

RESEARCH ARTICLE

A multi-view learning approach with diffusion model to synthesize FDG PET from MRI T1WI for diagnosis of Alzheimer's disease

Ke Chen^{1,2} | Ying Weng^{1,2} | Yueqin Huang³ | Yiming Zhang¹ | Tom Dening⁴ | Akram A. Hosseini⁵ | Weizhong Xiao⁶

¹School of Computer Science, University of Nottingham Ningbo China, Ningbo, China

²School of Computer Science, University of Nottingham, Nottingham, UK

³Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing, China

⁴School of Medicine, University of Nottingham, Nottingham, UK

⁵Neurology Department, Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham, UK

⁶Neurology Department, Peking University Third Hospitals, Beijing, China

Correspondence

Ying Weng, School of Computer Science, University of Nottingham Ningbo China, 199 Taikang East Road, Yinzhou, Ningbo, Zhejiang 315100, China; School of Computer Science, University of Nottingham, Wollaton Road, Nottingham, Nottinghamshire, NG8 1BB, UK. Email: ying.weng@nottingham.edu.cn; ying.weng@nottingham.ac.uk

Funding information

Ningbo Major Science & Technology Program, Grant/Award Number: 2022Z126; Nottingham Short-term RKE Project, Grant/Award Number: I05220500009; Medical Research Council, UK, Grant/Award Number: MR/T005580/1; National Institute of Health/NIA, USA, Grant/Award Number: 1R56AG074467-01

Abstract

INTRODUCTION: This study presents a novel multi-view learning approach for machine learning (ML)-based Alzheimer's disease (AD) diagnosis.

METHODS: A diffusion model is proposed to synthesize the fluorodeoxyglucose positron emission tomography (FDG PET) view from the magnetic resonance imaging T1 weighted imaging (MRI T1WI) view and incorporate two synthesis strategies: one-way synthesis and two-way synthesis. To assess the utility of the synthesized views, we use multilayer perceptron (MLP)-based classifiers with various combinations of the views.

RESULTS: The two-way synthesis achieves state-of-the-art performance with a structural similarity index measure (SSIM) at 0.9380 and a peak-signal-to-noise ratio (PSNR) at 26.47. The one-way synthesis achieves an SSIM at 0.9282 and a PSNR at 23.83. Both synthesized FDG PET views have shown their effectiveness in improving diagnostic accuracy.

DISCUSSION: This work supports the notion that ML-based cross-domain data synthesis can be a useful approach to improve AD diagnosis by providing additional synthesized disease-related views for multi-view learning.

KEYWORDS

Alzheimer's disease, data synthesis, diffusion model, fluorodeoxyglucose positron emission tomography, magnetic resonance imaging T1 weighted imaging, multi-view learning

Highlights

- We propose a diffusion model with two strategies to synthesize fluorodeoxyglucose positron emission tomography (FDG PET) from magnetic resonance imaging T1 weighted imaging (MRI T1WI).
- We raise multi-view learning with MRI T1WI and synthesized FDG PET for Alzheimer's disease (AD) diagnosis.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

- We provide a comprehensive experimental comparison for the synthesized FDG PET view.
- The feasibility of synthesized FDG PET view in AD diagnosis is validated with various experiments.
- We demonstrate the ability of synthesized FDG PET to enhance the performance of machine learning-based AD diagnosis.

1 | BACKGROUND

In practical scenarios, a multitude of information manifests in diverse formats as heterogeneous data sources describe an entity. In the medical field, data acquisition frequently encompasses a variety of domains or modalities. For instance, magnetic resonance imaging (MRI) and positron emission tomography (PET) can be categorically conceptualized as multi-view data.¹

Multi-view learning can exploit complementary information of multiple views of data or information to improve the performance of the machine learning (ML) model.^{2–4} The medical data in the artificial intelligence (AI) domain is limited, making it challenging to generalize. By exploiting multi-view learning, different medical data modalities can be treated as a view, and many researchers have found that it can improve model performance and generalization.⁵ In most medical applications, using multi-modality medical images can produce better results than using only one modality medical image, because different modalities of medical images reflect different pathologies and also provide different views for specific diseases.^{6,7}

The World Alzheimer Report 2018⁸ estimated that the number of people suffering from dementia was ≈ 50 million in 2018 and will more than triple to 152 million by 2050. Alzheimer's disease (AD) is the most common cause of dementia, accounting for 60% to 80% of dementia cases, and is one of the main causes of physical and mental health disorders especially among older people.⁹ Preliminary cognitive tests using multiple scales, such as the Mini-Mental State Examination (MMSE)¹⁰ and Montreal Cognitive Assessment (MoCA),¹¹ are most commonly used for AD diagnosis. However, these cognitive tests are affected by education level, language, and so forth, which implies they are insufficiently objective as a means of diagnosis. The neuroimaging examination is a more objective and precise test that can be used for the diagnosis or early diagnosis of AD as well as to distinguish it from other types of dementia.

Structural MRI (sMRI) is the most common modality for AD diagnosis due to its low cost of acquisition, non-invasiveness, and accessibility. Various ML methods using sMRI have been proposed for early diagnosis and prediction of AD as well as for exploring brain structural pathological patterns.^{12–20} PET is a relatively new modality of neuroimaging with fewer data existing due to its high expense and limited access to scanners compared to the sMRI, though the usage of PET is steadily increasing worldwide. Recent studies have shown the effectiveness of using PET with the assistance of ML techniques for the early

diagnosis and prediction of AD.^{12,20–22} Multi-view learning approaches can also assist in AD diagnosis and prognosis.^{21,23}

The multi-view learning of sMRI and PET has also been recognized as a useful method for the early diagnosis and prediction of AD.^{24–26} The combination of sMRI and PET scans can provide both structural and functional information related to AD, therefore enabling ML methods to enhance the effectiveness and robustness of the diagnosis.

The motivation behind this study lies in addressing the limitations posed by the availability and cost of PET scans, which are valuable yet often inaccessible for many patients. Given the high cost and limited accessibility of PET imaging, there is an imminent need to develop alternative methods that can provide similar diagnostic benefits. We explore the possibility of synthesizing fluorodeoxyglucose PET (FDG PET) images from sMRI data, which is more widely available and less costly. FDG PET can reflect neuronal activity by measuring cerebral glucose metabolism and hence can offer valuable insights into metabolic changes, even beyond obvious anatomical brain changes such as brain atrophy, which makes it a powerful tool for early diagnosis and monitoring of AD progression. The early diagnosis of AD is crucial for identifying the risk of AD progression and for earlier intervention. Furthermore, the use of FDG PET combined with other imaging modalities, such as sMRI, can provide a comprehensive view of both functional and structural changes in the brain, hence enriching the understanding of disease mechanisms and progression. By leveraging multi-view learning and advanced image synthesis techniques, we aim to enhance the accuracy and accessibility of AD diagnosis. Recent studies^{27–33} have also shown the feasibility of synthesizing FDG PET scans from sMRI scans with various ML-based methods.

