Journal of Soft Computing Paradigm (ISSN: 2582-2640)
www.irojournals.com/jscp/

Survey on Feature Detection of Alzheimer's disease in Early Stage by Particle Size Distribution in Feature Selection Optimization Procedure

B. Kedarnath

HOD, department of ECE, Guru Nanak Institutions Technical Campus, Hyderabad, India

E-mail: bkedarnath@gmail.com

Abstract

The rising incidence of Alzheimer's Disease (AD) and moderate impairments of scan results in the modern world have piqued the interest of scientists in the field of neuroimaging-based diagnostics. Neuroimaging allows for the quantification of pathological alterations in the brain that has been associated to AD. Through the use of categorization frameworks, which provide diagnostic and prognostic tools, these measurements have been quickly included into the signatures of AD in recent years. The purpose of this article is to summarise research on Alzheimer's disease that used optimization strategies for feature selection. To address the problem of excessive model complexity when using ML techniques, this work presents a novel approach to feature selection. Several stages of Alzheimer's disease and a state of altered brain function that is clinically similar to AD but less severe have been described. The effectiveness of an AD's classification in these approaches is evaluated using a wrapper-based feature selection mechanism. Then, a suggested Social Spider Metaheuristic algorithm has been employed to zero down on the most crucial characteristics for making a correct AD diagnosis.

Keywords: Alzheimer's disease, feature selection, optimization process, Mild Cognitive Impairment (MCI), Machine Learning (ML)

1. Introduction

Recently, it's the leading cause of dementia (60%-80% of all cases) in the world's elderly people. A person's intelligence and social abilities deteriorate gradually, and the disease eventually shuts down the body's systems one by one. The average age of the onset of

DOI: https://doi.org/10.36548/jscp.2022.3.005
Received: 28.07.2022, received in revised form: 22.08.2022, accepted: 06.09.2022, published: 16.09.2022

Alzheimer's disease is 60. The risk increases with age. Having an Alzheimer's disease patient in the family also poses a major threat. The gene expression data were utilised in a number of research. The number of features is often more important than the number of samples, making this data high-dimensional. As a result, dimensionality must be lowered to make better use of the data and avoid the risk of over-fitting. Exploring and analysing the data that allows to construct a basic model using dimensionality reduction, is the act of lowering the number of features under consideration by the combination, transformation, or selection of the numerous features [1- 4]. In feature reduction, generalization on dimensionality reduction may be made such that, new features are generated by linearly combining all the previous features, using just a subset of available features (feature selection).

1.1 Feature Selection

One of the most crucial stages of the pre-processing phase in data mining and machine learning is feature selection [5-8]. Feature selection is often used in situations when there are thousands of characteristics but only tens or hundreds of samples are required, such as in the analysis of DNA microarray data or Alzheimer's disease clinical data. The goals to achieve via the process of feature selection are, enhancing the predictors' ability to make accurate predictions while making them more efficient and less expensive, and offering insight into the mechanism that has resulted from the data being collected.

1.2 Detection at an Early Stage

One of the main causes of AD is the disconnection of neurons. The neuron link is gradually lost, leading to the AD symptoms. A neuron's primary job is to relay messages between different parts of the body. It is the synapses that play this part. Damage from Alzheimer's disease makes them ineffective, and eventually they pass away. Areas of the brain that experience widespread neuronal death often atrophy and decrease. Brain atrophy describes this deterioration of brain tissue. Significant and widespread shrinking of brain tissue occurs as a result [9, 10].

1.3 Alzheimer's Disease Origins

Brain alterations associated with AD are difficult to understand due to their complexity. Researchers study and learn more about its development every day. It's possible that brain damage starts ten or more years before symptoms like memory loss and difficulty in thinking become apparent. While it may seem like there are no symptoms at this point, it is

clear that harmful alterations are happening constantly in the brain. Amyloid plaques and tau tangles, and aberrant protein deposits, may be seen throughout the brain. This results in, the production of healthy neurons which are subsequently blocked by other neurons, resulting in the death of the neurons.

