SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 7, No. 3, 2024, 416-426
DOI: 10.35377/saucis...1493368

RESEARCH ARTICLE

Ensemble-Based Alzheimer's Disease Classification Using Features Extracted from Hog Descriptor and Pre-trained Models

Nedim Muzoğlu[1*], Enver Akbacak[2]

[1] University of Health Sciences, Department of Bioengineering, Istanbul, Türkiye
[2] Fenerbahçe University, Department of Computer Engineering, Istanbul, Türkiye

Corresponding author:
Nedim Muzoğlu, University of Health Sciences, Department of Bioengineering, Istanbul, Türkiye, nedim.muzoglu@sbu.edu.tr

Article History:
Received: 31.05.2024
Accepted: 30.09.2024
Published Online: 23.12.2024

ABSTRACT
Alzheimer's Disease is the most common type of dementia and is a progressive, neurodegenerative disease. The disease worsens over time, and the patient becomes bedridden, unable to move or understand what is happening around him. The main concern of medicine is to slow down the progression of the disease for which no treatment has yet been developed. Artificial intelligence studies have achieved significant success in detecting many diseases. In this study, an artificial intelligence-based approach that uses MR images of the early stage of Alzheimer's Disease to detect the disease at an early stage is presented. Initially, a new dataset was created through the application of the fuzzy technique, thereby expanding the feature space. Then, an ensemble learning-based hybrid deep learning model was developed to reduce the misclassification rate for all classes. The features derived from the inception module, residual modules, and histogram of oriented gradients descriptor are subjected to classification through bagging and boosting algorithms. The proposed model has surpassed many state-of-the-art studies by achieving a high success rate of 99.60% in detecting Alzheimer's disease in its early stages.

Keywords: Alzheimer disease, Hog features, Imbalance dataset, Bagging and boosting, Fuzzy image enhancement

1. Introduction

Alzheimer's Disease (AD) is the most common form of dementia. According to 2020 data, more than 55 million AD patients worldwide. This number is estimated to reach 139 million in 2050 [1]. Moreover, it is also reported that most of the disease will occur in developing countries, where the elderly population is increasing, and this poses a worldwide public health problem. According to TUIK data, and cause-of-death statistics in our country, the rate of older people dying from AD was 3.2% in 2022 [2].

AD is a neurological disease that affects neurons in the brain [3]. In patients with AD, abnormal proteins called amyloid plaques and neurofibrillary tangles accumulate in the brain, interrupting the communication between neurons and causing neurons to lose their functions and die. As AD progresses, these plaques and tangles spread and seriously affect some areas of the brain. This spread causes shrinkage in the hippocampus, which stores memories from the patient's recent past and reduces the ability to remember. Another area of the brain affected is the parietal lobe [4]. The shrinkage of the parietal lobe negatively affects various abilities, such as maintaining body balance, perception ability, and calculation. Thus, the disease destroys brain functions. It prevents the patient from performing tasks such as swallowing and breathing due to the weakening of the respiratory muscles, thus causing complications in the organs and causing the death of the patient.

The clinical evaluation of the disease encompasses three main phases. Symptoms observed in the mild phase include memory and speech problems, difficulty in routine daily tasks, loss of balance, impairment in abstract thinking, challenges in planning and problem-solving, misplacement of items, decision-making difficulties, and manifestations of depression. In the intermediate phase, notable are the deficits in executive functions, particularly in personal care. In the advanced stage, there is an increasing dependence on the bed and family members, accompanied by an increased risk of infection and embolism [5]. Due to its gradual progression spanning numerous years, AD may necessitate up to two decades for early-stage symptoms to manifest. Early diagnosis of the disease is crucial for initiating pharmacological treatment methods. Although early treatment methods cannot stop the progression of the disease, they can slow it down.

Cite as: N. Muzoğlu, E. Akbacak, "Ensemble-Based Alzheimer's Disease Classification Using Features Extracted from Hog Descriptor and Pre-trained Models," *Sakarya University Journal of Computer and Information Sciences*, vol.7, no. 3, pp. 416-426, 2024. Doi: 10.35377/saucis…1493368

PET, CT/PET, MRI, and PET/MRI imaging technologies are medical imaging techniques used to detect disease. Among these health technologies, magnetic resonance (MR) imaging emerges as an economically viable modality, devoid of radiological effects, and widely accessible across numerous health institutions. This method requires the expertise of radiologists to discern the course of the disease through examination of anatomical changes identified in MRI images. However, examining high-dimensional data obtained from neuroimaging modalities requires experienced radiologists, and it is time-consuming. Therefore, in the last twenty years, significant successes have been achieved in disease detection through artificial intelligence methods, aiming to support this issue. Researchers have also undertaken artificial intelligence studies focusing on the cerebral alterations induced by this disease, leveraging the features extracted from MRI images.

Traditional feature extraction methods and machine learning algorithms are commonly employed in studies aimed at detecting AD are as follows. Many machines learning models have been developed to classify brain tissue types, including white matter, cerebrospinal fluid (CSF), and gray matter (GM), with the aim of detecting AD. Features extracted from MRI images have been commonly employed in these studies. A local feature descriptor was obtained by [6] combining the fast Hessian detector to detect important voxels from different sections of MRI slices and the local binary pattern texture operator to determine variations in different sections, resulting in an 88.99% success rate achieved through training with a CNN network. [7] utilized the bag-of-feature method in conjunction with support vector machines (SVM) to classify MRI images within the AD Neuroimaging Initiative (ADNI) dataset. This approach yielded an accuracy rate of 93%. Another deep learning team is utilized ADNI dataset to detect the stages of AD using machine learning methods. They applied k-nearest neighbours, decision tree, Naive Bayes, and generalized linear model GLM with 88.24% accuracy [8]. In a different study [9], initially input the features obtained from the ADNI dataset into the kernel principal component analysis module and subsequently projected the Principal Component Analysis (PCA) coefficients onto the linear discriminant space. Then, they employed SVM for classification and achieved a disease detection accuracy of 93%. In the study [10] utilized the semi-supervised machine learning method using the ADNI dataset to detect the transition stage from mild cognitive impairment to AD. MRI images underwent a 10-fold cross-validation after the feature selection process, resulting in a 90% AUC score. [11] attained a success rate of 98.17% in the early diagnosis of Alzheimer's disease (AD) by employing a combined feature method. This method integrates the extraction of Histogram of Oriented Gradients (HOG) features and PCA feature selection, followed by classification using decision trees. Unlike many other studies, Liu at all. utilized spectrogram features extracted from speech data instead of image data to detect AD. In the study comparing speech data from AD patients with the Dem@Care dataset, they achieved an accuracy of 83% using the logistic regression method [12].

