the brain. The mental balance then diminished and the attitude shifted erratically. So, anxiety and visions are triggered (top of the brain). All the antibodies function together at the rear of the brain to remove the darkest thoughts of the subconscious. Finally, control centers (heart and breathing governance) are overpowered, which leads to death. Not clear though. The illness is incredibly damaging [5].

(ii) Faction of Risk:

While the specific cause of the disorder is still unknown, risk factors specifically contribute to Alzheimer's growth.

(i) Age: Alzheimer's age is quite specifically related. At 65 years of age, the risk of Alzheimer's disease rises every 5 years to almost 50% at 85 years.

(ii) Family History: It is more likely to be produced by a similar family member with Alzheimer's. The incidence rises if more than one family member becomes sick. Where diseases tend to take place in families, both hereditary and environmental factors could play a role.

(iii) Genetics: Throughout the production of AD, evolution may also play a significant role. Two gene types affect the progression of Alzheimer's disease:

- Gene threat (probability increase)
- Gene deterministic (direct disease cause)

(iii) Techniques for Detection:

The effects of Alzheimer's should be researched to enhance the outcomes of past interventions or to develop better and more accurate methods to detect Alzheimer's using modern inexpensive technology. Alzheimer's techniques of diagnosis are listed as invasive and non-invasive under two separate groups [6]. Invasive approaches include data from the inside of the patient's body through procedures such as a lumbar puncture or blood removal. These methods are designed to identify potential biomarkers that demonstrate an exact Alzheimer's indicator. Several biomarkers have been confirmed as AD measures such as the volume of beta-amyloid and tau in brain spinal blood. Validated biomarkers are therefore complicated and costly, and existing work continues to attempt to identify easier and more practical alternatives. However, these measures are not necessarily secure and relaxed, whereas others are insupportable [7]. On the other side, during the treatment process, non-invasive monitoring is safe and easy. To diagnose the condition of Alzheimer's in the early phases, such as neuroimaging, behavior and emotional examination, or cognitive interventions, a range of strategies based on non-invasive techniques are used.

Neuroimaging, for example, is used to diagnose disease-induced improvements in the patient's brains through MRI or Computed-Tomography (CT). These techniques compare the patient's head MRI data with the relevant data from Alzheimer's patients as shown in Figure 2. However, because buying such photographs usually requires not readily available medical devices and an inconvenient (claustrophobic or noisy) procedure, such methods can not be utilized when measuring vast numbers of individuals on costs and times. In other words, the method is not suitable. This can only be used by advanced machine accuracy learning methods [8].

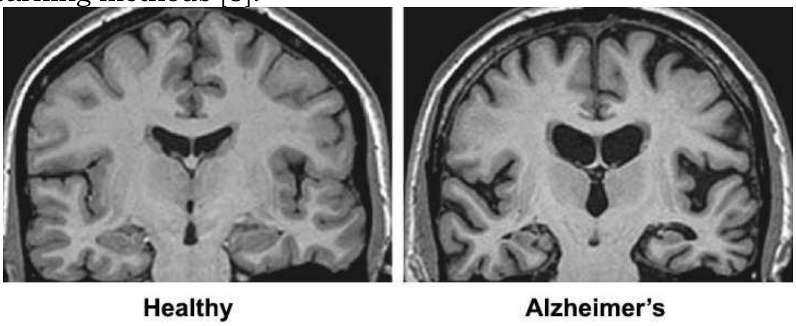

Figure 2. MRI Images for healthy and Alzheimer

The ***problem statement*** of this survey article is to deal with the Big-Data's Dimensional issue in Alzheimer's Disease Health-Care Application. Generally, the MRI data was higher dimensional data leading to computational issues. It was a significant task to analyze the larger dimensionality data by minimal test for investigating the brain-based imaging modality. The current strategies in these methods are isolated into some classifications such as reducing the dimensionality and selecting the proper features from the image. Numerous strategies have been proposed to decrease the dimensionality of the image [9]. But the strategies by lessening the dimensionality of the information bring about the numerous loss in features while loading the data. The techniques involved in selecting the features also lack opting for the best optimal feature to classify the cases.

For this purpose, the ***scope*** of this survey work is to review the ML techniques for Ad classification which were boomed in a great deal of progress in recent years. By utilizing ML algorithms in diagnosis the AD will solve the classification problems in AD's earlier stage. Recently wide techniques based on ML, especially in the investigation of MRI and CT images based became effective when contrasted with traditional techniques, hence it is used as a primary and useful component for analyzing imaging modalities [10].

This contribution of this review article was to survey various methods involved in AADD by its every module such as MRI-Preprocessing, Hippocampus-Segmentation, Feature-Extraction to extract brain structural features, Feature-Selection to select optimal features from the feature vector, and finally, Classification to classify the individual for CN or AD from MRI image datasets.