2. Literature Survey

Saputra et al. [11] predicted that by 2020, the situation will have improved. The Random Forest (RF) technique achieved the highest accuracy (91.15 percent) in studies that used ten-fold cross-validation is assessing a machine learning model to conduct extracted feature selection values. The kappa rate of the Particle Swarm Optimization-based RF approach was 0.884, and its accuracy value was 93.56 percent; it was also tested several times more often than the Decision Tree (DT) algorithm. This work is constrained by the difficulties of small sample sizes and poor precision. In order to enhance precision, a number of methods are employed to select characteristics and conduct local searches.

Eventually (around 2020), C. Park et al., proposed Deep Learning in the publication. A learning strategy was used here. Prediction of AD was achieved by combining large-scale Gene Expression (GE) data with DNA methylation information. It was challenging to model because of the need to integrate various omics datasets and cope with vast amounts of data from a limited sample size. A novel, straightforward method was developed to reduce the complexity of the by focusing on methylated locations. AUC ranged from 0.797 to 0.756, then to 0.773, and finally to 0.775. The paper's limitations are outlined. Maximal efficiency was in the processing speed [12].

A group of researchers led by K. Mohammed Niyas in 2020 proposed using a mix of choose characteristics for Alzheimer's disease diagnostics. The proposed method utilised SVM, KNearest Neighbor, etc. to classify Normal Controls, MCI, and obtained 90% and 91% balanced classification accuracy, respectively, as well as curve values of 0.97 and 0.98. This method improved both sensitivity and specificity [13].

N. Le et al. [14], used gene expression microarray data to build a machine learning model to make use of 35 expression properties. On average, the 35-feature model outperformed classifiers by 3.3 percentage points (AUC 98.3). However, the method used was inadequate for forecasting survival outcomes and even produced the reverse of the actual occurrence.

3. Method of Categorization

After cortical and subcortical MRI features have been retrieved, the combined components are normalised and put through the feature selection process. Once the chosen traits are obtained, they are utilised to distinguish the AD from other groups.

3.1 Symptoms and Indicators of Alzheimer's disease

Although AD symptoms seem to progress over a consistent timeline, they may not always manifest in the same way from one individual to the next. The typical stages are "early," "middle," and "late," corresponding to the onset, progression, and end of Mild Cognitive Impairment (MCI). Individuals may have a wide range of symptoms, some of which include memory loss, impaired cognitive abilities, poor eyesight, and difficulty concentrating or making judgments [15-18].

3.1.1 Early Onset Stage

In the latter stages of the illness, patients may have memory loss and other cognitive issues such as,

1. Difficulties with money management and bill paying.
2. Having a habit of misplacing or losing stuff.
3. Inquiring again and over again about changes in character or conduct.
4. Disoriented by wanderings and end up all over the place.

3.1.2 The Middle Ground

Reasoning, sensory processing, language, and conscious cognition are all impaired due to brain injury. Some of the signs of this destruction are:

1. Memory problems and brain fog have worsened.
2. Friends and family members are hard to identify.
3. The inability to pick up new skills or information.
4. Problem with multi-step processes.
5. Psychiatric problems include symptoms like hallucinations, paranoia, and delusions.
6. Abrupt shift in character

3.1.3 Severity Level

When a person reaches this point, they are completely reliant on the care of others. They may spend most of their time bedridden and unable to interact with the outside world until their bodies shut down entirely. These are some of the signs and symptoms:

1. Seizures
2. Slimming down
3. Illnesses of the skin
4. Improved restraint
5. Problems swallowing
6. Difficulty in expressing oneself
7. Making a low, grunting sound
8. Involuntary bowel and bladder movements

3.2 Fine-tuning and Classification

An error-minimizing classification function is used in this strategy, which is a suite of complex procedures from the characteristics learnt during training for a large domain [19]. Next, the classification algorithm is further changed to decrease the errors in the target dataset.