Although machine learning studies have been carried out to analyze high-dimensional data, they could be time-consuming since feature extraction, feature selection, dimensionality reduction, and classification steps could be computationally intensive. Given that these studies demand researchers to possess subject-specific skills, deep learning methods, which offer numerous advantages surpassing those of machine learning studies, have become increasingly popular in recent years. The LeNet deep learning architecture constitutes a pivotal milestone in the evolution of deep learning models [13]. It is based on providing the features extracted from convolutional networks as input to increase artificial neural networks' training and classification success. Subsequently, numerous deep learning models have emerged, showcasing their efficacy through notable achievements in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competitions. AlexNet and VGG16, with their sequential layer structure, were among the most prominent models in the following years [13]. However, these models could not overcome some problems, such as vanishing gradient problems. The solution to this problem has been ResNet models [14] developed with the help of residual blocks. While these models have addressed the issue of updating weights in deep layers, the increase in dimensionality with increasing depth remains another problem that needs to be resolved. GoogLeNet [15] , developed by Google Researcher, solved this problem with its Inception modules. Deep learning models continue to develop by providing solutions to issues encountered.

On the other hand, various studies have shown that machine learning applications have been successfully used to classify the features extracted by deep learning methods. Machine learning methods indeed have the potential to be highly successful in classifying datasets with unbalanced data distributions between classes. [16] introduced a deep learning-based pipeline model for AD diagnosis. This model implements an average strategy, concatenating bagging and majority voting methods, and combining features extracted from fine-tuned models, namely AlexNet, ResNet101, and InceptionResNetV2. It achieved a success rate of 98.51%. In the capsule CNN-based model proposed by [17] for the early diagnosis of AD, the preprocessing steps include histogram equalization and Gaussian filtering. Subsequently, segmentation is performed using UNET, while feature extraction is based on the Improved Gray-Level Run Length Matrix and the Gray-Level Size Zone Matrix. In addition, the model integrates Equestrian Ecosystem Optimization for feature selection. This ensemble approach yields to an overall success rate of 94.52%. [18] employed an adaptive synthetic oversampling technique on the highly imbalanced ADNI dataset to efficiently discern AD. Subsequently, the combined prediction values of the EfficientNet-B2 and VGG-16 models were obtained. This approach resulted in a prediction success rate of 94%.

In the literature review, the efficacy of methodologies leveraging Histogram of Oriented Gradients (HOG) features, deep learning models, and ensemble machine learning methods stands out. Based on this observation, the current study proposes the classification of features derived from both HOG and deep learning models employing ensemble machine learning methods. Since the AD data set, we used to be imbalanced between classes, the proposed approach as a solution is as follows. The distinctive features extracted from both inception and residual deep learning models, which capture detailed and coarse features from images at various scales and exhibit superior performance in handling gradient problems, were merged. These were later combined with HOG features. Subsequently, these combined features were classified using ensemble machine learning methods, among the leading approaches for handling imbalanced datasets between classes. The contributions of this study are:

- The reconstruction of the AD data set utilizing the Fuzzy logic image enhancement method to expand the feature space.
- The problem of features that cannot be learned sufficiently has been tried to be overcome with the help of multi-scale deep learning models and residual structures.
- The Discriminant HOG features extracted from the the dataset were also utilized to capitalize on their potential to enhance classification accuracy.
- Ensemble learning-based bagging and boosting methods achieved high classification success as a solution to the misclassification problem in minority classes in imbalanced datasets.

The remaining sections of this research are organized as follows: Next section will provide an exposition on the dataset, models, and the proposed method, while the subsequent section will detail the experimental results and a comparison of the success achieved in similar studies pertaining to AD classification will be conducted. Lastly, the concluding section will include the discussion and offer conclusive remarks.

2. Materials and Methods

2.1. Dataset

The AD dataset utilized in this study was obtained from Kaggle and is openly accessible [19]. The dataset has 6400 MRI images and consists of four classes of early stages of the disease: non-dementia, very mild dementia, mild dementia, and moderate dementia. The ratio of training to testing was set at 80:20, a common choice observed in numerous successful studies of a similar nature. Brain MRI images are sized at 176x208 pixels and are in .jpg format. Sample images from each class of both the Alzheimer Dataset and the AlzheimerF Dataset, which underwent enhancement using the fuzzy method, are depicted as seen in Figure 1. Additionally, details regarding the number of images in the datasets are as seen in Table 1.

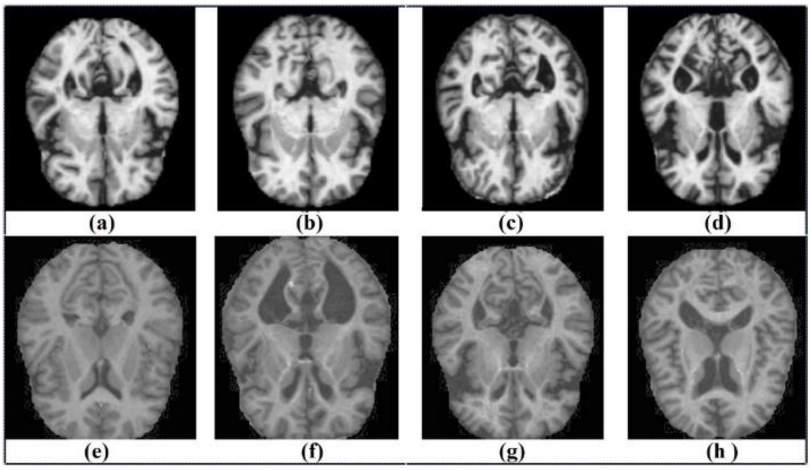

Figure 1. (a) Alzheimer Non-dementia, (b) Alzheimer Very Mild Dementia, (c) Alzheimer Mild Dementia, (d) Alzheimer Moderate Dementia, (e) AlzheimerF Non-dementia, (f) AlzheimerF Very Mild Dementia, (g) AlzheimerF Mild Dementia, (h) AlzheimerF Moderate Dementia

