AI Engineering
https://www.sciltp.com/journals/aieng

Article

Check for updates

Evaluating the Effectiveness of Graph Neural Networks on an Augmented Dataset for Melanoma Skin Cancer Detection

Benedetto Lozzano [1,†], Anvi Kumar [2,†], Meet Patel [2,†], Arav Kumar [2,†], Savya Vats [2,†] and Avimanyou Vatsa [2,†,*]

[1] School of Information Sciences, University of Illinois Urbana-Champaign, Champaign, IL 61820, USA
[2] Deep Chain (DC) Lab, Fairleigh Dickinson University, Teaneck, NJ 07666, USA
* Correspondence: avatsa@fdu.edu
† Authors contributed equally to this article.

How To Cite: Lozzano, B.; Kumar, A.; Patel, M.; et al. Evaluating the Effectiveness of Graph Neural Networks on an Augmented Dataset for Melanoma Skin Cancer Detection. *AI Engineering* **2026**, *2*(1), 7. https://doi.org/10.53941/aieng.2026.100007

Received: 24 May 2026
Revised: 16 June 2026
Accepted: 17 June 2026
Published: 29 June 2026

Abstract: Early detection of malignant melanoma is critical for patient survival, but even trained clinicians achieve only moderate accuracy with traditional visual inspection. Recent work demonstrates that deep convolutional neural networks (CNNs) can rival dermatologists in dermoscopy-based melanoma classification. Nevertheless, CNNs require large, balanced datasets, which are lacking for rare skin lesion classes. This study considered augmenting the ISIC 2019 skin lesion dataset with GAN-generated images to balance the benign and malignant classes and proposed a graph-based classification approach. The dermoscopic images are segmented into super pixels (using Simple Linear Iterative Clustering (SLIC)), a region-adjacency graph of super pixel nodes is constructed, and graph neural networks, including Graph Convolutional Network (GCN), Graph Attention Network (GAT), and a Dynamic Attention Variant graph, are trained. To mitigate the imbalance, we used focal loss and balanced resampling approaches. On the augmented dataset, GNN models achieve higher sensitivity and AUC than a CNN baseline (e.g., GAT ROC = 0.92 vs. CNN ROC = 0.87), indicating the benefit of graph-based relational features. Although based on other experimental results, all models tend to overfit to synthetic images, as performance drops when evaluated on a held-out set of real images. Additionally, training on a smaller, balanced, and real-only subset improves real-world performance and continues to favor GNNs over CNNs. Therefore, these findings suggest that graph representations can enhance melanoma detection in augmented data regimes but also highlight the pitfalls of synthetic augmentation and the need for robust evaluation.

Keywords: melanoma skin cancer; CNN; GNN; GCN; GAT; image classification; superpixels; GAN augmentation; focal loss

1. Introduction

Melanoma is a highly lethal form of skin cancer, and its incidence is rising globally. Early diagnosis dramatically improves prognosis – for example, a small early-stage melanoma can often be cured by a simple excision [1]. However, standard clinical practice relies on visual inspection (e.g., the ABCD rule) and dermoscopy, both of which remain error prone. In primary care settings, even dermoscopy increases melanoma detection sensitivity from ∼54% (naked eye) to about 76–79% [2, 3]. Furthermore, rural regions have sparse access to dermatologists, which leaves residents largely to self-diagnosis [4]. This results in a substantial false-negative rate, particularly in rural areas and other regions where expert dermatologists are unavailable [1,3,4].

Advances in data-driven deep learning models have shown promising results, including CNNs trained on large dermoscopy datasets that have achieved dermatologist-level accuracy. In [5], it was demonstrated that a ResNet-based CNN, trained on ∼130,000 skin images, matched board-certified dermatologists at distinguishing

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

benign from malignant lesions. As technology improves, recent studies further report that CNN models, such as ResNet-50, significantly outperform dermatologists in ROC AUC on public datasets [6]. However, CNNs are limited in that they cannot capture non-Euclidean relationships within images because convolutional layers aggregate information in a Euclidean manner. As a result, they often fail to capture relations between arbitrarily shaped regions in an image [7]. This may be detrimental in cases where multiple afflicted regions contribute to the diagnosis of skin melanoma. Such relations are more easily captured with Graph Neural Networks (GNNs). Originally introduced by Scarselli et al. [8] to process arbitrary graph-structured data, GNNs have recently been adapted to images by representing them as region graphs (e.g., superpixel graphs) and applying graph convolutions or attentions [5,9–11].

To facilitate the training of melanoma classification models, researchers have compiled labeled datasets, one of the most popular being the ISIC 2019 dataset, assembled from other datasets [12,13]. However, the dataset suffers from class imbalance, with benign cases being much more frequent than malignant cases. Thus, previous studies have attempted to augment the dataset using Generative Adversarial Networks (GANs) to reduce this imbalance [14]. Using a CNN model on such a dataset, a prediction ROC score of 88.35% is achieved. This study builds upon previous research, using a new GAN-augmented dataset and exploring the potential of GNNs to compare their performance with CNN models for melanoma prediction on the GAN-augmented ISIC 2019 dataset.