The article was categorized as Section 1 briefs about the introduction of Alzheimer's, its causes and impact, symptoms, and diagnosis methods, Section 2 lists out some recent research publications on AD classification, Section 3 briefs about the AADD methodologies modules with different techniques, Section 4 covers the discussion about the surveyed work and finally, Section 5 concludes the article with future scope.

Related Works

In [11] the authors proposed a modern multi-task technique for selecting features to retain additional information between modes. The selection of features from each modality forced the limit to establish an intermodality relationship as a separate task by separating the smallness of the features selected from each modality. The selected features in each classification modality were integrated with a Multi-kernel based "Support Vector Machine (SVM)". The "Alzheimer's Disease Neuroimaging Initiative (ADNI)" database has been used to acquire the PET and MRI reference images of subjects that attain 94.37% precision and an area of 0.9724 under the ROC-Curve for AD recognition, 78.80% accuracy, and an MCI recognition AUC of 0.8284.

In [12] the authors proposed a Gene-Selection technology that has been built based on both "Genetic Algorithm (GA)" and SVM. The GA / SVM model is used to pick multiple gene subsets that use a variety of training sets with the most detail. The findings have shown that the procedure proposed is excellent in both selection and classification, achieving 100 percent precision in just 15 genes. In biochemical words, Alzheimer's genes are correlated with at least eight (53%). The genes can also serve as a disease model and can be used to identify potential candidate genes.

In [13] the authors proposed an innovative tool using MRIs to diagnose MCI to AD in earlier times. The model consists of two main stages: selection and classification features. Different forms of monthly variations are used to produce the best results. The findings from the model study in terms of precision, flexibility, and specificity values were higher than those of earlier studies. The results of the work suggested showed that there is a great potential for earlier predictions of an individual from the MCI to AD conversions in their proposed model.

In [14] the authors analyzed the SVM through the MRI brain image data for earlier diagnosis and differential diagnosis of AD-related diseases which began with initial phases such as MRI-Preprocessing, extraction of features, selecting optimal feature subset, and finally classification. The verification strategies and extraction of biomarkers were associated with the MRI. The main advantages and disadvantages were explored for various techniques in their research. Results obtained in the study of classification results and biomarker results were reported to clarify the parameters associated with normal and pathological aging. Finally, unresolved problems and possible future directions were highlighted.

In [15] the authors analyzed an early diagnostic model of the MCI-AD conversion used by MRI. This model has two main steps: the collection and classification of functions. A search algorithm for variable neighborhoods with the most predictive classification features was selected in the first step. In the second step, Linear-SVM was used to categorize the features selected. The study involved 1435 cases and 226 ADNI database features in each cycle and eight separate months. Different monthly combinations have been applied to get the best-predicted results. The results obtained from the proposed model exceeded those of the previous research in terms of precision, sensitivity, and specificity. The results of this work showed that the model was extremely likely to predict and detect early conversion from MCI to AD.

Methodologies

It is important to know how this normal individual condition progresses over time to AD. There must be a certain generic technology. The words advancement and migration refer to the current analysis of the process of progressing to another level in the stage of the disease. CN is the first stage. No symptoms of cognitive impairment are present in a CN individual. If cognitive differences exist that are too severe for the individual or those around him or her to consider but not enough of their everyday life, the individual is converted from CN to "Mild Cognitive Impairment (MCI)". MCI is a mid-stage of age-related AD. The people at MCI are at risk for severe AD transition. MCI samples are moved to the final stage of AD disease. The aim of the current research is focused on reviewing the ML methods over AD classification. Many trained samples are generated by the trained data. The input objects and their target or label can be included in any example. In the case of AD detection, the input item consists of the subject's diagnostic records, and the determination of the subject is expected to move towards AD. The technique with learning capability analyses the training data and generates a model for any input item to determine the correct mark. This allows the technique to render unknown data fairly accessible from the training data. The overall methodologies module for AD based Computer-Aided-Diagnosis (CAD) was shown in Figure 3.