This study is clinically significant as it proposes a novel approach that could improve the early detection and diagnosis of AD. By synthesizing FDG PET images from sMRI data, specifically MRI T1 weighted imaging (MRI T1WI) scans, we provide a cost-effective, accessible, and non-invasive method to enhance diagnostic accuracy. This approach is beneficial for settings with limited access to FDG PET imaging, and will ultimately make contributions to better patient outcomes by enabling earlier and more accurate diagnosis. The overall pipeline is shown in Figure 1. The raw sMRI scans are first preprocessed to derive the corresponding skull removed and standard space registered scans. Then the preprocessed scans are fed into the diffusion model to generate the synthesized PET scans. Last, the preprocessed scans and the synthesized PET scans are jointly fed into the classifier to make a prediction of AD progression stage.

2 | METHOD

2.1 | Datasets

Data used in the preparation of this work were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI is to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD.

In this research, we have downloaded the paired MRI T1WI and FDG PET scans for the task of synthesizing FDG PET scans from MRI T1WI scans. The pairing procedure is based on the metadata of the downloaded images and the terms for pairing are the same subject ID and the interval of the date of obtaining the image < 30 days. The demographic analysis of the dataset with MRI PET pairs is shown in Table 1.

2.2 | Data preprocessing

The data in the ADNI dataset were collected from different sources, which introduces some differences existing in voxel size, orientation, and so forth. To make the data representation consistent, data preprocessing was carried out using the FMRIB Software Library (FSL).³⁴ For both MRI T1WI scans and FDG PET scans, we removed non-brain tissue at the first step to reduce the bias in the following data preprocessing steps. The images were then registered to the standard Montreal Neurological Institute 152 (MNI152) space to help further data preprocessing and experiments. Moreover, to extract more refined feature representation for AD diagnosis, brain region segmentation was applied to the registered images to obtain white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF).

RESEARCH IN CONTEXT

- 1. Systematic review:** Literature on machine learning (ML)-based Alzheimer's disease (AD) diagnosis and data synthesis in the medical field has been investigated. Many studies support the effectiveness of using neuroimaging for AD diagnosis as well as involving multi-view data. Some studies also show the potential to use synthesized data in the medical field.
- 2. Interpretation:** We propose a novel multi-view learning approach for the diagnosis of AD using ML, which involves data synthesis through a diffusion model. Our proposed two-way diffusion model generates high-quality synthesized data, and extensive experiments show the effectiveness of this synthesized data in ML-based AD diagnosis.
- 3. Future directions:** The results of our study demonstrate the potential of using synthesized data to augment the performance of computer-assisted diagnosis (CAD). Expanding the synthesis of diverse data modalities, particularly those that are costly yet valuable, would greatly benefit CAD. Further advancements in high-quality data synthesis methods are imperative for enhancing the diagnostic precision of CAD.

2.3 | Diffusion model

Diffusion models, also known as denoising diffusion probabilistic models (DDPMs),³⁵ are latent variable models that define a diffusion process and a reversed diffusion process, also known as the sample process. In the diffusion process, Gaussian noise is continuously added to the input data to obtain its latent structure in the Gaussian

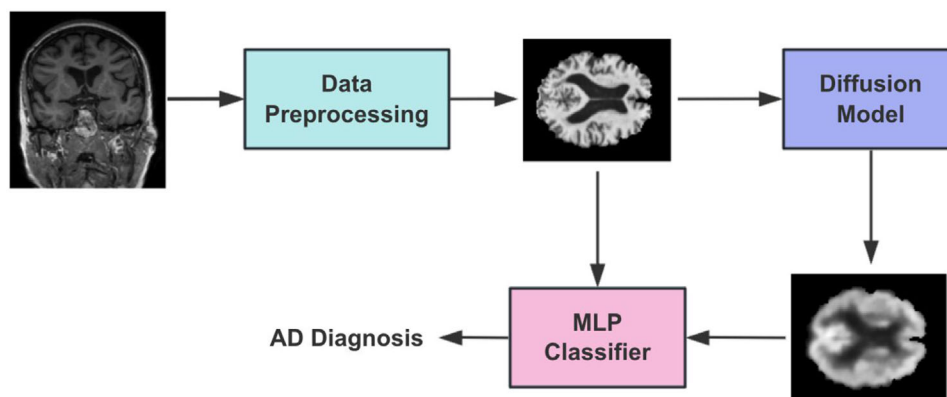


FIGURE 1 Pipeline of the framework. First, raw MRI T1WI scans are preprocessed to derive standardized preprocessed scans. Then the preprocessed MRI T1WI scans are fed into the diffusion model to synthesize the FDG PET. Furthermore, both preprocessed MRI T1WI scans and synthesized FDG PET scans are jointly fed into the classifier to provide an AD diagnosis. AD, Alzheimer's disease; FDG PET, fluorodeoxyglucose positron emission tomography; MLP, multilayer perceptron; MRI T1WI, magnetic resonance imaging T1 weighted imaging.

TABLE 1 Demographic analysis of paired dataset with MRI T1WI scans and FDG PET scans as pairs.

Age	76.59 ± 6.65
Sex	
Male	208
Female	124
Subject	332
Scan	1035

Abbreviations: FDG PET, fluorodeoxyglucose positron emission tomography; MRI T1WI, magnetic resonance imaging T1 weighted imaging.

distribution space while the noisy data are gradually denoised to obtain the expected output in the input data distribution space in the sample process. Figure 2 shows an example of diffusion progress in the second row including a diffusion process from right to left that diffuses the input data by gradually adding noise and a sample process from left to right that denoises the noisy input to reverse generate the desired noise-free image. The goal of a diffusion model is to learn the way of modeling the data into the latent Gaussian distribution space, equivalently the diffusion process.

In this research, the aim is to generate FDG PET scans from MRI T1WI scans, which implies that the MRI T1WI scans should be the input and the FDG PET scans should be the denoised output image. However, instead of directly learning the way to map the input to the latent Gaussian distribution space for both modalities of images with a uniform model, we use the MRI T1WI scans as condition images that are used to guide the diffusion and denoising progress as shown in Figure 2, where the first channel of the input is always MRI T1WI and the noising/denoising is only applied on the second channel. Specifically, MRI T1WI scans are used in an additional channel apart from their corresponding FDG PET scan in the diffusion process to train the model for the learning of space mapping from FDG PET scans to Gaussian noises. In the sampling process, MRI T1WI scans are also used in an additional

channel apart from noisy data to guide the denoising steps to recover their corresponding FDG PET scans.