The major contribution of this study is to compare and untangle GNNs' ability to perform image classification tasks, and to evaluate three of the most common GNN architectures for skin lesion classification and melanoma detection using the ISIC 2019 dataset. Consistent with its overall performance evaluation when proposed, the GAT with dynamic attention mechanisms performed the best across the board, although not by a very large margin. In conclusion, on the GAN-augmented dataset, the GNN models outperformed the VGG16 CNN model.

2. Background

The ISIC 2019 challenge dataset comprises over 30,000 dermoscopic images across nine lesion types, including melanoma (MEL) as one class [15]. As this study focuses only on predicting melanoma vs benign growths, only instances labeled either MEL or benign keratosis (BKL) are used. A sample of the images from the original ISIC 2019 dataset is shown below in Figure 1.

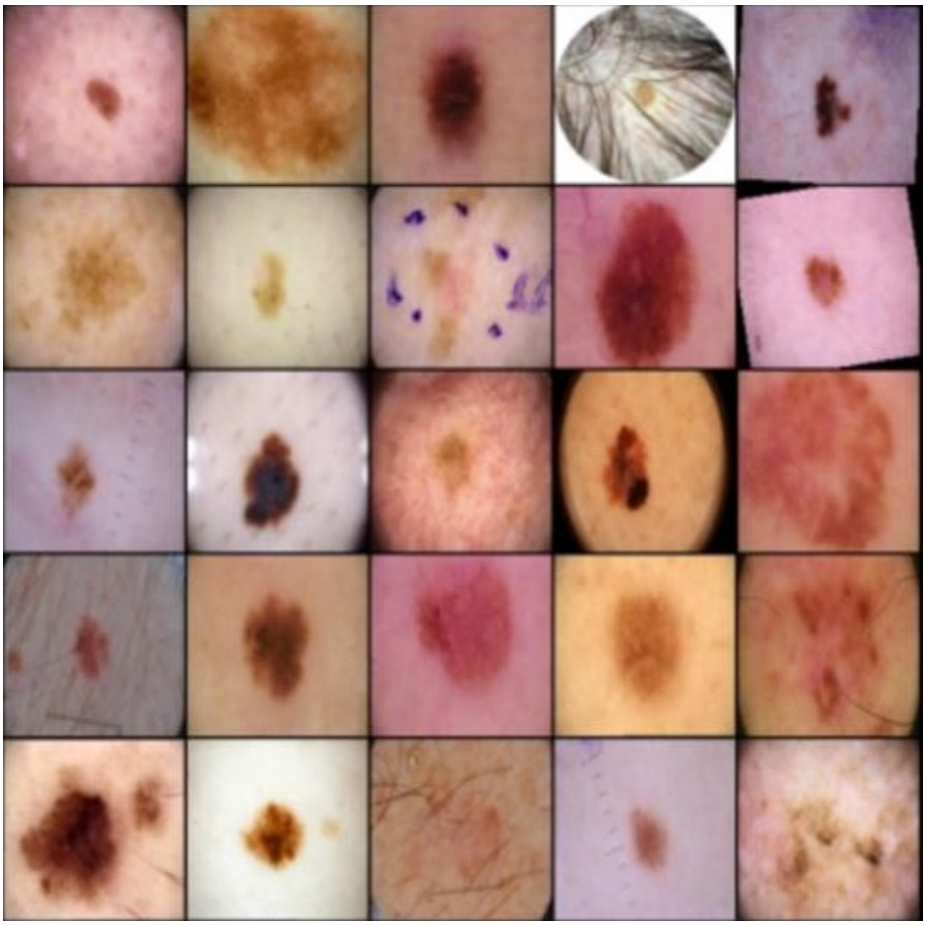

Figure 1. Sample of input images taken from the ISIC 2019 dataset, with a combination of malignant and benign lesions.

The original dataset comprises over 32,500 instances of BKL but only around 532 of MEL, resulting in a highly imbalanced dataset. This imbalance causes machine learning models to be biased towards majority classes. In practice, unbalanced training on ISIC data leads to poor recall for melanoma, which is clinically unacceptable. Many studies, including our own prior work, have therefore explored data augmentation or reweighting strategies to mitigate this [15,16].

GANs have been used to generate medical images and can improve classifier training by increasing data diversity [16,17]. Ref. [16] developed a DCGAN to generate synthetic dermoscopic images, demonstrating that GAN-based augmentation can boost melanoma classifier accuracy on ISIC datasets [16]. However, GANs may not fully capture the true variability of medical data, with even state-of-the-art GANs failing to reproduce the full richness of medical images, limiting their benefit [17]. As such, this study also seeks to test the efficacy of the GAN-generated images created in the previous study.

