JBE Journal of Biostatistics and Epidemiology
J Biostat Epidemiol. 2025;11(2): 166-176

ORIGINAL ARTICLE
DOI: *https://doi.org/10.18502/jbe.v11i2.20556*
*Received **February** 15, 2025; Accepted **April** 01, 2025;*
*Published online **June** 15, 2025.*

A Multifunctional Approach to Feature Extraction from fMRI Images in Alzheimer's Disease

Shahriar **Mohammadi**[1*], Soraya **Zarei**[1], Ali Mohammad **Mosadeghrad**[2]

[1]Information Technology Group, Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran.
[2]Department of Health Management and Economics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.

ABSTRACT

Introduction: The use of fMRI imaging in medical science has led to the diagnosis of diseases at the very first stages before the disease get advancedwhich plays a significant role in some diseases such as Alzheimer's. Extracting useful information from these images is the first step in the initial diagnosis of the disease that the accuracy in extracting as much of this information as possible contributes significantly to the initial diagnosis. Increases the speed of processing and estimation accuracy which was done in the present study using a multi-purpose method. While in recent studies, simpler methods with a limited number of features were used.

Methods: The information of 140 patients with Alzheimer's disease was obtained, and the stable multipurpose feature extraction method was used to extract the information. In this way, two-level wavelet, modeling of wavelet coefficients, normalization method and feature selection are applied.

Results: The results obtained from the examination of 285 features in five categories showed that some of the information contained in the features overlapped and lacked useful information. In addition, dimensionality and noise reduction using the PCA algorithm showed that about 41% of the relevant features are outliers or missing information.

Conclusion: In general, increasing the speed of processing and estimation accuracy which was done in the present study using a multi-purpose method. While in recent studies, simpler methods with a limited number of features were used.

Key words: Medical imaging; Brain diseases; fMRI images; Feature extraction

*Corresponding Author:
mohammadi@kntu.ac.ir

INTRODUCTION

One of the leading dementias in the globe that causes death is Alzheimer's disease (AD).[1] The most prevalent signs of AD include significant memory loss, cognitive decline, changes in mood or personality, etc.[2] The limbic system in the human brain, which includes the hippocampus, amygdala, and other structures, is responsible for many of the symptoms of AD. As a result, these structures suffer the most.[3]

The early detection, diagnosis, and treatment of diseases have all benefited from the widespread use of magnetic resonance imaging (MRI), a non-invasive medical imaging method.[4] MRI in studying the human brain allows researchers to learn not only about the structural makeup of the brain but also about its function and metabolic rate.[5] Due to their excellent spatial resolution, structural magnetic resonance imaging (sMRI) and functional magnetic resonance imaging (fMRI) have achieved significant advancements in the investigation of human brain structure and function, respectively.[6] A brain region's indirect impacts on neuronal activity are measured by fMRI because of variations in the magnetic susceptibility of blood with and without oxygen. The blood oxygenation level-dependent (BOLD) signal is the name given to this kind of signal. When compared to the diamagnetic oxyhemoglobin, the paramagnetic deoxyhemoglobin creates a little bit greater local inhomogeneity dis the local magnetic field, which leads to a little bit quicker dephasing of the spins (of hydrogen ions in the water molecules).[7-8]

The initially, fMRI data are processed by advanced statistical analysis to obtain the relevant information. Any amount of noise or artifact can dramatically alter the outcome of the analysis. When it comes to clinical analysis, this becomes a significant problem. A little artifact may cause erroneous interpretation and have an impact on the diagnosis. The fMRI data structure is also highly intricate. Medical practitioners can easily see and evaluate a conventional x-ray or anatomical MRI image. Without efficient processing of the statistically evaluated data, it is a challenge to fully visualize fMRI.[9-10]

In the realm of medical imaging, machine learning (ML), as a pattern recognition tool, is used. ML often begins by choosing attributes that are thought to be crucial for providing diagnoses or predictions. The ML system then determines which combination of these chosen features is most effective for categorizing or calculating certain metrics for the provided image.[11]

Numerous studies have been done in the area of picture feature extraction using various pieces of data to identify various illnesses. Particularly, those that affect the neurological system.[12-14] The accuracy of the estimation performed has been impacted by the use of various features, dimensions of feature extraction, the existence or absence of correlation between various characteristics, as well as the removal of noise elements in the features.[8-15] Predicting Alzheimer's disease requires effective feature extraction from fMRI images. On the other hand, the construction of numerous models is facilitated by advancements in the fields of artificial intelligence and nervous system imaging. Therefore, the objective of the current study was to increase the extraction's accuracy by applying efficient models in conjunction with the DWT feature extraction method from fMRI of persons with Alzheimer's disease, modeling of 2D DWT coefficients, normalization and extraction of robust multifunctional features.

