The Visual Computer (2023) 39:2235–2244
https://doi.org/10.1007/s00371-021-02354-5

ORIGINAL ARTICLE

Pairwise feature-based generative adversarial network for incomplete multi-modal Alzheimer's disease diagnosis

Haizhou Ye[1] · Qi Zhu[1] · Yuan Yao[1] · Yichao Jin[2] · Daoqiang Zhang[1]

Accepted: 27 October 2021 / Published online: 10 January 2022

Abstract

Magnetic resonance imaging (MRI) and positron emission tomography (PET) are widely used in diagnosis of Alzheimer's disease (AD). In practice, incomplete modality problem is unavoidable due to the cost of data acquisition. Deep learning based models especially generative adversarial networks (GAN) are usually adopted to impute missing images. However, there are still some problems: (1) there are many regions unrelated to the disease and have little significance in the actual diagnosis in brain images, which are very cumbersome to generate. (2) The image generated by GAN would introduce noises causing the poor performance in the diagnostic model. To address these problems, a pairwise feature-based generation adversarial network is proposed. Specifically, features from the original brain images are extracted firstly. For the paired data without modality loss, the extracted MRI features are used as input to generate its corresponding PET features, which not only reduces the scale of the model, but also ensures the direct correlation between the generated features and the diagnosis. In addition, the available real PET features of the paired samples are added as label to constrain the generated ones. Finally, the attention mechanism is adopted in both the generator and discriminator, which can effectively retain the structural information of the feature itself. A large number of experiments have demonstrated that our proposed method has achieved promising results in the diagnosis of AD.

Keywords Incomplete multi-modal · GAN · AD · Medical imaging

1 Introduction

Alzheimer's disease is a severe, irreversible syndrome with an increasing incidence over the years [1]. A slow and progressive decline in cognitive abilities is the main character of this disease, including memory, language, executive function and visuospatial function, which interferes with the daily life of patients [2].

Meanwhile, Alzheimer's disease is the most common cause of dementia by far, accounting for 80 percent of all dementia diagnoses [3]. Early diagnosis of AD is of great significance for subsequent treatment because the disease has usually progressed for many years by the time it is confirmed [4, 5].

Magnetic resonance imaging (MRI) and positron emission tomography (PET) are the main methods used in clinical diagnosis of AD [6, 7]. They provide comprehensive information about the brain of the patient. In recent years, it has been generally adopted in computer-aided AD diagnosis. A large number of studies have shown that the two modalities can complement each other and have a mutual promotion in the diagnostic accuracy [8–10]. However, there are many subjects who do not have all the MRI and PET images due to the situation that they do not complete all the examinations in the same hospital or they never take the complete examinations, as shown in Fig. 1, which leads to a lot of incomplete data problems.

✉ Qi Zhu
zhuqinuaa@163.com

Haizhou Ye
haizhouye@163.com

Yuan Yao
y_yao@nuaa.edu.cn

Yichao Jin
jinyichao@indeed.com

Daoqiang Zhang
dqzhang@nuaa.edu.cn

[1] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

[2] Indeed Singapore Pte Ltd, Singapore 049315, Singapore

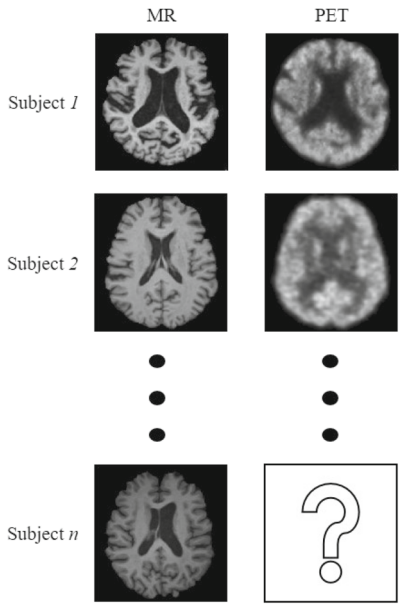


Fig. 1 Illustration of the incomplete multi-modal dataset problem. All subjects have the images of MRI, but some subjects do not have the images of PET

Most of the traditional multi-modal-based methods discard the incomplete subjects directly and utilize the complete modalities to train the diagnostic model [6, 11]. When the number of subjects is limited, discarding these available subjects can lead to poor performance of the trained model. In order to make the full use of all available subjects, many methods have been proposed to deal with the problem of missing data. Non-negative matrix factorization based methods usually adopt matrix decomposition to obtain the hidden space representation of different modalities after alignment [12]. In addition, a reverse mapping can be used to impute the missing values [13]. However, these methods can only deal with the missing of random values, but do not perform well in the condition of the whole modality missing. Later, partial multi-view clustering (PVC) and other variants improved from non-negative matrix factorization have been proposed to deal with multi-modal incomplete problem. Multi-modal data are mapped to a common hidden space representation by preserving the structural information of original space between samples by Laplacian matrices. Nowadays, some deep learning methods are used to deal with the incomplete multi-modal problem. They learn the mapping from one modality to another by deep neural networks, especially GAN [14, 15]. The generator synthesizes the target image and the discriminator distinguish the fake images from real ones. These methods aim to generate brain images [16, 17]. But even preprocessed brain images still contain a lot of useless information, for example, pattern noise and regions not associated with the disease. In the process of synthesizing images, all of the image need to be generated and identified even the uninterested regions, which greatly increases the scale of the model and the cost of training, and may also lead to the deviation of the regions of interest in diagnosis.

However, for the generation of extracted features, smaller model size is required and the training is relatively simple, and the generation of irrelevant regions is omitted, which is a worth studying problem.

