



OPEN Multimodal multitask learning for predicting MCI to AD conversion using stacked polynomial attention network and adaptive exponential decay

Ngoc-Huynh Ho¹, Yang-Hyung Jeong^{1✉} & Jahae Kim^{1,2}

Early identification and treatment of moderate cognitive impairment (MCI) can halt or postpone Alzheimer's disease (AD) and preserve brain function. For prompt diagnosis and AD reversal, precise prediction in the early and late phases of MCI is essential. This research investigates multimodal framework-based multitask learning in the following situations: (1) Differentiating early mild cognitive impairment (eMCI) from late MCI and (2) predicting when an MCI patient would acquire AD. Clinical data and two radiomics features on three brain areas deduced from magnetic resonance imaging were investigated (MRI). We proposed an attention-based module, Stack Polynomial Attention Network (SPAN), to firmly encode clinical and radiomics data input characteristics for successful representation from a small dataset. To improve multimodal data learning, we computed a potent factor using adaptive exponential decay (AED). We used experiments from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort study, which included 249 eMCI and 427 IMCI participants at baseline visits. The proposed multimodal strategy yielded the best c-index score in time prediction of MCI to AD conversion (0.85) and the best accuracy in MCI-stage categorization (83.19%). Moreover, our performance was equivalent to that of contemporary research.

Memory loss and cognitive decline are clinical symptoms of Alzheimer's disease (AD)¹, which is a progressive neurodegenerative sickness that affects the brain. It is the most prevalent cause of dementia in individuals over the age of 65. In 2021, over 55 million people live with dementia worldwide². AD accounts for 60 to 80 percent of all dementia cases according to reports³. AD is defined by the accumulation of beta-amyloid and tau proteins in the brain⁴, which hinder normal cognitive activities. This often emerges as alterations in memory, analytical thinking, perception, language, mood, and emotions, and eventually impairs physical control over the body. Several studies have been conducted in recent years in an attempt to identify early biomarkers that may be used to assess Alzheimer's disease risk prior to the onset of symptoms in a speedy and thorough manner^{5–7}.

Currently, mild cognitive impairment (MCI), which is a prodromal stage of Alzheimer's disease (AD), has attracted much attention because of the high likelihood that it will advance to AD. MCI is characterized by a minor but perceptible and quantifiable reduction in cognitive ability, including memory and reasoning abilities. An individual diagnosed with MCI may be at risk of acquiring Alzheimer's disease in the future, or the condition may be due to age-related cognitive deterioration, underscoring the necessity of early detection of the condition. Research has shown that the risk of late MCI (IMCI) conversion to AD is higher than that of early MCI (eMCI)⁸. Identification of potentially sensitive diagnostic indicators that change in response to illness development may aid the physician in making a diagnosis. If detected early in the course of MCI, individuals can significantly lower their risk of developing AD by approximately one-third with rehabilitation activities and medication⁹. Regrettably, sensitive indicators vary according to disease development¹⁰, and there are presently no definite diagnostic biomarkers or viable therapies for Alzheimer's disease¹¹.

There is widespread agreement on the importance of and benefits of early detection of the condition. Numerous neurologists and medical researchers are currently devoting significant work to developing procedures for

¹Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju 61186, South Korea. ²Department of Nuclear Medicine, Chonnam National University Hospital, Gwangju 61469, South Korea. ✉email: hjyang@jnu.ac.kr

early diagnosis of AD, with consistently encouraging findings¹². In recent decades, several studies have been proposed for automatic detection of AD^{6,13–16}. Various neuroimaging signals such as magnetic resonance imaging (MRI)^{17,18}, functional magnetic resonance imaging (fMRI)^{19,20}, positron emission tomography (PET)^{21,22}, electroencephalography (EEG)^{23–25}, and magnetoencephalography (MEG)^{26,27} have been investigated to determine if there are any anomalous clustering coefficients or distinctive path lengths in the brain networks of AD patients. The ability to diagnose and categorize MCI at an early stage helps physicians to make better informed judgments about clinical intervention and treatment planning at a later stage, which has a significant influence on cost-effectiveness of long-term care services²⁸. However, only a few studies on the features of brain networks in MCI patients have explored the properties of brain networks at different phases since brain abnormalities are so subtle^{29,30}.

Feature fusion strategies have gained significant attention in the medical field for their ability to integrate diverse information sources and enhance diagnostic accuracy. Multimodal image fusion techniques, as highlighted by Wang et al.³¹, enable the combination of complementary information from different imaging modalities, such as MRI, CT, and PET, to improve disease interpretation. Deep learning-based approaches, exemplified by Li et al.³², leverage feature fusion to enhance medical diagnosis by effectively integrating multimodal information. Moreover, Tong et al.³³ demonstrated the potential of feature fusion in Alzheimer's disease diagnosis, utilizing hybrid weighted multiple kernel learning to integrate clinical assessments, genetic profiles, and neuroimaging data. By leveraging feature fusion strategies, medical researchers and practitioners can harness the power of multiple data modalities to improve disease detection, localization, and overall patient outcomes. In this paper, we will present a straightforward and efficient fusion equation designed to combine multimodal data.

In this paper, we proposed a novel attention-based mechanism for multimodal multitask learning of AD progression. We employed MRI scans and clinical data to distinguish eMCI from IMCI while also predicting the time to AD conversion. We extract three brain regions, in particular, such as gray matter (GM), white matter (WM), and Cerebrospinal Fluid (CSF) from T1-MRI image using the statistical parametric mapping (SPM) toolbox (<https://www.fil.ion.ucl.ac.uk/spm/>). Then, we estimated the texture and shape features from the masked regions using the PyRadiomics toolbox (<https://pyradiomics.readthedocs.io/>). Consecutively, we introduced a novel deep learning (DL) approach called stacked polynomial attention network (SPAN) for learning a more accurate approximation basis for all polynomials of bounded degree³⁴. Two branches with SPAN and dropout layers are employed to encode the clinical and radiomics representations, and the prominent characteristics of both branches are effectively merged using our proposed formula, adaptive exponential decay (AED). The composite representation is scaled using a series of fully connected layers. Finally, the probability of IMCI and the hazard rate of AD conversion are calculated simultaneously as multitask learning. The main contributions of our studies are as follows:

- We proposed a multimodal multitask learning based approach to synchronously classify eMCI and IMCI stages in AD patients and predict the time to AD conversion from these MCI patients for early diagnosis of AD. To the best of our knowledge, this is the first study to integrate two tasks: the categorization of the MCI stage and the prediction of the period from the MCI stage to the onset of AD.
- Technically, we proposed a novel attention-based mechanism, SPAN, to learn data representations from finite sample datasets in a practical and effective manner.
- We carried out analysis of the exploratory investigation of radiomics characteristics for predicting the course of AD in three brain areas (GM, WM, and CSF).
- We proposed the use of a decay factor, AED, to aid in the acquisition of the dominant representation across modalities.
- We experimented on a public dataset and employed cross-validation to show the generalization of the proposed system. Several aspects of disease analysis were exploited to understand the course of AD better.

Results

Study participants. To evaluate the efficiency of the proposed framework, we used the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort, which includes diagnosis of 1,737 patients (ages 54.5 to 98.6 years) from 2005 to 2017. According to our scopes, which focus on the tasks of MCI-stage classification and time-to-AD prediction, we only selected patients who are diagnosed as either eMCI or IMCI at baseline timepoint. Furthermore, we cleaned up the raw data through removing timepoints that had been either duplicated or had implausible measurements, and we screened out irreversible individuals who had altered their condition from AD to MCI or from MCI to cognitive normal (CN) during the course of the study's history.