Table 1. AD Early-Stage Dataset Details

Classes	Train	Test	Total
NonDemented	2560	640	3200
VeryMildDemented	1792	448	2240
Mild dementia	717	179	896
ModerateDemented	51	13	64
Total	5120	1280	6400

2.2. Deep Learning Models

GoogLeNet [20] and ResNet18 [21] were employed in the deep architecture of this study. GoogLeNet was initially introduced in 2014 [22] at the ILSRVC, where it competed on behalf of Google Research’s. Before GoogLeNet, researchers assumed that the deeper a network is, the more efficient it is. However, as the layers deepen, it causes the gradients to approach zero, and the learning process becomes saturated, known as the vanishing gradient problem [23]. In addition, the computational costs increase. With GoogLeNet, it is ensured that although a CNN network is more complex and deeper, its computational cost is not excessively high. It is based on using parallel paths, i.e., initial modules, each with different filter sizes, instead of the gradual growth of the network depth. In an inception module [24], the outputs of CNN paths with filters of different sizes are concatenated. Thus, features at different scales are captured at the same time. In the inception module in a layer, each of the various size filters, such as 1x1, 3x3, and 5x5, is located on a different parallel path. Paths with 3x3 and 5x5 filters allow both small and global patterns to be captured. 1x1 convolutions performed before those performed by 3x3 and 5x5 filters reduce the number of input channels and provide dimension reduction. In this way, the computational cost is also reduced. As a result, although the network is deeper, both efficiency increases and computational cost is reduced. GoogLeNet consists of multiple stacked inception layers, each with an average pooling or fully connected layer at the end. Additionally, auxiliary classifiers were used in intermediate layers. These classifiers reduce the vanishing gradient problem by providing additional supervision information for backpropagation. Overall, the GoogLeNet architecture achieves multi-scale feature extraction, decreased effectiveness of the vanishing gradient, and computational efficiency. GoogLeNet has 22 layers, and inception modules are stacked. Each inception module contains Convolution, pooling, and concatenation, normalization operations, respectively. There are different versions of inception deep learning. GoogLeNet is also known as InceptionV1.

In deep networks, the gradients calculated for previous layers tend to shrink exponentially as the network gets deeper, and updating the weights becomes difficult, thus losing the network's ability to learn complex relationships in the data. As a solution to this challenging problem, Residual Network was introduced in the 2015 paper “Deep Residual Learning for Image Recognition” by [25]. This groundbreaking architecture, which enables easier flow from previous to subsequent layers by directly skipping several layers, has now solved the gradient problem. In the ResNet structure, the block that passes the input value through layers and connects the input back to the output is called the residual block, and it is shown as seen in Figure 2.

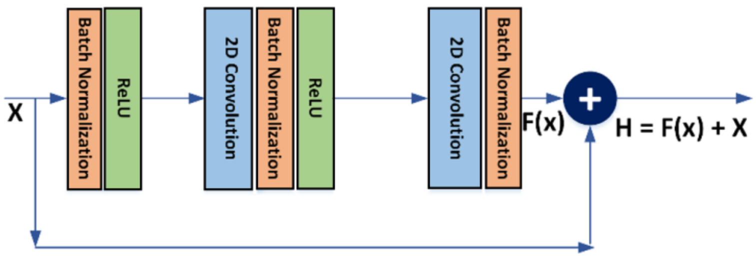

Figure 2. The Residual Blocks Architectures

ResNet18 and GoogLeNet were used in this study to benefit from these superior features. Moreover, given the imbalance within the dataset and the limited number of images in the minority classes, the first 10 layers of the models were frozen, and weight transfer based on transfer learning was implemented. Gradient Descent [26] was preferred for training the models and optimizing the error function.

2.3. Histogram of Oriented Gradients

The Histogram of Oriented Gradients (HOG) method continues to be recognized as a robust feature descriptor, maintaining relevance due to its resilience against illumination variations and transformations induced by medical devices in tissue imaging [27]. HOG can basically be defined as a feature descriptor that discards unnecessary elements while selectively extracting relevant information. The utilization of HOG descriptors on images is linked to the concept of gradients and directions. Gradients represent sudden transitions between pixels within an image, where the shift from darker to lighter shades is denoted as a positive gradient. To accomplish this task, the gradient magnitude and gradient angle of each pixel must first be computed. Subsequently, the histogram of gradients for each cell should be derived within the image partitioned into cells. Thus, the descriptors of the image are obtained by aggregating the HOG features from each cell [28]. In this study, the cell size for the Histogram of Oriented Gradients (HOG) feature extraction algorithm was set to 32 x 32 pixels, while all other parameters were maintained at their default values as provided by MATLAB 2022b. In this study, a total of 780 features were extracted per image from both the Alzheimer and AlzheimerF datasets, employing the HOG descriptor according to the parameters.

2.4. Fuzzy Image Enhancement Method

In medical imaging, image enhancement methods increase hidden image details or contrast, allowing medical professionals to identify abnormalities, tumors, and other critical information more effectively. Fuzzy image enhancement is achieved by mapping the gray levels of the image onto a fuzzy plane [29]. This technique creates an image with higher contrast than the original image by giving more weight to the image's average gray levels than those further away from the average value. Fuzzy image processing consists of fuzzification, where the image is encoded, fuzzy layer operations, and defuzzification, where the results are decoded [30]. The fuzzy image enhancement was performed by utilizing the Python Open CV library. The epsilon, gamma and variance values used in the fuzzy image generation are 0.00001, 1, and 0.35, respectively.

We aim to obtain rich features by using the Alzheimer Dataset, and AlzheimerF Dataset which is improved with the fuzzy logic image enhancement method. This approach based on the deep integration of GoogLeNet, which can extract multi-scale features, and the residual neural network, which successfully vanishes gradient problems, and also HOG technique.