With the FDG PET channel denoted as x_{PET} and the MRI T1WI channel as x_{MRI} , the input image is denoted as X , which contains two channels (x_{PET}, x_{MRI}). The general idea is gradually adding a small amount of noise to the input data to generate a series of noisy data $\{X_0, X_1, \dots, X_T\}$ for many time steps T . However, in our design, the noise is only added on the x_{PET} at each time step while the x_{MRI} remains the same. The noise level of X is then steadily increased from 0 to T . A U-NET Diff is used to predict the noise used for generating X_{t-1} from X_t . During the training, with the ground truth for X_0 and a Gaussian noise ϵ , an X_t can be derived as following:

$$X_t = \sqrt{\alpha_t} X_0 + \sqrt{1 - \alpha_t} \epsilon \quad (1)$$

where $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$, β_1 to β_t is a set of predefined variances.

The U-NET model then takes the X_t as input to predict its corresponding noise and is trained using an MSE loss with ϵ .

$$\text{MSE (Diff)} = \text{mean} \left(\epsilon - \text{Diff} \left(\sqrt{\alpha_t} X_0 + \sqrt{1 - \alpha_t} \epsilon \right) \right)^2 \quad (2)$$

Meanwhile, because the parts of the image other than brain tissue are not used for the FDG PET synthesis, we restrict the loss only to the brain tissue by applying a brain mask $Mask_{brain}$ as follows:

$$\text{Loss (Diff)} = \text{mean} \left(Mask_{brain} \left(\epsilon - \text{Diff} \left(\sqrt{\alpha_t} X_0 + \sqrt{1 - \alpha_t} \epsilon \right) \right) \right)^2 \quad (3)$$

During the sample process used for generating the FDG PET, according to the formulation for DDPMs, X_{t-1} can be derived from X_t as follows:

$$x_{t-1} = \sqrt{\bar{\alpha}_{t-1}} \left(\frac{x_t - \sqrt{1 - \bar{\alpha}_t} \text{Diff}(X_t)}{\sqrt{\alpha_t}} \right) + \sqrt{1 - \bar{\alpha}_{t-1} - \sigma_t^2} \text{Diff}(X_t) + \sigma_t \epsilon \quad (4)$$

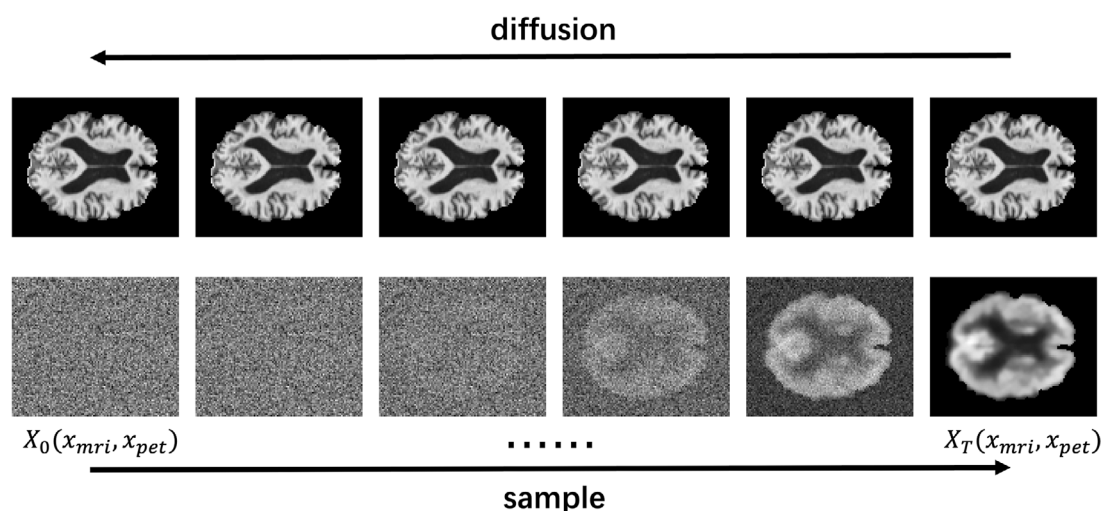


FIGURE 2 Illustration for diffusion and sample process.

where $\sigma_t = \sqrt{\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}} \sqrt{1-\frac{\bar{\alpha}_t}{\bar{\alpha}_{t-1}}}$, which is used to control the degree of adding stochastic noise at each step. Unlike a normal image generating task that requires diversity in the generated images, it is expected to be deterministic in FDG PET synthesized from MRI T1WI. Therefore, we set the σ_t to 0 as described in denoising diffusion implicit models (DDIMs)³⁶ to produce a deterministic sampling. Meanwhile, this enables speeding up sample progress by rescaling the total time steps T to a smaller value.

The backbone of the diffusion model is also a U-NET-style model with several steps of down-sampling modules to extract high-level features and several steps of up-sampling modules to reconstruct the high-resolution output. A skip connection is also included to enhance the model's output by merging lower level but higher resolution features into the expanding path, equivalent to the continuous up-sampling process. Moreover, special embeddings are added to the network as additional inputs, namely class embedding and time step embedding. In our work, time step embedding is a cos-sin embedding defined as following:

$$\text{seq} = [e^{-\log(\text{mp}) \times i}] \times \text{timestep} \quad (5)$$

$$\text{emb}_{\text{timestep}} = \text{concat}([\sin(\text{seq}), \cos(\text{seq})]) \quad (6)$$

where mp denotes for maximum period which is set to 10,000, i ranges from 0 to 255 in this work, and concat denotes the concatenation operation. To summarize, the time step embedding in this work has a dimension of 512 where the first 256 elements are related to the sin function and the remaining are related to the cos function. An example visualization of time step embedding is shown in Figure 3B with blue referring to the first 256 elements and green referring to the remaining 256 elements. The order of the elements is from right to left.

After generating the sin-cos time step embedding based on a specific time step, the time step embedding would further be fed into fully connected networks consisting of linear layers and non-linear activation layers to deeply embed the time step:

$$\text{emb}_{\text{timestep}}^{\text{deep}} = \text{fc}(\text{emb}_{\text{timestep}}) \quad (7)$$

where fc denotes the fully connected networks for deeply embedding the time step.

Class embedding is used for identifying which modality of data is going to be sampled or diffused. In this study, the data to be used could be either MRI T1WI or FDG PET scans. The class embedding is defined as a set of learnable parameters which has the same dimension as the time step embedding. In other words, both of the classes, MRI T1WI and FDG PET, have their corresponding learnable class embedding variable with a length of 512.

The final embedding fed into Down and Up modules as shown in Figure 3A is a combination of time step embedding and class embedding as follows:

$$\text{emb}_{\text{timestep}}^{\text{deep}} = \text{emb}_{\text{timestep}}^{\text{deep}} + \text{emb}_{\text{class}} \quad (8)$$

Therefore, the final embedding allows the model to receive information on which time step it is and refines the prediction with the assistance of time step embedding. Moreover, the final embedding also determines the model's output direction because it contains class embedding, which is an output prompt.

