



# BMNet: A New Region-Based Metric Learning Method for Early Alzheimer's Disease Identification With FDG-PET Images

Wenju Cui<sup>1,2†</sup>, Caiying Yan<sup>3†</sup>, Zhuangzhi Yan<sup>1</sup>, Yunsong Peng<sup>2,4</sup>, Yilin Leng<sup>1,2</sup>, Chenlu Liu<sup>3</sup>, Shuangqing Chen<sup>3\*</sup>, Xi Jiang<sup>5\*</sup>, Jian Zheng<sup>2</sup> and Xiaodong Yang<sup>2</sup>

<sup>1</sup> Institute of Biomedical Engineering, School of Communication and Information Engineering, Shanghai University, Shanghai, China, <sup>2</sup> Medical Imaging Department, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China, <sup>3</sup> Department of Radiology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China, <sup>4</sup> School of Biomedical Engineering, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China, <sup>5</sup> School of Life Sciences and Technology, The University of Electronic Science and Technology of China, Chengdu, China

## OPEN ACCESS

### Edited by:

Xiang Li,  
Massachusetts General Hospital  
and Harvard Medical School,  
United States

### Reviewed by:

Lu Zhang,  
University of Texas at Arlington,  
United States  
Mingxia Liu,  
University of North Carolina at Chapel  
Hill, United States

### \*Correspondence:

Shuangqing Chen  
sznaonao@163.com  
Xi Jiang  
xijiang@uestc.edu.cn

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 08 December 2021

**Accepted:** 11 January 2022

**Published:** 24 February 2022

### Citation:

Cui W, Yan C, Yan Z, Peng Y,  
Leng Y, Liu C, Chen S, Jiang X,  
Zheng J and Yang X (2022) BMNet:  
A New Region-Based Metric Learning  
Method for Early Alzheimer's Disease  
Identification With FDG-PET Images.  
*Front. Neurosci.* 16:831533.  
doi: 10.3389/fnins.2022.831533

18F-fluorodeoxyglucose (FDG)-positron emission tomography (PET) reveals altered brain metabolism in individuals with mild cognitive impairment (MCI) and Alzheimer's disease (AD). Some biomarkers derived from FDG-PET by computer-aided-diagnosis (CAD) technologies have been proved that they can accurately diagnosis normal control (NC), MCI, and AD. However, existing FDG-PET-based researches are still insufficient for the identification of early MCI (EMCI) and late MCI (LMCI). Compared with methods based other modalities, current methods with FDG-PET are also inadequate in using the inter-region-based features for the diagnosis of early AD. Moreover, considering the variability in different individuals, some hard samples which are very similar with both two classes limit the classification performance. To tackle these problems, in this paper, we propose a novel bilinear pooling and metric learning network (BMNet), which can extract the inter-region representation features and distinguish hard samples by constructing the embedding space. To validate the proposed method, we collect 898 FDG-PET images from Alzheimer's disease neuroimaging initiative (ADNI) including 263 normal control (NC) patients, 290 EMCI patients, 147 LMCI patients, and 198 AD patients. Following the common preprocessing steps, 90 features are extracted from each FDG-PET image according to the automatic anatomical landmark (AAL) template and then sent into the proposed network. Extensive fivefold cross-validation experiments are performed for multiple two-class classifications. Experiments show that most metrics are improved after adding the bilinear pooling module and metric losses to the Baseline model respectively. Specifically, in the classification task between EMCI and LMCI, the specificity improves 6.38% after adding the triple metric loss, and the negative predictive value (NPV) improves 3.45% after using the bilinear pooling module. In addition, the accuracy of classification between EMCI and LMCI achieves 79.64% using imbalanced FDG-PET images, which illustrates that the proposed method yields a state-of-the-art result of the classification accuracy between EMCI and LMCI based on PET images.

**Keywords:** early Alzheimer's disease, mild cognitive impairment, FDG-PET images, bilinear pooling, inter-region representation, metric learning, embedding space

## INTRODUCTION

Alzheimer's disease (AD), a brain degenerative disorder, is harming the health of thousands of old people now, and its rate of prevalence is expected to increase rapidly in the coming decades (Wang et al., 2013; Alzheimer's Association, 2018, 2019). Mild cognitive impairment (MCI) is considered to be a preclinical precursor of AD, but it is difficult to predict whether it will convert to AD or not (Gauthier et al., 2006; Dubois et al., 2016; Hampel and Lista, 2016). Considering the unpredictable process of MCI, it is crucial to develop relevant methods for diagnosing the early MCI and AD.

<sup>18</sup>F-fluorodeoxyglucose (FDG)-positron emission tomography (PET) can reveal altered brain metabolism in individuals with MCI and AD (Sørensen et al., 2019; Zhou et al., 2019; Wang et al., 2020). Various recent studies have proved that biomarkers derived from FDG-PET by computer-aided-diagnosis (CAD) technologies of machine learning and deep learning can accurately diagnose NC, MCI, and AD (Pagani et al., 2017; Choi et al., 2018; Blazhenets et al., 2019). Liu et al. (2018) proposed a new classification framework for AD diagnosis with 3D PET images. They decomposed 3D images into 2D slices to learn the intra-slice and inter-slice features and achieved a promising classification performance of AUC of 83.9% for MCI vs. NC classification. Zhou et al. (2021) developed a new deep belief network model for AD diagnosis based on sparse-response theory, which identified a better classification result than that of other models. To solve the multimodal data missing problem, Dong et al. (2021) proposed a high-order Laplacian regularized low-rank representation method for the classification tasks of NC, MCI, and AD. Pan et al. (2021) developed a disease-image-specific deep learning (DSDL) framework which can achieve neuroimage synthesis and disease diagnosis simultaneously using incomplete multi-modality neuroimages.

Many studies have achieved good performance on the classification of NC, MCI, and AD based on FDG-PET images. However, when it comes to the more refined task like classification of early MCI (EMCI) and late MCI (LMCI), the studies with FDG-PET images are still insufficient. Hao et al. (2020) proposed a novel multi-modal neuroimaging feature selection method with consistent metric constraint (MFCC) and obtained an accuracy (ACC) of 73.87% for the classification between EMCI and LMCI based on MRI and FDG-PET but only 64.69% when just using FDG-PET. Singh et al. (2017) proposed a multilayer neural network involving probabilistic principal component analysis for binary classification and only achieved an F1 score of 68.44%. Nozadi et al. (2018) used learned features from semantically labeled PET images to perform group classification and got an ACC of 72.5%. Forouzannezhad et al. (2018, 2020) applied a novel deep neural network and a random forest model respectively, and both models got a moderate ACC. Fang et al. (2020) introduced a supervised Gaussian discriminative component analysis (GDCA) algorithm for the effective classification of early Alzheimer's disease with MRI and PET. Yang and Liu (2020) applied the Convolutional Architecture for Fast Feature Embedding (CAFFE) as the framework of the deep learning platform for early Alzheimer's disease diagnosis.

By comparison, based on fMRI and DTI images, Lei et al. (2020) got an ACC of 78.05% for the classification between EMCI and LMCI *via* proposing a new joint multi-task learning method by combining low-rank self-calibrated functional and structural brain networks. Song et al. (2021) constructed a new graph convolution network (GCN) and got an ACC of 79.26% based on fMRI and 82.92% based on DTI for the same classification task. With MRI images, Lian et al. (2018) developed a hierarchical fully convolutional network that can achieve an ACC of 81% for the classification between progressive MCI (pMCI) and stable MCI (sMCI).

To sum up, the refined classification performance for early AD based on FDG-PET images still has some room for improvement. One of the reasons might be that existing classification methods based on FDG-PET have not fully explored the inter-region representation among different brain regions. For example, based on fMRI, there are many methods like Pearson's correlation and sparse representation for functional brain network (FBN) estimation (Huang J. et al., 2020). However, several studies have proved that brain metabolism connectivity has value in the diagnosis of early AD (Huang et al., 2010; Sanabria-Diaz et al., 2013; Titov et al., 2017), but few PET-based studies are using the inter-region features to improve classification performance. In addition, another reason might be that the number of PET images is generally much more than that of fMRI images in most researches. The bigger dataset might increase the variety of individuals and the probability of special samples which are hard to distinguish, thus causing complexity of the problem for classification tasks.