In this paper, a pairwise feature-based generative adversarial network has been proposed to address the incomplete multi-modal problem. Its framework is shown in Fig. 2. The available subjects with paired MRI and PET images are utilized to train the model. Different from the other works, the extracted features instead of the original images are synthesized after the feature extraction of the image by deep network. At the same time, in order to better retain the structure of the extracted features, the attention layer is adopted after the input layer of both generator and discriminator. In addition, since the training of GAN is often unstable and prone to gradient disappearance or gradient explosion, $Smooth\,L1$ loss is adopted to constraint the distance between available real PET features and the forged PET features output by generator. Our major contributions are as follows:

(1) The proposed method is able to deal with the missing of whole modality, not just the missing in random values, which is more practice for clinical.
(2) Extracted features rather than original images are synthesized by the generator, which effectively avoids noise and areas unrelated to disease in the images, greatly improves the efficiency of generation and strengthening the regions associated with the disease to improve the diagnostic accuracy.
(3) The available real features are taken as the label into the training process of generator, which make up for the unstable GAN and make the model converge more quickly. The experimental results show that our proposed method obtains promising diagnosis performance with incomplete MRI and PET data.

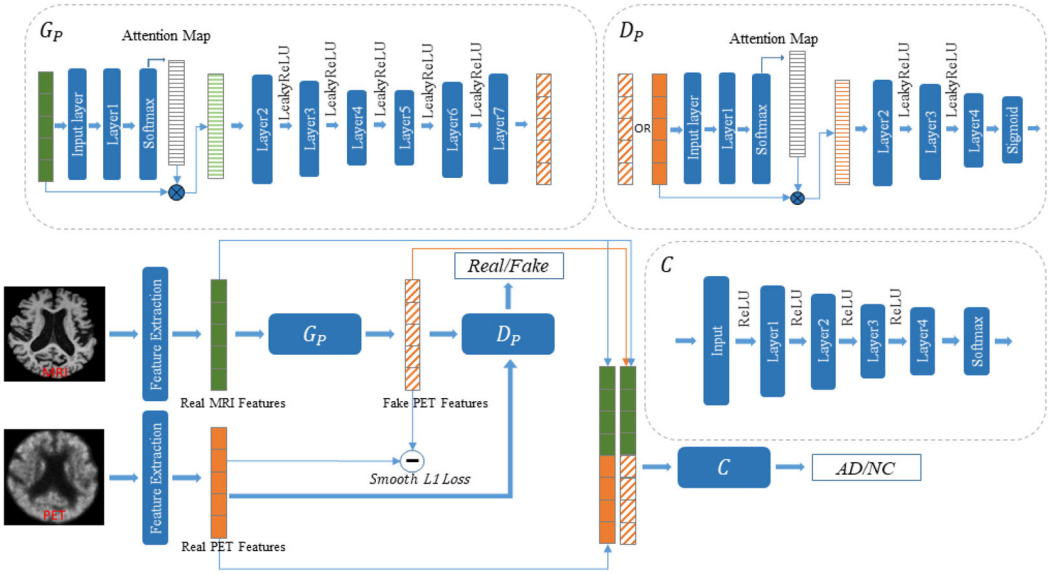


Fig. 2 Illustration of the framework of the proposed method, i.e., pairwise feature-based generative adversarial network. G_P is the generator which takes real MRI features as input and synthesized fake PET features as output. *Smooth L*1 loss is adopted in the training procedure. D_P distinguishes PET features from fake to real. Attention mechanism is added to both G_P and D_P to retain the structure of the extracted features. All layers are fully connected and updated by backpropagation

2 Methods

2.1 Generative adversarial networks

Original GAN architecture includes a generator G and a discriminator D. G tries to forge the real subjects according to a random noise z, and D distinguishes the forged subjects from real ones. Precisely, D outputs 1/0 when the input subject is real/fake. With the improvement of discriminator's capabilities, the generator will be further strengthened. In the end, G is able to fit the data distribution of the real subjects and D outputs 0.5 for both real and fake subjects. The loss function is as follows:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x-P_{data}(x)}[\log \mathrm{D}(x)] + \mathbb{E}_{z-P_x}\left[\log(1-D(G(z)))\right] \quad (1)$$

Among them, x is the real subject and z is the random noise.

In our proposed method, the generator and discriminator are retained as basic structure. For the extracted features, MRI data are the input to forge the corresponding PET data. Our purpose is to learn a mapping which can transfer real MRI features to fake PET features. Since GAN is able to synthesize any input into a subject that fits the target data distribution, when our input is a real sample, it can also obtain a generated subject that conforms to the probability distribution of the real data. Specifically, X_M and X_P denote the extracted real MRI and PET features, respectively. We aim to learn a feature generator G_P: $X_M \rightarrow \hat{X}_P$, which can impute $\hat{X}_P$ by $G_P(X_M)$ to complete the missing PET. $\hat{X}_P$ denotes the generated PET features. Besides, discriminator D_P is trained to distinguish $\hat{X}_P$ from X_P. To be specific, for the forged features $\hat{X}_P$, D_p outputs 0 and for the real features X_P, D_p outputs 1. The adversarial loss is defined as follows:

$$L_{adv} = \log(D_P(X_P)) + \log(1 - D_P(G_P(X_M))) \quad (2)$$

In the beginning of the training procedure, G_P is at a disadvantage to D_p in the adversariality. With epoch iteration, the output of G_P is getting closer to the real subjects and D_p will getting stronger to identify the forged input. In the end, Nash equilibrium was reached between two networks [18]. G_P is able to outputs high quality of forged PET features.

2.2 Attention layer

At the beginning of our network, input features of both G_P and D_P would go through an attention layer. Inspired by human attention mechanism theories, attention mechanism has achieved great success in natural language processing tasks and computer vision [19]. Attention mechanism is able to learn the importance of different parts of the input and assign different weights to these parts according to the importance [20], by which the parts that are more important for the diagnosis of the disease in input features are highlighted and original feature structure also be retained.

