7

Smart healthcare: fusion of multimodal medical signals

Nasir Rahim[1], Shaker El-Sappagh[1, 2, 3] and Tamer Abuhmed[1]

[1]COLLEGE OF COMPUTING AND INFORMATICS, SUNGKYUNKWAN UNIVERSITY, SUWON, SOUTH KOREA; [2]FACULTY OF COMPUTER SCIENCE AND ENGINEERING, GALALA UNIVERSITY, SUEZ, EGYPT; [3]INFORMATION SYSTEMS DEPARTMENT, FACULTY OF COMPUTERS AND ARTIFICIAL INTELLIGENCE, BENHA UNIVERSITY, BANHA, EGYPT

1. Introduction

Cells and entire organisms serve as prime examples of intricate systems, characterized by numerous interacting components that result in emergent behaviors (Walter et al., 2010). Understanding these interactions is essential, particularly for predicting complex diseases. Data modality is derived from capturing such phenomena with specific sensors (Duan et al., 2022), which alone provides limited information. On the other hand, multimodal data offers a broader perspective by revealing both individual components and their collective behaviors. With the swift advancements in high-throughput technologies, we now have exceptional access to comprehensive multimodal biomedical data, enabling us to make better use of this extensive information. Data fusion entails integrating information from various modalities, each offering unique perspectives on a shared phenomenon, to address inference challenges. This method is expected to deliver more accurate results compared to single-modality approaches (Lahat et al., 2015). The advantages of data fusion can be classified into complementary, redundant, and cooperative features (Duan et al., 2022; Lahat et al., 2015), which can overlap.

Cells and entire organisms exemplify complex systems composed of numerous interacting components that give rise to emergent behaviors (Walter et al., 2010). Gaining insights into these interactions is particularly crucial for predicting and understanding complex diseases. Data modality refers to the means of capturing such biological phenomena using specific sensors (Duan et al., 2022), though a single modality often provides only a limited perspective. In contrast, multimodal data integrates multiple sources of information, offering a more comprehensive understanding by capturing both individual elements and their collective dynamics. With rapid advancements in high-throughput technologies, the availability of extensive multimodal biomedical data has significantly increased, allowing for deeper and more precise biological analyses. Data fusion is the

Cybersecurity for Healthcare Systems in the Internet of Medical Things Era. https://doi.org/10.1016/B978-0-443-27448-0.00007-X

process of integrating information from diverse modalities, each contributing distinct insights into a common phenomenon, to improve inference and predictive accuracy. This approach has demonstrated superior performance compared to single-modality methods (Lahat et al., 2015). The benefits of data fusion can be categorized into three types, i.e., complementary, redundant, and cooperative features. Each of these categories overlaps to enhance analytical robustness and reliability (Duan et al., 2022; Lahat et al., 2015).

In biomedical research, data fusion plays a critical role in multimodal data analyses. For instance, genomic profiling of a tumor enables the identification of cancer-associated genes, while whole-slide imaging (WSI) of a biopsy provides detailed morphological and microenvironmental characteristics of the tumor. These modalities complement each other by offering distinct but interconnected insights into the disease, capturing aspects that would be indiscernible through a single modality alone. The integration of transcriptomic and proteomic data serves both complementary and redundant purposes. Since not all messenger RNAs (mRNAs) are translated into proteins, transcriptomic data alone may not fully capture functional molecular activity. However, proteomic data can confirm translation events, thereby providing redundancy that is particularly beneficial in cases where data is noisy or incomplete. Similarly, combining microRNA (miRNA) and mRNA sequencing data from the same tumor demonstrates cooperative fusion, as it enhances the biological context by explaining regulatory interactions that influence protein expression. This integration is especially valuable for understanding variations in the data, which can be critical for predicting a patient's response to treatment. The primary goal of fusion strategies is to effectively leverage the complementary, redundant, and cooperative attributes of different modalities to enhance predictive accuracy and biological interpretation. Achieving this requires the application of machine learning (ML) techniques capable of integrating both structured and unstructured data, each characterized by distinct statistical properties, sources of nonbiological variation, dimensionality differences (Mishra & Misra, 2018), and patterns of missing values (Marvasti et al., 2013). By employing advanced ML-driven fusion methodologies, researchers can extract deeper insights, improve the robustness of predictive models, and facilitate more precise clinical decision-making.

In recent years, the study and application of multimodal ML techniques have gained significant attention across various domains (Fard et al., 2021; Lahat et al., 2015). Among these techniques, multimodal deep learning (DL) presents notable advantages for data fusion compared to traditional shallow learning methods. Fully connected neural networks (FCNNs), a fundamental type of deep neural network (DNN), can be represented as a directed acyclic graph, where an input x is transformed into an output label y through multiple hidden layers, each performing nonlinear computations (Rahim et al., 2018). The primary objective of these algorithms is to learn high-level feature representations from the input data, thereby improving the predictive accuracy of the final classifier by capturing complex dependencies among the underlying, disentangled factors. The initial layers of these networks extract fundamental abstractions, while deeper layers progressively combine these abstractions into more intricate representations that

are essential for the learning task (Pachynski et al., 2021). Notably, multimodal DL models are capable of capturing both nonlinear relationships within individual modalities and complex interactions across multiple modalities. This ability has facilitated their widespread adoption across a range of applications, further demonstrating their effectiveness in enhancing data-driven decision-making (Lahat et al., 2015).

2. Alzheimer's disease as multimodal medical case study

Alzheimer's disease (AD), widely known as senile dementia, poses a significant challenge in finding a cure due to its complex pathogenesis. This degenerative neurological disorder gradually erodes cognitive function and memory, primarily attributed to the formation of neurofibrillary tangles (NFT) and the accumulation of extracellular amyloid-β (Aβ) plaques (Abrol et al., 2020; Theofilas et al., 2018). These pathological hallmarks lead to the loss or damage of neurons and synapses, contributing to the progressive decline in cognitive abilities. AD ranks as the fourth leading cause of global mortality, following cardiovascular disease, cancer, and stroke, emphasizing the urgent need for effective interventions. It not only imposes a substantial financial burden but has also surpassed cancer as the most feared affliction, claiming more lives than the combined burden of breast and prostate cancer. The devastating consequences of AD are characterized by a gradual loss of bodily functions, ultimately resulting in death. In 2018, an estimated minimum of 50 million individuals worldwide were affected by AD, highlighting the scale of its impact. The prevalence of AD is particularly high in the elderly population, with 4%–8% of adults aged 65 and above experiencing the disease. Beyond the age of 85, the risk escalates significantly to 35% (Wang et al., 2018). These statistics highlight the crucial need for early detection and treatment strategies. Mild cognitive impairment (MCI), often considered a precursor stage of AD, serves as a transitional phase between normal aging and the onset of AD. Unfortunately, the symptoms of MCI are frequently misdiagnosed as typical signs of aging, leading to delays in appropriate care. However, it is crucial to note that approximately 44% of individuals with MCI may eventually progress to AD within a few years (Wang et al., 2018). Addressing MCI and its progression to AD requires a comprehensive approach. Medications targeting the underlying neurodegenerative processes, such as cholinesterase inhibitors and N-methyl-D-aspartate (NMDA) receptor antagonists, have shown promise in slowing down cognitive decline and improving the quality of life for affected individuals. Given the profound impact of AD on individuals and societies worldwide, research on AD remains at the forefront of medical research. Scientists are actively working to understand the intricate mechanisms underlying AD's pathogenesis, identify novel therapeutic targets, and develop interventions to prevent or halt the progression of this devastating disease. The ongoing pursuit of knowledge and innovative solutions hold promise for improving the lives of those affected by AD and bringing the research community closer to finding a cure.

The field of AD research has witnessed the application of various ML methods to differentiate between different stages of AD (Bazargani et al., 2024; Zhang & Shen, 2012).

However, due to the complex and progressive nature of AD, it is necessary to integrate multiple modalities to obtain a comprehensive characterization of patients. Traditional approaches for AD diagnosis and progression detection have primarily focused on using single-modality data, such as structural MRI scans or positron emission tomography (PET) imaging, to classify individuals into AD or cognitively normal categories (Qiu et al., 2018). However, the complex nature of AD calls for a more comprehensive understanding of the disease by integrating multiple modalities, including not only neuroimaging data but also functional imaging, genetic markers, cerebrospinal fluid biomarkers, and clinical assessments (Fan et al., 2008). By leveraging ML techniques, such as deep neural networks and ensemble methods, and incorporating multimodal data, researchers aim to improve the accuracy and reliability of AD diagnosis and prediction, enabling early detection and personalized treatment strategies tailored to individual patients (Zhang & Shen, 2012). Furthermore, the utilization of time-series multimodal data holds great promise in enhancing the classification and prediction of AD progression. AD is a progressive disease that evolves over time, and capturing its dynamic nature is essential for developing effective monitoring and intervention strategies. Integrating longitudinal data, such as sequential neuroimaging volumes (e.g., MRI or PET scans taken at different time points), cognitive scores obtained from neuropsychological tests, and demographic information, with advanced ML models can potentially unlock new avenues for understanding AD progression and developing targeted interventions (Ho et al., 2022). Longitudinal analysis allows for the identification of subtle changes and patterns that occur during different stages of the disease, improving the accuracy of predicting future cognitive decline and enabling early interventions.

This study incorporates multimodal longitudinal data, such as sequential neuroimaging volumes, cognitive scores obtained at different time points, and demographic information collected over time. This integration of time-series data will enable the modeling of disease progression trajectories and the identification of temporal patterns that are indicative of disease progression (Chincarini et al., 2016). By capturing the dynamic changes in brain structure and function over time, the framework can provide a more accurate representation of AD progression dynamics and enable more precise predictions. Therefore, the main goal in this chapter is to propose a novel DL-based framework that leverages multimodal time-series data to accurately detect and predict the progression of AD. The framework aims to integrate neuroimaging modality (MRI), cognitive scores, and demographic information to capture the multifaceted nature of the disease and provide a comprehensive representation of AD progression. The findings of this research have the potential to significantly advance our understanding of AD progression and support clinical decision-making for early intervention and personalized treatment strategies.

The following key points summarize the primary contributions of this chapter.

- This study introduces a novel deep ensemble learning framework that effectively fuses multimodal longitudinal medical signals, specifically sequential MRI volumes and cognitive scores, for precise AD progression detection. The proposed approach

enhances disease characterization by integrating both structural and functional biomarkers.
- A lightweight convolutional autoencoder (CAE) is developed to extract high-dimensional feature representations from 2D MRI slices, reducing spatial redundancy. The study further employs principal component analysis (PCA) to refine these extracted features, optimizing storage and computational efficiency while preserving critical disease-relevant information.
- The study employs a heterogeneous ensemble of deep learning models, including 1D convolutional neural networks (1D CNN), long short-term memory (LSTM), and gated recurrent units (GRU). These models are processed through Bayesian hyperparameter optimization to fine-tune network configurations, ensuring optimal performance in capturing complex temporal dependencies within AD progression.
- The study integrates Bayesian-optimized time-series models into a heterogeneous ensemble framework, leveraging a weighted averaging strategy to aggregate class-wise probabilities across multiple models. This ensemble approach significantly improves the stability and generalizability of AD progression detection, outperforming individual deep learning models.
- The proposed framework is comprehensively validated using longitudinal data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Performance is benchmarked against single-modality and multimodal approaches, demonstrating superior accuracy, area under the curve (AUC), and F1-score, thus confirming the efficacy of the proposed method in clinical applications.

The rest of the chapter is organized as follows. Section 3 reviews existing literature on AD detection and progression, highlighting the strengths and limitations of existing ML techniques, and setting the context for the chapter's contributions. Section 4 details the building blocks of the proposed framework, including autoencoders, neural networks, and optimization techniques, laying the foundation for the proposed model in Section 5, which introduces an optimized ensemble framework for AD progression detection. Section 6 covers the dataset and model configuration, followed by experiments and results in Section 7, which validates the proposed methods, demonstrating their effectiveness in advancing AD research. Section 8 discusses the proposed method and provides a comparative analysis of state-of-the-art (SOTA) methods in AD progression detection. The conclusion of the chapter, with the current limitations and future work are discussed in Section 9.

3. Related work

AD is a serious type of cognitive decline, and specialists evaluate a patient's mental health through multiple modalities (Zhang et al., 2021). For instance, Ortiz et al. (2016) introduced a deep belief network and applied it to 3D gray matter patches. These patches were presegmented from 3D MRI volumes using the automated anatomical labeling (AAL) atlas. Their approach significantly improved the accuracy of

distinguishing between MCI and AD, as indicated by higher AUC values. However, a major drawback is the substantial computing power needed to run their method. Nanni et al. (2019) proposed an SVM-based hybrid ML approach to classify cognitively normal (CN) from AD patients. The authors focused on addressing the issue of reducing the high dimensionality of data while working with 3D MRI. They combined an SVM trained on texture descriptors with an SVM trained on voxel-based biomarkers. Although their approach reported high classification accuracy, a notable drawback is that they did not compare traditional ML approaches with deep neural network-based approaches. Feng et al. (2020) developed an AD classification method based on wavelet transform energy features (WTEF), which leverages the energy distribution in structural MRI (sMRI) signals. Their approach aimed to address the limitations of conventional spatial analysis techniques. However, a limitation of their model was that it processes redundant information due to the wavelet transformation's lack of down sampling nature, which can affect computational efficiency. Table 7.1 list studies in the AD domain based on MRI modality.

