Vol. 12-2 (2025) 47–59
International Journal of Computers and Information (IJCI)
Available online at https://ijci.journals.ekb.eg/

Enhanced Gan-Based Data Augmentation for Multimodal Medical Images Registration

Sameh Zarif [a*], Khalid Amin [b], Maryam George [c], Mohiy Hadhoud [d], Marian Wagdy [e]

a Information Technology Department Menofia University, Egyptian Russian University, sameh.shenoda@ci.menofia.edu.eg, sameh-zarief@eru.edu.eg

b Information Technology Department Menofia University, Menofia, EGYPT, kh.amin.0.0@gmail.com

c Information Technology Department Menofia University, Menofia, EGYPT mariam_it@ci.menofia.edu.eg

d Information Technology Department Menofia University, Menofia, EGYPT, mmhadhoud@ci.menofia.edu.eg

e Information Technology Department Tanta University, Tanta, EGYPT, marian_wagdy@ics.tanta.edu.eg

Abstract

Data augmentation is a crucial technique for enhancing the generalization of deep learning registration models, especially in the medical imaging domain, where high-quality and diverse multimodal data are often scarce. Prior multimodal registration approaches faced multiple limitations, such as intricate implementation processes and overfitting, which reduce the generalizability of the models and impact registration accuracy. These limitations underscore the need for improved methodologies to enhance the effectiveness of medical image analysis by advancing the progressive GAN framework that synthesizes high-quality multimodal medical images. In this study, we propose an approach based on Generative Adversarial Networks (GANs) to improve the quality and diversity of multimodal medical images. Our methodology includes preprocessing steps, real-time modifications during GAN training, and post-processing techniques to enhance the generated images. The results demonstrate that the proposed method outperforms traditional registration techniques, achieving a mean Dice Similarity Coefficient of 0.78, indicating a significant improvement in registration accuracy. These findings support the potential application of our approach in clinical settings, enhancing the effectiveness of medical image analysis.

Keywords: Gan Augmentation, Image Registration, Multimodal Images, Unsupervised Learning, Medical Image Analysis.

1. Introduction

Data augmentation is a technique used to increase the diversity of the training dataset by applying various transformations to the original images, such as rotation, scaling, and adding synthetic images [1] ,which helps improve the robustness, accuracy, and generalization capabilities of deep learning models in medical image analysis [2]. This technique broadens the model's ability to identify a wide array of features within augmented datasets, as opposed to relying on a limited feature set for object recognition in images [3]. In medical image registration, there are several challenging issues; a prevalent problem in this domain is the scarcity of normal data, as hospital-collected datasets predominantly consist of abnormal cases. Additionally, the propensity for overfitting caused by excessively complex models limits their ability to generalize to new data. Furthermore, problems such as inadequate training datasets, lack of accurate ground truth annotations, and susceptibility to adversarial attacks present considerable hurdles [4]. Given these challenges, there is a pressing need for data augmentation, particularly concerning paired multimodal medical images, as the limited diversity and quantity of such images severely constrain the effectiveness of contemporary deep learning techniques [5].

Synthesized data holds potential for enhancing classification, segmentation, and registration tasks involving multimodal MRI images [6]. In the medical field, the rarity of images related to uncommon diseases restricts the training capabilities of neural networks, thereby adversely affecting classification outcomes. Therefore, there is an urgent need for research focused on improving augmented and coherent data, establishing necessary evaluation standards, and increasing access to larger annotated datasets [7]. Data augmentation addresses these issues by significantly enhancing the clarity and resolution of medical images, thereby reducing blurriness in CT and MRI scans [8] , resulting in better outcomes at lower costs. This paper is based on integrating traditional augmentation techniques with GAN which can significantly improve the generation of synthetic images. This can be achieved through three main approaches: First, preprocessing augmentation involves applying methods such as rotation, scaling, and upsampling to the training dataset, enhancing input diversity. Second, augmentation incorporates real-time modifications during GAN training, including random cropping to focus on various image segments, adding Gaussian noise to introduce variability [9], and applying elastic transformations to simulate realistic deformations in medical images. Finally, post-processing augmentation enhances the generated synthetic images through traditional techniques, thereby increasing their variability and robustness. This research aims to synthesize high-quality and clinically meaningful multimodal medical images, ensuring that GANs learn from a diverse dataset and achieve improved generalization across different conditions for better registration.

The paper is organized as follows: Section 2 describes related work, Section 3 presents the proposed work, Section 4 introduces experimental results, and Section 5 concludes the paper.

2. Related works

2.1. GANs in Data Augmentation

In recent years, several papers have explored the use of Generative Adversarial Networks (GANs) for medical image augmentation and synthesis, each presenting unique advantages and disadvantages. For instance, DAGAN, proposed in [10], primarily focuses on data augmentation for medical imaging. While it successfully generates synthetic images, it faces challenges regarding the quality of these images. Additionally, its evaluation scope is limited, which may restrict its effectiveness across diverse medical applications. In [11] a biomedical data augmentation method using GANs has been demonstrated, generating high-quality biomedical images. However, its focus is predominantly on brain images, limiting the diversity of the data and potentially hindering the generalizability of its findings to other modalities. Similarly, the authors in [12] addressed medical image synthesis and data augmentation utilizing public datasets for image generation. While this method offers some benefits, relying on such datasets may restrict the **diversity** of the generated images, which is essential for training robust models across different medical imaging tasks. In [13], a study on generative adversarial networks for medical imaging provided practical code implementations, showcasing the accessibility of GANs for generating synthetic medical images, though it might face challenges in terms of generalizability to different medical imaging tasks. In the context of few-shot learning, the authors in [14] proposed a Few-shot 3D multimodal medical image segmentation using generative adversarial learning, effectively handling multimodal segmentation tasks. However, it risks overfitting due to the limited number of training samples, which may affect its performance on unseen data.The research on [15] introduced 3D conditional GANs for PET image estimation, providing high-quality estimations of PET images while reducing radiation exposure. However, it is heavily dependent on conditional inputs, which may limit its flexibility in practical applications. The study in [16] addressed bimodal medical image synthesis using semi-supervised GANs, offering high-quality synthetic images but being complex to implement, which may hinder its adoption in clinical settings and lead to challenges in generalizability across different datasets. The Vox2Vox method in [17] targeted enhancing training data for brain tumor segmentation but did not extensively address the evaluation metrics, which could impact the assessment of model performance. The research in [18] introduced multi-contrast MRI images, providing enhanced data diversity but facing challenges in evaluation metrics and the complexity of multi-contrast imaging.