METHODS

Data and processing

The ADNI (Alzheimer's disease neuroimaging initiative) online database was used to obtain all of the fMRI volumes used in this study. 675 patients' resting-state fMRI volumes were collected in total.[16] A 3T Philips MRI Medical System for to obtain each resting-state volume was used. With an echo time and repetition time of 30 ms and 3000 ms, respectively, the flip angle was 80 degrees. A time series of 140-time points recorded for each voxel, or as well as around 140 complete brain volumes.

Before feeding the data into the algorithm, all of the fMRI data by the normal preprocessing pipeline were processed. The SPM12 toolbox in MATLAB 2019a was applied for all of the preprocessing. Motion correction, intensity normalization, coregistration, removal of the global mean and temporal signal drift, and thresholding and masking based on intensity were all included in the preprocessing. Motion correction was the initial operation in the preprocessing. All of the fMRI volumes were adjusted for motion, using the temporal mean volume as the reference.

Then, the intensity of the fMRI volume normalized. Then, to make sure that all of the subject's data was in the same spatial size and dimensions, all of the fMRI volumes cogistered to the Montreal Neurological Institute (MNI) space. The 64 x 78 x 64 voxel size of the MNI atlas used for coregistration corresponded to a 64 x 64 x 64 voxel size. The background outside the region of interest was removed using a binary mask when all the volumes were in the same MNI space. To ensure uniformity for further analysis, the same binary mask in the MNI space applied to all volumes. Finally, a PCA-based method was used to eliminate temporal signal drift and global signal variations.

A reliable algorithm is described for magnetic resonance imaging (fMRI) to identify the disease type. As the input image, the two-dimensional discrete wavelet transform D DWT (2) is first calculated. The multifunctional feature extraction and normalization process is then established, and feature selection is carried out. A Spin-Echo sequence with weight T1 (axial), TR = 400 milliseconds, TE = 9 mm, flip angle = 90 °, and voxel size = 1 1 6 mm was utilized to produce high-resolution structural images of the brain.

Using the Echo / Echo Flat Imaging Protocol (EPI) (axial, TE = 60.3 ms, TR = 3125 ms, flip angle = 90°, field of view = 22 square centimeters, number of pieces = 15, piece thickness = 6 mm, distance = 0 mm, bandwidth = 15.62 kHz, voxel size = 4 4 6 mm), FMRI data with the same dimensions and orientation of structural images were obtained.

Due to hardware constraints, the EPI procedure was repeated in order to gather a sufficient number of image volumes for data analysis (up to 32-time points and 15 brain slices per MR scanner). To create 64 functional image volumes, 64 EPI sequences of 100 seconds each were performed. The block diagram of the proposed method as Figure 1.