3.3 Feature Selection Method

The wrapper-based strategy is used for choosing the number of features. An excessive quantity of memory is needed if there are many pieces. So, feature selection goes through the subsets and picks the best one based on how effectively it fulfils its task. The filleting technique is used to pick out the most relevant characteristics that may be utilised to positively identify the AD. This strategy will remove all the useless characteristics, saving time spent recognising beneficial information.

Mostly, feature extraction and classification approaches, i.e., classification frameworks, construct prediction models [20, 21] to aid in decision-making and improve automation of medical choices. Classification frameworks with enhanced affectability and an individual's specificity may also be used to build imaging markers or recordings. The ability to care each unique patient in this way is crucial in today's age of personalised medicine.

Furthermore, it is conceivable to promote the concept of genetic or lifestyle risks by the use of sophisticated computational control.

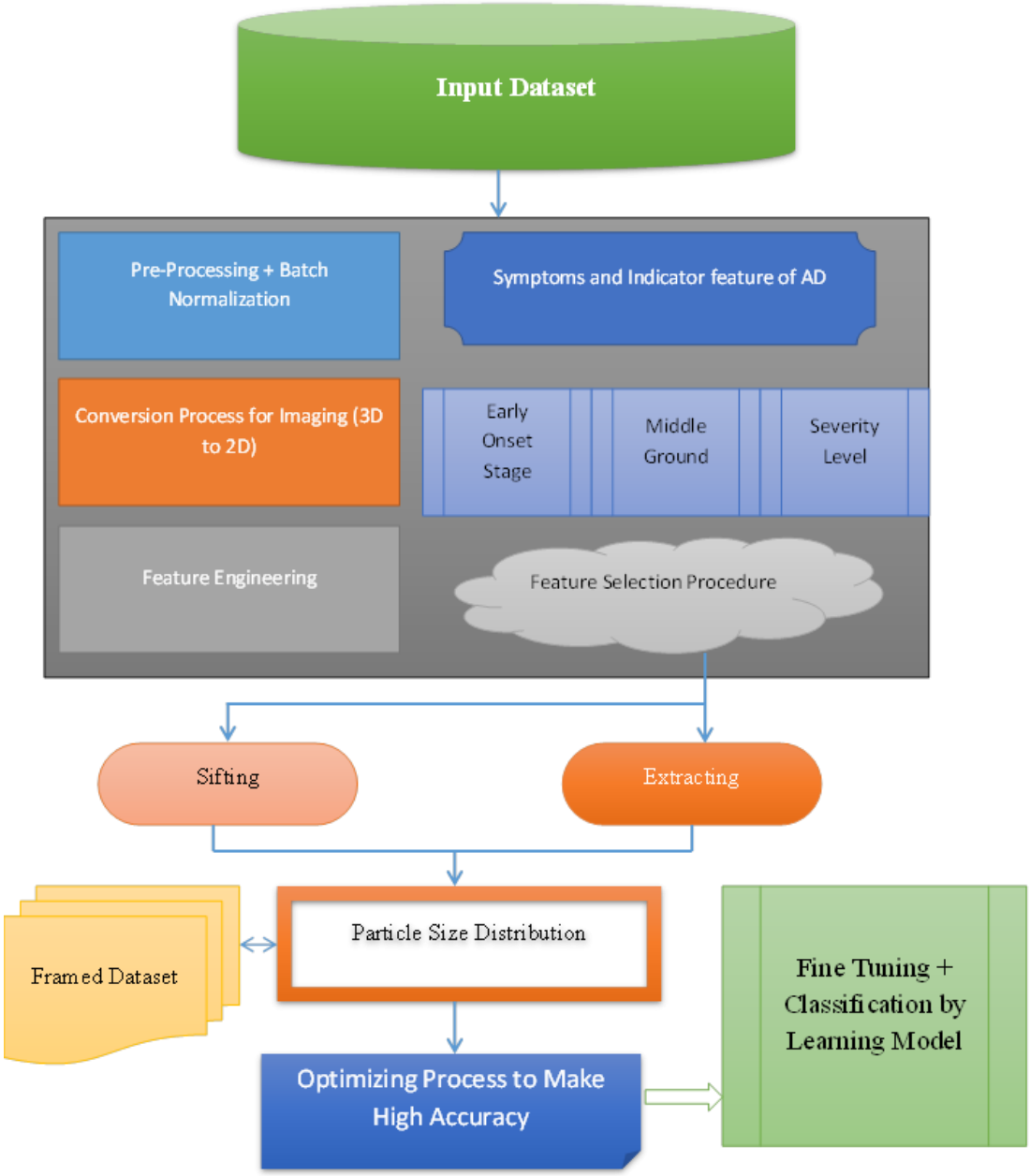

Figure 1. Basic functional diagram of AD detection by feature selection

There has been a lot of research over the last several decades towards neuroimaging-based classification of AD in an effort to facilitate earlier diagnosis. In light of the recent surge in research into AD, a summary of the many relevant activities that have been done, from feature extraction using different types of neuroimaging data to alternative categorization strategies is compiled. Moreover, concerns linked to the restricted sample size and data setup heterogeneity is explored by examining other research. The most important and fundamental test in AD evaluation is to determine whether a person has MCI and to

anticipate if a MCI person will grow into neurodegenerative disease. Figure 1 shows the basic functional diagram of AD detection by feature selection. Early-stage and late-stage MCI, for example, should be diagnosed with the same urgency. Due to:

1. Pre-processing photos that are of poor quality.
2. Unavailability of publicly accessible huge data samples for study.
3. Datasets with AD labels are scarce.
4. Insufficient information, particularly for localising ROI in the brain.

3.4 Particle Size Distribution by Sifting and Extracting

Automated diagnosis begins with image pre-processing; specifically feature extraction, which draws out distinct characteristics from pre-processed pictures belonging to several aberrant categories. It's accomplished by getting rid of superfluous details in the data. The classifier then makes use of these characteristics in order to assign classes to the data. It is assumed that the computational speed of a classifier will improve, and its accuracy will increase