GAN-based synthetic brain PET image generation

Jyoti Islam[*] and Yanqing Zhang

Abstract

In recent days, deep learning technologies have achieved tremendous success in computer vision-related tasks with the help of large-scale annotated dataset. Obtaining such dataset for medical image analysis is very challenging. Working with the limited dataset and small amount of annotated samples makes it difficult to develop a robust automated disease diagnosis model. We propose a novel approach to generate synthetic medical images using generative adversarial networks (GANs). Our proposed model can create brain PET images for three different stages of Alzheimer's disease—normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

Keywords: Synthetic medical image generation, Positron emission tomography (PET), Generative adversarial networks, Brain imaging, Alzheimer's disease
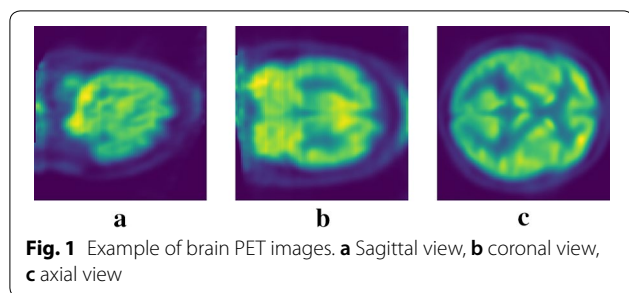
1 Introduction

Developing AI-assisted automated disease diagnosis systems using medical images often requires a large training dataset with annotated samples, especially for supervised learning methods. Experts with good knowledge of the specific data and task are needed for performing such annotations. So, medical image annotation process is expensive in terms of time, money and effort. It becomes more challenging for precise annotations, such as for identifying different stages of Alzheimer's disease. If diagnostic images are intended to be made public, patient consent may be necessary depending on the institutional protocols [1]. So there are very few public medical datasets available online, and they are still limited in size and quality. Collecting medical images for developing automated computer-aided diagnosis system is a complicated and expensive procedure and requires adequate funding, handling privacy concern, and collaboration of researchers, physicians, and hospitals. Medical datasets are often imbalanced as pathologic findings are usually rare, and it creates another challenge to train the automated diagnosis system (Fig. 1).

Data augmentation is one way to overcome the problem of limited dataset. There are several data augmentation techniques, such as translation, rotation, scale, flip, etc. But these techniques are not as useful for medical image analysis as they are for natural image dataset. On the contrary, techniques such as translation and rotation might change the pattern useful for the diagnosis. Besides, these images resemble a great extent to the original ones. So the ML model using these augmented data gain little performance improvements due to the lack of generalization abilities. Another type of data augmentation strategy is synthetic data generation. A synthetic dataset is generated programmatically. Such dataset is highly beneficial for medical image analysis. There is no patient data handling or privacy concerns as the data are produced synthetically. The dataset can contain samples from both positive and negative classes for diagnosis purpose and help build a generalized model.

Generating synthetic images for building a large-scale dataset for training deep learning model is an active research area. Image-to-image translation methods have made it possible to create such synthetic images. Image-to-image translation refers to the problem of translating a representation of an image into another, for example, converting an RGB image to BW image or vice versa. Both supervised and unsupervised

*Correspondence: jislam2@student.gsu.edu
Department of Computer Science, Georgia State University, Atlanta, Georgia 30302-5060, USA

Fig. 1 Example of brain PET images. **a** Sagittal view, **b** coronal view, **c** axial view

technologies have been used for image-to-image translation. Generative Adversarial Networks (GANs) [2] can generate synthetic data with good generalization ability. GAN has two different networks—Generator and Discriminator. The model is trained in an adversarial process where the Generator generates fake images, and the Discriminator learns to discriminate between the real and fake images. Computer vision community have widely used GAN for image-to-image translation and done some excellent research works to generate synthetic data [3–10]. Such papers have achieved impressive results on background removal, palette generation, sketch to portrait, pose transfer, semantic segmentation [3], super resolution [4], style transfer [5], image inpainting [6], future state prediction [8], and image manipulation guided by user constraints [9]. The success of the vision community for synthetic data generation using GAN and the limitation of medical data inspired us to explore methods suitable for medical image synthesis. In this study, we focus on synthetic brain positron emission iomography (PET) image generation for different stages of Alzheimer's disease—normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

Alzheimer's disease (AD) is a severe neurological disorder and the most common type of dementia. The prevalence of Alzheimer's disease is approximated to be around 5% after 65 years. In developed countries, the prevalence of Alzheimer's disease is staggering 30% for more than 85 years old. There is a high probability that around 0.64 billion people will be diagnosed with Alzheimer's disease by 2050 [11]. Alzheimer's disease is incurable. The effect of Alzheimer's disease is losing memory, ability to continue day-to-day activities and performing mental functions. At the initial stage, Alzheimer's disease affects the brain part controlling memory and language functionality. So, patients suffer from memory loss, confusion, and difficulty in speaking, reading or writing. Alzheimer's disease patients tend to forget about their life history and often cannot recognize family members. Alzheimer's disease patients have

difficulties in daily activities such as combing the hair or brushing the teeth. As a result, patients with Alzheimer's disease become anxious or aggressive. As they forgot things, they often wander away from home. In the advanced stage, the brain part controlling breathing and heart functionality get destroyed, and that causes death.

Since Alzheimer's disease is incurable, it is crucial to detect patients at the MCI stage before the disease progresses further. Earlier diagnosis can help in proper treatment and prevent brain tissue damage. Alzheimer's disease causes degeneration of brain cells. Such changes can be captured using different imaging modalities, e.g., structural and functional magnetic resonance imaging (sMRI, fMRI), positron emission tomography (PET), single photon emission computed tomography (SPECT), and diffusion tensor imaging (DTI) scans, etc. With the progression of Alzheimer's disease, the volume of abnormal proteins (amyloid-β [Aβ] and hyperphosphorylated tau) increases in the brain. The accumulation of these proteins causes gradual changes in the brain and leads to progressive synaptic, neuronal and axonal damage. There is a stereotypical pattern of these changes, including early medial temporal lobe (entorhinal cortex and hippocampus) involvement, followed by progressive neocortical damage [12]. These changes often occur years before the symptoms of Alzheimer's disease appear. The toxic hyperphosphorylated tau and/or amyloid-β [Aβ] seems to slowly erode the brain. Finally, amnestic symptoms start to develop when a clinical threshold is surpassed.

Hippocampus is a part of the brain that controls episodic and spatial memory. It is a small but vital organ that works as a relay structure between the brain and the body. Alzheimer's disease shrinks the hippocampus and cerebral cortex of the brain and enlarges the ventricles [13]. If the hippocampus is shrunk, it causes cell loss and damage to synapses and neuron ends. So neurons cannot communicate anymore via synapses. As a result, brain regions related to remembering (short-term memory), thinking, planning, and judgment are affected [13]. The degenerated brain cells can be captured using positron emission tomography (PET) for measuring these progressive changes. We propose a novel model to generate synthetic brain position emission tomography (PET) images exploiting Generative Adversarial Networks for three stages of Alzheimer's disease—normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

The rest of the paper is organized as follows. Section 2 discusses briefly about the related work on synthetic medical data generation. Section 3 presents the proposed model. Section 4 reports the experimental details and the

results. Finally, in Sect. 5, we conclude the paper with our future research direction.

