MDPI

Article

Utilizing Generative Adversarial Networks for Acne Dataset Generation in Dermatology

Aravinthan Sankar [1], Kunal Chaturvedi [1], Al-Akhir Nayan [2], Mohammad Hesam Hesamian [1], Ali Braytee [1] and Mukesh Prasad [1,*]

1   School of Computer Science, Faculty of Engineering and Information Technology,
    University of Technology Sydney, Sydney, NSW 2007, Australia; hesam.hesamian@uts.edu.au (M.H.H.)
2   Department of Computer Engineering, Chulalongkorn University, Bangkok 10330, Thailand
*   Correspondence: mukesh.prasad@uts.edu.au

Abstract: Background: In recent years, computer-aided diagnosis for skin conditions has made significant strides, primarily driven by artificial intelligence (AI) solutions. However, despite this progress, the efficiency of AI-enabled systems remains hindered by the scarcity of high-quality and large-scale datasets, primarily due to privacy concerns. Methods: This research circumvents privacy issues associated with real-world acne datasets by creating a synthetic dataset of human faces with varying acne severity levels (mild, moderate, and severe) using Generative Adversarial Networks (GANs). Further, three object detection models—YOLOv5, YOLOv8, and Detectron2—are used to evaluate the efficacy of the augmented dataset for detecting acne. Results: Integrating StyleGAN with these models, the results demonstrate the mean average precision (mAP) scores: YOLOv5: 73.5%, YOLOv8: 73.6%, and Detectron2: 37.7%. These scores surpass the mAP achieved without GANs. Conclusions: This study underscores the effectiveness of GANs in generating synthetic facial acne images and emphasizes the importance of utilizing GANs and convolutional neural network (CNN) models for accurate acne detection.

Keywords: acne; skin disease; GANs; synthetic; object detection

1. Introduction

Facial skin problems are common and often affect the quality of life of individuals. Among these problems, acne is one of the most prevalent and challenging to diagnose and treat. Acne is a chronic inflammatory skin condition that affects the pilosebaceous units of the skin, resulting in comedones, papules, pustules, nodules, and scars. According to the GBD study, acne affects about 9.4% of the global population [1], making it the eighth most prevalent disease worldwide. Acne can have significant psychological and social impacts on the affected individuals, such as reduced self-esteem, anxiety, depression, and impaired quality of life [2]. Diagnosing and analyzing acne are mainly based on clinical examination and subjective assessment of acne severity by dermatologists. However, these methods are prone to inter- and intra-observer variability and may not reflect the true extent of acne lesions [3]. Therefore, there is a need for objective and reliable methods to detect and quantify acne lesions on facial images.

In recent years, there has been a growing interest in the automatic classification of skin diseases, driven by the potential to enhance diagnostic accuracy and reduce mortality rates. This is due to the surge in growth and interest in computer vision-based approaches that are low-cost and transparent in decision-making [4–8]. These computer-aided diagnostics (CAD) methods, which are non-invasive, rely on accurately diagnosing dermoscopic images. Dermoscopic images provide a detailed visual interpretation of lesioned features, facilitating the detection of skin diseases. Deep learning methods, particularly those based on CNNs, have shown promising results in dermatology, including skin cancer detection,

skin lesion segmentation, and facial skin analysis. For instance, Shah Junayed et al. [9] introduced AcneNet, a deep Convolutional Neural Network (CNN)-based classification approach for acne classes, demonstrating the potential for automated diagnosis in dermatology. Srinivasu et al. [10], Wu et al. [11], and Xiang et al. [12] applied deep learning techniques for skin disease classification, utilizing neural networks like MobileNet V2, CNNs, and LSTM for improved diagnostic accuracy. Zhao et al. [13] and Liu et al. [14] focused on interpretable skin lesion classification using novel Convolutional Neural Network (CNN) algorithms and mole detection and segmentation software for mobile phone skin images.

Despite advancements in architecture leading to improved performance, there remains a crucial need for large, meticulously annotated datasets of superior quality [15]. However, in deep learning models, acquiring such datasets, especially in dermatology, where data and annotations are challenging, is rare. Existing datasets for acne detection are often small in size, lack diversity, or are not publicly accessible. Recently, Ghorbani et al. [16] introduced DermGAN, a method for the synthetic generation of clinical skin images with pathology, demonstrating the potential of generative models in creating realistic dermatological data. Baur et al. [17] and Zein et al. [18] explored the use of Generative Adversarial Networks (GANs) for generating highly realistic images of skin lesions, with a particular focus on anonymous acneic face dataset generation. Lin et al. [19] proposed a CGPG-GAN that uses a novel Generative Adversarial Network tailored for acne lesion inpainting, significantly enhancing the quality and reliability of subsequent diagnostic processes. AL-SAEDİ et al. [20] conducted a comparative analysis of six transfer learning networks (DenseNet, Xception, InceptionResNetV2, ResNet50, and MobileNet) for the task of skin cancer classification, leveraging the International Skin Imaging Collaboration (ISIC) dataset. Their findings underscored the efficacy of augmentation in enhancing classification performance, yielding notable improvements in accuracy and F-scores while mitigating false negatives. In a separate study, Alhudhaif et al. [21] proposed a deep learning approach incorporating attention mechanisms and supported by data balancing techniques. Leveraging the HAM10000 dataset comprising 10,015 labeled images across seven distinct skin lesion types, they initially observed accuracy rates of 85.73% for training, 70.90% for validation, and 69.75% for testing on the unbalanced dataset.