Considering these two limitations, we propose a novel bilinear pooling and metric learning network (BMNet) for early Alzheimer's disease identification with FDG-PET images, especially for the classification task between EMCI and LMCI. Our main contributions are as follows: (1) We propose a shallow convolutional neural network model to achieve the classification; (2) We introduce a bilinear pooling module into the model for exploring the inter-region representation features in the whole brain; (3) We introduce the deep metric learning to help model learn the hard samples in the embedding feature space; (4) We conduct our method on the dataset collected from the publicly released ADNI database and obtain a state-of-the-art result of the classification between EMCI and LMCI based on PET images.

The rest of this paper is organized as follows. In section "Materials and Methods," we present details of the materials and the proposed methods. Section "Results" presents the results of the experiments on the public ADNI database. Finally, we provide the discussions and conclusion of this paper in section "Discussion and Conclusion."

## MATERIALS AND METHODS

### Image Acquisition and Preprocessing

In this work, we use the data in the publicly released Alzheimer's Disease Neuroimaging Initiative (ADNI) database (Jack et al., 2008). We collect a cohort of subjects with FDG-PET images from the ADNI databases. The ADNI cohort includes FDG-PET

images from 898 subjects, including 263 NC, 290 EMCI, 147 LMCI, and 198 AD participants. **Table 1** lists the demographic characteristics of subjects.

We choose FDG-PET images which are in a state of rest with 30–35 min with  $185 \pm 18.5$  MBq FDG, and details of acquisition can be obtained from the study protocols in the ADNI database. Firstly, we normalize the images based on the template of the Montreal Neurological Institute (MNI). Then, we perform the smoothing with a Gaussian filter of 8 mm fullwidth at half-maximum (FWHM) (Wang et al., 2020). Finally, to verify the effectiveness of the proposed method, we do the main experiments using two different brain atlas. Based on the automated anatomical labeling (AAL) (Ashburner and Friston, 2000) atlas, we extract features of 90 regions of interest (ROIs) from FDG-PET images with intensity normalized averagely. Similarly, based on the Schaefer et al. (2018) atlas (Schaefer et al., 2018), we extract features of 400 regions. We perform all preprocessing steps by Statistical Parametric Mapping software (SPM12) (Tzourio-Mazoyer et al., 2002) and Matlab (2020).

## Methods

### Overview of the Proposed Network

**Figure 1** illustrates the method framework of this study. The left box is the preprocessing step of FDG-PET images, in which the left image is the raw PET image of the brain, and the right one is the AAL template. Then 90 features extracted based on the AAL template are input into the subsequent model. The model consists of two convolution layers, a bilinear pooling layer, and two fully connected layers.

After extracting the first-order features through two convolution layers, the bilinear pooling module is used to further extract the inter-region-based features. Finally, the metric learning loss is added to the classification loss to strengthen the ability to learn hard samples of the proposed model.

### Baseline Model

We construct a shallow neural network as the Baseline model, including two convolution blocks and three fully connected layers. Each convolution block includes a convolution layer, a batch normalization layer, and a Rectified Linear Unit (ReLU) activation layer.

Given a set of nodes (regions)  $R = \{r_1, r_2, r_3\}$ , and the features of each region is denoted as  $X_i$ . Each convolution block is defined as:

$$Y_i = \sigma[BN(f(X_i))] \quad (1)$$

**TABLE 1** | Demographic characteristics of the subjects in the ADNI database.

Subjects	NC	EMCI	LMCI	AD
Number	263	290	147	198
Gender (M/F)	130/133	160/130	80/67	119/79
Age	75.49 $\pm$ 6.47	71.40 $\pm$ 7.33	72.16 $\pm$ 7.55	75.05 $\pm$ 7.60
MMSE	29.06 $\pm$ 1.13	28.32 $\pm$ 1.57	27.62 $\pm$ 1.84	23.20 $\pm$ 2.17

The values are presented as mean  $\pm$  standard deviation. MMSE, Mini-Mental State Examination.

Where the  $f$  represents the convolution process,  $BN$  represents the batch normalization process,  $\sigma$  represents the activation process.

### Generating Inter-Region Representation via Bilinear Pooling Module

In this section, we propose to use a bilinear pooling module to further generate second-order features which may represent inter-region features among whole brain regions. Bilinear pooling is an effective feature fusion method, which has been widely used in various computer vision and machine learning tasks (Lin et al., 2015; Gao et al., 2020). Bilinear pooling captures the high-order statistical information of features by matrix operations and then generates an expressive global representation (Kim et al., 2016; Li et al., 2017; Gao et al., 2020). In the research of DTI and fMRI, this method is also used to extract connectivity-based features between brain regions (Huang F. et al., 2020). In theory, by using these features, the inter-region representation among the whole brain regions in FDG-PET images could be exploited to some extent, as the functional brain network of fMRI.

In this work, we introduce a new factorized bilinear pooling method (Gao et al., 2020) to capture inter-region features by fusing homogeneous features where the input features are from the same source. This new bilinear pooling method simplifies the complexity of calculation, reduces heavy computational redundancy issues. Based on factorized bilinear coding, it is proved that bilinear features are rank-one matrices whose rank is one. The bilinear features could be extracted by factorizing dictionary atoms into low-rank matrices and Hadamard product, instead of massive matrix operations, reducing the dimension of matrices and computational burden.

The main operations of bilinear features are as follows (Kim et al., 2016; Gao et al., 2020):

$$B = Y^T W Y = Y^T U V^T Y = P^T (U^T Y \circ V^T Y) \quad (2)$$

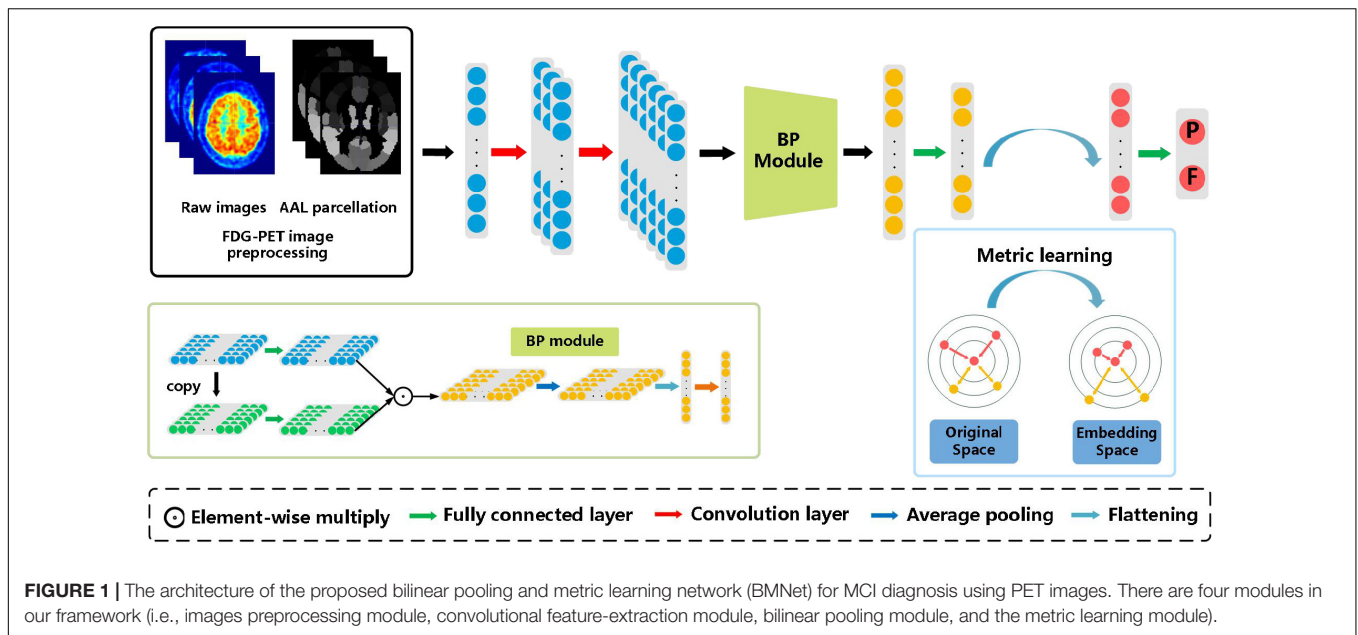
where  $B$  represents the bilinear features, and  $Y$  represents the input feature,  $U^T$ ,  $V^T$  and  $P^T$  are learnable parameters of the dictionary,  $\circ$  represents Hadamard product.

