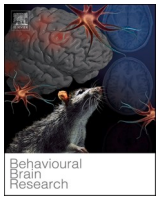
Behavioural Brain Research 463 (2024) 114900

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr



Longitudinal MRI analysis using a hybrid DenseNet-BiLSTM method for Alzheimer's disease prediction

Alireza Jomeiri [a], Ahmad Habibizad Navin [b,*], Mahboubeh Shamsi [c]

[a] Department of Computer Engineering, Qom Branch, Islamic Azad University, Qom, Iran
[b] Department of Computer Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran
[c] Department of Engineering, Qom University of Technology, Qom, Iran

ARTICLE INFO

Keywords:
Alzheimer's disease
Longitudinal analysis
Magnetic resonance imaging
DenseNet
BiLSTM
Deep learning

ABSTRACT

Alzheimer's disease is a progressive neurological disorder characterized by brain atrophy and cell death, leading to cognitive decline and impaired functioning. Previous research has primarily focused on using cross-sectional data for Alzheimer's disease identification, but analyzing longitudinal sequential MR images is crucial for improved diagnostic accuracy and understanding disease progression. However, existing deep learning models face challenges in learning spatial and temporal features from such data. To address these challenges, this study presents a novel hybrid DenseNet-BiLSTM method for Alzheimer's disease prediction using longitudinal MRI analysis. The proposed framework combines Convolutional DenseNet for spatial information extraction and joined BiLSTM layers for capturing temporal characteristics and relationships between longitudinal images at different time points. This approach overcomes issues like overfitting, vanishing gradients, and incomplete patient data. We evaluated the model on 684 longitudinal MRI images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, including normal controls, individuals with mild cognitive impairment, and Alzheimer's disease patients. The results demonstrate high classification accuracy, with 95.28% for AD/CN, 88.19% for NC/MCI, 83.51% for sMCI/pMCI, and 92.14% for MCI/AD. These findings highlight the substantial improvement in Alzheimer's disease diagnosis achieved through the utilization of longitudinal MRI images. The contributions of this study lie in both the deep learning and medical domains. In the deep learning domain, our hybrid framework effectively learns spatial and temporal features from longitudinal data, addressing the challenges associated with multi-dimensional and sequential time series data. In the medical domain, our study emphasizes the importance of analyzing baseline and longitudinal MR images for accurate diagnosis and understanding disease progression.

1. Introduction

1.1. Background and significance

Alzheimer's disease is a degenerative neurologic condition that destroys brain cells and causes the brain to shrink. Alzheimer's disease is the most prevalent cause of dementia, which is characterized by a steady deterioration in mental, behavioral, and social abilities and impairs a person's capacity for independent functioning. In the US, 5.8 million persons who are 65 or older and have Alzheimer's disease are affected (Matthews et al., 2018). 80 percent of them are 75 years of age or older. Between 60% and 70% of the 50 million or more persons with dementia globally are thought to have Alzheimer's disease. Alzheimer's disease treatment expenses in 2010 were estimated to be between $159 and $215 billion. According to projections, these expenses would increase to between $379 and more than $500 billion yearly by 2040 (Hurd et al., 2013). Alzheimer's disease still has no known clinical cure, and existing therapies simply halt disease progression (Masters and Beyreuther, November 2006).

In general, there are 4 categories of biomarkers to detect Alzheimer's disease: Neuroimaging, Cerebrospinal Fluid proteins, Blood and Urine Tests, Genetic Risk Profilers. Neuroimaging methods can be divided into structural and functional categories. Structural imaging methods include computed tomography (CT) and magnetic resonance imaging (MRI). CT imaging has a high resolution and can be used to distinguish two structures from one another. Due to its high spatial resolution, MRI

* Corresponding author.
E-mail address: A.habibizad@srbiau.ac.ir (A.H. Navin).

https://doi.org/10.1016/j.bbr.2024.114900
Received 30 July 2023; Received in revised form 16 December 2023; Accepted 6 February 2024
Available online 8 February 2024

imaging can, however, be utilized to discern between two tissues that are superficially similar but not identical. You (2022) proposed fine perceptive GANs for brain MR image super-resolution in the wavelet domain, which enhanced the resolution of MRI images, enabling more detailed analysis of brain structures. Positron emission tomography (PET), single photon emission computed tomography (SPECT), and functional MRI are examples of functional neuroimaging techniques (fMRI). Although functional imaging techniques offer some structural details, they have a lesser spatial resolution than structural imaging techniques (Márquez and Yassa, 2019). Images from various imaging modalities show the changes that cause AD due to brain cell degeneration. A number of researchers have used neuroimaging techniques to diagnose Alzheimer's. MRI (Khan, 2021; Cao, 2017; Long, 2017; Lama, 2017), fMRI (Ibrahim, 2021; Sarraf and Tofighi, 2016; Ramzan, 2020), PET (Chételat, 2020; Ding, 2019; Marcus et al., 2014), SPECT (Valotassiou, 2018; Świetlik and Białowąs, 2019; Górriz, 2011), and DTI (Khvostikov, 2018; De and Chowdhury, 2021) have all been used to diagnose or prognosis Alzheimer's disease. Moreover, Hu (2021) introduced bidirectional mapping generative adversarial networks for brain MR to PET synthesis, allowing the generation of PET images from MR images, which can provide complementary information for Alzheimer's disease diagnosis. Furthermore, data from multiple modalities was combined to enhance diagnosis performance (Alberdi et al., 2016; Zhang, 2021; Shi, 2019; Bi, 2020; Naik et al., 2020; Lin, 2023). Regarding the third marker, to date, the three most likely CSF biomarkers have been identified: A42, total tau (t-tau) and phosphorylated tau (p-tau) (Anoop, 2010). Since blood samples can be obtained less invasively, cheaper and more frequently than CSF samples, biomarkers of AD in the blood were also examined (Snyder et al., 2014). Genetic Risk Profilers are the fourth biomarker in the diagnosis of Alzheimer's disease. APOE-e4 is the first risk gene discovered and continues to be the gene that has the greatest impact on risk. According to researchers, the APOE-e4 gene is present in 40–65 percent of patients with Alzheimer's disease (Stocker, 2018).

Recent studies show that neuroimaging biomarkers provide much more accurate predictions in AD diagnosis than other biomarkers (Márquez and Yassa, 2019), providing hopeful methods for individualized diagnosis and prognosis (Rathore, 2017; Falahati et al., Jan. 2014; Haller et al., Jan. 2011). For instance, Yu (2020) proposed a multi-scale enhanced graph convolutional network for early mild cognitive impairment detection, which demonstrated promising results in accurately identifying cognitive impairment at an early stage. These markers can be very useful in the early detection as well as the conversion diagnosis of different stages of Alzheimer's disease (NC/MCI/AD). Lei (2022) proposed a joint and deep learning approach for predicting clinical scores in Alzheimer's disease, which effectively integrated multiple data sources and achieved accurate assessments of disease severity. Early and accurate diagnosis of Alzheimer's disease is critical for patient care and treatment. Mild Cognitive Impairment (MCI) is a transitional stage between normal cognition and Alzheimer's disease. Every year, about 10–25% of patients with MCI eventually progress to Alzheimer's (Lu et al., 2017). There are two types of MCI: progressive MCI (pMCI) and stable MCI (sMCI). pMCI denotes that MCI subjects will eventually convert to AD, whereas sMCI subjects are stable and will not convert. Since Alzheimer's disease cannot be cured or prevented at present, early detection of possible progression of pMCI before irreversible brain injury occurs is very important for preventive care. Machine learning algorithms can play an important role in helping specialists analyze patient data by predicting the conversion of MCI to AD (Kruthika et al., 2019; Uysal and Ozturk, 2020; Lu et al., 2017; Ito et al., 2011; Zhou et al., 2013; Liu et al., 2014; Duchesne et al., 2009; Beheshti, 2017; Cao, 2017; Tong et al., 2017). Furthermore, Hu (2020) developed a medical image reconstruction method using generative adversarial networks (GANs) to address the class-imbalance problem in Alzheimer's disease assessment, which improved the accuracy of disease diagnosis.

1.2. Previous approaches and limitations

Several modalities can be used to study the same subject, according to recent research. However, collecting different modalities for the same subject is difficult, resulting in a smaller number of subjects for study (Khan, 2021). On the other hand, there is the problem of modality-wise missing data in the use of multimodality (Liu et al., 2018 Apr). Therefore, to obtain better results in this paper, we have used single modality. Magnetic resonance imaging (MRI) is a non-invasive method of viewing brain atrophy changes that has been broadly used in AD research due to its high spatial resolution, increased accessibility, high contrast, low cost, and lack of radiation in the scanning process (Salvatore et al., 2018 May 24; Syaifullah et al., 2021 Feb 5; Wong, 2021). T1-weighted magnetic resonance (MR) imaging with high-resolution three-dimensional (3D) sequences provides detailed anatomical information of the brain, empowering a variety of brain imaging studies, including quantitative measurements of brain tissue volume and cortical thickness, and image classification for early disease diagnosis. Deep learning structures can be applied to three types of neuroanatomical methods: voxel-based, ROI--based, and patch-based. The voxel-based method assesses differences in local concentrations of brain tissues, primarily grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) (Zhang et al., 2019). This analytical method studies the entire brain structure rather than just the regional information in the brain, and therefore minor changes in the brain can be acquired with voxel-based morphometry. However, because the neuroimaging dataset is not large, this feature extraction method may suffer from overfitting issues, as there are fewer images available compared to millions of features in each image. The Region of Interest neuroanatomical computational method considers brain regions that are predefined structurally or functionally (Chitradevi and Prabha, 2020). It extracts features from T1/T2 weighted images and diffusion weighted images by utilizing the spatial information in these images. The hippocampus is a well-known region that has been considered in almost all ROI-based feature extraction studies on Alzheimer's prediction because there is signs of amyloid plaque deposition in this region (Uysal and Ozturk, 2020), (Grimm et al., 2015). One disadvantage of this ROI-based morphometry is that ROI specificity requires expert human knowledge. Patch-based feature extraction methods divide the entire brain into multiple patches from which several feature vectors are extracted (Zhang et al., 2016). In contrast to the ROI-based framework, this computational technique does not require any manual identification of ROIs. However, it becomes difficult to select the informative patches from the entire image because a vast amount of patches are generated from each image and the computational environment may not be compatible with all extracted patches (Wen et al., 2020).

Deep learning models have gained notoriety in recent years for their capacity to derive feature representations from the input raw data. Deep learning algorithms extract progressively complex feature representations from the data using a layered, hierarchical structure. Deep learning architectures construct complicated high-level features in a hierarchy by learning basic, low-level features from the input. In a number of fields, such as object tracking, visual object identification, natural language processing, human action recognition, image restoration, denoising, segmentation tasks, audio classification, and brain-computer interface, deep learning methods have shown revolutionary performance. Convolutional neural networks (CNNs), in particular, have recently shown great performance in the area of medical imaging, namely in the domains of segmentation, detection, registration, and classification (Litjens et al., 2017). Deep learning methods are capable of finding latent or hidden representations in neuroimaging data and effectively capturing disease-related disorders. Therefore, recently, many researchers have turned to using deep learning methods to diagnose Alzheimer's disease. (Zhang and Shi, (2020) developed DMFNet by concatenating 2D CNN with the attention model and combining both low-level and high-level characteristics. (Janghel and Rathore, 2020) used the VGG-16 network to classify AD from NC using fMRI and PET NIFTI 3D data in 2D format.

Li et al. (2020) used a 3D-CNN to extract spatial features from the volumes of 3D static neuroimages, and then transferred these feature maps to an LSTM method to retrieve temporal information. Convolutional Autoencoder and 3D CNN were used for Alzheimer detection by Oh et al. (2019). Separate 3D convolutional networks were built by Punjabi et al. (2019) for structural T1-weighted MRI and AV-45 Amyloid PET, and these data were then fused using a category cross entropy classifier. In order to predict pMCI using PET scans and a single-task model, (Choi and Jin, 2018) employed CNN. Using the late integration of MRI, demographics, neuropsychological, and apolipoprotein E4 (APOe4) data, (Spasov et al., 2018) suggested a multimodal single-task classification method based on a CNN to identify the development of AD However, because a medical professional often reviews the longitudinal patient data before making advancement decisions, these models are less efficient, less sufficient, and not medically acceptable (Ding et al., 2018). Based on MRI time-series data with five time intervals, (Cui et al., 2019) suggested a CNN-recurrent neural network (RNN) model for AD diagnosis. The majority of DL models for diagnosing AD use binary classifications based on a single time point and a single task to detect the pathological alterations (Amoroso et al., 2018).