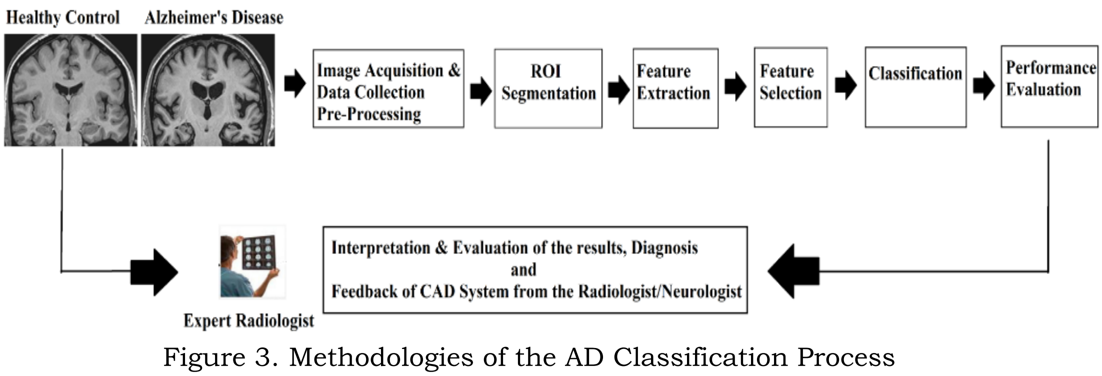

Figure 3. Methodologies of the AD Classification Process

Data Acquisition of MRI Brain-Images

The ADNI and KAGGLE are the two public open data sources for AD that are available for the researchers. These data sets combine containing T1 images of CN, MCI, and AD which were weighted as fast-field MRI echo images. The slices are taken from annual 2 years standardisation of 3T Philips medical systems by in-plan resolution 0.94 * 0.94 mm2, coronary contiguous slices with 160-170, TE = 4.6ms, matrix dimension = 256 = 256 cm2, TR = 2300ms, FOV = 256-260 = 240 cm2, thickness = 1.2mm and the flip angle = 8-9°. These datasets contain almost 500 images.

Pre-Processing of MRI Brain-Image

Datasets could include many combinations of information such as missing values and redundant information that may lead to misleading results. Therefore, before an analysis is carried out, the quality of the data must be improved. After the data set has been obtained from ADNI and KAGGLE. The filtering methods such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Median Filter, etc methods can be used for preprocessing the MRI brain image. The common steps in the preprocessing were smoothing and normalization. The measurements for brain volume were made by head size intracranial, including the brain cortex, white-matter, 3rd and 4th-valve, lower side ventricle, laterum ventricle, cerebellum cortex, cauthane, putamen, pallidum, white-matter brain, hippocampus, and amygdala. The first stage in the AADD is Preprocessing, which improves image data, eliminates redundant and discreet distortions from images, or improves image features by different techniques by this image pre-processing. The output of filtering techniques generates its Peak-Signal Noise Ratio (PSNR) metric. In the further segmentation phase, the better PSNR technique output will move.

Hippocampus-Segmentation

The process of hippocampus segmentation is to separate the hippocampus region from the MRI preprocessed image. The image segmentation methods can be used as an intermediate stage in high-level image processing. The process involved in the method for segmenting the MRI brain image will isolate hippocampus sections from the image into exposed areas that fit an agglomeration of pixels with identical characteristics. Based on the intensity of pixels or the position of the pixels, segmentation approaches can be performed. Information on the direction of pixels can be obtained from a given pixel set by calculating their distance. For such methods, segmentation strategies that are independent of their pixel location are referred to as non-contextual strategies and threshold-based. In the circumstances of non-uniform intensity, thresholding techniques are however not successful. Commonly, the methods of segmentation are defined as similarity and discontinuity based. The segmenting process on the edge model is based on the value of the intensity in which the MRI image in the boundaries is partitioned to separate areas. Due to the multiple artifacts having different intensity values in each MRI image. If the pixel values are identical, though, they are united to shape individual artifacts and this technique is called region-based segmentation. An example of this method is threshold-based segmentation.

Edge-based segmentation

The fundamental principle behind this method is to consider the abrupt change from the top left to the bottom right of the MRI image by its gray-level values. A difference in the value of amplitude correlates to an entity limit until the noise is present and a significant boundary arises as both of these points are joined together. The identification of the edge can be seen as a high-pass filtering mechanism that emphasizes transformations and avoids smooth regions. Connaissance of the variance of amplitude can be accomplished by a differential operator with a smooth area with zero, and sudden transitions with non-zero. Furthermore, a negative value would imply a decrease in strength and vice versa for the differential operator.

Region-based segmentation

This type of approach divides an image or class regions based on the general characteristics of separate artifacts comprising an image. The most common characteristics are spectral properties, object-specific patterns, intensity values, or intensity parameters. The two sub-approaches under this segmentation are the increase in the region and regions that are separating and combining. The first collection of small areas is clustered into larger regions based on regional development. It starts by choosing a pixel or a group of pixels from the random seed. The first seed pixels are to create a new area and are paired with all the next pixels that meet the requirement of similarity. Until the process is finished, another seed pixel will be chosen and the same process will be replicated that covers all pixels in the image. The option of seed points differs according to the data form of the image.

Color is regarded as the criterion for selecting a seed point in some applications, while the texture is used for other applications. However, for the area where the seed point is picked first, there will still be a distortion. The segmentation result will never be special and will be decided by the collection of the seed stage. Also, the seed point selected at the edges would worsen the problem. The collection of seed points often improves with prior experience of the image being segmented. In regions of identical probability distributions, where the distribution function is called the reference criterion.