if it employs small and relevant components [22]. This is especially important in real-time systems where memory constraints are a major concern. After features have been extracted, they may be refined using a process called feature selection. The goal of feature extraction and selection is a compact representation of the pattern, which improves measurement efficiency and classification precision. The final classifier will be more efficient and less demanding on system resources. As a result of feature selection, it is possible to merge components from several data models. Feature selection may be challenging due to factors such as a limited sample size and uncertainty about the appropriate criteria function to use. Methods for selecting features include:

3.4.1 Direct Instruction

In the case of direct instruction, predefined examples are annotated to indicate which class they belong to. The aim is to extend the class labelling of the training items to new things such that they can be reliably assigned the classes.

3.4.2 Unsupervised Learning

Since the object is not explicitly known in unsupervised feature selection, this space has seen substantially less exploration. Determining which items should be combined is a common task for unsupervised learning. To put it in another way, the student determines the

curriculum. Classifiers take features as input and output instances of the class they best represent. Feature extraction is a technique for gathering useful information from photographs in order to minimise the volume of raw data. This is accomplished by analysing and quantifying various characteristics that serve to differentiate between distinct patterns. Different kinds of characteristics are included in various methods. A filter allows to pick just the features to be seen. Techniques based on the wavelet transform including the correlation method is suitable to detect the AD very efficient.

3.5 Optimization via Particle Size Distribution

Based on the observations of flocking birds, Particle Size Optimization (PSO) is a stochastic metaheuristic algorithm. It may be inferred from the flocking behaviour of birds that individuals are influenced both by the leader (the global optima) and by one's own performance (the local optima). Eberhart and Kennedy (1995) introduced the PSO, a population-based optimization method that has been effectively used in a variety of global search situations. Since it is simple to build, takes a decent amount of time in terms of computing, does a global search, and has few parameters, it is taken into account in many feature selection situations. Every particle in PSO's starting population has a location and a speed. At each iteration, a fitness function is used to assess the particles' quality. The current location will be carried by every particle in the search volume.