(Ocasio and Duong, 2021) used the CNN network to predict the conversion of MCI to AD during 3 years. (El-Sappagh et al., 2020) used multimodal data and a combination of CNN and LSTM for disease classification. In (Zhu, 2021), a temporal structured support vector machine (TS-SVM) model was presented to investigate disease progression. In (Zhang et al., 2020), a linear SVM with nested LOOCV was proposed to classify four groups of images. On the other hand, most of the published works have used baseline patient data to diagnose and classify Alzheimer's and have not considered the spatio-temporal nature of the disease data. For example, (Márquez and Yassa, 2019) used MRI images and three categories based on SVM to predict Alzheimer's disease. Convolutional autoencoder (CAE)-based unsupervised learning and supervised transfer learning for AD versus NC classification were proposed in (Oh et al., 2019). Lin (2018) used CNN for disease classification and obtained a diagnosis accuracy of 81.4%. Consequently, models based on baseline data are less accurate than those based on longitudinal data from a patient, making this a suboptimal strategy for detecting progression. The problems that exist in previous works are: average detection accuracy, overfitting problem, vanishing gradients and incomplete patient data problem. To overcome these problems, a new hybrid model based on deep learning architectures is proposed, which can extract the spatio-temporal characteristics of the images with high accuracy and provide a suitable classification of the disease.

1.3. Importance of longitudinal analysis

Alzheimer's is a chronic disease. The state of the disease at a certain point in time is dependent on the state of the previous point in time. Considering that in treatment systems, the progress of Alzheimer's disease is determined over time, and usually the data obtained from the patient are longitudinal and are visited at different times, so the analysis of these data can help us lead to a more accurate and reliable diagnosis of the disease (Mingxia, 2018). According to our studies, until today, limited research has been done on longitudinal data of Alzheimer's disease with deep learning methods.

To model and measure the progression of Alzheimer's disease over time to achieve higher diagnostic accuracy, baseline and longitudinal analysis of sequential MRI images is particularly necessary. In this paper, we proposed a hybrid deep method for Alzheimer's disease classification using DenseNet network and BiLSTM network. Dense Convolutional Network (DenseNet) (Huang et al., 2017) benefits from feature map reuse by dense connections, minimizing dependency across layers by reusing feature maps from multiple layers, offering compact and differentiated input features by short connections of varying lengths, and successfully addressing the gradient vanishing problem that arises as the CNN layer deepens. It also outperforms CNN in training data from limited data. Furthermore, the Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) is a framework that addresses the long-term dependency issue, vanishing gradient, and exploding gradient of Recurrent Neural Network (RNN) and is primarily used to forecast time-series data, making it appropriate for identifying temporal properties of longitudinal MRI images, which are imaging data of several time points. The proposed method consists of two steps. In the first step, in order to learn and extract spatial features from structural MRI images and disease classification, a DenseNet network with different layers is built. In the second step and following the DenseNet network, a stacked deep network is created based on the BiLSTM structure in order to capture the output time series images of the DenseNet network and extract the temporal features of the images with the aim of accurately classifying the disease. BiLSTM architecture can process sequential images with different lengths and solve the problem of incomplete longitudinal data in different time points. In addition to solving the problems in previous works, this proposed hybrid deep model can automatically extract temporal and spatial features and internal correlations of baseline and longitudinal MRI images at different time points and provide accurate classification of Alzheimer's disease.

1.4. Objectives and contributions

The contributions of this study lie in both the deep learning and medical domains. In the deep learning domain, we propose a hybrid DenseNet-BiLSTM framework that overcomes challenges related to overfitting, vanishing gradients, and incomplete patient data in longitudinal MRI analysis. In the medical domain, our study emphasizes the importance of analyzing baseline and longitudinal sequential MR images for improved diagnostic accuracy and understanding of disease progression. By combining Convolutional DenseNet for spatial information extraction and joined BiLSTM layers for capturing temporal characteristics, our framework achieves high classification accuracy and significantly enhances the diagnosis of Alzheimer's disease using longitudinal MRI images. The main contributions of this study are: 1) proposing of a hybrid deep learning framework: we introduce a novel hybrid framework that addresses challenges in longitudinal MRI analysis for Alzheimer's disease prediction. 2) Improving diagnostic accuracy: The proposed framework demonstrates high classification accuracy in Alzheimer's disease diagnosis using longitudinal MRI images. 3) Addressing key challenges: our framework addresses crucial challenges in longitudinal MRI analysis, including overfitting, vanishing gradients, and incomplete patient data and 4) Utilization of longitudinal sequential MR images: we emphasize the importance of analyzing baseline and longitudinal sequential MR images for improved diagnostic accuracy and understanding disease progression in Alzheimer's disease. Previous research has primarily focused on cross-sectional data, and this study fills the gap by incorporating longitudinal data analysis. In this research work, we used the imaging data of 684 subjects who were scanned during 4 years in order to analyze Alzheimer's disease progress. These images include 193 NC, 132 pMCI, 185 sMCI and 174 AD. The proposed method was tested based on the important performance evaluation criteria. The results of various experiments showed that the proposed method has high efficiency and accuracy of diagnosis.

1.5. Paper organization

The rest of this paper is organized as follows. Section 2 describes the set of longitudinal MRI images used and the proposed methods and architectures. In Section 3, the implementation results, various experiments and comparison of the proposed deep method with the methods presented in the new papers are given. In Section 4, the discussion of the paper is presented. Study Limitations and Future Research Directions are discussed on Section 5 and finally, Section 6 deals with the summary and conclusion of the paper.

2. Materials and methods

The image dataset and the proposed hybrid deep model of disease classification are presented in this section. The set of MRI images is collected from the ADNI Comprehensive Dataset. The use of longitudinal and high-dimensional images is a promising attempt to improve the diagnosis of Alzheimer's disease. The proposed hybrid deep framework is shown in Fig. 1. In this model, we have presented a new combination based on DeseNet network architecture and BiLSTM network for classifying Alzheimer's disease. First, the MRI images are preprocessed. The pre-processed images are then fed into the DenseNet network. In this network, the spatial features of the images are extracted with high accuracy. After that, in order to extract temporal features, the time series data are entered into BiLSTM stack network. BiLSTM can model and measure disease progression using images taken at different time points. Finally, after integrating the set of features in the final layers of the model, the classification of the disease is done.

In our research, we employed a hybrid DenseNet-BiLSTM architecture for Alzheimer's disease prediction using longitudinal MRI analysis. The selection of this architecture was based on a combination of prior research and empirical experimentation. The DenseNet architecture was chosen for its ability to effectively extract spatial information from the input MRI images. DenseNet is a convolutional neural network (CNN) architecture that introduces dense connections between layers, enabling feature reuse and alleviating the vanishing gradient problem. Prior studies have demonstrated the effectiveness of DenseNet in various image-related tasks, including medical image analysis. To capture the temporal characteristics and relationships between longitudinal images, we incorporated Bidirectional Long Short-Term Memory (BiLSTM) layers into the architecture. BiLSTM is a type of recurrent neural network (RNN) that allows for the modeling of sequential data by considering both past and future information at each time step. This enables the model to capture temporal dependencies and changes in the longitudinal MRI data. The combination of DenseNet and BiLSTM components in our hybrid architecture aims to leverage the strengths of both models in learning spatial and temporal features from the longitudinal data, respectively.

Regarding the optimization and selection of the best architecture, we conducted an iterative process of experimentation and evaluation. We trained and evaluated multiple variations of the model by adjusting hyperparameters, such as the number of DenseNet blocks, the number of layers in the BiLSTM, and the size of the hidden states. We performed thorough training and validation to assess the performance of each architecture variant. Our selection of the final model architecture was based on a comprehensive analysis of the validation results, considering factors such as classification accuracy, convergence behavior, and generalization capability. We aimed to choose an architecture that achieved high prediction accuracy while avoiding overfitting and maintaining robust performance across different evaluation metrics.

2.1. Study cohorts

The Alzheimer's Disease Neuroimaging Initiative (ADNI) database was used for the study's data (http://adni.loni.usc.edu/). As a $60 million, five-year public-private partnership, the ADNI was established in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and nonprofit organizations. The main objective of ADNI has been to determine if serial MRI, PET, other biological markers, clinical, and neuropsychological evaluation may be used in conjunction to monitor the development of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Aiming to save the time and expense of clinical trials while also assisting doctors and researchers in the development of new therapies and the monitoring of their effects, the identification of precise and sensitive indicators of very early AD progression is a goal of many studies. The ADNI is the product of the work of several researchers from a variety of academic institutions and private businesses, and participants have been enrolled from more than 50 sites in the United States and Canada. All participating locations' IRBs have given their approval for the ADNI trial. All subjects and, if relevant, their legal representatives, gave their written informed consent. Visit www.adni-info.org for the most recent information.

In this research, we used longitudinal MRI brain scan images to diagnose Alzheimer's disease. It is known that MRI images are one of the most prominent and common modalities for accurate diagnosis of diseases. The set of baseline and longitudinal images used in this paper are obtained from 3 Tesla T1-weighted MR images using volumetric 3D MPRAGE with 240×256×176 voxels and 1 mm×1 mm×1.2 mm voxel size. In order to evaluate the proposed hybrid model, based on available data, we used brain scans of 684 ADNI participants. These images are time series and were taken from patients during 48 months of visits. The

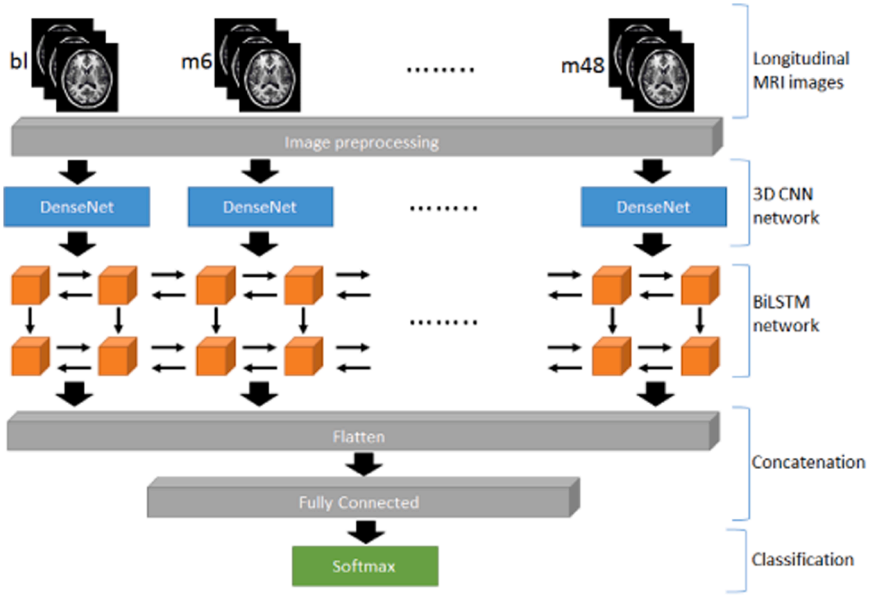


Fig. 1. - framework of proposed hybrid deep DenseNet-BiLSTM model.

number of time points is 7, namely: baseline, M6, M12, M18, M24, M36 and M48. The patient images, as shown in Table 1, are divided into four different classes. The first class, which is 193 subjects, includes NC who have remained healthy for 48 months. The second class, which is 185 subjects, includes sMCI patients (MCI patients who have not progressed to AD during a 4 years follow-up). The third class, which is 132 subjects, includes patients with pMCI (MCI patients who have converted to AD during a 4 years follow-up). And the fourth class, which is 174 subjects, includes patients who have AD throughout the visit. In this research, patients who had a reverse conversion during the visit (i.e. patients who changed from AD to MCI or from MCI to NC) and subjects who changed from a NC state to AD state were not considered. In three distinct planes, Fig. 2 displays representative MR scans of individuals with normal cognitive (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

2.1.1. Pre-processing

All structural MRI information utilized in this work was obtained from ADNI in the Neuroimaging Informatics Technology Initiative (NIfTI) format, which had previously undergone processing to account for spatial distortion brought on by gradient nonlinearity and B1 field inhomogeneity. By performing the standard techniques of anterior commissure (AC)-posterior commissure (PC) correction, skull-stripping, and cerebellum removal, we further processed the MR images. In particular, we employed the N3 algorithm (Sled et al., 1998) to adjust intensity inhomogeneity, smoothed images to 256×256×256, and used MIPAV software for AC-PC correction. Along with cerebellar removal, a precise and reliable skull stripping (Wang and Summers, 2014) was carried out. We manually examined the skull-stripped scans to verify the clean and dura removal. Then, the structural MRI image was segmented into three tissue types: gray matter (GM), white matter (WM), and cerebrospinal fluid using FAST in the FSL package (CSF). A mass-preserving deformable warping method named HAMMER will spatially normalize the three tissue volumes of each subject's various time-point images onto a standard space (Shen and Davatzikos, 2003). Next, we created the regional volumetric maps, known as RAVENS maps, by warping the images while still protecting the tissue (Davatzikos et al., 2001). Due of the spatially normalized GM densities' relative high significance to AD compared to WM and CSF, we exclusively took them into consideration in our study (Liu et al., 2012). When utilizing longitudinal GM volume map MRI images, the specific patch-level features that are extracted by Densent are:

Convolutional Features: DenseNet applies convolutional operations to the GM volume map patches to extract features that capture local patterns, edges, and structures. These convolutional features represent learned representations of the patches.