In our framework, the input is feature after extraction, which is itself an important part of the original image for diagnostic tasks. To improve the efficiency of generation, for the generator, key information of input MRI features which are crucial to produce the fake PET features will be emphasized by the attention mechanism. For the discriminator, more attention will be given to the important parts to identify the authenticity of input PET features which will greatly improve the identification ability of the discriminator. The attention layer can be formalized as follows:

$$q = \phi(\omega_1 x + b_1) \tag{3}$$

$$\alpha^i = \frac{\exp\left(\omega_2^i q + b_2^i\right)}{\sum_{i=1}^{n} \exp\left(\omega_2^i q + b_2^i\right)} \tag{4}$$

$$x_{adjusted} = [\alpha^1, \alpha^2, ..., \alpha^n] \bullet x \tag{5}$$

where x are the input MRI or PET features for generator and discriminator, respectively, q is the output of the hidden layer, $\phi(.)$ is the activation function to improve the nonlinear capability, α is the final attention scores for each feature in x, and $x_{adjusted}$ denotes the weighted features according to the attention scores. Besides, $\{\omega_1, b_1, \omega_2, b_2\}$ are the parameters of hidden layer and attention layer to be learned in the training process.

2.3 Pairwise feature-based GAN

In the general cases, the training of GAN is unsupervised, which bring about a lot of problems. For example, the discriminator tends to have the upper hand over the generator when the target is too complicated, which causes the generator to be unable to upgrade. Besides, the training procedure is easy to get caught up in the problem that gradient explosive and gradient vanishing [21]. In the diagnosis of AD, it is far from enough to fit the probability distribution of the target space due to the generated PET features will eventually be used for accurate diagnosis. Therefore, more discriminative information should be retained. To address these issues, pairwise features are adopted in our framework to constrain the producing procedure in the model. Different from the basic GAN, we emphasize the consistency of input MRI and output PET features in G_P. The forged PET features will be as similar as possible to the real PET features. A pure adversarial training is not able to achieve the ideal outcome. Hence, its corresponding real PET feature of each subject is adopted in the loss function when the input is MRI features. Specifically, $Smooth\ L1$ loss is adopted in the generator:

$$L_{pairwise}\left(\hat{X}_P, X_P\right) = \begin{cases} 0.5\left|\hat{X}_P - X_P\right|^2, & if\ \left|\hat{X}_P - X_P\right| < 1 \\ \left|\hat{X}_P - X_P\right| - 0.5, & otherwise \end{cases} \tag{6}$$

In this way, the generated PET feature and the real PET feature will be as close as possible in Euclidean distance.

We combine the adversarial loss with the pairwise loss. The overall loss is as follows:

$$L = L_{adv} + L_{pairwise} \tag{7}$$

In the end, a classifier C based on multi-layer perceptron is adopted in our model to make the final diagnosis according to the data completed by the generated PET features.

3 Experiments

The 913-ADNI dataset used in the experiments was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (www.loni.ucla.edu). The dataset includes five modalities, ID (serial number), single nucleotide polymorphism (SNPdata), voxel-based morphometry (VBM), fluorodeoxyglucose positron emission tomography (FDG) and F-18 florbetapir PET scans amyloid imaging (AV45) with AD, MCI and NC. Among them, there are 160 ADs, 542 MCIs and 211 NCs. The 542 MCI patients have three phases, like significant memory concern (SMC), early mild cognitive impairment (EMCI) and late mild cognitive impairment (LMCI). In our work, only VBM, FDG and AV45 modalities are adopted. The specific information is listed in Table 1.

3.1 Materials and experimental setup

Image pre-processing was performed to all MRI and PET images in 913-ADNI database which were obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). First of all, the multimodality image data of MRI (VBM) and PET (FDG, AV45) are aligned to the same visit scan with MNI152 template. Second, normalized gray matter density maps are created from VBM data in the standard Montreal Neurological Institute (MNI) space and utilized the SPM software package [22] to register the FDG and AV45 scans into same space. Based on the MarsBaR AAL atlas [23], the mean gray matter density of 116 regions of interest was measured. Then, the FDG glu-

Table 1 Information of studied samples in 913-ADNI dataset (NC = normal control; the values are denoted as mean ± standard deviation)

Subjects	Numbers	Gender(M/F)	Age
NC	210	109/101	76.1 ± 6.5
SMC	82	33/49	72.4 ± 5.7
EMCI	272	153/119	71.5 ± 7.1
LMCI	187	108/79	73.8 ± 8.4
AD	160	95/65	75.2 ± 7.9

Table 2 Diagnosis results with VBM/FDG of compared methods and our proposed pairwise feature-based generative adversarial network