2.2. GANs in Multimodal Registration Techniques

The application of GANs extends beyond augmentation to multimodal medical image registration. Recent advancements have introduced various techniques aimed at improving registration accuracy. For example , The introduction of 3D-StyleGAN in [19] allowed for high-quality 3D image generation but was computationally intensive and had limited practical application examples. Despite their potential, artificial data often loses its physiological structure, especially in ultrasound images. Various techniques have been developed to address these challenges, such as using Radon Transform to synthesize CT datasets (low-dose x-ray tomography through a deep convolutional neural network), employing autoencoders for PET image

reconstruction (deep reconstruction model for dynamic PET images), and implementing stochastic discriminator augmentation to prevent overfitting when training with limited data (training generative adversarial networks with limited data)[20]. Recent methods like the Progressive Generative Adversarial Method (PGAM) and Progressive Texture Generative Adversarial Network (PTGAN) focus on maintaining structural integrity while generating medical images, providing solutions for lesion repair and synthesis through mask-reconstruction strategies. The study in [21] proposed progressive GANs that improved model training for structurally inadequate datasets but had a complex training process and a risk of mode collapse.
To overcome these limitations, this paper presents a strategy that addresses the challenges associated with augmentation, seeking to achieve improved results in multimodal image registration through the incorporation and enhancement of data augmentation techniques.

3. Proposed Work

The related work in medical image augmentation and synthesis using Generative Adversarial Networks (GANs) reveals several limitations, including challenges with the quality of synthetic images, a limited evaluation scope focused on specific modalities, and complex implementation processes that hinder clinical adoption. Additionally, reliance on public datasets restricts the diversity of generated images, while issues like mode collapse and overfitting compromise the generalizability of models [22] as mentioned in the previous section. These limitations underscore the need for improved methodologies to enhance the effectiveness of medical image analysis by advancing the progressive GAN framework that synthesizes high-quality multimodal medical images, which improves the performance of multimodal medical image registration, as illustrated in Fig. 1.

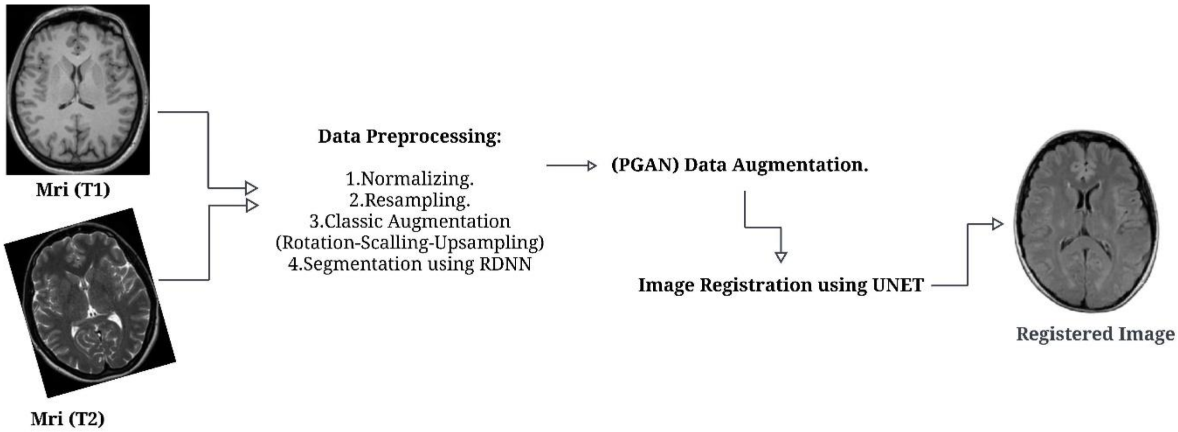

Fig. 1 Proposed Multimodal Medical Image Registration Approach

3.1 Data Pre-processing

This paper used the IXI dataset [23], which includes 1,100 MRI images of a single patient in both T1 and T2 modalities, and the RIRE dataset [24] which consists of multimodal (PET-MRI-CT) images from a single patient, for the testing stage. Data preprocessing consists of four steps, each described in detail below.

Normalizing MRI images is the first step to ensure consistent intensity values across the dataset, as shown in Equation 1:

$$I_{norm}=(I-\mu)/\sigma \quad (1)$$

where I is the original image, μ is the mean intensity, and σ is the standard deviation of the intensities. Fig. 2 illustrates an example of image normalization.

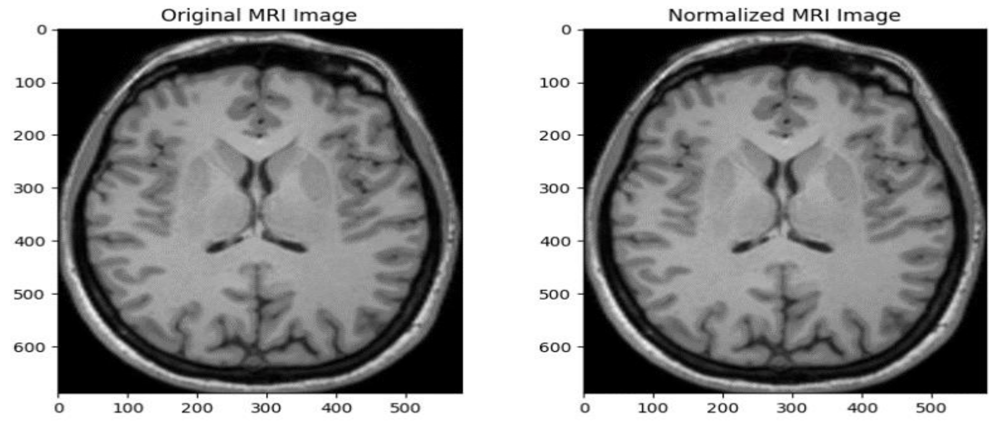

Fig. 2 Normalized MRI Image Enhanced Uniformity and Clarity of Anatomical Structures.