2 Related work

The recent advances of deep learning technologies have brought numerous breakthroughs in machine learning research and reached to a stage in some tasks where they provide similar or better performance than human. Some examples are image classification [14], intelligent driving [15], smart cities [16], voice recognition [17], playing Go [18], medical imaging [19–25], visual sentiment analysis [26], etc. The reason behind this success is significantly dependent on the size and the quality of the dataset being used to train the deep learning model. The scale and quality of the labeled or annotated data determine the performance of the deep learning model. Large-scale annotated dataset is required for training the model to achieve superior model performance. If the training-labeled dataset is small, the model fails to provide a good generalized performance. But obtaining such labeled data is difficult and expensive as it requires close and seamless collaboration from the outstanding experts in the field.

To address the insufficiency of the training dataset, researchers have proposed several oversampling methods. Duplicating the samples from minority class in a imbalanced dataset and adding artificial noise was proposed by DeRouin et al. [27]. Synthetic minority oversampling technique (SMOTE) was proposed by Chawla et al. [28] to create a synthetic dataset with samples from the minority class. Han et al. [29] proposed a Borderline-SMOTE method, considering neighboring instances and the minority instances near the borderline. Sample data generation using the weighted distribution for minority class instances based on the level of difficulty to learn them was proposed by He et al. [30]. Barua et al. [31] proposed a majority weighted minority oversampling technique using Euclidian distance-based clustering method to generate synthetic minority class samples. Xie et al. [32] introduced an oversampling technique by mapping the training samples in a low-dimensional space, assigning weights, and using the local densities. Their method addressed the problem of dimensionality that affected earlier methods. Douzas and Bacao [33] introduced a self-organizing map-based method using artificial data points in high-dimensional space. These oversampling methods helped to achieve more samples for the minority class for imbalanced datasets.

Some traditional approaches have been proposed to address the small sample size problem. Zhou and Jiang [34] trained a neural network and then employed it to generate a new training set, known as the neural-ensemble-based C4.5. Li and Lin [35] determined the probability density function of the training samples and used it

to generate new samples. Li and Fang [36] used group discovery and parametric equations of the hypersphere to propose a non-linear classification technique to generate samples for enlarging the training dataset. These traditional methods are limited in their ability to learn the inherent features of the samples.

Synthetic image generation methods can be classified into two major categories. The first category is the model-based approach where a model is formulated to observe the data and a dedicated engine renders the data. This approach has been used for increasing the training dataset of urban driving environment [37, 38], object detection [39], text segmentation [40], realistic digital brain-phantom generation [41], synthetic agar plate image generation [42]. Designing such specialized data generation engine requires accurate model and deep knowledge of the specific domain. The other category of synthetic image generation method is known as the learning-based approach. These methods can learn the intrinsic spatial variability of the training image dataset. The probability distribution of the real images in the training dataset is learned implicitly by the model, and new images are generated by mimicking the original samples. Generative Adversarial Network is a learning-based approach. For synthetic image generation, both supervised [3, 43, 44] and unsupervised [10, 45–47] approaches are being used. In supervised training, a set of pairs of corresponding images (s_i, t_i) are used, where s_i is an image of the source domain and t_i is a corresponding image in the target domain. For example, Pix2Pix [3] utilizes supervised training using a conditional GAN that learns to generate the output image based on the corresponding input image. The Generator network follows an encoder–decoder structure. The input of the Generator is the image from a particular domain A, and it learns to generate images in a different domain B. The Discriminator examines these generated images based on the training images from domain A and their corresponding images in domain B, and learns to distinguish between real and fake images. Based on the feedback from the discriminator, the Generator learns to generate more realistic images.

Generative adversarial network (GAN) [2] brought a breakthrough in the synthetic data generation research area. It can learn the distribution of the real dataset and generate synthetic samples conforming to that distribution. GAN have been successfully applied in image generation, image inpainting [48], image captioning [49–51], object detection [52], semantic segmentation [53, 54], natural language processing [55, 56], speech enhancement [57], credit card fraud detection [58] and supervised learning with insufficient training data [59]. From the experiments and results of these studies, it is

evident that GAN conforms to the distribution of the original data samples and can generate realistic synthetic data. These promising applications in different fields also emphasize that GAN is independent of the precise domain knowledge for generating synthetic data.

Medical image synthesis and Generative Adversarial Networks have got attention in recent years. Costa et al. [60] used a fully convolutional neural network to train on retinal vessel segmentation images and then applied GANs for generating synthetic retinal images. Dai et al. [61] used GANs for creating lung fields and heart segmentation images from chest X-ray images. Gou et al. [62] proposed a method to employ a GAN to generate and learn from synthetic eye images to improve eye detection accuracy. Shin et al. [63] utilized GANs for generating synthetic abnormal MRI images with brain tumors. Nie et al. [64] proposed an auto-context model for brain CT and MRI image refinement. Schlegl et al. [65] trained GANs for anomaly detection in retinal images. Frid-Adar et al. [66] applied GANs for synthesizing liver lesion ROIs to apply in liver lesion classification. Hu et al. [67] applied GANs to generate a MRI motion model. Mahapatra et al. [68] synthesized high-resolution retinal fundus images using Generative Adversarial Networks. Nie et al. [64] generated synthetic pelvic CT images using GANs. Liu et al. [69] synthesized HCC samples using an approach based on a generative adversarial network (GAN) combined with a deep neural network. Han et al. proposed [70] a two-step GAN-based DA to generate and refine brain magnetic resonance (MR) images with/without tumors separately. Andreini et al. [71] proposed a GAN-based approach for synthesizing high-quality retinal images, along with the corresponding semantic label.

In our previous research works, we had to handle the limited dataset problem for Alzheimer's disease diagnosis. There is a gap in research work for synthesizing brain images for Alzheimer's disease diagnosis. Besides, there are very few works done for PET image synthesis. To mitigate these gaps, we propose a novel model to generate synthetic brain positron emission tomography (PET) images exploiting Generative Adversarial Networks for three stages of Alzheimer's disease—normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

3 Methodology
3.1 Data selection
For our proposed model, we have used 411 PET scans (98 AD, 105 NC, 208 MCI) of 479 patients. We collected the data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Specifically we used ADNI1 baseline dataset for our model. The subjects

were in the age range 55–92. The ADNI was launched in 2003 as a public–private partnership, led by principal investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Up-to-date information related ADNI database can be found at http://www.adni-info.org [72].

3.2 Generative Adversarial Networks
Generative Adversarial Networks (GANs) is a deep learning architecture that consisted of two models—a generative model G and a discriminative model D. The generative model captures the data distribution. The discriminative model estimates the probability that the sample is drawn from the training data rather than the generative model. The two models are simultaneously trained via an adversarial process. The architecture is inspired by game theory and corresponds to a minimax two-player game. The training procedure of G is to maximize the probability of D making a mistake [2].