Clustering-based segmentation

This segmentation method is often based on the separation of each entity from its surroundings. Clustering can be described as a grouping of data points within a parameter space that separates various data sets with similar strengths. The clustering method senses the underlying data structure. A community of pixels with similar hippocampus characteristics may be identified and executed when a particular criteria feature is reached. The clustering mechanism is not special and is based on the quality of the data collection. Choosing a specific clustering algorithm depends on changing intensity environments, imagery properties and the specific object being imaged plays a key role. Regardless of whether a chosen clustering strategy is chosen, a threshold is often selected as the fundamental mechanism and objective feature is reduced. Techniques for clustering may be categorized as hard and soft based on how the data components are distributed to individual classes. In hard clusters, the fact that the data element belongs to a category is differentiated. In medical image processing, this form of prejudice is not valid owing in certain situations to uncertain limitations. Due to overlap borders with Amygdala, the process gets trivial when segmenting Hippocampus from a preprocessed MRI image for classifying AD.

Previously, numerous automatic methods had been created, which were intended for the morphometric analysis of the Hippocampus. An atlas was used in atlas-based approaches to extract the Hippocampus, and it's a reference template for a broad spectrum of populations. Different affine transformations were rendered to translate the reference template coordinates into the images for the test. Since, it is impossible to provide a single template that contains multi-atlas segmentation due to heterogeneity across age, gender and imagery modalities.

For deformable segmentation, Hippocampus was first described to continuously modify the shape to fit into the desired object and eventually eliminate it. Hippocampus segmentation strategies require prior expertise to delineate the Hippocampus as the most popular atlas-based and deformable model. In recent days for segmenting the Hippocampus region from the MRI brain image for AD classification, few researchers used K-Means Clustering, Fuzzy C-Means Clustering, etc. These algorithms help to identify AD at an early stage by the precise recognition and tracking of irregular tissue losses in entity borders. Also, multi-level thresholding has been established for the segmentation of Hippocampus length.

Hippocampus Feature Extraction

The extraction of the required features is the mechanism by which the segmented MRI data is interpreted or converted. The extraction of features is based on the following strategies. They were methods based on manual interpretation or soft computing (automation). A manual interpretation methodology tests feature dependent on interpretation such as color and choosing the perfect statistical portrayal. A soft computing framework for extracting such functionality of features is done through the machine-centered method.

Feature Extraction Based on Texture

Textures seem to be the key function for characterizing the structure and form of a segmented hippocampus region from the MRI brain-image. Textures have been described as an object composed of particles or pixel classes that relate to each other. Each set of pixels was called components of texture or texels. A texel is a statistical indicator of a hippocampus region's strength in the segmented image. The techniques of characterizing texture thus fall into two groups. It is quantitative and systemic. Quantitative approaches discern texture by statistical image-based strength distribution. One of the characteristics of distinguishing textures is the spatial distribution of grey values. Quantitative methods are used to analyze the spatial distribution of grey values. This is done by calculating local characteristics at every spot of a hippocampus area and a sequence of statistical data from local feature distribution is derived. By defining systemic primitives and their laws of positioning, the systemic methods will reflect texture.

Feature Extraction Based on Histogram

The key truth underlying the descriptors of the "Histogram Oriented Gradient (HOG)" is that it is by intensity gradients or boundary directions of the hippocampus area from the segmented MRI brain-image. This could be accomplished by splitting the segmented image into smaller interconnected areas, labeled cells, and HOG directions or edge orientations of the pixels in the cell are compiled for each cell. The descriptor is therefore a mixture of these histograms. For greater precision, the local histograms may be generalized by measuring the intensity calculation over a wider image region called a block. All cells inside the block are normalized for the use of this frequency. This normalization contributes to a greater invariance in the light or shadow shifts. The HOG descriptor has certain main benefits relative to various descriptor approaches. The process preserves an invariance for spatial and spectrometry conversions, apart from object-orientation, though this descriptor acts on localized cells. These improvements will appear only in broader spatial areas. This makes the HOG descriptor especially appropriate for AD based diagnosis.

Feature Extraction Based on Intensity Histogram

It's a graph that describes the given image with a different intensity value that indicates the number of pixels in one graph. There are 256 potential intensity values for an 8-bit grayscale image. The histogram of intensity is the statistics of the first order. The histogram is constructed for the segmented hippocampus region and from that the four characteristics are taken from the histogram. They are entropy, third moment, uniformity, and smoothness which are evaluated. The properties are created by counting the pixel number of every intensity level.

Feature Extraction Based on Second-Order Statistical Features

While the statistical characteristics of the first-order explain the grey "Region Of Interest (ROI)" distribution, these features don't provide details on the spatial distribution of the different levels of grey for the segmented hippocampus area within the ROI. These kinds of knowledge can be derived from the matrixes of co-occurrence and run-length. Functionalities focused on histograms are local. Such attributes do not take spatial details into account. For this reason, the space-based co-occurrence matrixes "hd (i,j)" are described with features known as histogram-based features of the second order. These properties are based on the mutual distribution of pixel pairs. For calculating the joint likelihood distribution between pixels, the distance 'd' and angle 'θ' residing at a given quarter is used. The most successful characteristics for the classification of CN, MCI, and AD are considered to include the ROI characteristics for the extraction of the Hippocampus Volume (HV) features from each MRI file.

