Semantic Fidelity Over Speed: A Comparative Study of Conditional vs. Unconditional GANs for Few-Shot Skin Disease Classification

Prabeen Singh Airee [1], Sarbin Sayami[2]

[1]Faculty Member, Divya Gyan College, Tribhuvan University

[2]Central Department of Computer Science and Information Technology, Tribhuvan University

Article History

Received: 11/12/2025

Revised: 27/12/2025

Accepted: 10/01/2026

Published: 15/01/2026

Corresponding Author

Name: Prabeen Singh Airee

Email: prabeenairee@divyagyan.edu.np

ORCiD: https://orcid.org/0009-0002-0131-0791

Abstract

This research aims to break the crucial bottleneck of data deficiency in developing AI for skin disease diagnosis by presenting a clear and evidence-based comparison of two different generative AI models for data augmentation. By using a small clinical data set, this research systematically compared a fast but unconditional model of generative AI, FastGAN, with a semantic-aware model of generative AI, cGAN. Synthetic data generated by these models was used for training and testing different classifiers for different kinds of medical diagnoses. The results of this research clearly show that there is a crucial dependency of AI models for medical data augmentation on semantic fidelity. The cGAN model, which is semantic-aware and preserves class-specific features of skin diseases, enabled classifiers to retain high accuracy (up to 93%) for different kinds of medical diagnoses. However, in stark contrast, the unconditional FastGAN model, despite being much faster in generating synthetic data, catastrophically failed in retaining accuracy as low as 49% due to semantic inconsistency in data augmentation.

Keywords: ***Skin Disease Classification, Data Augmentation, Generative Adversarial Networks (GAN), Conditional GAN, FastGAN, Few-Shot Learning, Medical Image Analysis, Deep Learning.***

Introduction

A lot of people worldwide suffer from skin conditions; therefore, precise diagnosis is essential to successful treatment. The availability of large datasets with annotations is vital for the success of deep learning techniques, particularly for Convolutional Neural Networks (CNNs), which have shown revolutionary possibilities for automating dermatological image processing. Due to the restrictions on patient privacies, the expensive cost of expert annotations, and the inherent scarcity of many diseases, medical imaging is marked by considerable data scarcity. As a consequence, the development of dependable and broadly applicable models is hindered by the small and unbalanced sets of data available. Data Augmentation plays a critical role in alleviating this problem, and this need is met by the presence of an effective replacement, termed as "Generative Adversarial Networks" (GANs), which are capable of generating new data that is similar to the existing data, based on the underlying distribution of the data available. Not all GANs are suitable replacements, however.

The main architectural difference is that, in the case of unconditional GANs, the image is generated using random noise, while in the case of conditional GANs (cGANs), the image is generated using class labels, thus allowing for targeted generation.

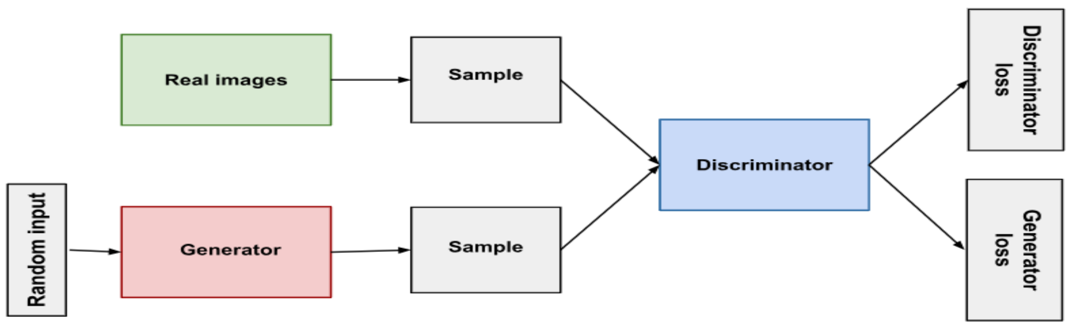

Figure 1: Architecture of Generative Adversial Network (GAN)

