Celal Bayar University Journal of Science

Binary Classification of Alzheimer's Disease Using Siamese Neural Network for Early Stage Diagnosis

Ruken Tekin[1]* , Tuğba Özge Onur[1]

[1] Dept. of Electrical-Electronics, Faculty of Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Türkiye
* r.tekin@fbe.karaelmas.edu.tr
* Orcid No: 0000-0003-4732-7580

Received: 7 August 2024
Accepted: 9 February 2025
DOI: 10.18466/cbayarfbe.1529546

Abstract

Alzheimer's Disease (AD) is a cognitive disease. In individuals with disease, increased brain cell loss is observed over time. This situation leads to deficiencies in memory and thinking ability over time. As a result, significant impairments occur in individuals ability to perform primary function. According to research results, the rate of thos disease doubles every five years among people aged between 65 and 85. The causes of AD are unknown and nowadays not definite cure. Early diagnosis of the disease in clinical cure as it has the potential to slow or stop progression. This study aimed to make a prediction based on Magnetic Resonance (MR) images. Images in the standard Alzheimer dataset obtained from the open access database Kaagle were enhanced by applying Gaussian and Median filters. Siamese Neural Network (SNN) categorizes disease stages by learning the similarity between these images. Two categories of images were used from the dataset: Very Mild Dementia (VMD) and Non-Dementia (ND). According yok this proposed study, the training accuracy was %99.62 and the validation accuracy %97.67.

Keywords: Siamese Neural Network, Machine Learning, Alzheimer Disease, Magnetic Resonance Imaging

1. Introduction

Alzheimer's disease is a neurodegenerative cognitive disorder [1]. Among people with AD, the continuous death of brain cells causes damage to memory and thinking functions, leaving the person unable to carry out even daily activities [2]. AD has a prevalence of 3-11% in people over the age of 65 and 20-47% over the age of 85. Studies in different regions of the world have shown that the incidence of people with AD doubles every five years [3].

Alzheimer's is generally divided four parts. These are Mild Cognitive Impairment (MCI), Mild Dementia, Moderate Dementia and Severe Dementia. MCI, as in this stage many people may experience memory loss due to age. This does not mean that every older person has MCI, but it can lead to dementia in others. Mild dementia (MID) is a condition seen in people who have cognitive impairments during the day, sometimes affecting their daily lives. In this phase of the disease, loss of memory, people' s personal behaivor status disorder manifests itself as the inability continue daily life or difficulty in doing so. In moderate dementia, it becomes noticeably more complex for people to carry out daily activities. The symptoms are mild but equivalent to high-level dementia. Individuals may become suspicious or angry for irrational reasons. Sleep problems are also likely to occur. In the severe phase of dementia, symptoms can be worse. Individuals may lose their ability to communicate and need care. People in this stage may have loss of control on bladder and unable to perform simple activities such as holding their own head in a normal posture or sitting in a chair [4].

Clinically, the cause of AD remains unclear and there is currently no treatment for the disease [5]. Diagnosing AD at an early state may provide an opportunity to slow or halt progression. Early prediction and diagnosis of AD is therefore significant in clinical treatment [6]. Research has revealed that there are different biomarkers in patients' brains that help early diagnosis of AD [7]. These changes in the brain are used to diagnose AD in clinical

Celal Bayar University Journal of Science
Volume 21, Issue 2, 2025, p 152-158
Doi: 10.18466/cbayarfbe.1529546

settings by magnetic resonance imaging (MRI) and positron emission tomography (PET)[8].

With MR imaging, structural and functional information of the brain can be obtained by utilising the tissue contrast of the images. PET imaging can provide metabolic and molecular information of the brain [9]. Additionally, with these methods, AD is diagnosed qualitatively, but also the sensitivity of the threshold can be measured, which can be identified at different stages of the disease [10].