Activation Maps: DenseNet generates activation maps that indicate the response of each filter in the convolutional layers to the input GM volume map patches. These activation maps can be used as features to capture the presence and strength of specific image features in different regions of the patches.

Concatenated Features: DenseNet utilizes dense connections, where the feature maps from previous layers are concatenated with the current layer's input. These concatenated features capture both local and global information about the GM volume map patches.

To leverage these features with a BiLSTM model, we reshape the extracted patch-level features into a sequence format suitable for sequential processing. Since we are working with longitudinal MRI images, it is important to consider the temporal dimension. One approach is to treat each longitudinal sequence of GM volume maps for a particular subject as a single input sequence for the BiLSTM model. In this case, we can extract the patch-level features from each GM volume map within the longitudinal sequence using DenseNet. The resulting features can then be organized into a sequence format, where each patch-level feature vector corresponds to a specific time point in the sequence. This sequence of patch-level feature vectors can be fed into the BiLSTM model to capture the temporal dependencies and patterns across the longitudinal GM volume maps for AD classification.

The BiLSTM component of the model extracts temporal dynamics, long-term dependencies, contextual information, and feature representations from the sequential patch-level features extracted by DenseNet. These features allow the model to effectively incorporate the temporal aspect of the longitudinal GM volume map MRI images and make accurate AD classification predictions.

2.2. Feature learning by DenseNet

This research is aimed at building a deep Densely Connected Convolutional Networks (DenseNet) to learn the meaningful and discriminative features from the 3D image density map for AD diagnosis, in contrast to the traditional approach that extracts hand-crafted features. After preprocessing the 3D MRI images and extracting the gray matter volume of the brain, these data are entered into the DenseNet network for training. It has been extensively researched how to train features for image classification using the DensNet architecture, which alternately stacks many convolutional and pooling layers followed by fully connected and softmax layers. While the input data travels through several layers before reaching the end of the network, CNNs develop deeper and deeper to increase the representational power. However, when the features changed from low to high levels, the information was lost. DenseNet, which increases direct connections between the low and high level layers, was offered as a solution to this issue. It connects each layer to every other layer in a feed-forward manner (Huang et al., 2017). DenseNet improves information flow across layers by densely connecting layers from various levels. DenseNets provide numerous enticing benefits over conventional CNN (Huang et al., 2017). Because there is a direct relationship between the low and high level layers, they can first address the vanishing gradient issue. To reuse the low-level features, feature propagation is further strengthened. Thirdly, they can greatly decrease the amount of learning parameters. As a result, we build the 3D DenseNet on the image map in order to learn the imaging intensity and spatial information for disease classification. Fig. 3 illustrates the DenseNet architectural layout. Using the associated GM density map alone, we train these DenseNets classifiers. With a convolutional layer, four dense blocks, three transition layers, an average pooling layer, and a softmax layer, our built-in deep 3D DenseNet model has a network design and parameters that are shown in Table 2. After the input layer, a convolutional layer with a stride of 2× 2 ×2 is inserted first, then the dense blocks. The dense block employs a dense connection as opposed to the conventional one-way connectivity from one layer to the next, via which the lth layer gets the output feature maps of all previous layer as described by Huang et al. (2017):

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \quad (1)$$

Table 1
Demographic and clinical information of the studied subjects from ADNI dataset.

Diagnosis	Number	Gender (M/F)	Age (Mean+SD)	MMSE (Mean+SD)	Education (year)	CDR (Mean+SD)
AD	174	95/79	75.6±7.8	23.3±2.0	15.1±2.8	0.8±0.8
pMCI	132	81/51	74.9±6.9	26.7±1.7	15.8±2.5	0.4±0.1
sMCI	185	128/57	75.1±7.8	27.3±1.8	15.0±2.8	0.4±0.1
NC	193	107/86	76.1±5.2	29.1±1.0	15.9±2.7	0.1±0.0

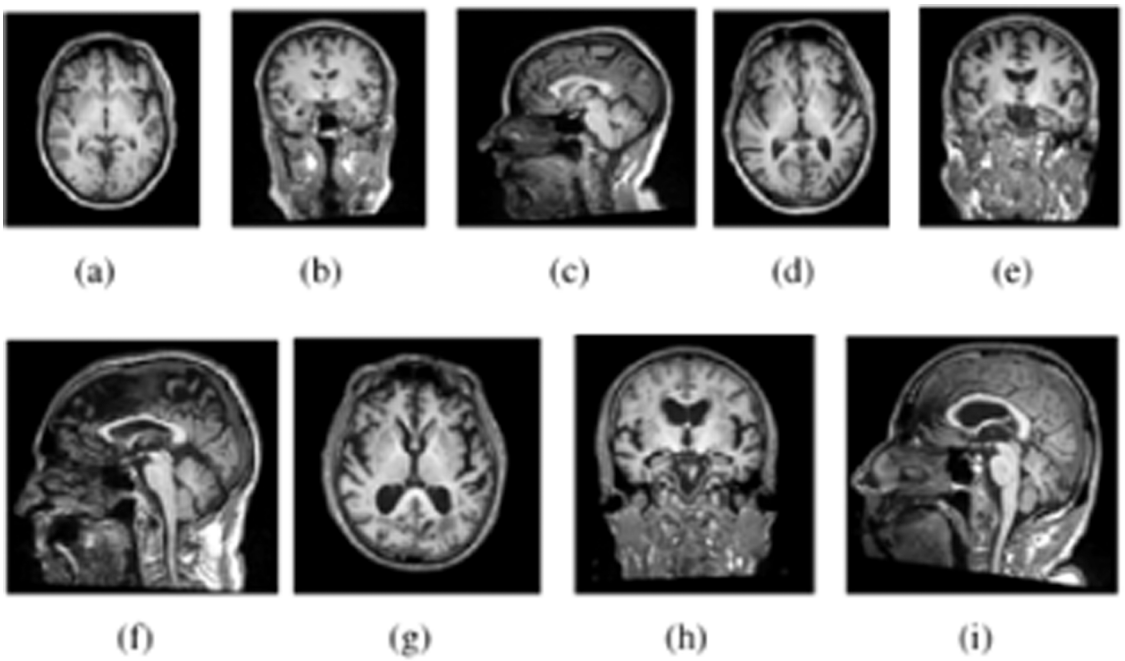


Fig. 2. Sample images from the ADNI cohort (a),(b),(c) are the images of a NC subject in axial, coronal and sagittal planes, respectively. (d),(e),(f) are the images of a MCI subject in axial, coronal and sagittal planes, respectively. (g),(h),(i) are the images of an AD subject in axial, coronal and sagittal planes, respectively.

where $[x_0, x_1, \ldots, x_{l-1}]$ represents the combination of the feature maps from all preceding layers into a single tensor, and $H_l$stands for a composite nonlinear transformation function made up of batch normalization, leaky rectified linear units, 3 ×3 ×3 convolution, and dropout. Every dense layer receives the feature maps of all preceding dense layers via shortcut links in the dense blocks. The dense layer consists of two batch normalization layers, two activation layers, one 1×1×1 and one 3×3×3 convolutional layer. All dense blocks have three dense layers. Concatenating the feature maps from one dense block results in the same dimensions for all of them. Thus, the transition layer is set between two dense blocks to achieve dimension reduction of the feature maps and it consists of five consecutive operations: batch normalization, leaky rectified linear units, a 1 × 1×1 convolution, dropout and a 3 × 3×3 convolution with a stride of 2 × 2×2. The last transition layer has the same structure as the previous transition layer except that the last convolution is of 4 × 4×3 with a 1 × 1×1 stride. Following the last dense block, an average pooling and a softmax classifier are appended to reduce the feature dimension and make the classification. The weights of DenseNet are updated through back-propagation for feature learning. In order to accomplish dimension reduction of the feature maps, the transition layer is placed between two dense blocks and comprises of five sequential operations: batch normalization, leaky rectified linear units, a 1×1×1 convolution, dropout, and a 3×3×3 convolution with a stride of 2×2×2. With the exception of the last convolution, which is a 4× 4× 3 convolution with a 1 ×1 ×1 stride, the final transition layer shares the same structure as the preceding transition layer. To decrease the feature dimension and perform the classification, an average pooling and a softmax classifier are added after the last dense block. Back-propagation is used to update DenseNet's weights as a result of feature learning. Through the shortcut connections, the loss function directly supervises each layer of DenseNet. A Softmax layer optimizes each DenseNet separately for the classification job before producing the class prediction score.

2.3. The structure of long short-term memory (LSTM) network

For the modeling, prediction, and classification of sequential/time series data, recurrent neural networks (RNNs) are one of the deep learning approaches employed. RNNs are very good at solving problems like speech recognition, keyword extraction, and machine translation, yet these networks have drawbacks like disappearing gradient and expanding gradient (Salehinejad, 2017). In order to address the vanishing and expanding gradient issues, the Long Short-Term Memory (LSTM) was developed as a group of RNN networks. The recurrent latent neurons are given an internal state in the LSTM structure so that, at each time point, the LSTM models use the output of the preceding state and the internal state as inputs to the new model. In this instance, all historical data pertaining to earlier sequences connected to the present state is regulated, updated, and kept (Yu, 2019). LSTM architecture consists of three gates. An input gate, a forget gate and an output gate, where x_t stands for the current input, C_t and C_{t-1} stand for the new and prior cell states, respectively, and $h_t$and h_{t-1} stand in for the current and past outputs. Fig. 4 depicts the internal structure of the LSTM. The following forms illustrate the LSTM input gate's basic operation.

$$i_t = \sigma(W_i.[h_{i-1}, x_t] + b_i) \tag{2}$$

$$\widetilde{C}_t = \tanh(W_i.[h_{i-1}, x_t] + b_i) \tag{3}$$

$$C_t = f_t C_{t-1} + i_t \widetilde{C}_t \tag{4}$$

where h_{t-1} and x_t are run through a sigmoid layer using Eq. (2) to determine which part of the information should be added. After h_{t-1} and x_t have traveled through the tanh layer, (3) is then used to retrieve new data. In (4), where W_i is a sigmoid output and $\widetilde{C}_t$ is a tanh output, the information from the present instant, $\widetilde{C}_t$, and long-term memory, C_{t-1} into C_t, are integrated. Here, b_i stands for the input gate bias of the LSTM, while W_i stands for weight matrices. The use of a sigmoid layer and a dot product together with the LSTM's forget gate then enables the selective transmission of information. The formula (5), where W_f is the weight matrix, b_f is the offset, and σ is the sigmoid function, is used to decide whether to forget relevant information from a previous cell with a certain probability.

$$f_t = \sigma(W_f.[h_{i-1}, x_t] + b_f) \tag{5}$$

The h_{t-1} and x_t inputs after (6) and (7) define the states necessary for

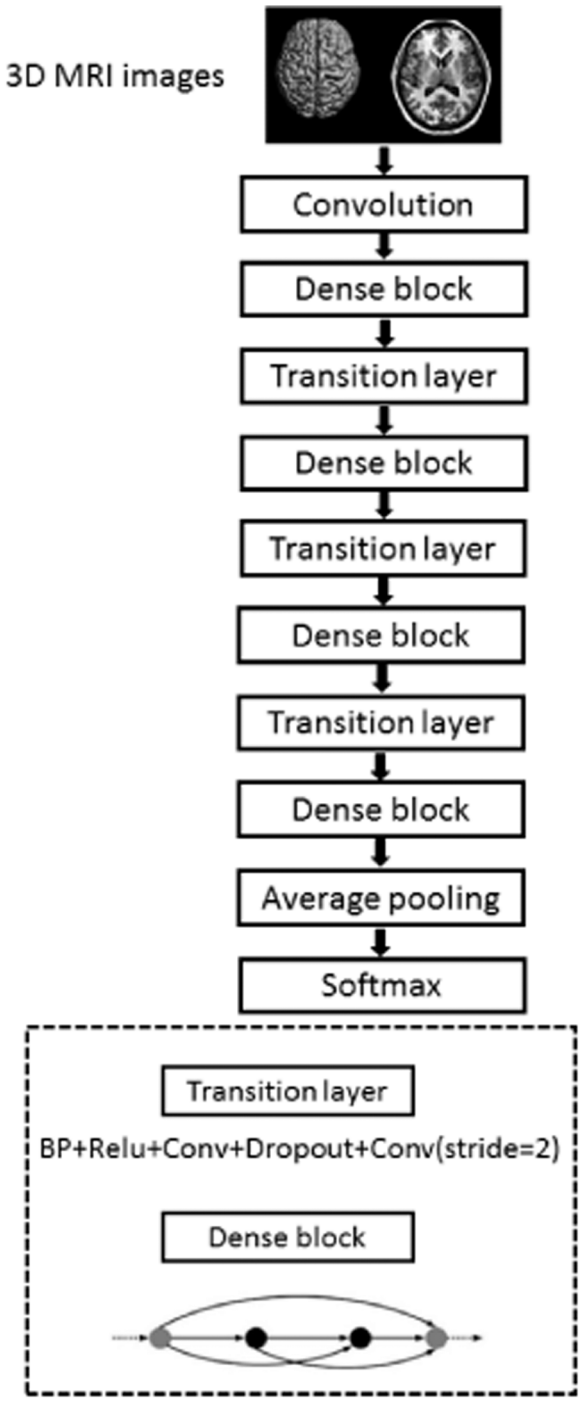


Fig. 3. The structure of DenseNet consisting of convolution layer, dense blocks, transition layers, average pooling and softmax layer.

continuation by the LSTM's output gate. The state decision vectors that convey new information,C_t, across the tanh layer are found and multiplied to get the final output.

$$O_t = \sigma(W_o.[h_{i-1}, x_t] + b_o) \tag{6}$$

$$h_t = O_t \tanh(C_t) \tag{7}$$

where b_o and W_o are the LSTM bias and the weighted matrices of the output gate, respectively.

