Bmed-H
e-ISSN: 2582-7405
Transaction on Biomedical Engineering Applications and Healthcare
Volume 6, Issue 1, April 2025

Enhancement of the Quality in PET Brain Images Using *AutoGANcoder* Algorithms

Reyhane Sadat Razavi Satvati [1], Naser vosoughi [2*], Pardis Ghafarian [3], Ali Jafari [4]

[1, 2] Faculty of Energy Engineering, Sharif University of Technology of Iran (SUT), Tehran, Iran
[3] Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran
PET/CT and Cyclotron Center, Masih Daneshvari Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4] Faculty of Engineering, IRIB University, Tehran, Iran
* Corresponding Author Email: nvosoughi@sharif.edu

Abstract

The utilization of Positron Emission Tomography (PET) in disease diagnosis has been increasing in recent years. The quality of images produced by PET scanners plays a significant role in accurate diagnosis. However, these images often contain substantial noise due to photon attenuation and scatter. Therefore, PET images require attenuation correction (AC) and scatter correction (SC) to provide precise metabolic information about the patient's organs. CT-based correction methods expose the patient to significant ionizing radiation. This study focuses on improving the quality of PET scan images by utilizing Generative Adversarial Networks (GAN), a revolutionary approach in modern medical imaging, to reduce errors and patient exposure to radiation. In this study, 92 epilepsy patients with an average weight of 72.15 kg were scanned. Brain imaging was conducted on the patients following the injection of an average activity of 347.13 MBq of FDG radiotracer over a duration of 1200 seconds. These brain images served as the dataset for our designed algorithm, a GAN-based model. Image quality metrics such as SSIM, PSNR, MSE, FID, and LPIPS were measured. Our AutoGANcoder algorithm, a unique combination of a GAN and an advanced autoencoder, demonstrated that it significantly improved PET imaging quality. When compared to other algorithms, the results show that the AutoGANcoder model is a promising choice for improving brain PET image quality.

Keywords

Brain Images, Deep Learning, Generative Adversarial Network, Positron Emission Tomography.

INTRODUCTION

Positron Emission Tomography (PET) images, due to ability to display tissue and organ metabolism, are a vital tool in medical imaging and the diagnosis of various diseases, including cancer. However, these images are often affected by noise and distortions such as photon attenuation and scattering, which degrade the image quality. These issues are especially prominent in non-attenuation-corrected (NAC) images. In contrast, images that have undergone attenuation and scatter correction (AC and SC) offer greater accuracy by reducing the noise effects caused by attenuation and scattering, thus providing better clinical information. [1]

In this regard, attenuation and scatter correction Significantly contributes to refining the quality of PET images. A common method to achieve CT data to generate attenuation correction maps (AC Maps), which is considered a necessary step in producing MAC images. [2]

Although standard PET/CT scans follow the ALARA principle and use low-dose CT, the radiation risk from CT remains a concern. It has been reported that even standard low-dose CT delivers a dose of 6.4 millisieverts, exposing patients to X-rays or gamma radiation. [3] [4] [5] [6] Additionally, for patients who need to undergo frequent PET scans, as well as for children and pregnant women who are sensitive to radiation, it is preferable to minimize the amount of time spent on CT imaging. [7] [8] Reducing the duration of CT imaging can halve the radiation dose received. [9] [10]

Reconstructing PET images based on CT is limited due to CT artifacts. Over the past few years, various methods Have been implemented to enhance PET scan resolution and reduce the dose delivered to the patient. One of these methods is the use of PET/MR, which can produce an attenuation-corrected PET image while delivering a lower dose to the patient. [11] [12] [13] [14]

Nevertheless, this approach faces challenges related to classification errors both within and beyond the atlas, as well as inherent anatomical variations. The primary limitation arises from the fact that MR intensity measures tissue proton density instead of electron density, which is a crucial factor for accurate PET imaging at an energy level of 511 keV. [15]

After the emergence of deep learning algorithms, numerous Several research efforts have focused on producing artificial CT images from MR and PET scans. [16] [17] [18] [19] [20] While it is true that this method reduces the radiation dose received by the patient, converting NAC images to MAC using MRI still requires CT generation.