Machine learning and deep learning methods serve as robust instruments for creating predictive models based on MRI [11]. There are similar works in the literature for the predictive diagnosis and diagnostic problem. Lu and others [12] suggested new multimodal deep neural network. The model achieved 82.4% accuracy in predicting MCI and detected that these patients developed Alzheimer's in the following three years. Gupta et al [13] suggested a diagnostic model for AD classification based on the Alzheimer's Disease Neuroimaging Initiative (ADNI) and National Questionnaire Scale dataset. By combining features from various parts of the brain in MRI data, this method classified people with AD with a success rate of 96.42% compared to Healthy Control (HC). Ahmed and others [14] suggested an ensemble Convolutional Neural Network (CNN) model. The study was designed to automate AD classification using sMRI. It adopted a patch-based approach specifically targeting the left and right hippocampus areas. In experiments on ADNI and the National Research Center for Dementia (NRCD) databases, accuracy rates of 85.55% and 90.05% were obtained, respectively. Nawaz et al [15] proposed the use of a trained AlexNet model for Alzheimer's detection. The model leveraging deep features attained a classification with a success rate of 99.21%. Shi and others [16] suggested deep polynomial network. This study achieved 55.34% success for AD. Liu and others [17] suggested deep SNNs to detect brain asymmetries related to AD and MCI. The dataset obtained from the ADNI database was used. In this study, the accuracy rate for MCI and AD classification is 92.72%. %. Xiaowang Bi and others [18] proposed deep metric learning and CNN for Alzheimer's disease-healthy control (AD-HC), healty control- Mild Cognitive Impairment (HC-MCI) binary classification. They used MRI in the ADNI database and obtained 65% accuracy for MCI-HC classification and 83% accuracy for AD-HC classification.

This paper proposes SNNs, a machine learning method for identification of early stage AD utilizing MR images. This study, a binary classification was made between people with very mild dementia, an early stage of AD, and people without dementia. Segmented 2-dimensional (2D) MRI were used and Median and Gaussian filters were applied. Multilayer CNN was built to augment the SNN. The results were evaluated with Accuracy (Acc), Sensitivity (Sens), and Specificity (Spec) success evaluation metrics and the results are explained in Section 4.

2. Materials and Methods

This part explains the material and methodology used to process and classify MRI datasets. Figure 1 represent processing steps of the proposed model. In this model, the dataset is acquired and preprocessed. Gaussian and Median filters are used in the preprocessing step. Then the dataset is trained and classified with SNNs.

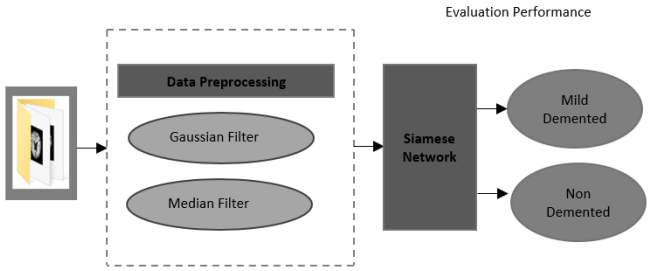

Figure 1. Stages of the study.

2.1. MRI Alzheimer's Dataset

The AD dataset was acquired by Kaggle. The dataset contains a total of 6400 MR images belonging to four classes: MID, Moderate Dementia (MOD), ND and VMD. The original size of the images in the dataset is 176 × 208.

In this study, 700 VMD samples and 700 ND samples were randomly selected and used. Figure 2 below shows some images resized to 176 x 176.

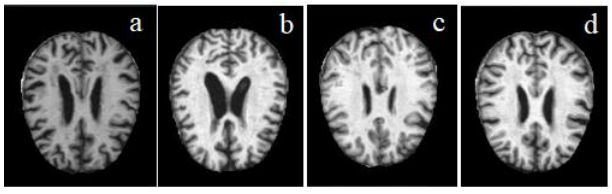

Figure 2. (a) Mild Demented, (b) Moderate Demented, (c) Non Demented, (d) Very Mild Demented

2.2. Data Preprocessing

In this study, Gaussian and Median filters were used to enhance MR images.

