International Journal of Computing Science and Information Technology, 2018
ISSN: 2278-9669, April 2018 (http://ijcsit.org)

Machine Learning Technique For Alzheimer's Disease (Ad) – A Comprehensive Review

Sumanth S [1].
Research Scholar,
Periyar University;
Salem, Tamilnadu, India

A Suresh [2].
Principal,
Maisurii Women's College of Arts and Science,
Kakapalayam, Salem, Tamilnadu, India

Received January 2018

Abstract

Alzheimer's disease (AD) has become a common neurodegenerative brain disease in elderly people for which exact cause is not known. AD is one type of dementia which causes memory loss. It is very important to obtain an accurate diagnosis of AD, so this paper presents literature review on the use of machine learning and pattern classification methods for accurate detection of AD.

Keywords: Alzheimer's Disease, neurodegenerative brain disease, machine learning and pattern classification methods

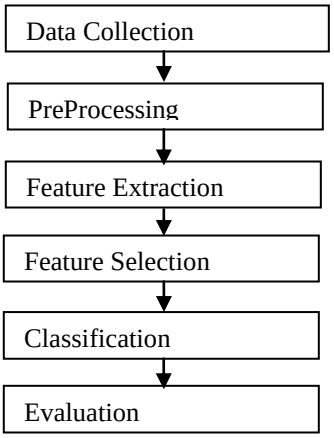
1. Introduction

Alzheimer's disease (AD) has become a common neurodegenerative brain disease in elderly people There are more than 80 different conditions that can cause dementia, to name few Alzheimer's disease(AD), Vascular Dementia, Dementia with Lewy Bodies or Lewy Body Dementia(LBD), Frontotemporal degeneration (FTD) or Frontal Lobe Dementia, HIV dementia etc., Among all AD is the most common cause of dementia in the elderly people, it also affects people with a lesser age group of 40s and 50s though it is less common. Also all elders are not demented. Again people of both sexes can be affected with dementia but it is more common in men than women.

According to a report published by Alzheimer's disease International 46.8 million people worldwide are living with dementia and this number will almost double every 20 years, and the number of dementia patients will be 74.7 million in 2030 and 131.5 million in 2050. The total estimated worldwide cost of dementia in 2015 is US$ 818 billion, by 2018; dementia will become a trillion dollar disease, rising to US$ 2 trillion by 2030.
User cooperation not require for gait recognition. But fingerprint, iris, signature, hand geometry and speech necessarily require the user cooperation.

2. Methodology

Fig. 1 shows an overview of steps in medical image processing.



Phase 1: Data Collection

The first phase in the medical image analysis involves collecting a huge number of medical images/data related to a particular target disease. The impact of medical imaging on medicine is rapidly growing with the recent development of advanced imaging techniques like Computed Tomography (CT) for diagnosing many cancers (lung, liver, and pancreatic cancer), Ultrasound, Magnetic Resonance Imaging (MRI) to differentiate normal and diseased soft tissues of the body, Functional MRI (fMRI), Single Photon Emission Computed Tomography (SPECT), X-Ray and Positron Emission Tomography (PET) to know the functioning of tissues and organs. The medical images taken from different imaging modalities

are stored in the medical database for further research by various research centers for better training the proposed system.

Data can also be obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database http://adni.loni.usc.edu/. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

Phase 2: Data Preprocessing

The images usually suffer from low contrast, blur, noisy and diminished colors and need to be pre-processed for enhancing the quality by removing the noise. The literature survey reveals that there are many improved algorithms are available for enhancing the image quality and also involves the separation of background from the object followed by partitioning the regions of interest with image segmentation based on various features like texture, color and depth measurements. The selection of a segmentation technique is decided by the type of image and characteristics of the problem (disease) selected.

Phase 3: Feature Identification and extraction

Third stage is feature identification and extraction. The accuracy of classification decreases with the increase in number of features extracted from the image. We can call it as the curse of dimensionality in the vision of classification. To deal with this problem, feature optimization is a suggestible choice.

