Medical Image Analysis 75 (2022) 102266

Contents lists available at ScienceDirect

Medical Image Analysis

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

Assessing clinical progression from subjective cognitive decline to mild cognitive impairment with incomplete multi-modal neuroimages

Yunbi Liu[a,b,d], Ling Yue[c,*], Shifu Xiao[c], Wei Yang[b], Dinggang Shen[a], Mingxia Liu[a,*]

[a] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[b] School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[c] Department of Geriatric Psychiatry, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200240, China
[d] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, P.R. China

ARTICLE INFO

Article history:
Received 30 October 2020
Revised 4 October 2021
Accepted 7 October 2021
Available online 14 October 2021

Keywords:
Subjective cognitive decline
Conversion prediction
Image synthesis
Multi-modal neuroimage

ABSTRACT

Accurately assessing clinical progression from subjective cognitive decline (SCD) to mild cognitive impairment (MCI) is crucial for early intervention of pathological cognitive decline. Multi-modal neuroimaging data such as T1-weighted magnetic resonance imaging (MRI) and positron emission tomography (PET), help provide objective and supplementary disease biomarkers for computer-aided diagnosis of MCI. However, there are few studies dedicated to SCD progression prediction since subjects usually lack one or more imaging modalities. Besides, one usually has a limited number (e.g., tens) of SCD subjects, negatively affecting model robustness. To this end, we propose a Joint neuroimage Synthesis and Representation Learning (JSRL) framework for SCD conversion prediction using incomplete multi-modal neuroimages. The JSRL contains two components: 1) a generative adversarial network to synthesize missing images and generate multi-modal features, and 2) a classification network to fuse multi-modal features for SCD conversion prediction. The two components are incorporated into a joint learning framework by sharing the same features, encouraging effective fusion of multi-modal features for accurate prediction. A transfer learning strategy is employed in the proposed framework by leveraging model trained on the Alzheimer's Disease Neuroimaging Initiative (ADNI) with MRI and fluorodeoxyglucose PET from 863 subjects to both the Chinese Longitudinal Aging Study (CLAS) with only MRI from 76 SCD subjects and the Australian Imaging, Biomarkers and Lifestyle (AIBL) with MRI from 235 subjects. Experimental results suggest that the proposed JSRL yields superior performance in SCD and MCI conversion prediction and cross-database neuroimage synthesis, compared with several state-of-the-art methods.

1. Introduction

Alzheimer's disease (AD) is characterized by progressive impairment of memory and cognitive functions, affecting an estimated 47 million people worldwide (Jack et al., 2011; Association et al., 2018). The Chinese Longitudinal Aging Study (CLAS) reported that 4.5% of people older than 60 have AD dementia in China (Xiao et al., 2016). There is no effective cure for AD-related dementia, while early intervention may delay the progression. Hence, identifying AD at its early or preclinical stage is essential for drug development and timely intervention of AD-related cognitive decline. Previous studies have proven that AD pathology has already been ongoing in a long preclinical phase, such as subjective cognitive decline (SCD) (Jack et al., 2010) during which individuals have self-reported cognitive decline but no impairment on standardized cognitive tests (Jessen et al., 2014a). Considering that SCD individuals are at increased risk of developing AD and its prodromal stage (i.e., mild cognitive impairment, MCI) (Jack et al., 2013; Barnes et al., 2006; Amariglio et al., 2012; Buckley et al., 2016b; 2016a; Kryscio et al., 2014; Liu et al., 2017), it is clinically meaningful to predict the progression of SCD for drug development and possible intervention of AD-related cognitive decline.

Increasing evidence has shown that SCD is related to AD biomarkers, such as reduced glucose metabolism uptake on positron emission tomography (PET) and brain changes revealed by structural magnetic resonance imaging (MRI) (van der Flier et al., 2004; Striepens et al., 2010; Yue et al., 2018). Several studies have shown that SCD individuals are biologically different from cognitively normal (CN) subjects in terms of neuroimaging biomarkers (Stewart et al., 2008; Scheef et al., 2012; Jessen et al., 2014b; Yue et al., 2021). To the best of our knowledge, this is among the first attempts dedicated to multi-modal neuroimaging-based SCD

* Corresponding authors.
E-mail addresses: bellinthemoon@hotmail.com (L. Yue), mxliu@med.unc.edu (M. Liu).

https://doi.org/10.1016/j.media.2021.102266
1361-8415/

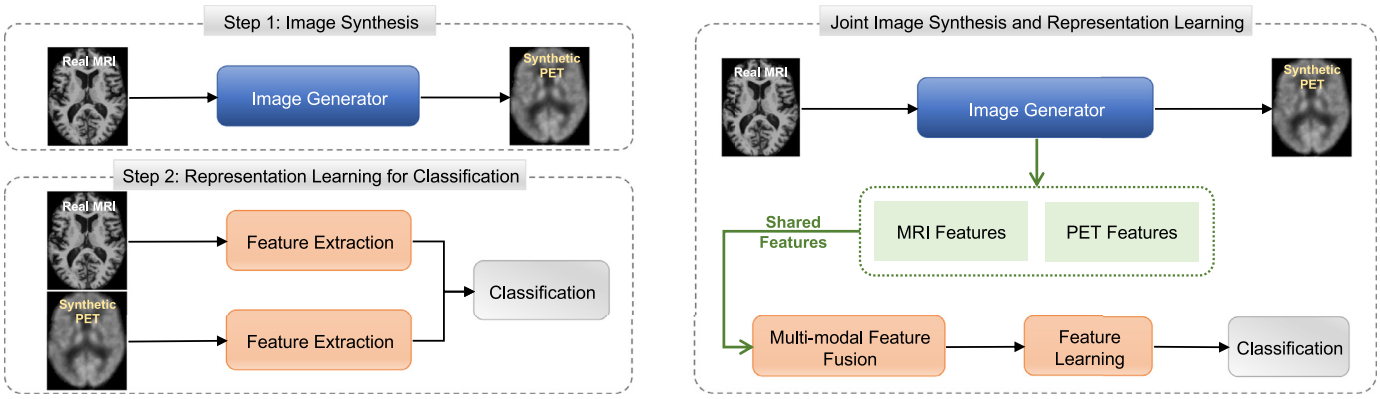

Fig. 1. Illustration of (a) conventional multi-modal methods for separate neuroimage synthesis and representation learning and (b) our Joint neuroimage Synthesis and Representation Learning (JSRL) framework using incomplete multi-modal neuroimages (i.e., MRI and PET).

progression/conversion prediction, i.e., predicting whether an SCD subject would convert to MCI within a period.

Existing studies have shown that MRI and PET contain complementary information for early diagnosis of brain disorders (Kawachi et al., 2006; Gray et al., 2013; Zu et al., 2016; Liu et al., 2018; Perrin et al., 2009; Zhou et al., 2018; Pan et al., 2021). It is usually challenging to collect complete multi-modal data from each subject in clinical practice, caused by patients' dropout or failed scans. Compared with MRI, it is generally more difficult to obtain PET data due to the relatively higher cost of PET scanning and other issues such as patients' concerns about radioactive exposure. As shown in Fig. 1(a), existing multi-modal methods usually first impute missing neuroimages/features, such as employing an image generator to impute the missing modality (Li et al., 2014; Sikka et al., 2018; Sharma and Hamarneh, 2019), and then perform feature learning and classification (Lassila et al., 2018; Cheng et al., 2015b; Campos et al., 2015; Cheng et al., 2017; Liu et al., 2014b; Yi et al., 2019; Pan et al., 2020). They treat imputation and multi-modal feature learning as two separate tasks, which may lead to suboptimal performance. Besides, we usually have a very limited number (e.g., tens) of subjects with SCD (Yue et al., 2021), bringing difficulty in constructing robust models and negatively affecting prediction performance (Cheng et al., 2015a; 2017).

To address these issues, we propose a Joint image Synthesis and Representation Learning (JSRL) framework for SCD conversion prediction based on incomplete multi-modal (e.g., MRI and PET) data. As shown in Fig. 1(b), our JSRL integrates image synthesis and multi-modal representation learning into a unified framework. The key idea is that these two tasks share the same multi-modal features (derived from the image generator) for image synthesis and classification, encouraging effective feature fusion for accurate prediction. To address the limited data problem, we further propose a transfer learning strategy for SCD conversion prediction by leveraging the JSRL learned from a large-scale ADNI database (with MRI and PET acquired from 863 subjects) to a small-scale SCD database (with only MRI from 76 subjects). The hypothesis is that, since SCD may be the preclinical stage of MCI/AD and AD is a progressive neurodegenerative disease, the discriminative brain changes between AD/MCI and cognitively normal subjects are potential biomarkers for SCD conversion prediction.