3.1 AD progression detection using multimodal data

Many studies mentioned in the earlier section were mainly focused on diagnosing AD using a single modality of data. However, research in the ML community has demonstrated that combining multiple modalities in a diagnostic process improves the accuracy of AD diagnosis significantly. Numerous research efforts have concentrated on classifying AD using multimodal data from various data sources, such as the ADNI or National Alzheimer's Coordinating Center (NACC) datasets. These approaches leverage the distinct and complementary information from various modalities to enhance classification accuracy, employing various feature fusion techniques (Dai et al., 2023). One common approach is to merge features from different modalities into a single cohesive representation. For instance, Jun et al.(Shi et al., 2018) proposed MM-SDPN to enhance

Table 7.1 MRI-based AD prediction.

Study	DM	# Of subjects	Framework	(Acc.%)
Wang et al. (2017)	MRI	CN = 400, MCI,	Binary classification (MCI vs. NC)	91
Wang et al. (2019)	MRI	MCI = 297, NC = 315	Binary classification (MCI vs. NC)	98
Pan et al. (2020)	MRI	NC = 162, MCI = 76	Binary classification (MCI vs. NC)	79
Mehmood et al. (2021)	MRI	NC = 85, EMCI = 70, LMC = 170	Multiclass classification (NC vs. EMCI vs. LMCI)	89
Yu et al. (2022)	MRI	CN = 826, AD = 422	Binary classification (CN vs. AD)	95
Sarraf et al. (2023)	MRI	CN = 91, MCI = 744, AD = 211	Muliclass classification (CN vs. MCI vs. AD)	97
Ours (2023)	MRI	CN = 282, Converted to AD = 282	Progression detection (CN vs. Converted)	86

DM: Data Modality.

the diagnosis of AD by integrating neuroimaging data from multiple sources, such as MRI and PET. The proposed approach consisted of a two-stage deep polynomial network, where the first stage extracts high-level features from each imaging modality, and the second stage fuses these features for improved classification. Experimental results confirmed that MM-SDPN effectively captures complementary information from different imaging modalities, improving the accuracy and reliability of AD diagnosis. Tong et al. (Qiu et al., 2020) proposed an interpretable DL model for AD classification using multimodal data, including MRI scans, age, gender, and cognitive assessment scores. By integrating a fully convolutional network with a multilayer perceptron, their proposed model generated high-resolution disease probability maps for enhanced diagnostic precision. Liu et al. (2015) employed a stacked autoencoder framework with a zero-masking technique to fuse MRI and PET data, aiming to extract complementary insights from both imaging sources. Table 7.2 lists studies that utilize multiple modalities for AD detection.

It is obvious that the aforementioned studies highlight the importance of multimodal data in disease diagnosis. However, AD being a neurological disorder, is primarily neurodegenerative in nature. These diseases progressively damage brain tissues, making it challenging to capture degenerative patterns from the baseline data alone. Additionally, overlooking the time-series aspect of these modalities misses crucial information about disease progression (Pachynski et al., 2021; Qiu et al., 2018).

3.2 Early prediction of AD's progression using longitudinal medical data

Medical data analysis over time is an essential step for diagnosing neurodegenerative diseases, as it uncovers patterns of neurological changes as they progress. Time-series

Table 7.2 AD progression detection using multimodal data.

Study	Mod.	Dataset (# of subjects)	Framework	Acc. (%)
Sivapriya et al. (2015)	NSB, FDG, PET	ADMI (750)	SVM, RF, MLP	(CN vs. AD: 93.75)
Albright (2019)	D, CSs	ADNI (1737)	MLP	(CN vs. Progress to AD: 96)
Hong et al. (2019)	MRI, PET, DTI	ADNI (1105)	LSTM	(Progression detection of AD:94)
Moore et al. (2019)	MRI, CS, CSF	ADNI (1737)	Random Forest	(MCI vs. AD: 73)
An et al. (2020)	MH, HIS, CVD, UPDRS, NPIQ, GDS, FAQ	NACC: (23,165)	Sparse Autoencoder	(NC vs. AD: 86)
Liu et al. (2021)	MRI, D, MMSE	ADNI: (2228)	Graph CNN	(AD progression detection: 83)
El-Sappagh et al. (2022)	CSs, NSB, Static	ADNI: (1371)	XG-Boost	(AD progression: 99)
Eslami et al. (2023)	MRI, PET, D, CS	ADNI, (610)	ML4VisAD	(CN, MCI, AD: 82)
Our (Rahim, El-Sappagh et al., 2023)	MRI, CS, D	ADNI (564)	3DCNN-BiRNN	(CN vs. Progressed to AD: 96)

Mod.: Data Modality.

data provide valuable insights into disease development, enabling earlier detection and more effective monitoring (Rahim, Abuhmed, et al., 2023). However, despite its promise, time-series medical data analysis over multiple timesteps remained very rare in ML applications for neurodegenerative disease research, highlighting the need for further investigation in this area. For instance, Chincarini et al. (2016) explored the use of longitudinal hippocampal volume measurements from MRI scans to enhance early detection of AD. By segmenting and analyzing hippocampal volume changes over time, the study introduces a refined approach that reduces variability and improves disease classification accuracy. Their proposed approach was evaluated on the ADNI dataset and reported significant improvement in the classification accuracy. The findings also highlighted the value of longitudinal imaging in identifying early neurodegenerative changes, offering a more reliable biomarker for AD's progression. In another study, Moradi et al. (2015) presented ML based framework for predicting AD conversion in MCI patients using MRI-based biomarkers. The approach integrates semi-supervised learning, feature selection, and the removal of age-related confounds to improve classification accuracy. By combining MRI data with cognitive assessments and age, the model achieves a high AUC score, demonstrating its effectiveness in early AD diagnosis. The findings emphasized the potential of integrating neuroimaging and ML for improved disease progression prediction. Table 7.3 provides an overview of studies focusing on AD progression analysis using longitudinal data.

Table 7.3 AD progression detection using longitudinal multimodal data.

Study	Mod.	# of TS	Dataset (# of subjects)	Framework	Acc. (%)
Hong et al. (2019)	MRI, PET, DTI	10	ADNI: (1105)	LSTM	(Progression detection of AD: 94)
Albright (2019)	D, CSs	3	ADNI: (1737)	MLP	(CN vs. Progress to AD: 96)
Zhu et al. (2021)	MRI	5	ADNI: (151)	Temporal SVM	(MCI converted to AD)
Abuhmed et al. (2021)	MRI, PET, CS, Comorbidities	4	ADNI: (1371)	BiLSTM	(AD progression detection: 84)
El-Sappagh et al. (2021)	Comorbidities, CS, brain disorder	4	ADNI: (1029)	BiLSTM	(AD progression detection: 99)
Liu et al. (2021)	MRI, D, MMSE	1	ADNI: (2228)	Graph CNN	(AD progression detection: 83)
El-Sappagh et al. (2022)	CSs, NSB, Static	4	ADNI: (1371)	XG-Boost	(AD progression: 99)
Eslami et al. (2023)	MRI, PET, D, CS	4	ADNI: (610)	ML4VisAD	(CN, MCI, AD: 82)
Our (Rahim, El-Sappagh et al., 2023)	MRI, CS, D	3	ADNI:(564)	3DCNN-BiRNN	(CN vs. Progressed to AD:96)

Modality, TS: Number of time steps

4. Materials and methods

We proposed a hybrid DL framework for detecting AD progression using longitudinal MRI and cognitive assessments. The framework starts with a lightweight CAE to extract high-dimensional features from 2D MRI slices. These features are then processed using PCA to reduce their dimensionality while retaining critical information. The optimized features are then used to fine-tune a set of time-series models through Bayesian optimization (BO) for hyperparameter tuning. Finally, various ensemble machine learning (EL) configurations are evaluated on the refined features to assess their ability to accurately predict AD progression, ensuring robustness and generalizability across diverse patient profiles. A visual representation of the proposed framework is provided in Fig. 7.1.

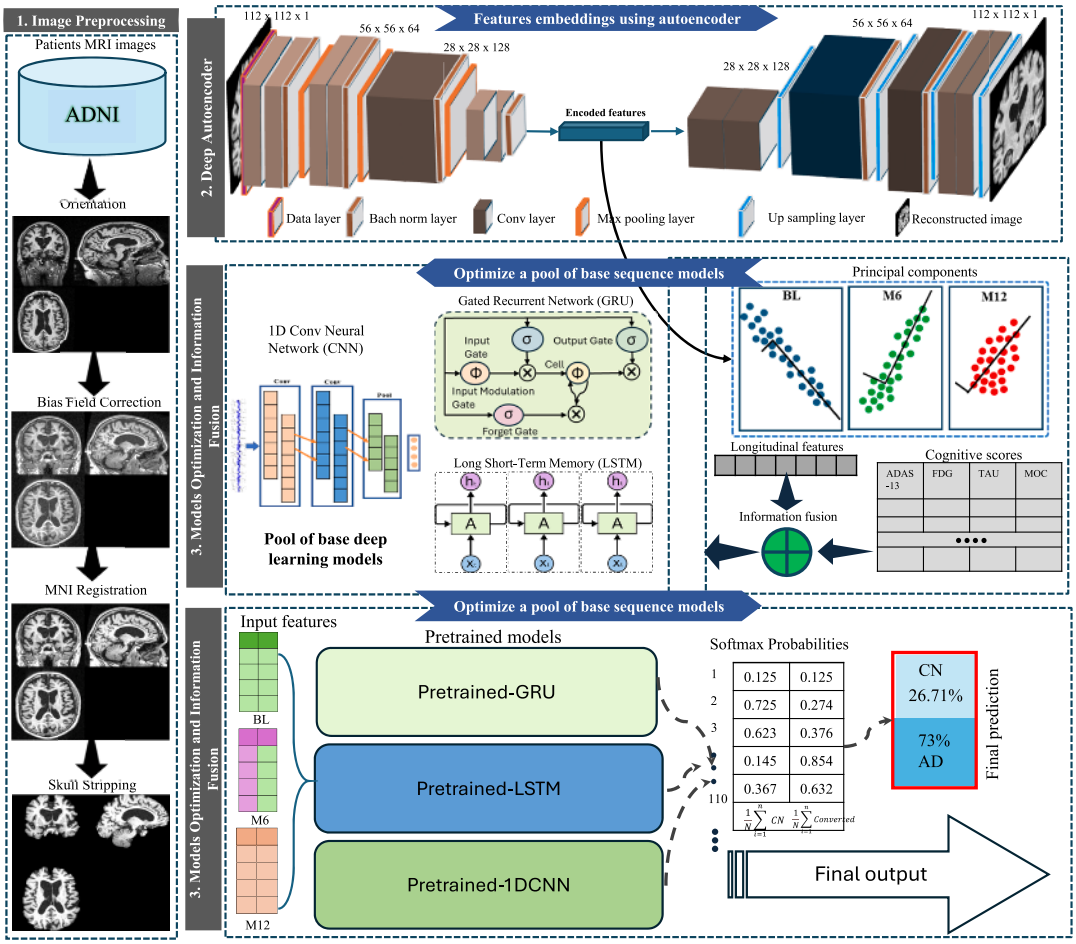

FIGURE 7.1 Proposed framework for AD progression detection.

4.1 Latent representations of deep features

An autoencoder (AE) is a type of neural network designed to learn efficient, compressed representations of input data in an unsupervised manner. It consists of two main components: the encoder and the decoder. The encoder compresses the input into a lower-dimensional latent space, while the decoder reconstructs the original input from this compressed representation. A convolutional autoencoder (CAE) is a variant of the standard autoencoder, where both the encoder and the decoder use convolutional layers instead of fully connected layers. This makes CAEs particularly effective for image-based tasks, as convolutional layers are capable of capturing spatial hierarchies in the data. In the encoder, convolution operations are applied to the input image, progressively reducing its spatial dimensions while increasing its depth, which results in a compressed, lower-dimensional representation. The decoder then employs transposed convolution to upscale the compressed representation back to the original image size. The training process optimizes these convolutional operations to minimize reconstruction loss, enabling the network to learn a more compact and meaningful representation of the image data.

4.2 Dimensionality reduction of deep features

Dimensionality reduction is critical for optimizing computational efficiency and reducing storage needs, especially when dealing with high-dimensional data. PCA is an effective method for simplifying data complexity while retaining essential information with high accuracy. PCA works by transforming the original high-dimensional space into a lower-dimensional subspace, ensuring that the most informative feature vectors are preserved (Jollife & Cadima, 2016). In 3D MRI analysis, where 2D slices are commonly used for feature extraction in CNNs, many features from inner layers show strong correlations. These interdependent features can cause redundancy and inefficiency. To mitigate this, PCA is applied to project feature vectors on an orthogonal basis, where each principal component captures the maximum variance, ensuring the extracted features are independent. By eliminating redundant or correlated information, PCA improves both accuracy and computational speed, creating a set of uncorrelated principal components that retain only the most discriminative information for further processing. Consequently, PCA is essential for enhancing feature representations, reducing redundancy, and improving the efficiency of CNN-based neuroimaging analysis.