Let the generator $G(z, \theta_x)$ is a differentiable function represented by a multilayer perceptron with parameters θ_g that depicts a mapping to the data space. To learn the generator's distribution ρ_g over the data space x, a prior ρ_z is defined on random input noise variables z. The discriminator $D(x, \theta_d)$ is also a neural network that gets a sample the real dataset or the generated synthetic dataset produced by G and outputs a single scalar value that the input data comes from the real training dataset. The training process focuses on the task that the discriminator D will maximize the probability of assigning correct labels to the training examples and generated samples from G. At the same time, G is trained to generate data samples similar to the real dataset so that D cannot differentiate them from actual data. Similar to game theory, the discriminator D and the generator G play a two-player mini–max game with following value function $V(G, D)$:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim \rho_{\text{data}(x)}}[\log D(x)] + \mathbb{E}_{z \sim \rho_{\text{data}(z)}}[\log(1 - D(z))], \quad (1)$$

where x is the real data and z is the input random noise. ρ_{data}, ρ_z represent the distribution of the real data and the input noise. $D(x)$ represents the probability that x came from the real data while $G(z)$ represents the mapping to synthesize the real data. The generator G is a deeper neural network and has more convolutional layers and nonlinearities. The noise vector z is upsampled while G

learns the weights through backpropagation. At some point, the generator starts producing data that are classified as real by the discriminator.

3.3 Deep Convolutional Generative Adversarial Networks (DCGANs)

Deep Convolutional Generative Adversarial Networks (DCGAN) [73] is a major improvement on the first GAN [2]. DCGAN can generate better quality images and have more stability during the training stage. In the synthetic image generation process using the DCGAN, there are two phases: a learning phase and a generation phase. In the training phase, the generator draws samples from an N-dimension normal distribution and works on this random input noise vector by successive upsampling operations, eventually generating an image from it. The discriminator attempts to distinguish between images drawn from the generator and images from the training set [73].

Two important features in DCGAN are BatchNorm ([74] for regulating the extracted feature scale, and LeakyRelu [75] for preventing dead gradients. DCGAN also replace all max pooling with convolutional stride and use transposed convolution for upsampling. It eliminates fully connected layers and uses batch normalization. DCGAN uses ReLU in the generator except for the output which uses Tanh and uses LeakyReLU in the discriminator.

3.4 Proposed model

We propose a novel approach to produce synthetic PET images using a Deep Convolutional Generative Adversarial Networks. Following the guidelines to construct the generator and discriminator, described in the paper written by Radford et al. [73], we implemented and trained them on PET scan images using the original discriminator and generator cost functions. Figure 2 shows the proposed synthetic PET image generator model.

3.4.1 Generator architecture

The input of the generator is a vector of random 100 numbers drawn from a uniform distribution, and the output is a brain PET image of size 128 * 128 * 3. The generator architecture is shown in Fig. 3. The network has a fully connected layer and five strided convolutional transpose (known also as 'deconv') layers. The strided convolutional transpose layers transform the latent vector into a volume with shape 128 * 128 * 3. Each convolutional transpose layer is paired with a 2d batch norm layer and a ReLU activation. The strided convolutional transpose layer inserts zeros in between the pixels of the input vector and expands it. The convolution operation is performed over the enlarged vector to create bigger output data. Normalizing responses to have zero mean and unit variance over the entire mini-batch are applied to stabilize the learning process. Figure 4 shows the output
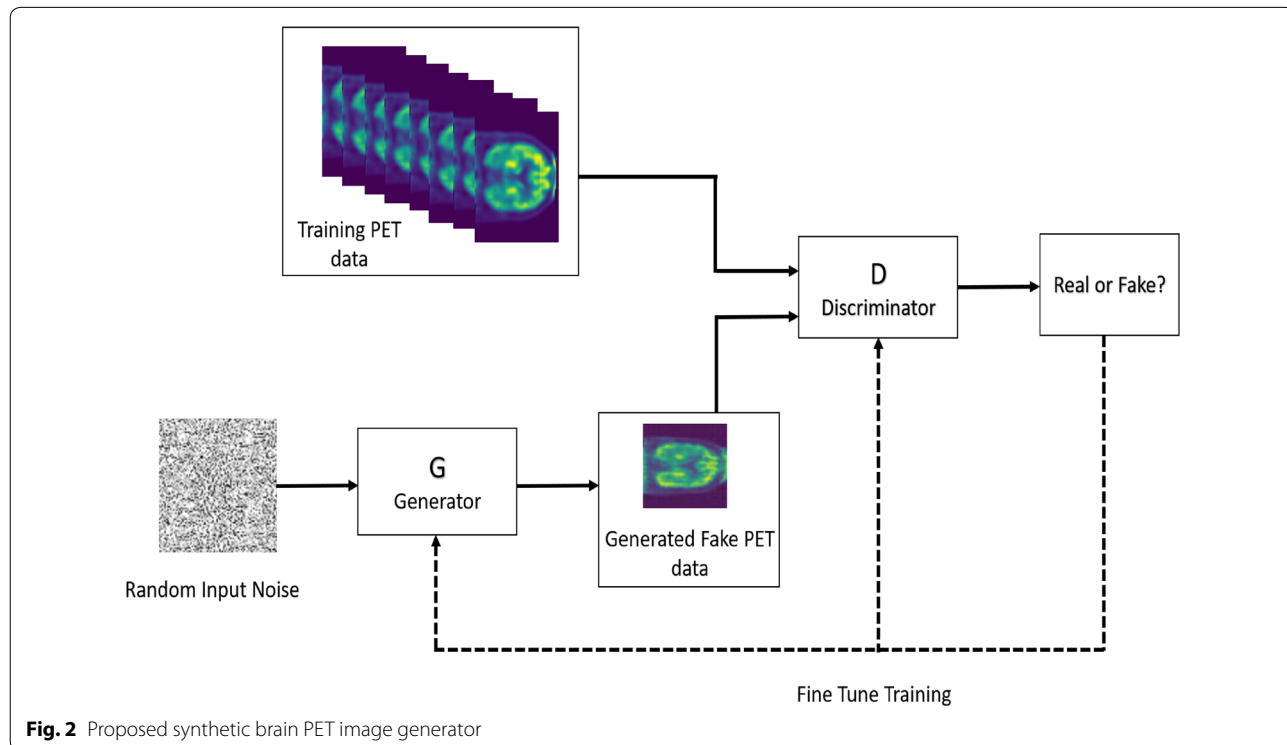

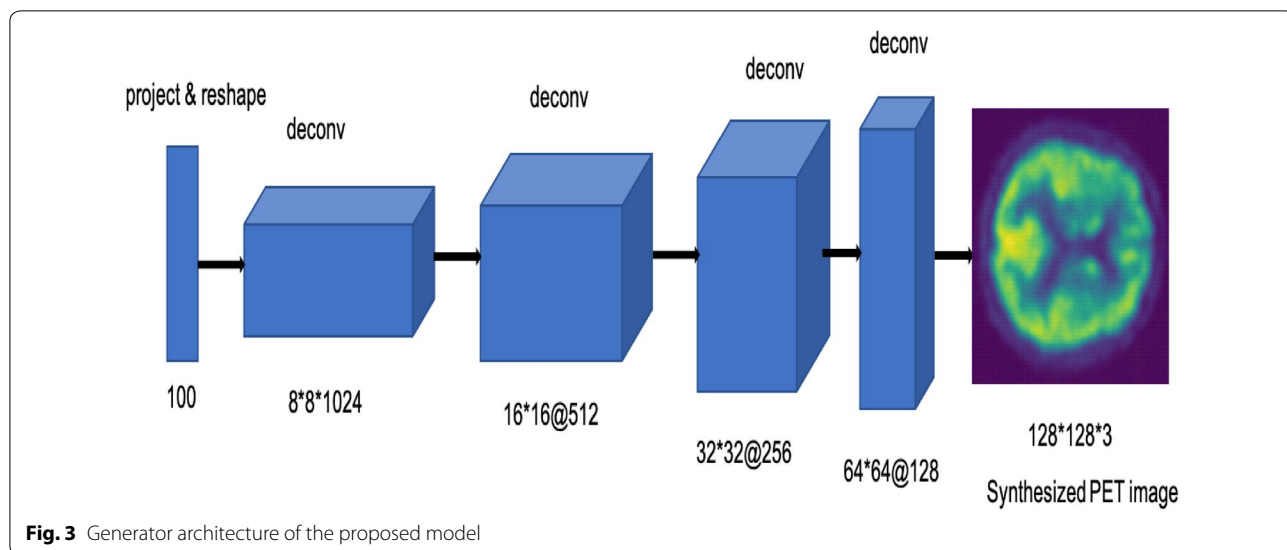**Fig. 2** Proposed synthetic brain PET image generator