2.2.1. Gaussian Filter

Gaussian filter was applied to the images. This linear filter assigns weighted values to each element, determined by the Gaussian function's shape. The filter is usually very effective in eliminating leakage noise [19]. Using Equation (1), the value of each member of the

Gaussian filter can be calculated or determined. In this equation, σ represents the standard deviation of the Gaussian Kernel and c represents the normalization constant.

$$h(x,y) = \frac{1}{c} e^{\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

2.2.1. Median Filter

Median filtering is an image processing method used to enhance images. This nonlinear method preserves helpful information in the image [20]. The preservation of helpful information is achieved by converting the original pixel gray value of an image into the average gray value of the pixels in a given neighbourhood. The numerical formula for the median filtered K(u,v) image for the J(x,y) image is shown in Equation (2),

$$J(x,y) = \underset{(u,v\in R_{xy})}{median}\{K(u,v)\} \quad (2)$$

2.3. Siamese Neural Network

Siamese Neural Network was proposed to solution signature authentication challenge [21]. The objective is verifying that the signatures belong to the person concerned. SNN is trained based on similarity discrimination between two input images. Figure 3 expresses the way this model works. The similarity of the images is determined by associating input images based on weight parameters shared across the neural networks. In cases where the weight parameters extracted on two images give approximately analogous results, the images are decided to be identical.

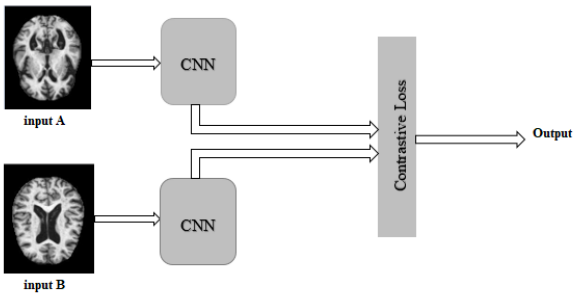

Figure 3. Working principle of SNN.

In the SNN model, a threshold value is utilised for this purpose and the gap between the output produced by the two neural networks is calculated. If this diff is less than the threshold, the output indicates that the images are analogous or identical; otherwise, the model produces an output indicating that the results are different [22].

Each pixel in the image has a numerical value. The distance between the pixels is usually calculated using the Euclidean distance given in Equation (3). Here, the variable U symbolises the interval, and the variables x and y symbolise the two vectors generated within the neural networks.

$$U = \sqrt[2]{\sum_{i=1}^{n}(x^i - y^i)^2} \quad (3)$$

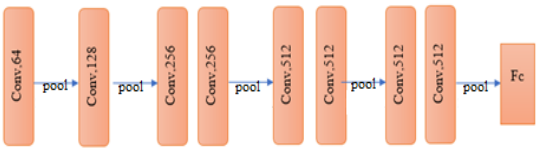

Figure 4. CNN architecture.

2.3.1. Contrastive Loss Function

The comparative loss function computes the loss for every pair of placements, aiming for reduced distances between similar pairs and increased distances between dissimilar pairs. The comparative loss is calculated by the formula given by Equation (4),

$$L = (1-y) * D^2 + y * \max(0, m-D)^2 \quad (4)$$

where:
L: Comparative loss for the pair.
D: Distance or difference between embeddings.
y: Label indicating whether the pair is similar or dissimilar (0 for similar, 1 for dissimilar).
m: Margin parameter defining the dissimilarity threshold.

2.4. Evaluation Index

The study's performance was assessed using the Acc, Sens, and Spec metrics.

In the equations below, letters are given according to the situations of the people. A refers to patients with a correct diagnosis. B, C and D refer to healthy individuals who have been correctly diagnosed; healthy individuals who were misdiagnosed; sick individuals who have been misdiagnosed, respectively.

Acc is calculated by Equation (5) and expresses the ratio of the number of people that can be properly predicted and categorized.

$$Acc = \frac{A+B}{A+B+C+D} \quad (5)$$

The term Sens denotes the proportion of correctly identified patients in the overall number of patients and is calculated by Equation (6).

$$Sens = \frac{A}{A+D} \quad (6)$$

The term Spec denotes the proportion of correctly identified healthy individuals relative to the total number of individuals without any health issues, computed as outlined in Equation (7).

$$Spec = \frac{B}{B+C} \quad (7)$$

Precision, recall and F1-score are the criteria that analyse the performance of the classification models. These measures are formulated with reference to the confusion matrix in Table 1.