A preliminary version of this manuscript was previously published (Liu et al., 2020). This paper extends the preliminary version substantially with the following improvements. 1) We have added more experiments to compare our JSRL with its five variants as well as two data imputation methods to demonstrate the effectiveness of the proposed approach. 2) Besides the original task of SCD conversion prediction, we have applied our method to a related but different task (i.e., MCI conversion prediction). 3) Besides the initially used two datasets, i.e., ADNI (Jack et al., 2008) and CLAS (Xiao et al., 2016), we have extended our work to generate ADNI-like PET scans for the Australian Imaging, Biomarkers and Lifestyle (AIBL) dataset with original amyloid PET (Ellis et al., 2009).

The contributions of this work are summarized as follows.

- A joint image synthesis and representation learning (JSRL) framework is developed for SCD conversion prediction using incomplete multi-modal data. This is different from previous approaches that discard data-missing subjects (Young et al., 2013; Suk et al., 2014) or treat image/feature imputation and multi-modal feature learning as two separate tasks (Campos et al., 2015; Pan et al., 2020).
- A transfer learning strategy is developed to handle the limited data problem by leveraging JSRL learned from ADNI to CLAS and AIBL. Due to the lack of SCD subjects in ADNI, a reasonable label transfer strategy is proposed to partition training data in ADNI.
- Extensive experiments have been performed to validate the effectiveness of JSRL in SCD conversion prediction, MCI conversion prediction, and neuroimage synthesis.

The remainder of this paper is organized as follows. Section 2 reviews the most relevant studies. In Section 3, we introduce the materials and proposed method. In Section 4, we compare the proposed method with several competing methods for SCD conversion prediction, analyze the influence of several major components of JSRL, and apply JSRL to MCI conversion prediction and cross-database neuroimage synthesis. We further compare our method with previous studies and analyze the limitations of the current work in Section 5. Finally, this paper is concluded in Section 6.

2. Related work

This section briefly reviews the most related work, including studies on subjective cognitive decline (SCD), multi-modal neuroimage analysis, and transfer learning for early diagnosis of Alzheimer's disease (AD) and related disorders.

2.1. Studies on subjective cognitive decline

In recent years, more attention has been drawn to subjective cognitive decline. Some interesting findings have been found about

SCD. Jessen et al. (2014a) proposed to list several features associated with SCD, *i.e.*, age at onset of SCD $\geq$ 60, which increases the presence likelihood of preclinical AD. Caselli et al. (2014) compared self-based and informant-based SCD to distinguish emotional factors from early-stage AD, and found that those with incident MCI generally self-endorsed decline earlier than informants (e.g., their family members or friends). Some studies also found that SCD subjects with normal cognition might undergo AD pathology changes related to their neuroimaging biomarkers, which may gradually progress to the next stage, such as MCI (i.e., the prodromal phase of AD) (Yue et al., 2018; Striepens et al., 2010; Jack et al., 2010). With structural MRIs, Striepens et al. (2010) reported total hippocampus volume loss in SCD cohorts. Yue et al. (2018) found a ladder-shaped difference of left-larger-than-right asymmetry in the amygdala (i.e., MCI>SCD>CN) using MRIs. However, only a few neuroimage-based studies attempted to predict the conversion from SCD to MCI within a follow-up time (Yue et al., 2021; Liu et al., 2020). In (Yue et al., 2021), the authors proposed a simple prediction model with only five (three clinical and two structural MRI-derived) features to identify pSCD from sSCD, which is the first MRI-based study for SCD analysis. Their experimental results suggested that SCD individuals who could develop to MCI might be identified from sSCD with a combination of a few clinical information and structural MRI-based neuroimaging biomarkers. In this work, we propose to construct a multi-modal neuroimage-based model for SCD conversion prediction.

2.2. Multi-Modal neuroimage analysis

Previous studies have shown that different modalities (e.g., MRI and PET) can provide complementary information for identifying disease-related brain changes (Kohannim et al., 2010; Jack et al., 2008; Liu et al., 2015). In recent years, many researchers focused on multi-modal studies for computer-aided diagnosis of brain diseases (Pan et al., 2020; Cheng et al., 2015b; Liu et al., 2014b; Young et al., 2013; Hinrichs et al., 2011; Suk et al., 2014; Shi et al., 2017; Zu et al., 2016; Zhou et al., 2020; Thung et al., 2017; Hor and Moradi, 2016). For example, a Gaussian process classification method was developed, by using structural MRI, PET, cerebral spinal fluid (CSF) and apolipoprotein E (APOE) data to identify MCI patients who will convert to AD within a period (Young et al., 2013). Suk et al. (2014) developed a systematic method for a joint feature representation from paired patches of MRI and PET with a multi-modal Deep Boltzmann Machine (DBM). These methods make full of multi-modal information and achieve good performance. However, they usually need complete multi-modal neuroimages, making their methods less practical since collecting complete multi-modal data for each subject is usually expensive and time-consuming in clinical practice. To address the issue of incomplete data, some imputation/completion techniques have been utilized, such as Zero completion, *k*-nearest neighbor (KNN) and image generation by employing an image generator (Campos et al., 2015; Van Tulder and de Bruijne, 2015; Li et al., 2014; Sikka et al., 2018; Sharma and Hamarneh, 2019). Based on synthesized data, multi-modal prediction models can handle those subjects with incomplete data (Campos et al., 2015; Cheng et al., 2017; Pan et al., 2020). For example, Pan et al. (2020) proposed a hybrid generative adversarial network (HGAN) to impute missing neuroimages, and then developed a spatially-constrained Fisher representation network (SCFR) to extract statistical descriptors of neuroimages for disease diagnosis. However, these methods usually perform feature selection of each modality individually, ignoring the underlying association between MRI and PET.

2.3. Transfer learning for AD-related disease diagnosis

Computer-aided methods for identifying stable SCD (sSCD) and progressive SCD (pSCD) individuals are similar to those used for identifying MCI subjects who will convert to AD within a period. For early diagnosis of AD (e.g., SCD and MCI), the number of training samples is usually very limited, while the feature dimension is very high. This so-called small-sample-size problem has been one of the main challenges in neuroimaging data analysis, which may lead to overfitting issues. Many machine learning methods have been proposed to address this issue via transfer learning (Cheng et al., 2015b; 2017; Filipovych and Davatzikos, 2011; Da et al., 2014; Lian et al., 2020; Khan et al., 2019). Cheng et al. (2015b) developed a domain transfer feature selection method for MCI conversion prediction by using auxiliary data (i.e., AD and cognitively normal subjects) for discriminative feature selection. They further proposed a multi-domain transfer learning framework for the early diagnosis of AD by using multiple auxiliary domains (Cheng et al., 2017). Lian et al. (2020) verified the effectiveness of transferred knowledge for model training, in which the task of MCI conversion prediction was implicitly enriched by utilizing the supplementary information of AD and NC subjects. These transfer learning methods suggest that discriminative brain changes from cognitively normal (CN) to AD are potentially related to MCI, considering AD is a progressive neurodegenerative disease. Therefore, knowledge learned from AD-related data can also be used to promote the performance of SCD conversion prediction.

3. Materials and methodology

In this section, we first introduce studied subjects and image pre-processing, and then present the proposed method in detail.

3.1. Studied subjects and image pre-processing

Three datasets are used in this work: 1) the Chinese Longitudinal Aging Study (CLAS) database, 2) the Alzheimer's Disease Neuroimaging Initiative (ADNI-1 and ADNI-2) database, and 3) Australian Imaging, Biomarkers and Lifestyle (AIBL) database (Ellis et al., 2009). Baseline data in ADNI are used in this work, while subjects that appear in both ADNI-1 and ADNI-2 are removed from ADNI-2 for independent evaluation.

CLAS. It contains only 3T T1-weighted MRIs acquired from 76 SCD subjects. Following (Abdulrab and Heun, 2008), the SCD group in CLAS meets the following criteria: 1) the onset age of >60 years old; 2) the presence of gradual memory decline persisting for $\geq$ 6 months; 3) objective memory performance within normal range. According to follow-up outcomes after 7 years, SCD subjects are categorized into stable SCD (sSCD) and progressive SCD (pSCD). At the follow-up visit time (i.e., 7 years), 24 SCD subjects converted to MCI within 84 months after baseline, 52 SCD subjects kept cognitive normal, and none of them converted to AD at the follow-up visit. Therefore, there are 24 sSCD and 52 pSCD in CLAS.