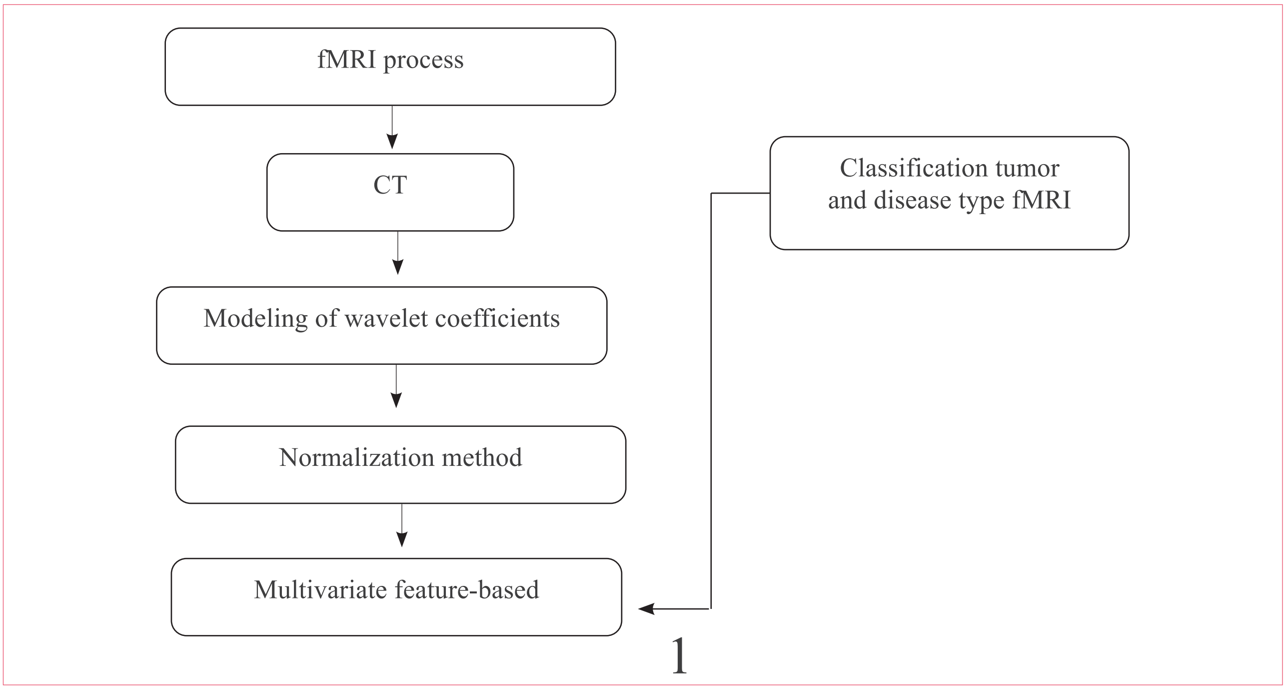

Figure 1. The block diagram of the proposed method

Machine algorithm

Several related tasks simultaneously plotted as common features and out-of-task tasks identified with the help of a robust multi-purpose feature-learning algorithm, and it is given learning tasks related to educational data $\{(X_1, y_1),\cdot\cdot\cdot,(X_m, y_m)\}$,,

Where:

$X_i \in R^{d\times n_i}$ is the i-th data matrix Is. Work with each column as an example.

$y_i \in R^{n_i}$ is the i-th answer

y_i has continuous values for regression and discrete values for classification

d is the dimensions of the data

n_i is the number of i-th examples of work.

The data is normalized so that the input (j, k) - amine X_i denoted by $x_{jk}^{(i)}$, is satisfied:

$$\sum_{k=1}^{ni}(x_{jk}^{(i)})^2 = 1, j \in \mathbb{N}_d$$

The linear learning function is defined as follows:

$$y_{ii} \approx f_i(x_j^{(i)}) = (x_j^{(i)})^T w_i, i \in \mathbb{N}_m, j \in \mathbb{N}_{n_i}$$

For each task and decompose the weight matrix W =[$w_1, \cdot\cdot\cdot, w_m$] ϵ $R^{d\times m}$ to the sum of the two components P and Q. We use different regularization terms in P and Q. Relationships between tasks formally, the proposed model formulated as follows:

$$\min_{W,P,Q} \sum_{i=1}^{m} \frac{1}{mn_i} \| X_i^T w_i - y_i \|^2 + \lambda_1 \| p \|_{1,2} + \lambda_2 \| Q^T \|_{1,2},$$
$$s.t. W = P + Q$$

Where λ1 and λ2 are non-negative parameters to control these two expressions, where P identifies common features among tasks and the second expression in Q finds outliers.

The minimum length description algorithm, which takes into account a kind of probability function and minimizes it to estimate the number of independent sources, was used to estimate the number of independent components. To achieve a mean and zero variance, all retrieved components transformed into Z space. The suitable threshold for each independent component was then determined using a Gaussian mixture in the following step. This threshold was used for the probability component maps to extract the final ICs. Following function analysis, 285 features were identified based on 140-time stages of the brain and other sections, and all functional images were recorded in their anatomical T1 images.

The first expression limits the optimal answer row P with all zero or non-zero elements by ordering on the P-group groupings. The presumption that all tasks have the same properties is not always true in real-world applications since garbage tasks occur. We introduce the second regularization expression based on the same group rope penalty but applied to the Q-column groups in order to identify these outliers. Similar to this, the columns of the ideal response Q are all zero or non-zero, with the non-zero columns representing the tossing tasks. i-th work is recognized as a skewed job because, intuitively, if i-th column Q is non-zero, the i-th column is also non-zero. Meanwhile, for the remaining tasks connected to column zero Q, they share a common set of qualities drawn from the non-zero rows P. The training ratio for the fMRI data was set at 20% and 30%, respectively.