Detecting small and subtle acne lesions on facial images poses an additional challenge, often overlooked or mistaken for other skin features by deep learning models [22,23]. To tackle these hurdles, we present a novel approach comprising two primary steps: (1) generating a realistic synthetic dataset of human faces exhibiting varying degrees of acne severity using a conditional GAN framework; and (2) assessing the efficacy of the synthetic dataset by training and testing three CNN models for acne lesion detection on facial images. We have employed YOLOv5, YOLOv8, and Detectron2 models due to their consistent performance with minimal background errors, robust accuracy across variations in lighting, skin tones, and image quality, and their effectiveness in detecting small objects in this study.

The primary contributions of this paper are outlined as follows:

- Introduce the utilization of a conditional GAN framework for generating a large and diverse synthetic dataset of human faces afflicted with acne. The conditional GAN framework enables the modulation of acne severity levels in the generated faces.
- Conduct a comprehensive investigation into state-of-the-art object detection models, including YOLOv5, YOLOv8, and Detectron2, to compare their performance with and without integrating the synthetic dataset as an additional training data source.
- Perform extensive experiments that showcase the capability of the proposed approach to enhance accuracy and robustness in acne detection on real facial images. Furthermore, we demonstrate that the synthetic dataset enhances the generalization ability of deep learning models.
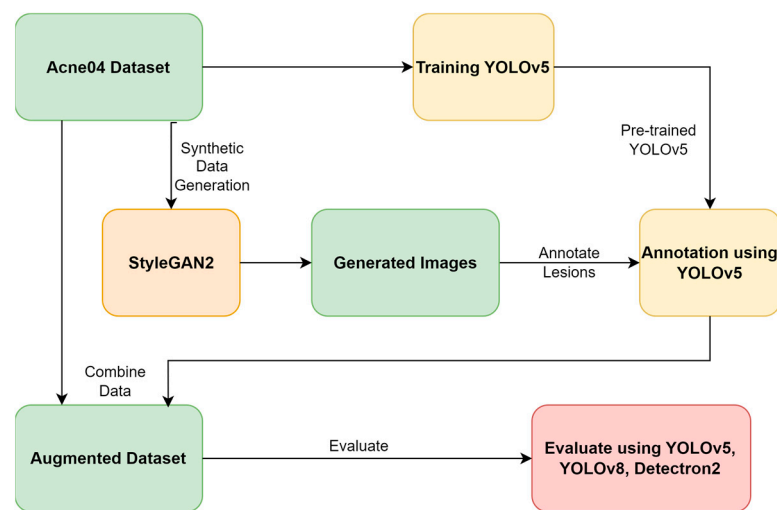
The subsequent sections of this paper are structured as follows: Section 2 outlines the methodology of the proposed approach, encompassing the dataset-gathering process

and the algorithms employed for model training. Section 3 elaborates on the experimental details and the preprocessing steps to enhance performance. Section 4 showcases and deliberates on the outcomes derived from the experiments. Finally, Section 5 concludes the paper and delineates potential avenues for future research.

2. Materials and Methods

2.1. Proposed Method

Figure 1 illustrates the block diagram representing the methodology followed in this study to achieve the research objective. The process begins by gathering facial acne images from various sources. These images are then subjected to data processing to ensure compatibility with the input layer of the StyleGAN2 model. The StyleGAN2 models are trained with different levels of facial acne severity using transfer learning. The output images generated by the GANs are then categorized into three classes: mild, moderate, and severe.

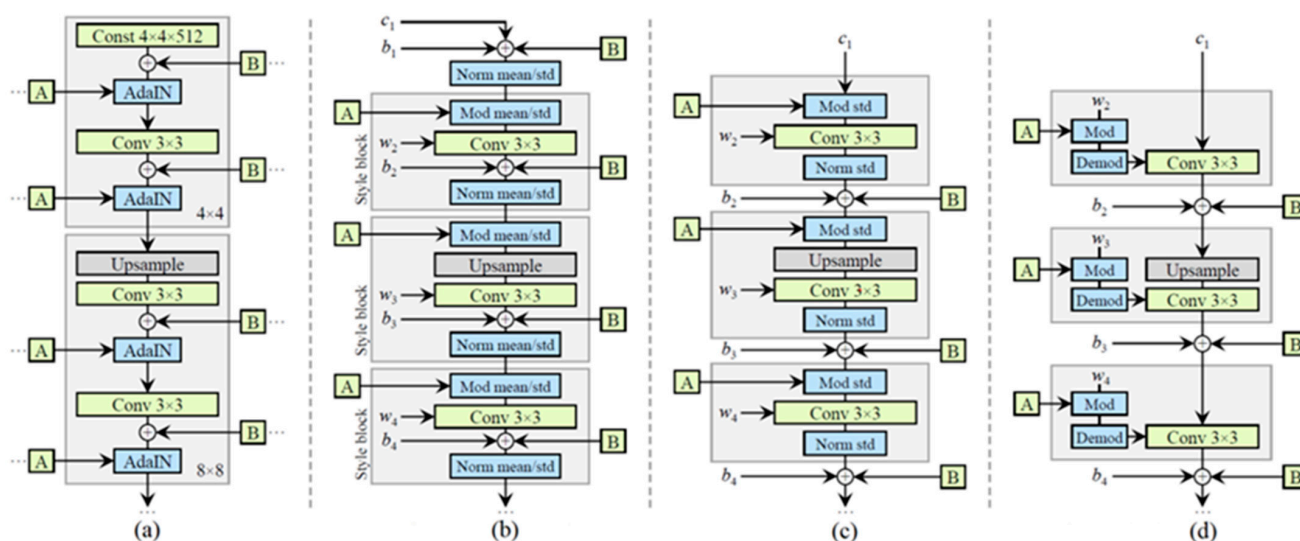**Figure 1.** Methodology diagram.