Table 2
Parameters and architecture of our deep 3D DenseNet model.

Layers	Output Size	Filter size, stride, number
Input layer	72 × 64 × 64	-
Convolution	64, 33 × 29 × 29	7 × 7 × 7, 2, 64, conv
Dense block (1)	96, 33 × 29 × 29	$\begin{bmatrix} 1\times1\times1, 1, 64, \text{conv} \\ 3\times3\times3, 1, 16, \text{conv} \end{bmatrix} \times 2$
Transition layer	48, 17 × 15 × 15	$\begin{bmatrix} 1\times1\times1, 1, 48, \text{conv} \\ 2\times2\times2, 2, 48, \text{conv} \end{bmatrix}$
Dense block (2)	80, 17 × 15 × 15	$\begin{bmatrix} 1\times1\times1, 1, 64, \text{conv} \\ 3\times3\times3, 1, 16, \text{conv} \end{bmatrix} \times 2$
Transition layer	40, 9 × 8 × 8	$\begin{bmatrix} 1\times1\times1, 1, 40, \text{conv} \\ 2\times2\times2, 2, 40, \text{conv} \end{bmatrix}$
Dense block (3)	72, 9 × 8 × 8	$\begin{bmatrix} 1\times1\times1, 1, 64, \text{conv} \\ 3\times3\times3, 1, 16, \text{conv} \end{bmatrix} \times 2$
Last transition layer	36, 6 × 5 × 6	$\begin{bmatrix} 1\times1\times1, 1, 36, \text{conv} \\ 4\times4\times3, 1, 36, \text{conv} \end{bmatrix}$
Dense block (4)	68, 6 × × 6	$\begin{bmatrix} 1\times1\times1, 1, 64, \text{conv} \\ 3\times3\times3, 1, 16, \text{conv} \end{bmatrix} \times 2$
Average pooling	68, 1 × 1 × 1	-
Softmax layer	2	-

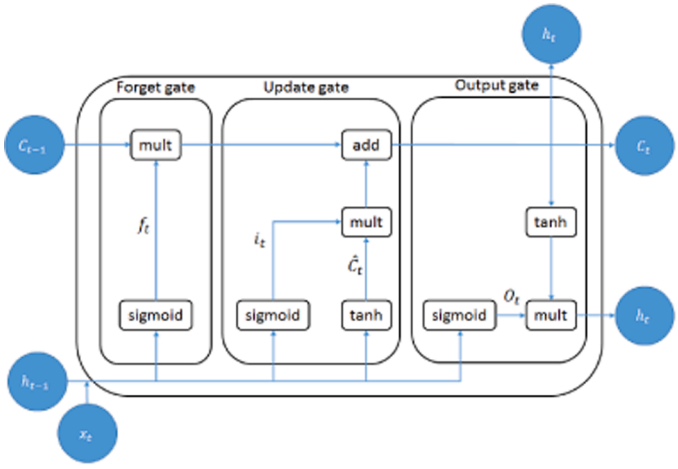


Fig. 4. LSTM cell structure (Hochreiter and Schmidhuber, 1997).

2.3.1. *Deep Bidirectional LSTMs (BiLSTM)*

The abovementioned LSTM models are extended by the deep-bidirectional LSTMs (Schuster and Paliwal, 1997), which employ two LSTMs to process the input data. An LSTM is used on the input sequence in the first round (i.e., forward layer). In the subsequent cycle, the LSTM model is given the input sequence's reverse version (i.e., backward layer). The accuracy of the model will increase as a result of applying the LSTM twice because it improves learning long-term relationships between time steps of time series or sequence data (Siami-Namini et al., 2019).

To effectively capture the temporal variations in correlations within AD, we utilize the BiLSTM technique. The BiLSTM employs both a forward LSTM and a backward LSTM, enabling us to capture the underlying dependencies and patterns present in the MRI sequence. In our approach, we train separate BiLSTM subnetworks for each individual time point in the MRI sequence. Since we are dealing with 7 time points, we train 7 independent one-layer BiLSTM networks concurrently. This allows us to focus on capturing the specific temporal features unique to each time point. Subsequently, we combine the learned features from these 7 networks and direct them towards an additional fully connected layer. This layer is responsible for extracting the shared features that are common among the different time points in the sequence. The architecture of the BiLSTM network used for extracting these temporal features is depicted in Fig. 5.

At each time point, the outputs of the forward and backward

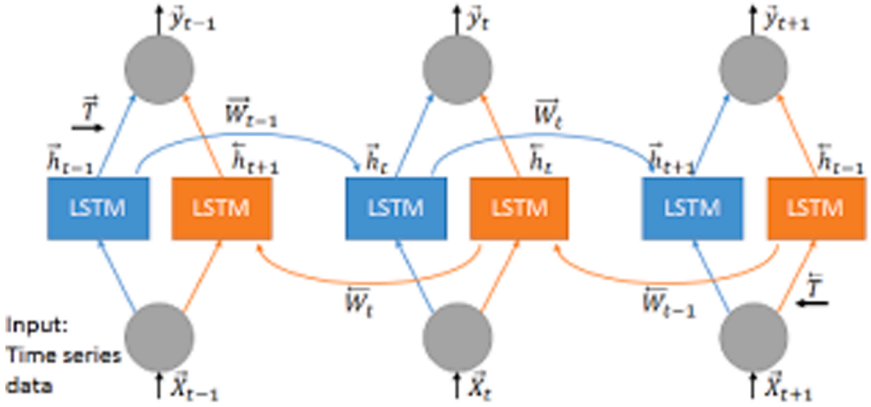


Fig. 5. Structure of BiLSTM network (Hochreiter and Schmidhuber, 1997).

BiLSTMs are merged by concatenating them together to form the output of the BiLSTM for that specific time point, as depicted in Fig. 5. Unlike a regular LSTM unit that captures information solely from the preceding input sequence in a time series, it does not account for the relationship between future and past sequences. In contrast, a BiLSTM(Schuster and Paliwal, 1997) combines two independent hidden LSTM layers, one in the forward direction and the other in the backward direction, to capture overall dependencies within a time series.

In the BiLSTM, an input sequence $X = (X_0, X_1, \ldots, X_{t+1})$ is processed by the forward hidden sequence $h_{tf} = (h_{0f}, h_{1f}, \ldots, h_{(t+1)f})$ and the backward hidden sequence $h_{tb} = (h_{0b}, h_{1b}, \ldots, h_{(t+1)b})$. The output vector of a hidden layer, denoted as $y_t = (y_0, y_1, \ldots, y_{t+1})$ for time steps t = 1, 2,., t, is formed by combining h_{tf} and h_{tb}, resulting in $y_t = [h_{tf}, h_{tb}]$, as shown in Eqs. (8) to (11).

$$h_{tnf} = \sigma\left(\theta_{h_{nf}} \bullet \left[h_{(tn-1)f}, X_{tn}\right] + b_{h_{nf}}\right) \tag{8}$$

$$h_{tnb} = \sigma\left(\theta_{h_{nb}} \bullet \left[h_{(tn-1)b}, X_{tn}\right] + b_{h_{nb}}\right) \tag{9}$$

$$\left(h_{t0f}, h_{t0b}\right) \ldots \left(h_{tnf}, h_{tnb}\right) = BiLSTM(X_0, X_1, \ldots, X_{t+1}) \tag{10}$$

$$y_t = \sigma(\theta_{y_t h_{nf}} h_{tnf} + \theta_{y_t h_{nb}} h_{tnb} + b_{y_t}) \tag{11}$$

The output of a BiLSTM y_t is fed as an input to the next layer.

2.4. The final classification

We build unique hybrid convolutional and recurrent neural networks for the GM volume map using DenseNets and BiLSTM. The connected BiLSTMs are employed to extract the high level correlation and temporal properties between images in multiple time points, while the 3D DenseNets are formed to learn more specific image and spatial information of Gray Matter for classification. In order to improve the final classification, one full connected layer is added after the hybrid neural networks merge the information they learnt from the GM density. Pre-training of individual 3D DenseNet and fine-tuning of the connected BiLSTM networks for the task-specific classification are both included in the training of the proposed hybrid convolutional and recurrent neural network. The DenseNets architecture is pre-trained using MR images from all time points in our implementation. Then, using a range of images, the pre-trained DenseNets are finetuned. The outputs of the fully connected layer are automatically mapped to the prediction scores of all class labels using the softmax classifier. The parameters of stacked BiLSTMs are fine-trained in conjunction with the higher fully connected and softmax prediction layers, whereas the initial-trained parameters of 3D DenseNets are fixed for all convolution and pooling, and fully connected layers. When the validation error rate stops dropping, the training iteration is complete. We determine the classification prediction probability for each test participant using the suggested hybrid deep network.

3. Experimental results

We initially describe the datasets and method implementation in this section. The comprehensive experiments we conducted to evaluate our methodology on the classifications of AD vs. NC, NC vs. MCI, sMCI vs. pMCI and MCI vs. AD are then presented. Finally, we discuss our findings after comparing our approach to other approaches.

3.1. Implementation details and hyperparameters

The proposed classification method based on the combination of DenseNet architecture and BiLSTM was tested in order to analyze the longitudinal MRI images obtained from the ADNI dataset. The size of the gray matter (GM) density map after preprocessing was 256 ×256×256 voxels. The proposed method was evaluated to classify AD vs. NC, NC vs. MCI, sMCI vs. pMCI and MCI vs. AD and the effectiveness results were obtained. The image set used for the evaluation was obtained from longitudinal scans of 684 ADNI participants over 4 years. These images include 193 NC, 132 pMCI, 185 sMCI and 174 AD. To obtain images with the size of 200 ×168×168 voxels, we removed pixels with zero density value from the images. These images are the inputs of the proposed classification algorithm. For each experiment, several deep learning models were trained with baseline and longitudinal data. An important advantage of time series data analysis is increasing the accuracy and reliability of the system when increasing the number of time steps. In this section, we implemented and performed tests on the deep learning model using DenseNet alone, BiLSTM alone, and the combination of DeneNet and BiLSTM architectures. In order to get better results and compare the performance of different architectures, training and testing of networks were done separately. DenseNet model was used to learn local features from longitudinal data, then BiLSTM model was used to learn temporal features. The results of the experiments showed that the combined model based on two architectures outperformed all other models, and that's why we used this model for Alzheimer's classification.