Resampling the images using B-spline interpolation is the second step. This is a crucial step to ensure that the input to the RDNN (Recurrent Deep Neural Network) model is uniform. An example of image sampling is illustrated in Fig. 3.

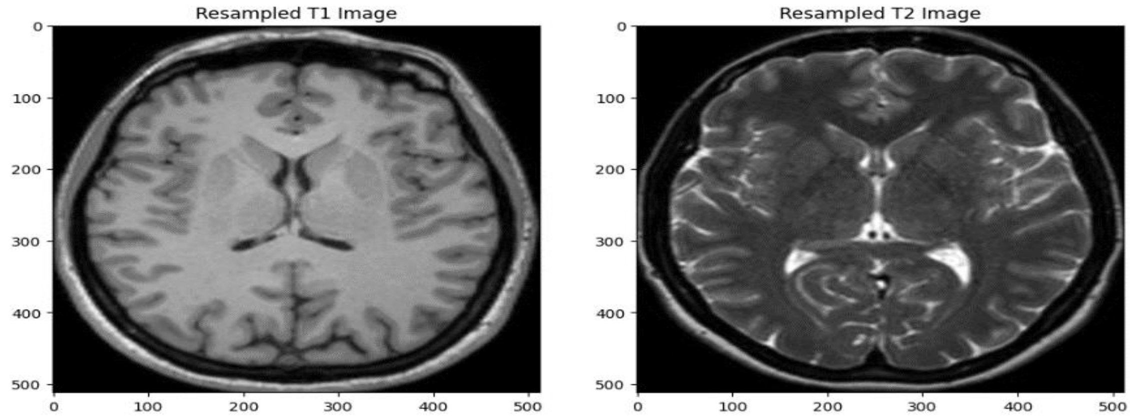

Fig. 3 Image resampling which enhanced detail through B-Spline interpolation.

Classic augmentation is the third step. Classic augmentation methods, such as scaling, rotation, and downsampling, are utilized on the 3D MRI and CT images, as demonstrated in Fig. 4.

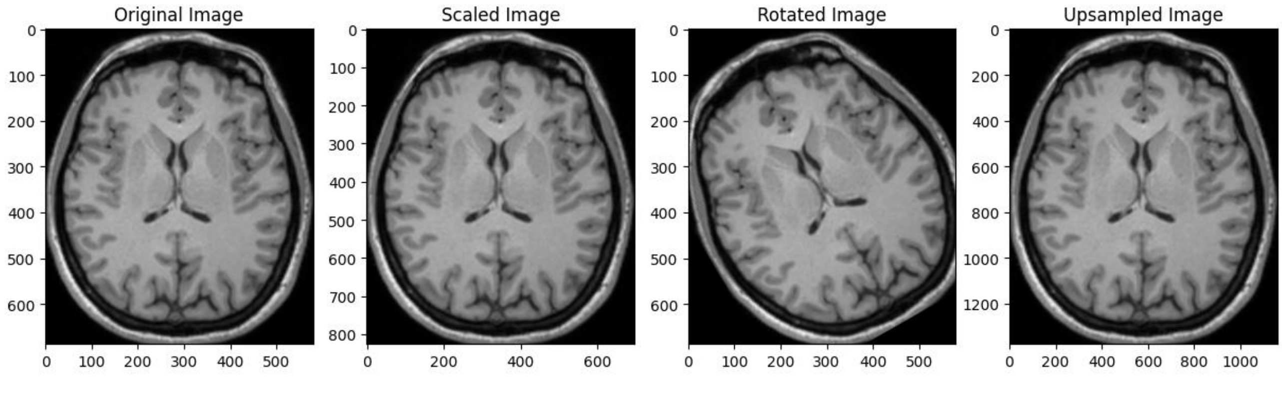

Fig. 4 Classic augmentation methods which enhancing dataset diversity for improved learning.

Segmentation is the fourth step. Extracting anatomical and functional details from the IXI dataset is performed using RDNN (Recurrent Deep Neural Network) due to its effectiveness in extracting functional and anatomical information from complicated medical datasets, including the IXI and MRI datasets.

The ability of RDNNs to recognize complex patterns and characteristics in the data is essential for precisely identifying various structures in medical imaging. Because of this feature, RDNNs are a good option for the proposed work in this study. The RDNN takes 3D medical images as input, and the input size is specified as (160 × 192 × 224 × 2). The loss function used in training can be defined by Equation 2:

$$L(y, \grave{y}) = -\sum_{i=1}^{N} y_i \log(\grave{y}_i) \quad (2)$$

where y is the true label and ỳ is the predicted label. Fig. 5 shows an example of image segmentation by RDNN.