Additionally, a synthetic healthy class is added to the dataset. Subsequently, a CNN model is trained using the synthetic dataset created by the GANs. The objective is to evaluate the CNN model's performance on real facial acne images. In this study, StyleGAN2 was chosen due to its ability to generate high-resolution images (up to 1024×1024) with realistic results. Although the initial dataset comprised 1473 images representing various facial acne severities, training individual GANs for each severity level was insufficient. To address this limitation, the following steps were taken: the image size was reduced to 256×256, and the mild, moderate, and severe classes were combined into a single class called "acne." By following this methodology, the research aims to leverage the capabilities of StyleGAN2 to generate synthetic facial acne datasets and subsequently train a CNN model for evaluating its performance on real-world facial acne images.

2.2. Generative Adversarial Networks (GANs)

We use GANs to create synthetic dermatological images that resemble real data. Goodfellow et al. [24] presented the foundational work on Generative Adversarial Networks (GANs), laying the groundwork for subsequent advancements in generating realistic dermatological images. GANs are composed of two models: a generator and a discriminator. The generator produces practical samples from random noise while the discriminator distinguishes between real and fake samples. The two models are trained adversarially, using backpropagation to minimize their losses. Karras et al. [25] contributed to the field by analyzing and improving the image quality of StyleGAN. We leverage StyleGAN2,

a state-of-the-art GAN model developed by NVIDIA, which introduces several improvements over its predecessor, StyleGAN.

StyleGAN was proposed by NVIDIA in 2018, and it significantly enhanced the generator architecture. Some of these enhancements are Baseline Progressive GAN, Tuning and Bilinear Up/Down Sampling, Mapping, Styles, and Adaptive Instance Normalization, Removal of Latent Vector Input, Noise Addition at Each Block, and Addition of Mixing Regularization. These features improved the quality and diversity of the synthetic images generated by StyleGAN. Figure 2 compares the architectures of StyleGAN and StyleGAN2 and highlights the changes made in StyleGAN2 to improve its performance further. The main change is the replacement of adaptive instance normalization (AdaIN) with weight demodulation and the relocation of noise addition outside the generator blocks.

Figure 2. Architecture of StyleGAN2 [25] (**a**) StyleGAN (**b**) StyleGAN (detailed) (**c**) Revised architecture (StyleGAN2) (**d**) Weight demodulation of StyleGAN2.

Figure 2a,b depict the original StyleGAN architecture, while Figure 2c shows the modified StyleGAN2 architecture, which does not use AdaIN. Figure 2d shows the Style-GAN2 uses weight demodulation, which divides the weight of the convolutional layer by its standard deviation. This modification helps achieve a normalization effect similar to AdaIN but with better performance and stability. The transition from StyleGAN to Style-GAN2 demonstrates the superiority of weight demodulation over AdaIN for normalization purposes. Using this change, we obtained remarkable results in generating high-quality synthetic images by StyleGAN2 in this study.

2.3. Acne Detection

2.3.1. YOLOv5 and YOLOv8

We use two state-of-the-art object detection models, YOLOv5 [26] and YOLOv8 [27], to compare their performance on various datasets and tasks. These models are based on the YOLO algorithm, which stands for You Only Look Once, and use a single neural network to process an entire image.

YOLOv5 is a model developed by Ultralytics, a company specializing in computer vision and deep learning. It is built on the PyTorch framework, making it easy for developers to use and deploy. YOLOv5 has five different sizes: n, s, m, l, and x, which vary in the number of layers and parameters. The larger the size, the higher the accuracy and the slower the speed. YOLOv5's architecture consists of three main parts: Backbone, Neck, and Head. The Backbone is the main body of the network, which uses the CSP-Darknet53 structure to extract features from the image. The Neck connects the Backbone and the Head and uses SPPF and PANet structures to generate feature pyramids for different scales.

The Head is responsible for generating the final output, which includes classes, objectness scores, and bounding boxes.

YOLOv8 is a model proposed by Augmented Startups, which provides online courses and tutorials on artificial intelligence and computer vision. It is also built on the PyTorch framework but claims to have some improvements over YOLOv5. YOLOv8 provides a unified framework for training models for object detection, instance segmentation, and image classification. It also introduces new features, such as Dynamic Anchor Boxes, which allow the model to adjust the anchor boxes according to the input image size and aspect ratio. This improves accuracy and reduces the computational cost. Attention Mechanism: This feature enables the model to focus on the most relevant parts of the image and ignore the background noise. This enhances the performance and reduces false positives. Multi-Head Output: This feature allows the model to output multiple types of predictions, such as bounding boxes, masks, and labels. This enables the model to handle different tasks with a single network.

We use these models to experiment with various datasets and the custom dataset. We evaluate their performance on metrics such as mAP, inference time, memory usage, and power consumption. We also compare their results on object detection, instance segmentation, and image classification tasks. We aim to determine which model suits different scenarios and applications.