Due to the complexity of extracting anatomical details from images, several studies have focused on generating MAC images directly from NAC images, thereby bypassing the requirement for CT-based synthesis. [21] [22]

In recent years, deep learning (DL) techniques, with an emphasis on Convolutional Neural Networks (CNNs) and the UNET architecture, have become increasingly vital in the field of medical imaging due to their exceptional

performance and reliability. [23] [24] Deep learning techniques allow for a reduction in patient radiation exposure by bypassing the need for CT and directly deriving MAC images from NAC scans.

Building on prior research, this study employs an enhanced UNET-based Generative Adversarial Network (GAN) to transform uncorrected PET scans (NAC) into refined PET images (MAC), thereby minimizing reliance on CT imaging.

MATERIALS AND METHODS

This section provides an overview of the dataset utilized for this research, the preprocessing steps applied to the images, and the architecture of the proposed model. The following subsections provide a detailed explanation of the dataset, preprocessing procedures, and the AutoGANcoder model implemented in the study.

Datasets

A dataset consisting of NAC and MAC images from 92 patients undergoing brain PET scans was utilized for this research. Participants in the study included 43 men and 49 women, with an average height of 1.66 ± 0.12 m and an average weight of 72.15 ± 18.28 kg. The injected radiotracer was 18F-FDG with an average dose of 347.13 MBq.

Image Preprocessing

To prepare the input data for the model, the images were first normalized to values between 0 and 1 by scaling the image with respect to its peak activity value. Then, the images were fed into the model in a 2D format with a size of 256×256.

AutoGANcoder

The designed AutoGANcoder framework includes two primary elements: the Generator and the Discriminator, which are used for training the model, as illustrated in Figure 1.

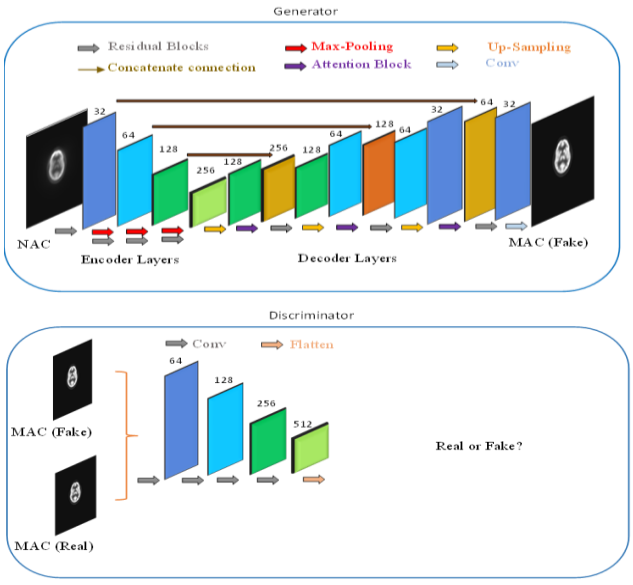

Figure 1. Architecture of the Generator and Discriminator in the AutoGANcoder model.

In the first stage, NAC images are fed into the Generator, which reconstructs the generated MAC image with the help of Residual Blocks and Attention Blocks. The Generator consists of four encoding stages (Encoder) and three decoding stages (Decoder). Each decoding stage accesses information from the corresponding encoding stage using the Attention Mechanism, allowing for more effective decoding of the information.

In the Generator, an improved UNet Autoencoder network is used, which, unlike previous studies that employed standard UNet or ResNet, utilizes a combination of Residual Blocks and Attention Blocks to improve accuracy and convergence speed. This combination enables the model to improve the resolution of generated images by focusing on important regions.

In the Discriminator section, a Conv2D network is used, which learns to differentiate actual MAC images from those produced by the Generator.

Loss Functions

For training the GAN model, three types of loss functions are used:

Mean Squared Error (MSE) Loss Function

MSE loss evaluates the squared deviation between the generated image and the reference image.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2 \quad (1)$$

Where x_i is the real MAC image, $\hat{x}_i$ is the output image produced by the Generator, and N is the number of samples.

Wasserstein GAN (WGAN) Loss Function

The WGAN function analyzes the distribution of actual and synthetic images, pushing the Generator to create images that more closely resemble the real ones. This function is defined for the Discriminator using the following formula:

$$L_D = E[D(x)] - E\big[D\big(G(x)\big)\big] \quad (2)$$

In this case, $D(x)$ refers to the Discriminator's output for real samples, while $D\big(G(x)\big)$ corresponds to its output for generated ones.

