Mitigating Biases in Dermatological Diagnosis through GAN Augmentation and Classification

Ashwitha C Thomas
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
ashwitha.thomas@nitte.edu.in

Preethi Salian K *
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
preethi.salian@nitte.edu.in

Prajwal P
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
4nm21is105@nmamit.in

Nihal
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
4nm22is405@nmamit.in

Kishor S Naik
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
4nm21is068@nmamit.in

Neil Mascarenhas
Department of ISE
Nitte (Deemed to be University)
NMAM Institute of Technology
Nitte, India
4nm21is095@nmamit.in

***Abstract*— Skin conditions affect individuals of all skin tones, yet biases in dermatological diagnosis disproportionately impact those with darker skin, leading to significant healthcare disparities. This paper proposes a comprehensive solution that integrates advanced data preprocessing, Generative Adversarial Network (GAN) augmentation, and Convolutional Neural Network (CNN) models to mitigate these biases. The initial phase involves meticulous preprocessing of diverse dermatological image datasets, ensuring broad coverage of skin tones. Images are standardized, noise is removed, and critical features are preserved for accurate classification. To address data imbalance, particularly for underrepresented skin tones, a GAN with spectral normalization generates synthetic images of dark skin lesions, enriching the dataset and enhancing model generalization. The preprocessed and augmented dataset is used to train three CNN architectures are ResNet50, NASNetLarge, and InceptionResNetV2, each fine-tuned to optimize performance. Class weights handle imbalances, and an ensemble approach improves classification accuracy, particularly in detecting challenging conditions like melanoma across all skin tones. This combined approach of GAN augmentation and CNN-based classification offers a promising path toward more equitable dermatological diagnostics.**

Keywords*— *dermatological diagnosis, healthcare disparities, preprocessing, GAN augmentation, CNN, synthetic images.

I. Introduction

Skin lesions, encompassing a spectrum from benign moles to malignant melanomas, represent a formidable diagnostic challenge in dermatology. The diverse morphological characteristics of these lesions, coupled with their potential for malignancy, necessitate accurate and timely diagnosis for optimal patient care. However, traditional diagnostic approaches reliant solely on human expertise often suffer from subjectivity and interobserver variability, leading to delays in diagnosis and treatment initiations [1].

In recent years, deep learning techniques have emerged as promising tools for automating the analysis and classification of skin lesions, offering the potential to augment the diagnostic capabilities of healthcare professionals [2]. Among these techniques, Generative Adversarial Networks (GANs) have garnered significant attention for their ability to generate synthetic data that closely resembles real world examples. By leveraging the adversarial training paradigm, GANs can learn to generate diverse and realistic images of skin lesions, thereby addressing data scarcity and enhancing dataset diversity[3].

This paper presents an approach for augmenting dermatological datasets using GANs, aimed at overcoming the limitations of traditional data augmentation methods. By generating synthetic images that capture the intricacies of real skin lesions, the methods seek to enrich dataset representation and improve the robustness of deep learning models for skin lesion analysis. The architecture and training procedure of the GAN model employed in the study are also discussed, along with insights into the generated images' quality and diversity.

In addition to GAN augmentation, the work utilizes CNN architectures like ResNet, NASNetLarge, and InceptionResNetV2 for classification. These CNN models, trained on the augmented dataset, are fine-tuned to ensure better generalization across diverse skin tones. The ensemble of these models further enhances the accuracy and fairness of the classification, balancing performance across all skin tones. Through this combined approach, the research aims to improve the accuracy and fairness of dermatological diagnoses for individuals with darker skin tones. By reducing biases, it seeks to advance healthcare equity and accessibility, ultimately contributing to better healthcare outcomes for marginalized communities. By addressing data scarcity and enhancing dataset diversity, this method has the potential to improve the accuracy and generalizability of deep learning models for skin lesion classification, ultimately benefiting patient care and clinical decision making.

II. Literature Survey

Ian Goodfellow et al came with the paradigm of Generative Adversarial Networks that changed the generative modelling significantly in machine learning [4]. They proposed a framework that is essentially based on two neural networks as much similar as a generator and discriminator, playing a game of sorts that will make the learning of the generator to produce very realistic data points that can fool the discriminator while the discriminator learns to classify real data from fake data. GANs can therefore generate highly realistic samples without any explicit use of probability models, but also suffer with among others stability in training, mode collapse and high computational requirements.