2.5. Ensemble Methods

The AD data sets were classified with Bagging and Boosting structures, which are ensemble algorithms [31]. In the bagging method, random samples are selected from the original dataset for each of the different classifiers. However, the total number of samples in the subsample groups must be equal to the total number of samples in the original data set. But the selected samples continue to remain in the original dataset. This means that the same samples can be present in different classifiers [32]. The average of the prediction of each trained classifier is taken, so the goal is to reduce the variance with the bagging method, and thus overfitting. In this study, decision trees were used as learners in the bagging method, and the number of learners and the maximum split were set to 20 and 5000, respectively. The Boosting method is based on a series of models that works sequentially and focus on examples that are difficult to classify [31]. Subsequently, the training process of the method proceeds by assigning greater weight to predictions that were misclassified by the preceding classifier.

2.6. Artificial Intelligence Approaches to Tackle the Imbalance Data Distribution

The ratio of samples in the four classes of the data set is 1:13:35:50. Therefore, the dataset is imbalanced between classes. In imbalanced data sets, the majority classes dominate in determining the weights during training. This situation results in misclassification during the prediction process. There are three different machine learning approaches to dealing with the classification of imbalanced data sets: data level, cost-sensitive, and ensemble methods [33]. In data-level methods, imbalance problems are dealt with by modifying the data distribution. Under-sampling, a method of removing data from the majority classes, and oversampling, a technique of adding synthetic data to minor classes, are data-level imbalance methods. Since new samples are created by copying existing data in the oversampling method, these samples cause memorization of specific patterns and noise in the training data, which results in overfitting rather than accurate prediction of the test data. On the other hand, since under sampling methods discard potentially valuable data from the majority class, the relevant class cannot be represented correctly, causing errors in the analysis [34]. Another approach, cost-sensitive learning aims to minimize misclassification of the minority classes by assigning cost values to different classes, usually appropriate to the class's number of samples [35]. On the other hand, ensemble machine learning approaches, which combine multiple classifiers to reduce errors and increase their accuracy, achieve better results in classifying imbalance datasets than the other two methods. However, these approaches may not guarantee high classification scores unless the features are considerably representative. Obtaining the features that best represent the dataset is an essential issue in classification problems, and convolutional deep learning models are one of the most effective methods for extracting representative features. However, feature extraction techniques from imbalanced datasets may result in biased features favoring the majority class and poor classification performance for the minority class [36]. From this point of view, it is inevitable to use the rich feature extraction power of deep learning models and the powerful structure of ensemble classifiers in the classification of imbalanced data sets.

Various valuable studies have been conducted on hybrid structures in which deep learning models are used together with ensemble machine-learning methods. Yuan et al. reported that they achieved a 24.7% increase in maximum accuracy by applying regularization methods to the unbalanced endoscope video dataset in the detection of bowel cancer [37]. In the study, the dataset they employed is balanced with synthetic data. [34] integrated the synthetic minority oversampling technique (SMOTE) method with the cross-validated committee filter (CVCF) technique to overcome the problems that cause fatal errors in imbalanced datasets [34]. In their work, they used the ensemble support vector machine introduced as a classifier, together with the majority voting strategy. [38] used pre-trained convolutional neural networks to detect skin cancer. They proposed an ensemble model to solve the imbalance class problem due to the scarcity of representative images from malignant tumors. Considering the proposed technique's area under the ROC curve (AUC-ROC) performance, the ensemble SVM structure has been reported to provide a 16% increase over the pure CNN structure. So, pure CNN feature extraction techniques used with imbalanced datasets may result in biased features favoring the majority class and poor classification performance for the minority class. To address this problem, this study proposes an approach based on the deep integration of GoogLeNet, which can extract multi-scale features for handling the imbalanced AD dataset and the residual neural network, which is successful in mitigating vanishing gradient problems. Moreover, ensemble models were employed to mitigate the impact of biased features favoring the majority class and attain high classification accuracy.

2.7. Proposed Approach

The proposed method entails utilizing 6400 images from an open-access database containing four classes for the early-stage diagnosis of AD. To enrich the feature space, the AlzheimerF dataset was generated by applying the fuzzy enhancement technique to this dataset. Features from both datasets were extracted from the final layers of the GoogLeNet and ResNet18 models, resulting in 1000 features per image in each dataset. Additionally, features from both datasets were extracted using the Histogram of Oriented Gradients (HOG) technique, renowned for its efficacy in feature extraction studies, thereby yielding 5560 features obtained through concatenation of the features extracted from the pre-trained models. To mitigate the misclassification error of minority classes in imbalanced datasets, Bagging and RUS-Boosting tree algorithms were employed, owing to their specific operational algorithms. During the training of these algorithms, the 5-fold cross-validation method, known for yielding optimal results, was favored to alleviate the potential high error stemming from the train and test distribution of the unbalanced dataset. They were then classified to predict the early stages of AD with the highest accuracy. The block diagram of the proposed model is given as seen in Figure 3.

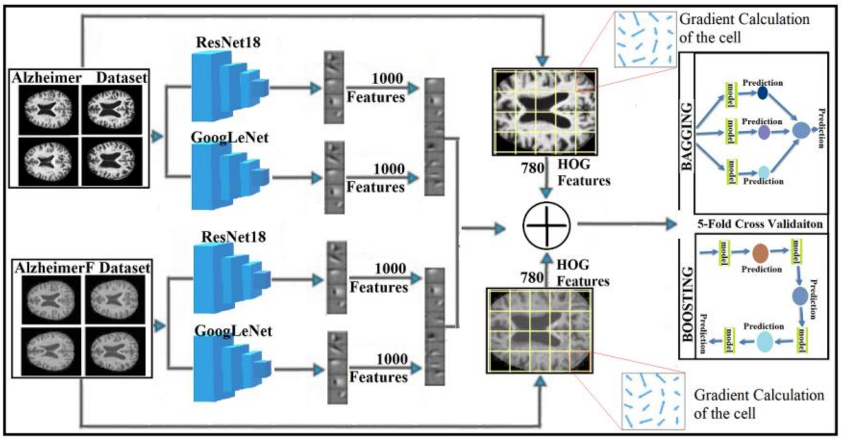

Figure 3. The Functional Block Diagram of The Proposed Method

3. Experimental Results

3.1. Experimental Setup

The confusion matrix values were utilized to evaluate the model. The accuracy metric (Acc) is obtained by dividing the number of correctly predicted samples by the total number of predicted samples. To obtain recall and F-Score (F-Scr), which are other essential evaluation metrics, true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values must be obtained from the confusion matrix [39]. Evaluation metrics are obtained as shown in Eq. 1-3.