Task		AD vs. NC				LMCI vs. EMCI			
Missing Rate	Method	ACC(%)	SEN(%)	SPE(%)	AUC(%)	ACC(%)	SEN(%)	SPE(%)	AUC(%)
10%	PVC	79.46 ± 1.96	73.60 ± 1.82	83.91 ± 1.19	78.75 ± 0.93	67.33 ± 2.45	77.53 ± 1.06	62.41 ± 3.80	71.44 ± 1.97
	UEAF	90.07 ± 1.58	87.04 ± 1.72	92.19 ± 1.93	89.62 ± 1.79	84.22 ± 1.34	92.70 ± 1.36	70.83 ± 1.42	82.44 ± 1.77
	GAIN	91.43 ± 2.01	85.47 ± 2.35	**96.67 ± 1.67**	91.07 ± 1.36	82.61 ± 1.35	91.19 ± 1.74	76.55 ± 3.88	84.13 ± 3.47
	Ours	**93.27 ± 1.45**	**92.95 ± 2.11**	93.41 ± 1.21	**93.18 ± 1.28**	**87.34 ± 1.29**	**93.57 ± 1.78**	**80.19 ± 1.48**	**86.47 ± 1.55**
30%	PVC	78.11 ± 1.54	74.56 ± 1.72	81.44 ± 1.53	78.00 ± 1.49	65.99 ± 2.96	80.30 ± 1.52	59.92 ± 3.84	69.17 ± 2.38
	UEAF	88.75 ± 2.08	84.64 ± 2.66	92.09 ± 1.72	88.36 ± 1.87	82.67 ± 1.51	88.75 ± 1.03	73.94 ± 1.99	82.49 ± 2.56
	GAIN	90.06 ± 1.59	85.99 ± 1.40	**94.98 ± 1.39**	90.49 ± 1.33	80.65 ± 1.06	90.68 ± 1.55	68.02 ± 4.52	80.46 ± 2.47
	Ours	**92.45 ± 2.15**	**92.49 ± 1.61**	92.52 ± 1.41	**92.51 ± 1.57**	**86.45 ± 1.23**	**92.66 ± 1.94**	**76.99 ± 1.94**	**85.62 ± 1.33**
50%	PVC	75.94 ± 2.76	70.75 ± 1.27	80.11 ± 2.14	75.43 ± 2.38	63.56 ± 1.46	70.50 ± 1.19	57.63 ± 3.49	67.56 ± 2.56
	UEAF	87.30 ± 0.63	84.69 ± 0.39	90.09 ± 1.56	87.39 ± 0.73	81.33 ± 1.63	86.87 ± 1.10	69.18 ± 1.06	79.80 ± 1.44
	GAIN	89.58 ± 1.14	84.62 ± 2.10	94.26 ± 1.38	89.44 ± 1.39	75.87 ± 1.36	91.82 ± 1.57	65.31 ± 4.17	78.14 ± 3.09
	Ours	**91.37 ± 1.87**	**88.06 ± 1.06**	**92.47 ± 1.57**	**91.35 ± 0.84**	**85.11 ± 1.42**	**93.63 ± 1.52**	**75.16 ± 1.46**	**83.73 ± 1.39**

Table 3 Diagnosis results with VBM/AV45 of compared methods and our proposed pairwise feature-based generative adversarial network

Task		AD vs. NC				LMCI vs. EMCI			
Missing Rate	Method	ACC(%)	SEN(%)	SPE(%)	AUC(%)	ACC(%)	SEN(%)	SPE(%)	AUC(%)
10%	PVC	80.00 ± 1.47	71.75 ± 1.89	86.45 ± 1.93	79.10 ± 1.73	72.67 ± 2.18	80.48 ± 1.99	53.70 ± 1.36	65.62 ± 2.28
	UEAF	90.03 ± 1.73	84.82 ± 2.06	**94.75 ± 1.54**	89.41 ± 1.87	83.78 ± 2.83	**94.05 ± 1.92**	71.36 ± 2.17	82.03 ± 1.60
	GAIN	91.16 ± 1.80	87.76 ± 1.44	94.18 ± 1.65	90.97 ± 1.01	85.56 ± 3.33	91.71 ± 2.37	71.57 ± 3.28	81.17 ± 1.44
	Ours	**92.52 ± 2.14**	**91.79 ± 2.26**	92.26 ± 1.38	**92.27 ± 2.11**	**87.56 ± 1.35**	92.76 ± 1.79	**77.52 ± 1.25**	**85.55 ± 1.52**
30%	PVC	77.30 ± 1.56	73.04 ± 2.04	81.44 ± 1.69	77.24 ± 0.56	71.11 ± 3.19	78.42 ± 1.68	46.04 ± 1.60	63.17 ± 2.19
	UEAF	88.65 ± 1.89	84.03 ± 0.97	92.90 ± 1.77	88.47 ± 1.63	82.97 ± 1.41	91.05 ± 1.93	74.22 ± 2.35	81.49 ± 1.62
	GAIN	90.03 ± 1.90	82.88 ± 1.04	**95.17 ± 1.49**	89.02 ± 1.66	82.75 ± 2.43	92.90 ± 1.05	63.92 ± 3.22	77.30 ± 1.28
	Ours	**91.64 ± 1.43**	**92.66 ± 1.10**	91.25 ± 1.33	**91.95 ± 1.82**	**87.33 ± 1.20**	**94.25 ± 0.90**	**77.44 ± 1.20**	**85.05 ± 1.35**
50%	PVC	74.86 ± 2.04	73.23 ± 1.14	76.33 ± 1.98	74.78 ± 0.71	68.89 ± 2.07	77.50 ± 1.21	53.06 ± 2.91	61.78 ± 1.50
	UEAF	88.38 ± 1.46	84.91 ± 1.67	92.17 ± 1.57	88.13 ± 1.43	80.81 ± 1.41	90.41 ± 1.53	72.53 ± 1.81	79.70 ± 1.82
	GAIN	88.98 ± 1.93	81.68 ± 1.31	**94.44 ± 1.71**	88.06 ± 1.97	80.62 ± 2.34	90.96 ± 1.15	61.83 ± 1.98	71.83 ± 1.66
	Ours	**90.83 ± 1.89**	**89.35 ± 0.81**	92.36 ± 0.94	**90.85 ± 1.72**	**84.22 ± 1.40**	**92.31 ± 1.74**	**73.38 ± 1.26**	**83.50 ± 1.47**

