Journal of HyoJeong Academia

Article

Enhancing AI-Driven Dermatology with CycleGAN: Synthetic Clinical-to-Dermoscopic Image Generation

Sarreha T. Rikta and Wonsang You*

AIIP Lab, Department of Information and Communication Engineering, Sun Moon University, Asan 31460, Korea
*Correspondence: wyou@sunmoon.ac.kr

Abstract: Dermoscopy and clinical imaging are vital for diagnosing skin conditions, as they provide complementary perspectives that enhance diagnostic insight. However, while clinical images are easy to acquire, dermoscopic imaging often faces limitations due to equipment costs, expertise requirements, and data scarcity. This imbalance restricts the development of robust deep learning models in dermatology. To overcome this challenge, we propose a CycleGAN-based bidirectional image translation framework capable of generating realistic synthetic dermoscopic and clinical images from their respective counterparts. The model effectively preserves key pathological structures while bridging the modality gap between the two imaging domains. Quantitative evaluation demonstrates promising results, with FID scores of 153.93 (clinical) and 117.03 (dermoscopic), and mean LPIPS scores of 0.6368 (clinical) and 0.6421 (dermoscopic), confirming the visual realism and structural consistency of the generated images. By reducing dependence on costly data acquisition and improving dataset diversity, this approach establishes a foundation for integrating synthetic data into dermatological deep learning, ultimately enhancing diagnostic accuracy and clinical applicability.

Keywords: Dermoscopy, Clinical skin image, Deep Learning, CycleGAN, Synthetic image generation

Citation: Sarreha T. Rikta and Wonsang You. Enhancing AI-Driven Dermatology with CycleGAN: Synthetic Clinical-to-Dermoscopic Image Generation. *JHJA 3*(2), **2025**.

Received: September 9, 2025
Accepted: October 8, 2025
Published: October 24, 2025

Publisher's Note: JHJA stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Skin diseases are among the most prevalent health issues worldwide and a significant cause of mortality [1]. Diagnosing these conditions usually requires a thorough dermatological examination and sometimes a biopsy. Dermoscopes, which are handheld devices that can reveal skin structures not visible to the naked eye, are frequently employed in this process. Many of these devices now come equipped with cameras to capture images of lesions. These images can be valuable for obtaining second opinions, including through automated computer-aided diagnosis (CAD) systems. However, this diagnostic process can be both time-consuming and expensive. To address these limitations, researchers have been exploring AI-based diagnostic systems in recent years. The use of artificial intelligence (AI) in dermatology offers great potential for improving the accuracy and efficiency of diagnosing skin conditions [2,3]. However, the effective implementation of AI algorithms heavily relies on having access to comprehensive and diverse training datasets. In dermatology, gathering large amounts of high-quality labeled images can be quite challenging, especially for rare skin diseases. This lack of sufficient data often results in imbalanced datasets, which can negatively impact the performance of AI models, leading to lower generalization and accuracy when diagnosing conditions that are not well-represented in the training data. To mitigate these issues, recent studies have explored generative models—particularly CycleGAN—as a promising solution for dataset enhancement

in dermatology. CycleGAN is well-known for its capability to create new images by learning from existing unsupervised data, allowing for transformations between various image types without the need for paired images [4]. For example, CycleGAN can successfully produce synthetic clinical photos from dermoscopic photos, or the other way around, thus achieving a more balanced representation of skin conditions across different demographics.

In this study [5], CycleGANs are utilized to generate realistic dermoscopic skin lesion images specific to different classes, serving as training data for supervised machine learning models. They proposed two innovative methods for generating these synthetic pictures, which are subsequently assessed for their effectiveness in enhancing a convolutional neural network (CNN) tasked with a complex lesion classification challenge. In another study, data augmentation techniques such as CycleGAN are employed to address the data scarcity issue by creating detailed synthetic training samples [6]. The results show that integrating CycleGAN-augmented data with the EfficientNet B1 model not only lowers the costs of manual annotation but also significantly enhances classification accuracy. Irungu et al. currently present a controllable image synthesis framework inspired by CycleGAN to tackle the challenges of breast lesion detection [7]. In addition, a CycleGAN-based image transformation method that utilizes convolutional autoencoders for domain translation improves privacy in deep learning models without sacrificing their utility [8]. In another paper, the author tackles the challenge of obtaining sufficient labeled data by utilizing Generative Adversarial Networks to artificially expand the dataset, addressing the problem of limited training samples [9]. The CNN model is then used for automatic classification, which enhances the model's ability to distinguish between different categories. Experimental results indicate that this method improves classification performance, achieving an average accuracy of 94.5%, which is 1.9% higher than previous techniques.