The WGAN loss function for the Generator is defined by the following formula:

$$L_G = -E\big[D\big(G(x)\big)\big] \quad (3)$$

Through this loss function, the Generator seeks to optimize the Discriminator's response to the generated images.

Combination of Loss Functions for the Generator:

In our algorithm, By integrating MSE and WGAN Loss functions, the Generator ensures improved output quality. This combination helps the Generator ensure that the generated images not only resemble real images but also match them in terms of structure and details. The combined loss function for the Generator is expressed as follows:

$$L_G = MSE + \lambda * WGAN \quad (4)$$

Bmed-H
e-ISSN: 2582-7405
Transaction on Biomedical Engineering Applications and Healthcare
Volume 6, Issue 1, April 2025

Where $\lambda = 10^{-3}$ is the weighting factor for the **WGAN Loss** to reduce its impact relative to the **MSE Loss**.

In this study, PSNR, SSIM and MSE metrics are applied to quantitatively analyze both the generated and original MAC-PET slices. For further validation, perceptual indicators such as FID and LPIPS have been incorporated into the analysis. FID calculates the disparity in feature distributions between real and generated images through the Inception network. Lower FID scores suggest improved similarity and higher-quality synthetic images. Utilizing features from pre-trained models, LPIPS measures perceptual discrepancies across images. While conventional methods such as MSE focus on pixel-based differences, LPIPS examines perceptual dissimilarities between images by employing features obtained from pre-trained neural networks. A reduced LPIPS score denotes greater perceptual similarity between the compared images. The computation procedures for FID and LPIPS are detailed below:

$$FID = \|\mu_r - \mu_g\|_2^2 + Tr(\sum_r + \sum_g - 2\sqrt{\sum_r \sum_g}) \quad (5)$$

In this context, μ_r and $\sum_r$ correspond to the mean and covariance values of real image features, while μ_g and $\sum_g$ signify the mean and covariance for generated images. The Euclidean distance $\|\mu_r - \mu_g\|_2^2$ quantifies how far apart the mean values of the two distributions are, with Tr indicating the trace, or the aggregate of diagonal elements.

$$LPIPS(x_1, x_2) = \sum_l \frac{1}{H_l W_l} \sum_{h,\omega} \|\phi_l(x_1)_{h,\omega} - \phi_l(x_2)_{h,\omega}\|_2^2 \quad (6)$$

The function ϕ_l represents the extracted features at layer l within the neural network; x_1, x_2 act as the input images, and H_l and W_l indicate their respective height and width.

EXPERIMENTS AND RESULTS

We compared the proposed AutoGANcoder with the advanced UNet model [24], whose architecture is illustrated in Figure 2. This UNet model is specifically designed to improve PET image quality or generate medical imaging data. The advanced UNet model is composed of encoder and decoder layers, incorporating an addition operation at the final stage that merges the source image with the generated one. To assess the influence of various components within the UNet model, the addition operation was omitted, and the model was trained based on the structure outlined in Figure 3. This enables a comparative analysis of the obtained performance, as detailed in the 'UNet without addition' section of Table 1.

Additionally, for a more thorough evaluation of the AutoGANcoder model and to demonstrate the effectiveness of the GAN architecture, we removed the Discriminator component and obtained the results. These results are presented in the " AutoGANcoder without Discriminator " column of Table 1.

In the UNet framework, the L1Loss function served as the objective function for model training, while the Adam optimizer, with a learning rate of 0.0002, was utilized to enhance optimization efficiency

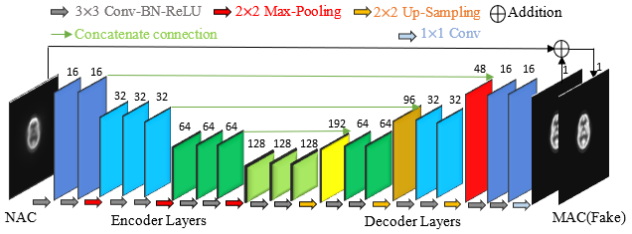

Figure 2. The architecture of the implemented UNet model [24].

To evaluate the impact of different parts of the UNet model, we removed the Addition section and trained the model as shown in Figure 3, in order to compare the resulting outcomes.