Table 1. Cofusion matrix

	PREDICTED	
ACTUAL	True Posivites (TP)	False Negatives (FN)
	False Posivites (FP)	True Negatives (TN)

The areas where the model makes correct predictions are expressed as True Positive and True Negative, while the areas where the model makes incorrect predictions are False Positive and False Negative.

Precision indicates the accuracy of Positive class predictions, calculated as the ratio of True Positives to all Positive predictions.

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

Recall (Rec) is a metric that shows how much of what is expected to be positively predicted is positively predicted.

$$Rec = \frac{TP}{TP+FN} \quad (9)$$

The F1 Score represents the harmonic mean of Precision and Recall.

$$F_1 = 2 * \frac{precision*recall}{precision+recall} \quad (10)$$

3. Training and Evaluation

In this study, Very Mild Demented and Non Demented MR images were taken as two different inputs and used for comparison. The images were preprocessed before being trained with the neural network. The images were resized as 176 x 176. Then Gaussian and Median filters were applied to improve the images.

To obtain a successful model in machine learning, the parameters must be estimated correctly. In order to accomplish this, the dataset is segmented into a training set and a validation set. Training set is the set in which the model is taught what is desired to be taught, while the validation set is the set utilized to measure how well the model has learned what has been taught [22]. Here the dataset is divided 80% training data and 20% validation data.

Within framework of the SNN used in differentiating the stages of AD, positive and negative image pairs were generated for similarity learning. For this pair generation step, two different images are selected; if the images belong to the same class, they are considered as positive pairs, and if they belong to different classes, they are considered as negative pairs. The features of the images are then extracted with CNN. The evaluation of the similarity relationship among images is conducted through the utilization of the Euclidean distance metric. Error estimation of the relationship is calculated with Contrastive loss function.

The structure of the CNN model consists of 8 convolution layers, 5 pooling layers, 3 fully dependent layers, 1 dropout layer and 2 normalization layers. Rectified Linear Unit (ReLU) is used as the activation function. In the CNN model of this study, average pooling was used as the pooling layer.

4. Results

Python programming language was utilized to create the model. The experiments were conducted using TensorFlow and Keras libraries in Google Colab Pro (Google) environment, a cloud service for deep learning research. The proposed model was tested using 32GB RAM and Tensor Processing Unit (TPU) speed.

The model was trained with 50 epoch parameters, 16 batch sizes and a learning rate of 0.0001. Training was performed on 700 very mild dementia samples and 700 non-dementia samples randomly selected from the dataset. Adaptive moment (Adam) was used as the optimization method.

Figure 5 shows the values obtained during the training for the model. Figure 5 (a) graphs the values obtained from training and validation data. Figure 5 (b) shows the accuracy and 5 (c) shows the sensitivity/specificity graphs. When the accuracy graph is analyzed; the fact that the performances of the training and validation sets have similar increases shows that the model does not face the issue of overlearning. The training accuracy of the model was 99.62% and the validation accuracy was 97.67%. Table 2 indicate other outcome of the model.

(1)

Celal Bayar University Journal of Science
Volume 21, Issue 2, 2025, p 152-158
Doi: 10.18466/cbayarfbe.1529546

