Enhancing Alzheimer's Disease Diagnosis: Leveraging Convolutional Neural Networks for Feature Extraction and KNN Classification

Marwah Nafea Saeea [1*] 🟢, Arash Deldari [2] 🟢

[1] College of Physical Education and Sport Science Wasit University, Iraq
[2] Imam Reza International University, Mashhad, Iran

*Corresponding Author: Marwah Nafea Saeea

ABSTRACT: Detecting Alzheimer's disease (AD) accurately is crucial for early patient management, allowing for proactive steps to minimize irreversible brain damage. AD-related cognitive decline varies in severity, from mild cognitive impairment to mild, moderate, and severe dementia. Early AD detection serves various purposes, such as reducing healthcare costs, slowing brain degeneration, and enhancing treatment outcomes. Machine learning techniques, especially in MRI image analysis, have gained significance in AD diagnosis. In our research, we've employed a deep learning model with convolutional layers for extracting essential information from MRI scans. Additionally, we've utilized the gray wolf meta-heuristic algorithm to identify optimal features. These features are then used in a K-nearest neighbors (KNN) classifier to categorize AD into three distinct types (very mild, mild, and moderate) while distinguishing normal and healthy cases. Our approach achieved an impressive accuracy rate of 95.6% when tested.

Keywords: Alzheimer's diagnosis, MRI image analysis, gray wolf meta-heuristic algorithm, feature extraction, K-nearest neighbors (KNN) classifier.

1. INTRODUCTION

Alzheimer's disease (AD) represents a pressing global health challenge, with an increasing prevalence that poses a substantial burden on healthcare systems and societies worldwide [1]. Timely and accurate diagnosis of AD is pivotal for facilitating early intervention and optimizing patient care [2]. In recent years, the convergence of medical imaging and artificial intelligence has opened new avenues for improving the diagnostic accuracy and efficiency of AD. One promising approach involves the utilization of Convolutional Neural Networks (CNNs) for feature extraction from neuroimaging data, followed by K-nearest neighbors (KNN) classification to categorize patients based on their neuroimaging profiles.

This innovative approach represents a paradigm shift in the field of AD diagnosis, harnessing the power of deep learning to extract intricate patterns and structures from medical images, particularly magnetic resonance imaging (MRI) scans. By automating the feature extraction process, CNNs can unveil subtle abnormalities that might go unnoticed by human observers, thus potentially enhancing diagnostic accuracy [3]. Subsequently, the KNN classifier leverages these extracted features to categorize patients into different diagnostic categories based on the similarity of their neuroimaging profiles to those in the training dataset.

This paper reviews and synthesizes the current state of research in the domain of AD diagnosis, focusing on the integration of CNNs for feature extraction and KNN classification. We explore the foundational studies that have laid the groundwork for this approach and investigate the advancements that have propelled it into the forefront of AD diagnosis research. By analyzing the contributions, challenges, and future directions in this field, we aim to provide a comprehensive overview of the potential of leveraging CNNs and KNN in the enhancement of Alzheimer's disease diagnosis.

Developing a robust computer-assisted diagnostic system for interpreting MRI scans is crucial for early AD detection. Current deep learning systems use cortical surface data as input for convolutional neural networks (CNNs) in the raw MRI image-based classification of AD.

2. RELATED WORK

The application of deep learning techniques in Alzheimer's disease (AD) diagnosis has seen significant advancements. Researchers have employed convolutional neural networks (CNNs) for feature extraction and classification, aiming to improve the accuracy and efficiency of diagnosis [4]. CNNs have been widely used to automatically extract meaningful features from medical images, including magnetic resonance imaging (MRI) scans, which are crucial for AD diagnosis. These networks can capture intricate patterns and structures that are often imperceptible to the human eye [5]. KNN is a commonly used machine learning algorithm for classifying AD and related cognitive disorders. It leverages feature vectors extracted from medical images to classify patients into different diagnostic categories based on their similarity to neighboring data points [6]. The integration of multiple neuroimaging modalities, such as MRI and positron emission tomography (PET), has shown promise in improving the accuracy of AD diagnosis. Researchers have used CNNs to fuse information from these modalities for more robust classification [7]. Transfer learning techniques, where pre-trained CNN models are fine-tuned for AD diagnosis, have gained attention. This approach leverages the knowledge learned from large datasets in related tasks to enhance the performance of AD classifiers [8]. Ensemble methods, including combining multiple CNN models or other machine learning algorithms, have been explored to improve the robustness and accuracy of AD classification models [9]. Despite significant progress, challenges remain in AD diagnosis, including handling large and heterogeneous datasets, addressing class imbalance, and improving the interpretability of deep learning models. Future research aims to address these issues to further enhance AD diagnosis [10].