2.1. Image to Graph Construction

GNNs were introduced by Scarselli et al. [8] as a general framework for learning from graph-structured data. For computer vision tasks, images must be preprocessed before being fed to GNNs for prediction. This generally occurs in three steps [18]. First, the image is segmented into superpixels using an image segmentation algorithm. Then the superpixels are used to construct a region adjacency graph (RAG) that represents the image. The superpixels serve as nodes, and adjacent superpixels are connected by an edge in the graph representation. Node features can encode color, texture, or position; edges can encode boundary gradients or relative adjacency. Finally, the constructed RAG is fed to the model for prediction. The segmentation method used to separate an image into superpixels generally varies from study to study, although a 2024 analysis by [19] indicates that Simple Linear Iterative Clustering (SLIC) is generally the superior segmentation algorithm.

2.2. The Advantage of Graph Neural Networks for Images

Graph Neural Networks (GNNs) provide a compelling alternative to traditional convolutional neural networks (CNNs) by operating directly on non-Euclidean data structures—such as graphs—rather than requiring regular grid-based input typical of images. This flexibility enables GNNs to harness inherent relational and irregular structures in visual data that CNNs cannot easily capture [11,20–22].

While CNNs are constrained to fixed receptive fields on pixel grids, GNNs adapt seamlessly to arbitrary graph topologies, aggregating information across nodes based on connectivity rather than spatial proximity, enabling them to model irregular spatial relationships [23]. This becomes valuable when modeling complex image structures such as irregular lesion borders or region textures. Representing images as graphs attempts to capture the inherent hierarchical and relational nature of visual information, including but not limited to color or texture relations between regions.

In biomedical imaging, graph models have been applied to histopathology. Rivera Monroy et al. [24] constructed graphs from multi-epitope melanoma tissue (MELC) images, achieving 87% accuracy in melanoma classification by GNNs on cell graphs. These studies also motivate the use of GNNs on skin lesion images. Unlike CNNs, GNNs can explicitly model pairwise or higher-order relationships between regions, which may capture lesion border irregularity or color patterns indicative of melanoma. However, to date, GNNs have not been widely applied to dermoscopic image classification [11,20–22].

2.3. GANs for Skin Lesion Augmentation

The severe class imbalance in ISIC 2019 has prompted the use of GANs to generate additional melanoma images. Several studies report that GANs can generate high-quality images that are indistinguishable even to expert investigators [16,17]. GAN-generated images often look realistic, but validation on held-out real images is crucial since models can exploit subtle artifacts. Because of this, Ref. [17] actually finds that GANs cannot "[reproduce] the full richness of medical datasets" [17], even if the images are indistinguishable to the human eye. However, in prior work at FDU, CycleGANs were used to augment the ISIC2019 dataset, which improved CNN F1 from 0.19 with the original dataset to about 0.75 in cross-validation [2]. Because of this contradiction, this paper further explores the efficacy of GAN-augmented images to reconsider the value of GAN-generated images for medical dataset augmentation.

3. Materials and Methods

3.1. Addressing Dataset Imbalance

This study starts with the ISIC 2019 augmented dataset, which contains 32,500 images of malignant and 5532 benign skin lesions. As discussed previously, the original dataset is highly imbalanced, with a class ratio of roughly ($\frac{32500}{532} \approx$) 61:1. Sriram et al. [14] explores the use of different GAN architectures to augment that dataset, yielding the best results with CycleGANs. As such, the CycleGAN is used to generate 5000 synthetic images for the malignant class. Figure 2 shows the class imbalance before and after the data augmentation.