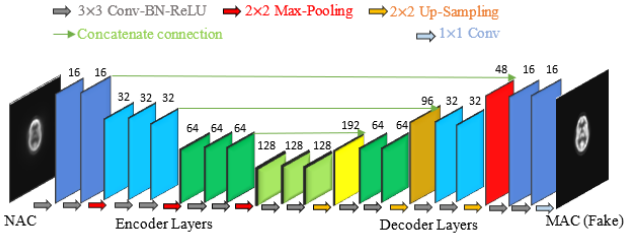

Figure 3. Architecture of the UNet model without Addition.

To prevent overfitting, the dataset was partitioned into 70% for training, 20% for validation, and 10% for testing purposes.

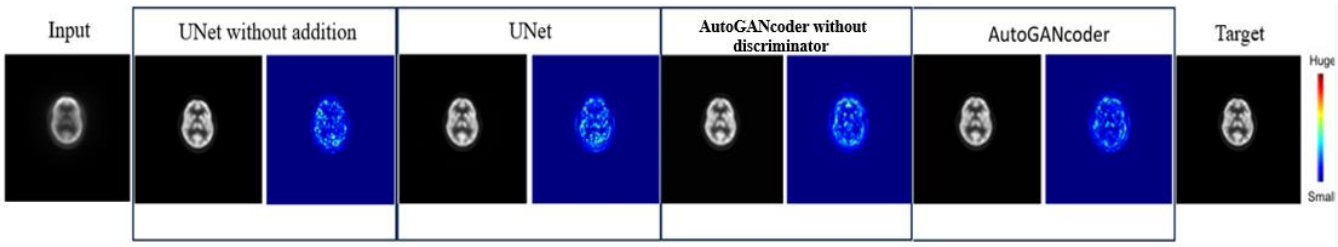

Figure 4. A comparative analysis of pseudo-color difference maps illustrating the variations between real and generated F-PET images using various methods, including our proposed approach. The magnitude of absolute differences ranges from minor to significant, visually represented by a color gradient transitioning from blue to red.

Table 1. Comparative analysis of AutoGANcoder and alternative methods for enhancing the quality of PET images

METHOD / *METRICS*	*UNet without addition*	*UNet[24]*	**Autoencoder (AutoGANcoder without discriminator)**	**AutoGANcoder**
SSIM↑	0.9960	0.9965	0.9974	**0.9982**
PSNR↑	45.80	46.45	47.75	**49.43**
MSE↓	2.7×10^{-5}	2.3×10^{-5}	1.7×10^{-5}	$\mathbf{1.2 \times 10^{-5}}$
FID↓	0.4592	0.4982	0.5154	**0.4620**
LPIPS↓	0.0062	0.0061	0.0024	**0.0016**

DISCUSSION

The results presented in Table 1 demonstrate that AutoGANcoder significantly outperforms other models in terms of PSNR (49.4374), SSIM (0.9982), and MSE (0.0000123089), indicating a higher resemblance of the generated PET images to MAC. Additionally, the low FID (0.4620) and LPIPS (0.0016) values further highlight the effectiveness of the proposed approach in generating high-quality images. These findings align with previous research emphasizing the advantages of generative models in medical image reconstruction. [25] [26] [27]

A key factor contributing to AutoGANcoder's superior performance is its advanced architecture, which incorporates Residual Blocks and Attention Blocks within both the encoder and decoder. These components enhance the feature extraction and image reconstruction processes, leading to improved image fidelity. Previous studies have also demonstrated the effectiveness of attention mechanisms in enhancing deep learning-based medical image synthesis [Add relevant reference].

Ablation studies, as shown in Table 1, provide further insights into the contribution of different architectural components. When the Discriminator is removed from AutoGANcoder, the generated image quality decreases significantly, underscoring the vital role of adversarial training in improving realism. This finding corroborates previous works demonstrating the necessity of the Discriminator in GAN-based architectures for medical imaging. [28] [29] [30] [31] [32] [33] [34]

Furthermore, the comparison between the standard UNet and the "UNet without addition" variant indicates that the residual connections within UNet contribute significantly to enhancing image quality. The removal of these connections leads to inferior MAC reconstructions, suggesting that the residual structure preserves critical spatial and contextual information during the encoding and decoding process. Similar observations have been reported in recent studies on deep learning-based medical image enhancement. [24]

Figure 4 illustrates pseudo-color difference maps comparing real F-PET images with generated ones, where AutoGANcoder demonstrates the lowest discrepancy. This outcome further supports the robustness of the model in preserving fine-grained details and reducing reconstruction errors.

