Unsupervised Medical Image Translation with CycleGAN: Applications and Clinical Boundaries

Zihao Ze[1*],

[1]School of Software Engineering, Shandong University, Jinan, China

Abstract. CycleGAN effectively mitigates the critical challenge of paired data scarcity in medical image-to-image translation. This comprehensive review examines its prominent applications across diverse modalities, particularly highlighting its success in MRI-to-CT synthesis for radiotherapy planning (e.g., reducing mean absolute error from 69.29 HU to 29.85 HU), ultrasound speckle noise reduction, and X-ray denoising. The review provides a comparative analysis with established methods, such as Pix2Pix (e.g., SSIM 0.82 vs. 0.85) and DualGAN, to underscore its strengths in robustness when leveraging unpaired datasets. Furthermore, the review identifies significant barriers to clinical deployment, primarily encompassing potential risks to diagnostic reliability induced by generative artifacts—such as creating clinically misleading features or missing subtle pathologies—and the unresolved ethical boundaries concerning the use of synthetic data. Future research directions should prioritize the development of robust artifact reduction techniques, the execution of large-scale multi-center clinical validation studies, and the advancement of hybrid models integrating anatomical priors.

1 Introduction

1.1 Challenges in Medical Imaging and the Motivation for Image Translation

The synthesis of multimodal medical images is fundamentally limited by the lack of precisely matched training data. This scarcity of training data is due to the fact that repetitive imaging is not very practical in clinical practice, and ethical restrictions further limit data acquisition. At the same time, the physiological changes between scans bring inherent biological noise [1-2]. In the medical setting, small-scale datasets, relatively large variances among modalities, and inconsistent annotations all exacerbate these challenges [1-2]. These limitations directly impede the development of data-intensive AI applications in medicine. For instance, the inability to assemble large, paired datasets can become a bottleneck in training robust models for rare diseases or for multi-institutional research, where data heterogeneity is the norm. This fundamental data constraint thus creates a compelling need for unsupervised translation methods that can bypass the strict requirement for pixel-aligned image pairs.

* Corresponding author: zezihao2024@outlook.com

1.2 Core Mechanism and Role of CycleGAN

CycleGAN is particularly valuable in medical imaging because it has the ability to learn from unpaired datasets, thus directly addressing the scarcity of precisely matched clinical data. This feature is achieved through a dual-generator framework, where G is from X to Y, F is from Y to X, and there is cycle consistency loss. That is, $\| F(G(x))-x \| + \| G(F(y))-y \|$, which is crucial for maintaining anatomical consistency among scans with inherent variations [1, 3]. This capability marks a significant departure from traditional supervised approaches. In clinical practice, it enables the leveraging of vast archives of existing, unpaired images, effectively turning previously unusable data into a valuable resource for model training and data augmentation, thereby unlocking new possibilities for building more comprehensive and generalizable AI systems.

One of the more significant applications is mri-ct synthesis, which can provide crucial electron density maps for radiotherapy plans. The implementation of CycleGAN in adaptive therapy has made considerable progress, such as reducing the mean absolute error from 69.29 HU to 29.85 HU. This improves the accuracy of dose calculation [4].

However, there are still some challenges in preserving high-frequency details such as microvessels. This limitation is driving the continuous development of hybrid models, which combine the adversarial robustness of CycleGAN with advanced architectures, such as transformers, to achieve better structural improvements [5].

1.3 Architectural Adaptations and Persistent Challenges

The standard CycleGAN architecture, while powerful, was not originally designed for the specific demands of medical imaging. Its baseline generators and discriminators can struggle with the high spatial fidelity and nuanced texture preservation required for accurate diagnostic interpretation. This means although CycleGAN's architecture provides a viable foundation for medical imaging, it must be adapted to address the unique challenges of biomedical data, namely its high spatial resolution and complex anatomy. Consequently, researchers have integrated specialized components such as residual blocks to improve gradient flow, deblurring modules to enhance texture recovery, and self-attention mechanisms to capture global contextual relationships. These innovations collectively work to preserve critical structural details more effectively [3, 6].

Nevertheless, these improvements fall short of resolving core issues. Physiological motion from respiration or cardiac activity often breaches the cycle-consistency premise, introducing artifacts that compromise final image quality [7]. Furthermore, model performance can be constrained when processing significant anatomical deviations (e.g., rare variants or post-surgical anatomy), highlighting the need for frameworks capable of integrating richer anatomical priors to ensure clinical reliability [4].