Al-Rasheed et al have used conditional GANs, which incorporated a number of image augmentation techniques;

979-8-3315-2751-8/25/$31.00 ©2025 IEEE

flip, affine transformation, contrast adjustment, Gaussian blur and multiplication [5]. A cGAN framework is proposed to train the generator along with the discriminator within a single architecture. This produces conditioned data on certain inputs but the discriminator determines real and generated data and improves the capacity of the model to generate contextual output. Ensemble algorithms were used to combine the predictions of multiple models to increase the precision in classification and minimize biasing. Even though cGANs have allowed controlled generation of data and increase accuracy by ensemble methods comparable to traditional GANs, it has certain issues such as unstable training and mode collapse.

Researchers H. Rashid et al have been working on making use of GAN to generate realistic dermoscopic images in enhancing the training datasets for skin lesions classification [6]. The approach is based upon GANs generative capabilities toward producing various representational synthetic data outperforms classic data augmentation techniques. Studies were also able to show improvements in the evaluation metrics such as precision, recall, and F1-score on the classification of skin lesions. As other GAN-based models, this approach also has some challenges, for example unstable training and susceptibility to mode collapse among others and computationally expensive production of high-quality synthetic images.

Alsaidi et al approached the problem of class imbalance in dermoscopic image classification by comparing a conventional CNN-based classification model with a GAN-augmented dataset model [7]. The authors tried answering if data augmentation and GAN-generated synthetic data can enhance the classification accuracy without class imbalance of a medical image. The accuracy of underrepresented classes improved with GAN-augmented data, that eased some problems caused by class imbalance happening with many medical image datasets and allowed for augmentation of generalization. There are some disadvantages discovered in the study: the complexity of training GANs as well as potentially low-quality synthetic data that might negatively influence classification if not well tuned for the discriminator. It is computationally expensive, hence involving very high overhead compared to traditional CNN-based methods.

Qichen Su et al proposed Self-Transfer GAN, STGAN, to do data augmentation of skin lesion classification data overcoming the data and class imbalance problem [8]. STGAN is much more efficient in improving the classification performance by connecting real images with synthetic images produced through the GAN model, thus supporting robust learning from the classifiers. During recent times, it is one of the good ways of image production. Highly useful for medical image related applications. Its benefit towards performing class distribution balancing improves its overall performance and generalizability, but the process is computationally expensive to implement with careful tuning in order not to collapse during training and may not capture the full diversity of real-world data, with this resulting in negative impacts on the general efficiency.

E. Goren et al experimentally evaluated various architectures of CNNs, which include CNN, MobileNet, and ResNet-18 for the classification of skin lesions and feature extraction [9]. One paper does recommend the hybrid dataset, a combination of synthetic images synthesized by GAN models, and combined with original datasets, which did perform well with promising improvements in the accuracy of classifications. The research also compares unconditional and conditional GANs, highlighting the need to attack bias in data as well as in models. Hybrid datasets make for a better classification model and a good feature extraction approach with the use of the same, but bias embedded within the models might arise because GANs are learned to perpetuate or amplify the already existing biases. Lastly, synthetic data usage requires wide computational resources and careful tuning, thereby further complicating its implementation.

It was further proposed by Shubham Innani et al that one could leverage E-GAN, or the Efficient GAN and M-GAN, or the Mobile-GAN for data augmentation [10]. EGAN enables the employment of a discriminator based on a patch GAN which has efficiency as it can differentiate between synthetic and real labels. It optimizes without losing the quality of the data. MGAN is targeted to real-time usage, for example, in the dermatoscopy machines with a lightweight architecture designed for low computational resources. The research study shows a huge amount of in-depth knowledge of the domain-specific challenges. The model EGAN focuses on high-quality data generation whereas MGAN follows the policy of real-time efficiency in resource-constrained domains for deployment. Although both the models come along with new solutions, the benefits and shortcomings are quite prominent. With their high-efficiency and versatility, both the models being at par with other GAN models possess some limitations such as the possibility of training instability and fine-tuning to avoid mode collapse.