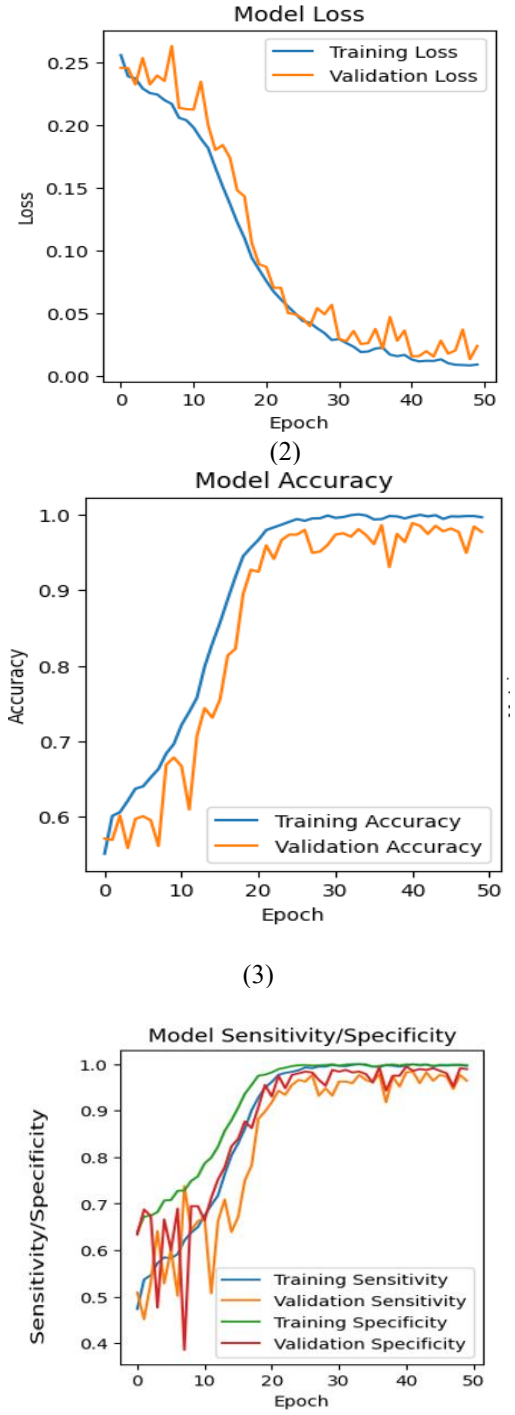

Figure 5. Training process (a) loss, (b) accuracy, (c) sensitivity/specificity of SNN model.

Table 2. AD diagnosis results

Metrics	Acc (%)	Loss (%)	Sens (%)	Spec (%)
Train Data	99.62	0.89	99.61	99.8
Validation Data	97.67	0.23	96.42	98.92

In order to compare the outputs of the study, a confusion matrix was created, shown in Figure 6. The performance of the model on the dataset is depicted by the confusion matrix. This matrix was created based on test data of two different categories. When the confusion matrix is analyzed, 554 data belonging to the Very Mild Demented class were true classified, while 6 data were mistake predicted. In addition, while 540 images in the Non Demented class were correctly predicted, 20 images were incorrectly predicted.

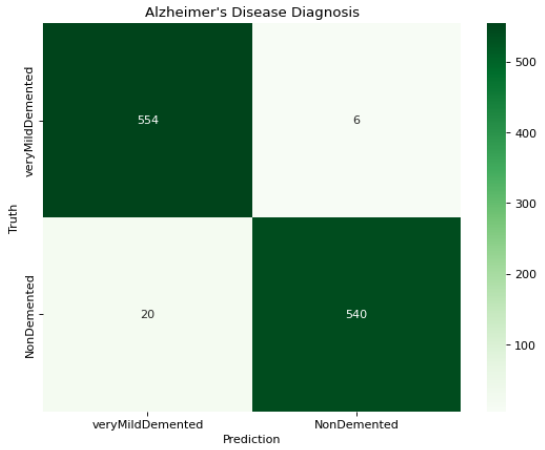

Figure 6. Confusion matrix of the model.

Table 3 shows the model's precision results, recall values and F1 score values calculated according to the confusion matrix.

Table 3. Performance indices of individual class.

Diseases	Precision	Recall	F1- score
Very Mild Demented	0.97	0.99	0.98
Non-Demented	0.99	0.96	0.98

5. Discussion

In the Discussion section, we evaluate different classification methods used for the early diagnosis of AD. As mentioned in the introduction, there are many studies in the literature that perform MRI classification for AD using different methods. Table 4 presents the classification studies in the literature for AD diagnosis together with the corresponding Acc scores obtained from the datasets, facilitating direct comparison.