The low-rank matrix  $U$  and  $V$  are used to approximate  $W$ , and the operation is simplified. Matrix  $P$  is used to control the length of the output. In the network, three fully connected layers are used to learn  $U^T$ ,  $V^T$  and  $P^T$ . Then, we use an average pooling layer to diminish the feature dimension and obtain the global information. Finally, the feature map is flattened to one-dimensional and a fully connected layer is used to diminish the feature dimension to facilitate subsequent learning processes.

We use this bilinear pooling method to capture inter-region representation with FDG-PET images. The homogeneous features achieve interaction of the whole brain by the bilinear pooling module, which needs complex and expensive computation before.

### Distinguishing Hard Samples in Embedding Space by Metric Learning

In this section, we introduce the deep metric learning strategy into the classification of different stages of AD. Metric learning is widely utilized with deep neural networks in classification



tasks, especially in problems affected by large intra-class sample changes (Liu et al., 2017; Sundgaard et al., 2021). Deep metric learning loss maps features to the embedded space, which is conducive to learning difficult samples and can effectively deal with the imbalance of data (Sundgaard et al., 2021). Inspired by these, we argue that deep metric learning might be suitable for our classification task. Thus, in this paper, we employ deep metric learning for the diagnosis of AD to help distinguish hard samples in the embedding space.

In deep neural networks, the loss function is a manifestation of metric learning, and there are a variety of different metric learning loss functions. In this paper, we employ two deep metric learning loss functions for automatic diagnosis of early AD, including contrastive loss and triplet loss, which are widely used in recent studies (Cheng et al., 2016; He et al., 2021; Sundgaard et al., 2021). Contrastive loss employs a pair of positive and negative samples for each training iteration. The contrastive loss function is measured by the Euclidian distance between two vectors in embedding space. The contrastive loss function is given as (Hadsell et al., 2006):

$$L_c(b_{1,i}, b_{2,i}) = \sum_{i=1}^N [y_i d_{1,2}^2 + (1 - y_i) \{ \max(0, m - d_{1,2}) \}^2] \quad (3)$$

$$d_{1,2} = \|f_{1,i} - f_{2,i}\|_2^2 \quad (4)$$

where  $y_i = 0$  for two positive vectors and  $y_i = 1$  for negative pairs,  $b_{1,i}, b_{2,i}$  is the training input from two classes,  $f_{1,i}, f_{2,i}$  represents the embedding vector of each training input generated by the network,  $N$  is the number of input samples, and  $m$  is the margin, usually set to 1.0.

When the input is a positive sample pair,  $d_{1,2}$  decreases gradually, and the same kind of samples will continue to form clusters in the feature space. On the contrary, when the network inputs a negative sample pair,  $d_{1,2}$  will gradually rise until it reaches the set  $m$ . By minimizing the loss functions, the distance

between positive sample pairs can be gradually reduced and the distance between negative sample pairs can be gradually increased, to meet the needs of the classification task.

Triplet loss is a widely used measure of metric learning loss, which is the basis of a large number of metric learning methods. Unlike contrastive loss, triplet loss requires three input samples including two positive samples and a negative sample. The three samples are named as fixed sample (anchor)  $b^a$ , positive sample (positive)  $b^p$  and negative sample (negative)  $b^n$  respectively.  $b^a$  and  $b^p$  form positive sample pairs, and  $b^a$  and  $b^n$  form negative sample pairs.

This triplet loss function simultaneously penalizes a short distance  $d_{a,n}$  between an anchor and a negative sample and a long distance  $d_{a,p}$  between an anchor and a positive sample, and is defined as (Schroff et al., 2015):

$$L_{triplet}(b_i^a, b_i^p, b_i^n) = \sum_{i=1}^N \max(0, m + d_{a,p} - d_{a,n}) \quad (5)$$

where  $b_i^a, b_i^p, b_i^n$  is the input from two training groups,  $N$  represents the number of samples, and  $m$  is the margin, usually set to 1.0.

$$d_{a,p} = \|f_i^a - f_i^p\|_2^2 \quad (6)$$

$$d_{a,n} = \|f_i^a - f_i^n\|_2^2 \quad (7)$$

$f_i^a, f_i^p, f_i^n$  represents the vector of training input in embedding space.

As shown in **Figure 1**, the triple loss can shorten the distance between positive sample pairs, while pushing away the distance between negative sample pairs. Finally, samples with the same class form feature clusters and embedding space to improve the performance of the classification tasks.



## Loss Functions

In addition, we use cross-entropy loss  $L_C$  for the classification task. Therefore, the final loss function includes a joint loss function  $L_{total}$  that contains metric loss  $L_M$  for the embedding space and cross-entropy loss for the classification task.

$$L_{total} = \lambda L_M + L_C \quad (8)$$

$$L_C = \frac{1}{N} \sum_i^N - [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (9)$$

Where  $y_i$  represents the label of the sample  $i$ , where  $p_i$  represents the probability that the sample  $i$  is projected to be a positive class,  $\lambda$  represents the coefficient which we define as 0.05 by experience.

## Performance Evaluation

We adopt six commonly used evaluation metrics to evaluate the performance of the models objectively, including accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), F1 score (F1), area under the receiver operating characteristic curve (AUC).

## Implementation Details

We implement the proposed network based on the public platform PyTorch 1.8 and Intel Core i5-9400 CPU with 16 GB memory. Besides, we adapt stochastic gradient descent (SGD) to optimize the model, in which momentum and weight decay are set to 0.9 and 0.001 respectively.

## Validation Strategies and Statistic Analysis Methods

To evaluate the effectiveness of the proposed model, we conduct a fivefold cross-validation strategy in all ablation and comparative experiments based on the AAL atlas. For each experiment, we

divide data into five groups, and each group maintains the same proportion of two classes. In each fold experiment, four groups are used as train groups and another group is used as the test group. The detailed classification results on the ADNI database are summarized in section “Ablation Experiments.”

In addition, we apply independent testing set strategy in the experiments based on Schaefer et al. (2018) atlas. We divide the collected dataset from the ADNI database into a training set (80%), validation set (10%), and testing set (10%). The corresponding detailed classification results are summarized in sections “Experiments on Different Atlases.”

Similarly, to evaluate the effectiveness of the proposed model, we use two methods to validate the statistical significance including the  $t$ -test and DeLong test. In the experiments on the AAL atlas, we use the  $t$ -test. In the experiments on Schaefer et al. (2018) atlas, we use the DeLong test.

## RESULTS

### Ablation Experiments

To verify the effect of the bilinear pooling module and the metric learning loss on the performance of the proposed model, we remove the bilinear pooling module and the metric learning mechanism loss from the proposed BMNet, respectively. In the first experiment (i.e., our method without a bilinear pooling module), we directly use a fully connected layer to replace the bilinear pooling module. In the second experiment (i.e., our method without metric learning losses), we just use the cross-entropy loss function. The details are as follows and the results are shown in **Tables 2–5** and **Figure 2**.

**TABLE 2** | Results of the ablation studies of BP module and metric learning losses for EMCI vs. LMCI classification (Mean  $\pm$  Standard Deviation).

Method	ACC (%)	PPV (%)	NPV (%)	SEN (%)	SPE (%)	AUC	F1 (%)	p
Baseline	75.74 $\pm$ 2.96	83.79 $\pm$ 2.32	59.84 $\pm$ 13.06	80.79 $\pm$ 4.89	64.92 $\pm$ 2.09	0.7332 $\pm$ 0.0602	82.26 $\pm$ 1.53	–
Baseline + BP	78.48 $\pm$ 3.44	86.21 $\pm$ 4.22	63.29 $\pm$ 4.18	82.24 $\pm$ 2.03	70.29 $\pm$ 7.26	0.7629 $\pm$ 0.0719	84.17 $\pm$ 2.70	0.068
Tri-loss	77.35 $\pm$ 5.28	87.93 $\pm$ 5.72	56.34 $\pm$ 19.67	80.49 $\pm$ 6.45	71.3 $\pm$ 8.60	0.7415 $\pm$ 0.0963	84.05 $\pm$ 3.41	0.342
Tri-loss + BP	<b>79.64 <math>\pm</math> 3.11</b>	<b>89.31 <math>\pm</math> 2.56</b>	60.55 $\pm$ 9.54	81.84 $\pm$ 3.66	<b>74.29 <math>\pm</math> 4.18</b>	0.7589 $\pm$ 0.0633	<b>85.41 <math>\pm</math> 2.13</b>	<b>0.013*</b>
Con-loss	77.81 $\pm$ 3.05	86.55 $\pm$ 3.57	59.17 $\pm$ 9.13	80.88 $\pm$ 3.37	70.53 $\pm$ 5.67	0.7387 $\pm$ 0.0801	83.94 $\pm$ 2.22	0.342
Con-loss + BP	79.40 $\pm$ 1.92	86.90 $\pm$ 6.41	<b>64.67 <math>\pm</math> 10.58</b>	<b>83.22 <math>\pm</math> 3.86</b>	72.43 $\pm$ 6.59	<b>0.7707 <math>\pm</math> 0.0848</b>	85.01 $\pm$ 1.89	0.079

The bold values represent the highest number.