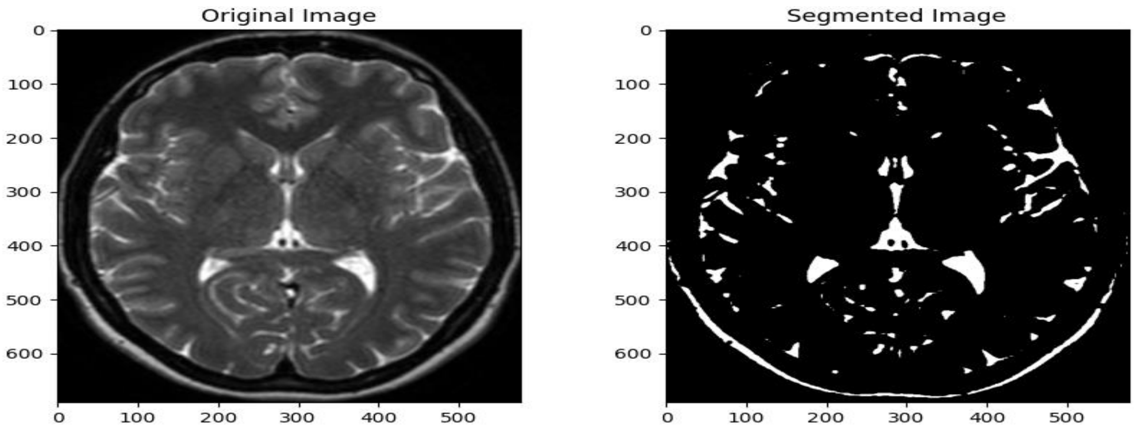

Fig. 5 Example of image segmentation by RDNN.

RDNNs maintain past inputs, known as the hidden state, which allows the network to consider the context of previous inputs, retaining crucial information within the image.

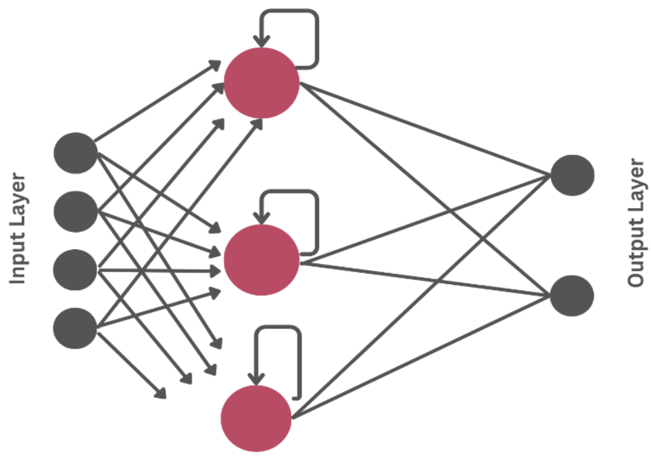

Fig. 6 Recurrent Deep Neural Network (RDNN) Architecture.

3.2 Synthetic Data Augmentation with PGAN

Progressive Growing of GANs (PGGANs) is a new technique for training GANs. It involves gradually increasing the resolution of both the generator and discriminator networks. This process starts with low-resolution images and then adds layers to capture finer details as training continues, contributing to the creation of more realistic medical images, as shown in Fig. 7.

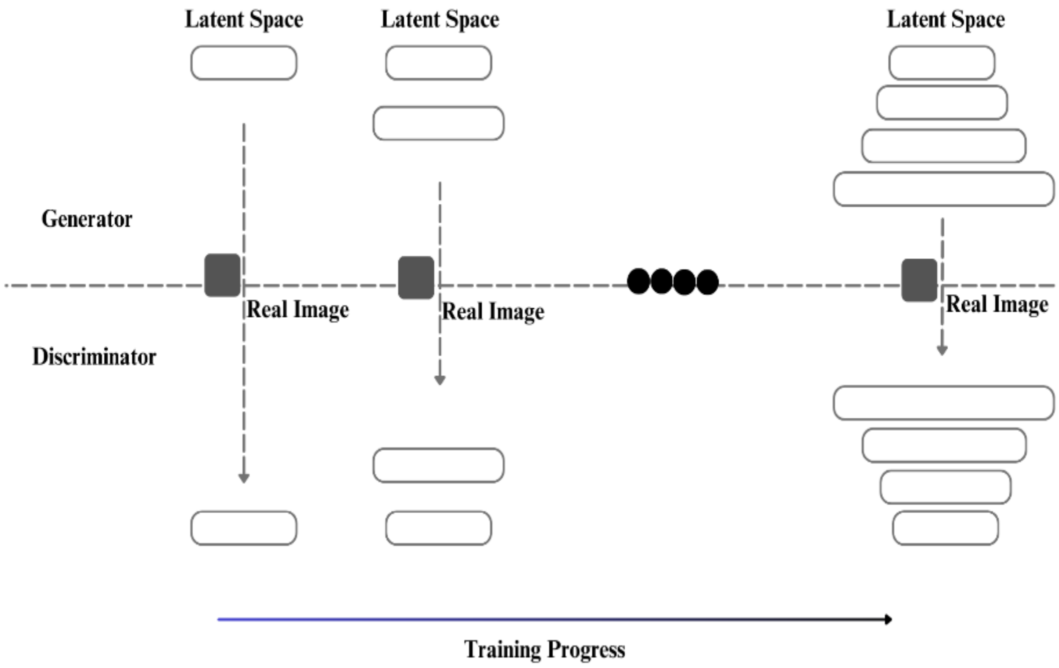

Fig. 7 PGAN architecture for synthetic image generation.

After extracting anatomical and functional details from the IXI dataset using RDNN, data augmentation is applied using PGAN. The RDNN's output, in the form of segmented images, will serve as a reference for the PGAN's augmentation process. The PGAN leverages these segmented outputs and uses them with the same output size to generate new synthetic images that preserve the anatomical structures identified by the RDNN. By learning from these segmented outputs, the PGAN produces synthetic samples that are consistent with the underlying anatomical features, thereby enriching the diversity and quality of the training dataset for subsequent medical image analysis tasks, as shown in Fig. 8.