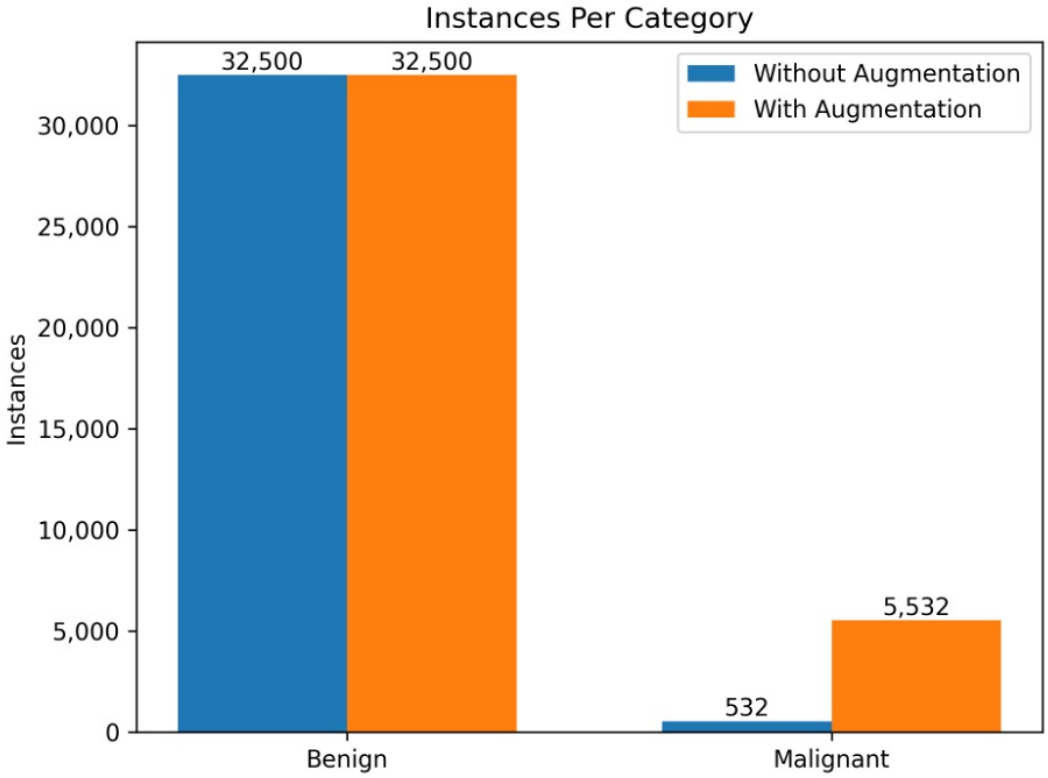

Figure 2. The distribution of benign vs. malignant instances of skin melanoma with and without augmentation .

Although adding the augmented image to the malignant class helps reduce class imbalance, it is not close enough to create a balanced dataset. As such, this study also explores data subsetting by randomly selecting 5500 images from the benign class in one run. To test the efficacy of the GAN-generated images, this study also uses data subsetting of benign images and malignant images available in the original dataset, creating a smaller, 1:1-balanced dataset of benign and malignant images [14,25].

Furthermore, the focal loss algorithm is used whenever a 1:1 ratio between malignant and benign images does not exist, with a gamma value of 2 and an alpha value of 0.75. These values are chosen due to the high-class imbalance and the optimal range of values given by [26]: a gamma value of around 2 and an alpha value between the range of 0.25 and 0.75. This algorithm is implemented as described in the original article creating focal loss [27]. The binary cross-entropy loss function from the PyTorch library is used when a 1:1 ratio exists.

3.2. Graph Construction and Attributes

Each dermoscopic image is first segmented into superpixels, which serve as graph nodes. As the SLIC algorithm is considered the best segmentation algorithm, it is adopted in this study [19]. The SLIC algorithm is then configured to generate 100 superpixels per image by clustering pixels based on color and spatial proximity. A RAG is then constructed from the 100 superpixels, with nodes representing superpixels and undirected edges drawn between two superpixels sharing a boundary. Node features include the mean and median color values, as well as the centroid x- and y-coordinates of the superpixel within the original image, creating a 5-dimensional feature vector. Edge attributes are two-dimensional, containing the mean and median color difference between adjacent superpixels. For a standard GNN architecture, edge features are not used in the model and are thus withheld.

4. Graph Neural Network Models

The study implements three GNN architectures using PyTorch Geometric [28]: (1) GCN (Graph Convolutional Network); (2) GAT (Graph Attention Network); and (3) Dynamic GAT (DyEdgeGAT). Each model takes the RAG as input and outputs a binary classification (malignant vs benign). The network consists of two graph-convolution/attention layers (hidden dimension 64) followed by a final fully connected classification layer. The GCN uses the standard spectral convolution (Kipf Welling style); the GAT uses multi-head self-attention with LeakyReLU and SoftMax weighting. The "dynamic" variant modifies GAT by allowing the model to adjust edge weights over layers, effectively relearning the adjacency via attention [29]. All models use ReLU activations and batch normalization between layers. The final graph representation is obtained by global mean pooling of node features, followed by a SoftMax classifier. All models were trained for 100 epochs using the Adam optimizer (learning rate = 0.001) and early stopping after 2 epochs without improvement on a validation split.

The training was initially allowed to run for 100 epochs (Figure 3). However, it was found that the model often overfitted, and that early stopping would not catch it in time to make an impact. As such, we decided to use the "elbow bend rule" and run it for only 20 epochs. Early stopping was still present but unused. We predict that the reason more training was not required was the model's smaller size and the relative simplicity of the binary classification task, especially when the model was predicting whether the image was generated rather than whether it contained a malignant lesion.