RESULTS

The fMRI data of 675 patients analyzed, in which 285 features evaluated separately in five, categories (mean cortical thickness, the standard deviation of cortical thickness, cortical packing volume, cortical packing volume, and cortical surface area. The results related to the learning algorithm with multi-purpose features are shown in Figure 2. The use of the used algorithm made it possible to check the common features and in this way, some outlier features were identified. In the used method, the actual weight given for each feature is used in predicting Alzheimer's disease, which greatly improves the estimate's accuracy. In Figure 3, an image of the decomposition of the weight matrix is presented, where the squares with a white background show zero entries. There were five features; the fourth feature was identified as an outlier, Figure 3. The results of checking all the features are shown in Figure 3.

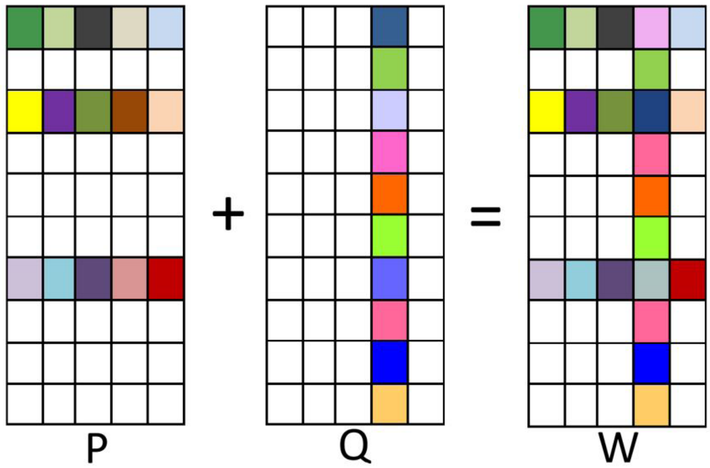

Figure 2. An image of the weight matrix analysis by the proposed multitasking method

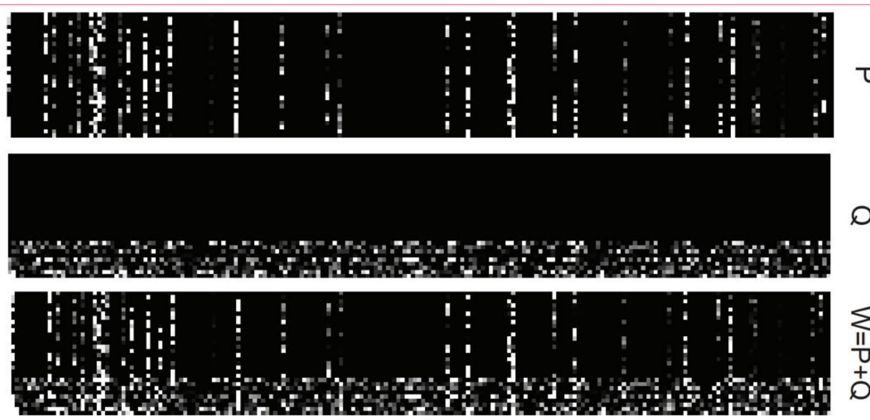

Figure 3. The values of P, Q and W = P + Q from the method presented on the fMRI data. The black dots correspond to the zero input. Note that the shapes are rotated 90 degrees clockwise.

We perform scalability studies on two algorithms that can identify outliers: the proposed method and RMTL. When the dimension and number of tasks increase. First we prove m = 20 and then let d increase as i 50i i = 1, 5. Then we perform the proposed method and RMTL on the fMRI data generated according to the same procedure. The computational time (CPU time) in front of the dimension diagram shown in the bottom left of Figure 4.