Our aim is to transform the diagnostic procedure and improve early detection to enhance patient outcomes significantly. Utilizing advanced deep learning methodologies, we strive to make noteworthy contributions to the field of Alzheimer's research and healthcare.

3. PROPOSED METHODOLOGY

Our primary objective is to improve the accuracy of Alzheimer's disease diagnosis using Convolutional Neural Networks (CNNs) applied to MRI images. Traditional diagnostic methods, such as medical history assessment, neuropsychological testing, and MRI scans, often suffer from limited efficiency due to their high sensitivity. In contrast, CNNs excel at identifying Alzheimer's disease-specific features by employing convolutional filters, resulting in enhanced diagnostic precision. However, after feature extraction, there can be challenges related to network under-fitting, particularly when dealing with voluminous data. To address this issue, we utilize the gray wolf optimization algorithm to select the most relevant features, mitigating redundancy and further enhancing the accuracy of Alzheimer's disease diagnosis.

Our methodology involves the following steps:
1. Retrieving pertinent information from the database.
2. Pre-processing the data to ensure quality and consistency.
3. Dividing the dataset into training and testing subsets.
4. Employing a neural network to extract informative features.

We leverage deep CNNs for robust feature extraction, thereby improving the accuracy of Alzheimer's disease diagnosis. Following the feature extraction process, the gray wolf optimization algorithm is utilized to pinpoint the most significant features, effectively reducing redundancy and elevating the accuracy of our diagnostic model. Ultimately, the K-nearest neighbors (KNN) algorithm is employed to provide the final diagnosis, with subsequent evaluation of the diagnostic results.

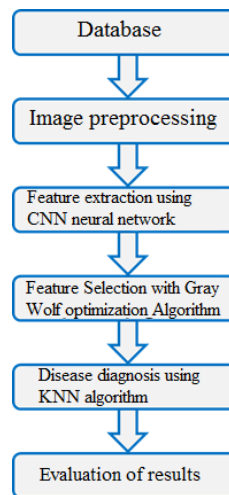For a visual representation of our proposed approach, please refer to Figure 1.

FIGURE 1. - Illustration of the Proposed Procedure

In our quest to diagnose Alzheimer's disease through the analysis of MRI images, we adhere to a structured three-step process:

1. Extraction of image features for the purpose of categorizing individuals into disease-afflicted and non-disease groups.
2. Streamlining the selection of these extracted features to enhance both accuracy and processing speed.
3. Application of a machine learning classification algorithm to effectively differentiate the identified features into patient and non-patient categories.

3.1 DATA PROCESSING

In this phase of our work, we initiate data preprocessing by resizing all images to a uniform size of 224x224 pixels. This standardization enables their compatibility as inputs for the CNN network.

3.2 FEATURE EXTRACTION USING CNN

The primary tool for extracting valuable features from our images is the deep Convolutional Neural Network (CNN). CNNs excel at capturing intricate image features through their convolutional layers, culminating in feature sets that serve as the foundation for subsequent classification tasks. However, in this particular study, our employment of CNN is solely focused on feature extraction from MRI images, while the actual classification is conducted through the K-Nearest Neighbors (KNN) algorithm.

The fundamental components of CNNs are the kernels or filters, which play a pivotal role in feature extraction from input data. These kernels facilitate the extraction of pertinent features from the input, accomplished through the convolution operation. After convolution, the resulting image features are discerned and highlighted.

Our proposed CNN architecture encompasses the following layers:

1. Convolutional Layer

2. Pooling Layer

3. Batch Normalization Layer

4. Dropout Layer

5. Fully Connected Layer

3.3 FEATURE SELECTION UTILIZING GRAY WOLF ALGORITHM

Subsequent to the extraction of features from images through the CNN network (with feature extraction occurring in the middle layers), we employ the meta-heuristic Gray Wolf Optimization (GWO) algorithm to identify and select the most valuable and influential features from the complete feature set. These selected features are subsequently employed in our analysis.