2.3.2. Detectron2

We use Detectron2 [28], a popular and powerful deep learning framework for computer vision tasks, to perform object detection and instance segmentation on the custom dataset. Detectron2 is an open-source framework developed by Facebook AI Research (FAIR) that builds on the PyTorch framework. It offers a flexible and modular design, allowing us to customize and extend it to meet needs quickly. We use object detection to identify and locate objects within an image. We use instance segmentation to provide pixel-level segmentation masks for each object instance. These tasks allow us to analyze and understand the visual content within the images in more detail.

We leverage the features and capabilities of Detectron2 to facilitate the development and evaluation process. The framework provides efficient data loading and augmentation techniques, distributed training across multiple GPUs, pre-processing and post-processing utilities, evaluation metrics, and visualization tools. The framework also provides a collection of pre-trained models, which we fine-tune on the dataset to achieve better performance. The highly modular framework allows us to experiment with different model components, loss functions, and network architectures. The framework offers comprehensive APIs for seamless integration with other deep learning frameworks and tools.

Detectron2 is a widely used and actively developed framework in the computer vision community. It has been applied to various applications, such as object detection in images and videos, instance segmentation, pose estimation, and more. The flexibility and extensibility of Detectron2 make it a valuable tool for the research.

3. Experiments

3.1. Datasets

Wu et al. [11] created the ACNE04 dataset, a public dataset of facial images with acne annotations. The dataset has 1457 pictures of different sizes, each with a label of acne severity and lesion count. Professional dermatologists marked each lesion with a rectangle. Figure 3 shows some examples from the ACNE04 dataset. The dataset is imbalanced, with more images having low lesion counts (less than 10) and fewer images having high lesion counts (between 40 and 50). The highest lesion count in an image is 65, and the lowest is 1.
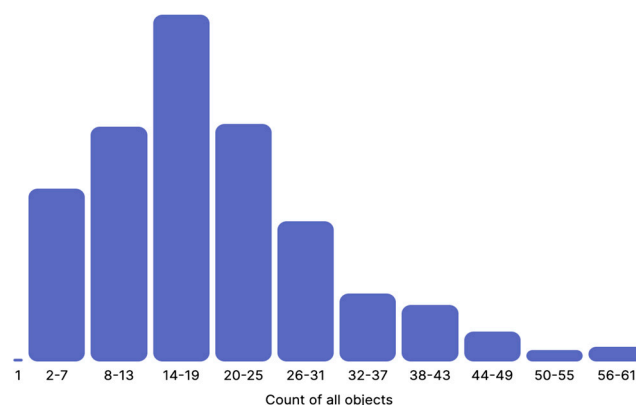
Figure 3. Images from ACNE04 dataset.

3.2. Dataset Expansion with GAN

After reclassifying the severity of the ACNE04 dataset, the next step in data preparation involved generating additional images using StyleGAN Ada. By leveraging the power of GANs, we aimed to expand the dataset and increase its size to 4000 images. Using the StyleGAN Ada model, we applied its generative capabilities to create new synthetic facial images. This process involved learning the underlying patterns and features from the original ACNE04 dataset and generating realistic-looking images that mimic the characteristics of the existing acne images. Using StyleGAN Ada, we introduced diversity to the dataset, including acne variations, skin types, and other facial attributes. The generated images were seamlessly integrated with the original ACNE04 dataset, resulting in a more extensive and diverse dataset for further analysis and model training. Figure 4 shows the histogram of lesion counts per image after using StyleGAN2.
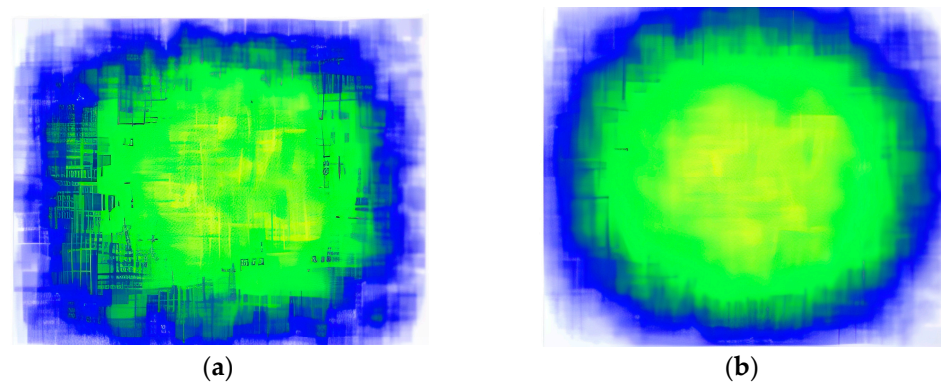

Figure 4. Histogram of number of acne lesions counted by images using StyleGAN2.

Expanding the dataset through GAN-generated images was crucial in addressing the limitations of the original dataset size, providing more samples for training and evaluation purposes. This augmented dataset would improve the performance and robustness of the acne detection models developed using the ACNE04 dataset.

3.3. Annotation Heatmaps

During the analysis of the ACNE04 dataset, we observed that the annotation heatmaps generated from this dataset were not optimized—exemplified by the incomplete filling of the bounding boxes with the green color in Figure 5a. The suboptimal heatmaps indicate that the annotations may not accurately capture the full extent of the acne lesions in the images. This issue can affect the performance of object detection models trained on this dataset. To address this limitation, we used various techniques to enhance the annotation process and improve the quality of the heatmaps. By implementing refined annotation methodologies and employing expert dermatologists' guidance, we aimed to ensure that the bounding boxes accurately encompassed the acne lesions, resulting in more precise and comprehensive heatmaps.

(**a**) (**b**)

Figure 5. Annotation heatmaps: (**a**) incomplete and (**b**) improved version.