Phase 4: Feature Selection

Feature selection is a process which transforms high dimensional data into low dimensional data. Feature selection is described as an optimization problem and aims to reduce the dimension of the dataset, decrease the computation time and improve the precision with a feature optimization algorithm by removing redundant, unrelated and noisy features. It consists of different fields such as machine learning, pattern recognition, data mining, medical data, etc.

The several feature selection approaches:

1. The Principal Component Analysis (PCA) is the most feature extraction method. It reduces the number of available variables
2. The Linear Discriminate Analysis (LDA) which reduces the dimensionality of the discriminatory information. The Latent Semantic Analysis finds co-occurrences in document. The main objective is to produce the mapping called Latent Topic Space.
3. The Independent Component Analysis (ICA) is used to find a linear representation of non-Gaussian data. It is used to extract features.

It consists of different fields such as machine learning, pattern recognition, data mining, medical data, etc.

It is classified into three categories: filter method, wrapper method and embedded methods.
The advantage is it has low computational complexity.

Filters work as a preprocessing step for classifiers and are completely independent of the classification. The popular filter approaches are:

1. Chi-square approach
2. Correlation approach
3. Correlation method
4. Information Gain

Wrapper method uses the machine learning to select the subset of variables based on the predictive power. The basic idea is to use the prediction to evaluate the effectiveness of the subset of variables.

The embedded approach, the idea is to jointly train the classifier and select the relevant features. The features are irrelevant when the scaling factor is small. The features are relevant when the scaling factor is large.

So, it is essential to develop an efficient algorithm to extract relevant good features from the medical images for classification.

Phase 5: Classification

Image classification is a process of labeling the images into a predefined category. Different types of classification techniques are being used by the researchers to classify and segmentation of medical image data especially MRI data to detect abnormalities found in different parts of the human body, this study confined to utilization of these techniques for classification and segmentation of medical image data as special case.

Phase 6: Evaluation

The research is to be carried out in this area to distinguish AD from No cognitive impairment and MCI, and also to find the best classification technique that provides good accuracy to distinguish between AD, MCI and NC.

3. Classification Techniques

Data classification is a two phase process:

Training Phase: Training phase is the first phase, where the classifier algorithm builds classifier with the training set of tuples.

Classification Phase: Classification phase is the second phase where the model is used for classification and its performance is analyzed with the testing set of tuples.

Various Classification algorithms are being used by the researchers to classify medical image data especially MRI data to detect abnormalities found in different parts of the human body (brain for AD) are explained below:

3.1. Decision Tree Algorithm

The decision tree is one of the classification algorithms. The learning algorithm applies a divide and-conquer strategy to construct the tree [1]. The sets of instances are associated by a set of attributes. A decision tree comprises of nodes and leaves, where nodes represent a test on the values of an attribute and leaves represent the class of an instance that satisfies the conditions. The outcome is „true‟ or „false‟. Rules can be derived from the path starting from the root node to the leaf and utilizing the nodes along the way as preconditions for the rule, to predict the class at the leaf. The tree pruning has to be carried out to remove unnecessary preconditions and duplications.

3.2 Support Vector Machine

SVM algorithms are based on the learning system which uses the statistical learning methodology and provides good generalization performance for classification. So, they are popularly used for classification. In SVM technique, the optimal boundary, known as hyper plane, of two sets in a vector space is obtained independently on the probabilistic distribution of training vectors in the set. This hyper plane locates the boundary that is most distant from the vectors nearest to the boundary in both sets. The vectors that are placed near the hyper plane are called supporting vectors. If the space is not linearly separable there may be no separating hyper plane. The kernel function is used to solve the problem. The kernel function analyses the relationship among the data and it creates a complex divisions in the space.

3.3 Random Forests

Random forest algorithm is one of the best among classification algorithms and is able to classify large amounts of data with high accuracy. It is an ensemble learning method building models that constructs a number of decision trees at training time and outputs the modal class out of the classes predicted by individual trees. It is a combination of tree predictors where each tree depends on the values of a random vector sampled independently with the same distribution for all the trees in the forest. The basic principle is that a group of "weak learners" can come together to form a "strong learner".