ADNI-1 & ADNI-2. The ADNI-1 and ADNI-2 datasets contain all baseline T1-weighted MR images and partial fluorodeoxyglucose positron emission tomography (FDG-PET) images from 1) AD patients, 2) cognitively normal (CN) subjects, 3) MCI individuals, and 4) SCD (also called subjective memory complaint) subjects. There are a total of 1,145 subjects with complete MRI and FDG-PET scans from the baseline ADNI-1 and ADNI-2 datasets. Following (Hor and Moradi, 2016; Lian et al., 2020; Pan et al., 2020), we use the 36-month window to define the conversion of SCD and MCI subjects in ADNI-1 and ADNI-2. For example, if an SCD subject converts to MCI at the 24-th month after baseline and does not have data at the 36-month, we treat it as pSCD because it converts to MCI within 36 months. Only 16 SCD subjects have definite conversion

Table 1
Demographic information of studied subjects in the task of SCD conversion prediction. The values are denoted as "mean ± standard deviation". F/M: Female/Male.

Phase	Dataset	Category	Gender(F/M)	Age	MMSE
Training	ADNI-1&ADNI-2	sCN	95/110	74.0±5.5	28.9±1.1
		pCN	11/18	78.2±5.9	29.2±0.9
		MCI	254/375	72.9±7.5	27.5±1.8
Test	CLAS	sSCD	27/25	68.6±7.2	27.7±2.2
		pSCD	13/11	71.3±6.6	26.8±2.6

Table 2
Demographic information of studied subjects in the task of MCI conversion prediction.

Phase	Dataset	Category	Gender(F/M)	Age	MMSE
Training	ADNI-1&ADNI-2	sCN	95/110	74.0±5.5	28.9±1.1
		sMCI	129/196	71.9±7.6	27.9±1.7
		pMCI	67/97	74.0±6.9	26.8±1.6
		AD	100/145	74.9±7.8	23.5±2.2
Test	ADNI-1&ADNI-2	sMCI	24/42	74.3±7.9	27.7±1.9
		pMCI	37/56	74.5±7.2	26.5±1.7

results within 36 months (i.e., 11 pSCD and 5 sSCD) in ADNI, and they only have MRI scans. Similarly, we divide CNs into stable CN (sCN) and progressive CN (pCN), where pCN would convert to MCI within 36 months and sCN remains stable. The MCI subjects in ADNI-1 and ADNI-2 are further divided into stable MCI (sMCI) and progressive MCI (pMCI), where subjects with pMCI would convert to AD within 36 months after baseline and subjects with sMCI remain stable.

AIBL. A total of 235 subjects with paired T1-weighted structural MRI and Flutemetamol/Pittsburgh compound B (Flute/PIB)-PET scans were randomly selected from AIBL. It is worth noting that no FDG-PET data is available in this dataset. So this dataset is only used to evaluate image synthesis performance of the proposed method, and is not involved in the task of SCD/MCI conversion prediction.

Tables 1 and 2 show demographic and clinical information of studied subjects in two tasks, *i.e.*, SCD conversion prediction and MCI conversion prediction. In the task of SCD conversion prediction, we add MCI subjects with uncertain conversion results as the negative category considering the prediction from SCD to MCI, which helps to increase the sample size. In MCI conversion prediction, we use only MCI subjects (i.e., pMCI and sMCI) that have clear definitions on whether they would convert to AD within 36 months after baseline time.

In this work, all training subjects have paired MRI and PET scans, but the test CLAS and ADNI data have only MRI scans. All MRI scans were pre-processed through a standard pipeline, including the following steps: 1) skull-stripping by calling the BET function in Freesurfer (Fischl, 2012), 2) intensity inhomogeneity correction using Freesurfer, and 3) spatial normalization to the Montreal Neurological Institute (MNI) space using the Statistical Parametric Mapping (SPM)[1] (Ashburner and Friston, 2005). For each PET image, we first linearly aligned it to its corresponding MRI scan, and then normalized its intensity (via min-max scaling), followed by spatial normalization to the MNI space (using the same transformation as the one used to spatially normalize the corresponding MRI). Thus, each pair of MRI and PET for the same subject will have spatial correspondence. After pre-processing, all MRI and PET scans will have the same size as the MNI template (i.e., $181 \times 217 \times 181$ with 1 *mm* isotropic voxels).

[1] fil.ion.ucl.ac.uk/spm/software/.

3.2. Proposed method

The proposed JSRL is illustrated in Fig. 2, consisting of 1) an image synthesis subnetwork (IS) to synthesize missing neuroimages and generate multi-modal features and 2) a representation learning subnetwork (RL) for multi-modal feature fusion, feature learning, and classification/prediction. These two subnetworks share the same imaging features, e.g., $[F_{M1}, F_{M2}]$ for MRI and $[F_{P1}, F_{P2}]$ for PET, encouraging that the association between MRI and PET can be effectively modeled and the multi-modal representations are prediction-oriented. At the training phase, our JSRL network is trained on 863 subjects with complete MRI and PET from the ADNI dataset. At the test phase, the trained model is directly applied to 76 subjects with only MRI from the CLAS dataset. Besides, a label transfer strategy is proposed to leverage the relatively large-scale training data in the SCD-related domain (i.e., ADNI) to aid the SCD conversion prediction task in the small-scale CLAS dataset.

3.2.1. Problem formulation

In this work, we aim to utilize the domain knowledge learned from ADNI with MRI and PET scans to aid the SCD conversion prediction with only MRI in CLAS. Let $\mathbf{I}_M$ denote the domain of MRI scans and $\mathbf{I}_P$ be the domain of PET images. We denote a set of subjects (with paired MRI and PET scans) as $\mathcal{D} = \{(\mathbf{x}_M, \mathbf{x}_P) \,|\mathbf{x}_M \in \mathbf{I}_M, \mathbf{x}_P \in \mathbf{I}_P\}$. If one has MRI but no PET data, an image generator G_M: $\mathbf{I}_M \rightarrow \mathbf{I}_P$ trained on paired multi-modal data can impute missing PET based on its corresponding MRI. With real MRI and synthetic PET, our multi-modal classification/prediction model is formulated as

$$y = C(\mathbf{x}_M, G_M(\mathbf{x}_M)), \tag{1}$$

where y is the class label (e.g., stable SCD or progressive SCD), and C is a classifier that tells whether an SCD subject will convert to MCI within a certain period. As shown in Fig. 1(a), conventional methods usually treat image synthesis and representation learning as two separate tasks, by first learning G_M and then training C independently. The proposed JSRL incorporated image synthesis and representation learning into a unified framework, by sharing multi-modal features (generated by G_M) between the IS and the RL subnetworks. Based on the shared multi-modal features (i.e., $[F_{M1}, F_{M2}]$ for MRI, and $[F_{P1}, F_{P2}]$ for PET), the prediction model can be formulated as

$$y = C(F_{M1}, F_{M2}, F_{P1}, F_{P2}). \tag{2}$$

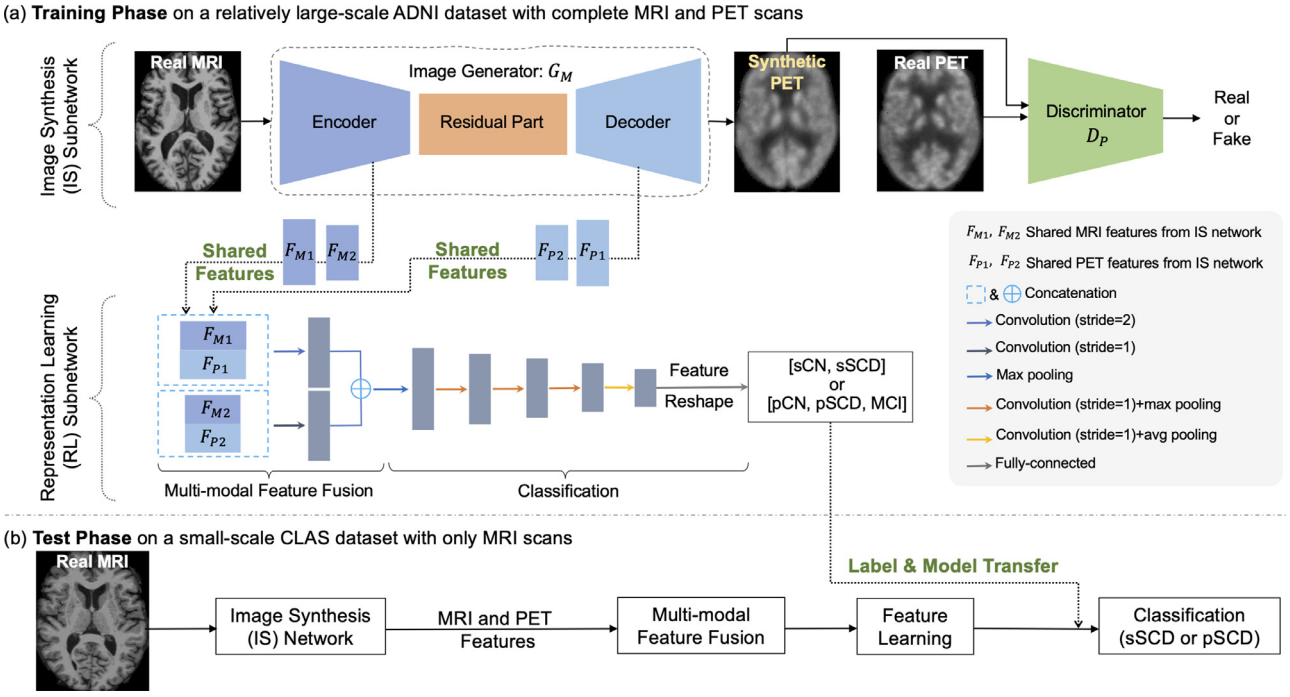

Fig. 2. Illustration of the proposed JSRL framework for SCD conversion prediction with an image synthesis (IS) subnetwork and a representation learning (RL) subnetwork at the training phase (a), and prediction for an unseen subject at the test phase (b). A transfer learning strategy is employed in JSRL.