The asterisk represents the results have the statistical significance.

**TABLE 3** | Results of the ablation studies of BP module and metric learning losses for NC vs. AD classification (Mean  $\pm$  Standard Deviation).

Method	ACC (%)	PPV (%)	NPV (%)	SEN (%)	SPE (%)	AUC	F1 (%)	p
Baseline	85.25 $\pm$ 2.50	92.01 $\pm$ 2.82	76.33 $\pm$ 6.77	83.91 $\pm$ 3.96	87.99 $\pm$ 3.52	0.9074 $\pm$ 0.0215	87.77 $\pm$ 1.96	–
Baseline + BP	88.94 $\pm$ 1.20	93.52 $\pm$ 3.48	82.79 $\pm$ 3.50	87.92 $\pm$ 1.77	90.93 $\pm$ 4.33	0.9286 $\pm$ 0.0218	90.63 $\pm$ 1.22	0.051
Tri-loss	88.29 $\pm$ 0.86	93.54 $\pm$ 0.99	81.33 $\pm$ 2.04	86.92 $\pm$ 1.44	90.47 $\pm$ 1.31	0.9279 $\pm$ 0.0127	90.35 $\pm$ 0.74	0.059
Tri-loss + BP	<b>89.80 <math>\pm</math> 0.62</b>	93.53 $\pm$ 3.70	<b>84.81 <math>\pm</math> 4.88</b>	<b>89.28 <math>\pm</math> 2.79</b>	91.20 $\pm$ 4.28	0.9281 $\pm$ 0.0192	91.11 $\pm$ 0.70	<b>0.032*</b>
Con-loss	89.14 $\pm$ 1.75	93.53 $\pm$ 2.56	83.35 $\pm$ 2.78	88.02 $\pm$ 1.59	90.76 $\pm$ 3.54	0.9281 $\pm$ 0.0253	90.69 $\pm$ 1.48	0.088
Con-loss + BP	<b>89.80 <math>\pm</math> 0.99</b>	<b>93.90 <math>\pm</math> 2.44</b>	84.32 $\pm$ 3.40	88.91 $\pm$ 2.00	<b>91.40 <math>\pm</math> 2.97</b>	<b>0.9334 <math>\pm</math> 0.0141</b>	<b>91.35 <math>\pm</math> 0.80</b>	<b>0.029*</b>

The bold values represent the highest number.

The asterisk represents the results have the statistical significance.

**TABLE 4** | Results of the ablation studies of BP module and metric learning loss for NC VS. LMCI classification (Mean ± Standard Deviation).

Method	ACC (%)	PPV (%)	NPV (%)	SEN (%)	SPE (%)	AUC	F1 (%)	p
Baseline	76.81 ± 4.27	78.86 ± 9.26	73.45 ± 10.68	84.50 ± 4.26	66.81 ± 6.33	0.7527 ± 0.0520	81.58 ± 4.79	–
Baseline + BP	80.00 ± 4.61	86.28 ± 6.49	68.71 ± 8.88	83.25 ± 3.83	74.45 ± 7.40	0.7871 ± 4.57	84.74 ± 4.03	<0.001*
Tri-loss	80.49 ± 2.86	89.33 ± 4.01	64.60 ± 5.42	81.91 ± 2.04	77.68 ± 6.17	0.7702 ± 5.77	85.46 ± 2.34	0.007*
Tri-loss + BP	<b>82.20 ± 4.36</b>	<b>89.36 ± 3.17</b>	69.42 ± 11.53	84.18 ± 5.23	<b>78.48 ± 5.76</b>	0.7985 ± 5.38	<b>86.69 ± 3.05</b>	<0.001*
Con-loss	79.03 ± 4.83	83.22 ± 7.72	71.36 ± 13.62	84.31 ± 5.15	71.28 ± 6.79	0.7841 ± 4.18	83.76 ± 4.23	0.016*
Con-loss + BP	81.46 ± 3.99	84.64 ± 6.62	<b>72.02 ± 9.80</b>	<b>84.96 ± 3.64</b>	76.05 ± 7.78	<b>0.8096 ± 4.22</b>	84.80 ± 4.35	0.001*

The bold values represent the highest number.

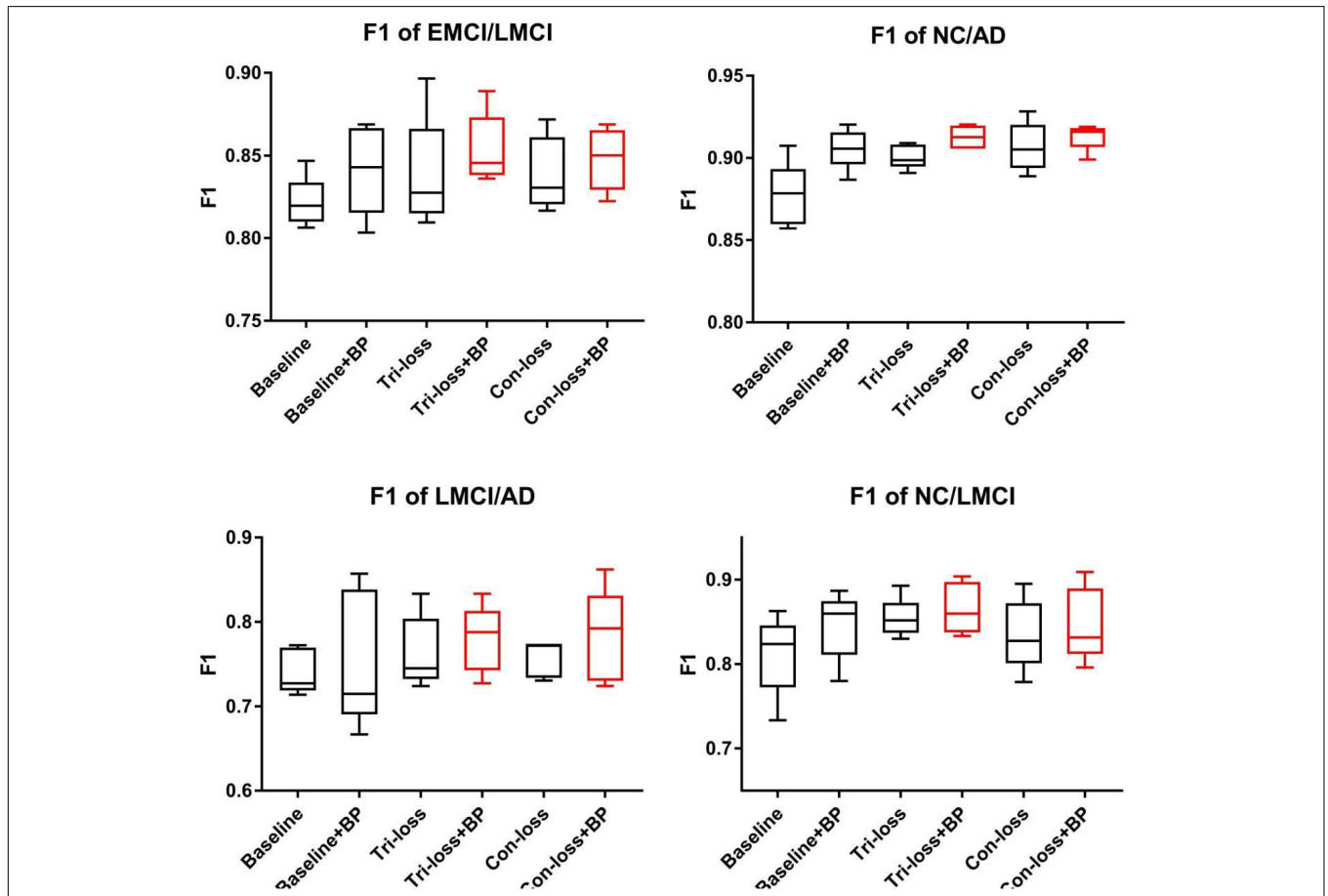
The asterisk represents the results have the statistical significance.