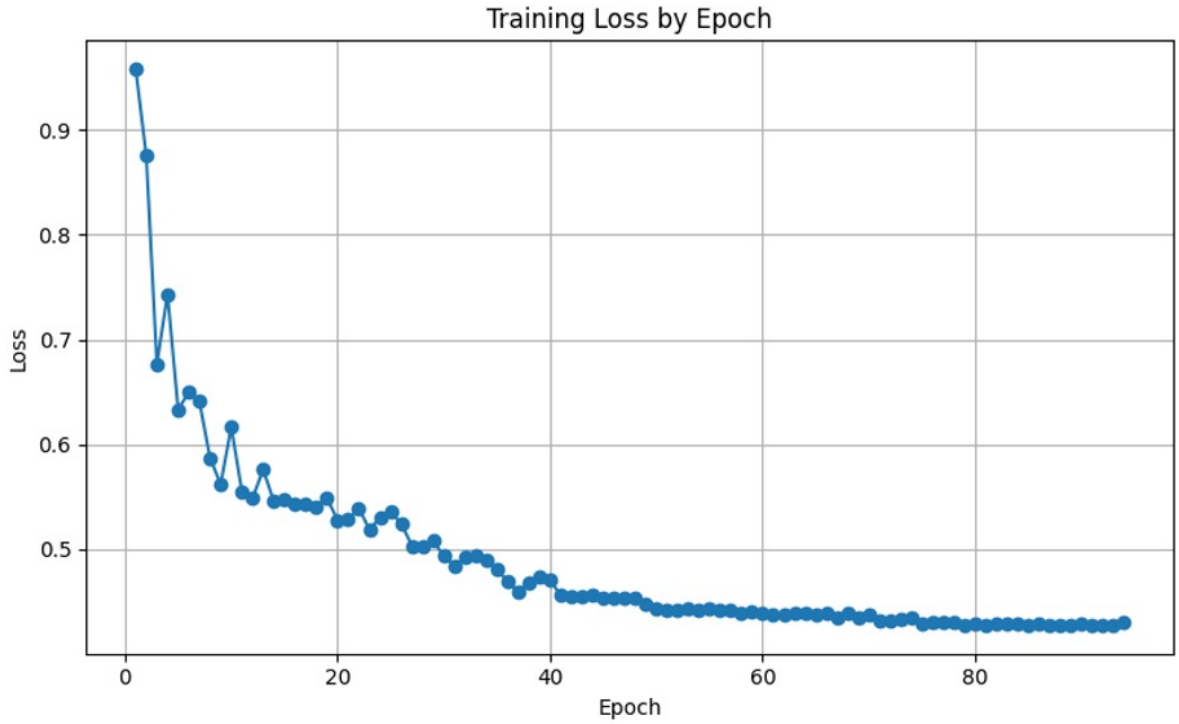

Figure 3. Training Loss by epochs.

5. Results and Discussions

Model performance is assessed using AUC-ROC , accuracy, precision, and recall. Figure 4 compares the performance of the various GNN architectures with that of the previously explored CNN architecture, both with and without augmented data from the GAN.

Most notably, the CNN model performs better without augmented data, but is outperformed by GNN models when augmented data is included. Recall for the VGG16 CNN model is notably higher than that of the GNN models without the augmented data. This suggests that the GNN models may be better suited to the GAN-augmented dataset than to the original dataset.

However, Figure 5 illustrates that when the GAN-generated images are taken out of the validation set, model accuracy drops significantly. A comparison of the performances of the various GNN architectures explored in this study shows a relatively high F1 score with augmented images in the validation set, but an F1 score close to 0 when they are removed. Such results indicate that the models seem to have been predicting which images were generated rather than classifying skin lesions as benign growths or melanoma.

However, data subsetting sees some success in improving model scores. Figure 6 shows a comparison of the three models when a subsetted dataset is used as opposed to the full, highly imbalanced dataset with focal loss. The GAT with dynamic attention mechanisms seems to do the best overall on the subsetted data, while all models perform comparably poorly on the full imbalance dataset.