Nonetheless, a number of earlier studies show that GANs are effective tools for creating synthetic data and can handle the lack of data in medical imaging by translating images realistically. The capacity of GAN-based frameworks to translate images in both directions while maintaining the source domain's structural integrity has drawn attention to them. This study suggests using CycleGAN to convert clinical images into dermoscopic images and vice versa, creating synthetic datasets that preserve important pathological features. This is because dermoscopic image acquisition is frequently limited by the high cost of imaging equipment and the requirement for skilled handling. By making dermoscopic datasets more accessible and diverse, this strategy seeks to close the data gap, advance AI-driven dermatology, and improve diagnostic accuracy.

The main contributions of this study are as follows:

1. A CycleGAN-based bidirectional image translation framework is developed to generate synthetic dermoscopic and clinical images from limited paired data without explicit ground-truth annotations.
2. To evaluate the generated images' structural consistency and perceptual realism, extensive qualitative and quantitative evaluations are carried out.
3. The suggested methodology lays the groundwork for using synthetic data in deep learning for dermatology, minimizing reliance on expensive data collection and opening the door for further study using different generative models like CUT, UNIT, and KAN.

2. Materials and Methods

2.1. Dataset Preprocessing

In this study, we utilized the 7-point criteria evaluation dataset, which comprises over 2,000 clinical and dermoscopic color images. For our analysis, we specifically focused on 887 paired clinical and dermoscopic images, though these do not include ground truth annotations. The dataset preprocessing includes resizing all images to 256×256 pixels, converting them into PyTorch tensors, and normalizing the pixel values to a range centered around 0 with a standard deviation of 0.5 for each channel. These steps help maintain consistency and ensure the data is suitable for deep learning models.

2.2. Overview of CycleGAN architecture

CycleGAN is an innovative framework created for unpaired image-to-image translation. Its goal is to learn a mapping between two domains, X and Y, allowing an image from domain X to be transformed into domain Y

and then back to domain X, thereby reconstructing the original image. The fundamental idea is based on using two mappings: one from X to Y and another from Y to X. This method effectively tackles the issue of lacking paired examples in numerous real-world situations. Figure 1 represents the architecture of CycleGAN.

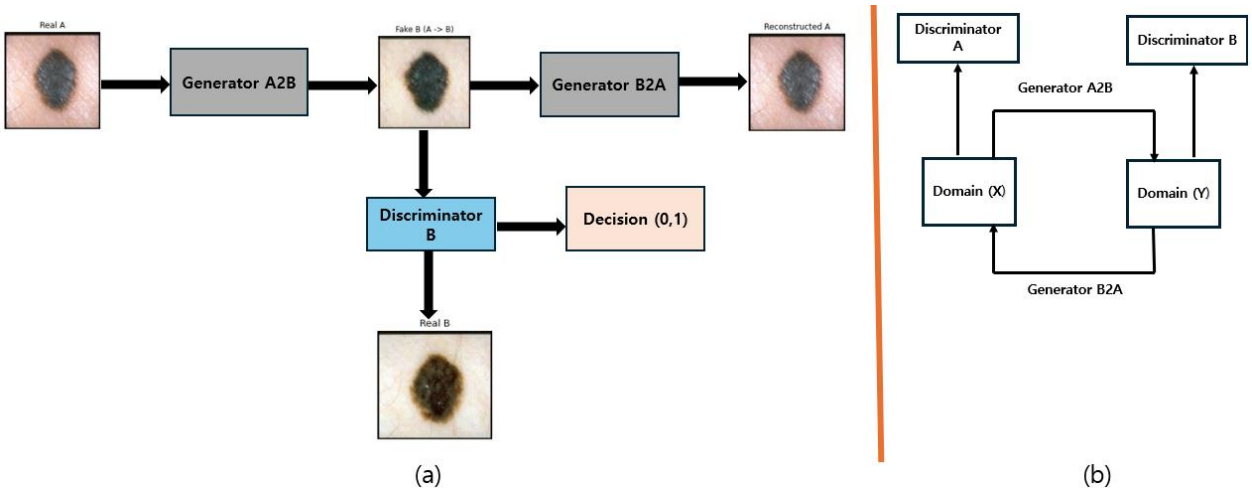


Figure 1. (a) Translation clinical (Domain X) to dermoscopic (Domain Y); (b) Full cyclic translation from Domain X to Domain Y and vice versa.