Bases on given patients' IDs, we manually collected their corresponding MRI scans from the ADNI site. In total, we obtained 249 eMCI and 427 IMCI patients at baseline diagnosis. Table 1 presents the collected data statistic from the ADNI cohort for two groups of eMCI and IMCI patients. There are significant differences between the two groups in terms of age, Clinical Dementia Rating Scale-Sum of Boxes (CDRSB), Mini Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale-13 (ADAS13), Rey Auditory Verbal Learning Test (RAVLT), Functional Activities Questionnaire (FAQ), volumetric and PET biomarkers ($p < 0.05$).

To determine time-to-AD conversion, for uncensored patients, we assume that the conversion time is the time span between the baseline diagnosis and the first observation of AD. When considering the censored patients, the conversion time is calculated by adding the delaying time to their most recent visits. The data distribution and conversion time are visualized in Fig. 1. For the eMCI cases, the censored patients outnumber the uncensored ones. In Fig. 1b, the uncensored patient event occurs when the patient is diagnosed with AD, while the censored patient event occurs at the end of the study, which is the last observation on this patient. The details of data

Variable	eMCI (n=249)	lMCI (n=427)	p value
Male, gender, n (%)	141 (56.6%)	262 (61.4%)	0.259*
Age, years, mean±std	71.8 ± 7.3	74.5 ± 7.3	4.1e-6 [⊥]
Education, years, mean±std	15.9 ± 2.7	15.9 ± 3	0.959 [⊥]
CDRSB, n (%)			0.003*
≤ 1	136 (54.6%)	181 (42.4%)	
MMSE, mean±std	28.2 ± 1.6	27.2 ± 1.8	2.2e-13 [⊥]
ADAS13, mean±std	13.0 ± 5.3	18.6 ± 6.7	6.7e-21 [⊥]
RAVLT, mean±std	5.2 ± 2.4	3.5 ± 2.4	2.5e-19 [⊥]
FAQ, mean±std	2.0 ± 2.9	3.7 ± 4.5	0.001 [⊥]
Volumetric biomarker, ml, median (IQR)			
Ventricles, × 10 ³	31.2 [19.6, 46.9]	39.1 [27.1, 57.7]	9.0e-6 [⊥]
Hippocampus, × 10 ³	7.2 [6.6, 7.8]	6.5 [5.8, 7.2]	4.6e-5 [⊥]
Entorhinal, × 10 ³	3.8 [3.3, 4.2]	3.4 [2.8, 3.9]	4.5e-4 [⊥]
Fusiform, × 10 ³	18.7 [16.8, 20.4]	16.9 [15.0, 18.4]	4.0e-7 [⊥]
Middle Temporal, × 10 ³	20.5 [19.1, 22.5]	19.0 [17.1, 21.0]	5.5e-5 [⊥]
Whole Brain, × 10 ⁵	10.6 [9.9, 11.3]	10.0 [9.3, 10.9]	1.3e-8 [⊥]
PET biomarker, median (IQR)			
FDG	1.3 [1.21, 1.35]	1.2 [1.13, 1.3]	2.5e-3 [⊥]
AV45	1.1 [1.01, 1.31]	1.3 [1.07, 1.5]	0. [⊥]

Table 1. Data statistic of the collected ADNI cohort in this study. IQR: Interquartile Range. * Chi-square test. [⊥] T-test. [⊥] Mann-Whitney U test

distributions are presented in the *Supplementary Materials*, section *Data distribution*. In addition, the model settings can be found in the *Supplementary Materials*, section *Experimental settings*.

For quantitative evaluation, we use c-index score (CI), Brier score (BS), and mean absolute error (MAE) as criteria of time prediction task, while using accuracy (Acc), average precision (AP), precision (Pre), recall (Rec), F1-score (F₁), and area under receiver operating characteristic curve (AUC) as criteria of classification task. For generalization, experiments are performed by 5-fold cross validation. The details of evaluation metrics are expressed in the *Supplementary Materials*, section *Evaluation metrics*.

Comparison to conventional studies. In this section, we present a comprehensive analysis of the performance of our proposed model and the existing research in terms of prediction of time-to-AD conversion and MCI-stage classification, as shown in Tables 2 and 3. Table 2 highlights the performance of various methods in prediction time-to-AD conversion, including our proposed approach, along with the modalities, data size, and evaluation metrics employed. The study by Polsterl et al.³⁵ utilized 3D hippocampus data along with clinical data to predict the conversion time to AD. Their approach achieved a CI of 0.803. Lu et al.³⁶ focused on MRI and genetic data as their modalities for conversion time prediction. They reported a CI of 0.681, indicating moderate predictive performance. The BS, a measure of calibration, was reported as 0.147, suggesting some room for improvement in calibration. Nakagawa et al.³⁷ employed gray matter (GM), patient age, and Mini-Mental State Examination (MMSE) scores as input features for their prediction model. Their approach achieved a CI of 0.83 when evaluating on both NC and MCI patients, and a CI of 0.75 when using MCI set only. Ho et al.³⁸ investigated the use of demographics and brain biomarkers for conversion time prediction. Their approach achieved a CI of 0.804, similar to the performance reported by³⁵. The BS value for this study was reported as 0.153, indicating good calibration. In comparison, our proposed method utilized radiomics features extracted from MRI scans along with clinical data for conversion time prediction of MCI patients. Our approach achieved a significantly higher CI of 0.846, indicating improved predictive performance compared to the other methods discussed. Additionally, the BS value for our approach was reported as 0.132, suggesting good calibration and accurate estimation of conversion time.

Table 3 summarizes the Acc and AUC values achieved by various methods for the classification of eMCI and late lMCI. Suk et al.³⁹ utilized MRI and PET modalities for the classification of eMCI and lMCI, achieving an accuracy of 75.92% and an AUC of 0.75. This approach performed reasonably well in distinguishing between the two classes, although the AUC suggests room for improvement in capturing the discriminatory power of the model. Nozadi et al.⁴⁰ also employed MRI and PET data for classification, with a dataset of 164 eMCI and 189 lMCI samples. Their approach achieved an accuracy of 65.2%, indicating moderate performance in distinguishing between the two classes. Jie et al.⁴¹ focused on resting-state functional MRI (rs-fMRI) as their modality for classification. Their approach demonstrated a higher accuracy of 78.8% and an AUC of 0.78. These results indicate better discrimination between eMCI and lMCI compared to the previous approaches utilizing MRI and PET. Zhang et al.⁴² also utilized rs-fMRI data for classification and achieved an accuracy of 83.87% and an impressive AUC of 0.9, demonstrating superior performance in accurately distinguishing between the two