Furthermore, we explored different augmentation techniques, such as 90° Rotate Clockwise, Counter-Clockwise, Shear ±9° Horizontal, ±23° Vertical, Hue Between −25° and +25°, and Bounding Box: Flip Horizontal, Vertical. These techniques helped increase the dataset's diversity and richness, providing a more robust and representative training set for the object detection models. Figure 5b shows the improved version of the heatmaps of the datasets.

3.4. Implementation Details

To evaluate the performance of the proposed model, we used a dataset of 1457 facial images. We use GAN to increase the diversity and robustness of the dataset. We then divided the augmented dataset into three subsets: training, testing, and validation, following a 70/20/10 ratio. We ensured each subset contained a proportional representation of images with and without acne. Table 1 shows the total number of images generated and the split used to train the model.

Table 1. Total number of images.

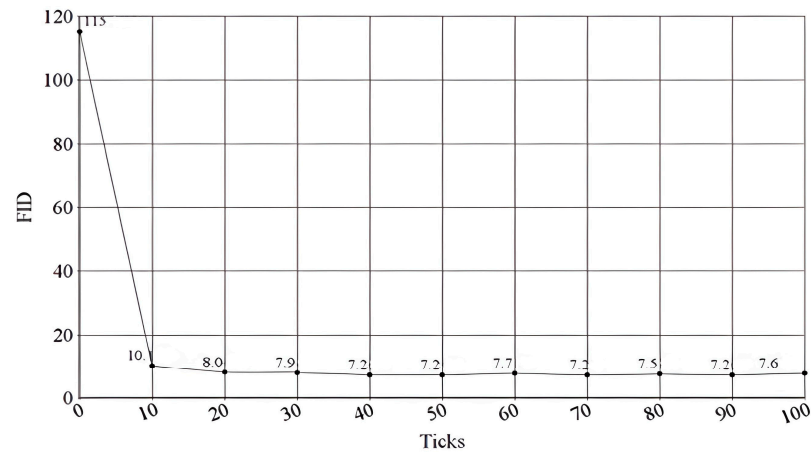| Type	Number of Images without GAN		Number of Images with GAN	
	Percentage	Total	Percentage	Total
Train	70%	1020 images	70%	3441 images
Test	20%	291 images	20%	983 images
Validation	10%	145 images	10%	491 images

3.5. Evaluation Metrics

To evaluate the performance of the StyleGAN2 models, we utilized the Fréchet Inception Distance (FID) as a metric for assessment. The FID metric measures the similarity between synthetic and real images at a deep convolutional layer known as InceptionV3. Rather than comparing images on a pixel-by-pixel basis, the FID metric calculates the mean and standard deviation to capture the differences in the distribution of synthetic and real images. The FID score provides a quantitative measure of the dissimilarity between the generated and real image sets, with lower scores indicating a higher level of similarity between the distributions.

In this paper, we conducted various experiments to compare the performance of custom-trained models using YOLOv5, YOLOv8, and Detectron2. The evaluation focused on precision, recall, mAP@0.5, and mAP@0.5:0.95, key metrics for assessing the overall detection performance.

4. Results and Discussion

4.1. Synthetic Data Analysis

After training the facial acne model using StyleGAN2, we generated a dataset of synthetic acne images. The results of the model are presented in Figure 6, showcasing the Fréchet Inception Distance (FID) at each iteration.

Figure 6. StyleGAN2 model FID graph.