3.6 Benefits

The following feature option benefits are now at disposal:

1. All unnecessary information is eliminated by this technique.
2. There is a clear correlation between the time it takes to analyse data and the speed with which learning algorithms may be executed.
3. Improvements in data precision.
4. Boosting the reliability of the final model.
5. A lessening of the number of features in order to save energy for the next cycle of data collection or application.
6. Capability expansion for better prediction.
7. Analyze information to discover its meaning or to have a sense of its visual presentation.

4. Future Challenges

Feature selection is being used to take on the problem of effective classification of AD, with promising outcomes thus far. To aid them in their feature selection efforts, researchers are increasingly looking for hybrid feature selection methodologies. It can classify these methods as filtering, wrapping, or embedding techniques. Since filtering methods need so little processing power, they dominate the field. Methods of encapsulation and embedding have been studiously avoided. Because of these methods, the classification model for Alzheimer's disease is more precise and the selected genes are more stable. The first problem is an insufficient number of features or knowledge-based characteristics in the model, both of which can be corrected in further research. The second problem is managing a mountain of information. Training a big quantity of data locally, which would require more technical and human labour, is not as effective as using cloud computing for this situation. The present biggest problem in this area is the scarcity of accessible datasets. Some high-quality gene expression datasets have been compiled for Alzheimer's disease.

5. Conclusion

Manual methods for detecting Alzheimer's disease in brain MRI scans are time-consuming and expensive. As a result, their analysis and categorization are used for the purposes of feature extraction and medical diagnosis. This publication provides a thorough overview of the various techniques used for MR image categorization. The basic feature extraction and picture classification findings are encouraging after examining all the detection techniques. This means that Alzheimer's sufferers might be accurately identified by MRI analysis using traditional approaches. This study makes a substantial addition to the field of automatic classification by demonstrating the use of several automated classification strategies to brain MRI scans. Using a system like this, radiologists and researchers may more accurately categorise AD.

References

[1] Wang, S. et al., Dual-Tree Complex Wavelet Transform and Twin Support Vector Machine for Pathological Brain Detection. Appl.Sci. 6(6):169, 2016. https://doi.org/10.3390/app6060169.

[2] Ayadi, W., Elhamzi, W., Charfi, I., and Atri, M., A hybrid feature extraction approach for brain MRI classification based on Bag-ofwords. Biomedical Signal Processing and Control 48:144–152, 2019. https://doi.org/10.1016/j.bspc.2018.10.010.

[3] Gudigar, A., Raghavendra, U., San, T. R., Ciaccio, E. J., and Acharya, U. R., Application of multiresolution analysis for automated detection of brain abnormality using MR images: A comparative study. Futur. Gener. Comput. Syst. 90:359–367, 2019.

[4] Acharya, U. R., Fujita, H., Lih, O. S., Adam, M., Tan, J. H., and Chua, C. K., Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowl.-Based Syst. 132:62–71, 2017.

[5] Jha, D., and Kwon, G.-R., Contourlet-based feature extraction for computer-aided classification of Alzheimer's disease. Alzheimers and Dementia 14(7):1473, 2018.

[6] Wang, S.-H., Du, S., Zhang, Y., Phillips, P.,Wu, L.-N., Chen, X.-Q. et al., Alzheimer's disease detection by pseudo Zernike moment and linear regression classification. CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders) 16:11–15, 2017.

[7] Gorji, H. T., and Haddadnia, J., A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. Neuroscience 305:361–371, 2015.

[8] Wang, S., Zhang, Y., Liu, G., Phillips, P., and Yuan, T.-F., Detection of Alzheimer's disease by three-dimensional displacement field estimation in structural magnetic resonance imaging. J. Alzheimers Dis. 50:233–248, 2016.

[9] Ramesh Kumar Lama, JeonghwanGwak,Jeong-Seon Park and Sang-Woong Lee, 2017. Diagnosis of Alzheimer's Disease Based on Structural MRIImages Using a Regularized Extreme Learning Machine and PCA Features,Hindawi, Journal of Healthcare Engineering.