2 Comparative Performance and Discussion

2.1 Applications in Medical Modalities

2.1.1 MRI and CT Translation

CycleGAN is particularly valuable for MRI-to-CT synthesis, a task critical to MRI-based radiotherapy planning as it provides the required electron density information without additional CT scanning [8]. This application exemplifies the core trade-off between supervised and unsupervised methods. While supervised models like Pix2Pix can achieve

excellent performance given perfectly paired data, CycleGAN offers superior robustness in the more common and realistic scenario of working with unpaired datasets [8].

However, challenges remain in preserving the finest anatomical details under certain conditions. Unsupervised processes can be susceptible to motion artifacts, with performance dips (e.g., reduced SSIM values) observed in motion-corrupted studies [7].

2.1.2 Ultrasound and X-ray Enhancement

Beyond MRI-CT translation, CycleGAN has been effectively adapted to enhance other key modalities. In ultrasound imaging, its primary value lies in speckle noise reduction while preserving critical anatomical textures. For example, CycleGAN-based denoising methods have demonstrated significant improvement over conventional techniques in quantitative metrics and visual assessment [9].

Similarly, the unsupervised paradigm of CycleGAN is being explored to address the challenge of radiation dose reduction in medical imaging. While most robust validations have been demonstrated in low-dose CT denoising [10], the same principle holds promise for generating diagnostic-quality images from low-dose X-ray acquisitions, particularly for rare cases where paired training data is unavailable.

2.2 Comparative Performance in Clinical Applications

A defining characteristic of CycleGAN's clinical utility is its inherent trade-off between precision and practicality. Supervised approaches, notably Pix2Pix, can deliver exceptional pixel-level accuracy when perfectly paired training data is available. CycleGAN, in contrast, foregoes this level of precision for a greater degree of robustness and broader applicability by operating effectively on unpaired datasets, which are far more common clinically [9,10]. This practical advantage is actually very prominent compared to other unsupervised models. The CycleGAN framework enables us to combine perceptual loss, which can significantly enhance stability and the fidelity of the results. This is reflected in applications with high requirements such as ultrasonic improvement [11]. However, Most of the time, alternative unsupervised paradigms encounter the problem of unstable training and are prone to generating confusing artifacts. This point has been particularly emphasized in the broader analysis of medical Gans [12].

2.3 Critical Clinical Challenges

While the comparative performance and architectural adaptations are crucial, the ultimate translation of CycleGAN into clinical workflows hinges on overcoming challenges beyond quantitative metrics. Applying CycleGAN to clinical practice faces challenges that are not limited to performance indicators. Among them, the most crucial challenge is the security risk. The model may synthesize features that could cause deceptive effects in clinical practice or fail to capture some subtle pathological indicators. Training the model on restricted and non-heterogeneous datasets will only exacerbate concerns about such security risks [12]. Equally critical are the ethical dilemmas introduced by deploying synthetic images. These encompass ambiguities in diagnostic accountability when AI-generated content is involved, the imperative for transparent patient consent regarding data usage in generative model training, and the current lack of mature regulatory frameworks to ensure rigorous oversight [12]. Therefore, addressing these technical and ethical challenges comprehensively is the key to unlocking the next stage of clinical translation for CycleGAN.

3 Conclusion

Overall, this review elaborately describes the significant impact that CycleGAN has brought to medical image synthesis, and particularly emphasizes its key advantages in learning from unpaired data across multiple imaging modalities. Our analysis has compared its performance with supervised methods, documented key architectural adaptations for biomedical data, and examined major translational challenges, including safety risks from synthetic image inaccuracies and persistent ethical issues in generative AI. The discussions throughout this review underscore that overcoming these challenges is a prerequisite for leveraging CycleGAN's capability in facilitating large-scale, multi-institutional studies, which often rely on heterogeneous and unpaired datasets.

To propel this technology toward safe and effective clinical integration, future endeavors should converge on several critical pathways. A primary focus should be the development of robust systems capable of detecting artifacts and errors directly within the generative pipeline. Such integrated safeguards, by monitoring generation results in real-time and flagging regions of high uncertainty, are essential for safeguarding diagnostic reliability. This capability would serve as a critical safety net in clinical settings. Concurrently, there is an urgent need to move beyond retrospective validations through the execution of large-scale, multi-institutional clinical trials. These trials should aim to establish definitive efficacy standards that transcend conventional image quality metrics, creating a rigorous evidence base for clinical approval grounded in diagnostic performance and patient outcomes. Such evidence is vital for convincing regulatory bodies and the broader medical community of the technology's real-world utility.