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \tag{1}$$

$$\text{F}-\text{Scr} = \frac{2xTP}{2xTP+FP+FN} \tag{2}$$

$$\text{Acc} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

This study performs experimental results in Keras API, Python, and MATLAB R2022b. The computer used in the study has a Windows 10 (64-bit) operating system and an NVIDIA GeForce RTX3070 graphics card. The mini-batch size, which is related to the suitability of the equipment, was preferred as 64. The learning rate and the number of epochs were set to 0.0001 and 100, respectively.

3.2. Experimental Results

In the first step of this study, deep learning models were trained with the AD Dataset, and AlzheimerF datasets obtained by the fuzzy logic image enhancement method. The training progress is presented as seen in Figure 4. This figure shows the overfitting problem commonly encountered during the training process of the imbalanced datasets. In this step, 85.26% and 86.72% accuracy are achieved in the GoogLeNet and ResNet18 models using the AD Dataset, respectively. For the AlzheimerF dataset, these accuracies are 86.02% and 87.26%, respectively.

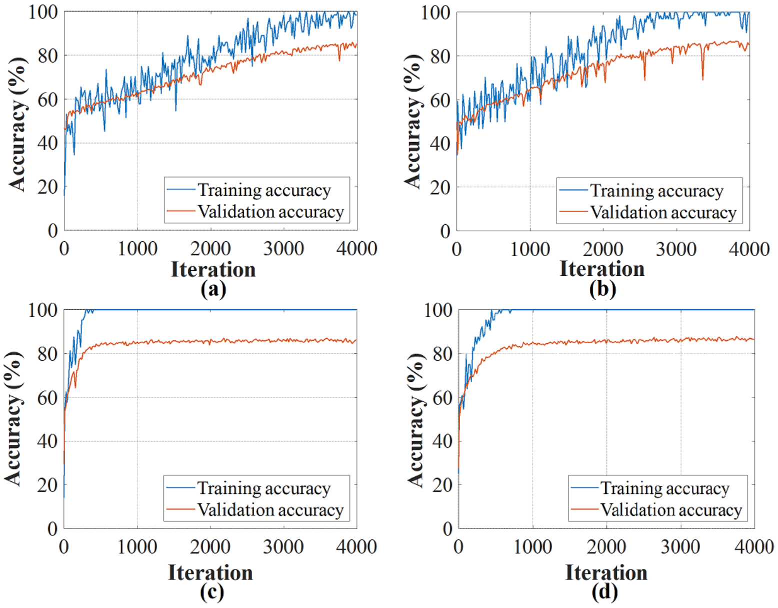

Figure 4. (a) GoogLeNet Alzheimer Dataset, (b) GoogLeNet AlzheimerF Dataset,
(c) ResNet18 Alzheimer Dataset, (d) ResNet18 AlzheimerF Dataset

Figure 5 shows the confusion matrix results of these classification processes. It is seen that the classification success of the Softmax classifier with 1000 features in each model is not sufficiently high, particularly when evaluating human health.

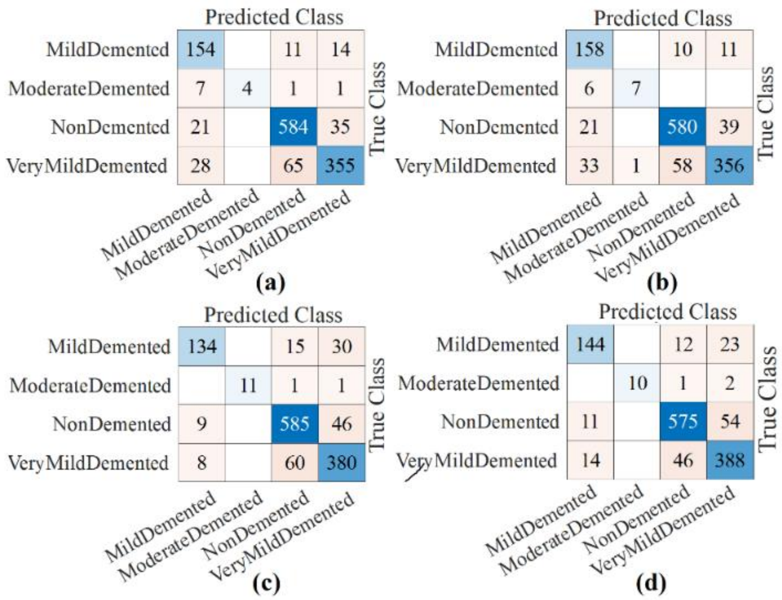

Figure 5. (a) GoogLeNet Alzheimer Dataset, (b) GoogLeNet AlzheimerF Dataset,
(c) ResNet18 Alzheimer Dataset, (d) ResNet18 AlzheimerF Dataset

Given that the method's primary aim is to minimize misclassification in minority classes by reducing overfitting, the features that best represent the datasets were initially extracted from the thoroughly trained GoogLeNet, ResNet18 models. These features were then fed into the Bagging and RUS-Boosting Tree modules. Bagged Trees and Boosted Trees classifiers, which are ensemble learning algorithms, are preferred in classifying the imbalance dataset to classify minority classes with low misclassification errors. In the first step of the proposed approach achieved classification success rates of 97.34% and 97.57% for the Bagging and RUS-Boosting Tree models, respectively, in the detection of AD solely through the utilization of features acquired from deep learning modules. Figure 6 illustrates the confusion matrices acquired subsequent to the classification process of Bagging and RUS-Boosting Tree models utilizing 4000 features extracted from Alzheimer and AlzheimerF datasets through deep learning models.