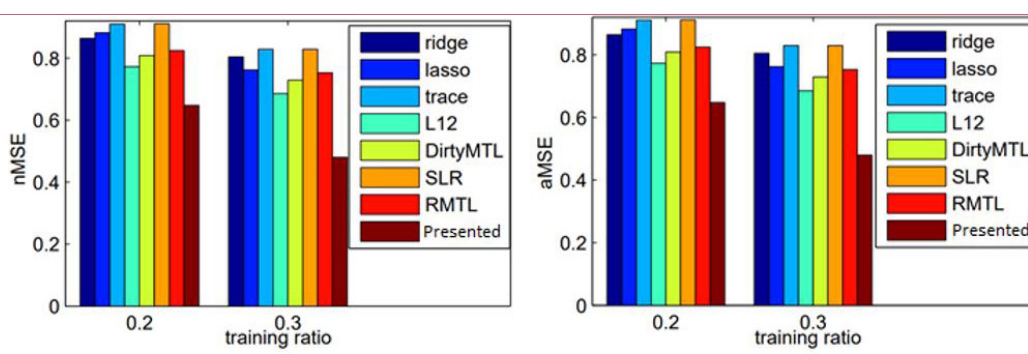

Figure 4. Mean test error (nMSE and aMSE) versus training ratio.

By extracting the features and passing it through the one-dimensional violet and removing the noise, it entered the dimensionality reduction stage with the PCA method, the results of which shown for two groups of people with Alzheimer's disease and healthy people in Figure 5. As can be seen in Figure 6, by reducing the dimensions of the features using the PCA method, from the total of 285 features used, 167 features have covered about 98% of the variance of the features, that is, about 41.4% of the total features have been reduced and They lack valuable information in predicting Alzheimer's disease.

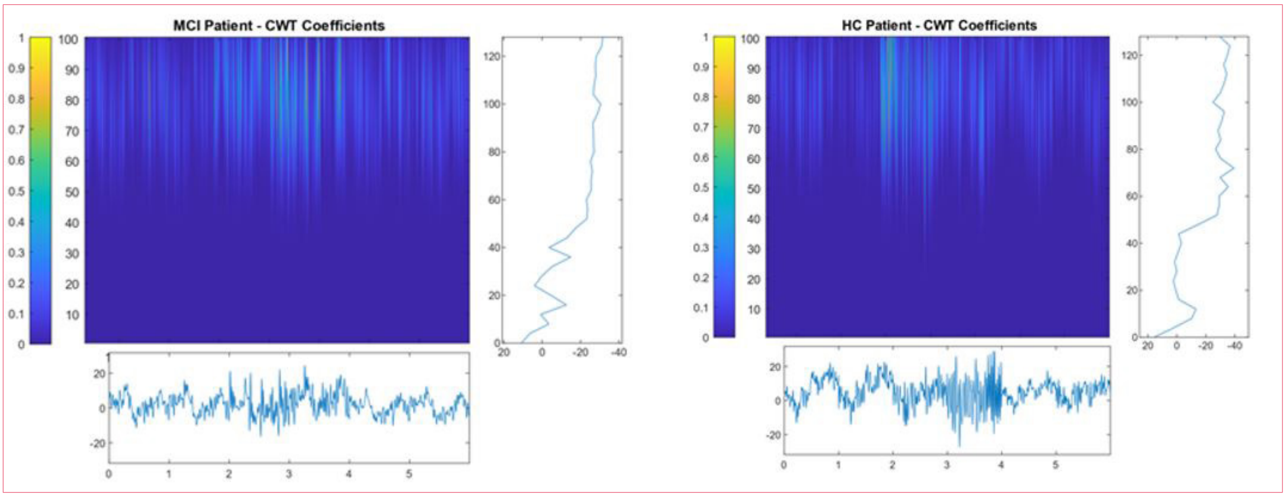

Figure 5. Wavelet results for fMRI sample for patient (left), for healthy person (rigth).