3.2.2. Joint image synthesis and representation learning

1) Image Synthesis (IS) Network : As shown in the top panel of Fig. 2, our image synthesis (IS) model is based on a generative adversarial network (GAN) to synthesize PET and generate multi-modal features, with a generator (i.e., G_M) and a discriminator (i.e., D_P). The generator has an encoder with 3 convolutional (Conv) layers, a residual part with 6 residual blocks, and a decoder with 2 deconvolutional (Deconv) layers and 1 output Conv layer. The channels of 3 Conv layers are 16, 32 and 64, respectively, with each Conv layer followed by instance normalization and *relu* activation. In the decoder, the channels of 2 Deconv layers are 32 and 16, respectively, and the output Conv layer is followed by instance normalization and *tanh* activation. The first and the output Conv layers in G_M have the filter size of $7 \times 7 \times 7$, while the remaining Conv and Deconv layers have the filter size of $3 \times 3 \times 3$. The discriminator contains 5 Conv layers with the channels of 32, 64, 128, 256 and 1, respectively. The first 3 Conv layers are followed by instance normalization and leaky relu activation with the leaky rate of 0.2.

Two types of loss functions are used in G_M, i.e., an adversarial loss and a reconstruction loss, defined as

$$\mathcal{L}(G_M) = \sum_{\{\mathbf{x}_M \in \mathbf{I}_M, \mathbf{x}_P \in \mathbf{I}_P\}} \|G_M(\mathbf{x}_M) - \mathbf{x}_P\|_1 + \log\left(1 - D_P(G_M(\mathbf{x}_M))\right), \tag{3}$$

where the 1st term is the reconstruction loss (with $\|\cdot\|_1$ denoting the l_1 norm), and the 2nd term is the adversarial loss. For D_P, we aim to minimize the following loss function

$$\mathcal{L}(D_P) = \sum_{\{\mathbf{x}_M \in \mathbf{I}_M, \mathbf{x}_P \in \mathbf{I}_P\}} \log\left(1 - D_P(\mathbf{x}_P)\right) + \log\left(D_P(G_M(\mathbf{x}_M))\right). \tag{4}$$

Constrained by $\mathcal{L}(G_M)$ and $\mathcal{L}(D_P)$, the trained IS network is encouraged to generate PET images that are visually similar to their corresponding real ones, and also generate multi-scale imaging features via the encoder and decoder parts. The multi-scale features will be shared with the following representation learning (RL) network for SCD conversion prediction. Note that the proposed IS model can also generate synthetic MRI scans based on PET images. In this work, we focus on generating synthetic PET images based on MRI. The reason is that in clinical practice, we usually have MRI but lack PET data, since PET imaging is relatively more expensive and invasive while MRI imaging is less costly and noninvasive.

2) Representation Learning (RL) Network : To capture the association of different modalities, we develop a representation learning (RL) network to fuse multi-scale MRI and PET features for SCD conversion prediction. Considering that the generator constructs a mapping between MRI and PET from the same subject, their features generated by the IS network naturally convey their underlying relevance. Accordingly, a *multi-scale feature sharing mechanism* is designed to facilitate knowledge transfer between the IS and the RL subnetworks. Specifically, the MRI feature maps (denoted as F_{M1} and F_{M2}) of the first two Conv layers in the encoder and the PET feature maps of the last two Deconv layers in the decoder (denoted as F_{P1} and F_{P2}) of G_M are shared by IS and RL networks.

As shown in the middle panel of Fig. 2, the RL network consists of two major components: 1) a *multi-modal feature fusion module* to fuse multi-scale feature maps generated by the IS network, and 2) a *classification module* to learn high-level semantic features and perform prediction. Denote the concatenation of F_{M1} and F_{P1} as $C1$, and the concatenation of F_{M2} and F_{P2} as $C2$. In the multi-modal feature fusion module, $C1$ is followed by a Conv layer with a stride of 2, while $C2$ is followed by a Conv layer with a stride of 1. Then, these features are concatenated, followed by a max-pooling operation. The fused features are then fed into the classification module. This classification module contains four successive Conv layers to learn high-level semantic features and a fully-connected (FC) layer for prediction, where the first 3 and the last Conv layers are followed by the max-pooling and average-pooling (stride: 2), respectively. These four Conv layers have 16 channels with the filter size of $3 \times 3 \times 3$. Feature maps of the last Conv layer are reshaped into a feature vector (with 1,280 neurons), followed by the FC layer for classification/prediction. We propose a hybrid loss function in RL

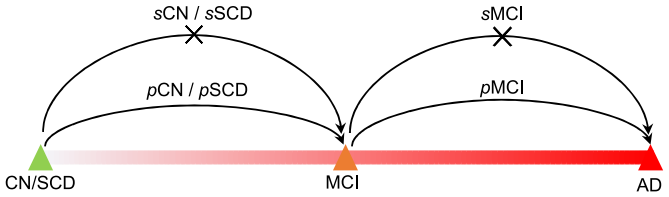

Fig. 3. Illustration of possible progression of AD-related categories, i.e., cognitive normal (CN), subjective cognitive decline (SCD), mild cognitive impairment (MCI) and Alzheimer's disease (AD). The term pCN/pSCD denotes CN/SCD subjects that would convert to MCI within a period, and pMCI denotes subjects that would convert to AD within a period. And sCN, sSCD, and sMCI subjects remain stable.

network for classification, defined as

$$\mathcal{L}_{hybrid} = \mathcal{L}(G_M) + \lambda L_C, \tag{5}$$

where L_C is the universal cross-entropy loss that is defined as

$$\mathcal{L}_C = -y \log p(\mathbf{x}) - (1-y) \log (1 - p(\mathbf{x})), \tag{6}$$

where $p(\mathbf{x})$ is the estimated probability of $\mathbf{x}$ belonging to the correct class y. We use the hybrid loss to jointly train the generator and the RL network, encouraging that the to-be-generated multi-modal representations by the IS network are prediction-oriented. The parameter λ is empirically set as 1 in this work.

3.2.3. Transfer learning

To deal with limited training data, we propose a transfer learning solution by leveraging knowledge from the relatively large-scale ADNI dataset to the small-scale CLAS dataset. A *label transfer* strategy is designed to augment training samples, and a *model transfer* strategy is used for neuroimage imputation and SCD conversion prediction.

Figure 3 illustrates the possible progression of four categories, including CN, SCD, MCI and AD. Since pCN and pSCD subjects would convert to MCI within a period, we assume that they have similar brain changes in their neuroimages. We denote subjects belonging to the five categories (i.e., sNC, pNC, sSCD, pSCD, and MCI) as *SCD-adjacent subjects*, and use them to aid SCD conversion prediction due to their close relationship (i.e., preclinical or prodromal stage of AD). We reasonably regard sNC and sSCD as positive samples, and treat pNC, pSCD and MCI as negative samples in ADNI. Through this label transfer strategy, we train JSRL for joint image synthesis and prediction on ADNI, and then apply the trained model to CLAS via model transfer. Thus, the knowledge learned from ADNI can be transferred to CLAS for SCD conversion prediction.

3.2.4. Implementation

The training of JSRL was performed via two steps. 1) In the 1st step, we trained the GAN (with G_M and D_P) for 50 epochs in the IS network to impute the missing neuroimages and generate multi-modal imaging features. We first trained G_M by minimizing $\mathcal{L}(G_M)$ with fixed D_P, and then trained D_P by minimizing $\mathcal{L}(D_P)$ with fixed G_M, iteratively. 2) In the 2nd step, we jointly trained the generator G_M and our RL network for 30 epochs for multi-modal feature learning and prediction. The Adam solver with a learning rate of 2×10^{-3} was used in the IS network. The gradient descent optimizer with a learning rate of 10^{-2} was used in the joint training of G_M and the RL network.