Feature Selection

Since many MRI samples have a 3D layout because every other MRI series often includes several images, contributing to a wide range of information, it is especially important to minimize the proportion of information collected by the feature extraction process. Compared to the comparatively limited range of patients, the large size of the space input is a well-recognized issue called the dimensionality curse, so a certain sort of functionality selection is sometimes implemented to overcome it. The selection of features often has advantages in speeding up the phase-in testing and in increasing the precision of the AD classification. The development in selecting features has two key instructions: Selection of subset and Rating of features. A training method is applied to the subject of selecting a specific subset of features to work on thus avoiding other rest. The rating of features is a simple term by rating higher for relevant and lower for irrelevant features. The collection of subsets is generally classified into 3 categories: Filtres, Wrapper, and Embedded.

Filtering Approaches

In an attempt to assess improper description, the ***filtering approach*** uses the statistical characteristics of features. Filter strategies determine the significance of characteristics mostly by analyzing the characteristic features of results. In certain instances, a characteristic significance scoring is determined, and poor scoring characteristics are discarded. The select group of features is then presented mostly to the classified model as an input. The benefits of filtering strategies are that they are quick and fast computation time and autonomous from the prediction model, efficiently scale into very large dimensional data sets. As a consequence, the feature choice must be achieved only once and multiple classification devices will then be tested. A typical downside to filtering methods is the ignoring of association with the classification model (search of the sub-set space of the function is isolated from the hypothesis space search), which ensures that the majority of techniques suggested are standardized. This ensures that each element is regarded independently and therefore it neglects functionality dependency. In contrast with other feature selecting methods, this may contribute to a poorer recognition rate.

Wrapper Approaches

Based on their prediction accuracy, ***Wrapper*** considers the entire collection of parameters to scoring functional subsets and optimizes a success criterion for the following process which employs the corresponding subset for AD classification.

Embedded Approaches

The final method was ***Embedded*** which was chosen in the ADAD model building phase. An integrated feature selection mechanism utilizing SVM was used by many researchers. This type of method uses the ranking data rather than directly experiencing rankings as a flat category, as regards feature ranking research methods. Finally, a collection of features with the highest significance and minimal resemblance is to be found, i.e. to avoid choosing repetitive characteristics which include Greedy-Evolutionary algorithms to overcome the optimal solution.

In response to the increasing number of missing feature dependencies, a variety of multi-variant filtering strategies were developed, aimed at the inclusion of feature relationships to a certain extent. Filters and wrappers always utilize their training method as an important tool for data collection. The preference of features can take care of its classifier features. To test subsets, wrappers be using the assessment feature of the classification error rate generated with the ML methods. That feature of wrappers provides a higher precision than basic filtering for subset filtering. For each sub-system assessment, the wrappers need to build a model which often takes time.

AD Classification Based On Ml Approaches

Classification is used to examine a given dataset and takes every instance of it for assigning the instance to a particular class where the classification error is minimum. Classification is used to accurately define the important data classes within a given dataset. The classification process comprises two steps. In the first step, the model is generated by relating the classification algorithm to the training data set. In the second step, the extracted model is verified against a predefined test dataset to improve the accuracy. Classification is defined as the process to allocate the class label from a dataset whose class label is not known. The classification technique forecast the target class for every data point. With classification for classifying the AD from MRI brain images, the risk factor is linked to the patients for medical database through examining patterns of diseases. In binary classification, two classes, namely AD patient and CN individual are taken while the multiclass approach includes more than two classes, namely AD, MCI, and CN individuals. Dataset is divided into training and testing datasets. Classification is used for predicting the outcome depending on the input. The training set comprises a set of attributes to predict the results. For predicting the output, it determines the relationship between the attributes. Goal and prediction are the outcomes. The prediction set comprises the same set of attributes similar to the training set. In a medical database of ADNI and KAGGLE, the training set comprises the information regarding a patient that is recorded at an earlier stage. The presence or absence of Alzheimer's Disease is the prediction attribute. Here we surveyed some main ML methods for classifying Alzheimer's.

Decision-Tree (DT) Classifier

The DT is a traditional approach for representing the classifier. It was constructed using the available data. It is similar to the flowchart where every non-leaf node denotes a test on a particular attribute. Every branch indicates an outcome of the test and every leaf node includes the class label. The root-node is the first-node of DT. Knowledge of the domain was not needed for deciding any problem. The DT is a research analyst for computing conditional probabilities. By using this, decision-makers select the best alternative, and traversal from root to leaf represents unique class separation depending on maximum information gain. The DT is employed in AD classification by few researchers. The DT in AD classification is self–explanatory. The set of rules is computed with help of DTs.

The DT representation is an essential one to represent any discrete-value classifier as it is capable to manage attributes and numeric input attributes. When datasets include the missing or the DT learning is the classification technique with better performance. Depending on input variables, DT designs the prediction model. Every internal node represents the test on attribute value. Every branch in the DT denotes the attribute value. A leaf node executes the classification process when it is reached. It provides exact predictions and idea descriptions. These are structured in a top-down manner. Many algorithms exist in DT construction termed DT inducers. The key aim is to recognize the optimal decision tree by reducing the generalization error.