Python 3.10.6 is used to implement the suggested technique, and the DenseNet and BiLSTM are built using the Keras library and the Tensorflow backend. There are seven deep DenseNets that have been individually trained to extract spatial information for specific time points with the goal of classifying diseases. We set the initial learning rate to 0.001 and use a learning rate scheduler to dynamically adjust it during training. We used a supervised learning approach and trained the hybrid DenseNet-BiLSTM model using the Adam optimizer. Adam combines the advantages of adaptive learning rate methods and momentum-based optimization. When using the Adam optimizer (Kingma and Ba, 2014) to train model, the initial weights for the whole network are uniform,

and after 200 iterations, the networks are stable. Each model neuron uses PReLU activation, and the batch sizes are set to 64. In our study, we employ a cross-entropy loss function for binary classification tasks. The overfitting issue is also handled using the L2 regulation with parameter 0.01 and dropout layers with probability 0.10. The training process of the proposed models involves optimizing the parameters to achieve the best performance in the classification task. To ensure unbiased evaluation, we divided our dataset into two parts: a stratified training and validation set, which accounted for 90% of the data, and a separate test set comprising 10% of the data. We employed a stratification procedure to randomize the instances during each execution, ensuring that both the training and testing datasets had a similar distribution of all classes. This procedure was repeated ten times throughout all the reported experiments to eliminate any potential bias. To determine the optimal hyperparameters, we utilized a technique called stratified 10-fold cross-validation. This involved evaluating the hyperparameters of the final models by dividing the data into ten equally-sized subsets while maintaining the class distribution. Each subset was then used as a validation set while the remaining data served as the training set. This process was repeated ten times, ensuring that every subset was used as a validation set exactly once. By conducting this cross-validation, we were able to assess the performance of different hyperparameter configurations before selecting the final set. During our experiments, we input each modality of the data at the seven different time points into a masking layer, which was then followed by seven stacked BiLSTM layers. This architecture allowed us to effectively process and extract features from the input data.

The validation phase is used to adjust the training process' iteration count in order to provide model weights with the best possible performance. Since the tasks for classifying MCI vs. NC, MCI vs. AD and pMCI vs. sMCI are more difficult than those for classifying AD vs. NC, we implemented this by transferring the network parameters learnt for classifying AD vs. NC to initialize training the network for classifying MCI vs. NC, MCI vs. AD and pMCI vs. sMCI. The effectiveness of the proposed classification method was evaluated by the criteria of accuracy (ACC), sensitivity (SEN), specificity (SPE), receiver operating characteristic (ROC) curve and area under ROC curve (AUC). The proportion of individuals that are accurately classified among the entire population is used to calculate ACC. SEN is calculated as the percentage of positively identified samples among all positively identified samples. The SPE is calculated as the percentage of accurately categorized negative samples among all negative samples. Plotting the true positive rate (TPR) versus the false positive rate (FPR) at different thresholds based on the class prediction scores produces the ROC curve.

3.2. Disease classification results

To evaluate the efficiency of the proposed Alzheimer's disease classification method, we conducted several experiments. In order to compare our proposed framework with other models, we also performed tests on popular deep learning models such as VGG19 (Han et al., 2015) and ResNet50 (He et al., 2016). In the first experiment, we trained the VGG19 model solely with the problem data. The VGG19 model utilized a transfer learning strategy for controlled 3D MRI tests, initializing its model weights using image data from ImageNet (Deng et al., 2009). Similarly, in the second experiment, we trained the ResNet50 model using the same approach. Moving on, we conducted an experiment where we exclusively trained the DenseNet network using the gray matter density map, without considering the second part of the model, BiLSTM. Patient data from all time points were given as input to the network. DenseNet successfully extracted high-level features from the brain images and performed disease classification. In another experiment, we fed patients' longitudinal data into a BiLSTM network and trained the network. By extracting temporal features and hidden correlations, BiLSTM effectively classified the images.

Finally, we tested the proposed classification algorithm based on the combination of DenseNet architecture and stacked BiLSTM to extract both spatial and temporal features of the images for disease classification. The results were evaluated using Receiver Operating Characteristic (ROC) curves. Fig. 6 depict the ROC curves of the methods used to classify AD vs. NC, NC vs. MCI, sMCI vs. pMCI, and MCI vs. AD, respectively. The comparison of these curves reveals two important findings. Firstly, the proposed method demonstrated higher detection accuracy compared to other methods. Secondly, the correct classification of samples was notably better in the AD versus NC comparison than in other disease states. Overall, these experimental results highlight the superiority of our hybrid DenseNet-BiLSTM method in Alzheimer's disease classification, showcasing its ability to leverage both spatial and temporal information for improved accuracy and performance.

3.2.1. Classification Accuracy

The performance comparison of VGG19, ResNet50, DenseNet, BiLSTM architectures and the proposed hybrid deep method to classify AD vs. NC, NC vs. MCI, sMCI vs. pMCI and MCI vs. AD is given in Table 3. The best straightforward measure for comparing different approaches is accuracy, which counts the number of properly identified samples in a test set. In terms of sensitivity and specificity, the likelihood of misdiagnosis for each clinical label decreases with increasing values of these metrics. The results show that the proposed method based on the combination of DenseNet and BiLSTM can improve the classification efficiency compared to the single use of DenseNet and BiLSTM. With the proposed framework, we obtained higher classification accuracy than other methods. The classification accuracy results for different tasks are as follows:

- AD/CN Classification:
  The proposed framework achieved a classification accuracy of 95.28% in distinguishing between Alzheimer's disease patients and normal controls. This high accuracy demonstrates the effectiveness of our method in accurately identifying individuals with Alzheimer's disease.
- NC/MCI Classification:
  For the task of differentiating normal controls from individuals with mild cognitive impairment, the proposed framework achieved an accuracy of 88.19%. This indicates the capability of our method to identify early cognitive decline in individuals with MCI.
- sMCI/pMCI Classification:
  In classifying individuals with stable mild cognitive impairment (sMCI) and progressive mild cognitive impairment (pMCI), the proposed framework achieved an accuracy of 83.51%. This highlights the potential of our method to differentiate different stages of cognitive impairment.
- MCI/AD Classification:

The classification accuracy for distinguishing individuals with mild cognitive impairment from those with Alzheimer's disease was 92.14%. This result demonstrates the effectiveness of our method in identifying individuals at risk of progressing to Alzheimer's disease.

The suggested framework, which directly used longitudinal 3D MRI data, provided a considerably superior result for AD classification than the other methods, as shown by the experiment results in Table 3 and Fig. 6. It proved our hypothesis that the conserved spatial and time series data in 3D MRI data is important for AD detection. Additionally, it has been demonstrated that the DenseNet-BiLSTM model put out in this study is a successful solution for classifying MRI images and fully extracting the spatio-temporal properties of MRI data for AD diagnosis.

The next experiment aimed to analyze the effect of longitudinal data at different time steps on the classification of AD vs. NC, NC vs. MCI, sMCI vs. pMCI, and MCI vs. AD using the proposed architectures. In this experiment, we tested the proposed method for longitudinal analysis of images by gradually adding MRI images from subsequent visits. The results of this experiment are presented in Fig. 7. The horizontal axis

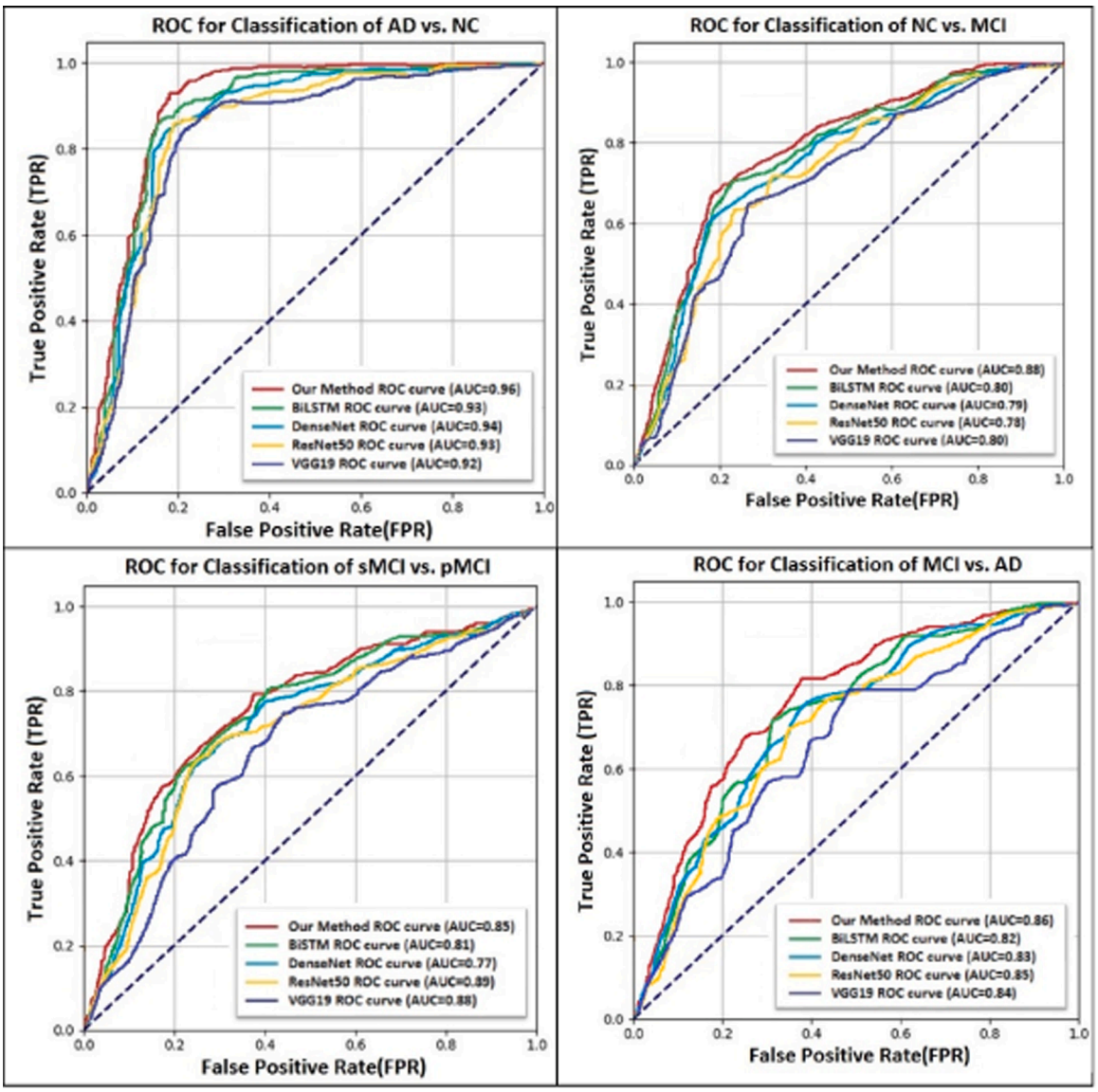


Fig. 6. - Comparison of ROC curves of the methods used to classify (a) AD vs. NC, (b) NC vs. MCI, (c) sMCI vs. pMCI and (d) MCI vs. AD, respectively.

represents the time steps, with m6, m12, m18, m24, m30, m36, and m48 denoting 6-month, 12-month, 18-month, 24-month, 30-month, 36-month, and 48-month intervals, respectively.

Fig. 7 illustrates that as the number of time steps increases and more longitudinal images are added to the model, the classification accuracy improves. It is evident that the proposed method successfully extracts spatio-temporal features from the longitudinal time series images, resulting in superior performance across all cases. Furthermore, the proposed method exhibits higher detection accuracy compared to other methods when considering the sequence of different inputs. These findings demonstrate the importance of incorporating longitudinal data and the effectiveness of the proposed method in capturing the spatio-temporal patterns present in longitudinal MRI images. By considering the progression of the disease over time, the proposed method achieves enhanced classification accuracy, providing valuable insights for early detection and monitoring of Alzheimer's disease.

3.3. Comparison with related studies

In this section, we compare the performance and effectiveness of our proposed method for classifying Alzheimer's disease with the methods used in recent studies. Table 4 provides a comparison of various studies based on eight main parameters. It is evident that several previous studies (El-Sappagh et al., 2021; Huang et al., 2019; Zandifar et al., 2020; Platero and Tobar, 2020; Dimitriadis et al., 2017; Beheshti et al., 2017; Ritter et al., 2015; Uysal and Ozturk, 2020; Bucholc et al., 2019) have utilized traditional classification methods such as K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) to analyze the progression of Alzheimer's disease. However, these methods generally do not achieve satisfactory diagnostic accuracy. Many studies have predominantly employed Magnetic Resonance Imaging (MRI) as the primary modality for Alzheimer's disease diagnosis (Cui et al., 2019; Dimitriadis et al., 2017; Beheshti et al., 2017; Uysal and Ozturk, 2020; Pan et al., 2020; Basaia et al., 2019; Abrol et al., 2020; Liu et al., 2020; Oh et al., 2019; Hong et al., 2019; Liu et al., 2019). Among the deep learning architectures used for MRI analysis, Convolutional Neural Networks (CNN) have been widely employed (Basaia et al., 2019; Oh et al., 2019; Liu et al., 2019; Spasov et al., 2019).

In contrast, the analysis of longitudinal data in the form of time series for Alzheimer's disease diagnosis has received less attention in the literature. This is primarily due to the challenges associated with obtaining complete patient data at regular time intervals and the difficulty in collecting such data. Only a limited number of studies have utilized longitudinal patient data (Cui et al., 2019; Platero and Tobar,

Table 3
The performance comparison of different methods with our proposed method. P-values are calculated using a paired t-test between the proposed and each competing method.