For efficient optimization, we used different training data in different tasks as shown in Tables 1-2. 1) For SCD conversion prediction (i.e., pSCD vs. sSCD classification), we used MRI and PET data of 863 subjects (including 205 sCN, 29 pCN and 629 MCI) from ADNI to train our JSRL model. 2) For MCI conversion prediction (i.e., pMCI vs. sMCI classification), we used 940 subjects (including 205 sCN, 325 sMCI, 164 pMCI and 245 AD) from ADNI to train our JSRL network. The code has been made publicly available[2].

4. Experiments

In this section, we first evaluate our method in the task of SCD conversion prediction, and then perform an ablation study to compare different variants of our proposed method. We further evaluate the effectiveness of our method on another challenging task, *i.e.*, MCI conversion prediction. Finally, we evaluate the cross-database neuroimage synthesis performance of our method on ADNI, CLAS and Australian Imaging, Biomarkers and Lifestyle (AIBL) databases (Ellis et al., 2009).

4.1. Results of SCD conversion prediction

Experimental Setup We first evaluate the performance of our JSRL in SCD conversion prediction, by comparing JSRL with three multi-modal methods using different handcrafted features, including 1) gray matter (GM) volume within 116 regions-of-interest (denoted as **ROI**) (Tzourio-Mazoyer et al., 2002; Rusinek et al., 1991), 2) patch-based morphology (**PBM**) method (Liu et al., 2014a), and 3) landmark-based local energy patterns (**LLEP**) (Zhang et al., 2017). Features of MRI and PET are concatenated in these three methods, followed by a linear support vector machine (SVM) for prediction. We also compare JSRL with three methods with different imputation ways, including Zero, KNN (Campos et al., 2015) and a state-of-the-art deep learning method called **HGAN+SCFR** (Pan et al., 2020) with a hybrid generative adversarial network (HGAN) and a spatially-constrained Fisher representation (SCFR) network. The HGAN+SCFR method performs image/feature imputation and neuroimaging feature learning separately, which is different from our JSRL model that treats image synthesis and classification jointly. Besides, the HGAN+SCFR method uses a late fusion strategy that first extracts imaging features of MRI and PET individually and then concatenates them at the last several Conv layers for prediction. In contrast, the proposed JSRL uses an early fusion strategy that fuses the multi-modal features through a multi-modal feature fusion module in the first several Conv layers of the RL network for prediction. In addition, the HGAN+SCFR inputs MRI and real/synthetic PET images, while our JSRL inputs multi-modal features generated from the same generator (in the image synthesis network) for prediction. We reasonably infer that the underlying relationship between the multi-modal features can be better conveyed by generating them from the same generator and fusing them using an early fusion strategy for prediction.

Note that three methods (i.e., ROI, PBM, and LLEP) cannot synthesize missing PET images, and they extract PET features from the real PET images from ADNI databases during the training phase and our synthetic PET scans via JSRL during the test phase. Besides, we compare our method with two different feature imputation methods based on ROI-based gray matter volume features, including **Zero** and **KNN** methods (Campos et al., 2015). In the Zero method, we simply fill the missing values with zeros. In the KNN method, we fill missing features with its k-nearest neighbor rows in the training set. We also compare our method with the single-modal counterpart of each competing method, including **ROI-M, PBM-M, LLEP-M** and **HGAN+SCFR-M** that use only MRI data. More details on the competing methods can be found in *Supplementary Materials*.

For a fair comparison, all competing methods use the same transfer learning strategy as our JSRL. That is, they use SCD-adjacent subjects as training data (i.e., label transfer), and their

[2] https://github.com/Candyeeee/JSRL

Table 3
Results achieved by different methods in SCD conversion prediction (i.e., pSCD vs. sSCD classification), with models trained on ADNI and tested on CLAS. Methods marked as "-M" denote that only subjects with MRI in ADNI are used for training, while the remaining methods employ all subjects with complete MRI and PET in ADNI for training. All 76 SCD subjects in CLAS are used as test data.

Method	AUC	BAC	SPE	SEN	F1S
ROI-M	0.621	0.642	**0.784**	0.500	0.511
PBM-M	0.592	0.516	0.490	0.542	0.413
LLEP-M	0.570	0.598	0.529	0.667	0.500
HGAN+SCFR-M	0.669	0.632	0.596	0.667	0.525
ROI	0.652	0.614	0.686	0.542	0.490
PBM	0.546	0.554	0.608	0.500	0.429
LLEP	0.611	0.536	0.529	0.542	0.426
Zero	0.648	0.647	0.711	0.583	0.528
KNN	0.652	0.646	0.750	0.542	0.520
HGAN+SCFR	0.632	0.553	0.539	0.583	0.452
JSRL (Ours)	**0.747**	**0.721**	0.692	**0.750**	**0.621**

models are trained on ADNI and tested on CLAS (i.e., model transfer). When training deep models, such as HGAN+SCFR-M, HGAN+SCFR and ours, we randomly partition 20% of each category in the training data as the inner validation set to optimize hyper-parameters (i.e., the number of epoch and learning rate) and determine their values when the corresponding model achieves the best performance (in terms of AUC) on the validation set. Then, we train the models on all training data by fixing these hyper-parameters and apply the trained models to the test set. The performance of SCD conversion prediction is measured by the area under the receiver operating characteristic (AUC), balanced accuracy (BAC), sensitivity (SEN), specificity (SPE), and F1-Score (F1S).

Experimental Results The results achieved by nine methods in SCD conversion prediction are reported in Table 3. As can be seen from Table 3, our JSRL method outperforms conventional hand-crafted feature based methods (i.e., ROI, PBM, and LLEP), the deep learning method (i.e., HGAN+SCFR), and two feature imputation methods (i.e., Zero and KNN) in most cases. This suggests the effectiveness of the proposed joint learning framework for image synthesis and multi-modal representation learning. Especially, the proposed JSRL achieves an improved SEN value (i.e., 0.750), which is 8.3 percent higher than the second-best SEN result (i.e., 0.667 achieved by LLEP-M and HGAN+SCFR-M). In real-world applications, the high sensitivity of JSRL may be very useful to accurately identify subjects with progressive SCD. Besides, compared with HGAN+SCFR that separately performs image synthesis and classification, JSRL with the proposed joint learning strategy significantly boosts the AUC result. This may be due to that joint learning can implicitly harness the complementary information of MRI and PET. Table 3 also suggests that most methods (trained on ADNI) work well in SCD conversion prediction on CLAS. This implies that the proposed transfer learning strategy (with label transfer and model transfer) effectively facilitates knowledge transfer from ADNI to CLAS.

4.2. Ablation study

Experimental Setup Our JSRL contains three key components, i.e., label transfer (LT), multi-modal representation sharing (MMRS) mechanism, and joint training (JT) of IS and RL subnetworks. To investigate their influence, we compare JSRL with its three variants for ablation analysis, including **JSRL-dLT** using a different label transfer strategy, **JSRL-oMMRS** without sharing features between IS and RL subnetworks, and **JSRL-oJT** that trains IS and RL subnetworks separately. Besides, to investigate the effectiveness of single-modal representations generated from our proposed model on SCD conversion prediction, we compare JSRL with its single-modal counterparts, namely **JSRL-M** and **JSRL-P**. The detailed network architectures of JSRL-oMMRS, JSRL-M and JSRL-P can be found in Figs. S1-S2 in the *Supplementary Materials*.

Table 4
Results achieved by our method with different strategies in SCD conversion prediction (i.e., pSCD vs. sSCD classification), with models trained on ADNI and tested on CLAS. The terms "dLT", "oMMRS", "oJT", "-M" and "-P" denote "different label transfer", "without multi-modal representation sharing", "without joint training", "with only MRI features" and "with only PET features", respectively.

Method	AUC	BAC	SPE	SEN	F1S
JSRL-dLT	0.711	0.683	0.615	**0.750**	0.581
JSRL-oMMRS	0.562	0.556	0.404	0.708	0.472
JSRL-oJT	0.676	0.601	0.596	0.625	0.500
JSRL-M	0.702	0.644	0.539	**0.750**	0.546
JSRL-P	0.593	0.543	0.462	0.462	0.448
JSRL (Ours)	**0.747**	**0.721**	**0.692**	**0.750**	**0.621**

1) JSRL-dLT: This method uses the same training data as JSRL (i.e., SCD-adjacent subjects in ADNI), but employs a different label transfer strategy by treating CN (with sCN and pCN) as the positive category and MCI as the negative category.
2) JSRL-oMMRS: In the JSRL-oMMRS method, those missing PET images are firstly generated from the image synthesis network. Then, to generate multi-modal feature inputs of the RL network, we perform two sequential convolution operations for MR and PET images individually, in which the strides of the first and the second Conv layers are 1 and 2, respectively. Finally, the separately generated multi-modal representation inputs are fed to the RL network, which is similar to the JSRL model. That is, our proposed multi-modal representation sharing mechanism is not employed in JSRL-oMMRS.
3) JSRL-oJT: The JSRL-oJT method trains the IS and the RL subnetworks separately. In JSRL, we first train the IS network, and then train IS and RL subnetworks jointly by optimizing the hybrid loss in Eq. (5). In JSRL-oJT, the hybrid loss in the RL subnetwork is replaced with a cross-entropy loss to facilitate the separate training of the IS and the RL subnetworks.
4) JSRL-M: In JSRL-M, only MRI representations (i.e., F_{M1} and F_{M2}) generated by the encoder part in the image synthesis network are fed into the RL network (with the same network architecture as JSRL).
5) JSRL-P: In JSRL-P, only PET representations (i.e., F_{P1}, F_{P2}) generated by the decoder part in the image synthesis network are fed into the RL network (with the same network architecture as JSRL).