Moreover, two approaches of synthesizing FDG PET from MRI T1WI using the above-described diffusion model are proposed, namely one-way synthesis and two-way synthesis as shown in Figure 4. The one-way synthesis approach starts with randomly generated Gaussian noise together with an MRI T1WI scan as the reference image. It only performs a sample process with an FDG PET-related class embedding, which tells the model to denoise the noised input to an FDG PET. The two-way synthesis approach starts with an MRI T1WI scan together with the same MRI T1WI scan as reference image. It consists of two steps, namely diffusion and sample. In the diffusion process, an MRI T1WI-related class embedding tells the model to continuously add noise to the initial MRI T1WI to generate Gaussian-distributed noise corresponding to the input, from which the input MRI T1WI can be sampled back. In the sample process, the model takes the noise derived from the diffusion process and continuously denoises the noise with an FDG PET-related class embedding to generate the synthesized FDG PET. The main difference between the one-way and two-way synthesis approaches is whether to introduce randomly generated noise for FDG PET synthesis. Due to no random noise introduced in the two-way synthesis approach, the synthesized results would be deterministic while the results of the one-way synthesis approach might change each time for model inferencing.

2.4 | Classifier

To evaluate the effectiveness of the synthesized FDG PET scans in ML-assisted AD diagnosis, we have designed a multilayer perceptron (MLP)-based classifier, which consists of several MLPs with each corresponding to one specific type of feature vector extracted from brain segments, namely WM, GM, and CSF, and returning the probability of each patient category (normal control [NC]/MCI/AD). The results from each MLP are then averaged to produce the final prediction. For a single modality, the final prediction is an average of three branches' outputs with WM, GM, and CSF feature vectors accordingly. With multiple modalities, the prediction scores are further averaged to obtain a prediction score with multi-modality data as joint input.

The input of the MLP-based classifier is designed for the brain atlases-based features extracted from different modalities with different brain segments, that is, GM, WM, and CSF, using AAL3³⁷ atlases. As for MRI T1WI, the feature extraction happens by counting non-zero voxels in each brain sub-region defined in the selected brain atlases as MRI T1WI mostly reflects the degree of structural atrophy of the brain in AD diagnosis. However, for the FDG PET, feature extraction happens by summing the values in each brain sub-region as the power reflects the degree of neuron activity, which is an important factor in the diagnosis of neurodegenerative diseases such as AD.

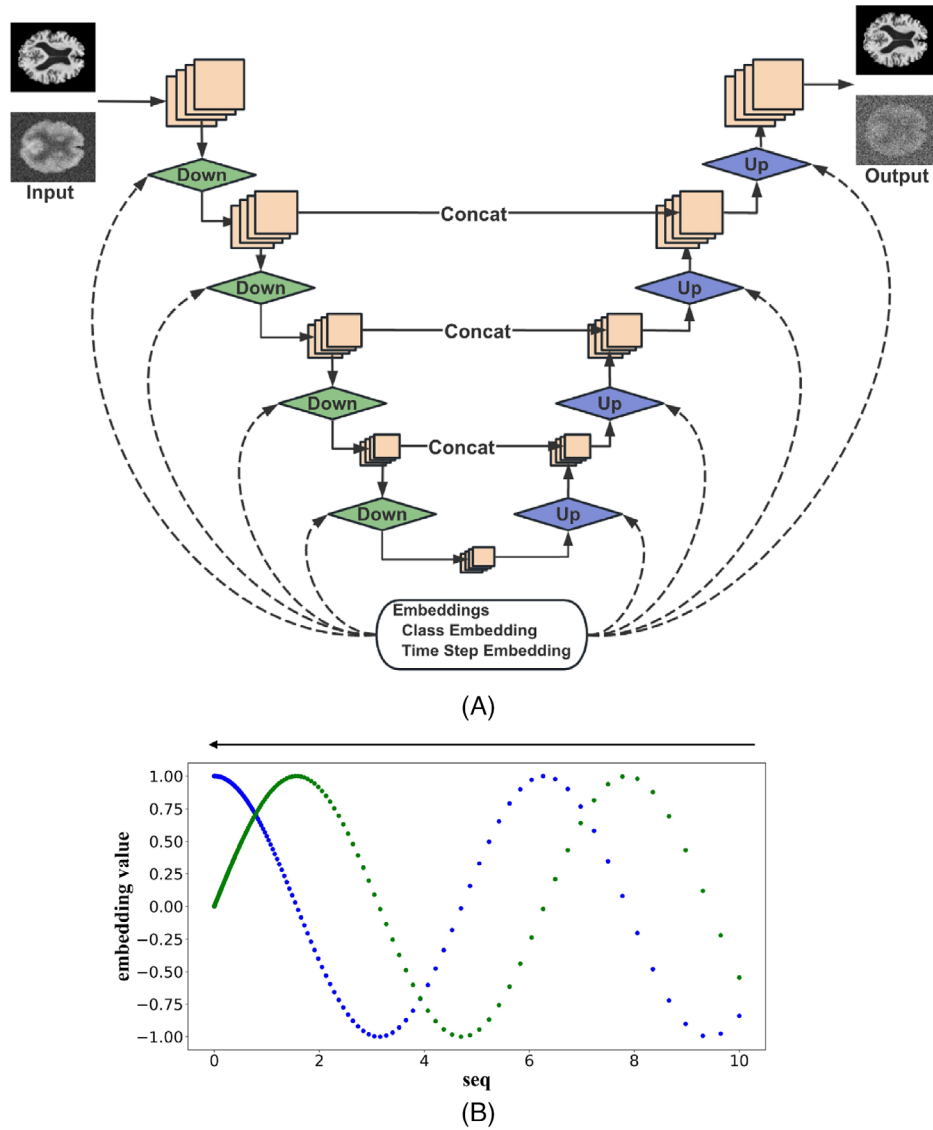


FIGURE 3 A, Structure of the diffusion model. B, An example of time step embedding at the 10th time step. The blue points denote the first 256 elements of the time step embedding related to the sin function and the green points denote the remaining 256 elements of the time step embedding related to the cos function. The y axis denotes the embedding value and the x axis denotes the *seq* variable in Equation (5). The order of the elements arranged in the time step embedding is from right to left as the arrow shows.