The training process for the StyleGAN2 model commenced with a random generation of facial acne images. At the start (tick 0), the model had an initial FID of 115.19. As the model learned and adapted during training, the FID steadily decreased. The lowest FID value of 7.209 was achieved at tick 70. The training duration for this model was approximately five days. These outcomes demonstrate the effectiveness of the StyleGAN2 model in generating synthetic facial acne images that exhibit reduced dissimilarity to real images, as indicated by the significant reduction in FID. The random generation of acne images contributed to the overall improvement and quality of the generated images by the StyleGAN2 model.
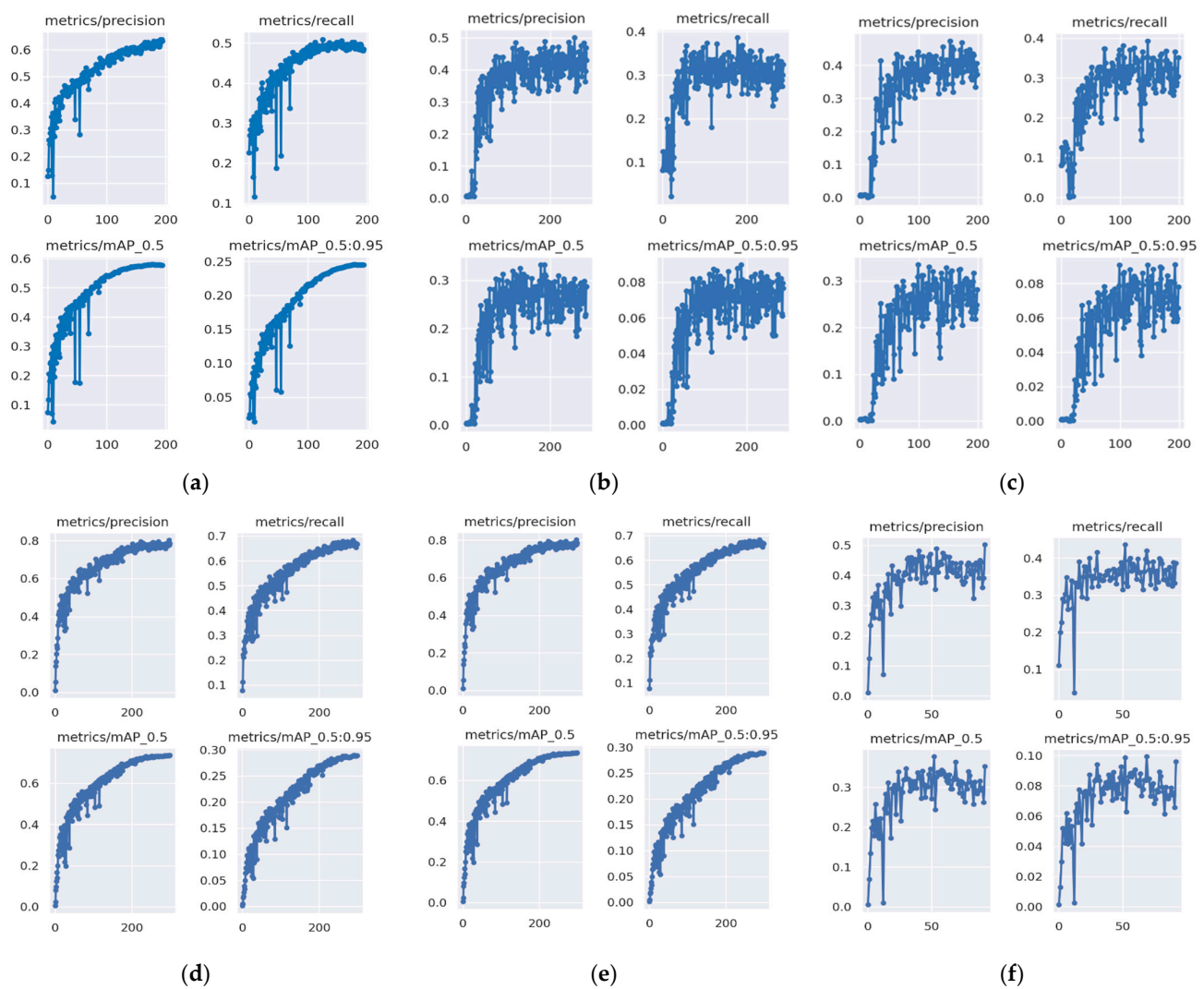
4.2. Synthetic Data Analysis

In this performance evaluation, we assess the capabilities of the YOLOv5, YOLOv8 and Detectron2 models for acne image detection. Unlike other assessments that incorporated GAN-generated datasets, this evaluation focuses solely on the performance of YOLOv5, YOLOv8, and Detectron2 without GAN. We compare the performance of these models to determine their effectiveness. The performance score is mentioned in Table 2, and more details are illustrated in Figure 7.

Table 2. YOLOv5, YOLOv8, and Detectron2 mAP, precision, and recall results.

Model Name	mAP@0.5:0.95	Precision	Recall	F1Score
YOLOv5	58.0%	60.6%	50.1%	54.85%
YOLOv8	66.2%	72.5%	60.8%	66.14%
Detectron2	33.5%	45.1%	35.3%	39.60%

Figure 7. Evaluation metrics with StyleGAN2 for (**a**) YoloV5, (**b**) YoloV8, (**c**) Detectron2 and without StyleGAN2 for (**d**) YoloV5, (**e**) YoloV8, (**f**) Detectron2.

The "metrics/precision" graph displays the precision metric over different evaluation iterations. Precision measures the proportion of correctly predicted positive samples (true positives) out of the total predicted positives (true positives + false positives). This graph provides insights into how well the model performs regarding precision, indicating the accuracy of the model's optimistic predictions.

The "metrics/recall" graph illustrates the recall metric throughout the evaluation process. Recall calculates the ratio of correctly predicted positive samples (true positives) to the total number of actual positives (true positives + false negatives). By analyzing this graph, we can assess the model's ability to identify and retrieve all relevant objects of interest, as recall captures the model's ability to find positive instances.

The "metrics/mAP_0.5" graph displays the mAP metric at an intersection over a union (IoU) threshold of 0.5. The mAP calculates the average precision across different object classes and IoU thresholds, providing an overall assessment of the model's detection accuracy. This graph allows us to track the model's mAP performance at the IoU threshold of 0.5.

The "metrics/mAP_0.5:0.95" graph represents the mAP metric across various IoU thresholds ranging from 0.5 to 0.95. This graph comprehensively evaluates the model's detection accuracy across a range of IoU thresholds. It allows us to analyze the model's

performance at different levels of object overlap and assess its robustness in detecting objects accurately.

4.3. CNN Models' Performance Evaluation with StyleGAN2

We conducted experiments with three CNN models for acne image detection: YOLOv5, YOLOv8, and Detectron2. We trained these models using a combination of synthetic images and real images from the ACNE04 dataset. The objective was to evaluate their accuracy, precision, recall, and mAP.

Table 3 presents the evaluation results. Among the models tested, YOLOv8 demonstrated the best performance, with a mAP of 73.6%, precision of 80.2%, and recall of 65.3%. These metrics indicate the model's ability to classify acne images accurately. However, it is worth noting that the YOLOv5 model also performed well, achieving a mAP of 73.5%, precision of 76.1%, and recall of 68.1%. Meanwhile, the Detectron2 model showed lower results, with a mAP of 37.7%, precision of 42.1%, and recall of 43.6%. These results show that YOLOv8 outperformed YOLOv5 and Detectron2 regarding acne image detection accuracy, precision, and recall. It exhibited the highest mAP, indicating a superior overall detection performance. The graphs in Figure 8 visually represent how the model's performance changes during evaluation.