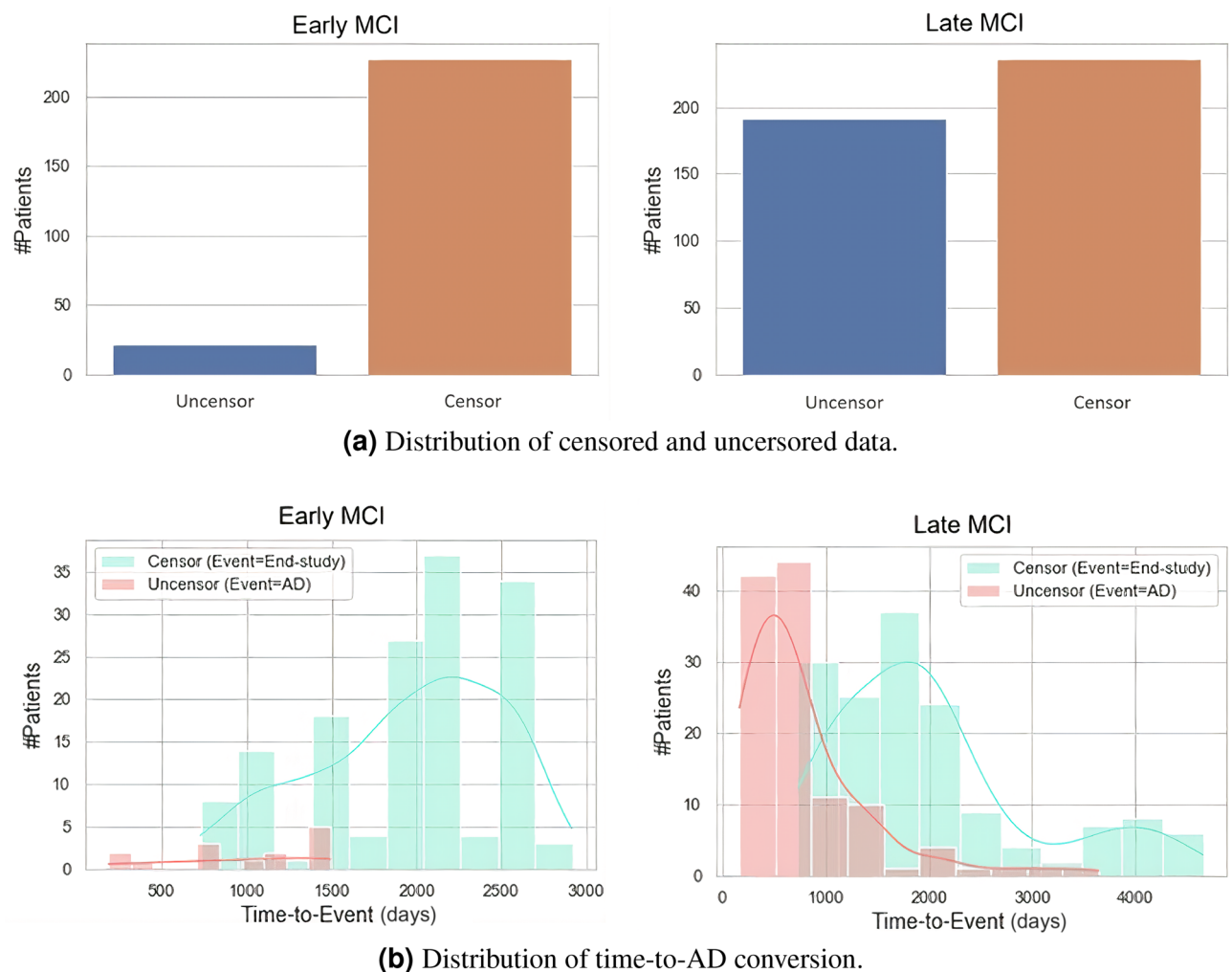


Figure 1. Data distribution.

Approach	Modality	Dataset	Prediction criteria	
			CI↑	BS↓
Polsterl et al. ³⁵	3D Hippocampus Clinical	397 MCI patients - Censoring rate: unknown	0.803	—
Lu et al. ³⁶	MRI Genetic	173 MCI patients - Censoring rate: 89%	0.681	0.147
Nakagawa et al. ³⁷	GM (MRI) Age, MMSE	2, 142 patients (NC, MCI) - Censoring rate: 75% 1, 211 MCI patients - Censoring rate: 58.6%	0.83 0.75	—
Ho et al. ³⁸	Demographics Biomarkers	1, 334 patients (NC, MCI, AD) - Censoring rate: 74%	0.804	0.153
Ours (SPAN-AED)	Radiomics (MRI) Clinical	249 eMCI patients - Censoring rate: 92% 427 lMCI patients - Censoring rate: 55.3%	0.846	0.132

Table 2. Performance for conversion-time-to-AD prediction in terms of CI and BS; where ↑ : higher is better; ↓ : lower is better. The boldface values indicate the best performance.

Approach	Modality	Dataset	Classification criteria	
		(eMCI / lMCI)	Acc↑	AUC↑
Suk et al. ³⁹	MRI & PET	128 / 76	75.92	0.75
Nozadi et al. ⁴⁰	MRI & PET	164 / 189	65.2	–
Jie et al. ⁴¹	rs-fMRI	56 / 43	78.8	0.78
Zhang et al. ⁴²	rs-fMRI	33 / 29	83.87	0.9
Mehmood et al. ¹⁸	GM (MRI)	70 / 70	83.7	–
Cui et al. ⁴³	MRI	45 / 51	76.13	–
Ours (SPAN-AED)	Radiomics (MRI) Clinical	249 / 427	83.19	0.91

Table 3. Performance for early MCI and late MCI classification in term of Acc and AUC; where ↑ : higher is better. The boldface values indicate the best performance.

classes. Nevertheless, their study was performed using an extremely limited dataset comprising only 33 eMCI and 29 lMCI patients, which makes it susceptible to issues of overfitting and lack of generalizability. Mehmood et al.¹⁸ focused on gray matter (GM) features derived from MRI scans for classification. Using a small dataset and fixing classes to be balanced with 70 eMCI and 70 lMCI samples, their approach achieved an accuracy of 83.7%. By training model on 2D images, they presented slower training and inference time compared to using input features as vector. Cui et al.⁴³ achieved an accuracy of 76.13% when evaluating with a small dataset of 45 eMCI and 5 lMCI patients. In comparison, our proposed method leveraged radiomics features extracted from MRI scans along with clinical data for eMCI and lMCI classification. Our approach achieved an accuracy of 83.19% and an AUC of 0.91, demonstrating robust performance in accurately distinguishing between the two classes. The large dataset size of 249 eMCI and 427 lMCI samples further enhances the reliability of our results. In general, the experimental results indicate that our proposed method demonstrates promising approach for multitask learning of predicting time-to-AD conversion and classifying MCI-stage.

Performance on prediction of MCI to AD conversion. We individually investigated clinical and radiomics characteristics to determine the effectiveness of our proposed model compared to the unimodal approach. In the *Performance on combination of radiomics features* subsection of the *Supplementary Materials*, we demonstrated that the optimal combination of [CSF][texture] features was utilized as the radiomics input for our proposed model. In comparison to the study conducted by Ho et al.³⁸, we replaced the SPAN module with a residual-attention (RA) module for feature encoding. Our SPAN algorithm exhibited superior performance in predicting conversion time-to-AD with clinical features, achieving a higher CI (0.82 compared to 0.8) and lower MAE (454 days compared to 510 days). Additionally, it slightly improved the performance of the MCI-stage classification task. Notably, the SPAN encoder outperformed the RA encoder in both tasks, resulting in a reduction of 56 days in MAE and an improvement of 0.89% in accuracy. Instead of the proposed AED fusion strategy, we utilized concatenation (Concat) for comparison. Experimental results showed that our proposed AED approach outperformed traditional concatenation in representation fusion for both tasks. It yielded a 0.02 increase in CI (using SPAN encoder), a reduction of up to 59 days in MAE (using RA encoder), a reduction of up to 8 days in MAE (using SPAN encoder), and an increase of up to 0.67% in accuracy (using RA encoder and SPAN encoder). The detailed results of the prediction and classification tasks for MCI to AD conversion are presented in Table 4. In conclusion, multimodal approaches surpassed the use of unimodal approaches, both for clinical and radiomics features. Moreover, the SPAN module consistently outperformed the RA module. Integrating SPAN with AED further enhanced performance compared to utilizing SPAN with the Concat operation.