3 | RESULTS

For the evaluation of FDG PET scans synthesized from MRI T1WI scans, five criteria metrics, namely mean absolute error (MAE), mean squared error (MSE), zero-normalized cross correlation (ZNCC), structural similarity index measure (SSIM), and peak signal-to-noise rate (PSNR) scores, are used to evaluate the performance of the synthetic results. While computing the criteria metrics, the ground truth FDG PET scans and the synthesized images are both normalized to 0 to 1 to produce a standard comparison. The equations of criteria metrics are shown as follows.

$$MAE(X, Y) = \frac{\sum_{i=1}^{\xi} |y_i - x_i|}{\xi} \quad (9)$$

$$MSE(X, Y) = \frac{\sum_{i=1}^{\xi} (y_i - x_i)^2}{\xi} \quad (10)$$

$$NCC(X, Y) = \frac{1}{\xi} \frac{1}{\sigma_X \sigma_Y} \sum_{i=1}^{\xi} (x_i - \mu_X)(y_i - \mu_Y) \quad (11)$$

$$SSIM(X, Y) = \frac{(2\mu_X \mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \quad (12)$$

$$PSNR(X, Y) = 10 \log_{10} \left(\frac{MAX_{X,Y}^2}{MSE} \right) \quad (13)$$

where $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$, and $c_3 = c_2 / 2$ are used to stabilize the division with weak denominator and k_1 is set to 0.01, k_2 is set to 0.03. Meanwhile, L denotes the range of voxel values, μ denotes the

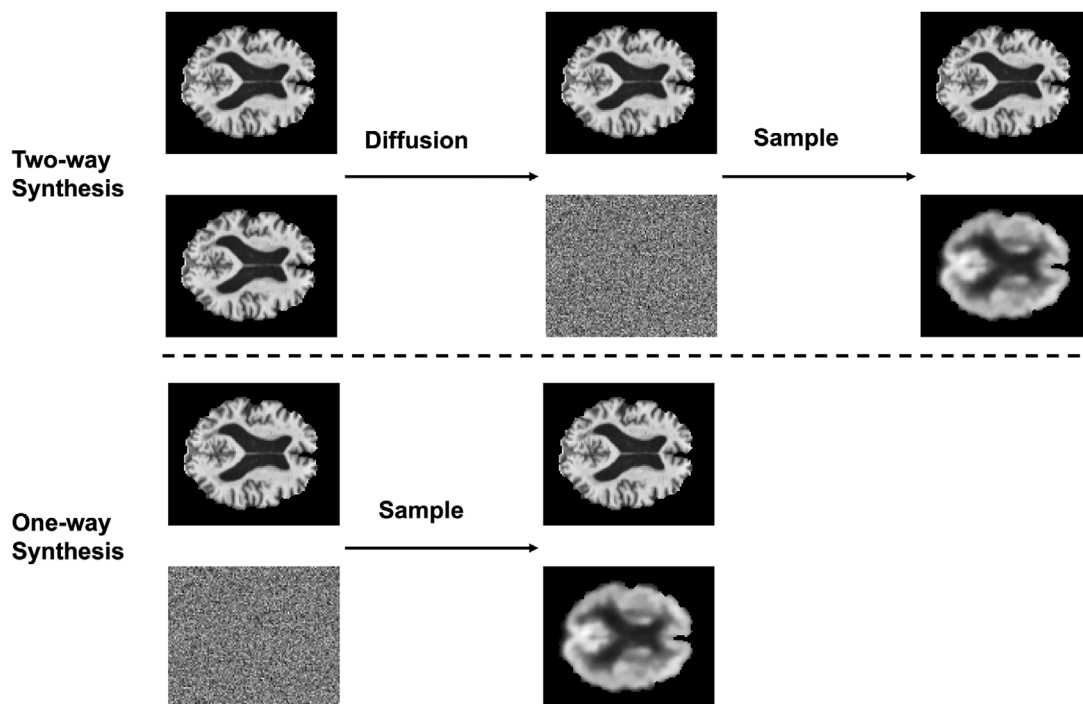


FIGURE 4 Illustration of the two approaches of synthesis, namely one-way synthesis and two-way synthesis.

mean value and σ denotes the standard deviation, and σ_{XY} denotes the covariance of X , Y .

3.1 | Experiments and results analysis for diffusion model

For the diffusion model, we have also conducted a five-fold cross-validation on the dataset. Moreover, instead of normalizing the image values to $0 \sim 1$, we have normalized the value of the input into the range $-1 \sim 1$ with the formula:

$$\text{Normalize}_{-1 \sim 1}(X) = \frac{X - \min(X)}{\max(X) - \min(X)} \times 2 - 1 \quad (14)$$

As aforementioned in section 2.3, the class embedding determines the model's output of whether to generate MRI T1WI-related noises or FDG PET-related noises. Therefore, with the same trained model, we have designed two experiments for FDG PET synthesis. One experiment is the one-way synthesis, which applies the sample process to the randomly generated noise data with MRI T1WI as the conditional image to obtain the desired output of synthesized FDG PET. The other experiment is the two-way synthesis, which applies the diffusion process to the input MRI T1WI data first to generate the noise data and then applies the sample process to the generated noise data to obtain the desired output of synthesized FDG PET. An illustration of the two experiments is shown in Figure 4.

The experimental results are shown in Table 2 with two experiments denoted as "Diffusion Two-way" and "Diffusion One-way." The two-way experiment achieves an MAE score of 0.0187, an MSE score of

0.0024, a ZNCC score of 0.9823, an SSIM score of 0.9380, and a PSNR score of 26.47. The one-way experiment has a slightly better performance than that of the two-way experiment except for the SSIM score. It achieves an MAE score of 0.0274, an MSE score of 0.0044, a ZNCC score of 0.9814, an SSIM score of 0.9282, and a PSNR score of 23.83.

A visual comparison of the three synthetic results from a randomly selected sample is shown in Figure 5

3.2 | Experiments and results analysis for classifier

To evaluate the effectiveness of synthesis results, apart from comparing with the ground truth images using various criterion metrics, we have also designed and trained MLP-based classifiers using different data modalities as input. Specifically, we experiment on a single modality including MRI T1WI, ground truth FDG PET, and synthesized FDG PET, and experiment on different combinations of joint input involving MRI T1WI.

The results of classifier performance for different types of input are shown in Table 3. For the classification of NC/MCI/AD, with MRI T1WI scans as input, the MLP classifier achieves an accuracy of 80.23%. With FDG PET as input, the MLP classifier achieves an accuracy of 81.31%, which outperforms the input setting using MRI T1WI only. With MRI T1WI and FDG PET as joint input, the MLP classifier also outperforms that of using MRI T1WI as input with an accuracy of 81.31%. For the two types of synthesized FDG PET scans as input, the MLP classifiers achieve accuracy of 79.16% and 81.41% for synthesized FDG-PET scans derived from the one-way diffusion model and two-way diffusion model, respectively. The synthesized FDG PET scans derived from

TABLE 2 Comparison of works with the same task as our proposed synthesis models for the prediction of FDG PET from MRI T1WI.

	MAE	MSE	ZNCC	SSIM	PSNR
Li et al. ²⁹	0.0862	–	–	0.5419	–
Sikka et al. ³⁰	0.0422	–	–	0.8211	–
Hu et al. ²⁷	–	–	–	0.89	27.88
Zhang et al. ²⁸	0.0318	–	–	0.7294	26.92
Vega et al. ³²	–	–	–	0.918	18.23
Vega et al. ³¹	–	–	–	0.905	22.69
Ouyang et al. ³³	–	–	–	0.88	27.83
Diffusion one-way	0.0274	0.0044	0.9814	0.9282	23.83
Diffusion two-way	0.0187	0.0024	0.9823	0.9380	26.47

Note: For ZNCC, SSIM, and PSNR metrics, the higher value indicates the better performance for the prediction and for MAE and MSE metrics, the lower value indicates the better performance of the prediction. The metrics with the best performance are highlighted in bold.