The training set consists of 2,400 images, the validation set contains 800 images, and the test set includes 800 images. The augmentation process achieves 4,000 pairs of T1 and T2 modalities.
Generating New Samples: Use the generator part of PGAN to create new MRI images based on the learned distribution of the training data.
Progressive Growing: Start training with low-resolution images and progressively increase the resolution, which helps stabilize the training process. The generator can be defined as in equation 3:

$$G(z)=\mathrm{ConvTranspose}(z) \tag{3}$$

where z is the latent vector sampled from a Gaussian distribution.

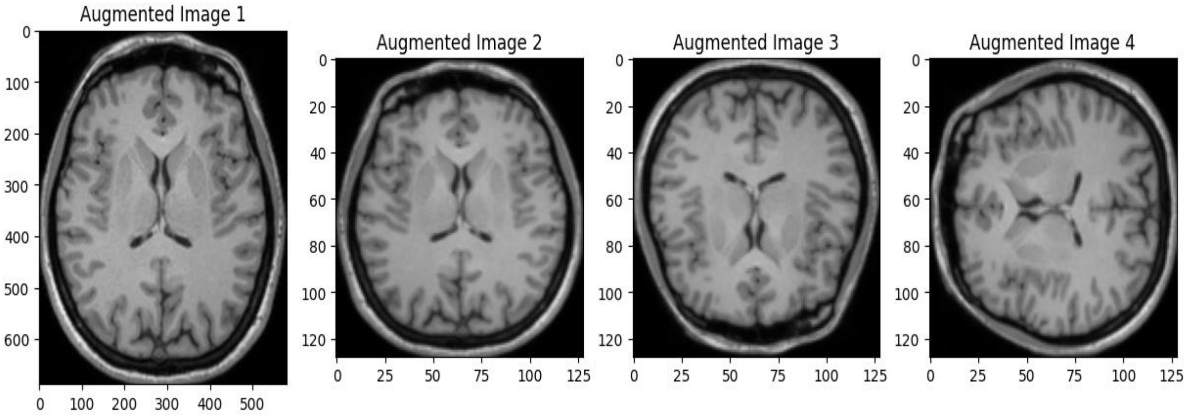

Fig. 8 PGAN generating synthetic detailed medical images output.

By combining synthetic and classic data augmentation techniques, this approach demonstrates superior performance in medical imaging, enabling a more robust training process. Furthermore, when trained on a dataset that maintains the important segmented features, PGAN can generate images suitable for medical image registration tasks, producing highly realistic, high-resolution images that surpass the limitations of earlier GAN models.

By employing a multi-stage training approach, PGGANs gradually increase image resolution, leading to more stable training and significantly improved image quality. This ability to create realistic images, especially in fields like medical imaging, is invaluable for applications such as image registration.

3.3 Image Registration using UNET

The proposed registration architecture, illustrated in Fig. 9, is inspired by U-Net. It utilizes an encoder-decoder design with skip connections, allowing flexibility in input dimensions, although the experiments use a specific input size of ($160 \times 192 \times 224 \times 2$). Both encoder and decoder stages employ 3D convolutions with a 3x3 kernel and a stride of 2, followed by Leaky ReLU activations. The encoder reduces spatial dimensions to capture global deformations, akin to traditional image pyramid techniques. The decoder combines upsampling, convolutions, and skip connections to integrate encoded features, refining the spatial scale for accurate anatomical alignment.

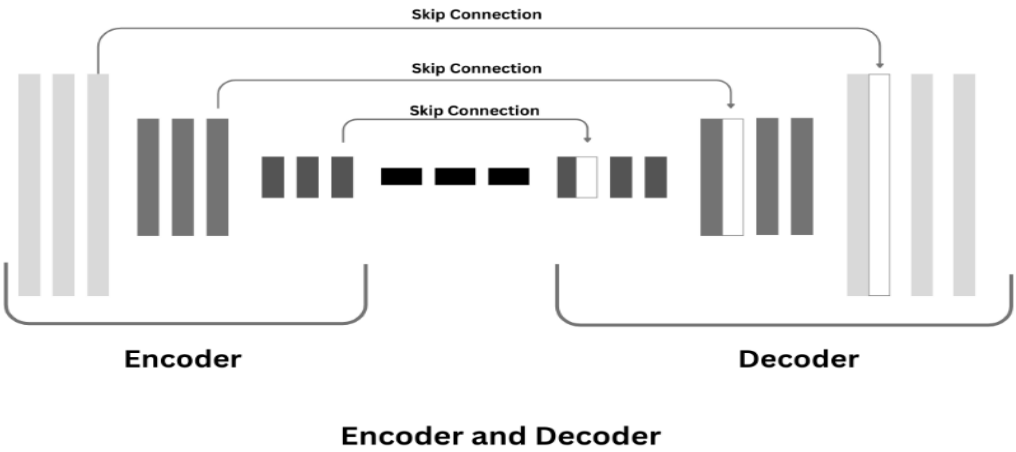

Fig. 9 UNet architecture for medical image segmentation: A dual pathway model utilizing encoder-decoder structure with skip connections for enhanced feature extraction and high-resolution reconstruction.

Loss Function

The loss function for training the network is formulated as:

$$L(I_f, I_m, \phi)) = Lsim(I_f, I_m, \phi) + R(\phi) \tag{4}$$

Here, L_{sim} measures the similarity between the deformed moving image (I_m) and the fixed image (I_f), while R is a regularization term that enforces smoothness in the deformation field.