**TABLE 5** | Results of the ablation studies of BP module and metric learning loss for LMCI vs. AD classification (Mean ± Standard Deviation).

Method	ACC (%)	PPV (%)	NPV (%)	SEN (%)	SPE (%)	AUC	F1 (%)	p
Baseline	77.69 ± 2.67	74.16 ± 5.54	80.30 ± 6.73	74.57 ± 5.91	80.86 ± 2.54	0.7964 ± 0.0216	74.37 ± 2.67	–
Baseline + BP	80.06 ± 5.78	72.92 ± 11.12	<b>85.37 ± 4.11</b>	78.62 ± 5.95	81.29 ± 6.67	0.8108 ± 0.0581	75.66 ± 8.02	0.243
Tri-loss	80.60 ± 3.04	74.25 ± 9.76	85.31 ± 6.67	<b>79.77 ± 6.85</b>	82.23 ± 5.29	0.8040 ± 0.0422	76.91 ± 4.30	0.307
Tri-loss + BP	81.18 ± 2.72	<b>78.90 ± 7.80</b>	82.85 ± 4.39	77.53 ± 3.26	<b>84.41 ± 4.42</b>	0.8167 ± 0.0295	78.21 ± 3.97	0.022*
Con-loss	79.71 ± 0.97	74.87 ± 7.56	83.31 ± 5.89	77.53 ± 5.36	82.00 ± 3.37	0.8018 ± 0.0332	76.18 ± 2.19	0.327
Con-loss + BP	<b>81.77 ± 4.50</b>	77.54 ± 8.19	84.91 ± 6.27	79.64 ± 6.88	83.84 ± 4.97	<b>0.8297 ± 0.0346</b>	<b>78.57 ± 5.53</b>	0.028*

The bold values represent the highest number.

The asterisk represents the results have the statistical significance.



**FIGURE 2** | The F1 scores of experiments for EMCI vs. LMCI classification, NC vs. LMCI classification, LMCI vs. AD classification and NC vs. AD classification.

**TABLE 6** | Results of the main studies based on the Schaefer et al. (2018) atlas.

Class	Method	ACC	SEN	SPE	F1	AUC	<i>p</i>
EMCI-LMCI	Baseline	0.7500	0.7647	0.7000	0.8254	0.7379	<b>0.0358*</b>
	Con-loss	0.7727	0.7714	0.7778	0.8438	0.7609	0.1090
	Baseline + BP	0.7955	0.8125	0.7500	0.8525	0.7425	0.0990
	Con-loss + BP	<b>0.8409</b>	<b>0.8235</b>	<b>0.9000</b>	<b>0.8889</b>	<b>0.8529</b>	–
NC-AD	Baseline	0.8298	0.8519	0.8000	0.8519	0.8796	<b>0.0395*</b>
	Con-loss	0.8511	0.8571	0.8421	0.8727	0.9139	0.3212
	Baseline + BP	0.8511	0.8333	0.8824	0.8772	0.9259	0.3548
	Con-loss + BP	<b>0.8936</b>	<b>0.8929</b>	<b>0.8947</b>	<b>0.9091</b>	<b>0.9574</b>	–

The bold values represent the highest number.

The asterisk represents the results have the statistical significance.

### The Ablation Experiments of the Bilinear Pooling Module

Firstly, we conduct the experiments based on the Baseline model. Then we conduct the experiments of adding a bilinear pooling (BP) module to the Baseline model. According to the results, after the BP module is added, the four groups of classification experimental results have been improved to a certain extent. Specifically, in classification experiments between EMCI and LMCI, ACC increases by 2.74%, and AUC increases by 2.97%. In classification experiments between NC and AD, the results are the best, where ACC increases by 3.69% and AUC increases by 2.12%. In addition, we also conduct experiments in the classification between NC and LMCI, LMCI, and AD. The results illustrate that the BP module has a good generalization ability in the different classification tasks.

Furthermore, we conduct comparative experiments to verify the effectiveness of the BP module based on metric learning loss. For example, in the classification experiments of EMCI and LMCI, after adding the BP model to the triplet loss (Tri-loss), ACC increases by 2.29%, and AUC increases by 1.74%.

### The Ablation Experiments of Metric Learning Losses

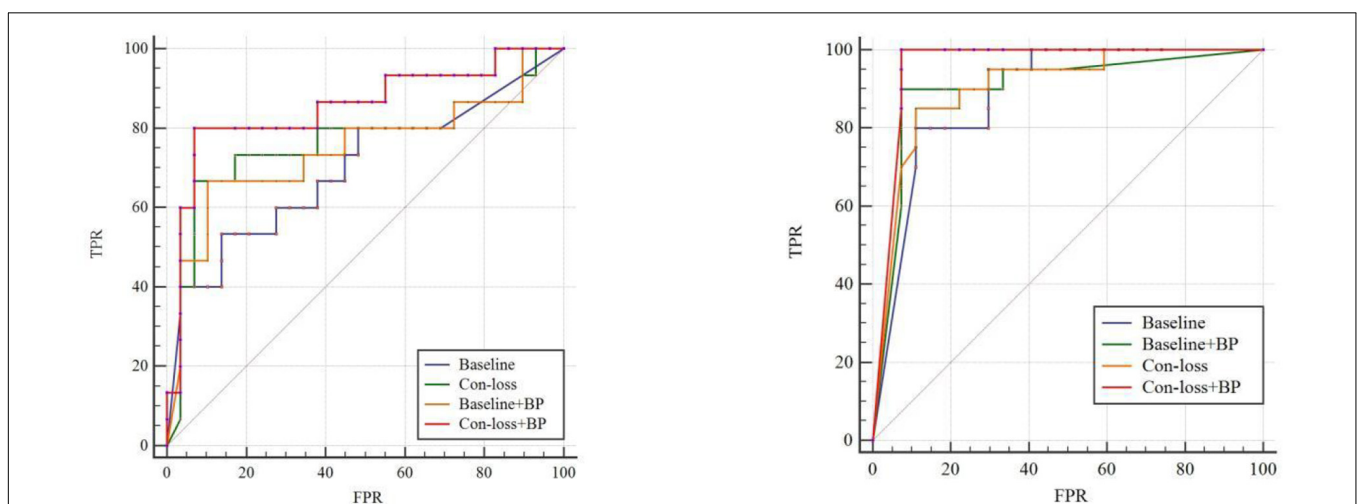
In this sub-section, we perform comparative experiments in terms of metric learning losses, including the triplet loss (Tri-loss) and the contrastive loss (Con-loss). We use two kinds of metric learning losses respectively, and the results illustrate that the two metric learning losses are both effective in different experiments. Specifically, in the classification experiments between EMCI and LMCI, ACC increases by 2.07% after adding the contrastive loss, which is a little higher than that of triplet loss. Similarly, in the classification experiments between NC and AD, ACC increases by 3.89%. In the classification experiments between NC and LMCI, the results of triplet loss improve more than these of contrastive loss, and ACC reaches 0.8049. On the contrary, in the classification experiment between LMCI and AD, the results of contrastive loss are better, where ACC reaches 0.8177 and AUC reaches 0.8297.

Finally, we use the *t*-test to measure the statistical significance comparing AUCs and the results are shown as *p*-value in **Tables 2–5**. We can see that most results of the two final models (Con-loss + BP and Tri-loss + BP) are statistically significant. In addition, we can also see that most F1 scores of the two final models are higher than these of other models in **Figure 2**.

### Experiments on Different Atlases

In this section, we evaluate the performance of our method (Con-loss + BP) based on the Schaefer et al. (2018) atlas. We conduct two groups of experiments for EMCI vs. LMCI classification and NC vs. AD classification and the results are shown in **Table 6** and **Figure 3**. As stated earlier, we apply independent testing set strategy in these experiments and use the DeLong test to validate the statistical significance.

The results illustrate that both BP module and contrastive loss are effective based on the Schaefer et al. (2018) atlas. In the experiments for EMCI vs. LMCI classification, ACC increases by 2.27% after adding the contrastive loss, which is a little



**FIGURE 3** | Receiver operating characteristic (ROC) curves of experiments for EMCI vs. LMCI classification and the ROC of experiments for NC vs. AD classification based on the Schaefer et al. (2018) atlas. TPR, true positive rate; FPR, false-positive rate; AUC, area under the receiver operating characteristic curve. Please see the web version for the complete colorful picture.