Bevan and Atapour-Abarghoei proposed an improvement in dataset annotation that would ultimately decrease biases in providing dermatological diagnoses especially if GAN-based augmentation techniques were to be employed [11]. The methodologies such as 'Learning Not To Learn' and 'Turning a Blind Eye' ,the author applied to the study then shows how that minimizes skin tone bias in the classification of skin lesions. This comes with the use of advanced GAN methodology leading to an excellent probability of results that are accurate and balanced for diagnosis. Though generalizing in a quest to balance dataset annotation bias with classification bias, significant scientific breakthroughs have been witnessed in the application of fair and accurate dermatological diagnosis from this research. Such an approach allows one to reduce bias on medical AI applications towards fairer and more inclusive models of diagnoses. This has come with further challenges-new ones-by asking the question of how such debiasing techniques might be implemented across the diversity of datasets and some potentially important features in data going unnoticed in the quest to eliminate bias.

Mikołajczyk et al emphasized the desirability of evaluation metrics to estimate the performance of GANs in beating the biases [12]. The experiments have been mainly on fidelity, diversity metrics, training speed, as well as the performance of classifiers on datasets containing a mix of real and synthetic data. Any form of training models on augmented datasets displayed gigantic improvements in performance. This approach helps in getting a better accuracy for the machine learning model and hence, well-balanced datasets and lesser biases. Increased complexity while evaluating takes place because GAN might produce synthetic data that contains unknown biases, in addition to that huge amounts of computational requirements take place during the training process.

The paper by Alankrita Aggarwal et al discusses how CNN can be utilized for image segmentation and how it goes a long way in the extraction of features and pixel-level classification, that significantly enhances the performance in applications like medical image analysis, traffic management, etc. [13]. This synthesis of realistic images is further refined with the integration of CNNs with GANs, thereby upgrading the methodologies of segmentations. Combining CNNs with GANs may capture advanced patterns and yield high quality output in image segmentation, but this comes with challenges of computational complexity, a possible instability of training, and the necessity of using a large dataset with annotations to achieve optimal performance.

A skin lesion classification study by Sara Atito Ali Ahmed et al, utilized the advanced CNNs; architectures Xception, InceptionResNetV2, and NasNetLarge within the framework of ISIC2019 designed to improve classifying correctness in detecting skin cancers [14]. The researchers had planned to put ensemble techniques such as LightGBM and one-class classifiers into action to handle class imbalance and anomaly detectors. In the words of the authors, the technique developed would, indeed, enhance the accuracy in diagnosis. While CNNs and ensemble methods provide several great advantages by boosting performance in medical imaging, the approach is still often limited by several issues: mainly complexity in training deep learning models, the necessity of large datasets, and susceptibility to overfitting where small or imbalanced data may be encountered.

Gouda et al highlights the use of deep learning techniques, in particular CNNs, to detect skin cancer at a very early stage [15]. With large datasets available, such as ISIC and others like this, the models achieve excellent classification performance on skin lesions and melanoma in particular. Innovations such as data augmentation and super-resolution GANs also improve output from such models by taking better input quality from such networks. Such developments make the diagnosis process better because they overcome weaknesses in what is essentially a dermatologist's assessment- lengthy and untrustworthy. Results from CNN-based models are excellent but with some limitations: demand huge labelled datasets, prone to biases through the training data, and overfitting while considering the variety of types of skin and lesions.

A study by Albahar et al uses a deep CNN model with a novel regularizer, aimed at classifying skin lesions as benign or malignant with a high average accuracy, surpassing several existing methods [16]. Tested on multiple lesion types, the model achieved notable AUC scores (up to 0.93 for certain comparisons), demonstrating potential utility in aiding dermatologists. However, CNN models for skin cancer diagnosis face limitations, including dependence on high-quality annotated datasets, potential overfitting in heterogeneous data, and computational demands, which can hinder accessibility and real-world deployment in diverse medical settings.