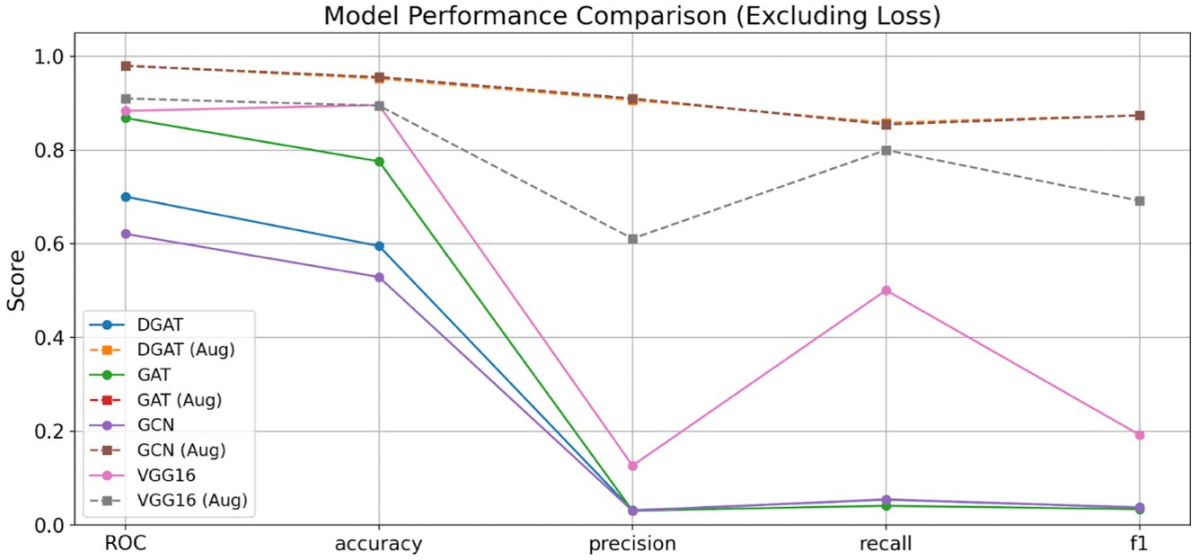

Figure 4. A comparison of the models with augmented data in the validation set (dotted line) and those without (solid line).

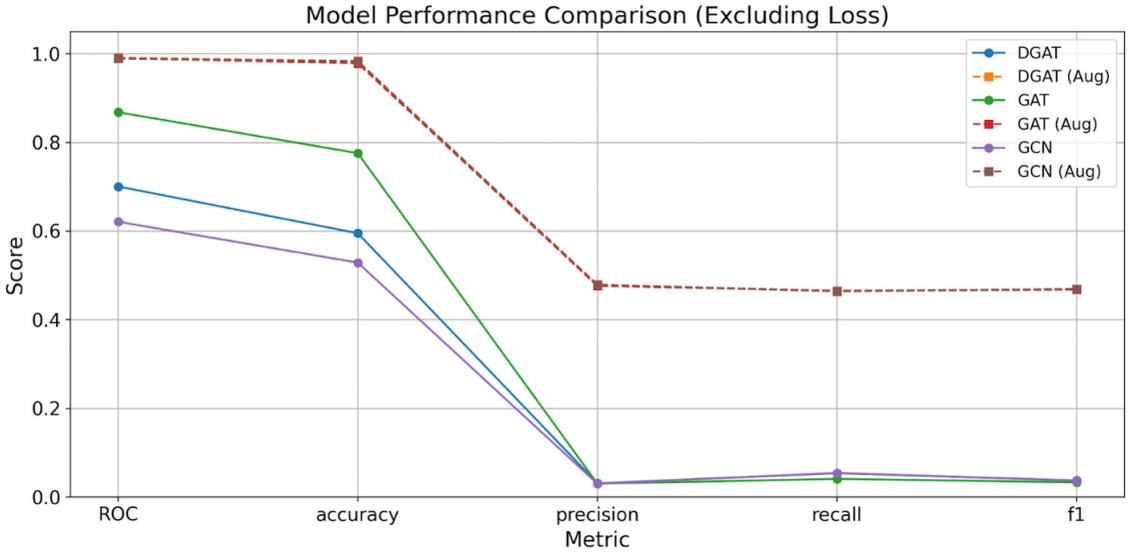

Figure 5. A comparison of the models without augmented data in the validation set (dotted line) and those without (solid line).

On the GAN-augmented dataset, the GNN models performed better than the VGG16 CNN, which was previously reported to achieve the best results on the same dataset [2]. However, on an unaugmented dataset, the CNN architecture outperformed all three GNN architectures. These mixed results indicate that neither architecture is inherently better than the other, and that the choice between them is task specific.

However, results also indicate that the GNN models were not actually predicting whether an image was melanoma or a benign growth, but rather whether it was GAN-generated. It has been observed that the performance of all three GNN models decreased significantly when the augmented data was removed from the validation datasets. This is because the GAN was trained only to generate malignant images, and thus, only malignant images were augmented. Furthermore, since the new dataset included 5000 generated images and only 500 original images, the model was encouraged to separate generated from non-generated images, rather than its original task of predicting skin cancer. This result is consistent with previous studies using GAN augmentation, which have found that augmentation may introduce artifacts that reveal the status of a generated image [16,17].