The architecture of CycleGAN features has two generator networks and two discriminator networks [10]. The generators create images in the target domain based on input images from the source domain, while the discriminators assess the authenticity of these generated images. A key component of CycleGAN is the cycle consistency loss, which ensures that when an image is transformed from X to Y and then back to X, it should closely match the original image. This relationship is mathematically expressed as F(G(X)) ≈ X and G(F(Y)) ≈ Y. In this model, we have utilized multiple loss functions, including Adversarial Loss and Cycle Consistency Loss, to ensure effective training. Adversarial Loss helps ensure that the images generated from one domain appear realistic in the other domain. The mathematical equation for this is as follows [11]:

$$L_{GAN}(G, D_{y,}X, Y) = E_{y\sim p_{data(y)}}[log(D_Y(y))] + E_{x\sim p_{data(x)}}[log(1 - D_Y(G(x))]$$

The adversarial loss equation describes how domain X is mapped to domain Y using generator G, and domain Y follows the same process. In this context, G represents generator A, F denotes generator B, D(x) refers to discriminator A, and D(Y) indicates discriminator B. Another important aspect is Cycle Consistency Loss, which guarantees that the relationships between two domains, X and Y, remain consistent. This loss function promotes the idea that the transformation should maintain the original content when transitioning back and forth between the domains. Mathematically, it is represented as follows-

$$L_{cyc}(G, F) = \mathbb{E}_{x\sim P_{data}}(x)[\| F(G(x)) - x \|_1] + \mathbb{E}_{y\sim P_{data(y)}}[\| G(F(y)) - y \| 1]$$

So, the total loss will be [4]

$$L(G, F, D_x, D_y) = L_{GAN}(G, D_{y,}X, Y) + L_{GAN}(F, D_{x,}X, Y) + \lambda L_{cyc}(G, F)$$

Where λ controls the importance of cycle consistency.

2.3. Experimental Setup

The experiments were conducted using PyTorch 1.10.0a0+ecc3718 with CUDA 11.4 and cuDNN 8202 on an Intel Xeon Gold 6426Y (x86_64) processor, 503.58 GB RAM, and an NVIDIA RTX A6000 GPU. All images were resized to 256 × 256 and normalized before training. CycleGAN was trained for 200 epochs with the Adam optimizer (learning rate 0.0001) and a batch size of 8. Table 1 represents a clear summary of the experiment.

Table 1. Experimental Setup Summary

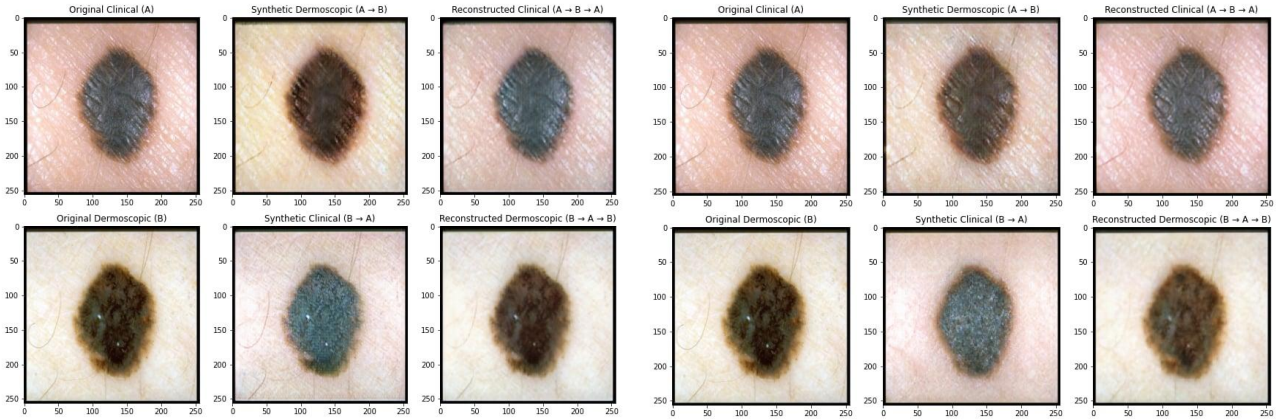
Category	Setup
Framework	PyTorch 1.10.0a0+ecc3718
CUDA Version	11.4
cuDNN Version	8202
Processor	Intel Xeon Gold 6426Y (x86_64)
RAM	503.58 GB
GPU	NVIDIA RTX A6000
Image Size	256 × 256
Training Epoch	200