A paper by T. Alkarakatly et al presents a CNN-based approach for classifying skin lesions into melanoma, atypical nevus, and common nevus, using the PH2 dataset with 200 images [17]. It explores preprocessing methods, including resizing and pixel normalization, and tests various configurations of CNN layers and filter sizes to optimize classification performance. Notably, the study uses weight reinitialization and dropout to address overfitting, achieving an accuracy of up to 93%. However, the model struggles with limited data, which can hinder generalizability and robustness. Unlike prior studies, this approach tests CNNs with smaller dataset variations, providing insights into CNN optimization under data constraints.

Elgamal et al. explored a hybrid technique for skin cancer classification by integrating Discrete Wavelet Transformation (DWT) for feature extraction, Principal Component Analysis (PCA) for dimensionality reduction, and classification through feedforward-backpropagation neural networks (FP-ANN) and k-nearest neighbors (k-NN) classifiers [18]. This study aimed to determine if combining DWT and PCA could improve classification accuracy, addressing the complexities often encountered in diagnosing diverse skin conditions. Results demonstrated that this hybrid model achieved significant diagnostic performance with good classification accuracies. However, challenges were noted in computational efficiency, as the need for detailed tuning during feature extraction increased processing demands. Such limitations could affect generalization across diverse datasets, necessitating further optimization to handle skin texture variability effectively. These findings underscore the model's potential while highlighting the importance of balanced computational and diagnostic accuracy considerations.

In another study, Bhosale conducted a comparative analysis of machine learning and deep learning frameworks for skin lesion classification by employing MobileNetV2 with TensorFlow and PyTorch, alongside Random Forest algorithms. This research aimed to evaluate model performance on the HAM10000 dataset, specifically to determine if MobileNetV2's compact architecture could support effective classification within resource-limited mobile environments [19]. The findings revealed that the TensorFlow-based Mobile- NetV2 demonstrated the highest accuracy and excelled in detecting malignant melanoma with strong recall. While promising, the study also emphasized the considerable computational resources required to maintain model performance and noted the importance of effective dataset management to ensure model reliability across different environments. This work underscores the potential for deep learning models to make dermatological diagnostics more accessible in underserved regions.

A study by Maron RC et al compares the diagnostic accuracy of dermatologists and a convolutional neural network (CNN) model in detecting skin cancer types, including melanoma, squamous cell carcinoma, and basal cell carcinoma [20]. The CNN was trained on the ISIC dataset, focusing on images from the HAM10000 Dataset, which included biopsy-verified lesions. Dermatologists from German university hospitals assessed 300 test images via questionnaires, which recorded lesion diagnosis, certainty level, and image quality. Results revealed that dermatologists had a mean sensitivity and specificity of 56.5%, while the CNN, based on ResNet50, achieved comparable accuracy with higher consistency across cases, especially in differentiating benign and malignant lesions, underlining the need for a machine learning system to serve as helping hands.

Despite advancements in machine learning and generative models like GANs, a notable gap persists in the classification of skin lesions on darker skin tones. A detailed analysis of studies reveals that dermatological models, particularly those trained on imbalanced datasets, often fail to generalize well across diverse skin tones, leading to biases in diagnostic outcomes.

III. Methodology

The proposed system integrates both GAN-based data augmentation and CNN-based classification models to improve dermatological diagnosis across diverse skin tones. The system includes the below processes.

A. Dataset Sourcing

The ISIC 2019 dataset is sourced as the primary dataset for this work. The ISIC 2019 dataset consists of 25,331 dermoscopic images of skin lesions from various sources including researchers, medical professionals, and machine learning enthusiasts, hospitals, clinics, and private practice settings. The images were acquired using a variety of dermoscopic devices and represented a diverse range of skin lesion types, including melanoma, nevus (moles), seborrheic keratosis, and other benign and malignant conditions. This makes the ISIC 2019 dataset an ideal choice to work with, ensuring the representation of skin lesions on individuals from various ethnicities.

B. Image Resizing

The acquired images are resized to a standardized dimension of 224x224 pixels. This resizing step ensures uniformmity in image dimensions across the dataset, facilitating compatibility with subsequent processing stages and machine learning models.