**TABLE 7** | Comparison of the performance of different model algorithms in experiments for EMCI vs. LMCI classification with the related works.

Method	Modality	DATA (EMCI/LMCI)	ACC	SEN	SPE	AUC	F1
SVM	PET	290/147	0.6620	0.7769	0.4653	0.6329	–
Singh et al., 2017	PET	178/158	–	0.6482	–	–	0.6844
Nozadi et al., 2018	PET	164/189	0.7250	0.7920	0.6990	0.790	–
Forouzannezhad et al., 2018	PET	296/193	0.6230	0.7820	0.4000	–	–
Forouzannezhad et al., 2020	PET	296/193	0.6280	0.6150	0.6430	–	–
Yang and Liu, 2020	PET	–	0.7219	0.7382	0.7305	–	–
Hao et al., 2020	PET	273/187	0.6469	0.7817	0.4444	0.6300	–
	PET+MRI		0.7387	0.9055	0.4952	0.7000	–
Fang et al., 2020	PET+MRI	297/196	0.8333	0.8235	0.8966	0.8947	
Lei et al., 2020	fMRI	44/38	0.7805	0.7368	0.8182	0.8571	–
	DTI		0.5366	0.5789	0.5000	0.5260	–
Song et al., 2021	fMRI	44/38	0.7926	0.8421	0.7500	0.9067	–
	DTI		<b>0.8292</b>	<b>0.9473</b>	0.7272	<b>0.9414</b>	–
Our method (Tri-loss+BP)	PET	290/147	0.7964	0.8184	0.7429	0.7589	<b>0.8541</b>
Our method(Con-loss+BP on Schaefer atlas)	PET	290/147	<b>0.8409</b>	<b>0.8235</b>	<b>0.9000</b>	<b>0.8889</b>	0.8529

The bold values represent the highest number.

lower than that of the BP module. Similarly, in the classification experiments between NC and AD, ACC increases by 2.13%. Finally, combining the BP module and contrastive loss, the final model (Con-loss + BP) achieves much improvement in both two classification experiments. Specifically, in the classification experiments for EMCI and LMCI, ACC, SEN, SPE, F1 and AUC achieve 84.09%, 82.35%, 90%, 88.89% and 0.8529 with an improvement of 9.09, 5.88, 20, 6.35, and 11.5% respectively, compared with Baseline model. In the NC vs. AD classification experiments, ACC, SEN, SPE, F1 and AUC increases by 6.38%, 4.1%, 9.47%, 5.72%, 7.78% and reach 89.36%, 89.29%, 89.47%, 90.91% and 0.9574.

## Comparison With Other Methods

In this section, we compare the performance of our method (Tri-loss + BP) with that of several recent representative

**TABLE 8** | Comparison of the performance of different model algorithms in experiments for NC vs. AD classification with the related works.

Method	Modality	DATA (NC/AD)	ACC	SEN	SPE	AUC
SVM	PET	263/198	0.6213	0.8063	0.5547	0.8445
Hao et al., 2020	PET	211/160	0.8006	0.8602	0.7194	0.85
	MRI		0.8663	0.9028	0.8181	0.93
Lei et al., 2020	fMRI	44/38	0.7805	0.7368	0.8182	0.8571
	DTI		0.5366	0.5789	0.5000	0.5260
Song et al., 2021	fMRI	44/38	0.7926	0.8421	0.75	0.9067
	DTI		0.8292	0.9473	0.7272	0.9414
Lian et al., 2018	MRI	429/358	0.90	0.82	<b>0.97</b>	0.95
Dong et al., 2021	MRI+PET	440/367	<b>0.9305</b>	<b>0.9474</b>	<b>0.9091</b>	0.9732
Our method (Tri-loss+BP)	PET	263/198	<b>0.898</b>	0.8928	<b>0.912</b>	0.9281
Our method (Con-loss+BP on Schaefer atlas)	PET	263/198	0.8936	<b>0.8929</b>	0.8947	<b>0.9574</b>

The bold values represent the highest number.

methods. In addition, we apply the least absolute shrinkage and selection operator (LASSO) feature selection method and support vector machine (SVM) method for the contrast experiments. From **Table 7**, we can find that our method gets the highest performance in classification experiments between EMCI and LMCI based on FDG-PET images. Specifically, the proposed

**TABLE 9** | Comparison of the performance of different model algorithms in experiments for NC vs. LMCI classification with the related works.

Method	Modality	DATA (NC/LMCI)	ACC	SEN	SPE	AUC
SVM	PET	263/147	0.6415	0.7446	0.5437	0.6724
Hao et al., 2020	PET	273/187	0.6677	0.7545	0.5594	0.68
	MRI		0.712	0.7801	0.6332	0.76
Lei et al., 2020	Fmri	44/38	0.7805	0.7368	<b>0.8182</b>	0.8571
	DTI		0.5366	0.5789	0.5000	0.5260
Song et al., 2021	fMRI	44/38	0.7926	0.8421	0.75	0.9067
	DTI		<b>0.8292</b>	<b>0.9473</b>	0.7272	<b>0.9414</b>
Our method (Tri-loss+BP)	PET	263/147	0.822	0.8418	0.7848	0.7985

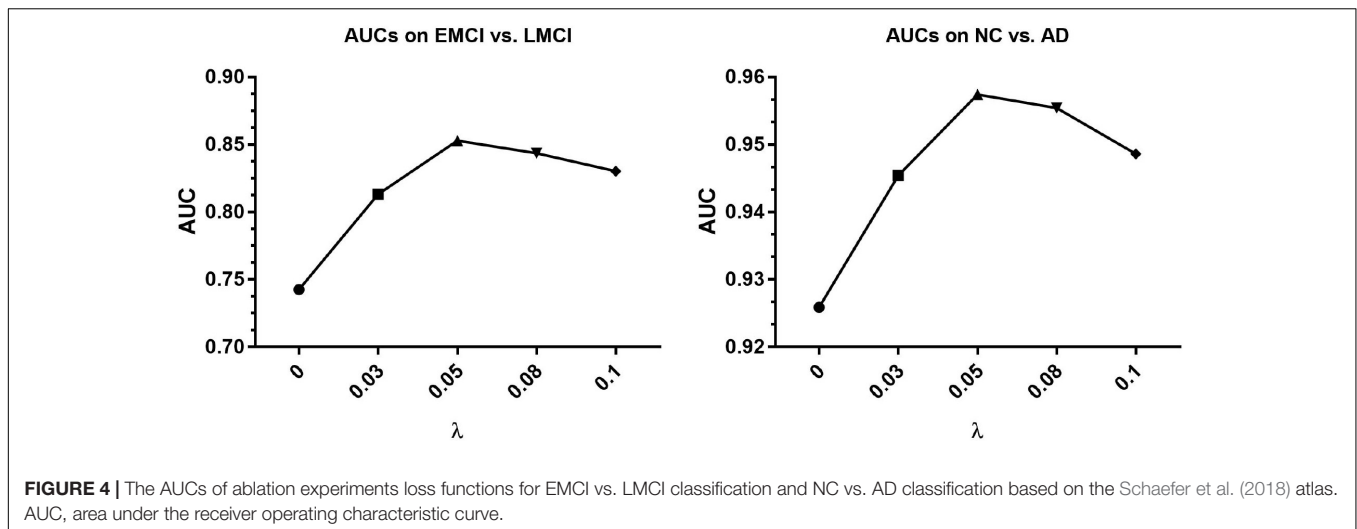
The bold values represent the highest number.

**TABLE 10** | Comparison of the performance of different model algorithms in experiments for LMCI vs. AD classification with the related works.