Experimental Results The results of our JSRL and its five variants in SCD conversion prediction are summarized in Table 4. Table 4 suggests that our JSRL with three strategies (i.e., LT, MMRS, and JT) achieves the best performance, while other variants achieve the overall degraded performance. Some other interesting observations can also be found in Table 4.

1) When we use different auxiliary domains (i.e., MCI vs. CN) to aid SCD conversion prediction, JSRL-dLT also achieves good results (e.g., AUC=0.711). This suggests that the task of pSCD vs. sSCD classification is implicitly related to the task of MCI vs. CN classification, since AD is a neurodegenerative disease that develops over time. However, part of CN subjects would convert to MCI within a period (similar to pSCD), which may introduce ambiguous information for SCD conversion prediction. Our proposed label transfer strategy deliberately regards sCN and pCN subjects as positive and negative categories, respectively, aiming to augment training samples. Even though only 29 pCN subjects in the ADNI dataset made the performance difference between JSRL and

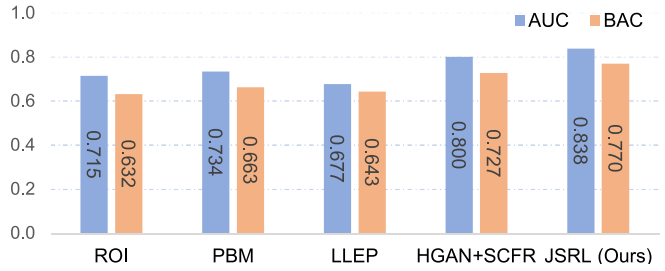

Fig. 4. Results of five multi-modal (with MRI and PET) methods in MCI conversion prediction (i.e., pMCI vs. sMCI classification) on ADNI, where models were trained on 939 subjects (with complete MRI and PET scans) and tested on 177 independent MCI subjects (with only MRI). AUC: Area under the ROC curve; BAC: Balanced Accuracy.

JSRL-dLT small, we can imagine the JSRL model will be more robust with the increased number of pCN subjects in the training set. 2) The JSRL-oMMRS method yields significantly degraded performance in comparison to JSRL. This confirms our assumption that the multi-modal representation shared between the IS and RL subnetworks conveys the potential relationship between different modalities, thereby helping to improve prediction performance. 3) Compared with JSRL-oJT, the joint training (via optimizing a hybrid loss) in JSRL helps improve the classification performance. The possible reason is that joint training encourages the model to generate prediction-oriented features, further constraining the IS model for efficient optimization. 4) The JSRL method is also superior to its single-modal counterparts (i.e., JSRL-M and JSRL-P), which validate the effectiveness of multi-modal representations generated by JSRL.

4.3. Application to MCI conversion prediction

We further evaluate our JSRL in MCI conversion prediction (i.e., pMCI vs. sMCI classification) on ADNI with incomplete multi-modal neuroimaging data. A total of 939 subjects (including 205 sCN, 325 sMCI, 164 pMCI and 245 AD) with complete MRI and PET scans are used as the training data for JSRL and the competing multi-modal methods. That is, sMCI and sCN are treated as the negative category, pMCI and AD are used as the positive category. And 177 subjects (including 81 sMCI and 96 pMCI) with only MRI are used as independent test data. All training data and test data are from the ADNI-1 and ADNI-2 datasets and the demographic and clinical information of the studied subjects are summarized in Table 2. The results of five methods in MCI conversion prediction are reported in Fig. 4. Note that the comparison between different methods in this Fig. 4 is fair, based on the same data and data partitions. Figure 4 suggests that our JSRL achieves the best AUC of 0.838 and ACC of 0.780, which is 4.11 and 5.65 percent higher than the second-best AUC and ACC achieved by HGAN+SCFR, respectively. This validates the effectiveness of the proposed JSRL in MCI conversion prediction. Results in Tables 3-4 and Fig. 4 imply our method can be used as a practical and general learning framework for early identification of AD-related disorders with modality-missing data. To further validate the generalization ability of the proposed method, we also report its results in MCI conversion prediction on ADNI using a 5-fold cross-validation strategy in Fig. S4 of the *Supplementary Materials*.

Besides, we summarize several state-of-the-art results reported in the literature for MCI conversion prediction using neuroimaging data of ADNI in Table 5. The results in Table 5 are not fully comparable, since these studies used different numbers of subjects and different data partitions. With the rough comparison, Table 5 shows that our JSRL method trained on a relatively large-scale training set obtains competitive performance in MCI conversion prediction on a relatively balanced test set. Even though the work in (Lian et al., 2020) shows a higher ACC, our method achieves the higher AUC result. The characteristic of yielding higher AUC values may be advantageous for more confident prediction, which is potentially very useful in practice (Huang and Ling, 2005).

4.4. Cross-database neuroimage synthesis

To investigate the effectiveness of the proposed image synthesis (IS) subnetwork in JSRL, we further perform cross-database image synthesis on three datasets, including 1) ADNI, 2) CLAS, and 3) AIBL (Ellis et al., 2009). Note that subjects from the ADNI have real FDG-PET images and those from the CLAS and AIBL datasets have no real/ground-truth FDG-PET images. Most subjects in the AIBL dataset have PET with different tracers (e.g., PIB/Flute-PET), and their appearance looks very different from FDG-PET. The IS model is trained on 863 subjects with complete structural MRI and FDG-PET from ADNI in the 1st step, and tested on 147 subjects with complete MRI and PET from ADNI (not involved in model training) for evaluation. Those 76 subjects with only MRI from CLAS and 235 subjects with complete MRI and Flute/PIB PET from AIBL are also used as independent test data for qualitative evaluation. The visual results can be found in Fig. S5 of the *Supplementary Materials*. To quantitatively evaluate our IS model, we compute the average structure similarity (SSIM) index and the average peak signal-to-noise ratio (PSNR) between the synthetic PET images and the corresponding real ones of 147 subjects from the ADNI dataset. The mean SSIM and PSNR values of synthetic PET in ADNI are 0.70 and 27.35 dB, respectively. The quantitative results and the visually reasonable results in Fig. S5 suggest that the proposed JSRL can generate ADNI-like FDG-PET across different datasets. It implies that our method could be potentially used for multi-site data harmonization in multi-center problems with different scanners or scanning protocols.

5. Discussion

In this section, we compare our method with relevant studies for early diagnosis of AD/MCI, and present the limitations of this work as well as future research directions.

5.1. Comparison with previous studies

In this work, we developed a multi-modal neuroimaging-based method to predict the 7-year's conversion from SCD to MCI of 76 SCD subjects with an unprecedentedly long follow-up time from the CLAS study. This task is very challenging due to the lack of PET data and the indistinguishable pathological changes between sSCD subjects and pSCD subjects who are at the very early stage of AD with normal cognition. According to the quantitative results, our proposed JSRL method is effective for the conversion prediction of SCD.

Several strategies may contribute to the effectiveness of early detection of SCD. *First*, our JSRL leverages the complementary information of multi-modal neuroimages for improving the prediction performances. This is among the first multi-modal neuroimaging-based studies for SCD conversion prediction. *Second*, the JSRL integrates image synthesis (IS) and representation learning (RL) into a unified framework. Meantime, a multi-modal representation sharing mechanism is presented in JSRL. That is, the inputs of the RL network are the multi-modal representations generated by the IS network rather than the individually extracted MRI and real/synthetic PET features as in previous work (Pan et al., 2019; Liu et al., 2014b). Due to the mapping relationship from MRI to PET learned by the image generator, the underlying relationship of multiple modalities from the same subject can be better preserved

Table 5
Comparison between the proposed JSRL method and the state-of-the-art neuroimaging studies in the task of MCI conversion prediction (i.e., pMCI vs. sMCI classification) on the ADNI database. N/P (training): Negative/Positive samples in the training set; N/P (test): Negative/Positive samples in the test set.