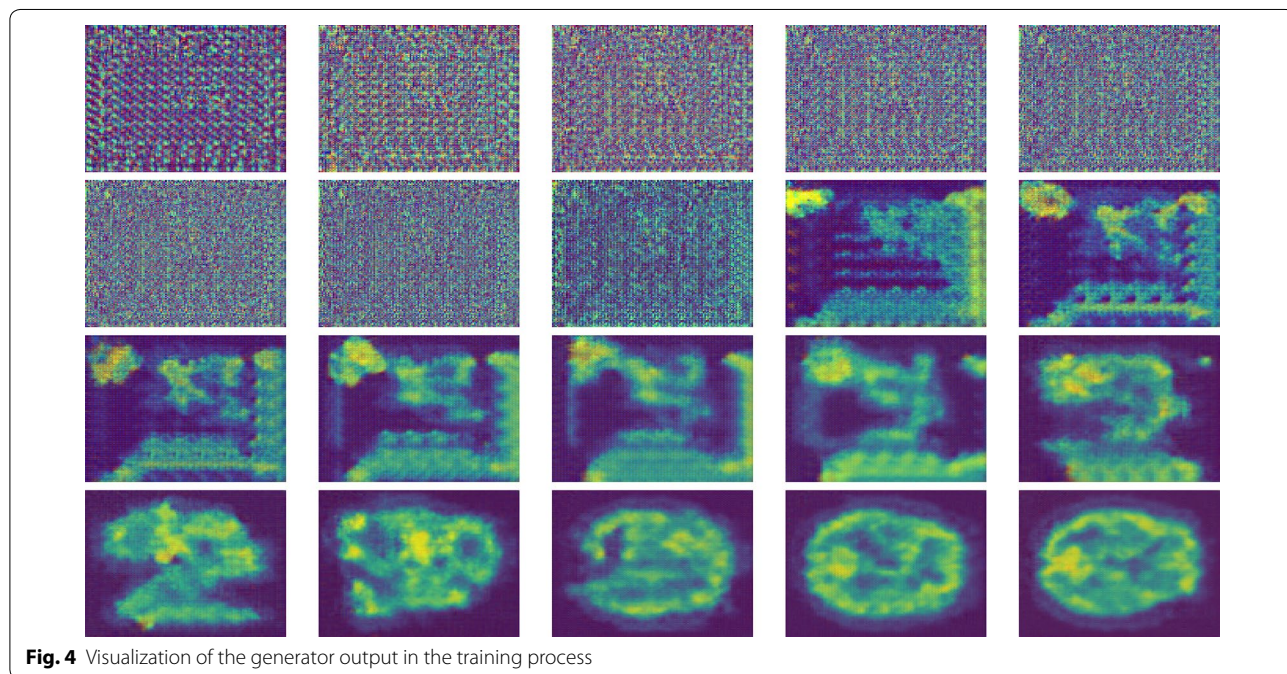**Fig. 3** Generator architecture of the proposed model

Fig. 4 Visualization of the generator output in the training process

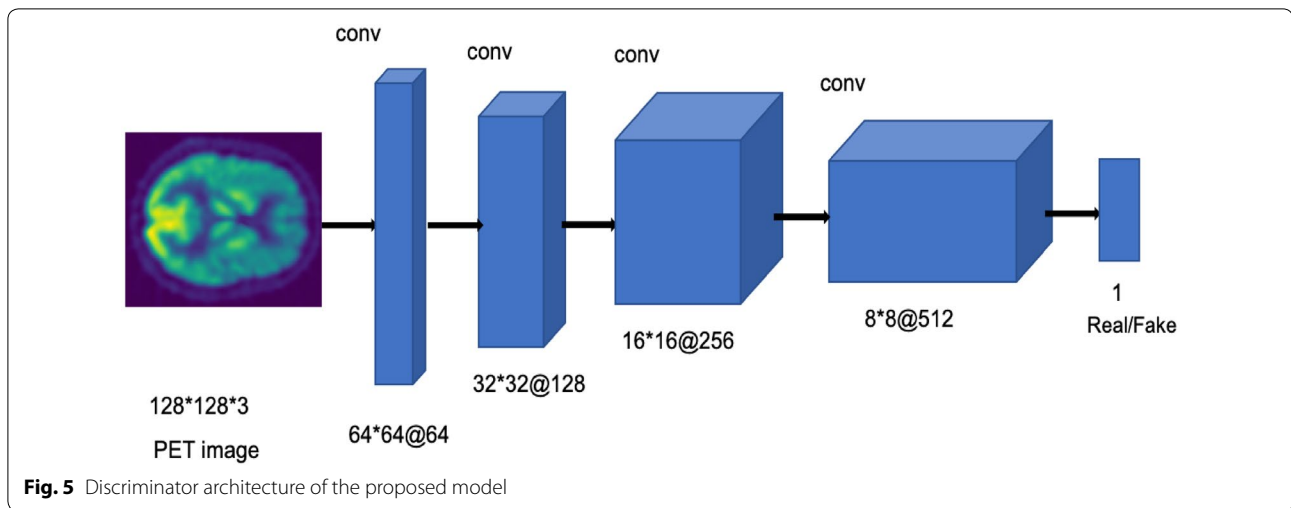of different steps from the generator of the proposed model.

3.4.2 Discriminator architecture
The discriminator network consists of a CNN architecture that takes an image of size 128 * 128 * 3 (brain PET image) as input. The discriminator analyzes the input brain PET image and decides if it is real or fake. The network consists of five convolution layers with a kernel size of 5 * 5 and a fully connected layer. Strided convolutions are applied to each convolutional layer to reduce spatial

dimensionality instead of using pooling layers. Batch-normalization and Leaky ReLU activation are applied to each convolutional layer of the network except the output layer. The fully connected output layer has a Sigmoid function to generate the likelihood probability (0,1) score of the input image to be real or fake. The discriminator architecture is shown in Fig. 5.

3.4.3 Training procedure
We trained the proposed model to synthesize brain PET images for three stages of Alzheimer's disease

Fig. 5 Discriminator architecture of the proposed model

separately. The training process was done iteratively for the generator and the discriminator. We used mini-batches of $m = 64$ brain PET examples for each stage (NC, MCI, and AD) and $m = 64$ noise samples drawn from a uniform distribution between $[-1, 1]$. In the Leaky ReLU, the slope of the leak was set to leak $= 0.2$. We initialized the weights to a zero-centered normal distribution with a standard deviation of 0.02. Stochastic gradient descent was used in the training process with the Adam optimizer, an adaptive moment estimation that incorporates the first and second moments of the gradients, controlled by parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$, respectively. We applied a learning rate of 0.0001 for 500 epochs.