Method	Modality	DATA (LMCI/AD)	ACC	SEN	SPE	AUC
SVM	PET	147/198	0.5841	0.7834	0.5044	0.6908
Hao et al., 2020	PET	273/187	0.6677	0.7545	0.5594	0.68
	MRI		0.712	0.7801	0.6332	0.76
Lei et al., 2020	fMRI	44/38	0.7805	0.7368	0.8182	0.8571
	DTI		0.5366	0.5789	0.5000	0.5260
Song et al., 2021	fMRI	44/38	0.7926	0.8421	0.75	0.9067
	DTI		<b>0.8292</b>	<b>0.9473</b>	0.7272	<b>0.9414</b>
Our method (Tri-loss+BP)	PET	147/198	0.8118	0.7753	<b>0.8441</b>	0.8167

The bold values represent the highest number.





method yields big improvement than the results of Singh et al. (2017) and Nozadi et al. (2018), although the dataset in our experiments is highly unbalanced. Based on a similar dataset, the proposed method still has better performance than the methods proposed by Forouzannezhad et al. (2018, 2020). In addition, compared with the method proposed by Hao et al. (2020), our method achieves an overall huge improvement with 14.95% in ACC, 3.67% in SEN, 29.85% in SPE, and 12.89% in AUC, respectively. Compared to the results of the fusion of PET and MRI (Singh et al., 2017; Forouzannezhad et al., 2018, 2020; Nozadi et al., 2018; Fang et al., 2020; Hao et al., 2020), our method also achieves an improvement in most metrics. Besides, our method gets a comparable performance compared to the methods based on fMRI and DTI adapted by Lei et al. (2020) and by Song et al. (2021), but the subjects in our research are much more than those they use.

Similarly, from **Table 8**, we can find that our method gets the highest performance of classification experiments between NC and AD based on PET images too. Specifically, compared with the method proposed by Hao et al. (2020) based on PET images, our method achieves an overall huge improvement with 9.74% in ACC, 3.26% in SEN, 19.26% in SPE, and 7.81% in AUC, respectively. Besides, our method gets a comparable performance compared to the methods based on other modalities (Lian et al., 2018; Lei et al., 2020; Song et al., 2021). ACC, SEN, SPE, AUC of our method based on PET images improve 10.76%, 4.69%, 16.83%, and 2.82% than those of their method based on fMRI. While SEN and AUC are slightly lower, ACC and SPE based on PET images improve 7.1% and 19.11% than those based on DTI.

In addition, we conduct the classification experiments between NC and LMCI, LMCI and AD, and the results compared with other methods are shown in **Tables 9, 10** respectively, to further validate the effectiveness of our method.

From those experiments above, we can see that our classification results between EMCI and LMCI have exceeded those of the existing methods overall based on FDG-PET images. In addition, our results are also comparable with those based on fMRI and DTI images.

## DISCUSSION AND CONCLUSION

### Comparison of Different Coefficients in Loss Functions

To select the proper coefficient of loss functions, we compare several numbers of coefficient  $\lambda$  in Equation 8, including 0, 0.03, 0.05, 0.08, and 0.1. We conduct the ablation experiments based on methods in section “Experiments on Different Atlases” and the corresponding AUCs are shown in **Figure 4**. It can be seen that the AUC turns out to be the highest when coefficient is around 0.05 and keep at a relatively high level in the range from 0.05 to 0.08. Therefore we set coefficient  $\lambda$  as 0.05 in most experiments.

### Comparisons With Previous Researches

In general, there are three major advances between the proposed method and previous methods. Firstly, current PET-based methods are deficient in extracting representation features among different brain regions, incurring poor performance for the classification of early AD. The proposed BMNet introduces a bilinear pooling module into the model to explore the inter-region representation features and get a good classification performance. Secondly, there are few methods to study hard samples to improve the classification results in the brain disorder analysis. By comparison, we apply two metric learning losses to our model which has been proved useful for hard samples classification and they both get a good performance in the experiments. Thirdly, brain metabolism is very important for AD diagnosis and can only be obtained by PET images. Based on PET images, the proposed method could extract region-based features which represent the brain metabolic connectivity network, excavate the potential of PET images, and improve the diagnosis performance. This is the main superiority of inter-region-based methods with PET images compare with other modalities. In addition, the proposed PET-based method is comparable to other modalities in classification tasks between EMCI and LMCI.

## Potential Applications in Other Modalities

Considering the good performance based on FDG-PET images, the proposed BMNet including the bilinear pooling module and the metric learning loss functions also has the potential capability of diagnosis for other neurological diseases with other kinds of brain images. Besides, the proposed method only requires features of each brain region as the input. This lightweight characteristic allows the model to be easily applied to fMR and DTI images. We will try to explore more applications of the proposed method in future work.

## Limitations and Future Works

While the proposed BMNet achieves good results for the diagnosis of early AD, there are still some limitations. Firstly, considering the characteristics of the convolution neural network, the models and results are hard to be interpreted and the inter-region representation of the brain regions is hard to be visualized. Secondly, the proposed method focus on region-based features, which are lightweight but only utilize the metabolism of brain regions, limiting the ability of the network. In future work, we will try to integrate whole 3D PET images into the network to achieve joint feature extraction and classification. Finally, there is still some potential in exploiting methods that can extract brain inter-region representation features based on FDG-PET images. In the future, we will try to design methods that could extract inter-region representation features more effectively. In addition, the proposed method directly combined the contrastive loss and triplet loss with the entropy loss to better distinguish the hard samples. In the future, we will some novel designs of these losses based on domain knowledge.

## Conclusion

We propose a novel neural network method for the diagnosis of early AD with FDG-PET. We firstly construct a shallow neural network as the Baseline model. Then we introduce a bilinear pooling module into the network to try to extract inter-region representation features among the whole brain. We also apply the deep metric learning losses into the final loss function to help distinguish hard samples in the embedding space. Finally, we conduct the BMNet on the ADNI database and the results show that

our method yields comparable classification performance with several representative methods. Especially, we get a satisfying classification performance in the experiment between EMCI and LMCI, which is the state-of-the-art result with FDG-PET.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

WC contributed to idea conceptualization, experiments, and wrote the first draft of the manuscript. CY and CL collected data and organized the database. YP contributed to idea conceptualization. YL contributed to experiments. SC contributed to fund support and manuscript revising. XJ and JZ contributed to manuscript revising and idea conceptualization. ZY contributed to manuscript revising. XY contributed to fund support. WC, YP, JZ, XJ, and CY contributed to conception and design of the study. WC and YL performed the statistical analysis. XJ, SC, JZ, XY, and ZY wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work is supported in part by the Key Project of Health Commission of Jiangsu Province (ZDB2020011), in part by the Technology Project of Diagnosis and treatment of key diseases in Suzhou (LCZX201909), and in part by the Suzhou Science and Technology Bureau (SJC2021023).

## REFERENCES

- Alzheimer's Association (2018). Alzheimer's disease facts and figures. *Alzheimer's Dementia* 14, 367–429.
- Alzheimer's Association (2019). 2019 Alzheimer's disease facts and figures. *Alzheimer's Dementia* 15, 321–387.
- Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821.
- Blazhenets, G., Ma, Y., Sörensen, A., Rücker, G., Schiller, F., Eidelberg, D., et al. (2019). Principal components analysis of brain metabolism predicts development of Alzheimer dementia. *J. Nucl. Med.* 60, 837–843. doi: 10.2967/jnumed.118.219097
- Cheng, D., Gong, Y., Zhou, S., Wang, J., and Zheng, N. (2016). “Person re-identification by multi-channel parts-based cnn with improved triplet loss function.” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ: IEEE), 1335–1344.
- Choi, H., Jin, K. H., Alzheimer's Disease, and Neuroimaging Initiative (2018). Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behav. Brain Res.* 344, 103–109. doi: 10.1016/j.bbr.2018.02.017
- Dong, A., Li, Z., Wang, M., Shen, D., and Liu, M. (2021). High-Order laplacian regularized low-rank representation for multimodal dementia diagnosis. *Front. Neurosci.* 15:634124. doi: 10.3389/fnins.2021.634124
- Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., et al. (2016). Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimer's Dementia* 12, 292–323. doi: 10.1016/j.jalz.2016.02.002