Abbreviations: FDG PET, fluorodeoxyglucose positron emission tomography; MAE, mean absolute error; MRI T1WI, magnetic resonance imaging T1 weighted imaging; MSE, mean squared error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; ZNCC, zero-normalized cross correlation.

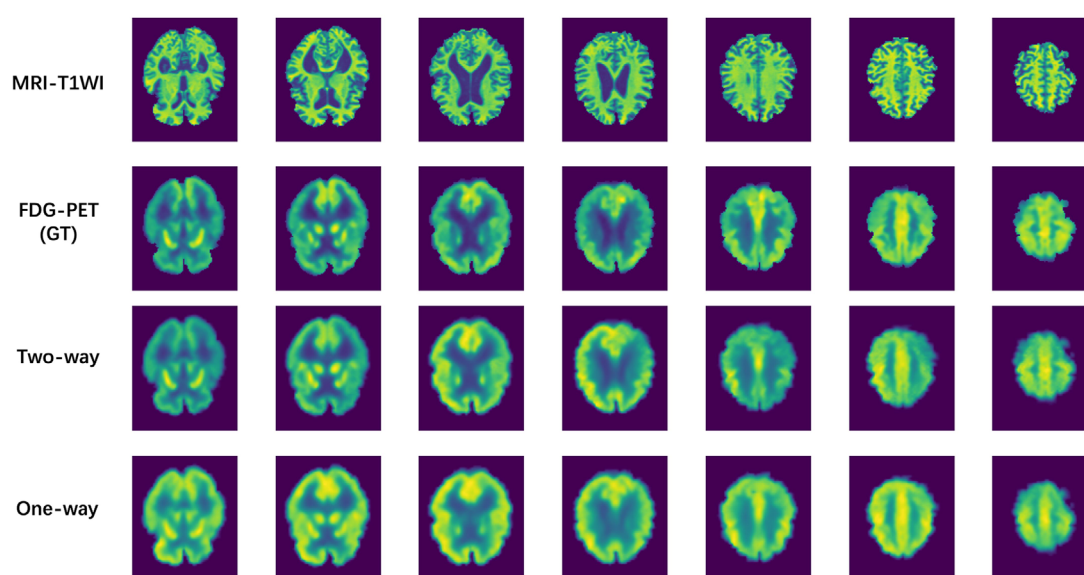


FIGURE 5 A visual comparison of the three synthetic results from a randomly selected sample. The first row is the conditional MRI T1WI, the second row is the ground truth FDG PET, the third and fourth rows are the synthetic results, two-way diffusion and one-way diffusion sequentially. FDG PET, fluorodeoxyglucose positron emission tomography; MRI T1WI, magnetic resonance imaging T1 weighted imaging.

one-way diffusion perform slightly worse than MRI T1WI but are still acceptable while that of two-way diffusion performs better than MRI T1WI, which indicates that using the synthesized FDG PET scans for AD diagnosis is feasible. Moreover, combining synthesized FDG PET scan results with the MRI T1WI scans can produce a better overall performance compared to individual results as shown in Table 3. With MRI T1WI and synthesized FDG PET from one-way diffusion, the MLP classifier achieves an accuracy of 81.90%, which even outperforms that of using MRI T1WI and ground truth FDG PET as input. Using MRI T1WI and synthesized FDG PET from two-way diffusion as joint input achieves the best performance among our experiments with different input settings with an accuracy of 82.19%.

We have also conducted experiments on binary classification, namely NC/MCI, MCI/AD, and NC/AD, in which the settings for each binary classification are the same as the three-class classification of NC/MCI/AD. The results are also shown in Table 3, in which using FDG PET as input achieves the best performance for NC/MCI with an accuracy of 91.49%, using MRI T1WI and FDG PET as joint input achieves the best performance for MCI/AD with an accuracy of 84.26%, and using MRI T1WI and synthesized FDG PET scans derived from the two-way diffusion model as joint input achieves the best performance for NC/AD with an accuracy of 94.05%. Though using MRI T1WI and synthesized FDG PET scans derived from the two-way diffusion model as joint input only achieves the best performance on NC/AD

TABLE 3 Experimental results of MLP-based classifiers with different types of input and their combinations.

NC/MCI/AD				
	Precision	Recall	F1 score	Accuracy
MRI T1WI				
NC	0.8893	0.8737	0.8814	80.23%
MCI	0.7874	0.8009	0.7941	
AD	0.7378	0.7296	0.7337	
FDG PET				
NC	0.8917	0.8667	0.8790	81.31% ↑
MCI	0.7910	0.8266	0.8084	
AD	0.7704	0.7333	0.7514	
One-way diffusion				
NC	0.8737	0.8737	0.8737	79.16%
MCI	0.7732	0.8030	0.7878	
AD	0.7341	0.6852	0.7088	
Two-way diffusion				
NC	0.8957	0.8737	0.8845	81.41% ↑
MCI	0.7926	0.8266	0.8092	
AD	0.7665	0.7296	0.7476	
MRI T1WI + FDG PET				
NC	0.8857	0.8702	0.8779	81.31% ↑
MCI	0.7918	0.8223	0.8067	
AD	0.7743	0.7370	0.7552	
MRI T1WI + one-way				
NC	0.8979	0.8947	0.8963	81.90% ↑
MCI	0.8004	0.8158	0.8081	
AD	0.7672	0.7444	0.7556	
MRI T1WI + two-way				
NC	0.8932	0.8807	0.8869	82.19% ↑
MCI	0.7967	0.8308	0.8134	
AD	0.7913	0.7444	0.7672	
NC/MCI				
	Precision	Recall	F1 score	Accuracy
MRI T1WI				
NC	0.8439	0.8912	0.8669	89.63%
MCI	0.9313	0.8994	0.9150	
FDG PET				
NC	0.8797	0.8982	0.8889	91.49% ↑
MCI	0.9371	0.9251	0.9310	
One-way diffusion				
NC	0.8464	0.9088	0.8765	90.29% ↑
MCI	0.9417	0.8994	0.9200	
Two-way diffusion				
NC	0.8741	0.9018	0.8877	91.36% ↑
MCI	0.9389	0.9208	0.9297	

(Continues)

TABLE 3 (Continued)