4.3 Sequential features analysis using time-series models

1D convolutional neural networks (1D CNNs), LSTM networks, and GRUs are widely used DL models designed for sequential data analysis, each with distinct features tailored to specific tasks.

1D CNNs excel at processing time-series data, signal processing, and text classification by applying convolutional filters along a single dimension to capture local patterns in sequential inputs. Pooling layers reduce dimensionality while retaining key features,

and nonlinear activation functions enable the network to learn complex relationships. These models are computationally efficient and suitable for large-scale sequential data tasks.

LSTMs are an advanced variant of recurrent neural networks (RNNs) that address the challenges of vanishing and exploding gradients, enabling the learning of long-term dependencies. LSTM units use gates (forget, input, and output gates) to regulate information flow, allowing them to maintain and update cell states effectively. This makes LSTMs ideal for applications such as natural language processing (NLP) and speech recognition.

GRUs are a simpler alternative to LSTMs, which use two gates (update gates and reset gates) to manage the flow of information. These gates balance retaining past information and incorporating new data, resulting in faster training and lower computational costs. Despite their differences, 1D CNNs, LSTMs, and GRUs are all crucial for extracting patterns from sequential data and enabling accurate predictive modeling in various domains.

4.4 Hyperparameter optimization of time-series models

The main objective of hyperparameter tuning is to optimize a time-series model by selecting the best hyperparameter configuration to enhance test set performance.

This optimization is represented by the following equation:

$$x_{opt} = argmax_{x \in X} f(x) \quad (1)$$

Here, $f(x)$ is the performance function, x represents the hyperparameter settings, and x_{opt} is the optimal set. The elements of x_{opt} can be continuous, integer-based, or categorical. The objective function $f(x)$, also referred to as a Gaussian Process, is what the model tries to minimize on a validation dataset. X denotes the hyperparameter search space. BO based on Bayes' theorem, is a method that uses prior knowledge to guide the search for optimal hyperparameters. Bayes' theorem is expressed as:

$$P(X|Z) \propto P(Z|X)P(X) \quad (2)$$

The posterior probability $P(X|Z)$ is obtained by multiplying the likelihood $P(Z|X)$ with the prior $P(X)$, indicating that $P(X)$ represents the initial assumptions. The commonly used hyperparameter tuning methods include grid and random search, but these approaches do not incorporate learning from the tuning process. To address this, BO is used, where a probability model is built to translate hyperparameter values into scores. Those hyperparameter values that receive the top scores are retained for further analysis.

4.5 Model aggregation using ensemble machine learning

EL enhances prediction accuracy by aggregating the outputs of multiple decision-making models (Guo et al., 2022; Muhammed & Thiyagarajan, 2021; Nanni et al., 2016). The primary techniques in EL include bagging, boosting, and stacking, each of

which contributes to improved generalization performance. In EL, models can be either heterogeneous or homogeneous. The heterogeneous approach entails employing multiple distinct algorithms on a single training dataset, whereas the homogeneous approach utilizes a single algorithm across multiple training datasets. The outputs from these models are combined using methods such as averaging, voting, or weighted learning, depending on the specific task. For regression tasks, predictions from different models are typically averaged or weighted to form the final output. Voting can be absolute, where the outcome is based on a majority agreement among models, or relative, where the most frequently predicted class is chosen as the final result. In classification tasks, voting is used to determine the final class label.

5. Proposed optimized ensemble model

The proposed EL framework was designed to optimize feature selection for precise detection of AD progression. To ensure high-quality input data, all MRI volumes underwent a standardized preprocessing pipeline. The preprocessed dataset was then partitioned into training (80%) and testing (20%) splits to facilitate a thorough evaluation. To enhance diagnostic accuracy, the most crucial middle slices were extracted from each 3D MRI volume, specifically from the coronal plane, as this region prominently captures AD-related neurodegenerative changes, particularly in the hippocampus, which is known to be a critical structure associated with memory and cognitive function. The training dataset comprised 451 patients, totaling 1353 MRI volumes, with each patient contributing scans at three distinct time points. The test dataset included 113 patients and 339 MRI volumes. To extract meaningful feature representations, multiple CAEs were trained on the 2D MRI slices. The high-dimensional feature maps generated by these CAEs were further refined using PCA to eliminate redundancy and improve computational efficiency. The resulting feature vectors, spanning all three-time steps, were then integrated with cognitive scores (CSs) through a late fusion approach to enhance predictive capability. Subsequently, these fused features were used to optimize the hyperparameters of various time-series models. In addition, the study explored different ensemble configurations by combining feature embeddings derived from multiple CAEs. In addition, a comparative analysis of homogeneous and heterogeneous ensemble networks was conducted to determine the most effective configuration for improving model performance and generalizability in AD progression detection. The following sections will detail each step of the proposed AD progression framework.

5.1 Preprocessing of 3D MRI

Image preprocessing involves eliminating irrelevant information from raw MRI volumes, which facilitates accurate comparison of brain scans of varying sizes by experts and researchers. In this study, we applied a predefined preprocessing pipeline to the raw MRI volumes. After processing, domain experts visually inspected the volumes to ensure their

quality and accuracy. This preprocessing approach led to significant improvements in model performance compared to using nonprocessed MRI data. In addition, the computational process was optimized since the data retained only disease-related information, excluding nonessential regions such as the skull and neck. The preprocessing steps we performed included correcting the bias field of the MRI volumes, separating the skull from the brain tissue, and registering the MRI slices to a standard template. Bias fields in MRI scanners are variations in image brightness that are caused by gradual, low-frequency signal changes within the scanner, leading to uneven brightness across the MRI images. These types of noises can negatively affect the overall quality of the MRI volume, particularly during the image reconstruction phase. One common cause of this issue is the use of outdated scanner technology. In this study, we employed the N4 bias field correction algorithm (Avants et al., no date) to address the bias field in our MRI volumes. N4 is an upgraded version of the N3 bias field correction algorithm, which utilizes a more advanced B-spline fitting algorithm in the correction process. After correcting the bias field, brain tissue can be separated from the nonbrain tissue. Nonbrain tissue refers to the extrameningeal layers surrounding the brain, such as the dura mater, arachnoid, and pia mater. In this study, we utilized the brain extraction tool (BET2) (Smith, 2002) to perform skull stripping on MRI volumes. Finally, all slices of the MRI volume were registered on the Montreal Neurological Institute (MNI) 152 template (Brett et al., 2002). MNI 152 template registration involves aligning an individual's brain imaging data with a standardized reference brain template to allow for comparison across subjects and studies by transforming the imaging data into a common coordinate system. We used the FLIRT tool from FSL to perform template registration.

5.2 Optimized deep autoencoder

Our proposed convolutional autoencoder (CAE) is designed to extract detailed, high-level representational features from an input image. The architecture of the proposed CAE is optimized for processing grayscale MRI slices with dimensions $112 \times 112 \times 1$. The encoder module in the proposed CAE is composed of convolutional layers, batch normalization layers, and pooling layers, forming the primary building blocks of the network. Specifically, the encoder is organized into three sequential blocks, with each block consisting of two consecutive convolutional layers followed by batch normalization and max pooling layers. The convolutional layers employ 3×3 kernels to capture spatial features at various scales, and each convolutional operation is followed by a batch normalization layer to ensure stability during training. The max pooling layers, which follow the convolutional and normalization layers, reduce the spatial dimensions by a factor of two at each block while expanding the channel dimension, thus enabling the model to capture more abstract features at deeper levels. At the final block of the encoder module, the feature maps have reduced spatial dimensions of size $14 \times 14 \times 128$. This compact latent representation encapsulates the most important characteristics of the original input image, while maintaining a high level of feature abstraction.

The decoder module mirrors the encoder's design, utilizing three consecutive blocks of convolution, batch normalization, and upsampling layers to reconstruct the original image from its latent representation. The upsampling layers increase the spatial dimensions of the feature maps, which are progressively refined by convolutional layers. This process gradually reduces the channel dimension of the feature maps, allowing for a detailed reconstruction of the input image. The decoder uses transposed convolution operations for upsampling, which helps in learning the precise spatial relationships and generating sharper, more accurate reconstructions. Throughout the encoder-decoder modules, the ReLU (Rectified Linear Unit) was used as activation function. Additionally, the binary cross-entropy loss function was used during training to quantify the pixel-wise differences between the reconstructed and original images, ensuring that the network learns to minimize reconstruction errors effectively.

The model was trained for 150 epochs using the Adam optimizer. The training process utilized a learning rate of 0.0001 and a batch size of 16, optimizing the network to balance convergence speed and accuracy. This architecture was finalized after a thorough process of experimentation and hyperparameter tuning, which involved balancing model complexity with performance to achieve optimal reconstruction results. The architectural design of the encoder-decoder modules, including layer configurations and detailed specifications, is presented in Table 7.4.

5.3 Enhanced features compression of the deep features

To enhance the quality and representativeness of the deep features extracted by the proposed CAE, we employed PCA to reduce the dimensionality of these features. This step was necessary because the kernels in the deeper layers of the CNN model often capture similar features across multiple feature maps, which contribute minimally to the model's accuracy but increase computational overhead. By applying PCA to the

Table 7.4 Technical details for the proposed CAE.

	[Size of kernel, number of kernels, stride, padding, number of consecutive layers], BN = batch normalizations, MP = max pooling/US = upsampling layer	**Output size**
Input layer	–	112 × 112 × 1
Block 1	[3 × 3, 32, 1, 1 × 2], BN, MP[2 × 2]	56 × 56 × 32
Block 2	[3 × 3, 64, 1, 1 × 2], BN, MP[2 × 2]	28 × 26 × 64
Block 3	[3 × 3, 128, 1, 1 × 2], BN, MP[2 × 2]	14 × 14 × 128
–	US[2 × 2]	28 × 28 × 128
Block 4	[3 × 3, 64, 1, 1 × 2], BN, US[2 × 2]	28 × 28 × 64
Block 5	[3 × 3, 32, 1, 1 × 2], BN, US[2 × 2]	56 × 56 × 32
Block 5	US, [3 × 3, 1, 1, 1 × 1]	112 × 112 × 1

encoder's output feature maps, we were able to eliminate redundant and less significant features, resulting in a more compact and informative data set. As a result, we reduced the dimensionality of the feature vector for the 2D slice embeddings at each time step to 1024 dimensions using PCA. This reduced-dimensional feature vector was then used for network optimization at a later stage.

5.4 Hyperparameter tuning of the time-series models

For each time-series model, we performed hyperparameter tuning via an automated optimization process, namely BO. This method was chosen due to its efficiency in exploring complex, high-dimensional search spaces while minimizing the number of evaluations required. We applied BO to fine-tune the hyperparameters for the LSTM, GRU, and 1D CNN models. The goal was to identify the most effective configurations of the models to enhance their performance on AD progression detection. Tables 7.5 and 7.6 provide comprehensive details of the architectural configurations for these

Table 7.5 Enhanced 1D CNN architecture developed through Bayesian optimization.

	Hyperparameter search space Number of filters, filter size, amount of dropout	**Optimized parameter using Bayesian optimizer [size of the kernel, number of kernels, stride, padding], BN = batch normalizations, MP = max pooling with stride, D = dropout**	**Output size**
Input layer		–	3072 × 1
Conv1D_1	96–128/3–5	[5 × 1, 96, 2, 1], BN, MP[2/2]	767 × 960
Conv1D_2	128–160/3–5	[3 × 1, 128, 1, 1], BN[128], MP[2 × 2]	382 × 1280
Conv1D_3	128–192/3–5	[3 × 1, 162, 1, 1], BN[162], MP[2 × 2], D[0.3]	190 × 162
Flatten_layer	–	–	30,780 + 14
Dense_1	512–1024, 0.1–0.5	Dense(832), D(0.3)	832
Dense_2	64–128, 0.1–0.5	Dense(567), D(0.1)	867
Dense_3	–	Softmax(2)	2

Table 7.6 Enhanced LSTM architecture developed through the Bayesian optimization process.

Input layer	**Hyperparameter search space LSTM units/Regularization/ Activation function**	**[# of LSTM cells, L1 regularizer, activation function]**	**Output size**
Input	–	–	3 × 1024
LSTM_1	416–576/0.1–0.001/[Sigmoid, Tanh]	[512, 0.003, *Tanh*],	3 × 512
LSTM_2	440–480, 0.1–0.001, [Sigmoid, Tanh]	[512, 0.005, *Tanh*],	3 × 512
LSTM_3	440–680, 0.1–0.001, [Sigmoid, Tanh]	[384, 0.002, *Tanh*], dropout(0.3)	3 × 384
Flatten_layer	–	–	1152 + 14
Dense_1	256–512	Dense(416), dropout(0.2)	416
Dense_2	–	Softmax(2)	2

models. Each table outlines the search space, which includes the specific parameter ranges and values that the Bayesian optimizer investigates during the optimization process. The *optimized parameters using the Bayesian optimizer* column in each table represent the specific set of hyperparameters that the Bayesian optimizer selected for each model. By leveraging Bayesian optimization, we ensured that the models were configured in a way that maximizes their performance while reducing the risk of overfitting (Table 7.7).