In the training process, the discriminator is trained to maximize the probability of assigning correct labels to the training examples and the generated samples. At first, the discriminator gets a batch of real samples from the training set. The batch is forward passed through D, and the loss ($log(D(x))$) is calculated. The gradients are calculated in a backward pass. Then, a batch of fake samples from the generator is forward passed through D. Similarly, the loss ($log(1 - D(G(z)))$) is calculated, and the gradients are accumulated with a backward pass. Finally, the gradients from both the all-real and all-fake batches are summed up, and a step of the Discriminator's optimizer is done.

The Generator is trained to generate better fake samples by minimizing $log(1 - D(G(z)))$. The training process maximizes $log(D(G(z)))$ to minimize the generator's loss $log(1 - D(G(z)))$. The output of the generator is passed to the discriminator, and the classification result is collected. The training process repeats unless the generator learns to generate samples labeled as real by the Discriminator.

4 Experiments and results

It is an open issue to develop objective metrics that correlate with perceived quality measurement. For quality evaluation of synthetic images, it should be specific for each application. Following previous state-of-the-art, we performed a quantitative and qualitative assessment of our proposed model. To our best knowledge, no previous works attempted to generate synthetic brain PET images using real PET images. We quantitatively compare the predicted results in terms of peak signal to noise ratio (PSNR) and structural similarity Iniex (SSIM). PSNR is used to measure the ratio between the maximum possible intensity value and the mean squared error of the synthetic and the real image:

$$\text{PSNR} = 10 \log_{10} \frac{(\max(y))^2}{\frac{1}{n}\sum_i^n (y_i - \hat{y}_i)^2}, \tag{2}$$

where n is the number of pixels in an image. For our proposed model, the mean PSNR of real and generated images is 32.83.

Structural similarity index (SSIM) finds the similarities within pixels of two image; i.e., if the pixels in the two images line up and or have similar pixel density values:

$$(x, y) = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{3}$$

where x is the estimated PET and y is the ground truth PET, μ_x is the average of x, μ_y is the average of y, μ_x^2 is the variance of x, μ_y^2 is the variance of y, σ_{xy} is the covariance of x and y. $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$ are used to stabilize the division with weak denominator, where L is the dynamic range of the pixel-values, $k_1 = 0.01$ and $k_2 = 0.03$. For our proposed model, the mean SSIM of real and generated images is 77.48.
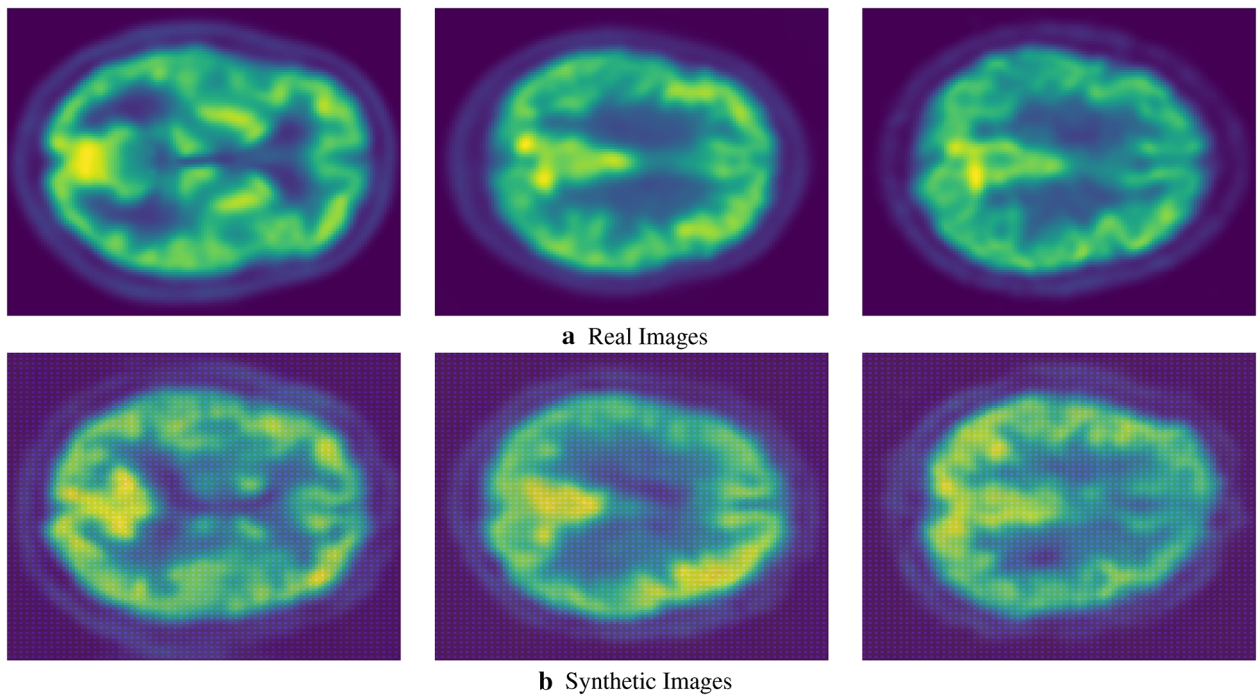
a Real Images

b Synthetic Images

Fig. 6 Real and synthetic brain PET images of normal patient: **a** real **b** synthetic