Recent developments have resulted in the formulation of new GAN variants, such as the FastGAN, designed for the stability and training speed of the model for use in small datasets (Biswas, 2023). Although the model is effective for the intended purpose, the absence of conditionality makes one ponder the applicability of the model in medical image synthesis, given the importance of disease-specific features. The cGAN, in turn, can be used as a DAGAN to address the class imbalance problem. Although the models have proved to be beneficial in the literature, there is no empirical study that uses the models for augmenting the same small clinical skin dataset using both traditional and data-efficient learning strategies.

This study addresses this gap by posing the following research questions:

1. What is the comparative impact of FastGAN (unconditional) and cGAN (conditional) augmentation on the classification performance of ResNet models and a Few-Shot Learning framework?
2. Which paradigm supervised learning or Few-Shot Learning is more effective at leveraging GAN-generated synthetic data under extreme data limitations?
3. What is the optimal combination of generative strategy and classifier for small-scale clinical skin disease classification?

Our contribution is a controlled comparison research that proves that the speed of unconditional generation is not remotely as useful for medical picture augmentation as the semantic fidelity ensured by conditional generation. We prove that FastGAN-generated data can even be detrimental when used, especially for Few-Shot Learning, while cGAN-generated data maintains classifier performance.

The rest of the sections of this paper will be structured in the following manner: Section 2 will deal with Related work. The Appendix outlines the methodology in Section 3. Section 4 presents and discusses its results. Section 5 concludes the study and gives recommendations to the ongoing research.

Related Work

Medical Data Augmentation using GANs: Medical data is missing, which has elicited the application of GANs. (Montenegro et al., 2021) introduced a privacy-preserving GAN, and it is case-based explain ability. An example of a lightweight GAN suitable to few-shot synthesis was released (Liu et al., 2021) and can process high-resolution image consumption using fewer data samples. To enhance the data efficiency of GANs, Zhao et al. introduced Differentiable Augmentation also known as DiffAugment. The studies demonstrate the potential of GANs, though the authors do not provide the comparison of the two approaches.

Few-Shot Learning in Dermatology: Few-Shot Learning (FSL) is a low-data-setting algorithm. Meta-transfer learning was used in the skin disease classification in the study of (Özdemir et al., 2024) under long-tailed settings. A few-shot model named CDD-Net was introduced in (Chen et al., 2024) and entails the use of multi-scale feature fusion to diagnose clinical data. Although FSL is applicable to dermatology, past studies tend to use original or otherwise augmented data, rather than GAN-generated ones. CNNs will stay in the core of skin lesion classification using deep learning. used a hybrid ResUNet++ and AlexNet-RF model on classification and segmentation challenges. Concatenated VGG-16 and ResNet-50 to detect both benign/malignant skin lesions. ResNet is renowned in that it is a depth-based model, and addresses the issue of vanishing gradients.