5.5 Ensemble of optimized time-series models for progression detection

After optimizing the hyperparameters for each time-series model using the refined feature set, we integrated all three Bayesian-optimized models into an ensemble framework. The ensemble framework then processed a 110-dimensional feature set obtained from the PCA, corresponding to the MRI slices per longitudinal timestep for each subject. Each individual model in the ensemble setting independently computes class-wise probabilities for each slice, resulting in a total of 330 probability values across the two classification categories: CN and progress to AD. To derive the final classification decision, the output probabilities across all three models are aggregated by computing the class-wise average probabilities for each subject. This way the proposed EL framework leverages the strengths of both convolutional and recurrent networks to capture long range temporal dependencies using a pool of time-series models. The final prediction for AD progression was determined by selecting the class with the highest average probability, ensuring a more robust and generalized classification outcome.

Table 7.7 Enhanced GRU architecture developed through the Bayesian optimization process.

Input layer	Hyperparameter search space LSTM units/Regularization/ Activation function	[# of GRU units, L1 regularizer, activation function]	Output size
Input	–	–	3 × 1024
GRU_1	256~512, 0.001~0.1, [Sigmoid, Tanh]	[512, 0.002, *Tanh*],	3 × 512
GRU_2	416~576/0.001~0.1, [Sigmoid, Tanh]	[384, 0.006, *Tanh*],	3 × 384
GRU_3	416~576/0.001~0.1, [Sigmoid, Tanh]	[320, 0.004, *Tanh*],	3 × 320
Flatten_layer	–	–	960 + 14
Dense_1	256~512, dropout(0.1~0.5)	Dense(512), dropout(0.4)	512
Dense_2	64~128	Dense(64)	64
Dense_3	–	Softmax(2)	2

6. Experimental design

The proposed framework was developed using TensorFlow 2.0 and executed on an NVIDIA company (NVIDIA) GeForce TITAN X GPU. To ensure robust evaluation, a 10-fold cross-validation was employed, where the training dataset was partitioned into a 75–25 train-test split at each fold at run-time during the training phase. Performance was assessed and reported using multiple evaluation metrics, including mean precision, mean recall, mean F1-score, mean AUC, and mean accuracy.

6.1 Dataset

The dataset used in this study was obtained from the ADNI, a widely known open-access database for neuroimaging research (Hojjati & Babajani-Feremi, 2022). A total of 564 subjects were included in this study, comprising 182 CN subjects and 182 individuals diagnosed with AD. Given that this study focuses on longitudinal MRI-based disease progression modeling, each subject had three imaging time points with a 6-month interval (Baseline (BL), Month 06 (M06), and Month 12 (M12)), leading to a total of 1692 3D MRI volumes in the dataset. The MRI scans were acquired using 3T T1-weighted anatomical sequences, adhering to the 3D MPRAGE protocol, with an isotropic voxel resolution of $1 \times 1 \times 1$ mm. An important aspect of the dataset is the inclusion of 100 subjects classified as converters within the AD cohort. These individuals were initially diagnosed as CN at BL but gradually developed AD by M48. Their progression from CN to AD typically started from M18, accompanied by continuous cognitive decline and structural brain changes, ultimately leading to an AD diagnosis at M48. Including both converted patients and stable AD cases enabled the proposed model to capture and learn the patterns of progressive neurodegeneration, helping to differentiate early-stage AD from fully developed cases. This distinction is essential for early diagnosis and intervention and is consistent with findings from previous longitudinal neuroimaging studies.

In addition, the study also incorporated CS commonly used in both clinical and research settings, including ADAS13, FAQ, MMSE, APOE4, hippocampal volume, and Rey Auditory Verbal Learning Test (RAVLT). These CS serve as quantitative measures of memory impairment, functional decline, and overall cognitive status. The integration of neuroimaging and CS provides a comprehensive multimodal dataset, enabling robust learning of disease progression dynamics and facilitating more precise AD progression detection.

6.2 Configuration details and proposed enhancements to existing autoencoders

To evaluate the effectiveness of the proposed CAE, we compared the representational quality of the features produced by proposed CAE with those of others, such as VGG-CAE

and UNET-CAE. The selection of these models was based on two primary considerations: (1) both architectures incorporate a symmetrical encoder-decoder structure, making them suitable for direct comparison; and (2) there are no officially available CAE implementations for architectures such as ResNet, Inception, or DenseNet. Consequently, we designed the VGG-CAE from the scratch, while the UNET model was adapted from its original implementation (Weng & Zhu, 2021), with minor modifications to better align with our feature embedding requirements. A more detailed explanation of these comparative models is provided in the following subsections.

The latent feature extraction using VGG-CAE: We implemented CAE based on VGG-16 architecture following the steps outlined in (Xu et al., 2021). The objective was to learn abstract representations of input images and generate compact feature sets. For training the VGG-CAE, we utilized 2D MRI slices of size $128 \times 128 \times 1$ grayscale images. Binary cross-entropy was used as a loss function using the Adam optimizer as optimization function with learning rate of 1×10^{-4} for 120 epochs. We implemented early stopping callbacks to stop the training process and prevent overfitting.

The latent feature extraction using UNET-CAE: We also implemented the UNET model to extract latent features from the input 2D MRI slices. For training the UNET-CAE, we treated the input image as the target label and used it for training the model using binary cross-entropy loss. This way the model learned meaningful representations of input image in a latent space. We applied the same hyperparameter settings used for training VGG-CAE to UNET-CAE to ensure consistency in the training process. To extract the latent feature set, we first processed the input images using only the encoder module of UNET-CAE to generate latent representations for the 2D MRI slices. The resulting feature embeddings were then subjected to PCA for dimensionality reduction while retaining essential information. These refined embeddings were used to optimize time-series models through a BO process. This procedure was applied to generate feature embeddings for all comparative CAEs.

7. Experiments and results

We conducted a comprehensive set of analysis of the proposed framework by performing a wide range of experiments to assess its robustness, accuracy, and generalizability. This evaluation involved comparing its performance against other CAE-based architectures, analyzing the impact of different feature embeddings, and fine-tuning hyperparameters for optimal results. In addition, we examined how integrating cognitive scores influenced classification outcomes and tested the model's ability to generalize across different data distributions. To further validate its effectiveness, we conducted cross-validation, measured key performance metrics such as accuracy, AUC, precision, recall, and F1-score to gain deeper insights into its decision-making process.

7.1 Exp. 1: single modality: network optimization using MRI only

In this experiment, various time-series models (1D CNN, LSTM, and GRU) were optimized using MRI feature embeddings extracted from different CAEs, including the proposed CAE, VGG-CAE, and UNET-CAE. These models were initially evaluated at the BL time step, with additional longitudinal data incorporated at six- and 12-month intervals (BL ~ M06 and BL ~ M12). Results in Table 7.8 indicate that the 1D CNN model performed the best at BL ~ M06 when using VGG-CAE embeddings, but UNET-CAE and the proposed CAE yielded greater stability and improved accuracy over time. The LSTM model achieved its highest accuracy (80.20%) with UNET-CAE embeddings, though the proposed CAE provided more consistent performance. The GRU model showed steady improvements across all metrics, demonstrating its ability to capture temporal patterns in AD progression. Notably, models utilizing the proposed CAE based features embeddings exhibited progressive performance improvement with additional time steps, highlighting the advantages of longitudinal data integration in improving disease prediction accuracy.

Figure 7.2 evaluates model efficacy using the mAUC as the primary metric due to its robustness. The 1D CNN model demonstrated improved performance when longitudinal MRI embeddings were incorporated, achieving its highest mAUC score (80%) at BL ~ M12 with CAE-based embeddings. LSTM performed best at BL with VGG-CAE embeddings (77%) and outperformed at BL ~ M12 for UNET (83%) and CAE (79%). The GRU model attained its highest mAUC scores at BL ~ M12 across all embeddings, though its trends were less consistent. To summarize, multimodality-based ML algorithms have shown enhanced diagnostic precision compared to single-modality based ML models. These findings motivated the subsequent experiment, which investigates the integration of multimodal data to further improve AD diagnosis.

7.2 Exp. 2: multiple modalities: network optimization using MRI + CS

In this experiment, the models from the previous experiment were further optimized by integrating longitudinal MRI features with CS through multimodal data fusion. BO was employed for hyperparameter tuning, leading to improved model performance and stability. This enhancement aligned with medical principles, as multimodal data provides a more comprehensive representation of disease progression by enriching feature representations and refining decision boundaries (El-Sappagh et al., 2020; Huang, 2011; Xu et al., 2015). Performance comparisons in Table 7.9 demonstrated that multimodal data fusion consistently improved classification accuracy over MRI-only models. The 1D CNN achieved its highest accuracy (85.01%) at BL ~ M12 using the proposed CAE-based embeddings, outperforming VGG-CAE and UNET-CAE. While LSTM exhibited inconsistent results with VGG-CAE embeddings, it showed an upward performance trend with UNET and CAE-based embeddings. The GRU model demonstrated the most stable improvement, achieving the highest AUC (90.86%) and accuracy (87.99%) at BL ~ M12 with the proposed CAE embeddings. These results confirmed that integrating CS with MRI features significantly enhances the accuracy and robustness of AD progression detection.

Table 7.8 Optimization of time-series DL models using MRI feature embeddings and their achieved performance.

		1D CNN			LSTM			GRU		
CAE	**Met.**	**BL**	**M6**	**M12**	**BL**	**M6**	**M12**	**BL**	**M6**	**M12**
VGG	Pre	75.53 ± 3.00	76.38 ± 2.90	73.21 ± 2.51	77.53 ± 4.69	74.32 ± 4.23	77.99 ± 2.23	78.42 ± 4.64	83.21 ± 3.45	84.53 ± 3.61
	Rec	75.01 ± 2.99	76.53 ± 2.90	72.31 ± 1.43	78.66 ± 3.72	72.14 ± 2.07	77.73 ± 1.93	78.35 ± 3.81	84.55 ± 1.77	85.43 ± 4.21
	F1	74.93 ± 1.28	77.83 ± 3.89	72.08 ± 1.41	78.93 ± 2.14	73.81 ± 2.29	76.03 ± 3.11	77.51 ± 3.37	83.32 ± 2.33	85.83 ± 3.65
	AUC	75.44 ± 0.99	77.15 ± 2.58	71.13 ± 3.54	78.63 ± 1.57	74.91 ± 3.41	76.61 ± 2.13	79.19 ± 3.51	82.75 ± 3.89	85.64 ± 2.31
	Acc	74.52 ± 2.91	75.82 ± 1.47	70.86 ± 2.22	76.89 ± 3.89	72.29 ± 4.62	76.98 ± 3.44	78.22 ± 4.45	81.87 ± 3.84	83.69 ± 1.41
UNET	Pre	75.85 ± 2.77	78.33 ± 2.93	80.83 ± 2.08	75.98 ± 2.42	75.14 ± 3.37	77.13 ± 3.93	77.41 ± 2.51	82.81 ± 2.81	86.28 ± 3.34
	Rec	74.72 ± 3.11	77.93 ± 3.86	80.71 ± 1.01	74.86 ± 2.52	75.81 ± 2.96	78.48 ± 2.28	79.23 ± 3.56	82.11 ± 2.64	85.82 ± 3.59
	F1	75.88 ± 3.55	77.82 ± 3.64	81.86 ± 3.29	75.85 ± 2.92	76.67 ± 3.94	77.06 ± 2.73	78.11 ± 1.94	83.10 ± 2.91	84.77 ± 1.59
	AUC	74.29 ± 2.19	77.21 ± 2.68	79.35 ± 2.94	74.49 ± 2.22	76.64 ± 1.44	78.24 ± 2.81	78.98 ± 3.41	82.35 ± 1.59	85.84 ± 1.22
	Acc	72.83 ± 2.94	75.62 ± 2.31	79.84 ± 3.06	73.43 ± 3.06	75.81 ± 2.95	76.55 ± 3.75	76.82 ± 2.16	80.92 ± 3.08	84.88 ± 3.35
Proposed	Pre	78.47 ± 2.40	83.91 ± 1.91	86.51 ± 1.41	82.24 ± 3.32	82.28 ± 1.59	86.12 ± 2.33	82.87 ± 2.76	84.11 ± 3.56	89.42 ± 2.16
	Rec	79.33 ± 2.29	84.91 ± 2.68	87.69 ± 2.82	81.02 ± 4.13	83.81 ± 2.93	86.59 ± 2.05	82.74 ± 2.15	85.64 ± 2.98	88.54 ± 3.27
	F1	78.94 ± 1.99	84.62 ± 1.17	86.22 ± 1.22	82.94 ± 4.26	84.95 ± 1.94	86.81 ± 1.97	83.39 ± 2.19	85.16 ± 4.62	88.02 ± 1.24
	AUC	79.65 ± 1.41	83.58 ± 2.58	87.58 ± 1.41	82.06 ± 2.13	84.92 ± 1.44	87.52 ± 3.56	84.69 ± 4.22	86.94 ± 2.85	90.86 ± 2.17
	Acc	77.53 ± 2.53	81.71 ± 1.38	85.01 ± 2.75	80.37 ± 1.99	81.54 ± 3.65	85.98 ± 2.27	81.87 ± 2.22	83.27 ± 1.96	87.99 ± 3.07

CAE = convolutional autoencoder, Met. = Evaluation metric, BL = baseline, m6 = BL ~ M6, m12 = BL ~ M12