CONCLUSION

In this study, we introduce AutoGANcoder as an effective approach to enhance PET image quality and generate MAC images from NAC images. Our proposed architecture achieves remarkable results across the mentioned metrics, demonstrating that combining UNET with Residual and Attention blocks within a GAN framework is highly effective.

To improve the AutoGANcoder model, it is suggested to incorporate the addition operation into its architecture. Additionally, increasing the number of layers and introducing multiple loss functions can further reduce the model's error rate.

REFERENCES

[1]. Y. Li and W. Wu, "A deep learning-based approach for direct PET attenuation correction using Wasserstein generative adversarial network," presented at the Journal of Physics, 2021.

[2]. R. D. Badawi et al., "First Human Imaging Studies with the EXPLORER Total-Body PET Scanner," J Nucl Med, vol. 60, no. 3, pp. 299-303, Mar 2019, doi: 10.2967/jnumed.119.226498.

[3]. D. Z. Solomon, B. Ayalew, S. T. Dellie, and D. Admasie, "Justification and optimization principles of ALARA in pediatric CT at a teaching hospital in Ethiopia," Ethiopian Journal of Health Sciences, vol. 30, no. 5, 2020.

[4]. Z. Huang et al., "Deep cascade residual networks (DCRNs): optimizing an encoder–decoder convolutional neural network for low-dose CT imaging," IEEE Transactions on Radiation and Plasma Medical Sciences, vol. 6, no. 8, pp. 829-840, 2022.

[5]. X. Dong et al., "Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging," Physics in Medicine & Biology, vol. 65, no. 5, p. 055011, 2020.

[6]. E. Prieto et al., "Ultra-low dose whole-body CT for attenuation correction in a dual tracer PET/CT protocol for multiple myeloma," Physica Medica, vol. 84, pp. 1-9, 2021.

[7]. D. J. Kadrmas and J. M. Hoffman, "Methodology for quantitative rapid multi-tracer PET tumor characterizations," Theranostics, vol. 3, no. 10, p. 757, 2013.