Method	N/P (training)	N/P (test)	Modality	AUC	ACC
Shi et al. (2017)	102/89	6/5	MRI+PET	0.801	0.789
Suk et al. (2014)	115/68	13/8	MRI+PET	0.747	0.759
Young et al. (2013)	36/35	42/30	MRI+PET+ APOE+CSF	0.763	0.722
Lian et al. (2020)	455/197	239/38	MRI	0.781	**0.809**
Pan et al. (2020)	477/308	256/89	MRI+PET	0.825	0.778
JSRL (Ours)	530/409	66/93	MRI+PET	**0.838**	0.780

in the RL network by leveraging the shared multi-modal representations from the image generator as inputs. *Besides*, a multi-modal feature fusion module is designed in the RL network for early fusion of multi-modal features, which is different from previous methods that extract MRI and PET features separately and fuse them via a late fusion strategy (Campos et al., 2015; Pan et al., 2020). *In addition*, to address the limited data problem, a transfer learning strategy is proposed by leveraging the JSRL learned from a large-scale ADNI database to a small-scale CLAS database, making the JSRL method feasible.

5.2. Limitations and future work

Several issues need to be considered in the future. *First*, we used a basic GAN architecture for image synthesis, and the improved performance of our method indicated that IS could produce informative multi-modal features for prediction. It is desired to design more advanced networks (e.g., conditional GAN (Mirza and Osindero, 2014) and bidirectional GAN (Zhu et al., 2017)) for IS to further improve the discriminative capability of multi-modal representations. *Besides*, we used only neuroimaging data for prediction, without considering the demographic information (e.g., age, gender, and education) that may also be related to brain status (Yue et al., 2021). In future work, we will employ demographic information to constrain the learned models. *Furthermore*, we designed a transfer learning strategy by leveraging knowledge learned from ADNI to CLAS, and did not explicitly consider the distribution gap between these two databases. Since the ADNI contains primarily Caucasian brains and the CLAS consists of Asian brains, it is interesting to develop advanced data harmonization methods (Kamnitsas et al., 2017) to reduce the distribution shift between two databases, which will also be our future work.

6. Conclusion

In this paper, we presented a joint neuroimage synthesis and representation learning (JSRL) framework for SCD conversion prediction based on incomplete and limited multi-modal neuroimages. The JSRL consisted of an image synthesis subnetwork for imputing missing images and a representation learning subnetwork for prediction. The two subnetworks were incorporated into a joint learning framework by sharing multi-modal features. A transfer learning strategy was also designed to facilitate the knowledge transfer between training and test data. Experiments were conducted on three databases, including the ADNI with MRI and FDG-PET, the CLAS with only MRI, and the AIBL with MRI and Flute/PIB PET. The experimental results suggested the efficacy of JSRL in the tasks of SCD and MCI conversion prediction and cross-database neuroimage synthesis, compared with several state-of-the-art methods.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yunbi Liu: Methodology, Software, Writing – original draft. **Ling Yue:** Data curation, Writing – review & editing. **Shifu Xiao:** Data curation. **Wei Yang:** Writing – review & editing. **Dinggang Shen:** Writing – review & editing. **Mingxia Liu:** Conceptualization, Validation, Writing – review & editing, Supervision.

Acknowledgment

We thank Dr. Yongsheng Pan for his valuable suggestions and help in the experiments. This work was finished when Y. Liu was visiting the University of North Carolina at Chapel Hill. M. Liu was supported in part by NIH grant (No. AG041721). Y. Liu and W. Yang were supported in part by National Natural Science Foundation of China (No. 81771916) and Guangdong Provincial Key Laboratory of Medical Image Processing (No. 2014B030301042). L. Yue and S. Xiao were supported in part by China Ministry of Science and Technology (No. 2009BAI77B03), National Natural Science Foundation of China (No. 82001123) and Shanghai Mental Health Center (Nos. CRC2017ZD02, 2018-FX-05, and 2020zd01). Part of the data used in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The investigators within the ADNI contributed to the design and implementation of ADNI and provided data but did not participate in analysis or writing of this article. A complete listing of ADNI investigators can be found online (https://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2021.102266.

References

Abdulrab, K., Heun, R., 2008. Subjective memory impairment. A review of its definitions indicates the need for a comprehensive set of standardised and validated criteria. Eur. Psychiatry 23 (5), 321–330.
Amariglio, R.E., Becker, J.A., Carmasin, J., Wadsworth, L.P., Lorius, N., Sullivan, C., Maye, J.E., Gidicsin, C., Pepin, L.C., Sperling, R.A., et al., 2012. Subjective cognitive complaints and amyloid burden in cognitively normal older individuals. Neuropsychologia 50 (12), 2880–2886.
Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26 (3), 839–851.
Association, A., et al., 2018. 2018 Alzheimer's disease facts and figures. Alzheimer's Dementia 14 (3), 367–429.
Barnes, L.L., Schneider, J.A., Boyle, P.A., Bienias, J.L., Bennett, D.A., 2006. Memory complaints are related to Alzheimer disease pathology in older persons. Neurology 67 (9), 1581–1585.

Buckley, R.F., Maruff, P., Ames, D., Bourgeat, P., Martins, R.N., Masters, C.L., Rainey–Smith, S., Lautenschlager, N., Rowe, C.C., Savage, G., et al., 2016. Subjective memory decline predicts greater rates of clinical progression in preclinical Alzheimer's disease. Alzheimer's Dementia 12 (7), 796–804.

Buckley, R.F., Villemagne, V.L., Masters, C.L., Ellis, K.A., Rowe, C.C., Johnson, K., Sperling, R., Amariglio, R., 2016. A conceptualization of the utility of subjective cognitive decline in clinical trials of preclinical Alzheimer's disease. J. Mol. Neurosci. 60 (3), 354–361.

Campos, S., Pizarro, L., Valle, C., Gray, K.R., Rueckert, D., Allende, H., 2015. Evaluating imputation techniques for missing data in ADNI: a patient classification study. In: Iberoamerican Congress on Pattern Recognition. Springer, pp. 3–10.

Caselli, R.J., Chen, K., Locke, D.E.C., Lee, W., Roontiva, A., Bandy, D., Fleisher, A.S., Reiman, E.M., 2014. Subjective cognitive decline: self and informant comparisons. Alzheimer's Dementia 10 (1), 93–98.

Cheng, B., Liu, M., Shen, D., Li, Z., Zhang, D., 2017. Multi-domain transfer learning for early diagnosis of Alzheimer's disease. Neuroinformatics 15 (2), 115–132.

Cheng, B., Liu, M., Suk, H.-I., Shen, D., Zhang, D., 2015. Multimodal manifold-regularized transfer learning for MCI conversion prediction. Brain Imaging Behav. 9 (4), 913–926.

Cheng, B., Liu, M., Zhang, D., Munsell, B.C., Shen, D., 2015. Domain transfer learning for MCI conversion prediction. IEEE Trans. Biomed. Eng. 62 (7), 1805–1817.

Da, X., Toledo, J.B., Zee, J., Wolk, D.A., Xie, S.X., Ou, Y., Shacklett, A., Parmpi, P., Shaw, L., Trojanowski, J.Q., et al., 2014. Integration and relative value of biomarkers for prediction of MCI to AD progression: spatial patterns of brain atrophy, cognitive scores, APOE genotype and CSF biomarkers. NeuroImage Clinical 4, 164–173.

Ellis, K.A., Bush, A.I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N.T., Lenzo, N., Martins, R.N., Maruff, P., et al., 2009. The Australian imaging, biomarkers and lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int. Psychogeriatrics 21 (4), 672–687.

Filipovych, R., Davatzikos, C., 2011. Semi-supervised pattern classification of medical images: application to mild cognitive impairment (MCI). NeuroImage 55 (3), 1109–1119.

Fischl, B., 2012. Freesurfer. NeuroImage 62 (2), 774–781.

van der Flier, W.M., Van Buchem, M.A., Weverling-Rijnsburger, A.W.E., Mutsaers, E.R., Bollen, E.L., Admiraal-Behloul, F., Westendorp, R.G.J., Middelkoop, H.A.M., 2004. Memory complaints in patients with normal cognition are associated with smaller hippocampal volumes. J. Neurol. 251 (6), 671–675.

Gray, K.R., Aljabar, P., Heckemann, R.A., Hammers, A., Rueckert, D., 2013. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. NeuroImage 65, 167–175.

Hinrichs, C., Singh, V., Xu, G., Johnson, S.C., 2011. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. NeuroImage 55 (2), 574–589.

Hor, S., Moradi, M., 2016. Learning in data-limited multimodal scenarios: scandent decision forests and tree-based features. Med. Image Anal. 34, 30–41.

Huang, J., Ling, C.X., 2005. Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowl. Data Eng. 17 (3), 299–310.