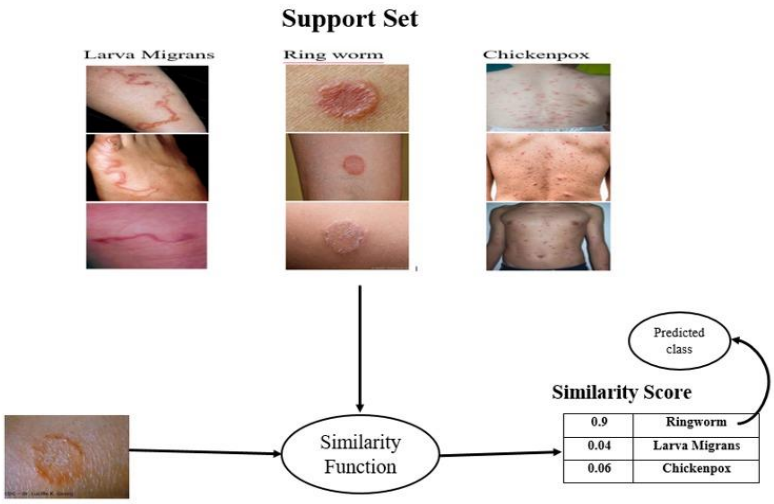

Figure 2: Architecture of FSL

The efficacy of the new unconditional GANs against their conditional counterparts, including their usage in supervised classifiers as well as in FSL classifiers, on the same clinical image set, has not been investigated although this has been demonstrated in earlier works regarding image augmentation and FSL using GANs and conditional GANs respectively. The paper addresses this research gap by offering guidelines of the selection of GAN variants, as the variants are used to generate FSL classifiers, to the cases of medical image datasets, which are affected by a lack of data.

Literature Review

The research (Özdemir et al., 2024) is focused on the classification of rare skin diseases in long-tailed data distributions by drawing a comparison between episodic and conventional training and few-shot learning and transfer learning. Using ISIC2018, Derm7pt, and SD-198 samples, the ImageNet pre-trained versions of DenseNet121 and a MobileNetV2 demonstrated significant improvements on limited labeled samples, demonstrating that traditional transfer learning using data augmentation performed better than the current methods, especially at 5-way classification and SD-198, and with a larger number of training examples, the classification rate increased.

CSDD-Net (Chen et al., 2024) is a plug-in module suggested by the authors to classify skin diseases on a few-shot basis, incorporating a context feature-fusion module to localize the details of lesions and a dual-attention mechanism to improve the relevant regions and channels and decrease the irrelevant areas. CDD-Net was validated by achieving an accuracy of +9.14 percentage points over baselines on the Derm104 dataset and ablation studies demonstrate the efficacy of both of its components.

The authors proposed a medical image classification privacy-preserving GAN that offers case-based interpretability in the form of factual explanations and counterfactual explanations and safeguards patient privacy by generating realistic images. The technique does not compromise personal identity as its explanatory power is found to be robust on both biometric and medical data. (Montenegro et al., 2021)

The authors (Liu et al., 2021) have created a lightweight GAN to perform few-shot image-based synthesis, which can produce high-resolution 1024x1024 images with low levels of data and computational-power. It trained in hours with a skip-layer excitation module and a self-supervised discriminator and achieved higher performance on thirteen datasets with limited data amounts than StyleGAN2.

In their work, the authors suggested Differentiable Augmentation (DiffAugment) (Zhao et al., 2020) which utilizes both real and generated images in differentiable augmentation to stabilize GAN training with small amounts of data. This avoids discriminator memorization and attains state-of-the-art results, such as large dividing FID on datasets such as FFHQ and LSUN, with only 20 percent of training data.

The authors (Mustafa Et al., 2024) introduced a hybrid deep learning method of skin lesion analysis with hair removal preprocessing, ResUNet++ segmentation, and AlexNet-Random Forest classification. They demonstrated a higher level of segmentation and classification performance on HAM10000 as their approach combined the medical segmentation capabilities of ResUNet++ with the powerful classification capabilities of AlexNet-RF.

The authors developed a hybrid CNN model for skin lesion classification. They combined the optimized VGG-16 and ResNet-50 architectures. The hybrid model achieved an accuracy of 85.65% in the classification of dermoscopic images from the ISIC 2016-17 data set. This proves the concept of combining different models for better performance. (Gupta et.al, 2022)
The article (Mukherjee, 2022) gives an overview of ResNet-50, a neural network model. It explains the importance of ResNet in the field of deep learning. It also gives a detailed explanation of the ResNet residual connection, which prevents the problem of vanishing gradients.

Methodology

1. Overall Workflow

Figure 3 showed that within a comparative experimental design, the research was carried out. Preprocessing of clinical data set was first performed. Subsequently, FastGAN and cGAN two variants of GANs were trained to perform image synthesis. Three other classifiers, namely ResNet-18, ResNet-50, and Prototypical Network, were then trained in three conditions of data namely Baseline, +FastGAN, and +cGAN.