To ensure realistic and smooth deformations, regularization terms are added. Diffusion Regularization (R) and Bending Energy are two examples presented in equations 7 and 8 where ∇u(p) is Gradient of the deformation field at position p.

$$R_{diffusion}(\phi) = \sum_{p\in\Omega} \|\nabla u(p)\|^2 \tag{5}$$

$$R_{bending}(\phi) = \sum_{p\in\Omega} \|\nabla^2 u(p)\|^2 \tag{6}$$

When anatomical segmentations are available, an additional loss function L_{seg} quantifying the overlap between segmentations is incorporated as shown in equation 9 as γ_{seg} (S_f–$_m$,Φ): Auxiliary penalty term.

$$L(I_f, I_m, \phi) = L_{sim}(I_f, I_m, \phi) + R(\phi) + \gamma_{seg}(S_f, S_m, \phi) \tag{7}$$

The method extends to include Bayesian uncertainty through Monte Carlo dropout, allowing the estimation of predictive variances for both transformation and appearance uncertainties as shown in equations 10 and 11 as u_t : Transformation output at sample T and $\bar{u}$ Mean of the transformation outputs across T samples and $\sigma^2_{transformation}$ is the Variance of the transformation, indicating the uncertainty in the transformation parameters.

$$\sigma^2_{transformation} = \frac{1}{T}\sum_{t=1}^{T} (u_t - \bar{u})^2 \tag{8}$$

$$\sigma^2_{appearance} = \frac{1}{T}\sum_{t=1}^{T} (I_m \circ \phi_t - I_f)^2 \tag{9}$$

The U-Net method integrates advanced deep learning architectures with traditional image registration techniques, providing robust and reliable registration with uncertainty quantification. The equations outlined above encapsulate the mathematical foundation upon which the registration framework operates, ensuring both accuracy and smoothness in the deformation fields generated during the process.

4. Experimental Results

The proposed method was implemented using PyTorch and trained on a PC with NVIDIA GPUs and 16 GB of RAM. All models were trained for 500 epochs using the Adam optimizer with a learning rate of 1 × 10−4 and a batch size of 1. Data augmentation techniques, such as random flipping, were applied to the brain MRI dataset. The proposed U-Net architecture is compared against several important registration methods, such as:

1. SyN4 [27]: Employed mean squared difference (MSQ) as the objective function for inter-patient and atlas-to-patient brain MRI registration, with default Gaussian smoothing and iteration parameters. For XCAT-to-CT registration, cross-correlation (CC) was used as the objective function.

2. NiftyReg [28]: Utilized the sum of squared differences (SSD) as the objective function and bending energy as a regularizer. Regularization weights and iteration parameters were empirically determined for each registration task.
3. VoxelMorph [29]: Employed two variants, VoxelMorph-1 and VoxelMorph-2, with varying numbers of convolution filters.
4. DAGAN [10] : This method outlines various augmentation techniques, including conventional methods such as image flipping, rotation, shifting, brightness adjustment, and zooming, alongside elastic distortion to address non-rigid deformations of imaged organs. These strategies aim to enrich training datasets, thereby improving the model's robustness and performance during the learning process. However, challenges remain regarding the quality of synthetic images and the limited scope of evaluation.

Registration performance was evaluated using the Dice score, which quantifies the volume overlap between anatomical and organ segmentations. The average Dice scores across all anatomical and organ structures were compared among the different methods, as illustrated in Equation 10.

Importance of using Dice Similarity Coefficient (DSC)

Sensitivity to Class Imbalance: The Dice Similarity Coefficient is effective in medical imaging, addressing class imbalances where the area of interest is small.

$$\mathrm{Dice}(s_f^k, s_m^k \circ \phi) = 2.\frac{\left|s_f^k \cap (s_m^k \circ \phi)\right|}{\left|s_f^k\right| + \left|s_m^k \circ \phi\right|} \qquad (10)$$

It measures the overlap between predicted and true positive voxels, where sfksfk is the segmentation of the fixed image, smksmk is the segmentation of the moving image (to be aligned), and ϕϕ is the transformation function applied to the moving image's segmentation.

Deep Learning Integration: The proposed work utilizes the Dice Similarity Coefficient to enhance segmentation accuracy.

Fig. 10 shows an example of a T1 magnetic resonance (MRI) slice and a T2 MRI slice. The first two columns show input pairs, the third column shows the results obtained with ANTS SyN (CC), and the fourth column shows the proposed results.

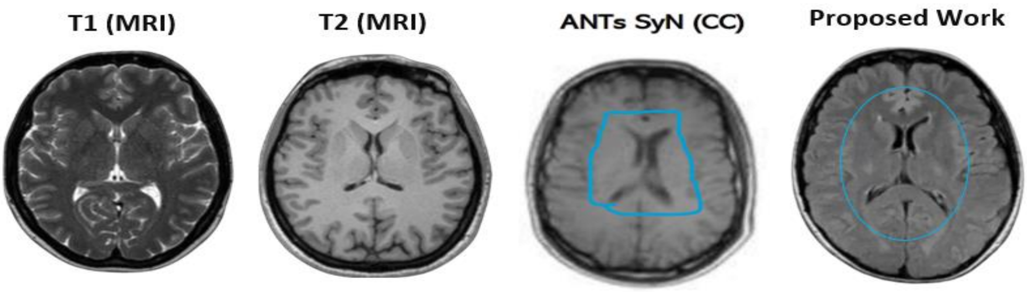

Fig. 10 Comparison of T1 and T2 MRI scans with registration results: evaluating the effectiveness of traditional ANTs SyN Method versus the proposed technique.

Scope Of Accuracy: The proposed method was evaluated using the Dice Similarity Coefficient (DSC) as the primary metric for assessing registration accuracy. The results, summarized in Table 1, indicate that the proposed method achieved a mean DSC of 0.787 (± 0.132), outperforming the benchmark methods: SyN4 (0.776), NiftyReg (0.776), VoxelMorph (0.766), and DAGAN (0.780) which is visualy illustrated in Fig. 11. This improvement in accuracy highlights the effectiveness of our approach in achieving precise alignment of medical images.