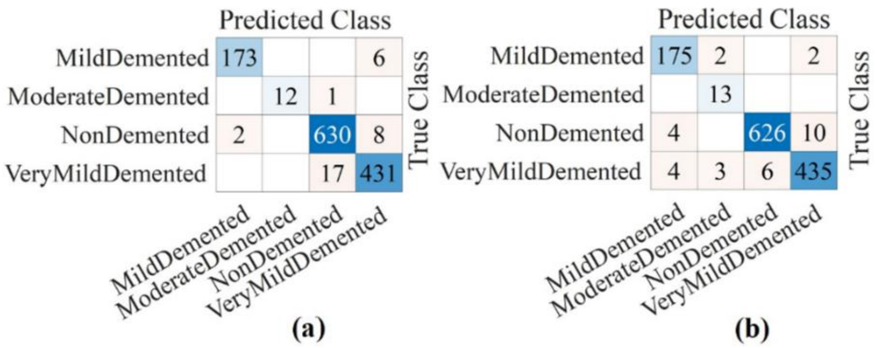

Figure 6. (a) 4000 Feature with Bagged Trees, (b) 4000 Feature with Rus-boosted Trees

In the Figure 7, the outcomes of the proposed approach are showcased, achieved through ensemble classification by enriching the features extracted from deep learning models with an additional set of 1440 HOG features.

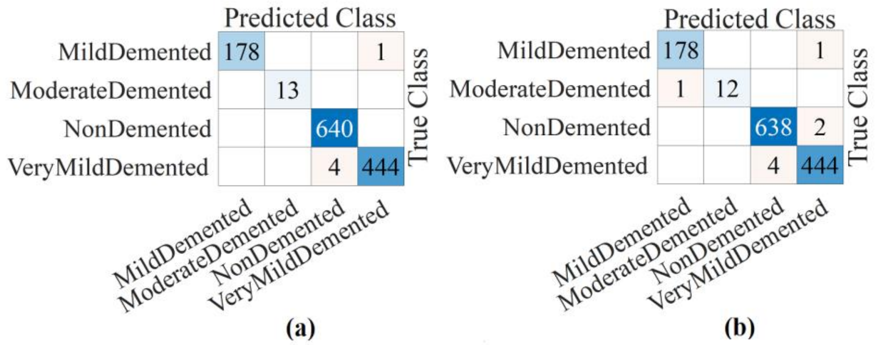

Figure 7. (a) 5440 Feature with Bagged Trees, (b) 5440 Feature with Rus-boosted Trees

The confusion matrices in Figure 7 clearly show that enriching the feature space through the inclusion of HOG features has facilitated ensemble classifiers in attaining notable success rates in reducing misclassifications. When the results scrutinized concerning misclassification errors within minority classes stemming from the imbalanced dataset, one of the primary objectives of this study, the Bagging method rendered predictions with only a single error in total for the mild and moderately demented classes. Moreover, in the very mild dementia class, closely associated with early-stage disease detection, both the Bagging and Boosting methods achieved remarkably high classification success, with only four misclassifications predicted in total. While the utilization of ResNet-18, a prominent deep learning model, attained an 87% success rate on the AlzheimerF dataset, the proposed model achieved a substantial 14.14% increase in accuracy, showcasing its efficacy and superiority in performance.

In conclusion, the application of the proposed approach resulted in notable accuracies for both bagging and boosting methods. Specifically, bagging methods yielded an accuracy of 99.60%, while boosting methods achieved an accuracy of 99.40%. The classification results are evaluated as seen in the Table 2.

Table 2. Performance Metrics Obtained with Ensemble Learning Classifier for the AD

Models	Classes	Recall (%)	Pre (%)	F-Scr (%)	Acc (%)
Bagged Trees	Mild Demented	100	99.44	99.72	99.60
	Moderate Demented	100	100	100	
	Non Demented	99.38	100	99.69	
	VeryMild Demented	99.76	99.11	99.44	
RUS-Boosted Trees	Mild Demented	99.44	99.44	99.44	99.37
	Moderate Demented	100	92.31	96	
	Non Demented	99.38	99.69	99.53	
	VeryMild Demented	99.33	99.11	99.22	

Numerous studies have been undertaken concerning the early-stage detection of AD employing a variety of methodologies, including traditional techniques, deep learning architectures, and hybrid models. Table 3 compares similar approaches using the AD dataset.

Table 3. The Comparison of the Proposed Approach with Similar Studies

Authors	Methods	Recall (%)	F-Scr (%)	Acc (%)
[19], 2018	Ensemble Learning, Convolutional Neural Networks.	93	92	93.18
[40], 2022	Transfer Learning, Alex Net.	94.21	94.12	94.53
[41], 2021	Oversampling, Occlusion Sensitivity Maps.	95.20	95.30	95.20
[42], 2022	Transfer Learning, Oversampling.	97	97	97.05
[43], 2022	Convolutional Neural Networks, Slice Based methods.	98	97	98.37
[44], 2022	Transfer Learning, Gradient Activation Maps	98.14	98.14	98.21
[45], 2023	Transfer Learning, PCA	99.87	99.89	99.88
Proposed approach	Ensemble Learning, Transfer Learning, Hog descriptor	99.79	99.64	99.60

4. Conclusion

This study proposes the detection of AD at an early stage through the utilization of an ensemble learning approach. The primary objective of this study is to mitigate the challenges posed by imbalanced data distributions through the expansion of the feature space. In this methodology, features extracted from the Alzheimer, and AlzheimerF datasets via traditional and transfer learning-based techniques are subsequently subjected to classification. The transfer learning models employed were chosen from two fundamentally different structures, the inception block, and the residual structure, to extract various features from the dataset. Moreover, the traditional hog feature extractor, which is still up to date, was preferred due to its outstanding success. Classifiers such as bagging and boosting, which obtain results with the predictive power of many classifiers, were used to minimize the misclassification error encountered in the imbalanced datasets. A closer look at the dataset reveals that the classes 'mild dementia' and 'moderate dementia' are significantly underrepresented and have only a small number of samples compared to the other classes. Especially the high imbalance moderately demented and mild demented classes were successful with the bagging method, except for one error.