Nearest-Neighbors (NN) Classifier

The NN rule differentiates the AD classification of unknown data points because of the closest neighbor whose class is known. The NN is calculated based on the estimation of 'k' that represents how many NNs are taken to characterize the data point class. It utilizes more than one closest neighbor to find out the class where the given data point belongs termed "KNN". The MRI data samples are required in memory at the run time called a memory-based technique. The training points are allocated weights based on their distances from the sample data point. However, the computational complexity and memory requirements remained a key issue for classifying the AD. For addressing the memory utilization problem, the size of MRI data gets minimized. The repeated patterns without additional data are removed from the training data set.

Naive-Bayes (NB) Classifier

The NB Classifier technique is functioned based on Bayesian-Theorem. For classifying the AD from CN the designed technique is used when the dimensionality of MRI features is high. Bayesian-Classifier is used for computing the possible classified AD output depending on the MRI features. It is feasible to add new raw data at runtime. An NB classifier represents the presence (or absence) of a feature (attribute) of a class (AD, MCI, or CN) that is unrelated to the presence (or absence) of any other feature when the class variable is known.

Support-Vector-Machine (SVM) Classifier

The SVM is used in many applications like medical, and military for classification purposes. It was employed for classification, regression, or ranking functions. It depends on statistical learning theory and the structural risk minimization principle. It determines the location of decision boundaries called hyperplane for optimal separation of classes for AD classification as described in Figure 4.

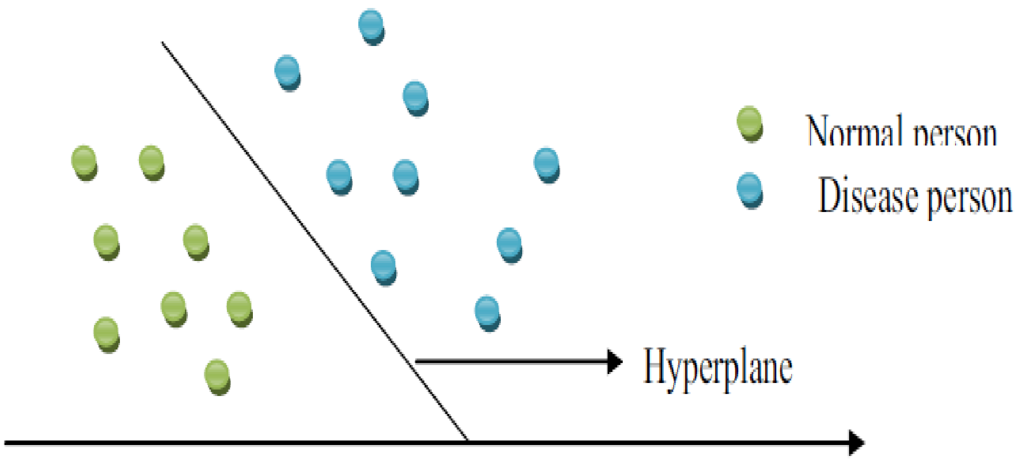

Figure 4. SVM Classification

Margin maximization through creating the largest distance between separating hyperplane and instances on either side is employed to minimize the upper bound on expected generalization error. The classification accuracy of SVM does not depend on the dimension of classified entities. The data analysis in SVM is based on convex quadratic programming. It is expensive as quadratic programming methods need large matrix operations and time-consuming numerical computations. Training time for SVM balances quadratically in the number of examples for a more efficient training algorithm.

SVM is extended to learn non-linear decision tasks by projecting the MRI features into high-dimensional feature space through kernel functions and devises linear AD classification problems in feature space. The resultant feature space is larger than the dataset size that is not feasible to store. For addressing these issues, decomposition-based algorithms are introduced. The key aim of the decomposition method is to divide the variables into two types, namely the working set and the fixed variable set. The set of free variables is a working set that gets updated for every iteration count. The values of variables that get fixed temporarily are fixed variable sets. The process gets repeated till the AD, MCI, and CN class conditions are met. SVM is used for binary classification and it is not easy for extending it to address multi-class classification problems. The key objective of multi-classification for classifying the AD from MRI features is to decompose multi-class issues into many two-class problems that are addressed using SVM.

Artificial-Neural-Network (ANN) Classifier

The ANNs are a kind of computer architecture inspired by biological Neural-Networks. The designed algorithm for classifying the AD approximates the functions based on a large number of unknown inputs. The ANN is the system of interconnected neurons that compute values from inputs. It is capable of ML and Pattern-Recognition because of its adaptive nature. An ANN functions by creating connections between different processing elements corresponding to a single neuron in the biological brain. The neurons are constructed or simulated by a digital computer system. Each neuron has many input signals. Then, a single output signal is produced depending on internal weighting. This output signal is taken as input for another neuron. The neurons are interconnected and arranged in different layers. The input layer collects the input and the output layer generates the final output. One or more hidden layers are inserted between the input layer and output layer. This structure failed to forecast or identify the exact data flow. The construction of ANN for AD classification is shown in Figure 5.