Table 3. YOLOv8, YOLOv5, and Detectron2 mAP, precision, and recall results.

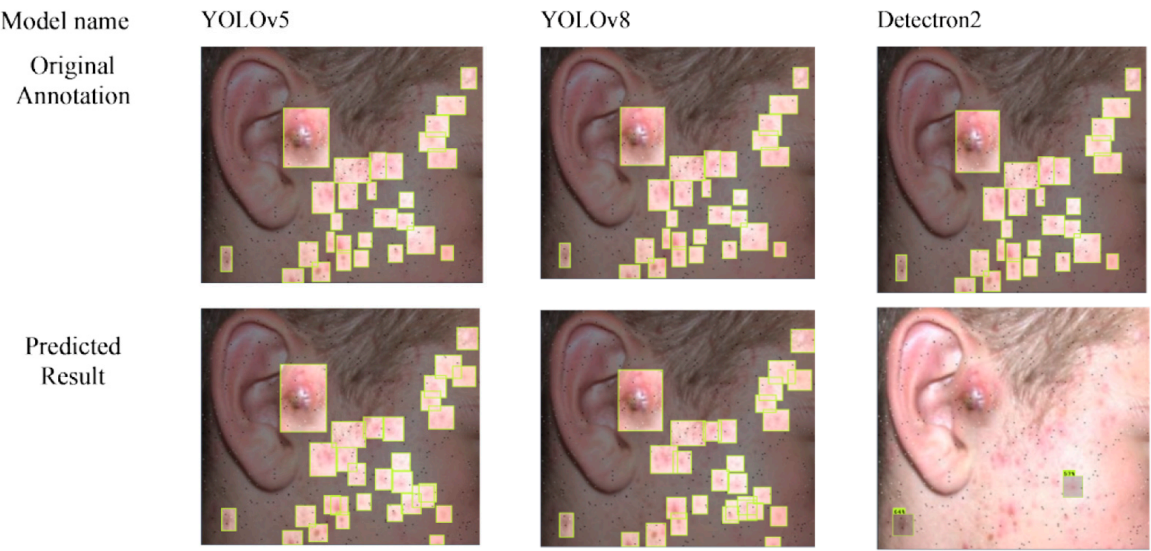| Model Name	mAP@0.5:0.95	Precision	Recall	F1Score
YOLOv5	73.5%	76.1%	68.1%	71.88%
YOLOv8	73.6%	80.2%	65.3%	71.99%
Detectron2	37.7%	42.1%	43.6%	42.84%

Figure 8. Original annotations and the predicted annotations for object detection with Yolov8, Yolov5, and Detectron2.

Figure 8 below showcases both the original and predicted annotations for object detection. The original annotations were manually added to accurately label the objects in the images, serving as a reference. Additionally, we utilized advanced models, namely, Yolov5, Yolov8, and Detectron2, to generate the predicted annotations. These models have proven to be effective in detecting and classifying objects. By comparing the predicted annotations with the original ones, we can evaluate the models' performance and assess their accuracy in detecting objects in the given images.

5. Conclusions

This study successfully addressed privacy concerns associated with acne datasets by utilizing Generative Adversarial Networks (GANs). A synthetic dataset of human faces with varying acne severity levels (mild, moderate, and severe) was created using StyleGAN2-ADA, providing a realistic and anonymous alternative to real face images. The scientific community can freely use this synthetic dataset. Three CNN-based models—YOLOv5, YOLOv8, and Detectron2—were trained and compared to evaluate acne detection accuracy with real face images. Notably, YOLOv8 emerged as the top performer in acne detection. Furthermore, the effectiveness of the StyleGAN2 model in generating synthetic facial acne images was demonstrated. These findings are foundational for future dermatology and computer vision research and applications.

Due to the lack of balanced data, the models struggled to maintain optimal accuracy. Nevertheless, the primary concern of strict privacy maintenance was upheld. Time constraints led us to conduct experimental analyses using only three models and one dataset; however, additional deep learning models and public datasets are available. In future endeavors, we aim to address the limitations of the current approach and improve the detection accuracy by addressing class imbalances through techniques such as oversampling or undersampling of acne severity levels.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available publicly.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tan, J.K.L.; Bhate, K. A global perspective on the epidemiology of acne. *Br. J. Dermatol.* **2015**, *172*, 3–12. [CrossRef]
2. Gallitano, S.M.; Berson, D.S. How Acne Bumps Cause the Blues: The Influence of Acne Vulgaris on Self-Esteem. *Int. J. Women's Dermatol.* **2018**, *4*, 12–17. [CrossRef]
3. Shen, X.; Zhang, J.; Yan, C.; Zhou, H. An Automatic Diagnosis Method of Facial Acne Vulgaris Based on Convolutional Neural Network. *Sci. Rep.* **2018**, *8*, 5839. [CrossRef]
4. Wei, X.; Zhang, L.; Zhang, J.; Wang, J.; Liu, W.; Li, J.; Jiang, X. Decoupled Sequential Detection Head for accurate acne detection. *Knowl. Based Syst.* **2024**, *284*, 111305. [CrossRef]
5. Gayap, H.T.; Akhloufi, M.A. Deep Machine Learning for Medical Diagnosis, Application to Lung Cancer Detection: A Review. *BioMedInformatics* **2024**, *4*, 15. [CrossRef]
6. Ali, M.A.; Dornaika, F.; Arganda-Carreras, I.; Ali, H.; Karaouni, M. Naturalize Revolution: Unprecedented AI-Driven Precision in Skin Cancer Classification Using Deep Learning. *BioMedInformatics* **2024**, *4*, 35. [CrossRef]
7. Hong, G.-J.; Li, D.-L.; Pare, S.; Saxena, A.; Prasad, M.; Lin, C.-T. Adaptive Decision Support System for On-Line Multi-Class Learning and Object Detection. *Appl. Sci.* **2021**, *11*, 11268. [CrossRef]
8. Zhang, L.; Li, J.; Huang, T.; Ma, Z.; Lin, Z.; Prasad, M. GAN2C: Information Completion GAN with Dual Consistency Constraints. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; IEEE: New York, NY, USA, 2018; pp. 1–8. [CrossRef]
9. Junayed, M.S.; Jeny, A.A.; Atik, S.T.; Neehal, N.; Karim, A.; Azam, S.; Shanmugam, B. AcneNet—A Deep CNN Based Classification Approach for Acne Classes. In Proceedings of the 2019 12th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 18 July 2019; IEEE: New York, NY, USA, 2019; pp. 203–208. [CrossRef]
10. Srinivasu, P.N.; Sivasai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors* **2021**, *21*, 2852. [CrossRef]