GWO is a cutting-edge optimization algorithm inspired by the predatory behavior of wolves in a well-organized pack. The algorithm shares similarities with actual wolf behavior, involving tactics of hunting and surveying the surroundings for potential prey. The hierarchical leadership within the GWO consists of four levels: alpha (α), beta (β), delta (δ), and omega (ω). Alpha (α) assumes decision-making responsibilities, with beta (β) providing support in the decision-making process. Delta (δ) represents guiding figures, possibly elder wolves or skilled hunters. Omega (ω) is the follower of the entire wolf pack and signifies the optimal solution. Alpha, beta, and delta represent the first, second, and third-best solutions, respectively. In our proposed approach, the decomposition of Algorithm 1 initializes a population of n wolves, where each wolf corresponds to an optimal solution, and n represents the number of features in the original dataset. As per the algorithm's logic, selected features that contribute to enhanced accuracy are represented with a binary value of 1, while those that do not significantly improve accuracy are assigned a binary value of 0.

3.4 K-NEAREST NEIGHBORS (KNN) CLASSIFIER

Ultimately, we employ the K-Nearest Neighbors (KNN) classifier to categorize the features chosen in the preceding step. Based on these selected features, the classifier distinguishes which images pertain to patients and which belong to healthy individuals.

Key aspects of the KNN algorithm include:

1. KNN assumes similarity between new data and existing data, placing new data into categories that closely resemble the existing ones.

2. The algorithm stores all existing data and classifies new data based on their similarity to the stored dataset. This allows new data to be easily categorized into well-defined clusters.

3. KNN is versatile and can be employed for both regression and classification, although it is primarily utilized for classification tasks.

4. It is categorized as a non-parametric algorithm, indicating that it makes no prior assumptions about the underlying data.

5. KNN is often referred to as a "lazy learning" algorithm as it retains the training dataset and performs operations on this dataset during classification. During training, KNN solely stores the dataset and classifies new data by finding the most similar data points in the stored dataset.

3.5 RESULTS EVALUATION

Upon receiving the results, we undertake network evaluation. Evaluation of the network's performance during training is conducted using the training data. Subsequently, the trained network is applied to the test dataset, and its outcomes are assessed against established criteria. This comprehensive evaluation process ultimately yields the final assessment of the network's performance.

4. RESULTS OF THE EXPERIMENT

In this work, the evaluation of the results obtained from the simulation of the proposed method is presented. All simulations were done by MATLAB software version 2021. In the rest of this chapter, first, the database used for the simulations is explained, then the evaluation criteria are reviewed, and finally, the numerical results of the simulations are presented.

4.1 DATASET DESCRIPTION

The dataset utilized for this study on Alzheimer's disease was sourced from Kaggle, an open-source platform. This dataset comprises 6400 MRI images categorized into four distinct classes: mild dementia (MID), moderate dementia (MOD), non-dementia (ND), and very mild dementia (VMD). The original image dimensions measured 176x208 pixels, which were subsequently adjusted to a uniform size of 176x176 pixels for consistency. To provide a visual representation of the four classes, sample images are presented in Figure 2.
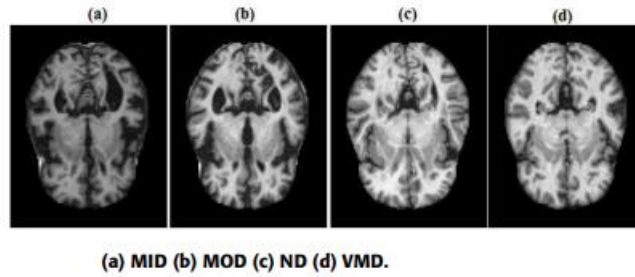
(a) MID (b) MOD (c) ND (d) VMD.

FIGURE 2. - Illustrations of four classes of dementia

4.2 EVALUATION CRITERIA

In this work, in order to evaluate the results of the proposed method, we will use the criteria of precision, accuracy, recall, and F-Score. The most important evaluation criterion in the work is prediction accuracy. In this work, the accuracy, i.e., how many percent of the images that belonged to the patient with Alzheimer's, was correctly recognized by our algorithm. The criteria are explained further. For a better understanding of these criteria, the following figure can be used.
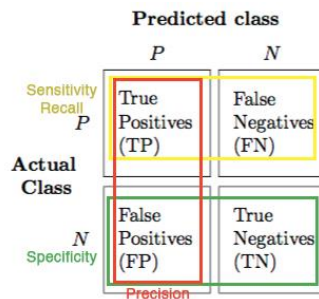

FIGURE 3. - Complexity matrix

1. Accuracy

Accuracy stands as a fundamental and readily comprehensible classification metric. Its simplicity makes it applicable to both binary and multi-class classification challenges. Accuracy measures the proportion of true results relative to the total number of cases examined. It serves as a valid evaluation metric, particularly well-suited for classification scenarios characterized by balanced distributions and the absence of skewness or class imbalance.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