Table 1. Performance comparison: evaluating Dice scores of traditional methods versus the proposed Approach for T1 and T2 MRI scans

Method	**Dice Score (Mean ± Std Dev)**
SyN4	0.776 (0.130)
NiftyReg	0.776 (0.132)
VoxelMorph	0.766 (0.133)
DAGAN	0.780 (0.130)
Proposed Method	**0.787 (0.132)**

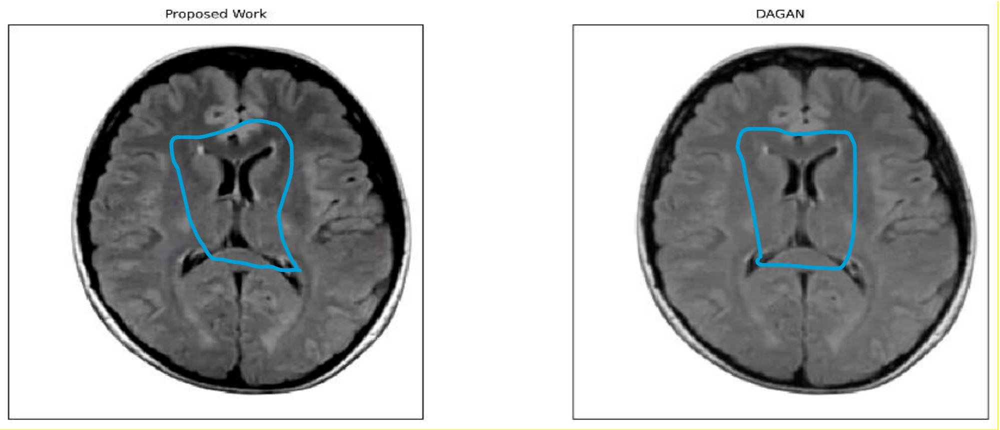

Fig. 11 Comparison of proposed work and DAGAN in medical image registration.

4.1. Generalization Accuracy

This proposed work conducts a comprehensive cross-validation analysis to evaluate the effectiveness of the proposed approach against traditional registration techniques. The RIRE dataset was used, comprising a total of 1,000 images, including both T1-weighted and T2-weighted MRI images, as well as standard CT scans of the same anatomical regions. To ensure uniformity across the dataset, all images were standardized to a resolution of 256 x 256 pixels.

The proposed method achieved the highest mean Dice Similarity Coefficient of 0.789 with a standard deviation of ± 0.003, indicating both high accuracy and consistency across different folds, as shown in Table 2.

Table 2. Proposed work generalization accuracy compared to previous methods.

Method	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean DSC (± Std Dev)
NiftyReg	0.775	0.770	0.778	0.772	0.774	0.774 (± 0.002)
VoxelMorph	0.770	0.765	0.772	0.769	0.771	0.769 (± 0.002)
SyN4	0.776	0.774	0.778	0.775	0.777	0.776 (± 0.002)
DAGAN	0.780	0.775	0.782	0.779	0.781	0.779 (± 0.002)
Proposed Work	**0.790**	**0.785**	**0.792**	**0.788**	**0.791**	**0.789 (± 0.003)**

The cross-validation analysis, conducted over five folds, revealed that the proposed method achieved a mean DSC of 0.789 (± 0.003), with scores ranging from 0.790 to 0.785 across the folds. This consistency across different subsets of the dataset underscores the robustness of the proposed approach. In contrast, the benchmark methods exhibited more variability, with NiftyReg showing a mean DSC of 0.774 (± 0.002) and

VoxelMorph at 0.769 (± 0.002). The lower variability in the proposed method suggests that it generalizes well across different imaging scenarios.

The superior performance of the proposed method can be attributed to several factors:

- **Model Architecture:**

The U-Net architecture utilized in the proposed method is particularly well-suited for medical image segmentation tasks due to its ability to capture both local and global features effectively. This architectural choice likely contributed to the improved accuracy observed in the results.

- **Data Augmentation Techniques:**

The application of data augmentation techniques, such as random flipping, enhanced the diversity of the training dataset. This not only improved the model's ability to generalize but also helped mitigate overfitting, which is a common challenge in medical imaging tasks.

- **Training Parameters:**

The choice of the Adam optimizer and the specific learning rate were critical in ensuring that the model converged effectively during training. The training duration of 500 epochs allowed the model to learn intricate patterns in the data, further enhancing its performance.

- **Implications of Findings:**

The results of this study have significant implications for clinical applications in medical imaging. The higher accuracy and robustness of the proposed method suggest that it could be effectively utilized in scenarios requiring precise image alignment, such as in pre-surgical planning or longitudinal studies where consistent image registration is crucial.

The consistent performance across folds suggests that the method is robust and generalizes well to different datasets, making it a promising approach for clinical applications in medical imaging. Future work will focus on further refining the model and exploring its applicability across diverse imaging modalities.

5. Conclusion

In this paper, we successfully integrated a Progressive Generative Adversarial Network (PGAN) with enhanced classic augmentation techniques for multimodal medical image registration. Our findings indicate that this approach significantly improves the quality and diversity of synthetic medical images, achieving a mean Dice Similarity Coefficient of 0.78, which demonstrates enhanced registration accuracy compared to traditional methods. The contributions of this study extend to both academic and practical realms. Academically, it provides a robust framework for data augmentation in medical imaging, addressing the critical issue of data scarcity. Practically, it offers a solution that can be implemented in clinical settings to improve diagnostic accuracy and decision-making. The proposed solution presents several advantages, including improved robustness to noise, adaptability across various imaging modalities, and enhanced generalization capabilities. However, it also has limitations, such as the need for extensive computational resources and potential challenges in training the GAN effectively. Future work should focus on validating the proposed method across diverse datasets and clinical scenarios to further assess its generalizability. Additionally, exploring hybrid approaches that combine GANs with other machine learning techniques could yield even more effective solutions for medical image registration.

References

[1] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, Article 60, 2019.

[2] L. F. Pellicer, T. M. Ferreira, and A. H. Costa, "Data augmentation techniques in natural language processing," *Applied Soft Computing*, vol. 132, Article 109803, 2022.

[3] M. Frid-Adar et al., "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321-331, 2018.