The development of hybrid architectures that embed domain knowledge is essential for better generalization. One effective approach is to formalize anatomical and biophysical principles — such as the standard value ranges of tissue parameters or the structural relationships between adjacent organs—as explicit constraints during model training. This shifts the paradigm from merely learning statistical patterns from data to generating images that actively conform to established medical knowledge, thus improving robustness on atypical or rare cases. This approach is expected to enhance the model's credibility among clinicians by aligning its outputs with their expert knowledge. Parallel to these technical strides, the establishment of concrete ethical guidelines and adaptable regulatory frameworks is indispensable. This involves clarifying accountability for diagnoses involving AI-generated content, ensuring transparent patient consent for data usage in model training, and implementing dynamic oversight that spans the entire model lifecycle from development to deployment and decommissioning. A clear regulatory pathway will provide the necessary guidance for innovation while protecting patient safety and rights. Progress across these interconnected fronts—technological robustness, clinical validation, architectural innovation, and ethical governance — is crucial for realizing the full and trustworthy potential of CycleGAN and related generative frameworks in practical medical applications.

References

1. J. Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc. IEEE Int. Conf. Comput. Vis.*, Venice, Italy, October 22-29, 2017, pp. 2223-2232.
2. N. K. Singh, K. Raza, Medical image generation using generative adversarial networks: A review, in Health informatics: a computational perspective in healthcare, R. Patgiri, A. Biswas, P. Roy (Eds.) (Springer Singapore, Singapore, 2021), pp. 77-96. https://doi.org/10.1007/978-981-15-9735-0

3. Y. Liu, Y. Lei, T. Wang, Y. Fu, X. Tang, W. J. Curran, X. Yang, CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy, *Med. Phys.* **47**, 2472-2483 (2020) https://doi.org/10.1002/mp.14121
4. X. Liang, L. Chen, D. Nguyen, Z. Zhou, X. Gu, M. Yang, S. Jiang, Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy, *Phys. Med. Biol.* **64**, 125002 (2019) https://doi.org/10.1088/1361-6560/ab22f9
5. Y. Hu, Y. Li, H. Zou, X. Zhang, An unsupervised fundus image enhancement method with multi-scale transformer and unreferenced loss, *Electronics* **12**, 2941 (2023) https://doi.org/10.3390/electronics12132941
6. H. J. Tien, H. C. Yang, P. W. Shueng, J. C. Chen, Cone-beam CT image quality improvement using Cycle-Deblur consistent adversarial networks (Cycle-Deblur GAN) for chest CT imaging in breast cancer patients, *Sci. Rep.* **11**, 1133 (2021)https://doi.org/10.1038/s41598-020-80803-2
7. G. Chen, H. Xie, X. Rao, X. Liu, M. Otikovs, L. Frydman, C. Liu, MRI motion correction through disentangled CycleGAN based on multi-mask K-space subsampling, *IEEE Trans. Med. Imaging* **44**, 1907-1921 (2024) https://doi.org/10.1109/TMI.2024.3523949
8. H. Yang, J. Sun, A. Carass, C. Zhao, J. Lee, J. L. Prince, Z. Xu, Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN, *IEEE Trans. Med. Imaging* **39**, 4249-4261 (2020)https://doi.org/10.1109/TMI.2020.3015379
9. J. Liu, C. Li, L. Liu, H. Chen, H. Han, B. Zhang, Q. Zhang, Speckle noise reduction for medical ultrasound images based on cycle-consistent generative adversarial network, *Biomed. Signal Process. Control* **86**, 105150 (2023)https://doi.org/10.1016/j.bspc.2023.105150
10. Z. Li, S. Zhou, J. Huang, L. Yu, M. Jin, Investigation of low-dose CT image denoising using unpaired deep learning methods, *IEEE Trans. Radiat. Plasma Med. Sci.* **5**, 224-234 (2020)https://doi.org/10.1109/TRPMS.2020.3007583
11. S. Athreya, A. Radhachandran, V. Ivezić, V. Sant, C. W. Arnold, W. Speier, Ultrasound image enhancement using CycleGAN and perceptual loss, *arXiv preprint* arXiv:2312.11748 (2023)https://doi.org/10.1109/TRPMS.2020.3007583
12. S. Islam, M. T. Aziz, H. R. Nabil, J. R. Jim, M. F. Mridha, M. M. Kabir, J. Shin, Generative adversarial networks (GANs) in medical imaging: Advancements, applications, and challenges, *IEEE Access* **12**, 35728-35753 (2024)https://doi.org/10.1109/ACCESS.2024.3370848