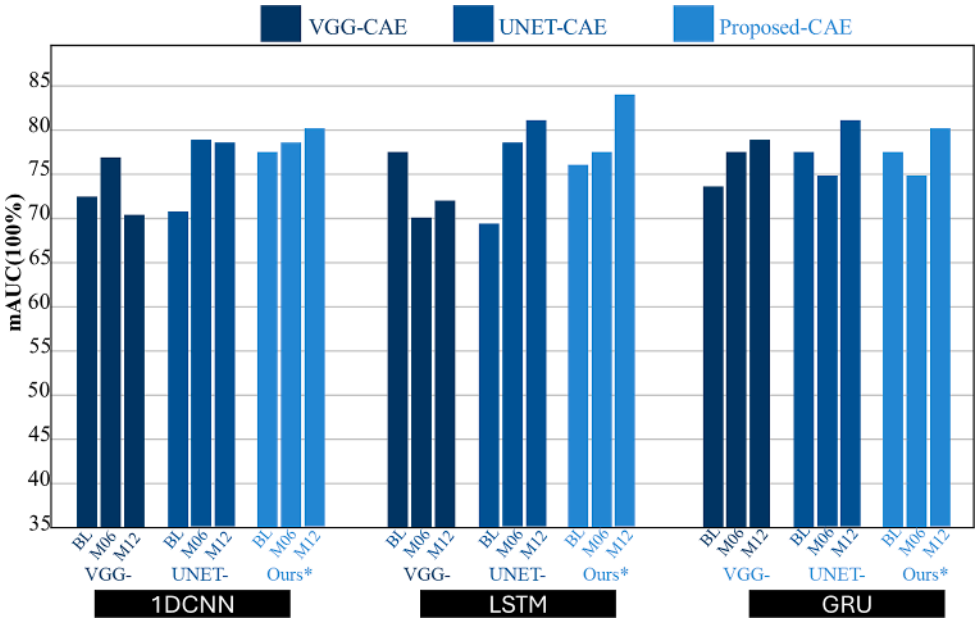

FIGURE 7.2 Comparison of AUC values across individual optimized time-series models and their respective CAEbased feature embeddings.

Figure 7.3 compares the mAUC of BO time-series models trained with multimodal data, showing a consistent increase in performance as more longitudinal data was included. The 1D CNN model with UNET and proposed CAE-based embeddings demonstrated improved predictive accuracy with additional time steps, with mAUC increasing by 3% and 7%, respectively, compared to single-modality training. However, VGG-CAE-based embeddings caused instability in performance. The LSTM model showed mixed results, improving with VGG-CAE and CAE-based embeddings, but not with UNET. The GRU model showed consistent improvement across all embeddings, particularly with CAE-based features, where mAUC rose from 79% to 90%. These results indicated that integrating CS with MRI data enhanced model stability and accuracy, with CAE-based embeddings offering the most consistent improvements. The next experiment will explore how combining multiple optimized models improves diagnostic accuracy over single-model approaches.

7.3 Exp. 3: evaluating the impact of homogeneous and heterogeneous deep EL models using multimodal data

In this experiment, we extended our analysis by investigating the effectiveness of EL models. To achieve this, we evaluated various EL configurations, including homogeneous and heterogeneous EL models. In the homogeneous EL setup, multiple instances of the same base model were combined, each trained with different parameters for the

Table 7.9 Performance evaluation of BO time-series models using feature embeddings extracted from various CAEs.

		1D CNN			LSTM			GRU		
CAE	**Met.**	**BL**	**M6**	**M12**	**BL**	**M6**	**M12**	**BL**	**M6**	**M12**
VGG	Pre	75.53 ± 3.00	76.38 ± 2.90	73.21 ± 2.51	77.53 ± 4.69	74.32 ± 4.23	77.99 ± 2.23	78.42 ± 4.64	83.21 ± 3.45	84.53 ± 3.61
	Rec	75.01 ± 2.99	76.53 ± 2.90	72.31 ± 1.43	78.66 ± 3.72	72.14 ± 2.07	77.73 ± 1.93	78.35 ± 3.81	84.55 ± 1.77	85.43 ± 4.21
	F1	74.93 ± 1.28	77.83 ± 3.89	72.08 ± 1.41	78.93 ± 2.14	73.81 ± 2.29	76.03 ± 3.11	77.51 ± 3.37	83.32 ± 2.33	85.83 ± 3.65
	AUC	75.44 ± 0.99	77.15 ± 2.58	71.13 ± 3.54	78.63 ± 1.57	74.91 ± 3.41	76.61 ± 2.13	79.19 ± 3.51	82.75 ± 3.89	85.64 ± 2.31
	Acc	74.52 ± 2.91	75.82 ± 1.47	70.86 ± 2.22	76.89 ± 3.89	72.29 ± 4.62	76.98 ± 3.44	78.22 ± 4.45	81.87 ± 3.84	83.69 ± 1.41
UNET	Pre	**77.53 ± 4.69**	74.32 ± 4.23	77.99 ± 2.23	75.98 ± 2.42	75.14 ± 3.37	**77.13 ± 3.93**	77.41 ± 2.51	82.81 ± 2.81	**86.28 ± 3.34**
	Rec	**78.66 ± 3.72**	72.14 ± 2.07	77.73 ± 1.93	74.86 ± 2.52	75.81 ± 2.96	**78.48 ± 2.28**	79.23 ± 3.56	82.11 ± 2.64	**85.82 ± 3.59**
	F1	**78.93 ± 2.14**	73.81 ± 2.29	76.03 ± 3.11	75.85 ± 2.92	76.67 ± 3.94	**77.06 ± 2.73**	78.11 ± 1.94	83.10 ± 2.91	**84.77 ± 1.59**
	AUC	**78.63 ± 1.57**	74.91 ± 3.41	76.61 ± 2.13	74.49 ± 2.22	76.64 ± 1.44	**78.24 ± 2.81**	78.98 ± 3.41	82.35 ± 1.59	**85.84 ± 1.22**
	Acc	**76.89 ± 3.89**	72.29 ± 4.62	76.98 ± 3.44	73.43 ± 3.06	75.81 ± 2.95	**76.55 ± 3.75**	76.82 ± 2.16	80.92 ± 3.08	**84.88 ± 3.35**
Proposed	Pre	78.47 ± 2.40	83.91 ± 1.91	**86.51 ± 1.41**	82.24 ± 3.32	82.28 ± 1.59	**86.12 ± 2.33**	82.87 ± 2.76	84.11 ± 3.56	**89.42 ± 2.16**
	Rec	79.33 ± 2.29	84.91 ± 2.68	**87.69 ± 2.82**	81.02 ± 4.13	83.81 ± 2.93	**86.59 ± 2.05**	82.74 ± 2.15	85.64 ± 2.98	**88.54 ± 3.27**
	F1	78.94 ± 1.99	84.62 ± 1.17	**86.22 ± 1.22**	82.94 ± 4.26	84.95 ± 1.94	**86.81 ± 1.97**	83.39 ± 2.19	85.16 ± 4.62	**88.02 ± 1.24**
	AUC	79.65 ± 1.41	83.58 ± 2.58	**87.58 ± 1.41**	82.06 ± 2.13	84.92 ± 1.44	**87.52 ± 3.56**	84.69 ± 4.22	86.94 ± 2.85	**90.86 ± 2.17**
	Acc	77.53 ± 2.53	81.71 ± 1.38	**85.01 ± 2.75**	80.37 ± 1.99	81.54 ± 3.65	**85.98 ± 2.27**	81.87 ± 2.22	83.27 ± 1.96	**87.99 ± 3.07**

CAE = convolutional autoencoder, Met. = Evaluation metric, BL = baseline, m6 = BL ~ M6, m12 = BL ~ M12

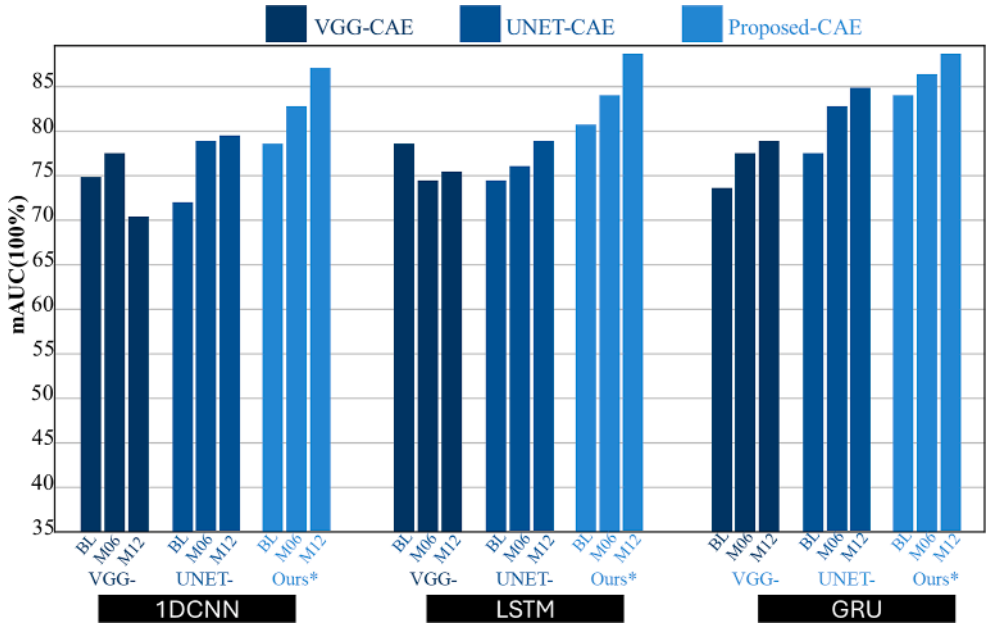

FIGURE 7.3 Comparison of mAUC scores across different time-series models using multimodal data.

same input modality. For example, we explored combinations such as VGG-CAE $\times$ 3 with 1D CNN $\times$ 3, UNET-CAE $\times$ 3 with 1D CNN $\times$ 3, and proposed CAE $\times$ 3 with 1D CNN $\times$ 3. A similar setup was applied to other time-series models, including LSTM and GRU, to assess their respective ensemble performances. The details of these configurations are presented in Table 7.10. For heterogeneous EL models, we combined different types of time-series models within an ensemble framework to leverage their complementary strengths. For instance, we evaluated combinations such as VGG-CAE $\times$ 3 with 1D CNN + LSTM + GRU, and similar heterogeneous setups were tested for UNET-CAE and Proposed CAE-based feature extractors. These configurations were designed to assess whether integrating diverse model architectures leads to improved AD progression detection compared to homogeneous ensembles. Table 7.10 presentes the performance comparison of homogeneous and heterogeneous EL models trained on MRI embeddings and cognitive scores.

In the homogeneous EL setting, models trained with identical network types exhibit moderate improvement over the BL time step. Among them, the LSTM ensembles generally perform better than 1D CNN and GRU ensembles, particularly in terms of recall and AUC. The UNET-CAE and proposed CAE-based feature extractors show a consistent trend of improved performance over time (from BL to M12), with proposed CAE achieving slightly better F1-scores and AUC values. However, accuracy gains remain limited compared to the heterogeneous models. In contrast, the heterogeneous EL setup, which combines 1D CNN, LSTM, and GRU, demonstrated significantly better performance across all metrics. Notably, the heterogeneous ensemble with the proposed CAE

Table 7.10 A multimodal ensemble model with homogenous and heterogenous networks trained on MRI embeddings and cognitive scores.