3. Results

In this study, we assessed the qualitative and quantitative performance of CycleGAN for bidirectional image translation between clinical and dermoscopic domains, emphasizing the model's capacity to retain lesion-specific features and structural integrity.

3.1. Qualitative Evaluation

We evaluated qualitative performance, with a focus on its ability to preserve important lesion characteristics. Qualitative analysis is used to assess the quality and effectiveness of the images generated. Qualitative results demonstrate how the model's output quality improved throughout the training process. Figure 2 displays synthetic and reconstructed images produced at various training stages (epoch 100 and epoch 200). These visual comparisons offer valuable insights into the model's capability to capture domain-specific features and maintain image fidelity across different training iterations. The synthetic images produced at epoch 100 show a good translation, effectively capturing key characteristics of lesions like shape, color, and texture. However, there are some minor artifacts and color tone inconsistencies, particularly in the synthetic dermoscopic images (A → B). By epoch 200, the model shows better consistency, with smoother and more polished textures in both synthetic and reconstructed images. The improved stability in the later epoch indicates that the CycleGAN model has converged well, minimizing artifacts while maintaining important clinical details. At epoch 200, the synthetic dermoscopic images (A → B) and synthetic clinical images (B → A) exhibit better color consistency and less artificial smoothing than those at epoch 100. The reconstructed clinical images (A → B → A) at epoch 200 closely match the original input, demonstrating improved cyclic consistency. Likewise, the reconstructed dermoscopic images (B → A → B) preserve the intricate structural details of lesions, showing enhanced contrast and fewer distortions.



(a) (b)

Figure 2. Qualitative analysis of CycleGAN-generated images at various training epochs: (a) The outputs at epoch 100 indicate early advancements in domain translation, though they still exhibit some minor artifacts and inconsistencies; (b) The outputs at epoch 200 reflect enhanced quality, showcasing improved textures, greater color consistency, and maintained structural details, underscoring the benefits of prolonged training.

The findings at epoch 200 indicate that prolonged training allows the model to recognize subtle differences across domains, which is essential for analyzing medical images. The enhanced ability to reconstruct clinical and dermoscopic features by the final epoch guarantees the preservation of lesion characteristics, a vital aspect for diagnostic purposes. The findings further highlight the model's ability to learn intricate relationships between clinical and dermoscopic images, effectively capturing specific features of each domain while preserving the key characteristics of the lesions.

3.2. Quantitative Evaluation

We calculated two perceptual measures that are frequently used in image-to-image translation tasks, Fréchet Inception Distance (FID) and Learned Perceptual Image Patch Similarity (LPIPS), in order to validate the model's performance. In comparison to actual samples, these measures evaluate the generated images' structural similarity and perceptual reality. The CycleGAN model obtained FID scores of 117.03 for dermoscopic images and 153.93 for clinical images, as indicated in Table 2, illustrating that the synthetic dermoscopic outputs are perceptually closer to genuine images in their respective domains.

Strong structural consistency and visual realism were also demonstrated by the mean LPIPS scores of 0.6368 (clinical) and 0.6421 (dermoscopic) in both modalities. Overall, these findings support the model's ability to produce artificial images that closely mimic actual dermatological data, thereby eliminating the gap between the dermoscopic and clinical worlds.

Table 2. Quantitative evaluation for the proposed study

Metric	Clinical Domain (A → B)	Dermoscopic Domain (B → A)	Interpretation
FID (↓)	153.93	117.03	Lower FID indicates better visual realism and similarity to real data.
LPIPS (↓)	0.6368	0.6421	Lower LPIPS reflects higher perceptual similarity and structural consistency.