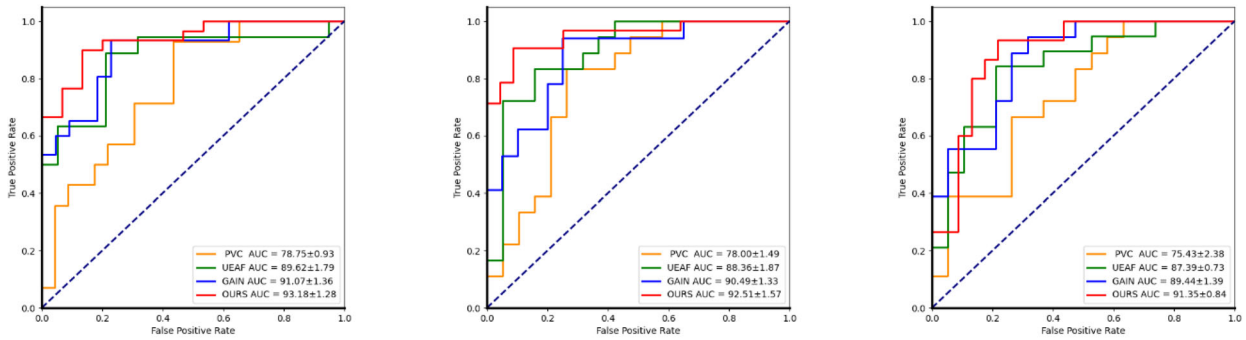
cose utilization and AV45 amyloid values were extracted. At last, the imaging measures on all modalities with 90 ROIs after the remove of cerebellum were adopted as quantitative traits (QTs) in our experiment.

Two kinds of diagnose tasks were performed in our experiments. The first is NC vs. AD and the second is LMCI vs. EMCI. We compared our method with the state-of-the-art methods in above two tasks. The dataset was randomly divided into 90% training set and 10% test set, and ten-fold cross-validation is utilized to adjust the parameters. It is worth noting that the PET modality in training set had varying degrees of missing from 10 to 50%. The diagnostic performance is evaluated by accuracy (ACC), sensitivity (SEN), specificity (SPE) and area under curve (AUC).

3.2 Performance Evaluation

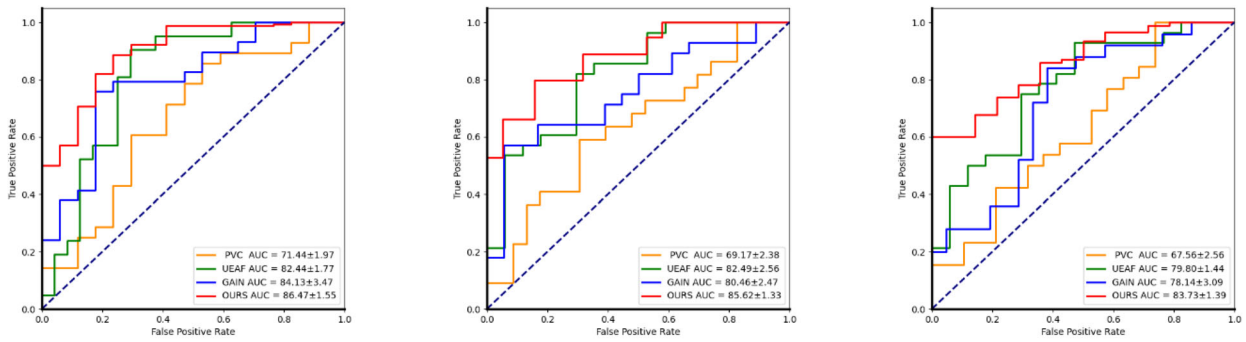
In this section, three methods are selected for comparison, namely partial multi-view clustering (PVC) [24], unified-embedding alignment framework (UEAF) [13] and generative adversarial imputation network (GAIN) [25]. Specifically, PVC builds the latent subspace based on matrix decomposition to obtain the common representation of each sample in subspace.

UEAF expands this work by using Laplacian matrix and manifold learning to build a framework that can complete the missing data according to the subspace and then optimize the common representation through the completed data, which have a mutual promotion. GAIN is a completion method



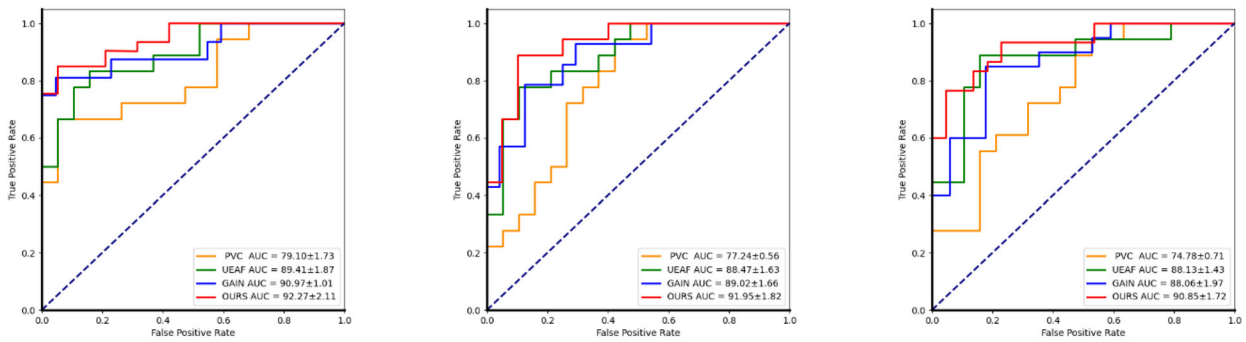
(1)Missing rate 10% (2)Missing rate 30% (3)Missing rate 50%

Fig. 3 ROC curves for different comparative experiments and different missing rates on task AD vs. NC with VBM/FDG. (1)~(3) The cases of missing 10%, 30% and 50%