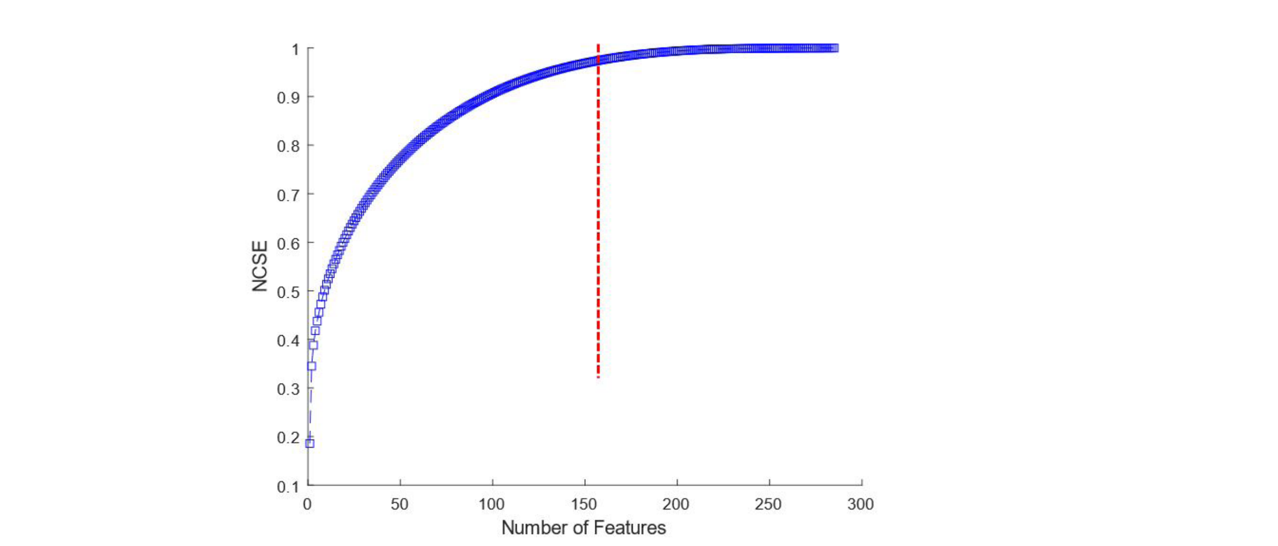

Figure 6. Cumulative sum of eigenvalues

DISCUSSION

Our results showed some features corelated and reduction them improve accuracy.various studies have used feature extraction to predict Alzheimer's disease. Most of these studies addressed this issue by using limited features and some with more features.[17-21] The results of the present research showed that some features are devoid of useful information in forecasting, and some of the information is outliers, and if this information is used, the possibility of reducing the accuracy and accuracy of forecasting increases.

Acharya et al (2019) developed a Computer-Aided-Brain-Diagnosis system by using MRI picture. For those purpose used several of quantitative techniques: filtering, feature extraction, Student's t-test based feature selection, and k-Nearest Neighbor (KNN) based classification. Their results showed the Shearlet Transform (ST) feature extraction technique improved Alzheimer's diagnosis. The proposed ST + KNN technique provided accuracy of 94.54%, precision of 88.33%, sensitivity of 96.30% and specificity of 93.64%.[22]

Different strategies were employed to utilize the complementary information of features extracted from structural MRI, rs-fMRI, and diffusion tensor imaging for Alzheimer disease classification. Dai et al (2012) used regional gray matter volumetric measures, and functional measures (amplitude of low-frequency fluctuations, regional homogeneit , and regional functional connectivity strength) as features.[23] They trained separate maximum uncertainty linear discriminant analysis classifiers on the structural and functional measures, and combined the output of the classifiers via weighted voting. Recently, Dyrba et al (2015) used regional gray matter volumetric measures, average tract intensity for fractional anisotropy, mean diffusivity, and mode of anisotropy, and network measures of weighted local clustering coefficient and the shortest weighted path-length calculated from rs-fMRI as features.[24] They adopted multi-kernel support vector machines for Alzheimer disease classification.

Silva et al (2019) evaluted the convolutional neural network architecture in three convolutional layers to extract the best features of the selected region. they put the selected attributes in a vector for learning and detection of patterns by another technique of computational intelligence. Their study were partitioned with the 10-folds cross-validation method and trained with the Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN) algorithms with different parameters for evaluation. The results of accuracy are 0.8832, 0.9607 and 0.8745, for the algorithms mentioned above, respectively.[25]

CONCLUSION

In general, the results of the present study showed that the use of robust multitasking feature learning algorithm and the simultaneous examination of common features play an important role in identifying noise factors and outliers. The present analysis showed that the reduction of feature dimensions and its simultaneous examination in fMRI images identifies and removes part of the data lacking useful information. About 41% of the features of the images in the diagnosis of Alzheimer's disease are among the non-useful and extraneous information, which leads to its identification in a precise evaluation with high accuracy and reduces the calculation time. In addition, by using the presented, better extraction of features from the data source can be optimized.

Conflicts of interest

The authors declare no conflicts of interest for this paper.