Modality	Approach	Prediction Criteria			Classification criteria					
		CI↑	BS↓	MAE↓	Acc↑	AP↑	Pre↑	Rec↑	F ₁ ↑	AUC↑
Clinical	RA	0.8	0.16	510	80.83	91.22	80.67	80.54	80.24	0.88
	SPAN	0.82	0.16	454	81.16	91.53	81.32	81.77	81.53	0.88
Radiomics	RA	0.67	0.2	725	77.21	92.37	76.47	78.23	77.54	84.42
	SPAN	0.67	0.19	669	78.1	92.64	77.81	78.65	78.13	85.46
Multimodal	RA-Concat	0.82	0.16	445	82.35	94.63	80.94	81.86	81.32	0.89
	RA-AED	0.82	0.14	386	83.02	93.68	82.25	83.44	82.51	0.89
	SPAN-Concat	0.83	0.13	355	82.89	94.14	83.23	82.36	82.23	0.9
	Ours(SPAN-AED)	0.85	0.13	347	83.19	94.76	82.31	82.85	82.47	0.91

Table 4. Performance for conversion time-to-AD prediction and eMCI vs lMCI classification for unimodal and multimodal approaches; where ↑ : higher is better; ↓ : lower is better. The boldface values indicate the best performance.

In addition, further analysis of performance on combination of radiomics features and visualization of time-to-AD conversion are described in the *Supplementary Materials*, section *Ablation studies*.

Discussion

Recent neuroimaging studies revealed that individuals diagnosed with MCI and AD have considerable disruption in either the structural network or the functional network when compared to a healthy control group^{17,44}. Few studies have investigated the features of whole brain networks in patients with MCI at various stages of the disease. Zhang et al.⁴² utilized the graph theory to measure the relationship between changes in the brain network connectivity from the resting-state fMRI. Then, the support vector machine (SVM) was used to distinguish eMCI from IMCI at different frequency bands, and achieved the best performance in slow-5 band with a 83.87% accuracy. Transfer learning approaches are usually used to overcome privacy and cost issues for a massive quantity of annotated data, which entails applying a pre-trained model to new problems using a smaller dataset. By taking the advantage of these facts, Mehmood et al.¹⁸ developed a layer-wise transfer learning model based on VGG architecture family⁴⁵ to segregates between eMCI and IMCI and achieved a 83.72% accuracy. Cui et al.⁴³ proposed two-stage algorithm based on particle swarm optimization (PSO) for removing redundant features and adaptive LASSO logistic regression model for selecting the most relevant features to predict AD stages. Experimental results have been shown a 76.13% accuracy on stable MCI (sMCI) vs converted MCI (cMCI) patients.

A survival analysis is a type of statistical study that examines time-to-event data, which describes the period between a time origin and an endpoint of particular interest⁴⁶. Polsterl et al.³⁵ proposed a wide and deep neural network for survival analysis that learns to detect individuals who are at a high risk of advancing to AD using information from 3D hippocampal geometry and tabular clinical data. According to their findings, tabular clinical makers with a median c-index of 0.750 are already good predictors of conversion from MCI to AD. In addition, in the hippocampus volume, the median c-index climbed to 0.803 when the hippocampus volume was included. Nakagawa et al.³⁷ discovered a deep learning method-based survival analysis could be used to assess the likelihood that an individual will get AD over a particular period of time. They approached the survival problem in a unique way and demonstrated encouraging results across many cohorts. Ho et. al.³⁸ proposed a modification of DeepSurv architecture⁴⁷, called RASurv to analyze the time-to-AD conversion for both cognitive normal and MCI patients. Their model achieved a competitive performance to other methods with a c-index score of 0.804.

The difficulty of precisely determining when an individual transitioned to AD can be attributed to the no studies on the topic of prediction of time-to-AD conversion. Typically, the occurrence happens before an individual is diagnosed as AD. However, we usually assume that the event occurs at the timepoint that the patient is diagnosed as conversion from MCI to AD to alleviate the problem. The above-conventional study focused exclusively on single tasks, despite the possibility of a correlation between MCI phases and time-to-AD conversion. In general, an eMCI patient has a lower risk of developing AD within a short period of time than an IMCI patient. As a result, it is essential to master two tasks concurrently: MCI-stage classification and conversion-time-to-AD prediction. In addition, the criteria for eMCI and IMCI can be found in the *Supplementary Materials*, section *Criteria for the MCI stages*.

This study presented a novel framework of multimodal multitask learning to discriminate eMCI patients from IMCI patients and forecast conversion time till the onset of Alzheimer's disease. The proposed model derived features from clinical representations (which include patient information, cognitive measurements, and biomarkers) as well as radiomics representations (which are estimated from brain MRI). The SPM program was used to normalize brain MRI dimension and segment three different brain regions: the GM, WM, and CSF, in particular. These regions' masks combined with brain image were used to determine the shape and texture of radiomics characteristics with the PyRadiomics program. We proposed SPAN (stacked polynomial attention network) to effectively and reliably capture the approximation basis for all polynomials of constrained degree. The clinical and radiomics characteristics were supplied into two branches of SPAN and dropout series, which were then used to encode the appropriate information in the patient's medical record. After that, we constructed an adaptive exponential decay (AED) factor to combine the encoded representations from both branches together. We evaluated the proposed model on the ADNI cohort that overcomes the state-of-the-art performance.

However, it is essential to obtain radiomics characteristics from MRI scans to lower the dimension of the 3D images due to the large batch size required for ranking optimization in this study. However, the performance of radiomics characteristics is significantly low when compared to clinical data as shown in Table 1. This means that radiomics features may contribute less to the multi-model and may potentially introduce bias into the overall network. As a result, in future studies, we will examine other strategies for extracting more robust representations of 3D images. Furthermore, we analyzed global brain regions (such as the GM, WM, and CSF) in this study, despite the fact that there are critical areas (such as frontal lobe, motor cortex, sensory cortex, parietal lobe, occipital lobe, and temporal lobe, etc.) that often influence AD conversion. Therefore, future studies will include examining the relationships between brain areas.

Overall, we firmly believe that our study holds important value in the field of AD diagnosis and understanding. Our proposed multimodal multitask learning approach, attention-based mechanism, and exploratory investigation of radiomics characteristics provide valuable insights and potential avenues for early diagnosis and improved understanding of the course of AD. By integrating tasks, improving data representations, and incorporating multimodal information, we aim to advance the field's understanding of AD progression and contribute to the development of more effective diagnostic tools.

Methods

We developed a new paradigm for identifying MCI stages and predicting time-to-AD conversion using clinical and radiomics characteristics. First, we preprocess the raw clinical data and estimate radiomics features from MRI scans. Then, we encode the clinical and radiomics representations using a succession of SPAN and dropout layers. The AED algorithm efficiently fuses the two branches' prominent characteristics predict the probability of eMCI vs IMCI phases and the hazard rate of AD conversion. The overall process is shown in Fig. 2. *Note that this article does not contain any studies involving animals or human participants performed by any of the authors.*

Preprocessing and feature extraction. Numerous studies on Alzheimer's disease dementia make use of information obtained through expensive and invasive techniques such as brain imaging or spinal taps to predict the risk of getting Alzheimer's disease dementia or fast cognitive decline in the future. A low-cost and noninvasive approach to studying the evolution of Alzheimer's disease dementia might be based on clinical data (e.g., demographics, vital signs, medicines, laboratory data, vital signs, and current medical problems). Because clinical data can be supplied in a variety of forms, it is necessary to do data preprocessing and transformation prior to training a model on clinical data. In this study, we perform one-hot encoding for transforming categorical data and z-score normalization for infinitive numerical values and maximum normalization for limited numerical values, excluding volumetric biomarkers, which are scaled by dividing the total intracranial volume (ICV) of each individual. Since clinical data commonly appear with missing data in medical studies when the value of the variables of interest is not measured or recorded for all of the participants in the sample, we utilize the Multiple Imputation by Chained Equations (MICE) algorithm⁴⁸ to impute the missing values.