(1)Missing rate 10% (2)Missing rate 30% (3)Missing rate 50%

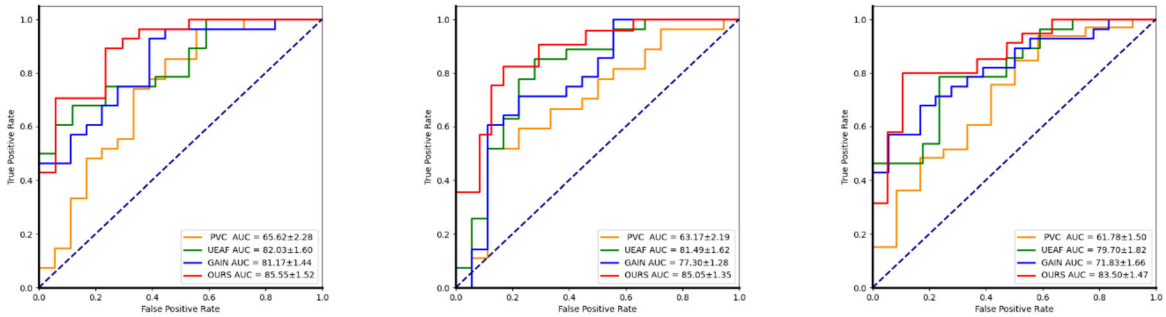
Fig. 4 ROC curves for different comparative experiments and different missing rates on task LMCI vs. EMCI with VBM/FDG. (1)~(3) The cases of missing 10%, 30% and 50%



(1)Missing rate 10% (2)Missing rate 30% (3)Missing rate 50%

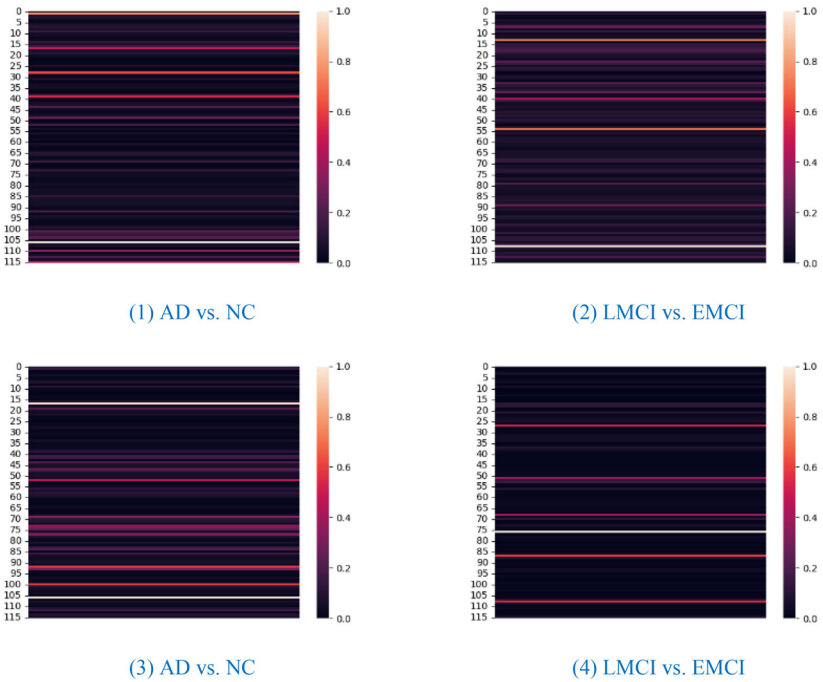
Fig. 5 ROC curves for different comparative experiments and different missing rates on task AD vs. NC with VBM/AV45. (1)~(3) The cases of missing 10%, 30% and 50%

based on generative adversarial network, and in this work, a hint matrix is added to improve the quality of generated data. The diagnosis result of VBM/FDG is shown in Table 2, and VBM/AV45 is shown in Table 3. Correspondingly, we show the ROC curves of all tasks in Figs. 3, 4, 5 and 6. Our method achieves the best results in two tasks with two different PET modalities.



(1)Missing rate 10% (2)Missing rate 30% (3)Missing rate 50%

Fig. 6 ROC curves for different comparative experiment and different missing rates on task LMCI vs. EMCI with VBM/AV45. (1) ~ (3) The cases of missing 10%, 30% and 50%



(1) AD vs. NC (2) LMCI vs. EMCI

(3) AD vs. NC (4) LMCI vs. EMCI

Fig. 7 Attention maps of generator with different modalities. The ordinate represents 116 ROIs of input MRI features. Specifically, (1), (3) on VBM/FDG; (2), (4) on VBM/AV45

4 Discussion

4.1 Analysis on comparative experiments

In the experiment, we compared the classification results of our method with the state-of-the-art incomplete multi-modal methods. The results of the classification tasks are reported in Tables 2 and 3, with the best results in bold. From these tables and figures, we can observe that our proposed method is superior to the compared methods in ACC and AUC. For the reason that PVC only considers the structural information between samples, it cannot learn the common representation well. Manifold learning and Laplacian matrix are adopted by UEAF, which can learn latent representation more effectively and achieve good results. GAIN uses a deep network to process missing values, which can better train the generator according to the hint mechanism and improve the quality of the generated data. Different from these methods, the structure of the feature itself is taken into consideration and real PET features are added into the training procedure. As the rate of missing data increases, the diagnostic accuracy tends to decrease, but our method is still effective. Due to the imbalance of samples in dataset, the SEN is significantly higher than SPE in the task of LMCI vs. EMCI. As we can see in the ROC curves, although our method does not outperform the comparative method in all cases, it still achieves the best AUC in all tasks. Extensive evidence proved the effectiveness of our proposed method.