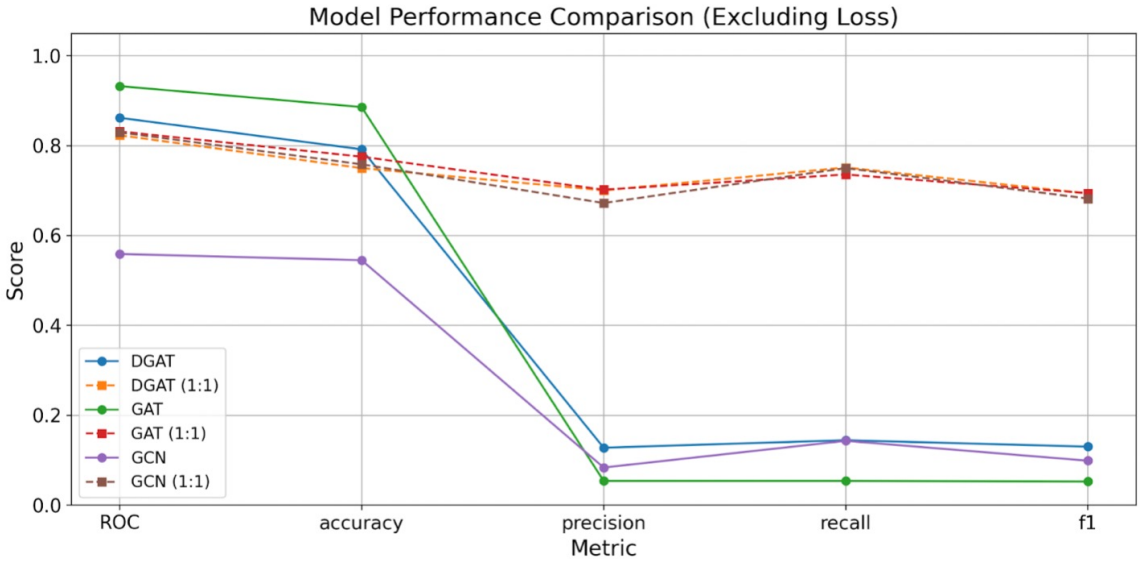

Figure 6. A comparison of performance metrics with a dataset ratio of 1:1 ratio benign to malignant (dotted), compared to utilizing the full imbalanced dataset (solid).

As the GNN models indicated a higher performance only when tested with augmented images in the validation set, it seems that GNN models may be better at determining generated images than the CNN model. As GANs work by using both a generator and a discriminator, which is usually a CNN model [30], further studies exploring the use of GNN models for the discriminator may help improve the efficacy of GANs.

Furthermore, future studies should explore the use of different types of segmentation algorithms. Although SLIC is generally considered the best algorithm [19], other algorithms may be better suited to this specific task than SLIC. This would be important in shaping the study's results, as the quality of the input graphs for GNNs derived from images is crucial for achieving good results with any GNN algorithm [23].

6. Conclusions

This study presents the use of GNN architectures on a GAN-augmented ISIC 2019 dataset, continuing earlier studies exploring the promise of GAN augmentation. This study reinforces the idea that the use of GANs for medical image substitution is not a viable solution at this time, contradicting the earlier promise of GAN augmentation explored in prior studies. Furthermore, it warns of the dangers of augmented data in the validation set, which may lead to reports of higher performance metrics than the model truly has on the given task. This study also suggests that GNNs may be better at detecting generated images than traditional CNN models, encouraging the exploration of GNNs for the discriminator component of a GAN.

Author Contributions

All authors contributed equally to this article. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Ethical review and approval were waived for this study because all dermoscopic images were sourced from publicly available datasets and do not contain any personally identifiable patient information.

Informed Consent Statement

No humans are involved for experimental or clinical trial purposes in this study.

Data Availability Statement

The dataset, source Python code, saved models, results, and instructins are available at GitHub: https://github.com/ben451P/FDUGNNs.

Acknowledgments

We gratefully acknowledge the support of the other lab members and collaborators of the Deep Chain (DC) lab at the Olsen College of Engineering and Science (OCES), Fairleigh Dickinson University (FDU), Teaneck, New Jersey 07666.

Conflicts of Interest

The authors and co-authors have no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