Radiomics, which is based on the high-dimensional quantification of medical scans and enables the retrieval of more precise features than standard visual interpretation, can reveal information for treatment interventions^{49,50}. There has been some investigation into the use of radiomics in identifying the progression of AD^{51,52}. These investigations revealed that radiomics biomarkers can be used to classify individuals with MCI who are at high risk of developing Alzheimer's disease in the future. Furthermore, radiomics biomarkers in combination with clinical analysis can vastly enhance the prediction accuracy of MCI to AD. To extract the radiomics features, we first utilize the "Normalization" module of SPM toolbox to scale the intensity and space of the three-dimension MRI image since the brain structure varies from person to person. Next, we segment the normalized brain into

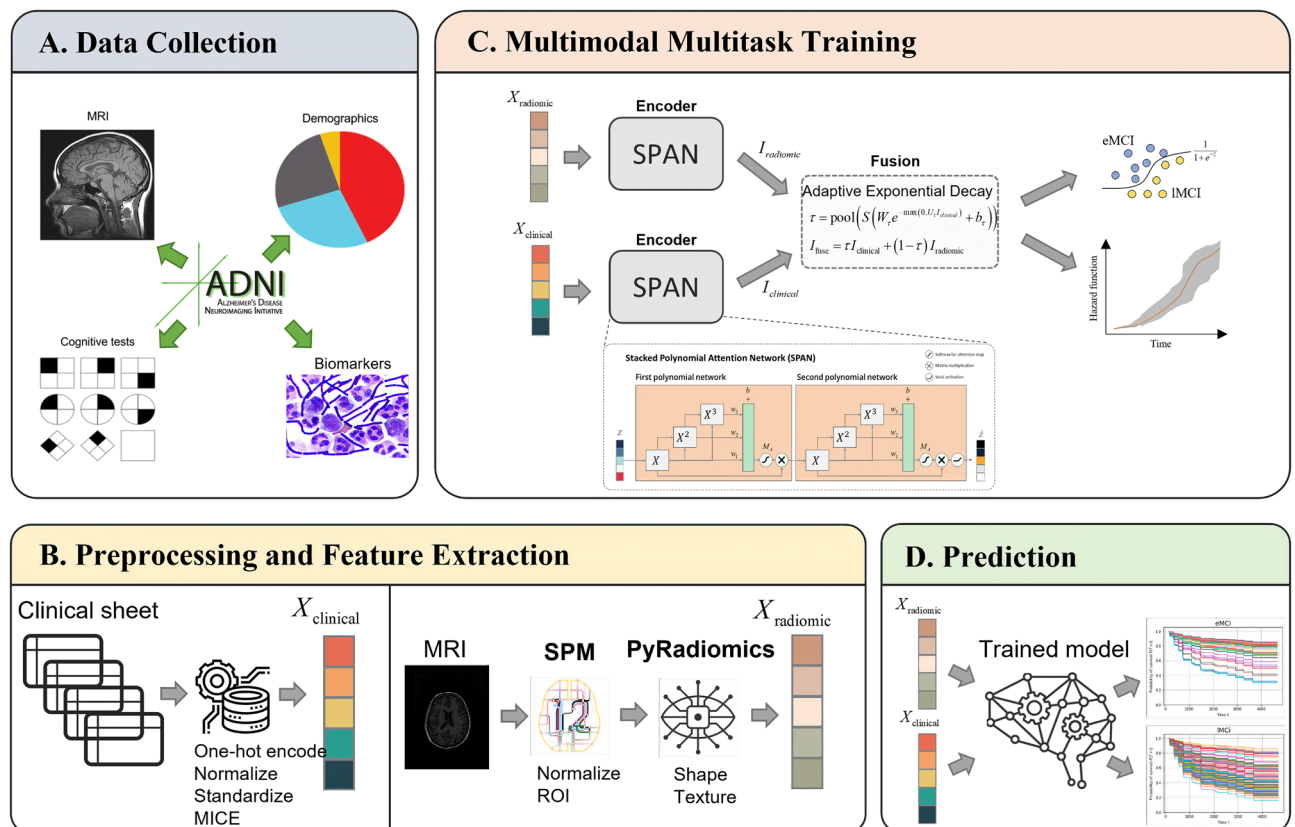


Figure 2. Overall process of the proposed method for MCI-stage classification and conversion-time-to-AD prediction. (A) Collecting data from ADNI cohort, (B) preprocessing clinical data and extracting radiomics features from MRI images, (C) training multimodal multitask model with the proposed SPAN network and AED fusion module, and (D) predicting time-to-AD conversion for both eMCI and IMCI patients.

three regions such as GM, WM, and CSF using the “Segmentation” module. Then, we extract the various types of radiomics features, which can be divided into shape and texture groups, using the PyRadiomics tool.

In our approach, we first normalize and standardize all features to ensure that features from different scales or with varying distributions were placed on a comparable scale, that prevent any single feature from dominating the learning process and promote fair contributions from all features. Next, we concatenate all features of each type of representation, namely radiomics and clinical, to create a single feature vector for each representation. This concatenation step ensures that all relevant information from the respective feature sets is preserved. The details of clinical data preprocessing and radiomics extraction can be found in the *Supplementary Materials*, section *Clinical and radiomics features preprocessing*.

Stacked polynomial attention network (SPAN). Recent years have seen an increase in the use of attention mechanisms to not only improve the performance but also the explainability of deep learning techniques. Initially, the attention mechanism was mostly employed to describe sequence dependencies independent of their real distances^{53,54}. Abd Hamid et al.⁵⁵ used an attention mechanism and a global average pooling (GAP) layer to extract the most prominent information from an MRI image for the purpose of differentiating between AD states. In previous study⁵⁶, researchers stated that the stacked deep polynomial network (S-DPN) can enhance the representation performance of the retrieved characteristics and held promise for the neuroimaging-based AD diagnosis. Based on these findings, we developed a novel attention mechanism based on S-DPN and dubbed the stacked polynomial attention network (SPAN) for exploiting attended representation from constrained indeterminates. Given an input feature Z , the sequential expressions of the first polynomial network of the SPAN module are as follows:

$$\begin{aligned} X_1^{(1)} &= W_Z^{(1)} Z + b_Z^{(1)} \\ X_2^{(1)} &= X_1^{(1)} \times X_1^{(1)} \\ &\dots \\ X_n^{(1)} &= X_{n-1}^{(1)} \times X_1^{(1)} \\ M_A^{(1)} &= \sigma \left(w_n^{(1)} X_n^{(1)} + \dots + w_1^{(1)} X_1^{(1)} + b^{(1)} \right) \\ \hat{Z}^{(1)} &= X_1^{(1)} \times M_A^{(1)} \end{aligned} \quad (1)$$

where $X_1^{(1)}, X_2^{(1)}, \dots, X_n^{(1)}$ represent indeterminates of polynomial function with n degrees of the first network, $W_Z^{(1)}, b_Z^{(1)}, w_1^{(1)}, \dots, w_n^{(1)}, b^{(1)}$ represent trainable parameters, $\sigma(\cdot)$ is softmax function for generating attention map $M_A^{(1)}$, and $\hat{Z}^{(1)}$ represents the attended representation. $\hat{Z}^{(1)}$ is then fed to the second polynomial network, which is similar to the first one, to stack up feature representation and yield a better and deeper structure. The scaled exponential linear unit (SELU) activation function⁽⁵⁷⁾ is then used to add non-linearity to the neural network. The sequence of the second polynomial network is expressed as follows:

$$\begin{aligned} X_1^{(2)} &= W_Z^{(2)} \hat{Z}^{(1)} + b_Z^{(2)} \\ X_2^{(2)} &= X_1^{(2)} \times X_1^{(2)} \\ &\dots \\ X_n^{(2)} &= X_{n-1}^{(2)} \times X_1^{(2)} \\ M_A^{(2)} &= \sigma \left(w_n^{(2)} X_n^{(2)} + \dots + w_1^{(2)} X_1^{(2)} + b^{(2)} \right) \\ \hat{Z}^{(2)} &= \text{SELU} \left(X_1^{(2)} \times M_A^{(2)} \right) \end{aligned} \quad (2)$$