a Real Images

b Synthetic Images

Fig. 7 Real and synthetic brain PET images of MCI patient: **a** real **b** synthetic
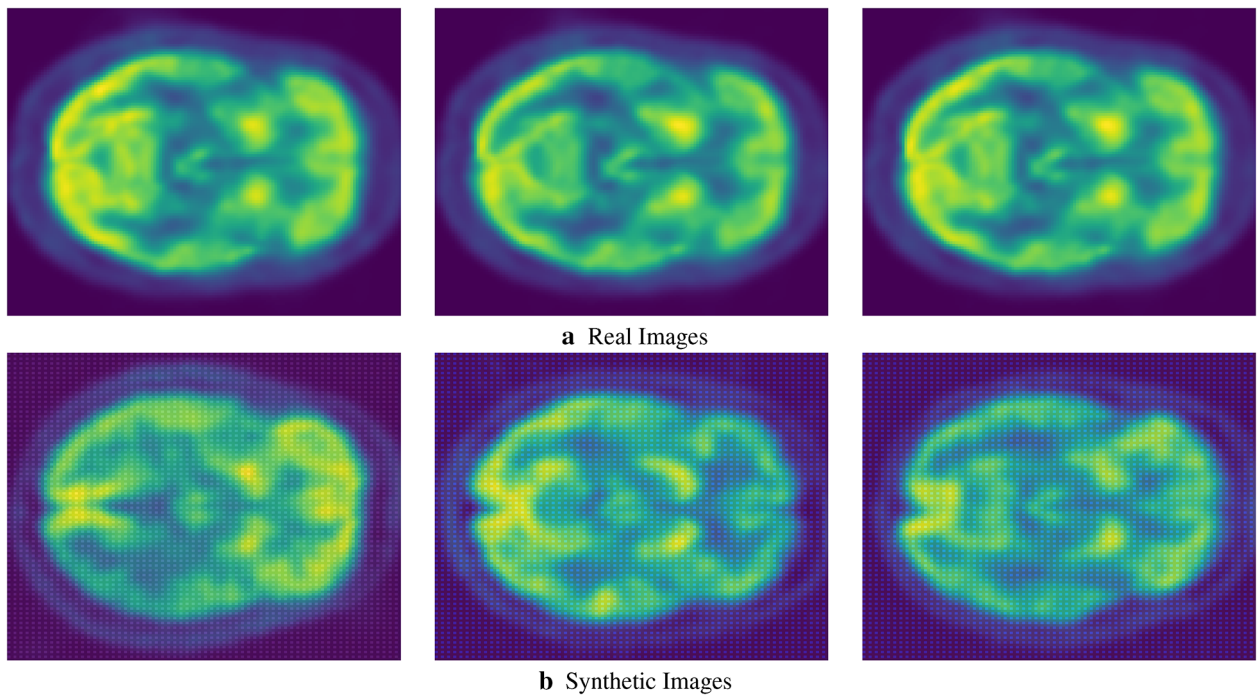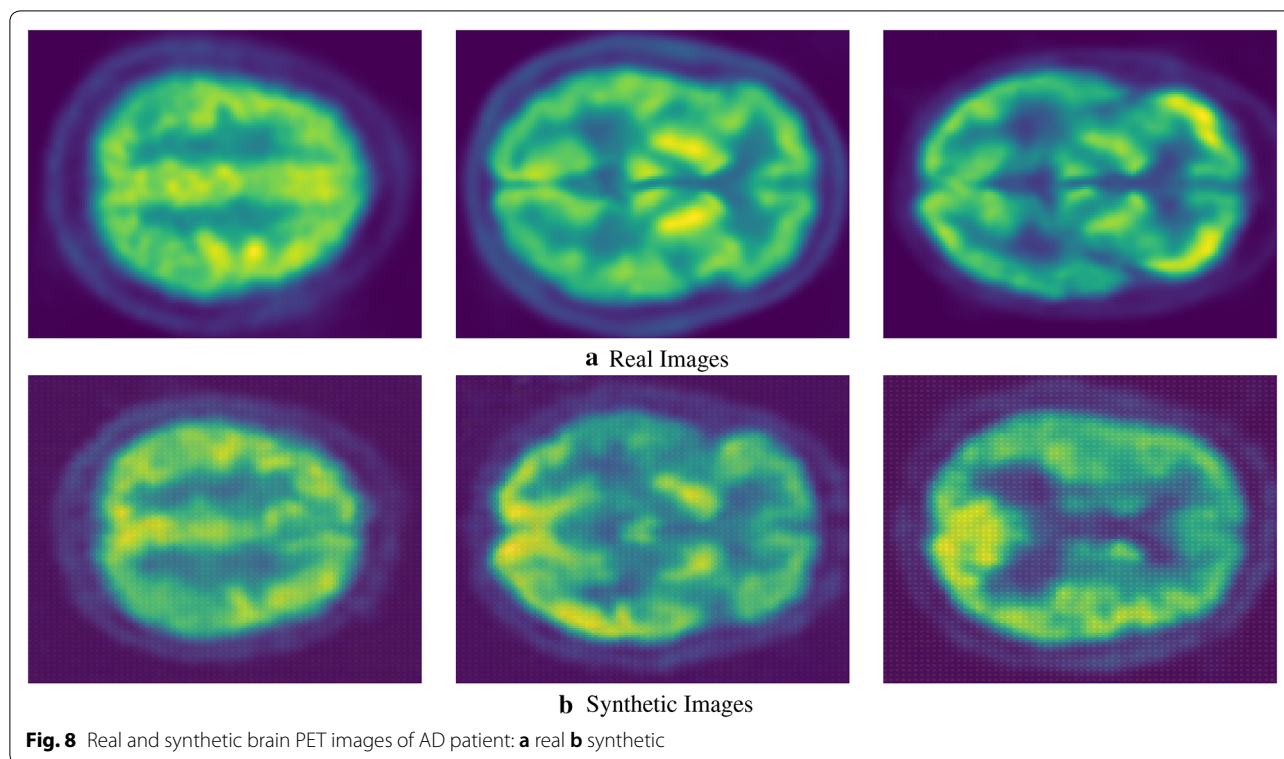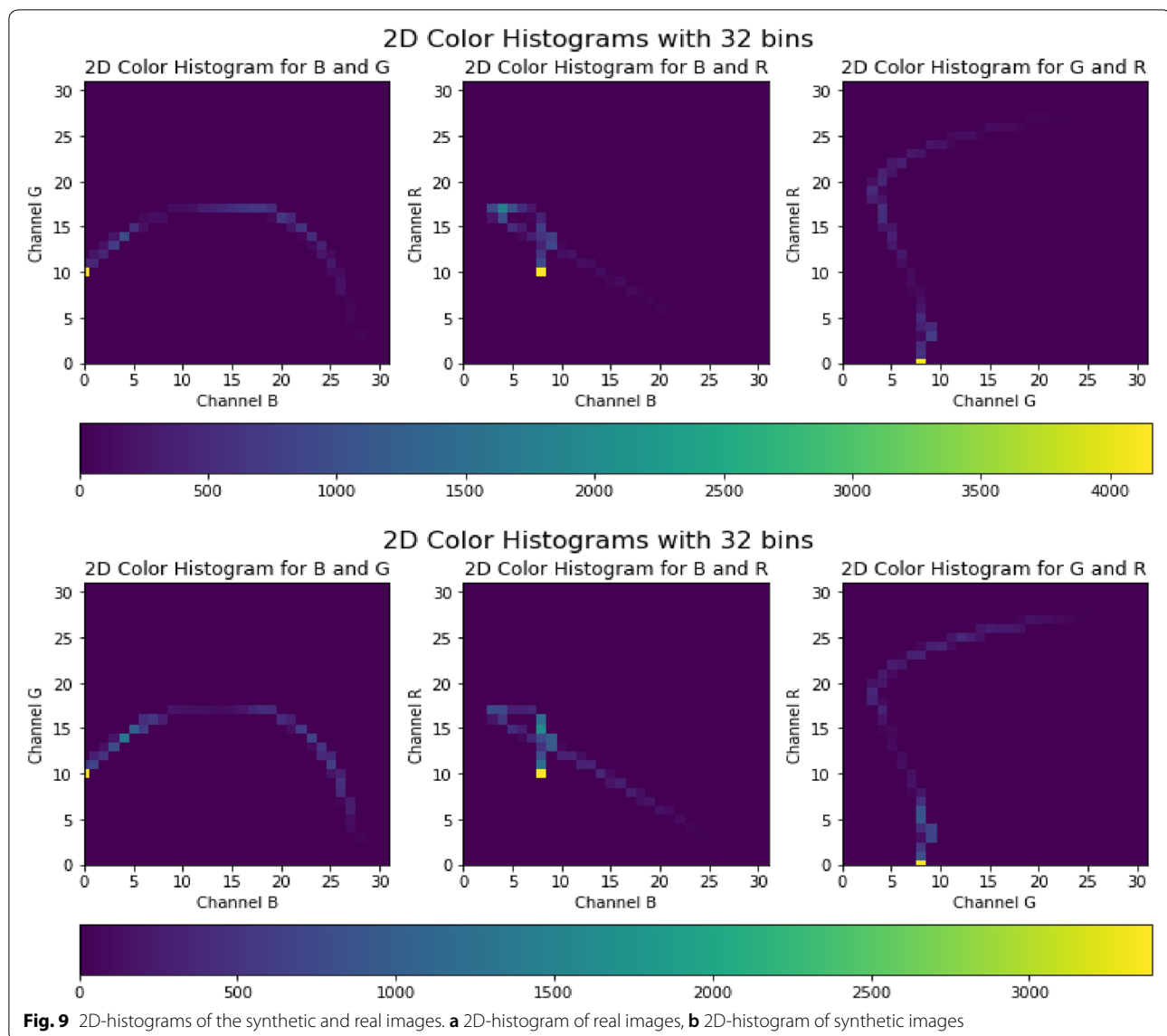
a Real Images

b Synthetic Images

Fig. 8 Real and synthetic brain PET images of AD patient: **a** real **b** synthetic