		Homogenous EL									Heterogenous EL		
		1D CNN £ 3			LSTM £ 3			GRU £ 3			1D CNN + LSTM + GRU		
CAE	Met.	BL	M6	M12	BL	M6	M12	BL	M6	M12	BL	M6	M12
VGG	Pre	75.13 ± 3.10	74.23 ± 3.12	76.17 ± 2.62	74.12 ± 2.12	76.32 ± 3.33	75.52 ± 3.35	75.32 ± 3.47	74.38 ± 3.41	73.12 ± 2.40	88.38 ± 3.41	89.38 ± 2.41	94.11 ± 2.10
	Rec	74.71 ± 3.09	73.63 ± 3.01	75.31 ± 1.53	73.43 ± 3.43	77.51 ± 4.62	76.33 ± 2.21	74.54 ± 3.34	75.12 ± 2.31	72.34 ± 2.42	87.42 ± 3.21	89.12 ± 2.14	94.24 ± 2.11
	F1	75.81 ± 1.38	73.82 ± 4.24	75.24 ± 1.51	74.43 ± 3.11	76.43 ± 2.43	76.49 ± 2.35	75.44 ± 2.73	75.72 ± 2.32	73.32 ± 2.33	88.31 ± 2.72	88.72 ± 2.12	93.02 ± 3.31
	AUC	74.31 ± 1.50	72.35 ± 2.38	75.43 ± 4.05	73.64 ± 2.22	76.23 ± 3.32	75.34 ± 2.24	75.12 ± 2.52	74.36 ± 3.42	73.16 ± 2.16	89.32 ± 2.12	89.42 ± 3.82	93.46 ± 2.10
	Acc	74.61 ± 3.01	72.51 ± 1.57	75.64 ± 2.31	72.64 ± 3.51	75.55 ± 3.55	75.74 ± 2.34	74.75 ± 3.55	73.66 ± 2.25	72.39 ± 3.53	87.20 ± 2.75	87.60 ± 2.05	93.29 ± 3.07
UNET	Pre	71.75 ± 2.37	75.44 ± 3.43	74.74 ± 2.45	76.46 ± 3.45	78.14 ± 3.13	79.56 ± 2.15	74.34 ± 3.43	76.67 ± 3.23	79.11 ± 2.28	88.32 ± 3.01	90.38 ± 3.41	95.11 ± 2.48
	Rec	72.42 ± 2.40	74.34 ± 3.15	76.04 ± 4.34	75.55 ± 3.64	78.51 ± 4.62	78.82 ± 3.32	75.24 ± 3.34	76.31 ± 2.44	80.44 ± 2.28	89.82 ± 3.54	90.12 ± 2.14	95.24 ± 2.18
	F1	73.07 ± 3.76	74.71 ± 3.74	75.63 ± 3.44	76.54 ± 4.25	79.52 ± 3.26	78.77 ± 3.62	74.43 ± 3.14	75.73 ± 3.22	79.33 ± 3.23	88.41 ± 2.44	90.72 ± 2.12	94.02 ± 2.31
	AUC	73.14 ± 1.30	74.14 ± 3.77	76.24 ± 2.45	76.41 ± 4.33	78.32 ± 2.23	79.88 ± 2.32	75.62 ± 3.42	77.42 ± 3.65	80.74 ± 2.35	89.35 ± 3.82	91.42 ± 3.17	95.84 ± 2.10
	Acc	72.52 ± 3.24	73.55 ± 1.23	75.33 ± 4.36	75.55 ± 3.39	77.23 ± 3.25	77.75 ± 2.65	74.70 ± 2.25	75.43 ± 3.55	79.33 ± 2.53	87.80 ± 2.75	89.50 ± 2.55	94.51 ± 2.02
Proposed	Pre	73.74 ± 3.45	74.79 ± 2.01	76.44 ± 2.52	72.88 ± 3.44	75.74 ± 3.64	77.45 ± 2.33	74.33 ± 3.33	76.42 ± 3.32	77.12 ± 2.21	89.13 ± 2.23	93.41 ± 2.31	96.11 ± 1.14
	Rec	72.04 ± 4.34	72.81 ± 3.78	75.59 ± 2.92	73.33 ± 3.34	74.02 ± 3.42	78.26 ± 2.44	74.44 ± 3.32	76.54 ± 2.72	78.41 ± 1.37	89.41 ± 2.22	93.65 ± 2.71	96.24 ± 1.38
	F1	73.43 ± 2.24	73.52 ± 2.28	75.16 ± 2.32	73.77 ± 3.15	74.47 ± 2.35	77.34 ± 1.43	75.36 ± 2.33	75.12 ± 3.23	78.43 ± 2.35	88.06 ± 1.23	94.42 ± 2.21	96.22 ± 1.30
	AUC	73.24 ± 3.45	73.48 ± 3.68	75.56 ± 1.54	72.34 ± 3.22	75.33 ± 2.32	78.39 ± 2.24	75.34 ± 2.82	77.24 ± 2.43	78.34 ± 2.83	88.24 ± 2.81	93.54 ± 1.41	97.04 ± 2.12
	Acc	72.33 ± 3.36	72.61 ± 2.44	74.94 ± 2.46	71.70 ± 3.25	74.45 ± 2.45	75.99 ± 2.47	73.55 ± 3.72	74.71 ± 2.32	77.43 ± 2.52	87.75 ± 2.75	92.51 ± 2.75	95.83 ± 2.32

CAE = convolutional autoencoder, Met. = Evaluation metric, BL = baseline, m6 = BL ~ M6, m12 = BL ~ M12.

feature extractor achieved the highest accuracy ($95.83 \pm 2.32\%$) at M12, followed closely by UNET-CAE ($94.51 \pm 2.02\%$) and VGG-CAE ($93.29 \pm 3.07\%$). A similar trend was observed for AUC and F1-score, where heterogeneous ensembles consistently outperform their homogeneous counterparts, highlighting the benefits of integrating diverse model architectures. Overall, the results indicated that heterogeneous ensembles provide a substantial boost in diagnostic accuracy, particularly at later time points (M12), where multimodal information is better leveraged to enhance performance. The proposed CAE-based heterogeneous EL model emerged as the most effective configuration, achieving the highest diagnostic accuracy and robustness across different evaluation metrics.

Figure 7.4 presents a comparative analysis of mAUC scores for homogeneous and heterogeneous deep EL models, highlighting the impact of model diversity on classification performance. The results suggest that the heterogeneous deep EL configuration, integrating three distinct base classifiers, consistently outperformed all homogeneous EL setups. Notably, homogeneous EL, despite leveraging multiple instances of the same model type, did not exhibit a clear trend of improved accuracy as longitudinal time steps increased. In addition, none of the homogeneous ensemble configurations surpassed the 80% mAUC threshold, suggesting limitations in their ability to generalize well across longitudinal embeddings. In contrast, heterogeneous ELs demonstrated a significant performance boost when incorporating three diverse classifiers, exhibiting not only higher mAUC scores but also enhanced stability across feature embeddings as the number of longitudinal time steps increased. This finding underscores the advantage of architectural diversity in EL for robust AD's progression detection.

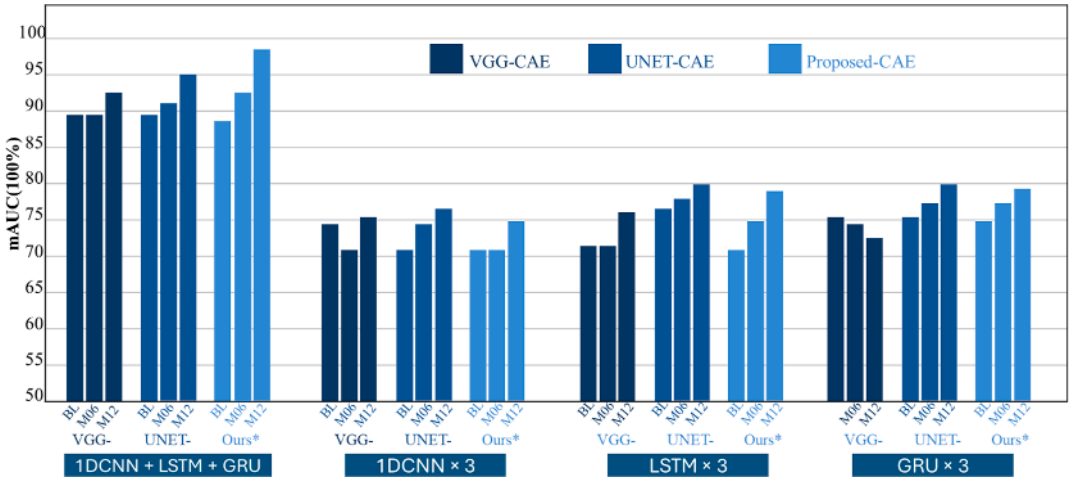

FIGURE 7.4 Comparison of homogeneous and heterogeneous deep EL models trained on multimodal data.

8. Comparative analysis of SoTA methods in AD progression detection

Table 7.11 provides a comprehensive summary of recent studies on AD progression detection, highlighting the key characteristics of each study. The table includes the following critical details: 1) whether the study utilizes longitudinal data, 2) if longitudinal data is employed, what are the number of longitudinal time step, 3) the data modalities used in a study (such as MRI, PET, or cognitive scores), 4) the number of classification categories (e.g., binary or multiclass classification), and 5) the reported performance metrics, including accuracy, precision, recall, and F1-score. The comparative analysis presented that our proposed network outperforms most existing state-of-the-art models in terms of accuracy, except for the model presented in (El-Sappagh et al., 2022), which achieved an impressive 99% accuracy across several evaluation metrics. This superior

Table 7.11 Comprehensive analysis of the proposed approach with SOTA progression detection models.

Ref.	Long. Modality		Data modalities	CC	Acc (%)	Pre (%)	Rec (%)	F1. (%)	AUC (%)
	Yes/ No.	# Of TS							
Farhan et al. (2014)	No	1	M1	CN versus AD	93.75	–	87.5	–	–
Sivapriya et al. (2015)	No	1	NSB, M4	CN versus AD	96.3	96.3	96.3	–	–
Nanni et al. (2016)	No	1	M1	CN versus AD	93.0	–	–	–	–
Armañanzas et al. (2017)	Yes	1	fMRI	CN versus AD	95.0	–	–	–	–
Moore et al. (2019)	Yes	–	M1, M2, M5	CN versus MCI versus AD	73.0	–	–	–	–
Ebadi (2017)	No	1	DTI	CN versus AD	83.3	80.0	82.5		
Qiu et al. (2018)	No	1	M1, MMSE, logical memory	CN versus MCI	91.50	93.56	95.93	95.24	–
Choi and Lee (2020)	No	1	M1	CN versus AD	93.84	–	–	–	–
Pan et al. (2020)	No	1	M1	AD progression	84.00				92.00
Ruiz et al. (2020)	No	1	M1	AD progression	83.33				
Muhammed and Thiyagarajan (2021)	No	1	M1, M4, M5, M2, M3	CN versus MCI versus AD	84.0	–	–	–	–
Bi et al. (2021)	No	1	fMRI	AD progression	83.47	–	90.90	–	–
Giovannetti et al. (2021)	Yes	–	MEG, M1	AD progression	87.00	–	–	–	–
Liu et al. (2021)	No	1	M1, M3, MMSE	AD progression	83.50	–	79.4	–	–
Tanveer et al. (2022)	No	1	M1	CN versus AD	85.27	–	87.32	–	–
Razzak et al. (2022)	No	1	M1	CN versus AD	87.11	–	–	–	–
El-Sappagh et al. (2022)	Yes	4	M2, NSB, Statistics	CN versus AD	99.56	99.56	99.56	99.51	–
Yiğit et al. (2022)	No	1	M4	CN versus AD	89.05	–	–	–	–
Ours*	Yes	3	MRI, M2	AD progression	**96.11**	96.24	96.22	97.06	95.43

Ref = Reference, MRI = M1, CS = M2, Demographics = M3, FDG-PET = M4, CSF = M5

performance in (El-Sappagh et al., 2022) is primarily attributed to the use of a larger number of longitudinal time steps and a more extensive combination of multimodal data inputs, offering a more granular and richer representation of the disease progression. These findings further support our hypothesis that the inclusion of diverse data modalities and a higher number of longitudinal time steps in the training process substantially improves a model's capacity to detect AD progression with greater precision. By incorporating such a comprehensive dataset, ML models can capture intricate temporal patterns and multimodal interactions, leading to enhanced robustness and generalization when applied to clinical scenarios.

9. Limitations and future research directions

This section highlights the constraints of the current study and outlines potential areas for future exploration and improvement.

- This study primarily utilized a single neuroimaging modality, MRI, for analysis. Expanding the approach to incorporate other imaging, such as functional MRI (fMRI) and PET, could provide a more comprehensive representation of brain activity and pathology. Future research will focus on integrating multimodal imaging data to enhance classification performance and improve the overall understanding of AD's progression detection.
- The dataset used for this study was obtained from the ADNI database. While ADNI provides high-quality neuroimaging and clinical data, relying on a single data source may limit the model's ability to generalize well across diverse populations. Our future studies will aim to incorporate data from additional sources, increasing sample diversity and enhancing the model's robustness in real-world applications.
- Although longitudinal MRI scans (BL, M06, M12) proved effective in predicting AD progression, incorporating additional longitudinal time points in future analyses will enable a deeper understanding of disease evolution. Capturing temporal changes in brain structure through sequential imaging data will allow the model to learn spatiotemporal patterns, potentially leading to earlier and more precise predictions of AD onset and progression.
- Future research will also explore transformer-based multimodal ML frameworks to enhance data fusion and processing. These models will be designed to efficiently integrate diverse data modalities, such as neuroimaging, genetic, and clinical information, leveraging their self-attention mechanisms to capture complex spatial and temporal dependencies. This approach will aim to improve both interpretability and predictive accuracy, ultimately contributing to more effective diagnostic and prognostic tools for neurodegenerative diseases.

10. Conclusion

In this work, we proposed a hybrid deep ensemble learning framework for detecting AD progression by integrating multimodal medical data. Traditional approaches primarily rely on single-modality data, limiting their ability to capture complex disease patterns. To address this limitation, our framework leveraged a combination of longitudinal MRI scans and cognitive scores, effectively enhancing predictive accuracy. By incorporating a lightweight convolutional autoencoder (CAE) for feature extraction and employing Bayesian optimization to fine-tune a heterogeneous ensemble of time-series models, we achieved a robust and computationally efficient solution for AD detection. Our findings suggest that multimodal data fusion significantly improves model performance, achieving state-of-the-art accuracy of 95% and an AUC of 97%. We anticipate that this research will contribute to the advancement of smart healthcare by facilitating early detection and personalized intervention for AD.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2021-NR058558), (Institute for Information & communications Technology Planning & Evaluation) (IITP) grant funded by the Korea government (MSIT) under the ICT Creative Consilience Program (IITP-2021–20-0-01821), and AI Platform to Fully Adapt and Reflect Privacy-Policy Changes (RS-2022-II220688).