11. Wu, Z.; Zhao, S.; Peng, Y.; He, X.; Zhao, X.; Huang, K.; Wu, X.; Fan, W.; Li, F.; Chen, M.; et al. Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images. *IEEE Access* **2019**, *7*, 66505–66511. [CrossRef]

12. Jha, A.; Ananthanarayana, V.S. An Interpretable Deep Learning Model for Skin Lesion Classification. In *Advances in Computing and Data Sciences, Proceedings of the 7th International Conference, ICACDS 2023, Kolkata, India, 27–28 April 2023*; Springer: Cham, Switzerland, 2023; pp. 543–553. [CrossRef]

13. Zhao, Z.; Wu, C.-M.; Zhang, S.; He, F.; Liu, F.; Wang, B.; Huang, Y.; Shi, W.; Jian, D.; Xie, H.; et al. A Novel Convolutional Neural Network for the Diagnosis and Classification of Rosacea: Usability Study. *JMIR Med. Inform.* **2021**, *9*, e23415. [CrossRef]

14. Liu, S.; Chen, Z.; Zhou, H.; He, K.; Duan, M.; Zheng, Q.; Xiong, P.; Huang, L.; Yu, Q.; Su, G.; et al. DiaMole: Mole Detection and Segmentation Software for Mobile Phone Skin Images. *J. Healthc. Eng.* **2021**, *2021*, 1–10. [CrossRef]

15. Tajbakhsh, N.; Jeyaseelan, L.; Li, Q.; Chiang, J.N.; Wu, Z.; Ding, X. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Med. Image Anal.* **2020**, *63*, 101693. [CrossRef]

16. Ghorbani, A.; Natarajan, V.; Coz, D.; Liu, Y. DermGAN: Synthetic Generation of Clinical Skin Images with Pathology. *Proc. Mach. Learn. Res.* **2019**, *116*, 155–170.

17. Baur, C.; Albarqouni, S.; Navab, N. Generating Highly Realistic Images of Skin Lesions with GANs. *arXiv* **2018**, arXiv:1809.01410.

18. Zein, H.; Chantaf, S.; Fournier, R.; Nait-Ali, A. Generative Adversarial Networks for anonymous Acneic face dataset generation. *arXiv* **2022**, arXiv:2211.04214.

19. Lin, Y.; Jiang, J.; Ma, Z.; Chen, D.; Guan, Y.; Liu, X.; You, H.; Yang, J.; Cheng, X. CGPG-GAN: An Acne Lesion Inpainting Model for Boosting Downstream Diagnosis. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 6–8 December 2022; IEEE: New York, NY, USA, 2022; pp. 1634–1638. [CrossRef]

20. Al-Saedi, D.K.A.; Savaş, S. Classification of Skin Cancer with Deep Transfer Learning Method. *Comput. Sci.* **2022**, *IDAP-2022*, 202–210. [CrossRef]

21. Alhudhaif, A.; Almaslukh, B.; Aseeri, A.O.; Guler, O.; Polat, K. A novel nonlinear automated multi-class skin lesion detection system using soft-attention based convolutional neural networks. *Chaos Solitons Fractals* **2023**, *170*, 113409. [CrossRef]

22. Li, H.; Pan, Y.; Zhao, J.; Zhang, L. Skin disease diagnosis with deep learning: A review. *Neurocomputing* **2021**, *464*, 364–393. [CrossRef]

23. Gulzar, Y.; Khan, S.A. Skin Lesion Segmentation Based on Vision Transformers and Convolutional Neural Networks—A Comparative Study. *Appl. Sci.* **2022**, *12*, 5990. [CrossRef]

24. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661.

25. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. *arXiv* **2019**, arXiv:1912.04958.

26. Ultralytics. YOLOv5. Available online: https://github.com/ultralytics/yolov5 (accessed on 8 August 2023).

27. Ultralytics. YOLOv8. Available online: https://github.com/ultralytics/ultralytics (accessed on 10 August 2023).

28. F.A.I. Research. Detectron2. Available online: https://github.com/facebookresearch/detectron2 (accessed on 2 September 2023).