Multimodal fusion network. Multimodal data can help improve the accuracy of diagnosis, prediction, and overall performance of learning systems⁽⁵⁸⁾. For instance, Venugopalan et al.⁵⁹ proposed multimodal deep learning models for AD data fusion to improve AD stage identification. Their trials established that the multimodal strategy outperformed the unimodal approach. De Jesus Junior et al.⁶⁰ described the discovery of multimodal indicators of AD severity for individuals in the early stages of the disease through combining Resting-State EEG and structural MRI data. In addition, their findings demonstrated the efficacy of the multimodal strategy. In this study, we present the multimodal multitask architecture for classifying MCI stage and predicting time-to-AD conversion. Our proposed model has two branches, as shown in the Fig. 3. The first branch operates on radiomics features, which are generated from a 3D MRI image using the SPM and PyRadiomics toolboxes, while the second branch operates on preprocessed clinical data. Each branch is connected to a SPAN block, which is comprised of a series of SPAN-followed dropout layers.

Assume that I_{radiomic} and I_{clinical} are the output features from SPAN blocks of radiomics and clinical features, respectively, we define an adaptive factor to select superior candidates from both representations. The operations for multimodal fusion are expressed as follows:

$$\begin{aligned} \tau &= \text{pool} \left(S \left(W_\tau e^{-\max(0, U_\tau I_{\text{clinical}})} + b_\tau \right) \right) \\ I_{\text{fused}} &= \tau I_{\text{clinical}} + (1 - \tau) I_{\text{radiomic}} \end{aligned} \quad (3)$$

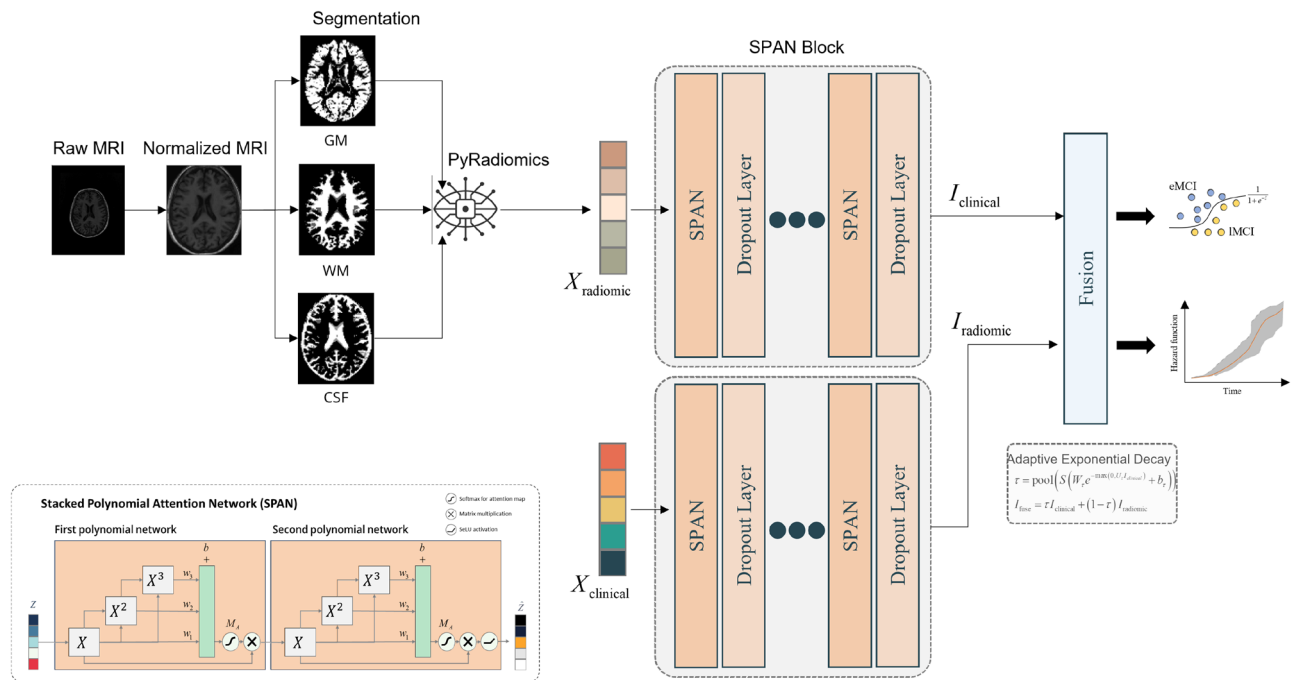


Figure 3. Architecture of the proposed multimodal multitask learning with SPAN.

where τ represents adaptive exponential decay (AED), W_τ , U_τ , b_τ represent trainable parameters, $\text{pool}(\cdot)$ represents maximum pooling operation, $S(\cdot)$ represents sigmoid function, and I_{fused} represents the fused representation from clinical and radiomics features. We used the exponential negative rectifier for decay factor τ to ensure that each decay rate decreases asymptotic within a tolerable range of 0 to 1. Lastly, the fused features is used to immediately predict the probabilities of eMCI vs IMCI, $\hat{y}_{\text{pr}}$, and the hazard rate of AD conversion, $\hat{y}_{\text{hr}}$, as follows:

$$\begin{aligned}\hat{y}_{\text{pr}} &= S(W_{\text{cl}}I_{\text{fused}} + b_{\text{cl}}) \\ \hat{y}_{\text{hr}} &= W_{\text{hr}}I_{\text{fused}} + b_{\text{hr}}\end{aligned}\quad (4)$$

where W_{cl} , b_{cl} , W_{hr} , b_{hr} represent trainable parameters.