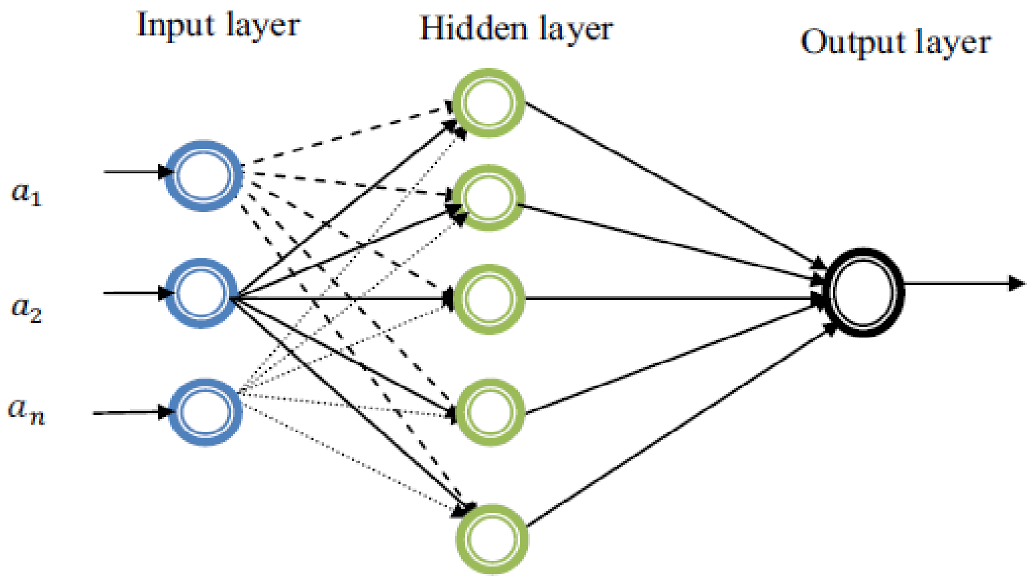

Figure 5. ANN Classification

Performance Metrics

The above classifiers identify the subject's MRI image as CN, MCI, and Pure AD. The performance of each classifier was evaluated using the following metrics Accuracy, Precision, Recall, and F-Measure.

Discussion of the Study

- In this survey we found out some observations of well-conducted research centered on ML for AD classification are provided in the preceding methodologies, there seem to be certainly numerous interesting research with similarly outstanding outcomes.
- Particularly in AD assessment and detection, numerous research demonstrated how the results must be evaluated and characterized.
- The ability to detect any flaws in MRI data input, sampling strategy, verification, or deployment is crucial among those who assess research and those who plan to apply ML approaches.
- Throughout this survey, the most prevalent issues we found were input sizes (big data), features handling, and affirmation, according to the research reviewed.

- These approaches cannot handle a big populace of input since they are simpler to get greater accuracy with limited datasets. Although smaller sample sizes might lead to excessive training, it's been shown that a big data set has numerous benefits for reliability, precision, and consistency.
- Although the reviewed approaches achieved a maximum of 94.6 percent accuracy, the model's resilience is questioned due to its usage of unverified data and its size. A significant amount of research is based on data that has not been proved pathologically, which might lead to inaccuracies in the findings.
- However, since such data may be collected from medical centers, there has been no justification for considering incorporating these data in the relevant research.
- The proportion of features to contexts affects the outcome. This research did not take into consideration the number of qualities or the basic knowledge about them.
- For ML to provide useful results, good quality of the data and careful selection of significant features are both essential.
- This survey finds that the quality of information and consistency were not adequately protected by the researchers' descriptions of the procedures they employed to protect the data.
- In addition to data quality, the selection of features is also crucial. For example, the MRI feature vector might not have been useful in the future because of histologist evaluations.
- To keep up with the times, classifiers must've been capable to upgrade selected features. In the same way, the training/testing outcomes must be stated properly.
- A large percentage of methods concentrate only on the AD classification module which leads to ignoring or incorrectly classifying the minorities.
- Selecting a dominating class with poor class forecasting leads to worse performance for AD classification as a consequence of this class imbalance.

Conclusion

In this article, we assess most current research articles on the prediction and prognosis of Alzheimer's classification employed by ML techniques. Accordingly, the most current developments in ML have been highlighted, covering the sorts of data that are being employed and the effectiveness of ML algorithms in detecting the initial phases of Alzheimer's disease. In comparison to traditional statistical methods, the ML seems to enhance the accuracy rate. In addition, the medical diagnostic wasn't 100% correct since pathological validation was not given, resulting in ambiguity in the expected findings. In this study, publications with pathologically proved facts were examined, as was the method for dealing with the class imbalance and overtraining difficulties. In an attempt to minimize expense, several models are constructed on a singular modality rather than merging multiple modalities, according to this analysis. If pathology evidence is available, we guess it would improve the efficiency and reliability of the findings, whereas a well-balanced class will improve the classification. Based on the review findings we suggest that the ADAD framework needs to be developed by concentrating on each module such as MRI-Preprocessing, Hippocampus-Segmentation, Feature-Extraction to extract brain structural features, Feature-Selection to select optimal features from the feature vector, and finally, Classification to classify the individual for CN, MCI and AD from MRI brain-image datasets for higher accuracy rate.