1. CDC. Skin Cancer: Melanoma of the Skin Statistics. Available online: https://www.cdc.gov/skin-cancer/statistics/index.html (accessed on 15 July 2025).
2. Kumar, A.; Vatsa, A. Untangling Classification Methods for Melanoma Skin Cancer. *Front. Big Data* **2022**, *5*, 848614. https://doi.org/10.3389/fdata.2022.848614.
3. Marghoob, A.A.; Usatine, R.P.; Jaimes, N. Dermoscopy for the Family Physician. *AFP* **2013**, *88*, 441–450.
4. Godinich, B.M.; Hensperger, V.; Guo, W.; et al. Barriers to Malignant Melanoma Diagnosis in Rural Areas in the United States: A Systematic Review. *JAAD Rev.* **2024**, *1*, 29–41.
5. Petrie, T.; Samatham, R.; Witkowski, A.M.; et al. Melanoma Early Detection: Big Data, Bigger Picture. *J. Invest. Dermatol.* **2019**, *139*, 25–30.
6. Mazhar, T.; Haq, I.; Ditta, A.; et al. The Role of Machine Learning and Deep Learning Approaches for the Detection of Skin Cancer. *Healthcare* **2023**, *11*, 415.
7. Wan, S.; Gong, C.; Zhong, P.; et al. Multi-Scale Dynamic Graph Convolutional Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3162–3177.
8. Scarselli, F.; Gori, M.; Tsoi, A.C.; et al. The Graph Neural Network Model. *IEEE Trans. Neur. Netw.* **2009**, *20*, 61–80. https://doi.org/10.1109/TNN.2008.2005605.
9. Marino, K.; Salakhutdinov, R.; Gupta, A. The More You Know: Using Knowledge Graphs for Image Classification. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 20–28.
10. Patel, M.; Khurana, D.; Mancinelli, L.; et al. Meta-Analysis on Image Classification Using GNN. https://www.researchsquare.com/article/rs-9817473/v1 (accessed on 16 June 2026).
11. Joseph, O.; Sukumar, P.; Ulloa, R.; et al. Sentiment Forecasting By Data-Driven Models. In Proceedings of the 35th Wireless and Optical Communications Conference (WOCC), Newark, NJ, USA, 8–9 May 2026.
12. Codella, N.C.F.; Gutman, D.; Celebi, M.E.; et al. Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC). In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018.
13. Tschandl, P.; Rosendahl, C.; Kittler, H. The HAM10000 Dataset, a Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions. *Sci. Data* **2018**, *5*, 180161. https://doi.org/10.1038/sdata.2018.161.
14. Sriram, A.; Vatsa, A.; Kumar, A.; et al. Challenges and Opportunities in Malignant Image Reconstruction Using GAN: A Review. In Proceedings of the 2024 IEEE Integrated STEM Education Conference (ISEC), Princeton, NJ, USA, 9 March 2024; pp. 1–6.
15. li, Z.; Chen, Z.; Che, X.; et al. A classification method for multi-class skin damage images combining quantum computing and Inception-ResNet-V1. *Front. Phys.* **2022**, *10*. https://doi.org/10.3389/fphy.2022.1046314.
16. Behara, K.; Bhero, E.; Agee, J.T. Skin Lesion Synthesis and Classification Using an Improved DCGAN Classifier *Diagnostics* **2023**, *13*, 2635.
17. Skandarani, Y.; Jodoin, P.M.; Lalande, A. GANs for Medical Image Synthesis: An Empirical Study. *J. Imaging* **2023**, *9*, 69. https://doi.org/10.3390/jimaging9030069.
18. Nazir, U.; Wang, H.; Taj, M. Survey of Image Based Graph Neural Networks. *arXiv* **2021**, arXiv:2106.06307.
19. Shoaib, A.; Vadiveloo, M.; Lim, S.P. Comparative Studies of Region-Based Segmentation Algorithms on Natural and Remote Sensing Images. In Proceedings of the 19th IMT-GT International Conference on Mathematics, Statistics and Their Applications (ICMSA 2024), Bangi, Malaysia, 27–28 May 2024; Volume 67.
20. Khurana, D.; Patel, M.; Mancinelli, L.; et al. Role of GNN in Skin Cancer Classification. In Proceedings of the 16th IEEE Integrated STEM Education Conference, Princeton, NJ, USA, 14 March 2026.
21. Khurana, D.; Mittal, H.; Joseph, O.; et al. Empirical Analysis of Video Segmentation for Sustainability. In Proceedings of the 16th IEEE Integrated STEM Education Conference, Princeton, NJ, USA, 14 March 2026.

22. Gautam, R.R.; Vatsa, A. Meta-Analysis of Image Classification Using Graph Neural Networks (GNNs). In Proceedings of the 16th IEEE Integrated STEM Education Conference, Princeton, NJ, USA, 14 March 2026.
23. Defferrard, M.; Bresson, X.; Vandergheynst, P. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. In Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 3844–3852.
24. Monroy, L.; Rist, L.; Eberhardt, M.; et al. Employing Graph Representations for Cell-level Characterization of Melanoma MELC Samples. *arXiv* **2022**, arXiv:2211.05884.
25. ISIC-Archive. ISIC Archive REST API Documentation. Available online: https://www.kaggle.com/c/siim-isic-melanoma-classification/data (accessed on 15 January 2025).
26. Cinar, U.; Cetin Atalay, R.; Cetin, Y.Y. Human Hepatocellular Carcinoma Classification from H&E Stained Histopathology Images with 3D Convolutional Neural Networks and Focal Loss Function. *J. Imaging* **2023**, *9*, 25. https://doi.org/10.3390/jimaging9020025.
27. Lin, T.Y.; Goyal, P.; Girshick, R.; et al. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 318–327.
28. Fey, M.; Lenssen, J.E. Fast Graph Representation Learning with PyTorch Geometric. *arXiv* **2019**, arXiv:1903.02428.
29. Brody, S.; Alon, U.; Yahav, E. How Attentive Are Graph Attention Networks? *arXiv* **2022**, arXiv:2105.14491.
30. Sharma, P.; Kumar, M.; Sharma, H.K.; et al. Generative Adversarial Networks (GANs): Introduction, Taxonomy, Variants, Limitations, and Applications. *Multimed. Tools Appl.* **2024**, *83*, 88811–88858. https://doi.org/10.1007/s11042-024-18767-y.