Objective functions. To optimize the model's cost, we joint two objective functions of two tasks: MCI-stage classification and conversion-time-to-AD prediction. For classification task of eMCI vs IMCI, each predicted probability to the actual class output is measured by the binary cross-entropy (BCE) ⁶¹. Once the score has been calculated, probabilities are penalized based on the distance from the predicted value. That indicates how near or far the actual number is from the estimate. Given the actual class y_{pr} (0 for eMCI and 1 for IMCI), the BCE formula is as follows:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{M} \sum_{i=1}^M (y_{\text{pr}} \cdot \log(\hat{y}_{\text{pr}}) + (1 - y_{\text{pr}}) \cdot \log(1 - \hat{y}_{\text{pr}})) \quad (5)$$

where M is the number of samples within an iteration. Besides, we use the negative log-likelihood function ⁶² to minimize model's loss for the conversion-time-to-AD prediction task. Its expression is as follows:

$$\mathcal{L}_{\text{NLL}} = -\frac{1}{M_{E=1}} \sum_{i:E_i=1} \left(\hat{y}_{\text{hr}}(i) - \log \left(\sum_{j \in \mathbb{R}(T_j)} e^{\hat{y}_{\text{hr}}(j)} \right) \right) \quad (6)$$

In order to optimize the $\mathcal{L}_{\text{NLL}}$, we need to maximize the term of $\left(\hat{y}_{\text{hr}}(i) - \log \left(\sum_{j \in \mathbb{R}(T_j)} e^{\hat{y}_{\text{hr}}(j)} \right) \right)$ for each patient i having event $E = 1$ (uncensored patient who is converted to AD) for every censored patient (non-converted to AD). It follows that we must raise the risk factor for every uncensored patient i while simultaneously lowering the risk factor for patients j who have not experienced the event until time T_i , which is the observed time-to-AD for patient i . Finally, we add both loss functions for simultaneously multitask learning and arrive at the following result:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{NLL}} \quad (7)$$

Data availability

The dataset generated and analysed during the current study are available in the *Test Data* section of the *Download | Study Data* section of the IDA website (<https://ida.loni.usc.edu/pages/access/studyData.jsp?categoryId=43&subCategoryId=94>) under the name of *Tadpole challenge data*.

Received: 2 November 2022; Accepted: 22 June 2023

Published online: 11 July 2023

References

- Kohannim, O. *et al.* Boosting power for clinical trials using classifiers based on multiple biomarkers. *Neurobiol. Aging* **31**, 1429–1442. <https://doi.org/10.1016/j.neurobiolaging.2010.04.022> (2010).
- Gauthier, S., Rosa-Neto, P., Morais, J. A. & Webster, C. World alzheimer report 2021 journey through the diagnosis of dementia. *Alzheimer's Disease Int.* (2021).
- Burns, A. & S., I. Alzheimer's disease. *BMJ* **338**, b75. <https://doi.org/10.1136/bmj.b75> (2009).
- Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **82**, 239–259. <https://doi.org/10.1007/BF00308809> (1991).
- Venkatraghavan, V. *et al.* Analyzing the effect of apoe on Alzheimer's disease progression using an event-based model for stratified populations. *Neuroimage* **227**, 117646. <https://doi.org/10.1016/j.neuroimage.2020.117646> (2021).
- Ho, N.-H. *et al.* Predicting progression of Alzheimer's disease using forward-to-backward bi-directional network with integrative imputation. *Neural Netw.* **150**, 422–439. <https://doi.org/10.1016/j.neunet.2022.03.016> (2022).
- Jang, I. *et al.* Multiscale structural mapping of Alzheimer's disease neurodegeneration. *NeuroImage: Clin.* **33**, 102948. <https://doi.org/10.1016/j.nicl.2022.102948> (2022).
- Jessen, F. *et al.* Ad dementia risk in late mci, in early mci, and in subjective memory impairment. *Alzheimer's Dementia* **10**, 76–83. <https://doi.org/10.1016/j.jalz.2012.09.017> (2014).
- Golob, E. J., Irimajiri, R. & Starr, A. Auditory cortical activity in amnesic mild cognitive impairment: Relationship to subtype and conversion to dementia. *Brain* **130**, 740–752. <https://doi.org/10.1093/brain/awl375> (2007).
- Aisen, P. S. *et al.* Clinical core of the alzheimer's disease neuroimaging initiative: Progress and plans. *Alzheimer's Dementia* **6**, 239–246. <https://doi.org/10.1016/j.jalz.2010.03.006> (2010).
- Chételat, G. *et al.* Relationship between atrophy and β -amyloid deposition in Alzheimer disease. *Ann. Neurol.* **67**, 317–324. <https://doi.org/10.1002/ana.21955> (2010).
- Association, A. 2019 alzheimer's disease facts and figures. *Alzheimer's Dementia* **15**, 321–387. <https://doi.org/10.1016/j.jalz.2019.01.010> (2019).
- Vu, T.-D., Ho, N.-H., Yang, H.-J., Kim, J. & Song, H.-C. Non-white matter tissue extraction and deep convolutional neural network for Alzheimer's disease detection. *Soft. Comput.* **22**, 6825–6833. <https://doi.org/10.1007/s00500-018-3421-5> (2018).
- Zhang, F. *et al.* Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease. *Neurocomputing* **361**, 185–195. <https://doi.org/10.1016/j.neucom.2019.04.093> (2019).
- Spasov, S. *et al.* A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* **189**, 276–287. <https://doi.org/10.1016/j.neuroimage.2019.01.031> (2019).
- Park, C., Ha, J. & Park, S. Prediction of Alzheimer's disease based on deep neural network by integrating gene expression and dna methylation dataset. *Expert Syst. Appl.* **140**, 112873. <https://doi.org/10.1016/j.eswa.2019.112873> (2020).
- Zhou, Y. & Lui, Y. W. Small-world properties in mild cognitive impairment and early Alzheimer's disease: A cortical thickness mri study. *Int. Scholarly Res. Notices* <https://doi.org/10.1155/2013/542080> (2013).
- Mehmood, A. *et al.* 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> (2021).
- Odusami, M., Maskeliūnas, R., Damaševičius, R. & Krilavičius, T. Analysis of features of alzheimer's disease: Detection of early stage from functional brain changes in magnetic resonance images using a finetuned resnet18 network. *Diagnostics* **11**, 1071. <https://doi.org/10.3390/diagnostics11061071> (2021).
- Zhang, T. *et al.* Predicting mci to ad conversion using integrated smri and rs-fmri: machine learning and graph theory approach. *Front. Aging Neurosci.* <https://doi.org/10.3389/fnagi.2021.688926> (2021).
- Marcus, C., Mena, E. & Subramaniam, R. M. Brain pet in the diagnosis of Alzheimer's disease. *Clin. Nucl. Med.* **39**, e413. <https://doi.org/10.1097/RLU.0000000000000547> (2014).
- Bao, W., Xie, F., Zuo, C., Guan, Y. & Huang, Y. H. Pet neuroimaging of alzheimer's disease: radiotracers and their utility in clinical research. *Front. Aging Neurosci.* **13**, 114. <https://doi.org/10.3389/fnagi.2021.624330> (2021).
- Ho, M.-C. *et al.* Detect ad patients by using eeg coherence analysis. *J. Med. Eng.* <https://doi.org/10.1155/2014/236734> (2014).
- Kim, D. & Kim, K. Detection of early stage alzheimer's disease using eeg relative power with deep neural network. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 352–355. <https://doi.org/10.1109/EMBC.2018.8512231>, IEEE (2018).
- Jiang, Z., Liu, P., Xia, Y. & Zhang, J. Application of cnn in eeg image classification of ad patients. In *The 2nd International Conference on Computing and Data Science*, 1–5. <https://doi.org/10.1145/3448734.3450473> (2021).
- López-Sanz, D., Serrano, N. & Maestú, F. The role of magnetoencephalography in the early stages of Alzheimer's disease. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2018.00572> (2018).
- Lopez-Martin, M., Nevado, A. & Carro, B. Detection of early stages of Alzheimer's disease based on meg activity with a randomized convolutional neural network. *Artif. Intell. Med.* **107**, 101924. <https://doi.org/10.1016/j.artmed.2020.101924> (2020).
- Stoeckel, J. & Fung, G. Svm feature selection for classification of spect images of alzheimer's disease using spatial information. In *Fifth IEEE International Conference on Data Mining (ICDM'05)* <https://doi.org/10.1109/ICDM.2005.141>, IEEE (2005).
- Stam, C. J., Jones, B., Nolte, G., Breakspear, M. & Scheltens, P. Small-world networks and functional connectivity in Alzheimer's disease. *Cereb. Cortex* **17**, 92–99. <https://doi.org/10.1093/cercor/bhj127> (2007).
- Khazaei, A. *et al.* Classification of patients with mci and ad from healthy controls using directed graph measures of resting-state fmri. *Behav. Brain Res.* **322**, 339–350. <https://doi.org/10.1016/j.bbr.2016.06.043> (2017).
- Wang, Q. *et al.* A review of multimodal fusion techniques for medical image analysis. *Med. Image Anal.* <https://doi.org/10.1016/j.media.2021.102065> (2021).
- Li, H. *et al.* Deepfuse: A deep unsupervised learning approach for multimodal image fusion in medical diagnosis. *IEEE Trans. Med. Imaging* <https://doi.org/10.1109/TMI.2018.2812070> (2018).
- Tong, T. *et al.* Multimodal feature fusion for Alzheimer's disease diagnosis using hybrid weighted multiple kernel learning. *IEEE Trans. Pattern Anal. Mach. Intell.* <https://doi.org/10.1109/TPAMI.2016.2531522> (2017).
- Livni, R., Shalev-Shwartz, S. & Shamir, O. An algorithm for training polynomial networks. arXiv preprint [arXiv:1304.7045](https://arxiv.org/abs/1304.7045) (2013).
- Pölsterl, S., Sarasua, I., Gutiérrez-Becker, B. & Wachinger, C. A wide and deep neural network for survival analysis from anatomical shape and tabular clinical data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 453–464. https://doi.org/10.1007/978-3-030-43823-4_37 (Springer, 2019).