Method	AD/CN				NC/MCI				sMCI/pMCI				MCI/AD			
	ACC (p-value)	SEN (p-value)	SPE (p-value)	AUC	ACC (p-value)	SEN (p-value)	SPE (p-value)	AUC	ACC (p-value)	SEN (p-value)	SPE (p-value)	AUC	ACC (p-value)	SEN (p-value)	SPE (p-value)	AUC
VGG19	94.42 (4.102e-3)	88.99 (3.666e-3)	94.52 (3.015e-3)	92.58	87.11 (2.45e-3)	74.94(2.4e-4)	82.49 (3.4e-5)	80.28	72.12 (2.7e-5)	72.55 (8.7e-5)	79.54 (3.4e-4)	88.02	91.12 (1.5e-4)	78.85 (1.4e-6)	85.59 (5.6e-6)	84.18
ResNet50	94.20 (3.911e-3)	87.95 (3.521e-3)	93.78 (2.95e-4)	93.26	87.37 (7.54e-3)	73.57(8.1e-5)	82.77 (5.2e-5)	78.76	77.54 (8.9e-4)	72.19 (5.5e-5)	80.48 (8.8e-4)	89.20	91.46 (2.1e-3)	79.66 (9.5e-5)	87.78 (5.9e-5)	85.75
DenseNet	94.68 (4.514e-2)	89.72 (6.425e-2)	94.71 (3.152e-3)	94.01	87.90 (4.791e-2)	75.01 (5.287e-4)	83.67 (6.54e-4)	79.54	76.94 (4.9e-4)	73.68 (4.7e-4)	79.83 (2.5e-4)	77.82	91.87 (5.5e-3)	79.15 (7.6e-5)	86.41 (3.9e-5)	83.99
BiLSTM	94.93 (4.8657e-2)	88.76 (4.1251e-2)	95.16 (3.572e-2)	93.87	87.88 (4.608e-1)	77.24 (1.58e-3)	84.05 (4.7e-3)	80.97	79.45 (5.7e-3)	75.14 (6.1e-3)	83.11 (1.5e-3)	81.36	91.84 (4.8e-3)	84.16 (5.4e-4)	91.57 (2.4e-4)	82.74
DenseNet-BiLSTM	95.28	89.18	97.84	96.49	88.19	85.42	89.96	88.65	83.51	79.27	87.24	85.47	92.14	89.08	95.37	86.69

2020; Hong et al., 2019; Lu et al., 2017; El-Sappagh et al., 2020). However, these studies have not achieved significant diagnostic accuracy. For instance, in one study (Lu et al., 2017), a Gated Recurrent Unit (GRU)-based model was presented for multi-data fusion using four time points, achieving an accuracy of 81%. Another study (Cui et al., 2019) utilized a combined model of CNN and Bidirectional Gated Recurrent Unit (BGRU) with 3D MRI images, achieving accuracies of 91.33% for Alzheimer's disease versus normal controls (AD vs. NC) and 71.71% for progressive MCI versus stable MCI (pMCI vs. sMCI). In a different study (El-Sappagh et al., 2020), 15 time points were used for investigating the disease, resulting in an accuracy of 92.62%. Similarly, (Platero and Tobar, 2020) utilized six time points, and (Hong et al., 2019) utilized ten time points for classification. Among these studies, Recurrent Neural Networks (RNN) have been the most commonly employed neural network architecture for longitudinal data analysis (Cui et al., 2019; Platero and Tobar, 2020; Hong et al., 2019; Lu et al., 2017). It is worth noting that all these studies have utilized the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset.

Our proposed model has demonstrated excellent performance in terms of its ability to extract specific features from longitudinal data and accurately classify the disease. Our model outperforms state-of-the-art studies and has the potential to be applied in treatment decision systems. However, further improvements and adaptations are necessary to make this model viable in real treatment environments. Additionally, as future work, incorporating other patient information can enhance the categorization of the disease and improve the overall diagnostic accuracy.

4. Discussion

Various methods for classifying Alzheimer's disease using longitudinal MRI data analysis as time series have been proposed in the papers. CNN, RNN, SVM, LDA and LSTM models are mainly used in these methods. In a small number of these works, the combination of these models has been used. Considering the nature of longitudinal brain data, our proposed framework, by combining DenseNet and BiLSTM architectures, was able to extract spatial and temporal features of 3D images well and provide a competitive classification of diseases. DenseNet deep architecture was used to extract the spatial features of 3D MRI images, and BiLSTM network architecture was used in DenseNet output to sequentially extract high-level temporal features of MRI images. At the end, all the features extracted by a Flatten layer were combined and after passing through a fully connected layer, they entered the Softmax layer for classification. In previous works, the problems of overfitting, vanishing gradients and the problem of incomplete patient data were not paid attention to. In this research, we tried to solve these problems by presenting a new hybrid deep model. A new hybrid model based on deep learning architectures is proposed, which can extract the spatio-temporal characteristics of the images with high accuracy and provide a suitable classification of the disease. Our proposed framework uses whole-brain MRI scans without extracting regional brain volumes and cortical thickness. As a result, the proposed method reduces the calculation time and speeds up the processing. Regarding the calculation cost, the proposed method includes two stages of training and testing. The time spent to train the proposed deep network model was 4.1 hours. In the second stage, about half a second was spent to test the specific image. All tests were performed on a PC with Ubuntu OS and NVIDIA GTX1080 GPU and 12 GB memory.

5. Study Limitations and Future Research Directions

While our study demonstrates promising results in Alzheimer's disease prediction using longitudinal MRI analysis, there are several limitations that should be acknowledged. These limitations provide opportunities for future research and improvement in the field.

Firstly, our study focused on utilizing a single modality, namely

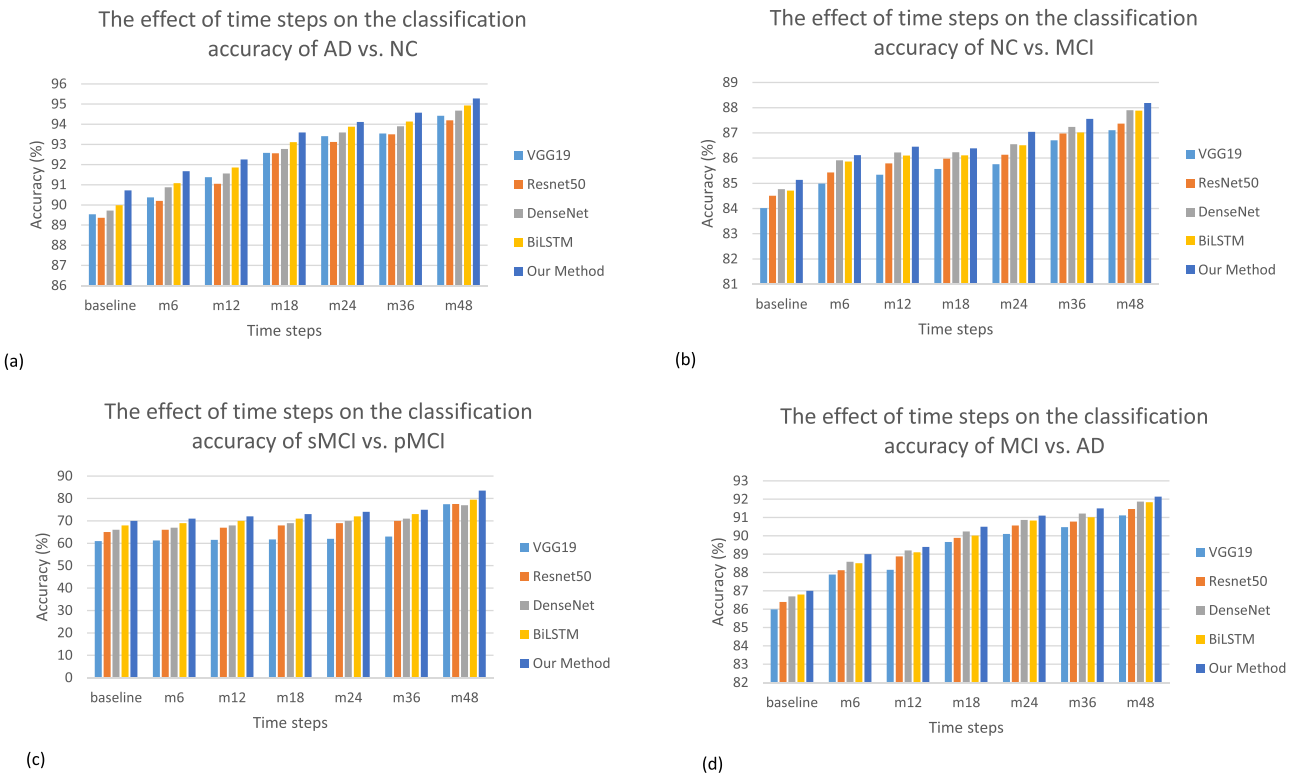


Fig. 7. The effect of increasing time steps on the classification performance of (a) AD vs. NC (b) NC vs. MCI (c) sMCI vs. pMCI (d) MCI vs. AD.

magnetic resonance imaging (MRI), for Alzheimer's disease prediction. While MRI provides high spatial resolution and detailed anatomical information, incorporating multiple modalities, such as positron emission tomography (PET) or cerebrospinal fluid (CSF) biomarkers, may further enhance the accuracy and reliability of the predictive models. Future research could explore the integration of multimodal data to improve the overall performance of Alzheimer's disease prediction models.

Secondly, the issue of missing data poses a challenge in the analysis of multimodality data. Collecting different modalities for the same subjects can be challenging, leading to a smaller number of subjects available for study. Additionally, there may be modality-wise missing data, which can impact the analysis and interpretation of the results. Future studies should aim to address these challenges by developing robust methodologies to handle missing data and investigate strategies to effectively combine information from diverse modalities.

Furthermore, our study focused on a specific population from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which may limit the generalizability of the findings to other populations or cohorts. Future research should aim to validate the proposed hybrid DenseNet-BiLSTM method on diverse and independent datasets, including different ethnicities, age groups, and disease stages, to assess its performance and generalizability across various populations.

In conclusion, while our study contributes to the field of Alzheimer's disease prediction using longitudinal MRI analysis, it is important to acknowledge the study limitations. The integration of multiple modalities, addressing missing data challenges and validating the proposed method on diverse datasets are all areas that warrant further investigation. Future research in these directions will provide valuable insights and advancements in the accurate diagnosis and understanding of Alzheimer's disease progression.

6. Conclusion

In this study, we presented a new method for Alzheimer's disease diagnosis based on the combination of DenseNet and BiLSTM architecture. Through feature reuse, DenseNet drastically decreases network parameters and avoids using numerous duplicate features. With lessened gradient vanishing, the convergence speed is also enhanced. The DenseNet network was used to extract the spatial features of the images at different time points, and the connected BiLSTM network was used to extract the longitudinal features of the images and to model high-level temporal variables for disease classification. The accuracy of the model will increase as a result of using the BiLSTM, which enhances learning long-term relationships between time steps of time series or sequences of MRI images. The results of various tests on the ADNI dataset have shown that firstly the proposed model is fully compatible with the characteristics of longitudinal MRI images and can be used for analyzing time series data and as a result classifying Alzheimer's disease and secondly the accuracy of diagnosing the disease with this deep method It has improved significantly compared to other methods used in the papers. The use of other imaging modalities such as fMRI, PET, EEG and their combination to diagnose the disease can be considered as future works. Also, in the future, we can work on topics such as image preprocessing, feature extraction, and applying this framework to diagnose other diseases.

Author Statement

First of all, we feel it is our duty to thank and appreciate the useful comments and suggestions of the esteemed referees. The requested corrections were made as much as possible in the article, and where there was no need to make more changes, they were answered with reasons. The places that have changed in the text of the article are marked with a special color. It means that each modification is highlighted with a color.

CRediT authorship contribution statement

Jomeiri Alireza: Writing – original draft. **Habibizad Navin Ahmad:**

Table 4
Comparison with other research works.