Early diagnosis of Alzheimer's disease is crucial because early intervention can slow down the progression of the disease. The disease's progression can be assessed from MRI images based on radiologists' reports. However, differences in decision-making among radiologists may occur due to variations in cognitive load. Hence, this study and similar approaches are recommended as computer-supported artificial intelligence assistants. In future studies with this dataset, the aim is to detect the disease in minority classes with the lowest misclassification using a cost-sensitive learning-based deep learning approach, one of the ensemble methods.

References

[1] L. C. Arevalo-Flechas, “Latino Alzheimer ’ s Caregiving Neither a Burden nor a Carga,” *Univ. ealth Sci. J. Nurs.*, vol. 1, no. 2, pp. 92–103, 2019.

[2] M. ŞENER and H. H. TEKİN, “Sosyal Belediyecilik Bağlamında Yaşlı Bakım ve Alzheimer Gündüz Yaşam Merkezleri,” *Geriatr. Bilim. Derg.*, vol. 3, no. 3, pp. 138–146, 2020, doi: 10.47141/geriatrik.737313.

[3] D. S. Knopman *et al.*, “Alzheimer disease,” *Nat. Rev. Dis. Prim.*, vol. 7, no. 1, pp. 1–21, 2021, doi: 10.1038/s41572-021-00269-y.

[4] E. Keleş, U. Fzt, S. Özalevli, İ. Kâtip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi, and F. ve Rehabilitasyon Bölümü, “Alzheimer Hastalığı ve Tedavi Yaklaşımları Alzheimer’s Disease and Treatment Approaches,” *İzmir*

Katip Çelebi Üniversitesi Sağlık Bilim. Fakültesi Derg., vol. 3, no. 2, pp. 39–42, 2018, [Online]. Available: https://www.alz.org/alzheimers_disease_stages_of_alzheimers.asp.

[5] P. S. Aisen, G. A. Jimenez-Maggiora, M. S. Rafii, S. Walter, and R. Raman, "Early-stage Alzheimer disease: getting trial-ready," *Nat. Rev. Neurol.*, vol. 18, no. 7, pp. 389–399, 2022, doi: 10.1038/s41582-022-00645-6.

[6] A. Francis and I. A. Pandian, "Early detection of Alzheimer's disease using local binary pattern and convolutional neural network," *Multimed. Tools Appl.*, vol. 80, no. 19, pp. 29585–29600, 2021, doi: 10.1007/s11042-021-11161-y.

[7] D. Bansal, K. Khanna, R. Chhikara, R. K. Dua, and R. Malhotra, "Classification of Magnetic Resonance Images using Bag of Features for Detecting Dementia," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 131–137, 2020, doi: 10.1016/j.procs.2020.03.190.

[8] M. Shahbaz, S. Ali, A. Guergachi, A. Niazi, and A. Umer, "Classification of Alzheimer's disease using machine learning techniques," *DATA 2019 - Proc. 8th Int. Conf. Data Sci. Technol. Appl.*, no. Data, pp. 296–303, 2019, doi: 10.5220/0007949902960303.

[9] S. Alam and G. R. Kwon, "Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM," *Int. J. Imaging Syst. Technol.*, vol. 27, no. 2, pp. 133–143, 2017, doi: 10.1002/ima.22217.

[10] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *Neuroimage*, vol. 104, pp. 398–412, 2015, doi: 10.1016/j.neuroimage.2014.10.002.

[11] J. A. Dinu and R. Manju, "A novel modelling technique for early recognition and classification of Alzheimer's disease," *2021 3rd Int. Conf. Signal Process. Commun. ICPSC 2021*, no. May, pp. 21–25, 2021, doi: 10.1109/ICSPC51351.2021.9451803.

[12] L. Liu, S. Zhao, H. Chen, and A. Wang, "A new machine learning method for identifying Alzheimer's disease," *Simul. Model. Pract. Theory*, vol. 99, no. November 2019, p. 102023, 2020, doi: 10.1016/j.simpat.2019.102023.

[13] X. Zhang, "The AlexNet, LeNet-5 and VGG NET applied to CIFAR-10," *Proc. - 2021 2nd Int. Conf. Big Data Artif. Intell. Softw. Eng. ICBASE 2021*, pp. 414–419, 2021, doi: 10.1109/ICBASE53849.2021.00083.

[14] C. Szegedy *et al.*, "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.

[15] C. Szegedy *et al.*, "Going Deeper with Convolutions," Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.4842.

[16] A. Loddo, S. Buttau, and C. Di Ruberto, "Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method," *Comput. Biol. Med.*, vol. 141, no. August 2021, p. 105032, 2022, doi: 10.1016/j.compbiomed.2021.105032.

[17] R. Butta, M. S. Shaik, and G. L. N. Murthy, "Ensemble deep learning approach for early diagnosis of Alzheimer's disease," *Multimed. Tools Appl.*, no. 0123456789, 2024, doi: 10.1007/s11042-023-18084-w.

[18] M. Mujahid, A. Rehman, T. Alam, F. S. Alamri, S. M. Fati, and T. Saba, "An Efficient Ensemble Approach for Alzheimer's Disease Detection Using an Adaptive Synthetic Technique and Deep Learning," *Diagnostics*, vol. 13, no. 15, pp. 1–20, 2023, doi: 10.3390/diagnostics13152489.

[19] J. Islam and Y. Zhang, "Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks," *Brain Informatics*, vol. 5, no. 2, 2018, doi: 10.1186/s40708-018-0080-3.

[20] L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, and G. Ramkumar, "GoogLeNet-Based Ensemble FCNet Classifier for Focal Liver Lesion Diagnosis," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 6, pp. 1686–1694, 2020, doi: 10.1109/JBHI.2019.2942774.

[21] M. Guo and Y. Du, "Classification of Thyroid Ultrasound Standard Plane Images using ResNet-18 Networks," *Proc. Int. Conf. Anti-Counterfeiting, Secur. Identification, ASID*, vol. 2019-Octob, no. i, pp. 324–328, 2019, doi: 10.1109/ICASID.2019.8925267.

[22] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/S11263-015-0816-Y.

[23] S. H. Noh, "Analysis of gradient vanishing of RNNs and performance comparison," *Inf.*, vol. 12, no. 11, 2021, doi: 10.3390/info12110442.