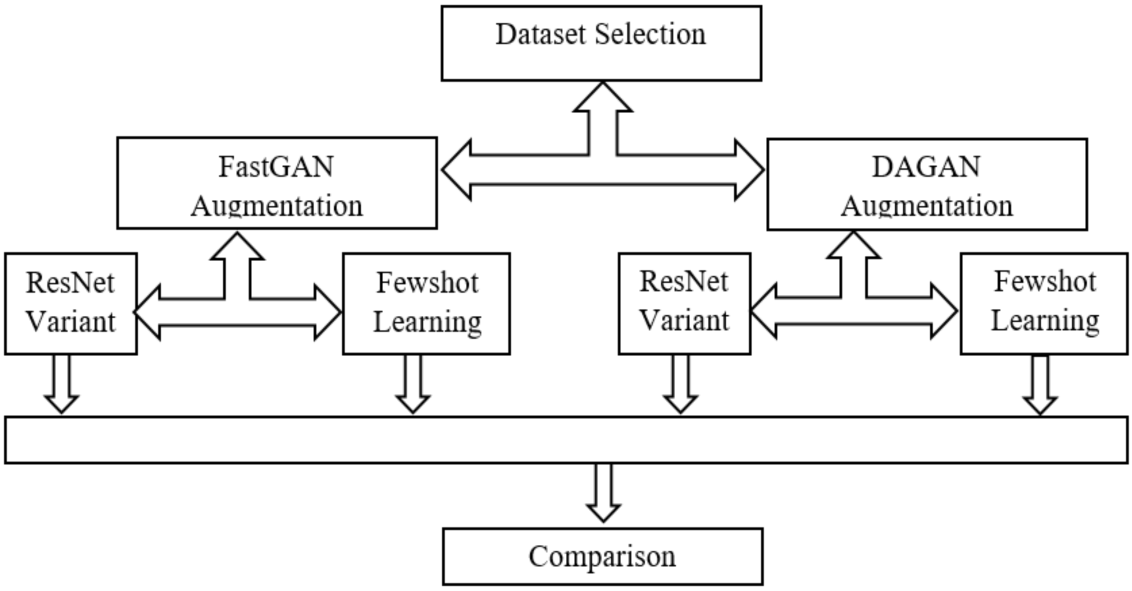

Figure 3: Research Workflow

2. Dataset and Preprocessing

It was based on the Skin Disease Dataset provided by Kaggle (Biswas, 2023), which included clinical images that have eight categories: bacterial diseases (Cellulitis, Impetigo), fungal diseases (Athlete's Foot, Nail Fungus, Ringworm), parasitic disease (Cutaneous Larva Migrans), and viral diseases (Chickenpox, Shingles). There were around 2,000 images that made up the dataset. The images were all resized to 224x224 and normalized. Training and testing were done according to an 80/20 split.

Table 1: Summary of skin Disease Dataset Categories

Disease Category	**Representative Diseases**
Bacterial infections	Cellulitis, Impetigo
Fungal infections	Athlete Foot, Nail Fungus, Ringworm
Parasitic infections	Cutaneous Larva Migrans
Viral Skin infections	Chickenpox, Shingles
Total Classes	8 distinct categories

To ensure consistency and compatibility with the deep learning models, all images underwent the following preprocessing pipeline:

Resizing: All images were resized to a uniform dimension of 224x224 pixels.

Normalization: Pixel values were normalized to a range of [0, 1] or [-1, 1] to stabilize and speed up the training process for the GANs and CNNs.

Data Splitting: The dataset was split into three subsets:

- Training Set (80%): Used to train the GANs and the classification models.
- Test Set (20%): Used only once for the final evaluation to report unbiased results. This set was completely hidden during all training phases

3. GAN Architectures for Augmentation

FastGAN Implementation: we used FastGAN architecture that was reported to rely on skip-layer channel-wise excitation in the generator and a multi-scale discriminator to train with small datasets with naturalness and efficiency. It was trained in the unconditioned mode with 50 epochs and Adam (lr=0.0002, 81=0.5, 82=0.999) to produce a balanced image set of all classes.