C. Image Darkening and Hair Removal

To introduce variability and enhance dataset diversity, images are subjected to a darkening process. With a specified probability of 0.25 and darkening factor of 0.5, a subset of images undergoes darkening, mimicking variations in lighting conditions commonly encountered in clinical settings. Subsequently, a hair removal algorithm is applied to the images to eliminate artifacts and distractions. Hair removal enhances the clarity and focus of the images, ensuring that the diagnostic features of the skin lesions remain prominent and unobstructed.

D. GAN Augmentation

Following preprocessing steps, the processed images are fed into Generative Adversarial Networks (GANs) for data augmentation. GANs, comprising a generator and discriminator network, engage in an adversarial training process to generate synthetic images resembling real dermatological cases. By leveraging the GAN framework, the system aims to further enrich the dataset with diverse and realistic synthetic images.

As shown in Fig. 1., the architecture of the Spectral Norm- alization GAN (SN-GAN) used in this work consists of two key components: the Generator and the Discriminator. The Generator receives a noise vector sampled from a Gaussian distribution as input and passes it through several layers to synthesize high-quality dermatological images. It starts with a Dense (fully connected) layer, which reshapes the noise into an initial feature map, followed by transposed convolutional layers that progressively upscale the image to the target resolution of 224×224. These layers are accompanied by batch normalization to stabilize the learning process and ReLU activation functions to introduce non-linearity. The final layer uses the tanh activation to scale the pixel values between -1 and 1.

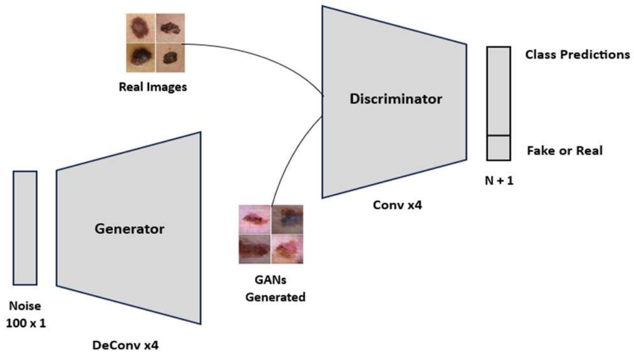

Fig. 1. Architecture of Generative Adversarial Network (GAN) consisting of generator and discriminator.

The Discriminator, which aims to distinguish real from synthetic images, takes an image as input and processes it through a series of convolutional layers that downsample the image while extracting meaningful features. To prevent overfitting and enhance stability during training, Spectral Normalization is applied to each layer, ensuring controlled gradients. LeakyReLU activation is employed after each layer to avoid dying neurons. The network also includes dropout layers to improve generalization. The output of the Discriminator is a single scalar, produced by a sigmoid activation, representing the probability that the input image is real.

For the loss functions, the GAN employs binary cross-entropy loss for both the Generator and Discriminator. The Generator's loss is designed to minimize the ability of the Discriminator to differentiate between real and generated images, while the Discriminator's loss aims to maximize its ability to distinguish real images from synthetic ones. The architecture and spectral normalization together help maintain balanced training between the two networks, reducing the risk of mode collapse and ensuring high-quality image generation.

E. Dataset Expansion

The synthetic images generated by the GAN are included in the dataset. This augmented dataset provides a more comprehensive representation across different skin tones, helping to reduce bias in training.

F. CNN Model Training

Convolutional Neural Networks (CNNs) are trained on the expanded dataset. In this experiment, CNN architectures such as ResNet50, InceptionResNetV2, and NASNetLarge are employed. These models are well-suited for image classification tasks, with the ability to extract deep features from images.

1) ResNet50: It is a deep learning model that employs residual learning, enabling the effective training of networks with considerable depth. The architecture incorporates skip connections, which are instrumental in mitigating the vanishing gradient problem. This allows for deeper networks with improved feature extraction capabilities, making it highly effective for skin lesion classification. The input tensor dimensions for ResNet50 are specified as (224,224,3).

2) InceptionResNetV2: This model combines the strengths of Inception modules and residual connections, offering both depth and efficiency in feature extraction. It is particularly adept at capturing fine-grained details in images, essential for accurate dermatological diagnoses. The input tensor dimensions are established at (299,299,3).