We present sample visual results of representative slices from the generated PET data for qualitative comparison. Figures 6, 7, and 8 show the synthesized PET images from NC, MCI, and AD patients, respectively. From the results, we could see that the synthesized brain PET images are quite similar to the real brain PET images. To analyze the similarity between synthetic and real images, we also obtained the 2D-histogram of real and synthetic images. Figure 9 presents the 2D-histogram of a sample real and synthetic image [76]. We also developed a 2D-CNN model using axial, coronal, and sagittal slices from the generated PET data. The model achieved 71.45% classification accuracy for CN/AD classification, that is 10% more than the classification model trained without the generated synthetic data.

5 Conclusions

We conclude that synthetic medical image generation is a promising research area and cost-saving approach for developing automated diagnostic technology. Our proposed model can be generalized in other disease diagnosis systems using PET images and can help to supplement the training dataset. The qualitative and quantitative evaluation of the proposed model demonstrates that the synthesized images are close to real brain PET images of different stages of Alzheimer's disease. We believe that our proposed model can help to generate labeled images and aid data augmentation for developing robust disease diagnosis systems, and eventually save lives. There are several limitations to the proposed work. One possible extension could be an increase from 2-D to 3-D input

Fig. 9 2D-histograms of the synthetic and real images. **a** 2D-histogram of real images, **b** 2D-histogram of synthetic images

volumes, using 3-D GAN, at the cost of a longer processing time and an increased memory usage. We trained separate GANs for each stage of Alzheimer's disease, which increased the training complexity. Future research can focus on the investigation of GAN architectures that generate multi-class samples together.

Authors' contributions
JI carried out the background study, proposed and implemented the approach to generate synthetic medical images using Generative Adversarial Networks (GANs), evaluated the result and drafted the manuscript. YZ

supervised the work, monitored result evaluation process, and drafted the manuscript. Both authors read and approved the final manuscript.

References
1. Dysmorphology Subcommittee of the Clinical Practice Committee (2000) Informed consent for medical photographs. Genet Med 2(6):353
2. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems. pp 2672–2680
3. Isola P, Zhu JY, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 1125–1134
4. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, et al (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 4681–4690
5. Li C, Wand M (2016) Precomputed real-time texture synthesis with markovian generative adversarial networks. In: European conference on computer vision. Springer. pp 702–716
6. Pathak D, Krahenbuhl P, Donahue J, Darrell T, Efros AA (2016) Context encoders: Feature learning by inpainting. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 2536–2544
7. Zhang H, Xu T, Li H, Zhang S, Wang X, Huang X, Metaxas DN (2017) Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp 5907–5915
8. Zhou Y, Berg TL (2016) Learning temporal transformations from time-lapse videos. In: European conference on computer vision. Springer. pp 262–277
9. Zhu JY, Krähenbühl P, Shechtman E, Efros AA (2016) Generative visual manipulation on the natural image manifold. In: European conference on computer vision. Springer. pp 597–613
10. Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp 2223–2232
11. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM (2007) Forecasting the global burden of alzheimer's disease. Alzheimer's Dementia 3(3):186–191
12. Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM (2010) The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 6(2):67–77
13. Sarraf S, Anderson J, Tofighi G (2016) Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri. bioRxiv, p 070441
14. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115(3):211–252
15. Al-Qizwini M, Barjasteh I, Al-Qassab H, Radha H (2017) Deep learning algorithm for autonomous driving using googlenet. In: 2017 IEEE intelligent vehicles symposium (IV). IEEE. pp 89–96
16. Wang L, Sng D (2015) Deep learning algorithms with applications to video analytics for a smart city: a survey. arXiv preprint arXiv:1512.03131
17. Ar Mohamed, Dahl GE, Hinton G (2011) Acoustic modeling using deep belief networks. IEEE Trans Audio Speech Lang Proces 20(1):14–22
18. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M et al (2016) Mastering the game of go with deep neural networks and tree search. Nature 529(7587):484
19. Islam J, Zhang Y, for the Alzheimer's Disease Neuroimaging Initiative* (2018) Deep convolutional neural networks for automated diagnosis of Alzheimer's disease and mild cognitive impairment using 3D brain MRI. In: Wang S et al (eds) Brain informatics. BI 2018. Lecture notes in computer science, vol 11309. Springer, Cham, pp 359–369
20. Islam J, Zhang Y (2018) Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. Brain Inform 5:2
21. Islam J, Zhang Y (2017) A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data. In: Zeng Y et al (eds) Brain informatics. BI 2017. Lecture notes in computer science, vol 10654. Springer, Cham, pp 213–222
22. Islam J, Zhang Y (2017) An ensemble of deep convolutional neural networks for Alzheimer's disease detection and classification. arXiv preprint arXiv:1712.01675
23. Islam J, Zhang Y (2018) The IEEE conference on computer vision and pattern recognition (CVPR) Workshops, pp 1881–1883
24. Islam J, Zhang Y (2018) Towards robust lung segmentation in chest radiographs with deep learning. arXiv preprint arXiv:1811.12638
25. Islam J, Zhang Y (2019) Understanding 3D CNN behavior for Alzheimer's disease diagnosis from brain PET scan. arXiv preprint arXiv:1912.04563
26. Islam J, Zhang Y (2016) Visual sentiment analysis for social images using transfer learning approach. In: 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), Atlanta, GA, pp 124–130
27. DeRouin E, Brown J, Beck H, Fausett L, Schneider M (1991) Neural network training on unequally represented classes. In: Intelligent engineering systems through artificial neural networks. pp 135–145
28. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
29. Han H, Wang WY, Mao BH (2005) Borderline-smote: a new over-sampling method in imbalanced data sets learning. In: International conference on intelligent computing. Springer. pp 878–887
30. He H, Bai Y, Garcia EA, Li S (2008) Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence). IEEE. pp 1322–1328
31. Barua S, Islam MM, Yao X, Murase K (2012) Mwmote-majority weighted minority oversampling technique for imbalanced data set learning. IEEE Trans Knowl Data Eng 26(2):405–425
32. Xie Z, Jiang L, Ye T, Li X (2015) A synthetic minority oversampling method based on local densities in low-dimensional space for imbalanced learning. In: International conference on database systems for advanced applications. Springer. pp 3–18
33. Douzas G, Bacao F (2017) Self-organizing map oversampling (somo) for imbalanced data set learning. Expert Syst Appl 82:40–52
34. Zhou ZH, Jiang Y (2004) Nec4. 5: neural ensemble based c4. 5. IEEE Trans Knowl Data Eng 16(6):770–773
35. Li DC, Lin YS (2006) Using virtual sample generation to build up management knowledge in the early manufacturing stages. Eur J Oper Res 175(1):413–434
36. Li DC, Fang YH (2009) A non-linearly virtual sample generation technique using group discovery and parametric equations of hypersphere. Expert Syst Appl 36(1):844–851
37. Richter SR, Vineet V, Roth S, Koltun V (2016) Playing for data: ground truth from computer games. In: European conference on computer vision. Springer. pp 102–118
38. Santana E, Hotz G (2016) Learning a driving simulator. arXiv preprint arXiv:1608.01230
39. Hodaň T, Vineet V, Gal R, Shalev E, Hanzelka J, Connell T, Urbina P, Sinha SN, Guenter B (2019) Photorealistic image synthesis for object instance

detection. In: 2019 IEEE international conference on image processing (ICIP). IEEE. pp 66–70