3.4 Evolutionary Algorithms

A Genetic Algorithm (GA) is an evolutionary and stochastic method for finding optimal solutions in large and complex search spaces. A GA is inspired by natural evolution: a population of encoded candidate solutions (called "chromosomes") is evolved through generations using genetic-like operations such as crossover and mutation. At each generation, solutions are probabilistically selected based on their fitness, in order to generate offspring and create the next generation. The initial population is randomly generated, and at each generation, every candidate solution is evaluated against an objective function in order to gain a fitness score. In a learning system, the objective function is typically the measure of the accuracy of a candidate over a training set of instances.

3.5 Swarm Intelligence

Swarm intelligence (SI) is a computational intelligence technique to solve complex real-world problems. It involves the study of collective behaviour of individuals in a population who interact locally with one another and with their environment in a decentralized control system. The inspiration often comes from nature, especially biological systems. The agents follow very simple rules, and although there is no centralized control structure dictating how individual agents should behave, local and to a certain degree random interactions between the agents lead to an "intelligent" global behaviour which is unknown to the individual agents. Some of the popular SI algorithms are Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) and Artificial Bee Colony (ABC).

3.6 Fuzzy Logic

Zadeh [2] introduced the fuzzy set theory; a major contribution of fuzzy set theory is its capability of representing vague data. Fuzzy sets and fuzzy logic are powerful mathematical tools for modeling: uncertain systems in industry, nature and humanity, and facilitators for common-sense reasoning in decision making in the absence of complete and precise information. A fuzzy number is characterized by a given interval of real numbers, each with a grade of membership between 0 and 1.Fuzzy logic based clustering technique is frequently utilized for classification of medical image data.

3.7 Artificial Neural Network (ANN)

Artificial neural network (ANN) [3] is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing. Some of the architectures of ANN are explained below:

3.7.1 Multi-Layer Neural Network (MLNN)

Multilayer Neural Networks [3] solve the classification problem for non-linear sets by employing hidden layers, whose neurons are not directly connected to the output. The additional hidden layers can be interpreted geometrically as additional hyper-planes, which enhance the separation capacity of the network. Multi-layer neural network is mostly used for classification of different categories of data .A popularly used MLNN is back propagation network (BPN) with gradient descent. Back propagation artificial neural network (BPANN) is a neural network technique which is able to train nonlinear data and is based on gradient descent. This network is trained with popular error back propagation algorithm (EBPA). This algorithm has two passes: feed forward phase in which output is calculated and feed backward phase in which the calculated error is propagated back to the network to adjust the weights.

3.7.2 Polynomial Neural Network (PNN)

Polynomial neural networks (PNN) are multilayer perceptrons of neuron-like units which produce high order multivariate polynomial mappings. These are tree structured hierarchical cascades of first-order and second order activation polynomials in the nodes, and input variables passed from the leaves. The activation polynomial outcomes are fed forward to their parent nodes, where partial polynomial models are made.

3.7.3 Radial Basis Function Neural Network (RBFNN)

Radial basis functions [3] are powerful techniques for interpolation in multidimensional space. A RBF is a function which has built into a distance criterion with respect to a center. Radial basis function (RBF) networks are feed-forward networks trained using a supervised training algorithm. It has one input layer, one output layer and one hidden layer generally with special type of activation function known as basis functions one can use a suitable basis function like radial basis, polynomial, sigmoid or linear basis function as per suitability of data pattern. These are also known as kernel type and can be changed to tune the network.

In this network, the determination of number of neurons in the hidden layer is very important. This is because it affects the network complexity and the generalizing capability of the network. If there are an insufficient number of neurons present in the hidden layer then the RBF network cannot learn the data adequately. Poor generalization will occur if there is more number of neurons.