- Fang, C., Li, C., Forouzaneshad, P., Cabrerizo, M., Curiel, R. E., Loewenstein, D., et al. (2020). Gaussian discriminative component analysis for early detection of Alzheimer's disease: a supervised dimensionality reduction algorithm. *J. Neurosci. Methods* 344:108856. doi: 10.1016/j.jneumeth.2020.108856
- Forouzaneshad, P., Abbaspour, A., Li, C., Cabrerizo, M., and Adjouadi, M. (2018). "A deep neural network approach for early diagnosis of mild cognitive impairment using multiple features," in *Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (Piscataway, NJ: IEEE), 1341–1346.
- Forouzaneshad, P., Abbaspour, A., Li, C., Fang, C., Williams, U., Cabrerizo, M., et al. (2020). A Gaussian-based model for early detection of mild cognitive impairment using multimodal neuroimaging. *J. Neurosci. Methods* 333:108544. doi: 10.1016/j.jneumeth.2019.108544
- Gao, Z., Wu, Y., Zhang, X., Dai, J., Jia, Y., and Harandi, M. (2020). "Revisiting bilinear pooling: a coding perspective," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34 (New York, NY), 3954–3961. doi: 10.1609/aaai.v34i04.5811
- Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., et al. (2006). Mild cognitive impairment. *Lancet* 367, 1262–1270.
- Hadsell, R., Chopra, S., and LeCun, Y. (2006). *Dimensionality Reduction by Learning an Invariant Mapping*, Vol. 2. Piscataway, NJ: IEEE, 1735–1742.
- Hampel, H., and Lista, S. (2016). The rising global tide of cognitive impairment. *Nat. Rev. Neurol.* 12, 131–132. doi: 10.1038/nrneuro.2015.250
- Hao, X., Bao, Y., Guo, Y., Yu, M., Zhang, D., Risacher, S. L., et al. (2020). Multimodal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. *Med. Image Anal.* 60:101625. doi: 10.1016/j.media.2019.101625
- He, K., Lian, C., Adeli, E., Huo, J., Gao, Y., Zhang, B., et al. (2021). MetricUNet: synergistic image- and voxel-level learning for precise prostate segmentation via online sampling. *Med. Image Anal.* 71:102039. doi: 10.1016/j.media.2021.102039
- Huang, F., Tan, E. L., Yang, P., Huang, S., Ou-Yang, L., Cao, J., et al. (2020). Self-weighted adaptive structure learning for ASD diagnosis via multi-template multi-center representation. *Med. Image Anal.* 63:101662. doi: 10.1016/j.media.2020.101662
- Huang, J., Zhou, L., Wang, L., and Zhang, D. (2020). Attention-diffusion-bilinear neural network for brain network analysis. *IEEE Trans. Med. Imaging* 39, 2541–2552. doi: 10.1109/TMI.2020.2973650
- Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., et al. (2010). Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *NeuroImage* 50, 935–949. doi: 10.1016/j.neuroimage.2009.12.120
- Jack, C. R. Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magnetic Resonance Imag.* 27, 685–691. doi: 10.1002/jmri.21049
- Kim, J. H., On, K. W., Lim, W., Kim, J., Ha, J. W., and Zhang, B. T. (2016). Hadamard product for low-rank bilinear pooling. *arXiv [preprint]*. Available online at: <https://arxiv.org/abs/1610.04325> (accessed March 20, 2021).
- Lei, B., Cheng, N., Frangi, A. F., Tan, E. L., Cao, J., Yang, P., et al. (2020). Self-calibrated brain network estimation and joint non-convex multi-task learning for identification of early Alzheimer's disease. *Med. Image Anal.* 61:101652. doi: 10.1016/j.media.2020.101652
- Li, Y., Wang, N., Liu, J., and Hou, X. (2017). "Factorized bilinear models for image recognition," in *Proceedings of the IEEE International Conference on Computer Vision* (Piscataway, NJ: IEEE), 2079–2087.
- Lian, C., Liu, M., Zhang, J., and Shen, D. (2018). Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 880–893. doi: 10.1109/TPAMI.2018.2889096
- Lin, T. Y., RoyChowdhury, A., and Maji, S. (2015). "Bilinear cnn models for fine-grained visual recognition," in *Proceedings of the IEEE International Conference on Computer Vision* (Piscataway, NJ: IEEE), 1449–1457.
- Liu, M., Cheng, D., Yan, W., Alzheimer's Disease, and Neuroimaging Initiative (2018). Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. *Front. Neuroinform.* 12:35. doi: 10.3389/fninf.2018.00035
- Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., and Song, L. (2017). "Sphereface: deep hypersphere embedding for face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ: IEEE), 212–220. doi: 10.1109/TPAMI.2019.2914680
- Matlab (2020). *MATLAB 9.8*. MathWorks Inc. Available online at: <https://www.mathworks.cn/products/matlab.html>
- Nozadi, S. H., Kadoury, S., Alzheimer's Disease, and Neuroimaging Initiative (2018). Classification of Alzheimer's and MCI patients from semantically parcelled PET images: a comparison between AV45 and FDG-PET. *Int. J. Biomed. Imag.* 2018:1247430. doi: 10.1155/2018/1247430
- Pagani, M., Nobili, F., Morbelli, S., Arnaldi, D., Giuliani, A., Öberg, J., et al. (2017). Early identification of MCI converting to AD: a FDG PET study. *Eur. J. Nucl. Med. Mol. Imag.* 44, 2042–2052. doi: 10.1007/s00259-017-3761-x
- Pan, Y., Liu, M., Xia, Y., and Shen, D. (2021). Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data. *IEEE Trans. Pattern Anal. Mach. Intell.* Online ahead of print. doi: 10.1109/TPAMI.2021.3091214
- Sanabria-Diaz, G., Martinez-Montes, E., Melie-Garcia, L., Alzheimer's Disease, and Neuroimaging Initiative (2013). Glucose metabolism during resting state reveals abnormal brain networks organization in the Alzheimer's disease and mild cognitive impairment. *PLoS One* 8:e68860. doi: 10.1371/journal.pone.0068860
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., et al. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* 28, 3095–3114. doi: 10.1093/cercor/bhx179
- Schroff, F., Kalenichenko, D., and Philbin, J. (2015). "Facenet: a unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ: IEEE), 815–823.
- Singh, S., Srivastava, A., Mi, L., Caselli, R. J., Chen, K., Goradia, D., et al. (2017). "Deep-learning-based classification of FDG-PET data for Alzheimer's disease categories," in *Proceedings of the 13th International Conference on Medical Information Processing and Analysis* (Bellingham, WA: International Society for Optics and Photonics). doi: 10.1117/12.2294537
- Song, X., Zhou, F., Frangi, A. F., Cao, J., Xiao, X., Lei, Y., et al. (2021). Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction. *Med. Image Anal.* 69:101947. doi: 10.1016/j.media.2020.101947
- Sörensen, A., Blazhenets, G., Rücker, G., Schiller, F., Meyer, P. T., Frings, L., et al. (2019). Prognosis of conversion of mild cognitive impairment to Alzheimer's dementia by voxel-wise Cox regression based on FDG PET data. *NeuroImage: Clin.* 21:101637. doi: 10.1016/j.nicl.2018.101637
- Sundgaard, J. V., Harte, J., Bray, P., Laugesen, S., Kamide, Y., Tanaka, C., et al. (2021). Deep metric learning for otitis media classification. *Med. Image Anal.* 71:102034. doi: 10.1016/j.media.2021.102034
- Titov, D., Diehl-Schmid, J., Shi, K., Pernecky, R., Zou, N., Grimmer, T., et al. (2017). Metabolic connectivity for differential diagnosis of dementing disorders. *J. Cereb. Blood Flow Metab.* 37, 252–262. doi: 10.1177/0271678X15622465
- Tzourio-Mazoyer, N., Landeau, B., Papatathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., et al. (2013). Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol. Psychiatry* 73, 472–481. doi: 10.1016/j.biopsych.2012.03.026
- Wang, M., Jiang, J., Yan, Z., Alberts, I., Ge, J., Zhang, H., et al. (2020). Individual brain metabolic connectome indicator based on Kullback-Leibler Divergence Similarity Estimation predicts progression from mild cognitive impairment to Alzheimer's dementia. *Eur. J. Nucl. Med. Mol. Imaging* 47, 2753–2764. doi: 10.1007/s00259-020-04814-x
- Yang, Z., and Liu, Z. (2020). The risk prediction of Alzheimer's disease based on the deep learning model of brain 18F-FDG positron emission tomography. *Saudi J. Biol. Sci.* 27, 659–665. doi: 10.1016/j.sjbs.2019.12.004
- Zhou, H., Jiang, J., Lu, J., Wang, M., Zhang, H., Zuo, C., et al. (2019). Dual-model radiomic biomarkers predict development of mild cognitive impairment progression to Alzheimer's disease. *Front. Neurosci.* 12:1045. doi: 10.3389/fnins.2018.01045
- Zhou, P., Jiang, S., Yu, L., Feng, Y., Chen, C., Li, F., et al. (2021). Use of a sparse-response deep belief network and extreme learning machine to discriminate Alzheimer's disease, mild cognitive impairment, and normal controls based

on amyloid PET/MRI images. *Front. Med.* 7:621204. doi: 10.3389/fmed.2020.621204

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Cui, Yan, Yan, Peng, Leng, Liu, Chen, Jiang, Zheng and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*