Study	subjects	Dataset	Modality	Algorithm	Training data	Classification results					
						AD/CN	CN/ MCI	sMCI/ pMCI	MCI/ AD	3 class	4 class
(El-Sappagh et al., 2021)	1029	ADNI	C, CSs, MH	LR, KNN, SVM, DT, RF	Baseline	100%	95.6%	-	-	87.69%	-
(Pan et al., 2020)	509	ADNI	MRI	Ensemble CNN	Baseline	84%	-	62%	-	-	-
(Huang et al., 2019)	290	ADNI	CSF, CSs, MRI	SVM	Baseline	-	-	86.4%	-	-	-
(Kruthika et al., 2019)	475	ADNI	MRI	Naïve Bayes, SVM, KNN	Baseline	-	-	-	81.3%	-	-
(Zandifar et al., 2020)	756	ADNI	MRI, CSs	Naïve Bayes	Baseline	-	-	87%	-	-	-
(Basaia et al., 2019)	229	ADNI	MRI	CNN	Baseline	99%	-	75%	-	-	-
(Lu et al., 2018)	1051	ADNI	FDG-PET	DNN	Baseline	93.58%	-	81.55%	-	-	-
(Abrol et al., 2020)	828	ADNI	MRI	ReseNet	Baseline	91%	-	77.8%	-	-	83.01%
(Platero and Tobar, 2020)	321	ADNI	MRI & N	LDA	Longitudinal	-	-	-	85%	-	-
(Dimitriadis et al., 2017)	400	ADNI	MRI	RF	Baseline	-	-	-	-	-	61.9%
(Liu et al., 2020)	449	ADNI	MRI	CNN, DenseNet	Baseline	88.9%	76.2%	-	-	-	-
(Fang et al., 2020)	906	ADNI	MRI, AV45 PET	GDCA	Baseline	-	-	-	-	75.28%	-
(Beheshti et al., 2017)	458	ADNI	MRI	SVM	Baseline	93.01%	-	75%	-	-	-
(Oh et al., 2019)	694	ADNI	MRI	CNN	Baseline	86.6%	-	-	-	-	-
(Oh et al., 2019)	1536	ADNI	MRI, PET, CSs, N	Stacked CNN-BiLSTM	Longitudinal	-	-	-	-	-	92.62%
(Cui et al., 2019)	830	ADNI	MRI	Stacked CNN-BGRU	Longitudinal	91.33%	-	71.71%	-	-	-
(Ritter et al., 2015)	237	ADNI	10 Modalities	SVM	Baseline	73%	-	-	-	-	-
(Uysal and Ozturk, 2020)	485	ADNI	MRI	LR, KNN, SVM,DT, RF	Baseline	94%	95%	-	87%	82%	-
(Bucholc et al., 2019)	488	ADNI	CSs, PET, MRI, CSF	SVM, Ridge	Baseline	-	-	-	-	83%	-
(Hong et al., 2019)	1105	ADNI	MRI	LSTM	Longitudinal	93.5%	73.9%	-	79.8%	77.7%	-
(Spasov et al., 2019)	785	ADNI	MRI, D, N, APOe4	CNN	Baseline	100%	-	86%	-	-	-
(Liu et al., 2019)	1984	ADNI	MRI	CNN	Baseline	-	-	-	-	-	51.8%
(Lu et al., 2017)	1618	ADNI	D, MRI, CSs, CSF	GRU	Longitudinal	-	90.53%	-	-	-	-
Our method	684	ADNI	MRI	Hybrid DenseNet & BiLSTM	Longitudinal	95.28%	88.19%	83.51%	92.14%	-	-

C: comorbidities, MH: Medication History, NP: Neuropathology, N: Neuropsychological, D: Demographics, 3-classes: CN/MCI/AD, 4-classes: CN/sMCI/pMCI/AD.

Supervision. **Shamsi Mahboubeh:** Writing – review & editing.

Data Availability

Data will be made available on request.

References

K.A. Matthews, W. Xu, A.H. Gaglioti, J.B. Holt, J.B. Croft, D. Mack, L.C. McGuire, Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged≥ 65 years, Alzheimer's & Dementia,, 2018.

M.D. Hurd, P. Martorell, A. Delavande, K.J. Mullen, K.M. Langa, Monetary costs of dementia in the United States, NEJM 368 (14) (2013) 1326–1334.

Colin L. Masters, Konrad Beyreuther, Alzheimer's centennial legacy: prospects for rational therapeutic intervention targeting the Aβ amyloid pathway, Brain Volume 129 (Issue 11) (2006) 2823–2839.

You, Senrong, et al. Fine perceptive gans for brain mr image super-resolution in wavelet domain. IEEE transactions on neural networks and learning systems (2022).

F. Márquez, M.A. Yassa, Neuroimaging Biomarkers for Alzheimer's Disease, Mol. Neurodegener. 14 (2019) 21.

Riyaj Uddin Khan, et al., A novel method for the classification of Alzheimer's disease from normal controls using magnetic resonance imaging, Expert Syst. 38 (1) (2021) e12566.

Peng Cao, et al., Sparse shared structure based multi-task learning for MRI based cognitive performance prediction of Alzheimer's disease, Pattern Recognit. 72 (2017) 219–235.

Xiaojing Long, et al., Prediction and classification of Alzheimer disease based on quantification of MRI deformation, PloS One 12 (3) (2017) e0173372.

Ramesh Kumar Lama, et al., Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features, J. Healthc. Eng. 2017 (2017).

Buhari Ibrahim, et al., Diagnostic power of resting-state fMRI for detection of network connectivity in Alzheimer's disease and mild cognitive impairment: A systematic review, Hum. brain Mapp. 42 (9) (2021) 2941–2968.

Sarraf, Saman, and Ghassem Tofighi. Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks. arXiv preprint arXiv:1603.08631 (2016).

Farheen Ramzan, et al., A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks, J. Med. Syst. 44 (2) (2020) 1–16.

Gaël Chételat, et al., Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias, Lancet Neurol. 19.11 (2020) 951–962.

Yiming Ding, et al., A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain, Radiology 290.2 (2019) 456.

Charles Marcus, Esther Mena, Rathan M. Subramaniam, Brain PET in the diagnosis of Alzheimer's disease, Clin. Nucl. Med. 39.10 (2014) e413.

Varvara Valotassiou, et al., SPECT and PET imaging in Alzheimer's disease, Ann. Nucl. Med. 32.9 (2018) 583–593.

Dariusz Świetlik, Jacek Białowąs, Application of artificial neural networks to identify alzheimer's disease using cerebral perfusion SPECT data, Int. J. Environ. Res. Public Health 16.7 (2019) 1303.

J.M. Górriz, et al., GMM based SPECT image classification for the diagnosis of Alzheimer's disease, Appl. Soft Comput. 11.2 (2011) 2313–2325.

Khvostikov, Alexander, et al. 3D CNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies. arXiv preprint arXiv:1801.05968 (2018).

Arijit De, Ananda S. Chowdhury, DTI based Alzheimer's disease classification with rank modulated fusion of CNNs and random forest, Expert Syst. Appl. 169 (2021) 114338.

Shengye Hu, et al., Bidirectional mapping generative adversarial networks for brain MR to PET synthesis, IEEE Trans. Med. Imaging 41.1 (2021) 145–157.

Ane Alberdi, Asier Aztiria, Adrian Basarab, On the early diagnosis of Alzheimer's Disease from multimodal signals: A survey, Artif. Intell. Med. 71 (2016) 1–29.

Yuanpeng Zhang, et al., Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion, Inf. Fusion 66 (2021) 170–183.

Yinghuan Shi, et al., Leveraging coupled interaction for multimodal Alzheimer's disease diagnosis, IEEE Trans. Neural Netw. Learn. Syst. 31.1 (2019) 186–200.

Xia-an Bi, et al., Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest, IEEE J. Biomed. Health Inform. 24.10 (2020) 2973–2983.