NC/MCI				
	Precision	Recall	F1 score	Accuracy
MRI T1WI + FDG PET				
NC	0.8737	0.8982	0.8858	91.22% ↑
MCI	0.9368	0.9208	0.9287	
MRI T1WI + one-way				
NC	0.8543	0.9053	0.8790	90.56% ↑
MCI	0.9400	0.9058	0.9226	
MRI T1WI + two-way				
NC	0.8610	0.8912	0.8759	90.43% ↑
MCI	0.9322	0.9122	0.9221	
MCI/AD				
	Precision	Recall	F1 score	Accuracy
MRI T1WI				
MCI	0.8647	0.8758	0.8702	83.45%
AD	0.7803	0.7630	0.7715	
FDG PET				
MCI	0.8628	0.8887	0.8755	83.99% ↑
AD	0.7969	0.7556	0.7757	
One-way diffusion				
MCI	0.8587	0.8587	0.8587	82.09%
AD	0.7556	0.7556	0.7556	
Two-way diffusion				
MCI	0.8607	0.8865	0.8734	83.72% ↑
AD	0.7930	0.7519	0.7719	
MRI T1WI + FDG PET				
MCI	0.8695	0.8844	0.8769	84.26% ↑
AD	0.7939	0.7704	0.7820	
MRI T1WI + one-way				
MCI	0.8650	0.8779	0.8714	83.58% ↑
AD	0.7833	0.7630	0.7730	
MRI T1WI + two-way				
MCI	0.8628	0.8887	0.8755	83.99% ↑
AD	0.7969	0.7556	0.7757	
NC/AD				
	Precision	Recall	F1 score	Accuracy
MRI T1WI				
NC	0.9531	0.9263	0.9395	93.87%
AD	0.9245	0.9519	0.9380	
FDG PET				
NC	0.9313	0.9509	0.9410	93.87%
AD	0.9470	0.9259	0.9363	
One-way diffusion				
NC	0.9437	0.9404	0.9420	94.05% ↑
AD	0.9373	0.9407	0.9390	

(Continues)

TABLE 3 (Continued)

NC/AD	Precision	Recall	F1 score	Accuracy
Two-way diffusion				
NC	0.9496	0.9263	0.9378	93.69%
AD	0.9242	0.9481	0.9360	
MRI T1WI + FDG PET				
NC	0.9371	0.9404	0.9387	93.69%
AD	0.9368	0.9333	0.9351	
MRI T1WI + one-way				
NC	0.9399	0.9333	0.9366	93.51%
AD	0.9301	0.9370	0.9336	
MRI T1WI + two-way				
NC	0.9532	0.9298	0.9414	94.05% ↑
AD	0.9278	0.9519	0.9397	

Note: The best accuracies of experiments with different tasks are highlighted in bold.

Abbreviations: AD, Alzheimer's disease; FDG PET, fluorodeoxyglucose positron emission tomography; MAE, mean absolute error; MCI, mild cognitive impairment; MLP, multilayer perceptron; MRI T1WI, magnetic resonance imaging T1 weighted; MSE, mean squared error; NC, normal control; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; ZNCC, zero-normalized cross correlation.

classification, it still indicates that involving synthesized FDG PET scans derived from the two-way diffusion model can enhance the classification performance compared to using MRI T1WI as input. Meanwhile, using synthesized FDG PET scans derived from the one-way diffusion model can also enhance the classification performance for NC/MCI and MCI/AD, except for NC/AD. In conclusion, the results show the feasibility of the synthesized model in computer-aided diagnosis for AD.

4 | DISCUSSION

From the view of FDG PET synthesis and according to Table 2 and Figure 5, two-way synthesis outperforms one-way synthesis by using diffusion models both from the visual check and criterion metrics. As described by Xiao et al.,³⁸ the diffusion models have the advantages of both "high-quality samples" and "model converge/diversity." The reason the two-way synthesis outperforms the one-way synthesis might be that the one-way synthesis introduces randomly generated noise in the synthesis process, which increases the bias and thus reduces the quality of the synthesized FDG PET views. Though the one-way synthesis performs worse than the two-way synthesis, it still looks similar to the ground truth FDG PET scans.

As using both the one-way synthesis and two-way synthesis for FDG PET views can produce convincing results for AD diagnosis, thus they are both effective in the clinic. Meanwhile, the complementary views of synthesized FDG PET scans with MRI T1WI scans can enhance the overall performance of ML-based AD diagnosis, which implies the synthesized FDG PET scans not only look similar to the ground truth FDG PET scans but also make contributions to AD diagnosis. Furthermore, comparing the performance of using MRI T1WI to syn-

thesized FDG PET scans derived from either the one-way or two-way diffusion model, it appears that the two-way synthesis produces better results compared to the one-way synthesis for NC/MCI/AD, MCI/AD, and NC/AD while opposite for NC/MCI. A similar situation exists for NC/AD when using synthesized scans as input only. Considering the overall classification performance together with the criterion metrics for image similarity measure, we still hypothesize that two-way synthesis generally outperforms one-way synthesis. However, according to the results, one-way synthesis is still valuable for enhancing the performance of ML-based AD diagnosis. Biases introduced by one-way synthesis might also help in the training process to overcome the overfitting problem, especially for the limited-scale dataset.

Our experimental results demonstrate that the involvement of an additional view, specifically FDG PET in this work, can enhance the performance of ML-based models for AD diagnosis, even though the involved view is synthesized instead of ground truth. Therefore, we consider the generality that more views, including but not limited to image data, can also enhance the model's performance. As many disease-related medical data are hard or costly to acquire, for future works, involving more views, either of real data or of synthesized data, this would be a feasible way to improve the performance of ML-based models for AD diagnosis.

5 | CONCLUSION

This paper proposes a diffusion model with two strategies to synthesize FDG PET scans from MRI T1WI scans for AD diagnosis. In our proposal, the two-way synthesis achieves state-of-the-art results with MAE at 0.0187, MSE at 0.0024, ZNCC at 0.9823, SSIM at 0.9380, and PSNR at

26.47. The one-way synthesis also shows convincing results with SSIM scores of 0.9282.

Further to the evaluation of the synthesized FDG PET scans compared to the ground truth FDG PET scans, we have also designed MLP-based classifiers to carry out validation. The experiments show positive results and prove the feasibility of synthesized FDG PET scans in AD diagnosis and the ability to improve the classifiers' performance.

This work gives support to the notion that ML-based cross-domain data synthesis can be a useful approach to improve the diagnosis of AD by providing additional synthesized disease-related views for multi-view learning.

ACKNOWLEDGMENTS

The work is supported in part by Ningbo Major Science & Technology Program under Grant 2022Z126, Nottingham Short-term RKE Project under Grant I05220500009. A.A.H. receives funding from the Medical Research Council, UK (grant MR/T005580/1), and National Institute of Health/NIA, USA (grant 1R56AG074467-01) through her institute. The funders had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Author disclosures are available in the [supporting information](#).