[24] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," *31st AAAI Conf. Artif. Intell. AAAI 2017*, pp. 4278–4284, 2017, doi: 10.1609/aaai.v31i1.11231.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[26] S. Ruder, "An overview of gradient descent optimization algorithms," pp. 1–14, 2016, [Online]. Available: http://arxiv.org/abs/1609.04747.

[27] C. I. Patel, D. Labana, S. Pandya, K. Modi, H. Ghayvat, and M. Awais, "Histogram of oriented gradient-based fusion of features for human action recognition in action video sequences," *Sensors (Switzerland)*, vol. 20, no. 24, pp. 1–33, 2020, doi: 10.3390/s20247299.

[28] T. Q. Bao, N. T. T. Kiet, T. Q. Dinh, and H. X. Hiep, "Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks," *J. Inf. Telecommun.*, vol. 4, no. 2, pp. 140–

150, 2020, doi: 10.1080/24751839.2019.1666625.

[29] M. Hanmandlu and D. Jha, "An optimal fuzzy system for color image enhancement," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2956–2966, 2006, doi: 10.1109/TIP.2006.877499.

[30] S. Banerjee, S. K. Singh, A. Chakraborty, A. Das, and R. Bag, "Melanoma diagnosis using deep learning and fuzzy logic," *Diagnostics*, vol. 10, no. 8, 2020, doi: 10.3390/diagnostics10080577.

[31] M. Singh, S. Bansal, S. Ahuja, R. K. Dubey, B. K. Panigrahi, and N. Dey, "Transfer learning–based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data," *Med. Biol. Eng. Comput.*, vol. 59, no. 4, pp. 825–839, Apr. 2021, doi: 10.1007/s11517-020-02299-2.

[32] L. Gang, Z. Haixuan, E. Linning, Z. Ling, L. Yu, and Z. Juming, "Recognition of honeycomb lung in CT images based on improved MobileNet model," *Med. Phys.*, vol. 48, no. 8, pp. 4304–4315, 2021, doi: 10.1002/mp.14873.

[33] Z. Chen, J. Duan, L. Kang, and G. Qiu, "Class-Imbalanced Deep Learning via a Class-Balanced Ensemble," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 10, pp. 5626–5640, 2022, doi: 10.1109/TNNLS.2021.3071122.

[34] N. Liu, X. Li, E. Qi, M. Xu, L. Li, and B. Gao, "A novel ensemble learning paradigm for medical diagnosis with imbalanced data," *IEEE Access*, vol. 8, pp. 171263–171280, 2020, doi: 10.1109/ACCESS.2020.3014362.

[35] V. López, A. Fernández, J. G. Moreno-Torres, and F. Herrera, "Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification. Open problems on intrinsic data characteristics," *Expert Syst. Appl.*, vol. 39, no. 7, pp. 6585–6608, Jun. 2012, doi: 10.1016/j.eswa.2011.12.043.

[36] F. Jia, S. Li, H. Zuo, and J. Shen, "Deep Neural Network Ensemble for the Intelligent Fault Diagnosis of Machines under Imbalanced Data," *IEEE Access*, vol. 8, pp. 120974–120982, 2020, doi: 10.1109/ACCESS.2020.3006895.

[37] X. Yuan, L. Xie, and M. Abouelenien, "A regularized ensemble framework of deep learning for cancer detection from multi-class, imbalanced training data," *Pattern Recognit.*, vol. 77, pp. 160–172, 2018, doi: 10.1016/j.patcog.2017.12.017.

[38] A. S. Qureshi and T. Roos, "Transfer Learning with Ensembles of Deep Neural Networks for Skin Cancer Detection in Imbalanced Data Sets," *Neural Process. Lett.*, vol. 55, no. 4, pp. 4461–4479, 2023, doi: 10.1007/s11063-022-11049-4.

[39] M. N. Hossin, M., Sulaiman, "A Review of Evaluation Metrics in Machine Learning Algorithms," *Med. Image Anal.*, vol. 80, no. 2, p. 102478, 2022, [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1361841522001256.

[40] M. H. Al-Adhaileh, "Diagnosis and classification of Alzheimer's disease by using a convolution neural network algorithm," *Soft Comput.*, vol. 26, no. 16, pp. 7751–7762, 2022, doi: 10.1007/s00500-022-06762-0.

[41] S. Murugan *et al.*, "DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia from MR Images," *IEEE Access*, vol. 9, pp. 90319–90329, 2021, doi: 10.1109/ACCESS.2021.3090474.

[42] M. M. S. Fareed *et al.*, "ADD-Net: An Effective Deep Learning Model for Early Detection of Alzheimer Disease in MRI Scans," *IEEE Access*, vol. 10, no. September, pp. 96930–96951, 2022, doi: 10.1109/ACCESS.2022.3204395.

[43] N. Goenka and S. Tiwari, "AlzVNet: A volumetric convolutional neural network for multiclass classification of Alzheimer's disease through multiple neuroimaging computational approaches," *Biomed. Signal Process. Control*, vol. 74, no. January, p. 103500, 2022, doi: 10.1016/j.bspc.2022.103500.

[44] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "An Intelligent System for Early Recognition of Alzheimer's Disease Using Neuroimaging," *Sensors*, vol. 22, no. 3, 2022, doi: 10.3390/s22030740.

[45] S. A. Kumar and S. Sasikala, "Enhanced Alzheimer's Disease Classification Using Multilayer Deep Convolutional Neural Network-Based Experimentations," *Iran. J. Sci. Technol. - Trans. Electr. Eng.*, vol. 47, no. 4, pp. 1595–1621, 2023, doi: 10.1007/s40998-023-00622-9.

Conflict of Interest

The authors have no conflict of interest to declare.

Author(s) Contributions

Nedim Muzoğlu: Conceptualization, Methodology, Writing – Original Draft, Software
Enver Akbacak: Methodology, review & editing

Acknowledgments

Not Applicable

Ethical Approval and Informed Consent

This study adheres to scientific and ethical principles, with all referenced works appropriately cited in the bibliography.

Availability of Data and Materials

This dataset is openly accessible via the provided link: https://www.kaggle.com/code/aashidutt3/alzheimer-classification-with-mri-images/input?select=OriginalDataset.

Plagiarism Statement

This article has been scanned by iThenticate ™.