Table 4. The performance of this study and other studies in the literature for the diagnosis of the disease

Ref	Dataset	Modality	Acc(%)
Binary Classification of Alzheimer's Disease Using Siamese Network for Early Stage Diagnosis (this study)	**Kaggle (2 class / VMD-ND)**	**MRI**	**97.67**
Ahmed and others [14]	ADNI and GARD (2 class / AD-HC)	sMRI	94.03
Liu and others [17]	ADNI (2 class / MCI-AD)	MRI	92.72
Bi and others [18]	ADNI (2 class / MCI-HC)	MRI	65

**Acc: Accuracy, AD: Alzheimer's Disease, ADNI: Alzheimer's Disease Neuroimaging Initiative HC: Health Care, GARD: Gwangju Alzheimer's and Related Dementia , MCI: Mild Cognitive Impairment, MRI: Magnetic Resonance Imaging, ND: Non Dementia sMRI: Structural Magnetic Resonance VMD:Very Mild Dementia*

The ensemble-based classifier method [14] trained the model on the ADNI dataset and performed validation on the GARD dataset. The method achieved 94.03% accuracy in AD-HC classification. The Siamese network method [17] used the ADNI dataset and obtained 92.72% accuracy on MCI-AD and MRI images. The deep metric learning method [18] provided 65% accuracy on MCI-HC MR images in the ADNI dataset. This study achieved 97.67% accuracy when run on the Kaggle dataset. In line with these results, it is proved that the proposed work gives better results.

6. Conclusion

Alzheimer is an irremediable disease that impact the aged individuals and profoundly impacts their lives. It's important to diagnose the disease early to effectively control it. Machine learning enables the creation of models to predict disease with MR images. Promising results are obtained for use in clinical treatment.

In this study, SNN is proposed for early detection of AD. In the literature, SNN model has been used for AD classification, for example, in Siamese network [17] and Deep Metric Learning [18], the features of images were extracted using CNN and the similarity between the images was evaluated by Euclidean distance. The proposed work differs in terms of the complexity of the data set and the CNN model. At the same time, the MR images used in this study are 2D segmented images enhanced by image processing. Compared to the other studies in the literature, the proposed work has proven to provide better results.

The model proposed in this study and tested using Kaggle data for the sorting of disease phases achieved an accuracy of 97.67%. According to the obtained results, the suggested model was found to be suitable for the binary classification problem.

In the future, it is believed that the success of the study can be increased if the dataset is enlarged, image processing algorithms are developed and hybrid with specialized CNN models are used.

Author's Contributions

Ruken Tekin: Methodology, investigation, software, resources, data curation, original draft preparation.

Tuğba Özge Onur: Methodology, formal analysis, validation, writing—review and editing, visualization, supervision.

Ethics

There are no ethical issues after the publication of this manuscript.

References

[1]. M. Vasileios, A. Alexiou, Biomarkers for Alzheimer's Disease Diagnosis, Current Alzheimer Research, vol. 14, no. 11, pp. 1149-1154, 2017.

[2]. S. Przedborski, M. Vila, V. Jackson-Lewis, Neurodegeneration: what is it and where are we?, The Journal of Clinical Investigation, vol. 111, no. 1, pp. 3-10, 2003

[3]. Türkiye Alzheimer Derneği Resmi Web Sitesi. (2 January 2024). Alzheimer Hastalığı Nedir?, https://www.alzheimerdernegi.org.tr/.

[4]. S. Murugan, C. Venkatesan, M. G. Sumithra, X. Gao, B. Elakkiya, M. Akila, M. Subramanian, DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images, IEEE Access, vol. 9, pp. 90319-90329, 2021.

[5]. S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, NeuroImage Clinical, vol. 21, 2019.

[6]. S. Qiu, P. Joshi, M. I. Miller, C. Xue, X. Zhou, C. Karjadi, G. H. Chang, A. S. Joshi, B. Dwyer, S. Zhu, M. C. Kaku, Y. Zhou, Y. J. Alderazi, A. Swaminathan, S. Kedar, M. Saint-Hilaire, S. H. Auerbach, J. Yuan, E. Sartor, R. Au, V. B. Kolachalama, Development and validation of an interpretable deep learning framework for Alzheimer's disease classification, Brain, vol. 143, pp. 1920-1933, 2020.

[7]. Z. Wan, Y. Dong, Z. Yu, H. Lv, Z. Lv, Semi-Supervised Support Vector Machine for Digital Twins Based Brain Image Fusion, Frontiers in Neuroscience, vol. 15, 2021.