3) NASNetLarge: It is a convolutional neural network architecture developed using Neural Architecture Search (NAS) techniques. This architecture integrates depth wise separable convolutions and feature normalization, which collectively enhance performance while maintaining computational efficiency. The input tensor is defined with dimensions (H, W, C), where H and W represent the height and width (specifically 331x331 pixels), and C denotes the number of channels (3 for RGB images). The architecture commences with a stem block composed of several convolutional and pooling operations that extract low-level features from the input data.

G. Iterative Improvement

The performance of both the GAN and CNN models are evaluated and iterative improvements are made. This could involve fine-tuning hyperparameters, adjusting augmentation techniques, or optimizing the network architecture.

H. Quality Assurance

Throughout the process, quality assurance measures are implemented to validate the integrity and fidelity of the processed images. This involves manual inspection of a subset of images to ensure that preprocessing and augmentation techniques preserves the diagnostic features and authenticity of the dermatological cases.

I. Optimization Strategies for High Computational Demands

While GAN training and CNN ensemble models yield high accuracy and robustness, they also pose significant computational challenges. This study employs several strategies to mitigate these demands.

1) Efficient Architectures: Lightweight models like NASNetLarge and optimization-focused GANs such as SN-GANs reduce computational overhead without compromising performance.

2) Distributed Training: Multi-GPU setups and distributed frameworks are utilized to parallelize computations, reducing training time.

3) Hyperparameter Tuning: Automated and manual optimization of learning rates, batch sizes, and regularization parameters ensures faster convergence.

4) Gradient Checkpointing: Memory efficient techniques are employed to recompute activations only when necessary, reducing GPU memory usage.

5) Model Quantization: The CNN ensemble employs quantized models for inference, reducing precision and memory requirements.

IV. IMPLEMENTATION

The Fig. 2. outlines the systematic approach followed beginning with project planning and research, which establishes the foundation for subsequent steps. Following this, the data collection and preprocessing stage ensures that high quality data is gathered and prepared for analysis. This leads to the GAN model development for augmentation, where Generative Adversarial Networks (GANs) are created to enhance the dataset by generating synthetic images, particularly targeting the underrepresentation of darker skin tones in dermatological datasets. Once the GAN model is developed, it moves into the GAN model training and optimization phase, where the model is refined for better performance in generating useful images. This training phase is crucial for producing high-quality synthetic data that can enhance the overall dataset.

Simultaneously, there's a focus on the preparation of the dataset for CNN classification, ensuring that the data is well-structured for the next phases. Convolutional Neural Networks (CNNs) are trained using both real and synthetic images. The CNN model development involves building the CNN architecture tailored for the task. After these models are optimized, model evaluation and testing are done, ensuring that the performance meets the required standards before proceeding to deployment and integration.