Jack Jr., C.R., Albert, M.S., Knopman, D.S., McKhann, G.M., Sperling, R.A., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. Introduction to the recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dementia 7 (3), 257–262.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Jennifer, L.W., Ward, C., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27 (4), 685–691.

Jack Jr., C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., et al., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol. 12 (2), 207–216.

Jack Jr, C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Trojanowski, J.Q., 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 9 (1), 119–128.

Jessen, F., Amariglio, R.E., Van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K.A., Van Der Flier, W.M., et al., 2014. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimer's Dementia 10 (6), 844–852.

Jessen, F., Wolfsgruber, S., Wiese, B., Bickel, H., Mösch, E., Kaduszkiewicz, H., Pentzek, M., Riedel-Heller, S.G., Luck, T., Fuchs, A., et al., 2014. AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. Alzheimer's Dementia 10 (1), 76–83.

Kamnitsas, K., Baumgartner, C., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., Nori, A., Criminisi, A., Rueckert, D., et al., 2017. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 597–609.

Kawachi, T., Ishii, K., Sakamoto, S., Sasaki, M., Mori, T., Yamashita, F., Matsuda, H., Mori, E., 2006. Comparison of the diagnostic performance of FDG-PET and VBM-MRI in very mild Alzheimer's disease. Eur. J. Nucl. Med. Mol. Imaging 33 (7), 801–809.

Khan, N.M., Abraham, N., Hon, M., 2019. Transfer learning with intelligent training data selection for prediction of Alzheimer's disease. IEEE Access 7, 72726–72735.

Kohannim, O., Hua, X., Hibar, D.P., Lee, S., Chou, Y.-Y., Toga, A.W., Jack Jr, C.R., Weiner, M.W., Thompson, P.M., 2010. Boosting power for clinical trials using classifiers based on multiple biomarkers. Neurobiol. Aging 31 (8), 1429–1442.

Kryscio, R.J., Abner, E.L., Cooper, G.E., Fardo, D.W., Jicha, G.A., Nelson, P.T., Smith, C.D., Van Eldik, L.J., Wan, L., Schmitt, F.A., 2014. Self-reported memory complaints: implications from a longitudinal cohort with autopsies. Neurology 83 (15), 1359–1365.

Lassila, T., Faria, H.M., Sarrami-Foroushani, A., Meneghello, F., Venneri, A., Frangi, A.F., 2018. Multi-modal synthesis of ASL-MRI features with KPLS regression on heterogeneous data. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 473–481.

Li, R., Zhang, W., Suk, H.-I., Wang, L., Li, J., Shen, D., Ji, S., 2014. Deep learning based imaging data completion for improved brain disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 305–312.

Lian, C., Liu, M., Zhang, J., Shen, D., 2020. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Trans. Pattern Anal. Mach. Intell. 42 (4), 880–893.

Liu, M., Zhang, D., Chen, S., Xue, H., 2015. Joint binary classifier learning for ECOC-based multi-class classification. IEEE Trans. Pattern Anal. Mach. Intell. 38 (11), 2335–2341.

Liu, M., Zhang, D., Shen, D., 2014. Hierarchical fusion of features and classifier decisions for Alzheimer's disease diagnosis. Hum. Brain Mapp. 35 (4), 1305–1319.

Liu, M., Zhang, J., Adeli, E., Shen, D., 2018. Landmark-based deep multi-instance learning for brain disease diagnosis. Med. Image Anal. 43, 157–168.

Liu, M., Zhang, J., Yap, P.-T., Shen, D., 2017. View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data. Med. Image Anal. 36, 123–134.

Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Feng, D., Fulham, M.J., 2014. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Trans. Biomed. Eng. 62 (4), 1132–1140.

Liu, Y., Pan, Y., Yang, W., Ning, Z., Yue, L., Liu, M., Shen, D., 2020. Joint neuroimage synthesis and representation learning for conversion prediction of subjective cognitive decline. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer International Publishing, Cham, pp. 583–592.

Mirza, M., Osindero, S., Conditional generative adversarial nets, 2014. arXiv preprint arXiv:1411.1784

Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D., 2019. Disease-image specific generative adversarial network for brain disease diagnosis with incomplete multi-modal neuroimages. In: MICCAI. Springer, pp. 137–145.

Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D., 2020. Spatially-constrained Fisher representation for brain disease identification with incomplete multi-modal neuroimages. IEEE Trans. Med Imaging 39 (9), 2965–2975.

Pan, Y., Liu, M., Xia, Y., Shen, D., 2021. Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data. IEEE Trans. Pattern Anal. Mach. Intell..

Perrin, R.J., Fagan, A.M., Holtzman, D.M., 2009. Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. Nature 461 (7266), 916–922.

Rusinek, H., De Leon, M.J., George, A.E., Stylopoulos, L.A., Chandra, R., Smith, G., Rand, T., Mourino, M., Kowalski, H., 1991. Alzheimer disease: measuring loss of cerebral gray matter with MR imaging. Radiology 178 (1), 109–114.

Scheef, L., Spottke, A., Daerr, M., Joe, A., Striepens, N., Kölsch, H., Popp, J., Daamen, M., Gorris, D., Heneka, M.T., et al., 2012. Glucose metabolism, gray matter structure, and memory decline in subjective memory impairment. Neurology 79 (13), 1332–1339.

Sharma, A., Hamarneh, G., 2019. Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. IEEE Trans. Med. Imaging 39 (4), 1170–1183.

Shi, J., Zheng, X., Li, Y., Zhang, Q., Ying, S., 2017. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE J. Biomed. Health Inform. 22 (1), 173–183.

Sikka, A., Peri, S.V., Bathula, D.R., 2018. MRI to FDG-PET: cross-modal synthesis using 3D U-Net for multi-modal Alzheimer's classification. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 80–89.

Stewart, R., Dufouil, C., Godin, O., Ritchie, K., Maillard, P., Delcroix, N., Crivello, F., Mazoyer, B., Tzourio, C., 2008. Neuroimaging correlates of subjective memory deficits in a community population. Neurology 70 (18), 1601–1607.

Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., Suliman, H., Wagner, M., Schild, H.H., Jessen, F., 2010. Volume loss of the medial temporal lobe structures in subjective memory impairment. Dement Geriatr Cogn. Disord. 29 (1), 75–81.

Suk, H.-I., Lee, S.-W., Shen, D., 2014. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage 101, 569–582.

Thung, K.-H., Yap, P.-T., Shen, D., 2017. Multi-stage diagnosis of Alzheimer's disease with incomplete multimodal data via multi-task deep learning. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 160–168.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15 (1), 273–289.

Van Tulder, G., de Bruijne, M., 2015. Why does synthesized data improve multi-sequence classification? In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 531–538.

Xiao, S., Lewis, M., Mellor, D., McCabe, M., Byrne, L., Wang, T., Wang, J., Zhu, M., Cheng, Y., Yang, C., et al., 2016. The China longitudinal ageing study: overview of the demographic, psychosocial and cognitive data of the shanghai sample. J. Mental Health 25 (2), 131–136.
Yi, X., Walia, E., Babyn, P., 2019. Generative adversarial network in medical imaging: a review. Med. Image Anal. 101552.
Young, J., Modat, M., Cardoso, M.J., Mendelson, A., Cash, D., Ourselin, S., 2013. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. NeuroImage Clinical 2, 735–745.
Yue, L., Hu, D., Zhang, H., Wen, J., Wu, Y., Li, W., Sun, L., Li, X., Wang, J., Li, G., et al., 2021. Prediction of 7-year's conversion from subjective cognitive decline to mild cognitive impairment. Hum. Brain Mapp. 42 (1), 192–203.
Yue, L., Wang, T., Wang, J., Li, G., Wang, J., Li, X., Li, W., Hu, M., Xiao, S., 2018. Asymmetry of hippocampus and amygdala defect in subjective cognitive decline among the community dwelling chinese. Front. Psychiatry 9, 226.
Zhang, J., Liu, M., An, L., Gao, Y., Shen, D., 2017. Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images. IEEE J. Biomed. Health Inform. 21 (6), 1607–1616.
Zhou, T., Thung, K.-H., Liu, M., Shen, D., 2018. Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model. IEEE Trans. Biomed. Eng. 66 (1), 165–175.
Zhou, T., Thung, K.-H., Liu, M., Shi, F., Zhang, C., Shen, D., 2020. Multi-modal latent space inducing ensemble SVM classifier for early dementia diagnosis with neuroimaging data. Med. Image Anal. 60, 101630.
Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232.
Zu, C., Jie, B., Liu, M., Chen, S., Shen, D., Zhang, D., 2016. Label-aligned multi-task feature learning for multimodal classification of Alzheimer's disease and mild cognitive impairment. Brain Imaging Behav. 10 (4), 1148–1159.