[8]. S. Pala, Alzheimer hastalığının erken teşhisi için biyobelirteçlere dayalı stratejik yol haritası derleme çeviri çalışması, Tıbbi Politika Yazısı, 2021.

[9]. Zhang, X. Y., Yang, Z. L., Lu, G. M., Yang, G. F., Zhang, L. J., PET/MR Imaging: New Frontier in Alzheimer's Disease and Other Dementias. Frontiers in Molecular Neuroscience, vol. 10, 2017.

[10].Barthel H., Schroeter M. L., Hoffmann K. T., Sabri O., PET/MR in dementia and other neurodegenerative diseases, Seminars in Nuclear Medicine, vol. 45, pp. 224–233, 2015

[11]. B. A. Mohammed, E. M. Senan, T. H. Rassem, N. M. Makbol, A. A. Alanazi, Z. G. Al-Mekhlafi, T. S. Almurayziq, F. A. Ghaleb, Multi-Method Analysis of Medical Records and MRI Images for Early Diagnosis of Dementia and Alzheimer's Disease Based on Deep Learning and Hybrid Methods, Electronics, vol. 10, no. 22, 2021.

[12]. D. Lu, K. Popuri, G. W. Ding, R. Balachandar, M. F. Beg, Multimodal and Multiscale Deep Neural Networks for the Early Diagnosis of Alzheimer's Disease using structural MR and FDG-PET images, Scientific Reports, vol. 8, 2017.

[13]. Y. Gupta, K. H. Lee, K. Y. Choi, J. J. Lee B. Kim, G. Kwon, Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images, PloS One, vol. 14, no. 10, 2019.

[14]. S. Ahmed, K. Y. Choi, J. J. Lee, B. C. Kim, G. Kwon, K. H. Lee, H. B. Jung, Ensembles of Patch-Based Classifiers for Diagnosis of Alzheimer Diseases, IEEE Access, vol. 7, pp. 73373-73383, 2019.

[15]. H. Nawaz, M. Maqsood, S. Afzal, F. Aadil, I. Mehmood, S. Rho, A deep feature-based real-time system for Alzheimer disease stage detection, Multimedia Tools and Applications, vol. 80, pp. 1-19, 2021.

[16]. J. Shi, X. Zheng, Y. Li, Q. Zhang, S. Ying, Multimodal Neuroimaging Feature Learning With Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer's Disease, IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 1, pp. 173-183, 2018.

[17]. C. F. Liu, S. Padhy, S. Ramachandran, V. Wang, A. Efimov, A. Bernal, L. Shi, M. Vaillant, J. T. Ratnanather, A. V. Faria, B. Caffo, M. S. Albert, M. I. Miller, Using deep Siamese neural networks for detection of brain asymmetries associated with Alzheimer's Disease and Mild Cognitive Impairment, Magnetic Resonance Imaging, vol. 64, pp. 190-199, 2019.

[18]. X. Bi, W. Liu, H. Liu, Q. Shang, Artificial Intelligence-based MRI Images for Brain in Prediction of Alzheimer's Disease, Journal of Healthcare Engineering, vol. 165, 2021.

[19]. T. M. Khan, D. G. Bailey, M. A. U. Khan, Y. Kong, Efficient Hardware Implementation For Fingerprint Image Enhancement Using Anisotropic Gaussian Filter, IEEE Transactions on Image Processing, vol. 26, no. 5, pp. 2116-2126, 2017.

[20]. Y.Lee, S. A. Kassam, Generalized median filtering and related nonlinear filtering techniques, IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 33, no. 3, pp. 672-683, 1985.

[21]. J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. Lecun, E. Sckinger, R. Shah, "Signature Verification Using A Siamese Time Delay Neural Network, Internatioanal Journal of Pattern Recognition Artificial Intelligence, vol. 7, pp. 669-688, 1993.

[22]. M. Toğaçar, Z. Cömert, B. Ergen, Siyam Sinir Ağlarını Kullanarak Türk İşaret Dilindeki Rakamların Tanımlanması, Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi, vol. 23, no. 68, pp. 349-56, 2021.