References

Abrol, A., Bhattarai, M., Fedorov, A., Du, Y., Plis, S., & Calhoun, V. (2020). Deep residual learning for neuroimaging: An application to predict progression to Alzheimer's disease. *Elsevier B.V., United States Journal of Neuroscience Methods, 339*. https://doi.org/10.1016/j.jneumeth.2020.108701

Abuhmed, T., El-Sappagh, S., & Alonso, J. M. (2021). Robust hybrid deep learning models for Alzheimer's progression detection. *Knowledge-Based Systems, 213*, Article 106688. https://doi.org/10.1016/j.knosys.2020.106688

Albright, J. (2019). Forecasting the progression of Alzheimer's disease using neural networks and a novel preprocessing algorithm. *Alzheimer's and Dementia: Translational Research and Clinical Interventions, 5*, 483–491. https://doi.org/10.1016/j.trci.2019.07.001

An, N., Ding, H., Yang, J., Au, R., & Ang, T. F. A. (2020). Deep ensemble learning for Alzheimer's disease classification. *Journal of Biomedical Informatics*, 105. https://doi.org/10.1016/j.jbi.2020.103411

Armañanzas, R., Iglesias, M., Morales, D. A., & Alonso-Nanclares, L. (2017). Voxel-based diagnosis of Alzheimer's disease using classifier ensembles. *IEEE Journal of Biomedical and Health Informatics, 21*(3), 778–784. https://doi.org/10.1109/JBHI.2016.2538559

B.B. Avants, N.J. Tustison, H.J. Johnson, ANTs by stnava .

Bazargani, J. S., Rahim, N., Sadeghi-Niaraki, A., Abuhmed, T., Song, H., & Choi, S. M. (2024). Alzheimer's disease diagnosis in the metaverse. *Computer Methods and Programs in Biomedicine, 255*. https://doi.org/10.1016/j.cmpb.2024.108348

Bi, X.a., Xie, Y., Wu, H., & Xu, L. (2021). Identification of differential brain regions in MCI progression via clustering-evolutionary weighted SVM ensemble algorithm. *Frontiers of Computer Science, 15*(6). https://doi.org/10.1007/s11704-020-9520-3

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience, 3*(3), 243–249. https://doi.org/10.1038/nrn756

Chincarini, A., Sensi, F., Rei, L., Gemme, G., Squarcia, S., Longo, R., Brun, F., Tangaro, S., Bellotti, R., Amoroso, N., Bocchetta, M., Redolfi, A., Bosco, P., Boccardi, M., Frisoni, G. B., & Nobili, F. (2016). Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease. *NeuroImage, 125*, 834–847. https://doi.org/10.1016/j.neuroimage.2015.10.065

Choi, J. Y., & Lee, B. (2020). Combining of multiple deep networks via ensemble generalization loss, based on MRI images, for Alzheimer's Disease classification. In , *Institute of electrical and electronics engineers Inc., South Korea IEEE signal processing letter: Vol 27. s* (pp. 206–210). https://doi.org/10.1109/LSP.2020.2964161

Dai, Y., Zou, B., Zhu, C., Li, Y., Chen, Z., Ji, Z., Kui, X., & Zhang, W. (2023). DE-JANet: A unified network based on dual encoder and joint attention for Alzheimer's disease classification using multi-modal data. *Computers in Biology and Medicine, 165*, Article 107396. https://doi.org/10.1016/j.compbiomed.2023.107396

Duan, S., Shi, Q., & Wu, J. (2022). Multimodal sensors and ML-Based data fusion for advanced robots. *Advanced Intelligent Systems, 4*(12). https://doi.org/10.1002/aisy.202200213

Ebadi, A. (2017). Ensemble classification of alzheimer's disease and mild cognitive impairment based on complex graph measures from diffusion tensor images. *Frontiers in Neuroscience, 11*. https://doi.org/10.3389/FNINS.2017.00056

El-Sappagh, S., Abuhmed, T., Riazul Islam, S. M., & Kwak, K. S. (2020). Multimodal multitask deep learning model for Alzheimer's disease progression detection based on time series data. *Neurocomputing, 412*, 197–215. https://doi.org/10.1016/j.neucom.2020.05.087

El-Sappagh, S., Ali, F., Abuhmed, T., Singh, J., & Alonso, J. M. (2022). Automatic detection of Alzheimer's disease progression: An efficient information fusion approach with heterogeneous ensemble classifiers. *Neurocomputing, 512*, 203–224. https://doi.org/10.1016/j.neucom.2022.09.009

El-Sappagh, S., Saleh, H., Sahal, R., Abuhmed, T., Islam, S. M. R., Ali, F., & Amer, E. (2021). Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data. *Future Generation Computer Systems*, (115), 680–699. https://doi.org/10.1016/j.future.2020.10.005

Eslami, M., Tabarestani, S., & Adjouadi, M. (2023). A unique color-coded visualization system with multimodal information fusion and deep learning in a longitudinal study of Alzheimer's disease. *Artificial Intelligence in Medicine, 140*, Article 102543. https://doi.org/10.1016/j.artmed.2023.102543

Fan, Y., Batmanghelich, N., Clark, C. M., & Davatzikos, C. (2008). Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage, 39*(4), 1731–1743. https://doi.org/10.1016/j.neuroimage.2007.10.031

Fard, A. S., Reutens, D. C., & Vegh, V. (2021). *CNNs and GANs in MRI-based cross-modality medical image estimation*. Australia: arXiv. https://arxiv.org.

Farhan, S., Fahiem, M. A., & Tauseef, H. (2014). An ensemble-of-classifiers based approach for early diagnosis of Alzheimer'S disease: Classification using structural features of brain images. *Computational and Mathematical Methods in Medicine, 2014*. https://doi.org/10.1155/2014/862307

Feng, J., Shao-Wu, Z., Luonan, C., & Alzheimer's Disease Neuroimaging Initiative (ADNI). (2020). Identification of Alzheimer's disease based on wavelet transformation energy feature of the structural MRI image and NN classifier. *Artificial Intelligence in Medicine, 108*, 101940. https://doi.org/10.1016/j.artmed.2020.101940

Giovannetti, A., Susi, G., Casti, P., Mencattini, A., Pusil, S., López, M. E., Di Natale, C., & Martinelli, E. (2021). Deep-MEG: Spatiotemporal CNN features and multiband ensemble classification for

predicting the early signs of Alzheimer's disease with magnetoencephalography. *Neural Computing & Applications, 33*(21), 14651–14667. https://doi.org/10.1007/s00521-021-06105-4

Guo, F., Ng, M., Kuling, G., & Wright, G. (2022). Cardiac MRI segmentation with sparse annotations: Ensembling deep learning uncertainty and shape priors. *Medical Image Analysis, 81*, Article 102532. https://doi.org/10.1016/j.media.2022.102532

Ho, N. H., Yang, H. J., Kim, J., Dao, D. P., Park, H. R., & Pant, S. (2022). Predicting progression of Alzheimer's disease using forward-to-backward bi-directional network with integrative imputation. *Neural Networks, 150*, 422–439. https://doi.org/10.1016/j.neunet.2022.03.016

Hojjati, S. H., & Babajani-Feremi, A. (2022). Prediction and modeling of neuropsychological scores in alzheimer's disease using multimodal neuroimaging data and artificial neural networks. *Frontiers in Computational Neuroscience, 15*. https://doi.org/10.3389/fncom.2021.769982

Hong, X., Lin, R., Yang, C., Zeng, N., Cai, C., Gou, J., & Yang, J. (2019). Predicting Alzheimer's disease using LSTM. *IEEE Access, 7*, 80893–80901. https://doi.org/10.1109/ACCESS.2019.2919385

Huang, S. (2011). Identifying Alzheimer's disease-related brain regions from multi-modality neuroimaging data using sparse composite linear discrimination analysis. *Advances in Neural Information Processing Systems, 24*.

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065). https://doi.org/10.1098/rsta.2015.0202

Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal data fusion: An overview of methods, challenges, and prospects. *Proceedings of the IEEE, 103*(9), 1449–1477. https://doi.org/10.1109/JPROC.2015.2460697

Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Feng, D., & Fulham, M. J. (2015). ADNI, multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Biomedical Engineering, 62*(4), 1132–1140. https://doi.org/10.1109/tbme.2014.2372011

Liu, J., Zeng, D., Guo, R., Lu, M., Wu, F. X., & Wang, J. (2021). MMHGE: Detecting mild cognitive impairment based on multi-atlas multi-view hybrid graph convolutional networks and ensemble learning. *Cluster Computing, 24*(1), 103–113. https://doi.org/10.1007/s10586-020-03199-8

Marvasti, M. A., Poghosyan, A. V., Harutyunyan, A. N., & Grigoryan, N. M. (2013). Pattern detection in unstructured data: An experience for a virtualized IT infrastructure. *Proceedings of the 2013 IFIP/IEEE International Symposium on Integrated Network Management*, 1048–1053.

Mehmood, A., Yang, S., Feng, Z., Wang, M., Ahmad, A. L. S., Khan, R., Maqsood, M., & Yaqub, M. (2021). A transfer learning approach for early diagnosis of alzheimer's disease on MRI images. *Neuroscience, 460*, 43–52. https://doi.org/10.1016/j.neuroscience.2021.01.002

Mishra, S., & Misra, A. (2018). Structured and unstructured big data analytics. In *International conference on current trends in computer, electrical, electronics and communication, CTCEEC 2017* (pp. 740–746). India: Institute of Electrical and Electronics Engineers Inc.. https://doi.org/10.1109/CTCEEC.2017.8454999. http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8440575

Moore, P. J., Lyons, T. J., & Gallacher, J. (2019). Random forest prediction of Alzheimer's disease using pairwise selection from time series data. *PLoS One, 14*(2). https://doi.org/10.1371/journal.pone.0211558

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., & Tohka, J. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *NeuroImage, 104*, 398–412. https://doi.org/10.1016/j.neuroimage.2014.10.002

Muhammed, M. N., & Thiyagarajan. (2021). Alzheimer's classification using dynamic ensemble of classifiers selection algorithms: A performance analysis. *Biomedical Signal Processing and Control, 68*. https://doi.org/10.1016/J.BSPC.2021.102729

Nanni, L., Brahnam, S., Salvatore, C., & Castiglioni, I. (2019). Texture descriptors and voxels for the early diagnosis of Alzheimer's disease. *Artificial Intelligence in Medicine, 97*, 19–26. https://doi.org/10.1016/j.artmed.2019.05.003

Nanni, L., Salvatore, C., Cerasa, A., & Castiglioni, I. (2016). Combining multiple approaches for the early diagnosis of Alzheimer's disease. *Pattern Recognition Letters, 84*, 259–266. https://doi.org/10.1016/j.patrec.2016.10.010

Ortiz, A., Munilla, J., Górriz, J. M., & Ramírez, J. (2016). Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *International Journal of Neural Systems*. . https://doi.org/10.1142/S0129065716500258

Pachynski, R. K., Kim, E. H., Miheecheva, N., Kotlov, N., Ramachandran, A., Postovalova, E., Galkin, I., Svekolkin, V., Lyu, Y., Zou, Q., Cao, D., Gaut, J., Ippolito, J. E., Bagaev, A., Bruttan, M., Gancharova, O., Nomie, K., Tsiper, M., Andriole, G. L., Ataullakhanov, R., & Hsieh, J. J. (2021). Single-cell spatial proteomic revelations on the multiparametric MRI heterogeneity of clinically significant prostate cancer. *Clinical Cancer Research, 27*(12), 3478–3490. https://doi.org/10.1158/1078-0432.CCR-20-4217

Pan, D., Zeng, A., Jia, L., Huang, Y., Frizzell, T., & Song, X. (2020). Early detection of alzheimer's disease using magnetic resonance imaging: A novel approach combining convolutional neural networks and ensemble learning. *Frontiers in Neuroscience, 14*. https://doi.org/10.3389/fnins.2020.00259

Qiu, S., Chang, G. H., Panagia, M., Gopal, D. M., Au, R., & Kolachalama, V. B. (2018). Fusion of deep learning models of MRI scans, mini–mental state examination, and logical memory test enhances diagnosis of mild cognitive impairment. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 10*, 737–749. https://doi.org/10.1016/j.dadm.2018.08.013

Qiu, S., Joshi, P. S., Miller, M. I., Xue, C., Zhou, X., Karjadi, C., Chang, G. H., Joshi, A. S., Dwyer, B., Zhu, S., Kaku, M., Zhou, Y., Alderazi, Y. J., Swaminathan, A., Kedar, S., Saint-Hilaire, M.-H., Auerbach, S. H., Yuan, J., Sartor, E. A., Au, R., & Kolachalama, V. B. (2020). Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain, 143*(6), 1920–1933. https://doi.org/10.1093/brain/awaa137

Rahim, N., Abuhmed, T., Mirjalili, S., El-Sappagh, S., & Muhammad, K. (2023). Time-series visual explainability for Alzheimer's disease progression detection for smart healthcare. *Alexandria Engineering Journal, 82*, 484–502. https://doi.org/10.1016/j.aej.2023.09.050

Rahim, N., Ahmad, J., Muhammad, K., Sangaiah, A. K., & Baik, S. W. (2018). Privacy-preserving image retrieval for mobile devices with deep features on the cloud. *Computer Communications, 127*, 75–85. https://doi.org/10.1016/j.comcom.2018.06.001

Rahim, N., El-Sappagh, S., Ali, S., Muhammad, K., Del Ser, J., & Abuhmed, T. (2023). Prediction of Alzheimer's progression based on multimodal deep-learning-based fusion and visual explainability of time-series data. *Information Fusion, 92*, 363–388. https://doi.org/10.1016/j.inffus.2022.11.028

Razzak, I., Naz, S., Ashraf, A., Khalifa, F., Bouadjenek, M. R., & Mumtaz, S. (2022). Mutliresolutional ensemble PartialNet for Alzheimer detection using magnetic resonance imaging data. *International Journal of Intelligent Systems, 37*(10), 6613–6630. https://doi.org/10.1002/int.22856

Ruiz, J., Mahmud, M., Modasshir, M., & Shamim Kaiser, M. (2020). 3D DenseNet ensemble in 4-Way classification of alzheimer's disease. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (pp. 85–96). United Kingdom: Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-59277-6_8. www.springer.com/series/55812241.