Binny Naik, Ashir Mehta, Manan Shah, Denouements of machine learning and multimodal diagnostic classification of Alzheimer's disease, Vis. Comput. Ind., Biomed., Art. 3 (1) (2020) 1–18.
Weiming Lin, et al., Multiclass diagnosis of stages of Alzheimer's disease using linear discriminant analysis scoring for multimodal data, Comput. Biol. Med. 134 (2023) 104478.
A. Anoop, et al., CSF biomarkers for Alzheimer's disease diagnosis, Int. J. Alzheimer'S. Dis. 2010 (2010).
H.M. Snyder, M.C. Carrillo, F. Grodstein, K. Henriksen, A. Jeromin, S. Lovestone, et al., Developing novel blood-based biomarkers for Alzheimer's disease, Alzheimer'S. Dement. 10 (1) (2014) 109–114.
Hannah Stocker, et al., The genetic risk of Alzheimer's disease beyond APOE ε4: systematic review of Alzheimer's genetic risk scores, Transl. Psychiatry 8.1 (2018) 1–9.
F. Márquez, M.A. Yassa, Neuroimaging Biomarkers for Alzheimer's disease, Mol. Neurodegener. 5 (2019) 1–14.
Saima Rathore, et al., A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages, NeuroImage 155 (2017) 530–548.
Falahati, Farshad, Westman, Eric, and Simmons, Andrew. Multivariate Data Analysis and Machine Learning in Alzheimer's Disease with a Focus on Structural Magnetic Resonance Imaging. 1 Jan. 2014, 685 – 708.
Haller, Sven, Lovblad, Karl O., and Giannakopoulos, Panteleimon. Principles of Classification Analyses in Mild Cognitive Impairment (MCI) and Alzheimer Disease. 1 Jan. 2011: 389 – 394.
Shuangzhi Yu, et al., Multi-scale enhanced graph convolutional network for early mild cognitive impairment detection. Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23, Springer International Publishing,, 2020.
Baiying Lei, et al., Predicting clinical scores for Alzheimer's disease based on joint and deep learning, Expert Syst. Appl. 187 (2022) 115966.
S. Lu, Y. Xia, W. Cai, M. Fulham, D.D. Feng, Early identification of mild cognitive impairment using incomplete random forest-robust support vector machine and FDG-PET imaging, Comput. Med. Imaging Graph 60 (2017) 35–41, https://doi.org/10.1016/j.compmedimag.2017.01.001 (Computational Methods for Molecular Imaging).
K.R. Kruthika, S. Rajeswari, H.D. Mahesappa, Multistage classiier-based approach for Alzheimer's disease prediction and retrieval. Inf. Med Unlocked 14 (2019) 34–42.
G. Uysal, M. Ozturk, Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods, J. Neurosci. Methods 337 (February) (2020) 1–9.
S. Lu, Y. Xia, W. Cai, M. Fulham, D.D. Feng, Early identification of mild cognitive impairment using incomplete random forest-robust support vector machine and FDG-PET imaging, Comput. Med. Imaging Graph 60 (2017) 35–41, https://doi.org/10.1016/j.compmedimag.2017.01.001 (Computational Methods for Molecular Imaging).
K. Ito, B. Corrigan, Q. Zhao, J. French, R. Miller, H. Soares, E. Katz, T. Nicholas, B. Billing, R. Anziano, T. Fullerton, Disease progression model for cognitive deterioration from alzheimer's disease neuroimaging initiative database, Alzheimer'S. Dement. J. Alzheimer'S. Assoc. 7 (2) (2011) 151–160, https://doi.org/10.1016/j.jalz.2010.03.018.
J. Zhou, J. Liu, V.A. Narayan, J. Ye, Modeling disease progression via multitask learning, NeuroImage 78 (2013) 233–248, https://doi.org/10.1016/j.neuroimage.2013.03.073.
F. Liu, L. Zhou, C. Shen, J. Yin, Multiple kernel learning in the primal for multimodal Alzheimer's disease classification, IEEE J. Biomed. Health Inf. 18 (3) (2014) 984–990, https://doi.org/10.1109/JBHI.2013.2285378.
S. Duchesne, A. Caroli, C. Geroldi, D.L. Collins, G.B. Frisoni, Relating one year cognitive change in mild cognitive impairment to baseline MRI features, NeuroImage 47 (4) (2009) 1363–1370, https://doi.org/10.1016/j.neuroimage.2009.04.023.
Iman Beheshti, et al., Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm, Comput. Biol. Med. 83 (2017) 109–119.
Peng Cao, et al., Sparse shared structure based multi-task learning for MRI based cognitive performance prediction of Alzheimer's disease, Pattern Recognit. 72 (2017) 219–235.
T. Tong, Q. Gao, R. Guerrero, C. Ledig, L. Chen, D. Rueckert, et al., A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease, IEEE Trans. Biomed. Eng. 64 (2017) 155–165.
Shengye Hu, et al., Medical image reconstruction using generative adversarial network for Alzheimer disease assessment with class-imbalance problem. 2020 IEEE 6th international conference on computer and communications (ICCC), IEEE,, 2020.
X. Liu, K. Chen, T. Wu, D. Weidman, F. Lure, J. Li, Use of multimodality imaging and artificial intelligence for diagnosis and prognosis of early stages of Alzheimer's disease, Transl. Res 194 (2018) 56–67.
C. Salvatore, A. Cerasa, I. Castiglioni, MRI Characterizes the Progressive Course of AD and Predicts Conversion to Alzheimer's Dementia 24 Months Before Probable Diagnosis, Front Aging Neurosci. 10 (2018 24) 135, https://doi.org/10.3389/fnagi.2018.00135. PMID: 29881340; PMCID: PMC5977985.
A.H. Syaifullah, A. Shiino, H. Kitahara, R. Ito, M. Ishida, K. Tanigaki, Machine Learning for Diagnosis of AD and Prediction of MCI Progression From Brain MRI Using Brain Anatomical Analysis Using Diffeomorphic Deformation, Front Neurol. 11 (2021 5) 576029, https://doi.org/10.3389/fneur.2020.576029. PMID: 33613411; PMCID: PMC7893082.
Raymond Wong, et al., Advances in computerized MRI-based biomarkers in Alzheimer's disease, Brain Sci. Adv. 7 (1) (2021) 26–43.
F. Zhang, S. Tian, S. Chen, Y. Ma, X. Guo, Voxel-based morphometry: improving the diagnosis of Alzheimer's disease based on an extreme learning machine method from the ADNI, Neuroscience 414 (2019) 273–279.
D. Chitradevi, S. Prabha, Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease, Appl. Soft Comput. J. 86 (2020) 105857.
O. Grimm, S. Pohlack, R. Cacciaglia, T. Winkelmann, M.M. Plichta, T. Demirakca, H. Flor, Amygdalar and hippocampal volume: a comparison between manual segmentation, Freesurfer and VBM, J. Neurosci. Methods 253 (2015) 254–261.
J. Zhang, Y. Wang, Q. Li, J. Shi, R.B. Iii, K. Chen, E.M. Reiman, R.J. Caselli, C. M. Stonnington, Patch-based sparse coding and multivariate surface morphometry for predicting amnestic mild cognitive impairment and Alzheimer's disease in cognitively unimpaired individuals, Alzheimer'S. Dement 12 (7) (2016) P947.
J. Wen, E. Thibeau-sutre, M. Diaz-melo, Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation, Med Image Anal. 63 (2020) 101694.
Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciompi F., Ghafoorian M., van der Laak J. A., van Ginneken B., Sánchez C.I. (2017) A survey on deep learn- ing in medical image analysis. arXiv:1702.057475.
T. Zhang, M. Shi, Multi-modal neuroimaging feature fusion for diagnosis of Alzheimer's disease, J. Neurosci. Methods 341 (2020) 108795.
R.R. Janghel, Y.K. Rathore, Deep convolution neural network based system for early diagnosis of Alzheimer's disease, IRBM 1 (2020) 1–10.
W. Li, X. Lin, X. Chen, Detecting Alzheimer's disease Based on 4D fMRI: an exploration under deep learning framework, Neurocomputing 388 (2020) 280–287.
K. Oh, Y. Chung, K.W. Kim, W. Kim, I. Oh, Classification and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning, Sci. Rep. 9 (2019) 1–16.
A. Punjabi, A. Martersteck, Y. Wang, T.B. Parrish, K. Katsaggelos, N. Initiative, Neuroimaging modality fusion in Alzheimer's classification using convolutional neural networks, PLoS ONE 14 (2019) 1–14.
H. Choi, K.H. Jin, Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging, Behav. Brain Res. 344 (2018) 103–109.
Spasov, S., Passamonti, L., Duggento, A., Lio, P., Toschi, N.: A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease 189, 383687 (2018).
X. Ding, M. Bucholc, H. Wang, D.H. Glass, H. Wang, D.H. Clarke, A.J. Bjourson, L.R. C. Dowey, M. O'Kane, G. Prasad, L. Maguire, K. Wong-Lin, A hybrid computational approach for efficient Alzheimer's disease classification based on heterogeneous data, Sci. Rep. 8 (2018) 1–10.
R. Cui, M. Liu, N. Initiative, RNN-based longitudinal analysis for diagnosis of Alzheimer's disease, Comput. Med. Imaging Graph. (2019).
N. Amoroso, D. Diacono, A. Fanizzi, M. La Rocca, A. Monaco, A. Lombardi, C. Guaragnella, R. Bellotti, S. Tangaro, Deep learning reveals Alzheimer's disease onset in MCI subjects: results from an international challenge, J. Neurosci. Methods 302 (2018) 3–9.
Liu et al. 2018) Liu Mingxia, et al., Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis, IEEE Trans. Biomed. Eng. 66.5 (2018) 1195–1206.
Ethan Ocasio, Tim Q. Duong, Deep learning prediction of mild cognitive impairment conversion to Alzheimer's disease at 3 years after diagnosis using longitudinal and whole-brain 3D MRI, PeerJ Comput. Sci. 7 (2021) e560.
Shaker El-Sappagh, Tamer Abuhmed, Kyung Sup Kwak, Alzheimer disease prediction model based on decision fusion of CNN-BiLSTM deep neural networks. Proceedings of SAI Intelligent Systems Conference, Springer,, Cham, 2020.
Yingying Zhu, et al., Long range early diagnosis of Alzheimer's disease using longitudinal MR imaging data, Med. Image Anal. 67 (2021) 101825.
Yingteng Zhang, Shenquan Liu, Xiaoli Yu, Longitudinal structural MRI analysis and classification in Alzheimer's disease and mild cognitive impairment, Int. J. Imaging Syst. Technol. 30 (2) (2020) 421–433.
Weiming Lin, et al., Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment, Front. Neurosci. 12 (2018) 777.
Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.
J.G. Sled, A.P. Zijdenbos, A.C. Evans, A nonparametric method for automatic correction of intensity nonuniformity in MRI data, IEEE Trans. Med. Imaging 17 (1998) 87–97.
S. Wang, R.M. Summers, Machine learning and radiology, Med. Image Anal. 16 (2014) 933–951.
D. Shen, C. Davatzikos, Very high resolution morphometry using mass-preserving deformations and HAMMER elastic registration, NeuroImage 18 (2003) 28–41.
C. Davatzikos, A. Genc, D. Xu, S.M. Resnick, Voxel-based morphometry using the RAVENS maps: Methods and validation using simulated longitudinal atrophy, NeuroImage 14 (2001) 1361–1369.
M. Liu, D. Zhang, D. Shen, Ensemble sparse classification of Alzheimer's disease, NeuroImage 60 (2012) 1106–1116.
Salehinejad, Hojjat, et al. Recent advances in recurrent neural networks. arXiv preprint arXiv:1801.01078 (2017).
Yong Yu, et al., A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (7) (2019) 1235–1270.
M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (11) (1997) 2673–2681.

Sima Siami-Namini, Neda Tavakoli, Akbar Siami Namin, The performance of LSTM and BiLSTM in forecasting time series. 2019 IEEE International Conference on Big Data (Big Data), IEEE,, 2019.
D. Kingma, J. Ba, Adam: a method for stochastic optimization, Comput. Sci. (2014).
S. Han, H. Mao, W.J. Dally et al., Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding arXiv:1510. 00149, 2015 - arxiv.org.
K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, IEEE Conference on Computer Vision and Pattern Recognition, 2016.
J. Deng, W. Dong, R. Socher, L. Li, K. Li, F. Li, ImageNet: A Large-Scale Hierarchical Image Database, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.
S. El-Sappagh, H. Saleh, R. Sahal, T. Abuhmed, S.R. Islam, F. Ali, E. Amer, Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data, Future Gener. Comput. Syst. 115 (2021) 680–699, https://doi.org/10.1016/j.future.2020.10.005.
K. Huang, Y. Lin, L. Yang, Y. Wang, S. Cai, L. Pang, X. wu, L. Huang, A multipredictor model to predict the conversion of mild cognitive impairment to Alzheimer's disease by using a predictive nomogram, Neuropsychopharmacology 45 (2019), https://doi.org/10.1038/s41386-019-0551-0.
A. Zandifar, V.S. Fonov, S. Ducharme, S. Belleville, D.L. Collins, MRI And cognitive scores complement each other to accurately predict Alzheimer's dementia 2 to 7 years before clinical onset, NeuroImage Clin. 25 (2020) 102121, https://doi.org/10.1016/j.nicl.2019.102121.
C. Platero, M. Tobar, Longitudinal survival analysis and two-group comparison for predicting the progression of mild cognitive impairment to Alzheimer's disease, J. Neurosci. Methods 341 (2020) 108698, https://doi.org/10.1016/j.jneumeth.2020.108698.
S. Dimitriadis, D. Liparas, M. Tsolaki, Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate healthy elderly, early/late MCI and Alzheimer's disease patients: from Alzheimer's disease neuroimaging initiative (ADNI) database, J. Neurosci. Methods 302 (2017), https://doi.org/10.1016/j.jneumeth.2017.12.010.
I. Beheshti, H. Demirel, H. Matsuda, Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm (http://dx.doi. org/), Comput. Biol. Med. 83 (2017) 109–119, https://doi.org/10.1016/j.compbiomed.2017.02.011.
K. Ritter, J. Schumacher, M. Weygandt, R. Buchert, C. Allefeld, J.-D. Haynes, Multimodal prediction of conversion to alzheimer's disease based on incomplete biomarkers. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 2015, pp. 206–215, 110.1016/j.dadm.2015.01.006.
G. Uysal, M. Ozturk, Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods, J. Neurosci. Methods 337 (2020) 108669, https://doi.org/10.1016/j.jneumeth.2020.108669.
M. Bucholc, X. Ding, H. Wang, D.H. Glass, H. Wang, G. Prasad, L.P. Maguire, A. J. Bjourson, P.L. McClean, S. Todd, D.P. Finn, K. Wong-Lin, A practical computerized decision support system for predicting the severity of Alzheimer's disease of an individual, Expert Syst. Appl. 130 (2019) 157–171, https://doi.org/10.1016/j.eswa.2019.04.022.
D. Pan, A. Zeng, L. Jia, Y. Huang, T. Frizzell, X. Song, Early detection of Alzheimer's disease using magnetic resonance imaging: A novel approach combining convolutional neural networks and ensemble learning, Front. Neurosci. 14 (2020) 259, https://doi.org/10.3389/fnins.2020.00259.
S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, NeuroImage Clin. 21 (2019) 101645, https://doi.org/10.1016/j.nicl.2018.101645.
A. Abrol, M. Bhattarai, A. Fedorov, Y. Du, S. Plis, V. Calhoun, Deep residual learning for neuroimaging: An application to predict progression to Alzheimer's disease (http://dx.doi.org/), J. Neurosci. Methods 339 (2020) 108701, https://doi.org/10.1016/j.jneumeth.2020.108701.
M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, M. Xu, A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease, NeuroImage 208 (2020) 116459, https://doi.org/10.1016/j.neuroimage.2019.116459.
K. Oh, Y.-C. Chung, K. Kim, W.-S. Kim, I.-S. Oh, Classification and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning, Sci. Rep. 9 (2019), https://doi.org/10.1038/s41598-019-54548-6.
X. Hong, R. Lin, C. Yang, N. Zeng, C. Cai, J. Gou, J. Yang, Predicting Alzheimer's disease using LSTM (http: //dx.doi.org/), IEEE Access 7 (2019) 80893–80901, https://doi.org/10.1109/ACCESS.2019.2919385.
M. Liu, J. Zhang, E. Adeli, D. Shen, Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis, IEEE Trans. Biomed. Eng. 66 (5) (2019) 1195–1206, https://doi.org/10.1109/TBME.2018.2869989.
S. Spasov, L. Passamonti, A. Duggento, P. Lió, N. Toschi, A parameterefficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease (http: //dx.doi.org/), NeuroImage 189 (2019) 276–287, https://doi.org/10.1016/j.neuroimage.2019.01.031.
S. Lu, Y. Xia, W. Cai, M. Fulham, D.D. Feng, Early identification of mild cognitive impairment using incomplete random forest-robust support vector machine and FDG-PET imaging, Comput. Med. Imaging Graph 60 (2017) 35–41, https://doi.org/10.1016/j.compmedimag.2017.01.001 (Computational Methods for Molecular Imaging).
S. El-Sappagh, T. Abuhmed, S. Riazul Islam, K.S. Kwak, Multimodal multitask deep learning model for Alzheimer's disease progression detection based on time series data (http: //dx.doi.org/), Neurocomputing 412 (2020) 197–215, https://doi.org/10.1016/j.neucom.2020.05.087.
K.R. Kruthika, H.D. Rajeswari, Maheshappa, Multistage classifier-based approach for Alzheimer's disease prediction and retrieval, Inform. Med. Unlocked 14 (2019) 34–42, https://doi.org/10.1016/j.imu.2018.12.003.
D. Lu, K. Popuri, G.W. Ding, R. Balachandar, M.F. Beg, Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease (http://dx.doi. org/), Med. Image Anal. 46 (2018) 26–34, https://doi.org/10.1016/j.media.2018.02.002.
C. Fang, C. Li, P. Forouzannezhad, M. Cabrerizo, R. Curiel, D. Loewenstein, R. Duara, M. Adjouadi, Gaussian Discriminative component analysis for early detection of Alzheimer's disease: A supervised dimensionality reduction algorithm, J. Neurosci. Methods 344 (2020) 108856, https://doi.org/10.1016/j.jneumeth.2020.108856.