2. Precision

Precision emerges as a reliable and appropriate evaluation metric when certainty in predictions is paramount. For instance, in scenarios such as developing a system to determine whether a specific account's credit limit should be reduced, ensuring a high level of confidence in predictions is crucial to prevent potential customer dissatisfaction.

$$Precision = TP(TP + FP)$$

3. Recall

Recall serves as a pertinent evaluation metric when the primary objective is to identify as many positive instances as feasible. For instance, in the context of developing a system for cancer prediction in individuals, the aim is to detect the disease even in situations where confidence levels may not be exceptionally high.

$$Recall = TP(TP + FN)$$

4. F1-Score

The F1-Score is calculated as the harmonic mean of Precision and Recall.

$$F1\ score = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

4.3 DATA PARTITIONING AND EVALUATION PROCESS

Within this section, we commence the process by furnishing the network with training data that includes specific class identifiers, facilitating the network's learning process based on these class distinctions. Ultimately, we proceed to evaluate the network's performance by presenting the training data without associated labels.

Figure (4) provides a visual representation of the confusion matrix involving four distinct classes, namely moderate, non-dementia, mild, and very mild dementia data. Remarkably, the proposed methodology attains a remarkable 100% accuracy rate when applied to the training dataset.

Furthermore, Figure (5) presents the numerical values corresponding to various evaluation metrics applied to the training dataset. These metrics offer a comprehensive insight into the network's performance on the training data.

4.3.1 OUTCOMES FROM TRAINING DATA

In our research, the dataset undergoes meticulous partitioning into two distinct segments: the training dataset and the testing dataset. A substantial 80% of the data is dedicated to training the proposed neural network, while the remaining 20% is earmarked for the comprehensive evaluation of our proposed methodology.

This critical stage encompasses several pivotal components, including the analysis of outcomes derived from both the training and testing datasets, their subsequent evaluation, and a comparative assessment.

To commence, we commence by presenting the simulation results associated with the training dataset. Following this, we proceed to evaluate the results pertaining to the test dataset, which the network has not previously encountered. Importantly, these test data are presented to the network without labels, and the results generated from this testing phase are regarded as the primary outcomes of this study. In the concluding phase, we undertake a meticulous comparison of the results obtained in our research with those presented in the reference article.

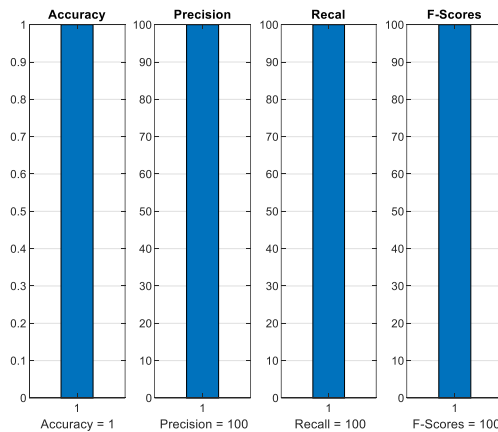**FIGURE 4. - Confusion matrix for training data**

FIGURE 5. - The numerical value of evaluation criteria for training data

61

Figures (6) and (7) show the regression and the ROC curve of the training, respectively.

The ROC (Receiver Operating Characteristics) curve, which is known as the performance characteristic curve, is one of the tools for identifying the performance of a system. By using the results of the ROC curve, concepts such as the cutoff point, sensitivity, and specificity of a test can be obtained.
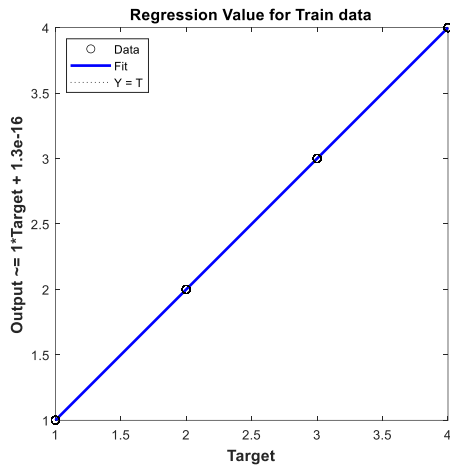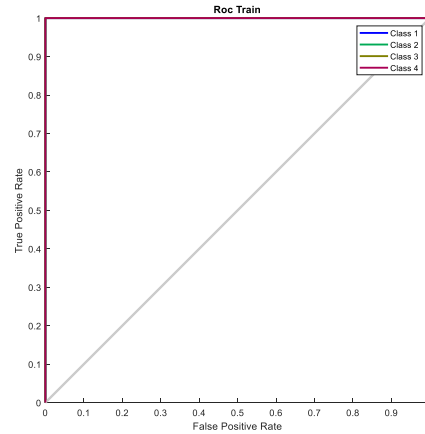