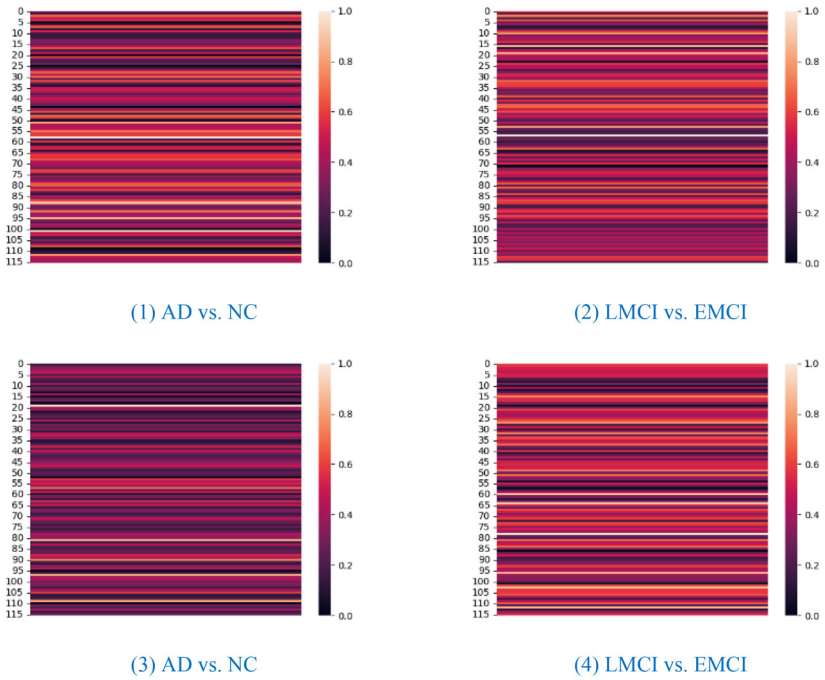


Fig. 8 Attention maps of discriminator with different modalities. The ordinate represents 116 ROIs of input PET features. Specifically, (1), (3) on VBM/FDG; (2), (4) on VBM/AV45

Table 4 Experiment results on four tasks without attention mechanism (VBM/FDG)

Task	Missing rate	ACC(%)	SEN(%)	SPE(%)	AUC(%)
AD vs. NC	0.1	91.54	92.33	93.01	92.28
	0.3	91.73	90.18	92.01	91.79
	0.5	90.36	89.12	91.98	90.53
LMCI vs. EMCI	0.1	87.02	92.99	78.64	85.49
	0.3	86.13	91.98	77.37	85.16
	0.5	82.34	92.32	72.88	81.06

Table 5 Experiment results on four tasks without attention mechanism (VBM/AV45)

Task	Missing rate	ACC(%)	SEN(%)	SPE(%)	AUC(%)
AD vs. NC	0.1	92.42	92.04	92.15	92.23
	0.3	91.22	91.47	91.34	91.51
	0.5	89.65	89.72	91.84	90.10
LMCI vs. EMCI	0.1	86.69	91.98	76.61	84.01
	0.3	85.93	91.04	76.13	83.56
	0.5	81.44	90.06	71.12	79.93

4.2 Analysis on attention mechanism

We adopted attention layer in the generator and discriminator, respectively. Inspired by the human attention mechanism, attention can assign different weights to each feature according to its importance in the current task, so as to emphasize important features.

As shown in Figs. 7 and 8, we extracted attention maps of the generator and discriminator for four different diagnostic tasks. As shown in Fig. 7, the input of discriminator is MRI features, which has a large difference in importance between different ROIs and relatively large variance. In Fig. 8, the difference between PET features is small, with small variance.

In addition, we removed the attention layer and conducted ablation experiments for all diagnose tasks. As shown in Table 4, it can be seen that in almost all cases, the precision of the model without attention layer is reduced to some extent, which can fully prove the importance of attention mechanism.

5 Conclusion

In this paper, we presents a novel pairwise feature-based generative adversarial networks for feature synthesis and diagnosis with incomplete multi-modal features extracted from brain images. Specifically, we designed a framework which can generate corresponding missing PET features by available MRI features. Besides, a linear attention mechanism is adopted to retain the structure of the features and highlight the disease-relevant parts. Experiments on 913-ADNI demonstrate that our proposed method generates discriminative features and achieves promising performance in the tasks of both AD vs. NC and LMCI vs. EMCI.

Authors' contributions QZ and DZ received the idea, identified and coordinated the study. HY partially designed the methods. QZ contributed to the method. HY contributed to programing generated results, and wrote the initial draft. QZ, YY, YJ and DZ revised the manuscript. All authors read, revised and approved the final manuscript.

Funding This work was supported in part by National Natural Science Foundation of China (Nos. 62076129, 61501230, 61732006, 61876082 and 61861130366), National Science and Technology Major Project (No. 2018ZX10201002), and the National Key R&D Program of China (Grant Nos.: 2018YFC2001600, 2018YFC2001602).

Declarations

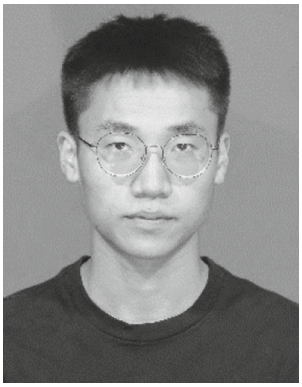
Conflicts of interest We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Consent for publication All the authors listed have approved the manuscript that is enclosed.