36. Lu, P. & Colliot, O. Multilevel survival analysis with structured penalties for imaging genetics data. In *Medical Imaging 2020: Image Processing*, vol. 11313, 113130K. <https://doi.org/10.1109/IBHI.2021.3100918> (International Society for Optics and Photonics, 2020).
37. Nakagawa, T. *et al.* Prediction of conversion to Alzheimer's disease using deep survival analysis of mri images. *Brain Commun.* **2**, fcaa057. <https://doi.org/10.1093/braincomms/fcaa057> (2020).
38. Ho, N.-H., Yang, H.-J., Kim, J., Dao, D.-P. & Pant, S. Rasurv: Residual attention-aware method for progression-free survival of Alzheimer's disease. In *The 5th International Conference on Big data, IoT, and Cloud Computing* (2021).
39. Suk, H.-I. *et al.* Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *Neuroimage* **101**, 569–582. <https://doi.org/10.1016/j.neuroimage.2014.06.077> (2014).
40. Nozadi, S. H. *et al.* Classification of Alzheimer's and mci patients from semantically Parcelled pet images: a comparison between av45 and fdg-pet. *Int. J. Biomed. Imaging* <https://doi.org/10.1155/2018/1247430> (2018).
41. Jie, B., Liu, M. & Shen, D. Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease. *Med. Image Anal.* **47**, 81–94. <https://doi.org/10.1016/j.media.2018.03.013> (2018).
42. Zhang, T. *et al.* Classification of early and late mild cognitive impairment using functional brain network of resting-state fmri. *Front. Psychiat.* <https://doi.org/10.3389/fpsyt.2019.00572> (2019).
43. Cui, X. *et al.* Adaptive lasso logistic regression based on particle swarm optimization for Alzheimer's disease early diagnosis. *Chemom. Intell. Lab. Syst.* **215**, 104316. <https://doi.org/10.1016/j.chemolab.2021.104316> (2021).
44. He, Y., Chen, Z., Gong, G. & Evans, A. Neuronal networks in Alzheimer's disease. *Neuroscientist* **15**, 333–350. <https://doi.org/10.1177/1073858409334423> (2009).
45. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014).
46. Kartsonaki, C. Survival analysis. *Diagn. Histopathol.* **22**, 263–270. <https://doi.org/10.1016/j.mpdhp.2016.06.005> (2016).
47. Katzman, J. L. *et al.* DeepSurv: personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC Med. Res. Methodol.* **18**, 1–12 (2018).
48. Royston, P. & White, I. R. Multiple imputation by chained equations (mice): implementation in stata. *J. Stat. Softw.* **45**, 1–20. <https://doi.org/10.18637/jss.v045.i04> (2011).
49. Lambin, P. *et al.* Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**, 749–762. <https://doi.org/10.1038/nrclinonc.2017.141> (2017).
50. Yip, S. S. & Aerts, H. J. Applications and limitations of radiomics. *Phys. Med. Biol.* **61**, R150. <https://doi.org/10.1088/0031-9155/61/13/R150> (2016).
51. Du, Y. *et al.* Radiomic features of the hippocampus for diagnosing early-onset and late-onset alzheimer's disease. *Front. Aging Neurosci.* **13**, <https://doi.org/10.3389/fnagi.2021.789099> (2021).
52. Shu, Z.-Y. *et al.* Prediction of the progression from mild cognitive impairment to Alzheimer's disease using a radiomics-integrated model. *Ther. Adv. Neurol. Disord.* **14**, 17562864211029552. <https://doi.org/10.1177/17562864211029551> (2021).
53. Shickel, B., Tighe, P. J., Bihorac, A. & Rashidi, P. Deep ehr: a survey of recent advances in deep learning techniques for electronic health record (ehr) analysis. *IEEE J. Biomed. Health Inform.* **22**, 1589–1604. <https://doi.org/10.1109/IBHI.2017.2767063> (2017).
54. Vaswani, A. *et al.* Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30** (2017).
55. Abd Hamid, N. A. *et al.* Incorporating attention mechanism in enhancing classification of alzheimer's disease. In *New Trends in Intelligent Software Methodologies, Tools and Techniques*, 496–509, <https://doi.org/10.3233/FAIA210048> (IOS Press, 2021).
56. Zheng, X., Shi, J., Li, Y., Liu, X. & Zhang, Q. Multi-modality stacked deep polynomial network based feature learning for Alzheimer's disease diagnosis. In *2016 IEEE 13th international symposium on biomedical imaging (ISBI)*, 851–854, <https://doi.org/10.1109/ISBI.2016.7493399> (IEEE, 2016).
57. Klambauer, G., Unterthiner, T., Mayr, A. & Hochreiter, S. Self-normalizing neural networks. *Adv. Neural Inf. Process. Syst.* **30** (2017).
58. Behrad, F. & Abadeh, M. S. An overview of deep learning methods for multimodal medical data mining. *Exp. Syst. Appl.* **117**, 117006. <https://doi.org/10.1016/j.eswa.2022.117006> (2022).
59. Venugopalan, J., Tong, L., Hassanzadeh, H. R. & Wang, M. D. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep.* **11**, 1–13. <https://doi.org/10.1038/s41598-020-74399-w> (2021).
60. De Jesus Junior, B. *et al.* Multimodal prediction of Alzheimer's disease severity level based on resting-state eeg and structural mri. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2021.700627> (2021).
61. DiPietro, R. A friendly introduction to cross-entropy loss. <https://rdipietro.github.io/friendly-intro-to-cross-entropy-loss> (2016).
62. Cox, D. R. Regression models and life-tables. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* **34**, 187–202 (1972).

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT). (RS-2023-00208397). This study was supported by a grant (HCRI 23038) Chonnam National University Hwasun Hospital Institute for Biomedical Science. This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government (MSIT).

Author contributions

N.-H.H conceived and conducted the experiments, analysed the results, and participated in the writing of the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-37500-7>.

Correspondence and requests for materials should be addressed to Y.-H.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023