3.8 Hybrid Techniques

To overcome problem of individual techniques hybridization is required, a suitable hybrid technique [4] with combination of two or more intelligent techniques like Neuro-Fuzzy, Neuro-Genetic or Neuro-Fuzzy-Genetic can be utilized. Authors [4] are currently using hybrid techniques for medical image data classification. Some very well known hybrid techniques are explained below:

3.8.1 Neuro-Fuzzy Technique

A Fuzzy Neural Network or Neuro-Fuzzy System [5] is a learning machine that finds the parameters of a fuzzy system by exploiting approximation techniques from neural networks. This means that the main intention of Neuro-Fuzzy approach is to create or improve a fuzzy system automatically by means of neural network methods. A Neuro-Fuzzy system based on an underlying fuzzy system is trained by means of a data-driven learning method derived from Neural Network theory. It can be represented as a set of fuzzy rules at any time of the learning process, i.e. before, during and after. Thus the system might be initialized with or without prior knowledge in terms of fuzzy rules. The learning procedure is constrained to ensure the semantic properties of the underlying fuzzy system. A Neuro-Fuzzy network is a fuzzy inference system in the body of an artificial neural network. Depending on the Fuzzy Inference System (FIS) type, there are several layers that simulate the processes involved in a fuzzy inference like fuzzification, inference, aggregation and defuzzification.

3.8.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS, developed by Jang [6] is an adaptive network incorporates the concept of fuzzy logic into the neural networks, and has been widely used in many applications. ANFIS largely removes the requirement for manual optimization of the fuzzy system parameters. An adaptive network is network of nodes and directional links. Associated with the network is a learning rule – for example back propagation. It's called adaptive because some, or all, of the nodes have parameters which affect the output of the node. These networks are learning a relationship between inputs and outputs. By using a hybrid learning procedure, the proposed ANFIS can construct an input-output mapping based on both human knowledge (in the form of fuzzy if-then rules) and stipulated input-output data pairs.

3.8.3 Neuro-Genetic Technique

The Neuro-Genetic [7] model is a hybrid model which exhibits the characteristics of both ANN and GA. It can be used as the tool for decision making in order to solve the complex nonlinear problems. In this method first we define a network structure with a fixed number of inputs, hidden nodes and outputs. Second we employed the GA in the learning phase of the network, as it is capable to search in a large search space. The hybridization of ANN and GA is able to select the optimal weight sets as well as the bias value for the classification.

CONCLUSION

This paper provides a comprehensive review on the use of machine learning and pattern classification techniques used for medical image processing. Research is to be carried out in this area to identify the best classification technique that provides good accuracy to distinguish AD from No cognitive impairment and MCI.

REFERENCE

[1] J. HAN, M. KAMBER, AND J. PEI, DATA MINING: CONCEPTS AND TECHNIQUES, 3RD REVISED ED. BURLINGTON, MA: MORGAN KAUFMANN PUBLISHERS IN, 2011.

[2]L.A. Zadeh, "Some reflection on soft computing, granular computing and their roles in the conception, design and utilization of information/ intelligent system", Soft computing,
vol. 2, 1998.

[3]Chuan-Yu Chang, Da-FengZhuang and et.al, "A Fuzzy-Based Learning Vector Quantization Neural Network for Recurrent Nasal Papilloma Detection",IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, VOL. 54, NO. 12, DECEMBER 2007.

[4]EL-Sayed, A. EL-Dahshan, and et.al, "A Hybrid Technique
for Automatic MRI Brain Images Classification", Studia Univ.
BABES_BOLYAI, INFORMATICA, Volume LIV, Number 1, 2009.

[5] D.Jude Hemanth1, C.KeziSelva, and et.al, "Comparative Analysis of Neural Model and Fuzzy Model for MR Brain Tumor Image Segmentation", IEEE, 2009.

[6] Jang, Sun, Mizutani(1997) Neuro-Fuzzy and Soft Computing-Prentice Hall,pp 335-368, ISBN 0-13-261066-3.

[7] N.Benamrane, A. Aribi, L.Kraoula and et. al, "Fuzzy Neural Networks and Genetic Algorithms for Medical Images

[8] Interpretation".IEEE Proceedings of the Geometric Modeling and Imaging, 2006.