FIGURE 6. - Regression of training data

FIGURE 7. - ROC of training data

4.3.2   OUTCOMES FROM TEST DATA

Within this section, we compile and present the simulation results derived from the analysis of the test data. It's noteworthy that the key distinction between the training and test datasets lies in the fact that the test data represents a subset that the network has not encountered during prior assessments. Consequently, the final evaluation of the network's performance hinges upon this previously unseen information.

Figure (8) visually represents the confusion matrix associated with the test data for a single run. This matrix encompasses four distinct classes: moderate deterioration, non-dementia, mild, and very mild. Impressively, the network attains a final accuracy rate of 95.8% when applied to the test dataset.

Furthermore, Figure (9) presents numerical values corresponding to various evaluation criteria applied to the test data, offering a comprehensive evaluation of the network's performance. In addition, Figure (10) portrays the test ROC curve, further enhancing our understanding of the network's performance during the evaluation process.
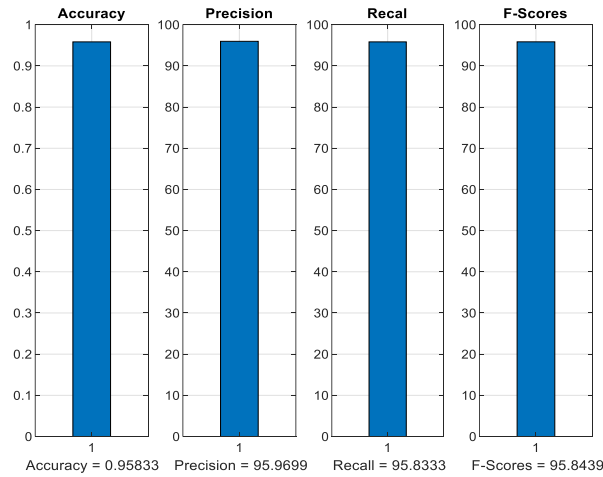

FIGURE 8. - Confusion matrix for test data

FIGURE 9. - The numerical value of evaluation criteria for test data
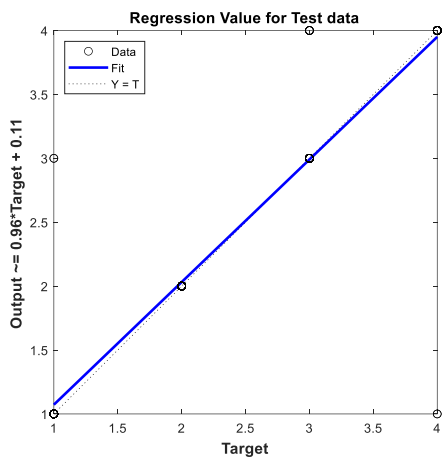

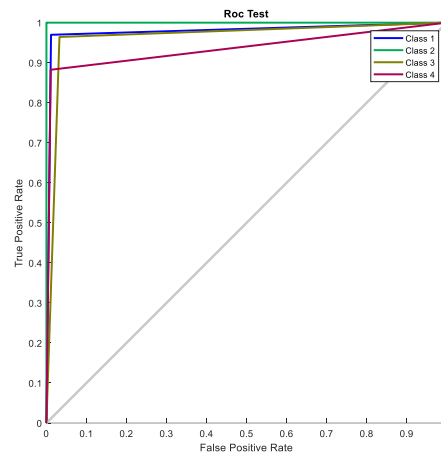**FIGURE 10. - Regression of test data**

FIGURE 11. - ROC of test data

4.4 COMPARING RESULTS

This specific section of the research is dedicated to conducting a thorough evaluation of the proposed methodology's performance. The study meticulously compares the outcomes achieved through the proposed approach with those of various other techniques cited in the preceding article [11]. This evaluation adheres to rigorous criteria to ensure an equitable and impartial assessment.

Central to this assessment is Table 1, which plays a pivotal role in presenting a concise summary of the comparative results. It is crucial to emphasize that the results featured in this table emanate from an exhaustive series of 50 simulation tests, each thoughtfully designed to rigorously scrutinize the capabilities of the proposed method. To bolster the credibility of these findings, the results are aggregated, offering a more dependable depiction of the method's overall performance.