Conditional GAN (cGAN) Implementation:

We adopted a cGAN that was used in which class labels (embedded vectors) were inputted to both the generator and the discriminator. This conditions the generation on certain diseases to allow imbalanced classes to be oversampled. It was trained on the same hyper parameters as FastGAN to compare with it fairly with attention to minority classes generating supplemental images.

4. Classification Models

ResNet based supervised learning:

We used ResNet-18 and ResNet-50, which were fined tuned on our task of 8 classes, using ImageNet pre-training. The last fully-connected layer was substituted. The models were trained on Cross-Entropy loss and Adam (lr=1e-4, batch size=32) in 30 epochs.

Few-Shot Learning: Prototypical Networks:

Our training regime was a 5-way 5-shot episodic training scheme (Snell et al., 2017). The embedding network was a 4 layer CNN. The support embeddings of the prototypes were computed on mean. In training, synthetic images were provided to help sets of underrepresented classes.

$$\boldsymbol{P_c} = \frac{\mathbf{1}}{|\boldsymbol{Sc}|}\sum_{(x_i, y_i)\in \boldsymbol{Sc}} \boldsymbol{f_0}(\boldsymbol{x_i})\ldots\ldots\ldots\ldots\ldots\textbf{(1)}$$

P_c -: Prototype vector for class c

S_c -: Support set for class c

The classification is done by choosing the closest prototype to the query point in the embedding space.

This metric space is used to query a sample of a prototype to classify it. Training of the model was done in stages. The support sets were used to add synthetic images on training to make them more diverse, especially on rare classes.

5. Experimental Setup & Evaluation

Three experimental data conditions were established:

1. Baseline: Original training data with traditional augmentations (flips, rotations).
2. FastGAN: Original data + synthetic images from the unconditional FastGAN.
3. cGAN: Original data + targeted synthetic images from the conditional GAN.

All models were evaluated on the same unseen test set using Accuracy, Precision, Recall, F1-Score, and Confusion Matrices.

Results and Discussion

1. Overall Performance Summary

The test accuracy for all models under the three data conditions is summarized in Table 1. The results reveal a consistent and striking pattern: performance with cGAN-augmented data, while lower than the high baseline, was consistently and significantly superior to performance with FastGAN-augmented data across all classifiers.

Table 2: Test Accuracy (%) for All Models Across Data Conditions. (Mean ± SD over 5 runs)

Models	**FewShot**	**ResNET 50**	**ResNET 18**
On Original Data	99% ± 0.3	94% ± 0.5	88% ± 1.1
On cGAN Augmented Data	93% ± 0.8	80% ± 1.2	76% ± 1.5