References

[1]. M. W. Weiner et al., “The Alzheimer’s disease neuroimaging Initiative: A review of papers published since its inception,” Alzheimer’s Dementia, J. Alzheimer’s Assoc., vol. 8, no. 1, Feb. 2012, Art. no. S168.

[2]. S. Adaszewski, J. Dukart, F. Kherif, R. Frackowiak, B. Draganski, and A. D. N. Initiative, “How early can we predict Alzheimer’s disease using computational anatomy?” Neurobiol. Aging, vol. 34, no. 12, pp. 2815–2826, Dec. 2013

[3]. B. Dubois et al., “Advancing research diagnostic criteria for Alzheimer’s disease: The IWG-2 criteria,” Lancet Neurol., vol. 13, no. 6, pp. 614–629, 2014.

[4]. F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, and A. Ortiz, “A spherical brain mapping of MR images for the detection of Alzheimer’s disease,”Current Alzheimer Res., vol. 13, no. 5, pp. 575–588, 2016.

[5]. G. H. Weissberger, J. V. Strong, K. B. Stefanidis, M. J. Summers, M. W. Bondi, and N. H. Stricker, “Diagnostic accuracy of memory measures in Alzheimer’s dementia and mild cognitive impairment: A systematic review and meta-analysis,” Neuropsychol. Rev., vol. 27, pp. 354–388, Sep. 2017.

[6]. Liu, F, Wee, CY, Chen, H & Shen, D 2014, 'Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification', NeuroImage, vol. 84, pp. 466-475. [6]. G. B. Frisoni et al., “Strategic roadmap for an early diagnosis of Alzheimer’s disease based on biomarkers,” Lancet Neurol., vol. 16, no. 8, 2017, Art. no. 661676.

[7]. A. Sarica, A. Cerasa, A. Quattrone, and V. Calhoun, “Editorial on special issue: Machine learning on MCI,” J. Neurosci. Methods, vol. 302, pp. 1–2, May 2018.

[8]. H. Li, M. Habes, D. A. Wolk, and Y. Fan, “A deep learning model for early prediction of Alzheimer’s disease dementia based on hippocampal magnetic resonance imaging data,” Alzheimer’s Dementia, vol. 15, no. 8, pp. 1059–1070, Aug. 2019.

[9]. L. Cai et al., “Functional integration and segregation in multiplex brain networks for Alzheimer’s disease,” Frontiers Neurosci., vol. 14, pp. 14–51, Feb. 2020.

[10]. H. Yu, X. Lei, Z. Song, C. Liu, and J. Wang, “Supervised network-based fuzzy learning of EEG signals for Alzheimer’s disease identification,” IEEE Trans. Fuzzy Syst., vol.28, no. 1, pp. 60–71, Jan. 2020.

[11]. Kumar, S. (2022). A quest for sustainium (sustainability Premium): review of sustainable bonds. Academy of Accounting and Financial Studies Journal, Vol. 26, no.2, pp. 1-18

[12] Allugunti, V.R. (2019). Diabetes Kaggle Dataset Adequacy Scrutiny using Factor Exploration and Correlation. International Journal of Recent Technology and Engineering, Volume-8, Issue-1S4, pp 1105-1110.

[13] Viswanatha KKRC, Reddy A, Elango N M (2019). Diabetes Kaggle Dataset Adequacy Scrutiny using Factor Exploration and Correlation, *International Journal of Recent Technology and Engineering (IJRTE)* Vol. 8.

[14] S. Z. Paylakhi, S. Ozgoli, and S. H. Paylakhi, "A novel gene selection method using GA/SVM and fisher criteria in Alzheimer's disease," 2015 23rd Iranian Conference on Electrical Engineering, Tehran, 2015, pp. 956-959, DOI: 10.1109/IranianCEE.2015.7146349.

[15]. Atas, PK, Tufan, K & Şevkli, AZ 2016, 'A variable neighborhood search based feature selection model for early prediction of the Alzhemier's disease', 2016 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), pp. 1-4.

[16]. C. Salvatore, P. Battista, I. Castiglioni Frontiers for the Early Diagnosis of AD by Means of MRI Brain Imaging and Support Vector Machines Curr. Alzheimer Res., 13 (5) (2016), pp. 509-533

[17]. Nusman CM, Hemke R, Lavini C, et al. Dynamic contrast-enhanced magnetic resonance imaging can play a role in predicting flare in juvenile idiopathic arthritis. European Journal of Radiology. 2017;88(Supplement C):77-81.