Note: ↓ indicates that lower values represent better performance.

4. Discussion

The technical results of this work have significant academic and practical ramifications. It contributes to the expanding field of AI-driven dermatology on an academic level by showing how generative models, like CycleGAN, may successfully alleviate dataset scarcity without the need for paired or annotated data. The suggested framework also creates a repeatable procedure for assessing structural and perceptual similarity metrics (FID, LPIPS) in activities involving the translation of medical images. High-quality synthetic dermoscopic and clinical image generation can, in practice, lessen reliance on expensive imaging equipment and professional labeling, allowing for the creation of diagnostic models that are more inclusive and accessible for a range of clinical contexts. Beyond the current experimental setup, we acknowledge that this study is limited to a single paired dataset, which may constrain the generalizability of the proposed approach across diverse populations, imaging environments, and skin conditions. To guarantee robustness and clinical reliability, future studies will try to assess the model on larger and more diverse datasets, taking into account differences in ethnicity, lesion type, illumination, and acquisition devices. Furthermore, cooperative validation with dermatologists from various medical facilities will be sought to evaluate the diagnostic utility and adaptability of the synthetic pictures in practical contexts. These actions are necessary to verify our suggested framework's cross-domain usability and scalability in real-world dermatological workflows.

Furthermore, in order to improve translation accuracy, feature preservation, and computing efficiency, thereby guaranteeing wider scalability and clinical reliability, continuing tests will compare the suggested approach with CUT, UNIT, and KAN models.

5. Conclusion and future work

In this study, we explored the effectiveness of CycleGAN for translating between clinical and dermoscopic image types, tackling the challenges posed by the limited availability of dermoscopic datasets. The synthetic images produced by CycleGAN effectively maintained essential pathological features, demonstrating their ability to enhance dataset variety and aid in the creation of strong diagnostic models. Future research will also investigate lightweight architecture and diffusion-based generative models to increase image fidelity while lowering processing requirements. All together, these continued initiatives should promote more dependable and easily available AI-driven diagnostic tools by expanding the use of synthetic data in dermatological imaging.

References

1. Rikta, S. et al. Xml-gbm lung: An explainable machine learning-based application for the diagnosis of lung cancer. *Journal of Pathology Informatics 14*:100307, (2023).
2. Lundervold, A. and Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik 29*(2), 102–127 (2019).
3. Litjens, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis 42*, 60–88 (2017).
4. Zhu J., Park, T., Isola, P. and Efros, A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision, Venice, Italy (22-29 Oct. 2017).
5. Fossen-Romsaas, S., Storm-Johannessen, A. and Lundervold, A. Synthesizing skin lesion images using cyclegans–a case study. *Medicine Computer Science* (2020).
6. Yixin Chen, Yifan Zhu, and Yanfeng Chang. Cyclegan based data augmentation for melanoma images classification. In Proceedings of the 2020 3rd international conference on artificial intelligence and pattern recognition, Xiamen, China (25-27 Sep. 2020).
7. Irungu, A. et al. Cyclegan-driven data augmentation for improved neural network disease detection. *Optics in Health Care and Biomedical Optics 14*(13242), 84–93 (2024).
8. Rodriguez, D. and Krishnan, R. A Cycle-GAN Based Image Encoding Scheme for Privacy Enhanced Deep Neural Networks. In *Lecture Notes in Computer Science,* vol. 14424 (Springer, 2023). https://doi.org/10.1007/978-3-031-49099-6_11
9. Tao, M. and Yan, Y. Identifying melanoma in lesion images using cycle-consistent adversarial networks-based data augmentation. In Proceedings of 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021) Medical Imaging and Computer-Aided Diagnosis, a virtual conference (25–26 March 2021).
10. Kawahara, J., Daneshvar, S., Argenziano, G. and Hamareh, G. Seven-Point Checklist and Skin Lesion Classification Using Multitask Multimodal Neural Nets. *IEEE Journal of Biomedical and Health Informatics 23*(2), 538-546 (2019).
11. An, T. and Joo, CycleGANAS: Differentiable Neural Architecture Search for CycleGAN. In Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, US (16-22 June 2024).