On Fast GAN Augmented Data	49% ± 2.1	68% ± 1.8	62% ± 2.0

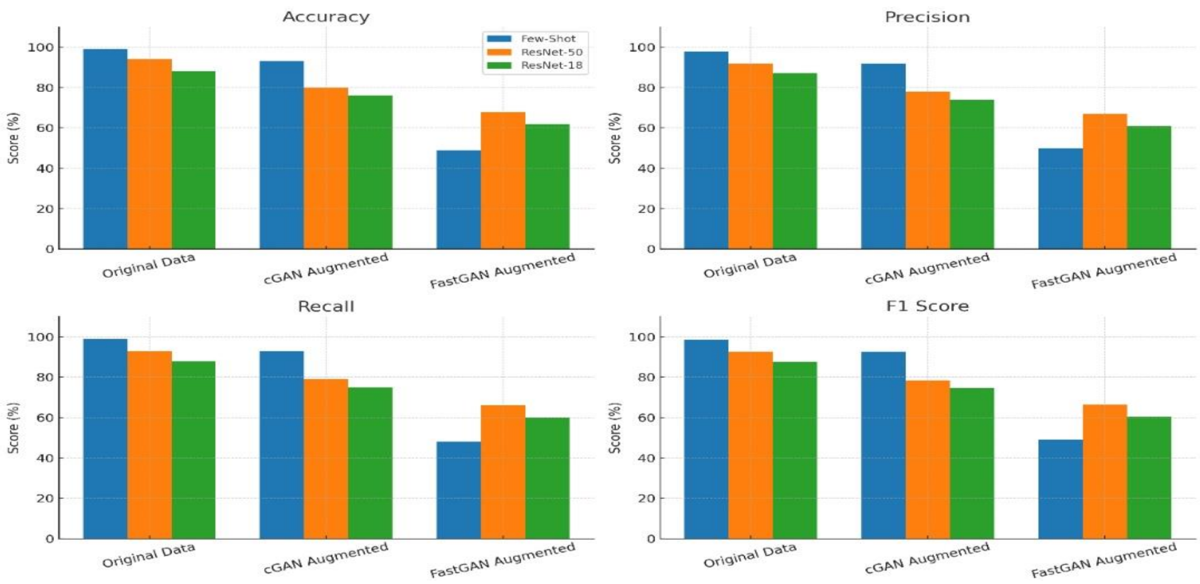

Figure 4: Comparision of Accuracy, Precision, Recall and F1-Score across models

2. Impact on Few-Shot Learning Performance

The highest extreme variance was found in the Prototypical Network (Fig. 4). It was almost perfect with 99% accuracy on the original data, and the confusion matrix is clean and there is great separation between classes. Accuracy using cGAN-augmented data was 93, but the confusion matrix showed that there were minor changes in inter-class confusion. Importantly, when using FastGAN-augmented data, performance reduced to 49% which is close to mere chance when dealing with an 8-class problem and the confusion matrix presented with many misclassifications. It means that the unconditional FastGAN produced features that were semantically incompatible with actual morphology of diseases, poisoning the metric embedding space on which few-shot classification was to occur.

3. Impact on Supervised Learning Performance

The ResNet models were similar in terms of the ordinal pattern: Baseline > cGAN > FastGAN (Fig. 4). As was anticipated, ResNet-50 performed better than ResNet-18 under all conditions because it is more profound. The fact that the performance between baseline and cGAN-augmented data declined by 13 percent and 12 percent, respectively, in ResNet-50 andResNet-18, indicates that cGAN data achieved valuable variance but might have introduced some distributional shift or noise. Nonetheless, the decrease in accuracy with FastGAN-augmented data was much more dramatic (-26% with ResNet-50, -27% with ResNet-18), which validates the unconditional synthetic images had an element of features harmful to learning that distincts a robust decision boundary.

4. Discussion

The Excellence of Semantic Fidelity: The higher results of the cGAN compared to the FastGAN makes the paramount importance of semantic fidelity to medical AI. FastGAN is a fast and stable model that can generate unconstrained and feasible semantically plausible images. This may appear in the medical field as biologically implausible lesion textures, boundaries or color distributions fantasy features that are damaging noise during training. The model is a clean-conditioned cGAN and can be regularised to generate images based on the actual appearance of a particular class, thus producing more pedagogically useful synthetic data.

The High Sensitivity of Few-Shot Learners: Prototypical Network with FastGAN data catastrophic failure underscores the extreme sensitivity of metric-based learning to the quality of the data. The FSL models are based on the building of a continuous embedding space in which distance is related to semantic similarity. Loud or distracting artificial data disastrously interferes with this space, giving rise to incorrect prototypes and unsuccessful generalization. This renders FSL an effective but challenging paradigm, which needs genuinely pristine or sincerely augmented information.