References

1. Weller, J., Budson, A.: Current understanding of Alzheimer's disease diagnosis and treatment. F1000Res **7**, 1161 (2018)
2. Kirova, A.-M., Bays, R.B., Lagalwar, S.: Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. Biomed. Res. Int. **2015**, 1–9 (2015)
3. Guarino, A., Favieri, F., Boncompagni, I., Agostini, F., Cantone, M., Casagrande, M.: Executive functions in Alzheimer disease: a systematic review. Front. Aging Neurosci. **10**, 437 (2019)
4. Kitamura, Y., Usami, R., Ichihara, S., Kida, H., Satoh, M., Tomimoto, H., Murata, M., Oikawa, S.: Plasma protein profiling for potential biomarkers in the early diagnosis of Alzheimer's disease. Neurol. Res. **39**(3), 231–238 (2017)
5. Cummings, J., Lee, G., Ritter, A., Sabbagh, M., Zhong, K.: "Alzheimer's disease drug development pipeline: 2019", Alzheimer's \& Dement. Transl. Res. Clin. Interv. **5**, 272–293 (2019)
6. Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., Alzheimer's Disease Neuroimaging Initiative et al.: Multimodal classification of Alzheimer's disease and mild cognitive impairment, *Neuroimage*, **55**(3), 856–867, (2011)
7. Lin, E., Lin, C.-H., Lane, H.-Y.: Deep learning with neuroimaging and genomics in Alzheimer's disease. Int. J. Mol. Sci. **22**(15), 7911 (2021)
8. Zhang, D., Shen, D., Alzheimer's Disease Neuroimaging Initiative et al.: Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. Neuroimage, **59**(2): 895–907, (2012).
9. Ieracitano, C., Mammone, N., Hussain, A., Morabito, F.C.: A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. Neural Netw. **123**, 176–190 (2020)
10. Zhou, T., Thung, K.-H., Liu, M., Shi, F., Zhang, C., Shen, D.: Multi-modal latent space inducing ensemble SVM classifier for early dementia diagnosis with neuroimaging data. Med. Image Anal. **60**, 101630 (2020)
11. Basavegowda, H.S., Dagnew, G.: Deep learning approach for microarray cancer data classification. CAAI Trans. Intell. Technol. **5**(1), 22–33 (2020)
12. Shao, W., He, L., Philip, S. Y.: Multiple incomplete views clustering via weighted nonnegative matrix factorization with $$ L_ {2, 1} $$ regularization, In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 318–334 (2015)
13. Wen, J., Zhang, Z., Xu, Y., Zhang, B., Fei, L., Liu, H.: Unified embedding alignment with missing views inferring for incomplete multi-view clustering. Proc. AAAI Conf. Artif. Intell. **33**(01), 5393–5400 (2019)
14. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks, (2014) arXiv Prepr. arXiv: 1406.2661.
15. Zhu, C., Yan, W., Cai, X., Liu, S., Li, T.H., Li, G.: Neural saliency algorithm guide bi-directional visual perception style transfer. CAAI Trans. Intell. Technol. **5**(1), 1–8 (2020)
16. Pan, Y., Liu, M., Lian, C., Zhou, T., Xia, Y., Shen,D.: Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis, In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 455–463 (2018).
17. Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D.: Disease-image specific generative adversarial network for brain disease diagnosis with incomplete multi-modal neuroimages, In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 137–145 (2019).
18. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training gans. Adv. Neural Inf. Process. Syst. **29**, 2234–2242 (2016)
19. Rensink, R.A.: The dynamic representation of scenes. Vis. cogn. **7**(1–3), 17–42 (2000)
20. Yang, J., Xing, D., Hu, Z., Yao, T.: A two-branch network with pyramid-based local and spatial attention global feature learning for vehicle re-identification. CAAI Trans. Intell. Technol. **6**(1), 46–54 (2021)
21. Mescheder, L., Geiger, A., Nowozin, S.: Which training methods for GANs do actually converge? In: International Conference on Machine Learning, pp. 3481–3490 (2018).
22. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage **15**(1), 273–289 (2002)
23. Ashburner, J., Friston, K.J.: Voxel-based morphometry—the methods. Neuroimage **11**(6), 805–821 (2000)
24. Li, S.-Y., Jiang, Y., Zhou, Z.-H.: Partial multi-view clustering, In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 28(1) (2014)

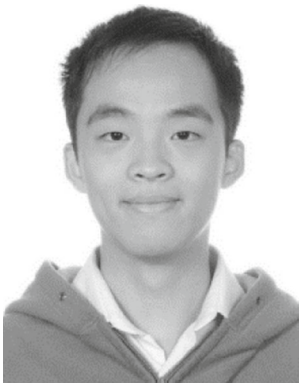
25. Yoon, J., Jordon, J., Van Der Schaar, M.: GAIN: Missing data imputation using generative adversarial nets, In: 35th International Conference Machine Learning. ICML 2018, vol. 13, pp. 9042–9051, (2018).



Haizhou Ye received the B.S degree from Nanjing University of Posts and Telecommunications, Jiangsu, China, in 2019. He is current working toward the M.S. degree in software engineering from Nanjing University of Aeronautics and Astronautics. His current research interests include machine learning, pattern recognition, medical image processing.

Qi Zhu received his B.S. degree, M.S. degree, and Ph.D. degree from Harbin Institute of Technology in 2007, 2010 and 2014, respectively. Recently, he is an associate professor at College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. His interests include pattern recognition, feature extraction and medical image analysis.

Yuan Yao received the B.S degree from Nanjing Forestry University, Jiangsu, China, in 2019. He is current working toward the M.S. degree in software engineering from Nanjing University of Aeronautics and Astronautics. His current research interests include machine learning, pattern recognition, medical image processing.



Yichao Jin received the B.E and M.E degree from Nanjing University of Posts and Telecommunications (NUPT), China, in 2008 and 2011, respectively. He is currently pursuing his Ph.D degree in the School of Computing Engineering, Nanyang Technological University (NTU), Singapore. His research interests are cloud computing and content delivery networks. He received IEEE Globecom Best Paper Award in 2013.

Daoqiang Zhang was born in 1978 and received his B.S. degree and Ph.D. degree in Computer Science from Nanjing University of Aeronautics and Astronautics in 1999 and 2004, respectively. And he is a Professor and a PhD supervisor in Nanjing University of Aeronautics and Astronautics at present. His research interests include machine learning, pattern recognition, data mining, medical imaging analysis, etc. In these areas, he has published over 100 scientific articles in refereed international journals such as NeuroImage, Pattern Recognition, Artificial Intelligence in Medicine, IEEE Trans. Neural Networks; and conference proceedings such as IJCAI, AAAI, SDM, and ICDM. He is a member of the Machine Learning Society of the Chinese Association of Artificial Intelligence (CAAI), and the Artificial Intelligence & Pattern Recognition Society of the China Computer Federation (CCF).