[10] Srinivasan, K., Garg, L., Datta, D., A. Alaboudi, A., Agarwal, N. R., and George Thomas, A. (2021). Performance Comparison of Deep CNN Models for Detecting Driver's Distraction. Comput. Mater. Continua 68 (3), 4109–4124. doi:10.32604/cmc.2021.016736

[11] R. A. Saputra et al., "Detecting Alzheimer's Disease by the Decision Tree Methods Based on Particle Swarm Optimization," J. Phys. Conf. Ser., vol. 1641, no. 1, 2020. https://doi.org/10.1088/1742-6596/1641/1/012025

[12] C. Park, J. Ha, and S. Park, "Prediction of Alzheimer's disease based on deep neural network by integrating gene expression and DNA methylation dataset," Expert Syst. Appl., vol. 140, p. 112873, 2020. https://doi.org/10.1016/j.eswa.2019.112873

[13] K. P. Muhammed Niyas and P. Thiyagarajan, "Feature selection using efficient fusion of Fisher Score and greedy searching for Alzheimer's classification," J. King Saud Univ.-Comput. Inf. Sci., no. xxxx, 2021.

[14] N. Q. K. Le, D. T. Do, T.-T.-D. Nguyen, N. T. K. Nguyen, T. N. K. Hung, and N. T. T. Trang, "Identification of gene expression signatures for psoriasis classification using machine learning techniques," Med. Omi., vol. 1, no. December 2020, p. 100001, 2021. https://doi.org/10.1016/j.meomic.2020.100001

[15] Srinivasan, K., Mahendran, N., Vincent, D. R., Chang, C.-Y., and Syed-Abdul, S. (2020). Realizing an Integrated Multistage Support Vector Machine Model for Augmented Recognition of Unipolar Depression. Electronics 9, 647. doi:10.3390/electronics9040647

[16] Sureshkumar, V., Chandrasekar, V., Venkatesan, R., and Prasad, R. K. (2020). Improved Performance Accuracy in Detecting Tumor in Liver Using Deep Learning Techniques. J. Ambient Intelligence Humanized Comput. 12, 1–8. doi:10.1007/s12652-020-02107-7

[17] Sanchez-Riera, J., Srinivasan, K., Hua, K.-L., Cheng, W.-H., Hossain, M. A., and Alhamid, M. F. (2018). Robust RGB-D Hand Tracking Using Deep Learning Priors. IEEE Trans. Circuits Syst. Video Technol. 28 (9), 2289–2301. doi:10.1109/TCSVT.2017.2718622

[18] Sarraf, S., and Tofighi, G. (2016). Classification of Alzheimer's Disease Using Fmri Data and Deep Learning Convolutional Neural Networks. arXiv preprint arXiv: 1603.08631.

[19] Sarraf, S., and Tofighi, G. (2016). "Deep Learning-Based Pipeline to Recognize Alzheimer's Disease Using fMRI Data," in 2016 future technologies conference (FTC) (San Francisco, CA, USA: IEEE), 816–820. doi:10.1109/ftc.2016.7821697

[20] D. Jha, S. Alam, J.Y. Pyun, K.H. Lee, and G.-R. Kwon, "Alzheimer's Disease Detection Using Extreme Learning Machine, Complex Dual Tree Wavelet Principal Coefficients and Linear Discriminant Dnalysis," Journal of Medical Imaging and Health Informatics, Vol. 8, No. 5, pp. 881-890, 2018.

[21] S.H. Nozadi, S. Kadoury, and The Alzheimer's Disease Neuroimaging Initiative, "Classification of Alzheimer's and MCI Patients from Semantically Parcelled PET

Images: a Comparison between AV45 and FDG-PET," International Journal of Biomedical Imaging, Vol. 2018, Article ID 12417430, 2018.

[22] M. Khajehnejad, F. Saatlou, and H. Mohammadzade, "Alzheimer's Disease Early Diagnosis Using Manifold-Based Semi- Supervised Learning," Brain Sciences, Vol. 7, No. 12, pp. 1-19, 2017.

Author's biography

B. Kedarnath is currently working a HOD in the Department of ECE, Guru Nanak Institutions Technical Campus, Hyderabad, India. His area of research includes wireless communications and image Processing.