Declaration of funding

None

REFERENCES

1. Association A. Alzheimer’s disease facts and figures. Alzheimer’s Dementia 2018; 14 (3): 367–429.

2. Korolev, I.O., 2014. Alzheimer’s disease: a clinical and basic science review. Med. Student Res. J. 4 (1), 24–33.

3. Moon SW, Lee B, Choi YC. Changes in the hippocampal volume and shape in early-onset mild cognitive impairment. Psychiatry Invest 2018; 15 (5): 531–537.

4. Brody H. Medical imaging. Nat. Cell Biol. 2013; 502, S81.

5. Yousaf T; Dervenoulas G; Politis M. Chapter Two-Advances in MRI Methodology. In International Review of Neurobiology; Politis, M., Ed.; Academic Press: Cambridge, MA, USA 2018; 141: 31–76.

6. Akhavan Aghdam M; Sharifi A; Pedram MM. Combination of rs-fMRI and sMRI Data to Discriminate Autism Spectrum Disorders in Young Children Using Deep Belief Network. J. Digit Imaging 2018; 31: 895–903.

7. Sarraf S, Tofighi G, Alzheimer's Disease Neuroimaging Initiative. DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI. BioRxiv. 2016; 1:070441.

8. Parmar HS, Nutter B, Long R, Antani S, Mitra S. Deep learning of volumetric 3D CNN for fMRI in Alzheimer's disease classification. InMedical Imaging. Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics.2020; 11317: 113170C.

9. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. J. Med. Imaging Radiat. Sci. 2019; 50: 477–487.

10. Duc NT, Ryu S, Qureshi MN, Choi M, Lee KH, Lee B. 3D-deep learning based automatic diagnosis

of Alzheimer's disease with joint MMSE prediction using resting-state fMRI. Neuroinformatics. 2020;18(1):71-86.

11. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. Radiographics 2017; 37: 505–515.

12. Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D. Early diagnosis of Alzheimer's disease with deep learning. In2014 IEEE 11th international symposium on biomedical imaging (ISBI) 2014; 29: 1015-1018.

13. Wang SH, Phillips P, Sui Y, Liu B, Yang M, Cheng H. Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. Journal of medical systems. 2018; 42(5): 1-1.

14. Ramzan F, Khan MU, Rehmat A, Iqbal S, Saba T, Rehman A, Mehmood Z. A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. Journal of medical systems. 2020; 44(2):1-6.

15. Suk HI, Lee SW, Shen D, Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage. 2014; 101:569-82.

16. ADNI. Alzheimer's disease Neuroimaging Initiative: ADNI. http://adni.loni.usc. edu/data-samples/access-data, accessed: 2020-07-13.

17. Kam TE, Zhang H, Shen D. A Novel Deep Learning Framework on Brain Functional Networks for Early MCI Diagnosis. In Medical Image Computing and Computer Assisted Intervention—MICCAI 2018; Springer Science and Business Media LLC: Cham, Switzerland, 2018; 11072: 293–301.

18. Zheng Y, Guo H, Zhang L, Wu J, Li Q, Lv F. Machine Learning-Based Framework for Differential Diagnosis Between Vascular Dementia and Alzheimer's Disease Using Structural MRI Features. Front. Neurol. 2019; 10.

19. Khanna A, Tanwar S, Rodrigues JJ, Roy NR. Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. Comput. Electr. Eng. 2019; 77: 230–243.

20. Amini M, Pedram M, Moradi A, Ouchani M. Diagnosis of Alzheimer's Disease Severity with fMRI Images Using Robust Multitask Feature Extraction Method and Convolutional Neural Network (CNN). Comput. Math. Methods Med 202;1: 1–15.

21. Al-Khuzaie FE, Bayat O, Duru AD. Diagnosis of Alzheimer's disease Using 2D MRI Slices by Convolutional Neural Network. Applied Bionics and Biomechanics. 2021; 2:2021.

22. Acharya UR, Fernandes SL, WeiKoh JE. Automated Detection of Alzheimer's Disease Using Brain MRI Images– A Study with Various Feature Extraction Techniques. J Med Syst 2019; 43: 30.

23. Dyrba M, Grothe M, Kirste T, Teipel SJ. Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. Hum Brain Mapp 2015; 36: 2118-2131.

24. Dai Z, Yan C, Wang Z, Wang J, Xia M, Li K, He Y. Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). Neuroimage 2012; 59: 2187-2195.

25. Silva IRR, Silva GSL, de Souza RG, dos Santos WP, Fagundes RA. "Model Based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease," 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary 2019; 1-7