40. Bonechi S, Andreini P, Bianchini M, Scarselli F (2019) Coco_ts dataset: pixel—level annotations based on weak supervision for scene text segmentation. In: International conference on artificial neural networks. Springer. pp 238–250

41. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC (1998) Design and construction of a realistic digital brain phantom. IEEE Trans Med Imaging 17(3):463–468

42. Andreini P, Bonechi S, Bianchini M, Mecocci A, Scarselli F (2018) A deep learning approach to bacterial colony segmentation. In: International conference on artificial neural networks. Springer. pp 522–533

43. Chen Q, Koltun V (2017) Photographic image synthesis with cascaded refinement networks. In: Proceedings of the IEEE international conference on computer vision. pp 1511–1520

44. Karras T, Aila T, Laine S, Lehtinen J (2017) Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196

45. Liu MY, Breuel T, Kautz J (2017) Unsupervised image-to-image translation networks. In: Advances in neural information processing systems. pp 700–708

46. Liu MY, Tuzel O (2016) Coupled generative adversarial networks. In: Advances in neural information processing systems. pp 469–477

47. Yi Z, Zhang H, Tan P, Gong M (2017) Dualgan: Unsupervised dual learning for image-to-image translation. In: Proceedings of the IEEE international conference on computer vision. pp 2849–2857

48. Yeh R, Chen C, Lim TY, Hasegawa-Johnson M, Do MN (2016) Semantic image inpainting with perceptual and contextual losses. arXiv preprint arXiv:1607.07539 2(3)

49. Chen TH, Liao YH, Chuang CY, Hsu WT, Fu J, Sun M (2017) Show, adapt and tell: adversarial training of cross-domain image captioner. In: Proceedings of the IEEE international conference on computer vision. pp 521–530

50. Liang X, Hu Z, Zhang H, Gan C, Xing EP (2017) Recurrent topic-transition gan for visual paragraph generation. In: Proceedings of the IEEE international conference on computer vision. pp 3362–3371

51. Zhao W, Xu W, Yang M, Ye J, Zhao Z, Feng Y, Qiao Y (2017) Dual learning for cross-domain image captioning. In: Proceedings of the 2017 ACM on conference on information and knowledge management. pp 29–38

52. Wang X, Shrivastava A, Gupta A (2017) A-fast-rcnn: Hard positive generation via adversary for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 2606–2615

53. Andreini P, Bonechi S, Bianchini M, Mecocci A, Scarselli F (2020) Image generation by gan and style transfer for agar plate image segmentation. Comput Methods Programs Biomed 184:105268

54. Luc P, Couprie C, Chintala S, Verbeek J (2016) Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408

55. Li J, Monroe W, Shi T, Jean S, Ritter A, Jurafsky D (2017) Adversarial learning for neural dialogue generation. arXiv preprint arXiv:1701.06547

56. Ren S, Deng Y, He K, Che W (2019) Generating natural language adversarial examples through probability weighted word saliency. In: Proceedings of the 57th annual meeting of the association for computational linguistics. pp 1085–1097

57. Pascual S, Bonafonte A, Serra J (2017) Segan: speech enhancement generative adversarial network. arXiv preprint arXiv:1703.09452

58. Fiore U, De Santis A, Perla F, Zanetti P, Palmieri F (2019) Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Inform Sci 479:448–455

59. Douzas G, Bacao F (2018) Effective data generation for imbalanced learning using conditional generative adversarial networks. Expert Syst Appl 91:464–471

60. Costa P, Galdran A, Meyer MI, Niemeijer M, Abràmoff M, Mendonça AM, Campilho A (2017) End-to-end adversarial retinal image synthesis. IEEE Trans Med Imaging 37(3):781–791

61. Dai W, Dong N, Wang Z, Liang X, Zhang H, Xing EP (2018) Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer. pp 263–273

62. Gou C, Wu Y, Wang K, Wang FY, Ji Q (2016) Learning-by-synthesis for accurate eye detection. In: 2016 23rd international conference on pattern recognition (ICPR). IEEE. pp 3362–3367

63. Shin HC, Tenenholtz NA, Rogers JK, Schwarz CG, Senjem ML, Gunter JL, Andriole KP, Michalski M (2018) Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In: International workshop on simulation and synthesis in medical imaging. Springer. pp 1–11

64. Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q, Shen D (2017) Medical image synthesis with context-aware generative adversarial networks. In: International conference on medical image computing and computer-assisted intervention. Springer. pp 417–425

65. Schlegl T, Seebock P, Waldstein SM, Schmidt-Erfurth U, Langs G (2017) Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: International conference on information processing in medical imaging. Springer. pp 146–157

66. Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H (2018) Synthetic data augmentation using gan for improved liver lesion classification. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE. pp 289–293

67. Hu Y, Gibson E, Vercauteren T, Ahmed HU, Emberton M, Moore CM, Noble JA, Barratt DC (2017) Intraoperative organ motion models with an ensemble of conditional generative adversarial networks. In: International conference on medical image computing and computer-assisted intervention. Springer. pp 368–376

68. Mahapatra D, Bozorgtabar B, Hewavitharanage S, Garnavi R (2017) Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis. In: International conference on medical image computing and computer-assisted intervention. Springer. pp 382–390

69. Liu Y, Zhou Y, Liu X, Dong F, Wang C, Wang Z (2019) Wasserstein gan-based small-sample augmentation for new-generation artificial intelligence: a case study of cancer-staging data in biology. Engineering 5(1):156–163

70. Han C, Rundo L, Araki R, Nagano Y, Furukawa Y, Mauri G, Nakayama H, Hayashi H (2019) Combining noise-to-image and image-to-image gans: brain mr image augmentation for tumor detection. IEEE Access 7:156966–156977

71. Andreini P, Bonechi S, Bianchini M, Mecocci A, Scarselli F, Sodi A (2019) A two stage gan for high resolution retinal image generation and segmentation. arXiv preprint arXiv:1907.12296

72. The Alzheimer's Disease Neuroimaging Initiative (ADNI) (2020). http://www.adni-info.org/. Accessed 10 Jan 2020

73. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434

74. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167

75. Maas AL, Hannun AY, Ng AY (2013) Rectifier nonlinearities improve neural network acoustic models. In: Proc. icml. vol 30, p 3

76. Islam J (2019) Towards AI-assisted disease diagnosis: learning deep feature representations for medical image analysis. Ph.D. thesis, Georgia State University

Publisher's Note