Sarraf, S., Sarraf, A., DeSouza, D. D., Anderson, J. A. E., & Kabia, M. (2023). OViTAD: Optimized vision transformer to predict various stages of alzheimer's disease using resting-state fMRI and structural MRI data. *Brain Sciences, 13*(2). https://doi.org/10.3390/brainsci13020260

Shi, J., Zheng, X., Li, Y., Zhang, Q., & Ying, S. (2018). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics, 22*(1), 173–183. https://doi.org/10.1109/JBHI.2017.2655720

Sivapriya, T. R., Kamal, A. R. N. B., & Thangaiah, P. R. J. (2015). Ensemble merit merge feature selection for enhanced multinomial classification in alzheimer's dementia. *Computational and Mathematical Methods in Medicine, 2015*. https://doi.org/10.1155/2015/676129

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Tanveer, M., Rashid, A. H., Ganaie, M. A., Reza, M., Razzak, I., & Hua, K. L. (2022). Classification of Alzheimer's disease using ensemble of deep neural networks trained through transfer learning. *Journal of Biomedical and Health Informatics, 26*(4), 1453–1463. https://doi.org/10.1109/JBHI.2021.3083274

Theofilas, P., Ehrenberg, A. J., Nguy, A., Thackrey, J. M., Dunlop, S., Mejia, M. B., Alho, A. T., Leite, R. E. P., Rodriguez, R. D., Suemoto, C. K., Nascimento, C. F., Chin, M., Medina-Cleghorn, D., Cuervo, A. M., Arkin, M., Seeley, W. W., Miller, B. L., Nitrini, R., Pasqualucci, C. A., … Grinberg, L. T. (2018). Probing the correlation of neuronal loss, neurofibrillary tangles, and cell death markers across the Alzheimer's disease Braak stages: A quantitative study in humans. *Neurobiology of Aging, 61*, 1–12. https://doi.org/10.1016/j.neurobiolaging.2017.09.007

Walter, T., Shattuck, D. W., Baldock, R., Bastin, M. E., Carpenter, A. E., Duce, S., Ellenberg, J., Fraser, A., Hamilton, N., Pieper, S., Ragan, M. A., Schneider, J. E., Tomancak, P., & Hériché, J.-K. (2010). Visualization of image data from cells to organisms. *Nature Methods, 7*(S3), S26. https://doi.org/10.1038/nmeth.1431

Wang, C., Saar, V., Leung, K. L., Chen, L., & Wong, G. (2018). Human amyloid β peptide and tau co-expression impairs behavior and causes specific gene expression changes in Caenorhabditis elegans. *Neurobiology of Disease, 109*, 88–101. https://doi.org/10.1016/j.nbd.2017.10.003

Wang, H., Shen, Y., Wang, S., Xiao, T., Deng, L., Wang, X., & Zhao, X. (2019). Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing, 333*, 145–156. https://doi.org/10.1016/j.neucom.2018.12.018

Weng, W., & Zhu, X. (2021). INet: Convolutional networks for biomedical image segmentation. *IEEE Access, 9*, 16591–16603. https://doi.org/10.1109/ACCESS.2021.3053408

Xu, L., Wu, X., Chen, K., & Yao, L. (2015). Multi-modality sparse representation-based classification for Alzheimer's disease and mild cognitive impairment. *Computer Methods and Programs in Biomedicine, 122*(2), 182–190. https://doi.org/10.1016/j.cmpb.2015.08.004

Xu, Z., Zhang, Q., Hao, F., Ren, Z., Kang, Y., & Cheng, J. (2021). VGG-CAE: Unsupervised visual place recognition using VGG16-Based convolutional autoencoder. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (pp. 91–102). China: Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-88007-1_8

Yiğit, A., Baştanlar, Y., & Işık, Z. (2022). Dementia diagnosis by ensemble deep neural networks using FDG-PET scans. *Signal, Image and Video Processing, 16*(8), 2203–2210. https://doi.org/10.1007/s11760-022-02185-4

Yu, L., Xiang, W., Fang, J., Phoebe Chen, Y. P., & Zhu, R. (2022). A novel explainable neural network for Alzheimer's disease diagnosis. *Pattern Recognition*, 131. https://doi.org/10.1016/j.patcog.2022.108876

Zhang, D., & Shen, D. (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage, 59*(2), 895–907. https://doi.org/10.1016/j.neuroimage.2011.09.069

Zhang, Y., Wang, S., Xia, K., Jiang, Y., & Qian, P. (2021). Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion. *Information Fusion, 66*, 170–183. https://doi.org/10.1016/j.inffus.2020.09.002

Zhu, Y., Kim, M., Zhu, X., Kaufer, D., & Wu, G. (2021). Long range early diagnosis of Alzheimer's disease using longitudinal MR imaging data. *Medical Image Analysis, 67*, Article 101825. https://doi.org/10.1016/j.media.2020.101825

Wang, S., Shen, Y., Chen, W., Xiao, T., & Hu, J. (2017). Automatic recognition of mild cognitive impairment from MRI images using expedited convolutional neural networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 373–380). China: Springer Verlag. https://doi.org/10.1007/978-3-319-68600-4_43. http://springerlink.com/content/0302-9743/copyright/2005/ 10613.

Further reading

Asl, E. H., et al. (2018). Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network. *Front. Biosci. - Landmark, 23*(3), 584–596. https://doi.org/10.2741/4606

Beekly, D. L., et al. (Jul. 2007). The National Alzheimer's Coordinating Center (NACC) database: The uniform data set. *Alzheimer Disease and Associated Disorders, 21*(3), 249–258. https://doi.org/10.1097/WAD.0b013e318142774e

Bordin, V., Coluzzi, D., Rivolta, M. W., & Baselli, G. (2022). Explainable AI points to white matter hyperintensities for alzheimer's disease identification: A preliminary study. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2022-July*, 484–487. https://doi.org/10.1109/EMBC48229.2022.9871306

Camacho, M., et al. (2023). Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets. *NeuroImage Clin, 38*(April), Article 103405. https://doi.org/10.1016/j.nicl.2023.103405

Cherubini, A., et al. (2010). Combined volumetry and DTI in subcortical structures of mild cognitive impairment and Alzheimer's disease patients. *Journal of Alzheimer's Disease, 19*(4), 1273–1282. https://doi.org/10.3233/JAD-2010-091186

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical evaluation of gated recurrent neural networks on sequence modeling* (pp. 1–9).

Devanand, D. P., et al. (Mar. 2007). Hippocampal and entorhinal atrophy in mild cognitive impairment: Prediction of Alzheimer disease. *Neurology, 68*(11), 828–836. https://doi.org/10.1212/01.wnl.0000256697.20968.d7

Essemlali, A., St-Onge, E., Descoteaux, M., & Jodoin, P.-M. (2020). Understanding Alzheimer disease's structural connectivity through explainable AI. *Proc. Mach. Learn. Res., 121*, 217–229.

Gotkowski, K., Gonzalez, C., Bucher, A., & Mukhopadhyay, A. (2021). M3d-CAM: A PyTorch library to generate 3D attention maps for medical deep learning. In *Informatik aktuell, Springer Vieweg, Wiesbaden* (pp. 217–222). https://doi.org/10.1007/978-3-658-33198-6_52

Heckemann, R. A., et al. (Jun. 2011). Automatic morphometry in Alzheimer's disease and mild cognitive impairment. *NeuroImage, 56*(4), 2024–2037. https://doi.org/10.1016/J.NEUROIMAGE.2011.03.014

Hochreiter, S. (Nov. 1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int. J. Uncertainty, Fuzziness Knowlege-Based Syst., 6*(2), 107–116. https://doi.org/10.1142/S0218488598000094

Holzinger, A., et al. (Mar. 2022). Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Information Fusion, 79*, 263–278. https://doi.org/10.1016/j.inffus.2021.10.007

Il Suk, H., Lee, S. W., & Shen, D. (2016). Deep sparse multi-task learning for feature selection in Alzheimer's disease diagnosis. *Brain Structure and Function, 221*(5), 2569–2587. https://doi.org/10.1007/s00429-015-1059-y

Leming, M., Das, S., & Im, H. (2022). Construction of a confounder-free clinical MRI dataset in the Mass General Brigham system for classification of Alzheimer's disease. *Artificial Intelligence in Medicine, 129*(February), Article 102309. https://doi.org/10.1016/j.artmed.2022.102309

Likeman, M., et al. (2005). Visual assessment of atrophy on magnetic resonance imaging in the diagnosis of pathologically confirmed young-onset dementias. *Archives of Neurology, 62*(9), 1410–1415. https://doi.org/10.1001/archneur.62.9.1410

Malone, I. B., et al. (Apr. 2013). MIRIAD-public release of a multiple time point Alzheimer's MR imaging dataset. *NeuroImage, 70*, 33–36. https://doi.org/10.1016/j.neuroimage.2012.12.044

Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2010). Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults. *Journal of Cognitive Neuroscience, 22*(12), 2677–2684. https://doi.org/10.1162/jocn.2009.21407

Muddamsetty, S. M., Jahromi, M. N. S., Ciontos, A. E., Fenoy, L. M., & Moeslund, T. B. (2022). Visual explanation of black-box model: Similarity Difference and Uniqueness (SIDU) method. *Pattern Recognit, 127*, Article 108604. https://doi.org/10.1016/j.patcog.2022.108604

Muhammad, G., Alshehri, F., Karray, F., El Saddik, A., Alsulaiman, M., & Falk, T. H. (Dec. 2021). A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. *Information Fusion, 76*, 355–375. https://doi.org/10.1016/j.inffus.2021.06.007

Nelson, P. T., et al. (2018). The amygdala as a locus of pathologic misfolding in neurodegenerative diseases. *Journal of Neuropathology and Experimental Neurology, 77*(1), 2–20. https://doi.org/10.1093/jnen/nlx099

Rahim, N., El-Sappagh, S., Rizk, H., El-serafy, O. A., & Abuhmed, T. (2024). Information fusion-based Bayesian optimized heterogeneous deep ensemble model based on longitudinal neuroimaging data. *Applied Soft Computing, 162*(December 2023), Article 111749. https://doi.org/10.1016/j.asoc.2024.111749

Schmahmann, J. D. (May 2016). Cerebellum in Alzheimer's disease and frontotemporal dementia: Not a silent bystander. *Brain, 139*(5), 1314–1318. https://doi.org/10.1093/brain/aww064

Senanayake, U., Sowmya, A., & Dawes, L. (2018). Deep fusion pipeline for mild cognitive impairment diagnosis. *Proc. - Int. Symp. Biomed. Imaging, 2018-April*(Isbi), 1394–1397. https://doi.org/10.1109/ISBI.2018.8363832

Shad, H. A., et al. (2021). Exploring alzheimer's disease prediction with XAI in various neural network models. In *IEEE Region 10 annual international conference, Proceedings/TENCON, IEEE* (pp. 720–725). https://doi.org/10.1109/TENCON54134.2021.9707468

Syed, A. H., Khan, T., Hassan, A., Alromema, N. A., Binsawad, M., & Alsayed, A. O. (2020). An ensemble-learning based application to predict the earlier stages of Alzheimer's disease (AD). *IEEE Access, 8*, 222126–222143. https://doi.org/10.1109/ACCESS.2020.3043715

Tinauer, C., et al. (2022). Interpretable brain disease classification and relevance-guided deep learning. *Scientific Reports, 12*(1), 1–14. https://doi.org/10.1038/s41598-022-24541-7

Van Hoesen, G. W., Augustinack, J. C., Dierking, J., Redman, S. J., & Thangavel, R. (2000). The parahippocampal gyrus in Alzheimer's disease. Clinical and preclinical neuroanatomical correlates. *Annals of the New York Academy of Sciences, New York Academy of Sciences*, 254–274. https://doi.org/10.1111/j.1749-6632.2000.tb06731.x

Zhang, X., Han, L., Zhu, W., Sun, L., & Zhang, D. (2022). An explainable 3D residual self-attention deep neural network for joint atrophy localization and alzheimer's disease diagnosis using structural MRI. *IEEE J. Biomed. Heal. Informatics, 26*(11), 5289–5297. https://doi.org/10.1109/JBHI.2021.3066832