One notable and compelling inference drawn from this comparative analysis is the pronounced superiority of the recommended method in terms of accuracy when juxtaposed with the alternatives discussed in the aforementioned article [11]. This signifies a notable advancement in the field, suggesting that the proposed approach possesses the potential to yield substantial enhancements within the investigated domain. These results stand as a testament to the efficacy of the method and hold promise for its application in future research endeavors

Table 1. - Comparative Analysis of the Proposed Method and the Referenced Article

No.	Method	F-Score	Recall	Precision	Accuracy
1	(DEMNET with SMOTE) [11]	95.27	95	96	95.23
2	(DEMNET Without SMOTE) [11]	83	88	80	85
3	Proposed method	96.1	96.01	96.23	95.6

5. CONCLUSION

Alzheimer's disease (AD) typically commences with a phase of mild cognitive impairment (MCI), followed by a gradual decline in cognitive function, behavioral alterations, loss of independence, and ultimately, fatality. The primary objective in AD treatment revolves around decelerating the progression during its early stages. The early detection of symptoms is of paramount importance in this endeavor.

AD entails a progressive deterioration of both brain structure and function, a process that can be tracked through non-invasive longitudinal magnetic resonance (MR) imaging. Longitudinal imaging proves to be particularly valuable in the early detection of AD, especially in cases of MCI where cognitive changes may not be profoundly disabling but still elevate the risk of developing dementia.

Our research is centered on the utilization of brain MRI to identify individuals with mild and moderate AD. Our methodology employs a combination of Convolutional Neural Network (CNN)-based feature extraction, feature selection guided by the gray wolf algorithm, and K-Nearest Neighbors (KNN) classification. Notably, our approach achieves a remarkable accuracy rate of 95.60%, representing an improvement over the reference models and reaffirming the significance of our research in advancing AD detection and early intervention efforts.

6. FUTURE WORK

In future research, several promising directions can be explored:
1. Hybrid Deep Learning (CNN-RNN): Investigating the integration of CNN-RNN architectures to extract both temporal and spatial features from data, enabling more robust analysis of spatiotemporal patterns in various applications.
2. Frequency Domain Feature Extraction: Leveraging Fourier or Wavelet transforms to extract valuable frequency domain information from images, offering enhanced insights for tasks like anomaly detection, signal processing, and pattern recognition.
3. Deep Learning for Classification: Continuously advancing deep learning techniques for classification tasks, including the development of innovative network architectures, optimization methods, and transfer learning approaches to improve accuracy and generalization across diverse datasets.
4. These directions hold significant potential for expanding the capabilities of deep learning across domains, fostering innovation, and addressing complex challenges.

REFERENCES

[1] G. Livingston, A. Sommerlad, V. Orgeta, S. G. Costafreda, J. Huntley, D. Ames, ... & N. Mukadam, "Dementia prevention, intervention, and care," The Lancet, vol. 390, no. 10113, pp. 2673-2734, 2017.

[2] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, ... & C. H. Phelps, "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 280-292, 2011.

[3] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," arXiv preprint arXiv:1502.02506, 2015.

[4] S. Sarraf and G. Tofighi, "DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks Using MRI and fMRI," BioRxiv, p. 070441, 2016.

[5] H. I. Suk, S. W. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," NeuroImage, vol. 101, pp. 569-582, 2017.

[6] J. H. Lee, S. H. Kim, Y. J. Kim, U. Yoon, and J. M. Lee, "Integration of Alzheimer's disease genetics and myelin content in schizophrenia and bipolar disorder," Journal of Alzheimer's Disease, vol. 48, no. 4, pp. 953-962, 2015.

[7] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," arXiv preprint arXiv:1502.02506, 2015.

[8] H. I. Suk and D. Shen, "Deep learning-based feature representation for AD/MCI classification," in International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 583-590, 2013.

[9] H. J. Park and J. J. Kim, "A convolutional neural network for Alzheimer's disease classification based on fusion of images and clinical markers," Journal of Neuroscience Methods, vol. 282, pp. 71-77, 2017.

[10] G. Livingston, A. Sommerlad, V. Orgeta, S. G. Costafreda, J. Huntley, D. Ames, ... & N. Mukadam, "Dementia prevention, intervention, and care," The Lancet, vol. 390, no. 10113, pp. 2673-2734, 2017.

[11] 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," Computers in Biology and Medicine, vol. 141, p. 105032, 2022.