Practical Implications and the challenge of High Baseline: The high baseline accuracies (89%-99%) on the original dataset would suggest that it was well structured and learnable This resulted in a ceiling effect so that it was hard to depict the simple improvement by augmentation. Thus, it is not important to find that augmentation was effective, but rather that cGAN augmentation did not affect performance but FastGAN augmentation was highly deleterious. This is an important lesson to practitioners: the conditional, targeted, generative model is a less risky, more stable option to use when augmenting small medical datasets than an unconditional one.

Limitations and Future Work: The research presented is limited to one clinical dataset and two GAN architectures. The findings on dermoscopic images and other areas of medicine should be supported by future work. It is reasonable to move on to more developed conditional generators, such as StyleGAN2-ADA or Latent Diffusion Models. Moreover, augmentation pipelines could be improved by creating ways of filtering or determining the usefulness of individual synthetic samples prior to incorporating them into training.

The excessive high baseline accuracies (i.e., up to 99 percent) indicate that there is a possibility of either the simplicity of the data sets, a lack of variability, or that the data is leaked, thus, affecting the validity and generalizability of the results. Even the absence of cross-dataset validation or stringent cross-validation prevents the confidence of the model to be applied to the real world.

Conclusion

The present research gives a direct, empirical response to the question concerning the GAN selection in medical image augmentation. Conditional GAN (cGAN) was superior to an unconditional, speed-optimized GAN (FastGAN) in a direct comparison on a small-scale, clinically-relevant task of classifying small clinical skin diseases, in both settings of standard supervised (ResNet) and Few-Shot Learning (Prototypical Network) classifiers.

The most important conclusion is that in medical use, the semantic faithfulness of conditional generation is inalienable. Although unconditional GANs provide performance, it is prone to be filled with semantically inconsistent samples that adversely affect model performance. This is especially disastrous to Few-Shot Learning paradigms. Thus, we firmly believe in the application of the selective, conditional generative models in increasing the limited medical data in constructing robust and dependable AI-based diagnostic instruments.
The intended work in the future will involve incorporating the most recent conditional generative models and building smart methods of synthetic data curation.

References

Biswas, S. (2023). *Skin-Disease-Dataset*. (Kaggle) Retrieved March 16, 2025, from https://www.kaggle.com/datasets/subirbiswas19/skin-disease-dataset

Chen et al. (2024). *Few-Shot Classification with Multiscale Feature Fusion for Clinical Skin Disease Diagnosis.* USA: National Library of Medicine (National Center for Biotechnology Information).

Gupta et.al. (2022). Skin Lesion Detection using VGG-16 and ResNet-50 based Hybrid CNN Model. *International Journal Of Creative Research Thoughts(IJCRT), 10*(6).

Liu et al. (2021). *TOWARDS FASTER AND STABILIZED GAN TRAINING FOR HIGH-FIDELITY FEW-SHOT IMAGE SYNTHESIS.* USA: Department of Computer Science, Rutgers University.

Montenegro et al. (2021). *Privacy-Preserving Generative Adversarial Network for Case-Based Explainability in Medical Image Analysis.* Porto, Portugal: Faculty of Engineering, Univercity of Porto.

Mukherjee, S. (2022, August 18). *The Annotated ResNet-50.* Retrieved June 18, 2025, from https://medium.com/data-science/the-annotated-resnet-50-a6c536034758

Mustafa Et al. (2024). *Deep learning-based skin lesion analysis using hybrid ResUNet++ and modified AlexNet-Random Forest for enhanced segmentation and classification.* Lahore, Pakistan: PLOS ONE.

Özdemir et al. (2024, April 25). *META-TRANSFER DERM-DIAGNOSIS: EXPLORING FEW-SHOT LEARNING AND TRANSFER LEARNING FOR SKIN DISEASE CLASSIFICATION IN LONG-TAIL DISTRIBUTION.* Retrieved from https://arxiv.org/abs/2404.16814

Zhao et al. (2020, June 18). *Differentiable Augmentation for Data-Efficient GAN Training.* Retrieved July 18, 2025, from arXiv:2006.10738