REFERENCES

- Shachor Y, Greenspan H, Goldberger J. A mixture of views network with applications to multi-view medical imaging. *Neurocomputing*. 2020;374:1-9.
- Zhao J, Xie X, Xu X, Sun S. Multi-view learning overview: recent progress and new challenges. *Inf Fusion*. 2017;38:43-54.
- Li Y, Yang M, Zhang Z. A survey of multi-view representation learning. *IEEE Trans Knowl Data Eng*. 2018;31(10):1863-1883.
- Yan X, Hu S, Mao Y, Ye Y, Yu H. Deep multi-view learning methods: a review. *Neurocomputing*. 2021;448:106-129.
- Jouirou A, Baâzaoui A, Barhoumi W. Multi-view information fusion in mammograms: a comprehensive overview. *Inf Fusion*. 2019;52:308-321.
- Azam MA, Khan KB, Salahuddin S, et al. A review on multimodal medical image fusion: compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Comput Biol Med*. 2022;144:105253.
- Haribabu M, Guriviah V, Yogarajah P. Recent advancements in multimodal medical image fusion techniques for better diagnosis: an overview. *Curr Med Imag Formerly Curr Med Imag Rev*. 2022;18.
- Patterson C. World Alzheimer report 2018. Alzheimer's Disease International. 2018.
- Mueller SG, Weiner MW, Thal LJ, et al. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin N Am*. 2005;15(4):869-877.
- Kurlowicz L, Wallace M. The Mini-Mental State Examination (MMSE). *J Gerontol Nurs*. 1999;25:8-9.
- Pinto TC, Machado L, Bulgacov TM, et al. Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's disease (AD) in the elderly?. *Int Psychogeriatr*. 2019;31(4):491-504.
- Liu M, Zhang D, Shen D, Initiative ADN. View-centralized multi-atlas classification for Alzheimer's disease diagnosis. *Hum Brain Mapp*. 2015;36(5):1847-1865.
- Wee CY, Yap PT, Shen D; Initiative ADNI. Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Hum Brain Mapp*. 2013;34(12):3411-3425.
- Chincarini A, Bosco P, Calvini P, et al.; Alzheimer's Disease Neuroimaging Initiative. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *Neuroimage*. 2011;58(2):469-480.
- Kruthika K, Maheshappa H; Initiative ADNI, others. CBIR system using capsule networks and 3D CNN for Alzheimer's disease diagnosis. *Inf Med Unlock*. 2019;14:59-68.
- Xia Z, Yue G, Xu Y, et al. A novel end-to-end hybrid network for Alzheimer's disease detection using 3D cnn and 3D CLSTM. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (Isbi)*. IEEE; 2020:1-4.
- Zhu W, Sun L, Huang J, Han L, Zhang D. Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI. *IEEE Trans Med Imaging*. 2021;40(9):2354-2366.
- Cheng J, Liu Z, Guan H, et al. Brain age estimation from MRI using cascade networks with ranking loss. *IEEE Trans Med Imaging*. 2021;40(12):3400-3412.
- Qiao H, Chen L, Ye Z, Zhu F. Early Alzheimer's disease diagnosis with the contrastive loss using paired structural MRIs. *Comput Methods Programs Biomed*. 2021;208:106282.
- Chen K, Weng Y, Hosseini AA, Dening T, Zuo G, Zhang Y. A comparative study of GNN and MLP based machine learning for the diagnosis of Alzheimer's disease involving data synthesis. *Neural Netw*. 2024;169:442-452.
- Pan X, Phan TL, Adel M, et al. Multi-view separable pyramid network for AD prediction at mci stage by 18 F-FDG brain PET imaging. *IEEE Trans Med Imaging*. 2020;40(1):81-92.
- Lu D, Popuri K, Ding GW, Balachandrar R, Beg MF; Alzheimer's Disease Neuroimaging Initiative. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease. *Med Image Anal*. 2018;46:26-34.
- Zhang X, Yang Y, Li T, Zhang Y, Wang H, Fujita H. CMC: a consensus multi-view clustering model for predicting Alzheimer's disease progression. *Comput Methods Programs Biomed*. 2021;199:105895.
- Xu L, Wu H, He C, et al. Multi-modal sequence learning for Alzheimer's disease progression prediction with incomplete variable-length longitudinal data. *Med Image Anal*. 2022;82:102643.
- Hao X, Bao Y, Guo Y, et al.; Alzheimer's Disease Neuroimaging Initiative. Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. *Med Image Anal*. 2020;60:101625.
- Huang Y, Xu J, Zhou Y, Tong T, Zhuang X; (ADNI) ADNI. Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network. *Front Neurosci*. 2019;13:509.
- Hu S, Lei B, Wang S, Wang Y, Feng Z, Shen Y. Bidirectional mapping generative adversarial networks for brain MR to PET synthesis. *IEEE Trans Med Imaging*. 2021;41(1):145-157.
- Zhang J, He X, Qing L, Gao F, Wang B. BPGAN: brain PET synthesis from MRI using generative adversarial network for multi-modal Alzheimer's disease diagnosis. *Comput Methods Programs Biomed*. 2022;217:106676.
- Li R, Zhang W, Suk HI, et al. Deep learning based imaging data completion for improved brain disease diagnosis. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2014:305-312.
- Sikka A, Peri SV, Bathula DR. MRI to FDG-PET: cross-modal synthesis using 3D U-Net for multi-modal Alzheimer's classification.

- International Workshop on Simulation and Synthesis in Medical Imaging. Springer; 2018:80-89.
31. Vega F, Addeh A, Ganesh A, Smith EE, MacDonald ME. Image translation for estimating two-dimensional axial amyloid-beta PET from structural MRI. *J Magn Reson Imaging*. 2024;59(3):1021-1031.
 32. Vega F, Addeh A, MacDonald ME. Amyloid-beta PET synthesis from structural MRI: a potential alternative method for Alzheimer's disease screening. *Alzheimers Dement*. 2023;19:e079258.
 33. Ouyang J, Chen KT, Duarte Armindo R, et al. Predicting FDG-PET images from multi-contrast MRI using deep learning in patients with brain neoplasms. *J Magn Reson Imaging*. 2024;59(3):1010-1020.
 34. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage*. 2012;62(2):782-790.
 35. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Adv Neural Info Proc Sys*. 2020;33:6840-6851.
 36. Song J, Meng C, Ermon S. Denoising diffusion implicit models. *arXiv preprint arXiv:201002502*. Published online 2020.
 37. Rolls ET, Huang CC, Lin CP, Feng J, Joliot M. Automated anatomical labelling atlas 3. *Neuroimage*. 2020;206:116189.
 38. Xiao Z, Kreis K, Vahdat A. Tackling the generative learning trilemma with denoising diffusion gans. *arXiv preprint arXiv:211207804*. Published online 2021.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chen K, Weng Y, Huang Y, et al. A multi-view learning approach with diffusion model to synthesize FDG PET from MRI T1WI for diagnosis of Alzheimer's disease. *Alzheimer's Dement*. 2025;21:e14421. <https://doi.org/10.1002/alz.14421>