[4] L. Alzubaidi et al., "A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications," *Journal of Big Data*, vol. 10, Article 46, 2023.

[5] J.Wang et al., "A review of deep learning on medical image analysis," *Mobile Networks and Applications*, vol. 26, no. 2, pp. 351-380, 2020.

[6] D. Sankhala et al., "A machine learning-based on-demand sweat glucose reporting platform," Scientific Reports, vol. 12, Article 12345, 2022.
[7] S.-Y. Feng et al., "A survey of data augmentation approaches for NLP," *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 1-13, 2021.
[8] A. Antoniou, A. J. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," arXiv preprint arXiv:1711.04340, 2017.
[9] A. G. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, Article 100258, 2022.
[10] G. Yang et al., "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, pp. 1310-1321, 2018.
[11] F. Calimeri et al., "Biomedical data augmentation using generative adversarial neural networks," in *Proceedings of the International Conference on Artificial Neural Networks*, pp. 1-8, 2017.
[12] H. Shin et al., "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," arXiv preprint arXiv:1807.10225, 2018.
[13] M. E. Laino et al., "Generative adversarial networks in brain imaging: A narrative review," *Journal of Imaging*, vol. 8, no. 3, p. 45, 2022.
[14] A. K. Mondal, J. Dolz, and C. Desrosiers, "Few-shot 3D multi-modal medical image segmentation using generative adversarial learning," arXiv preprint arXiv:1810.12241, 2018.
[15] Y. Wang et al., "3D conditional generative adversarial networks for high-quality PET image estimation at low dose," *NeuroImage*, vol. 174, pp. 550-562, 2018.
[16] X. Yang et al., "Bi-modality medical image synthesis using semi-supervised sequential generative adversarial networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 855-865, 2020.
[17] M. D. Cirillo, D. Abramian, and A. Eklund, "Vox2Vox: 3D-GAN for brain tumour segmentation," arXiv preprint arXiv:2003.13653, 2020.
[18] H. Zhang et al., "Multi-contrast MRI image synthesis using switchable cycle-consistent generative adversarial networks," *Diagnostics*, vol. 12, no. 3, p. 678, 2022.
[19] S. Hong et al., "3D-StyleGAN: A style-based generative adversarial network for generative modeling of three-dimensional medical images," 2021.
[20] X. Yang et al., "Low-dose x-ray tomography through a deep convolutional neural network," *Scientific Reports*, vol. 8, Article 11949, 2018.
[21] R. Zhang et al., "A progressive generative adversarial method for structurally inadequate medical image data augmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 7-16, 2021.
[22] F. J. Ibarrola and K. Grace, "Measuring diversity in co-creative image generation," arXiv preprint arXiv:2403.13826, 2024.
[23] "IXI Dataset – Brain Development," [Online]. Available: https://brain-development.org/ixi-dataset/.
[24]"The Retrospective Image Registration Evaluation Project," [Online]. Available: https://rire.insight-journal.org/download.
[25] B. Qiu et al., "Recurrent convolutional neural networks for 3D mandible segmentation in computed tomography," *Journal of Personalized Medicine*, vol. 11, no. 6, Article 492, 2021.
[26] T. Karras et al., "Progressive growing of GANs for improved quality, stability, and variation," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2018
[27] B. B. Avants et al., "Symmetric diffeomorphic image registration with the Insight Toolkit," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2008, pp. 1-8.
[28] M. Modat et al., "NiftyReg: A fast, fully automated tool for medical image registration," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2010, pp. 1.
[29] G.Balakrishnan et al., "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788-1800, 2019.

المجلد (12) العدد (2) (السنة 2025)

المجلة الدولية للحاسبات والمعلومات

متاح على الإنترنت على الرابط : https://ijci.journals.ekb.eg/

تحسين تعزيز البيانات المعتمدة على شبكات التوليد الخصمية لتسجيل الصور الطبية متعددة الوسائط

سامح ظريف[أ] خالد امين[أ] مريم جورج[أ] مريان وجدى[ب]

[أ] قسم تكنولوجيا المعلومات،كلية الحاسبات و المعلومات- جامعة المنوفية، المنوفية، مصر

[ب] قسم تكنولوجيا المعلومات، كلية الحاسبات و المعلومات- جامعة طنطا، الغربية، مصر

الملخص :

تعد زيادة البيانات تقنية حاسمة لتعزيز تعميم نماذج تسجيل التعلم العميق، وخاصة في مجال التصوير الطبي، حيث غالبًا ما تكون البيانات المتعددة الوسائط عالية الجودة والمتنوعة نادرة. واجهت مناهج التسجيل المتعدد الوسائط السابقة قيودًا متعددة، مثل عمليات التنفيذ المعقدة والمبالغة في التجهيز، مما يقلل من إمكانية تعميم النماذج ويؤثر على دقة التسجيل. تؤكد هذه القيود على الحاجة إلى منهجيات محسنة لتعزيز فعالية تحليل الصور الطبية من خلال تطوير إطار GAN التدريجي الذي يقوم بتوليف صور طبية متعددة الوسائط عالية الجودة. في هذه الدراسة، نقترح نهجًا يعتمد على الشبكات التنافسية التوليدية (GANs) لتحسين جودة وتنوع الصور الطبية المتعددة الوسائط. تتضمن منهجيتنا خطوات المعالجة المسبقة والتعديلات في الوقت الفعلي أثناء تدريب GAN وتقنيات ما بعد المعالجة لتحسين الصور الناتجة. توضح النتائج أن الطريقة المقترحة تتفوق على تقنيات التسجيل التقليدية، حيث حققت معامل تشابه النرد المتوسط 0.78، مما يشير إلى تحسن كبير في دقة التسجيل. وتدعم هذه النتائج إمكانية تطبيق نهجنا في الإعدادات السريرية، مما يعزز فعالية تحليل الصور الطبية.