[8]. J. H. Thrall, "Radiation exposure in CT scanning and risk: where are we?," vol. 264, ed: Radiological Society of North America, Inc., 2012, pp. 325-328.
[9]. A. Kaushik et al., "Estimation of patient dose in 18F-FDG and 18F-FDOPA PET/CT examinations," Journal of Cancer Research and Therapeutics, vol. 9, no. 3, pp. 477-483, 2013.
[10]. K. Khamwan, A. Krisanachinda, and P. Pasawang, "The determination of patient dose from 18F-FDG PET/CT examination," Radiation Protection Dosimetry, vol. 141, no. 1, pp. 50-55, 2010.
[11]. S. Ahangari et al., "A deep learning-based whole-body solution for PET/MRI attenuation correction," EJNMMI Physics, vol. 9, p. 55, 2022, doi: https://doi.org/10.1186/s40658-022-00486-8.
[12]. G. Wagenknecht, H.-J. Kaiser, F. M. Mottaghy, and H. Herzog, "MRI for attenuation correction in PET: methods and challenges," Magnetic Resonance Materials in Physics, Biology and Medicine, vol. 26, pp. 99-113, 2013.
[13]. J. H. Kim, J. S. Lee, I.-C. Song, and D. S. Lee, "Comparison of segmentation-based attenuation correction methods for PET/MRI: evaluation of bone and liver standardized uptake value with oncologic PET/CT data," Journal of Nuclear Medicine, vol. 53, no. 12, pp. 1878-1882, 2012.
[14]. S.-H. Hsu, Y. Cao, K. Huang, M. Feng, and J. M. Balter, "Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy," Physics in Medicine & Biology, vol. 58, no. 23, p. 8419, 2013.
[15]. I. Shiri et al., "Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC)," European Radiology, vol. 29, pp. 6867–6879, 2019, doi: https://doi.org/10.1007/s00330-019-06229-1.
[16]. H. Arabi, G. Zeng, G. Zheng, and H. Zaidi, "Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI," European Journal of Nuclear Medicine and Molecular Imaging, vol. 46, pp. 2746-2759, 2019.
[17]. A. Bahrami, A. Karimian, E. Fatemizadeh, H. Arabi, and H. Zaidi, "A new deep convolutional neural network design with efficient learning capability: Application to CT image synthesis from MRI," Medical Physics, vol. 47, no. 10, pp. 5158-5171, 2020.
[18]. X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method," Medical Physics, vol. 44, no. 4, pp. 1408-1419, 2017.
[19]. A. Bahrami, A. Karimian, and H. Arabi, "Comparison of different deep learning architectures for synthetic CT generation from MR images," Physica Medica, vol. 90, pp. 99-107, 2021.
[20]. Y. Lei et al., "PET attenuation correction using non-AC PET-based synthetic CT," in Medical Imaging 2020: Physics of Medical Imaging, vol. 11312: SPIE, pp. 1057-1063, 2020.
[21]. C. N. Ladefoged, L. Marner, A. Hindsholm, I. Law, L. Højgaard, and F. L. Andersen, "Deep learning based attenuation correction of PET/MRI in pediatric brain tumor patients: evaluation in a clinical setting," Frontiers in Neuroscience, vol. 12, p. 1005, 2019.
[22]. H. Arabi and H. Zaidi, "Truncation compensation and metallic dental implant artefact reduction in PET/MRI attenuation correction using deep learning-based object completion," Physics in Medicine & Biology, vol. 65, no. 19, p. 195002, 2020.
[23]. R. Jahangir, A. Kamali-Asl, and H. Arabi, "Deep Learning-Based Attenuation and Scatter Correction of Brain 18F - FDG PET Images in the Image Domain," arXiv, 2022, doi: https://doi.org/10.48550/arXiv.2206.14673.
[24]. K. T. Chen et al., "Ultra–Low-Dose 18 F-Florbetaben Amyloid PET Imaging Using Deep Learning with Multi-Contrast MRI Inputs," Radiology, vol. 00, pp. 1–8, 2019, doi: https://doi.org/10.1148/radiol.2018180940.
[25]. T.-A. Song, S. R. Chowdhury, F. Yang, and J. Dutta, "PET image super-resolution using generative adversarial networks," *Neural Networks*, vol. 125, pp. 83-91, 2020.
[26]. K. Chen *et al.*, "Generative Adversarial Network–Enhanced Ultra-Low-Dose [18F]-PI-2620 τ PET/MRI in Aging and Neurodegenerative Populations," *American Journal of Neuroradiology*, vol. 44, no. 9, pp. 1012-1019, 2023.
[27]. L. Bi, J. Kim, A. Kumar, D. Feng, and M. Fulham, "Synthesis of positron emission tomography (PET) images via multi-channel generative adversarial networks (GANs)," in Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment: Fifth International Workshop, CMMI 2017, Second International Workshop, RAMBO 2017, and First International Workshop, SWITCH 2017, Held in Conjunction with MICCAI 2017, QuÉBec City, QC, Canada, September 14, 2017, Proceedings 5, 2017: Springer, pp. 43-51.
[28]. J. Zhao, X. Hou, M. Pan, and H. Zhang, "Attention-based generative adversarial network in medical imaging: A narrative review," *Computers in Biology and Medicine*, vol. 149, p. 105948, 2022.
[29]. J. Fu, W. Li, J. Du, and L. Xu, "DSAGAN: A generative adversarial network based on dual-stream attention mechanism for anatomical and functional image fusion," *Information Sciences*, vol. 576, pp. 484-506, 2021.
[30]. Y. Xue, Y. Zhang, and F. Neri, "A method based on evolutionary algorithms and channel attention mechanism to enhance cycle generative adversarial network performance for image translation," *International Journal of Neural Systems*, vol. 33, no. 05, p. 2350026, 2023.
[31]. S. Kazeminia *et al.*, "GANs for medical image analysis," *Artificial intelligence in medicine*, vol. 109, p. 101938, 2020.
[32]. J. Huang, Z. Le, Y. Ma, F. Fan, H. Zhang, and L. Yang, "MGMDcGAN: medical image fusion using multi-generator multi-discriminator conditional generative adversarial network," *IEEE Access*, vol. 8, pp. 55145-55157, 2020.
[33]. L. Huang, H. Jiang, S. Li, Z. Bai, and J. Zhang, "Two stage residual CNN for texture denoising and structure enhancement on low dose CT image," *Computer methods and programs in biomedicine*, vol. 184, p. 105115, 2020.
[34]. Y. Zhang, J. Chi, C. Wu, and X. Yu, "Deep residual network based medical image reconstruction," in *2019 Chinese Control Conference (CCC)*, 2019: IEEE, pp. 8550-8555.