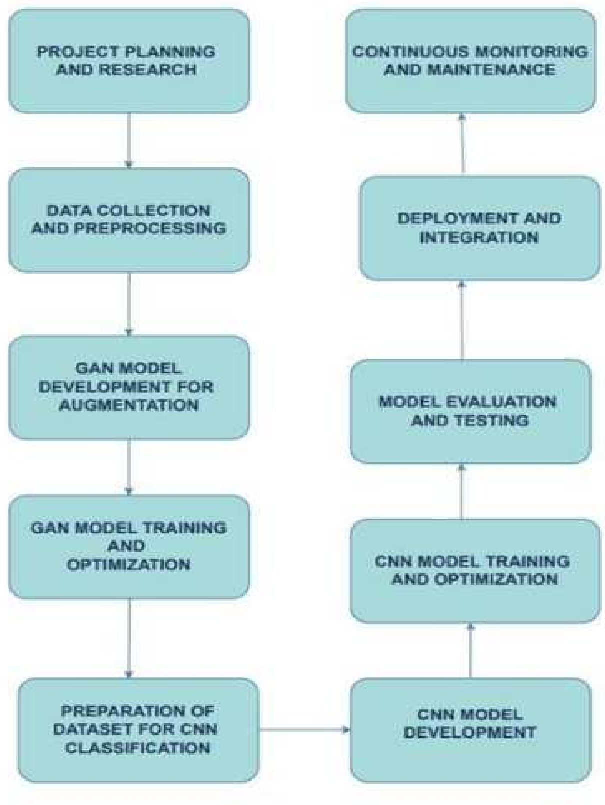

Fig. 2. Data Flow Diagram

The experiment utilized an advanced version of GAN called Spectral Normalized Generative Adversarial Networks (SN-GAN). Spectral Normalized Generative Adversarial Networks (SN-GANs) enhance the traditional GAN framework by incorporating spectral normalization, which stabilizes training and improves the quality of generated samples. As in standard GANs, SN-GANs consist of two neural networks—the generator and the discriminator—trained adversarial to learn the underlying distribution of the training data and generate samples that closely resemble real data.

A. Generator

The generator consists of a series of transposed convolutional layers, each followed by batch normalization layers and ReLU activation functions. The sequence begins with a latent vector of random noise, which is passed as a feature map to subsequent layers. Each transposed convolutional layer generates an output with varying channel counts. Batch normalization normalizes the output across the batch dimension, stabilizing the training process. The ReLU activation function is defined in (1)

$$f(x) = max\ (0, x) \tag{1}$$

Where x is the input value to the activation function. The ReLU function applies the following rule:

- If the input value x is positive, it remains unchanged (x is returned).
- If the input value x is negative, it is set to zero.

Unlike sigmoid or tanh activation functions, ReLU does not saturate for large positive values, which can help address the vanishing gradient problem during training. The final layer in the generator utilizes the tanh activation function, squashing the pixel values between -1 and 1, enabling the generator to produce output images that align with the data distribution.

B. Discriminator

The discriminator functions as a binary classifier that distinguishes between real and generated samples. Its architecture includes convolutional layers (denoted as Conv2d), spectral normalization, batch normalization, and Leaky ReLU activation functions. The architecture starts with a convolutional layer, immediately followed by a Leaky ReLU layer. This is succeeded by a series of convolutional, batch normalization, and Leaky ReLU layers. Leaky ReLU (Leaky Rectified Linear Unit) is a variant of the standard ReLU activation function. It is defined in (2)

$$f(x) = max\ (\alpha x, x) \tag{2}$$

Where α is a small positive constant, typically in the range of 0.01 to 0.3. The Leaky ReLU function works as follows:

- If the input value x is positive, it remains unchanged (x is returned), just like the standard ReLU.
- If the input value x is negative, instead of outputting zero, it outputs a small, non-zero value equal to αx.

The sigmoid layer is the final activation function that applies the sigmoid function to the output of the previous layer, squashing the values between 0 and 1. This output represents the probability of the input image being real or fake, as perceived by the discriminator.

C. Adam optimizer

In Generative Adversarial Networks (GANs), the optimizer plays a crucial role in training both the generator and discriminator networks. The optimizer is responsible for updating the weights of the networks in order to minimize their respective loss functions during training. Both the generator and discriminator use the Adam optimizer, which is a popular optimization algorithm for training deep learning models, including GANs. The Adam optimizer has several hyperparameters:

1)Learning rate: It determines the step size in the direction of the negative gradient during optimization. A smaller learning rate (e.g., 0.0002) is often used for GANs to ensure stable training. In the experiment the value of learning rate used is lr=0.0002.

2)Exponential decay rates : These are the exponential decay rates for the momentum and the squared gradient, respectively. The default values are used in the experiment (betas=(0.5, 0.999)).

D. Loss function

The generator's goal is to generate fake images that are indistinguishable from real images. Its loss function is based on the discriminator's predictions for the generated images. The generator aims to maximize the discriminator's prediction error for fake images, effectively minimizing its binary cross-entropy loss.The discriminator's goal is to distinguish between real and fake images. Its loss function is a typical binary cross-entropy loss, which measures the difference between its predictions and the true labels (real or fake). The discriminator aims to minimize this loss.Both the generator and discriminator use the Binary Cross-Entropy (BCE) loss function for both the generator and discriminator. The BCE loss is a common choice for binary classification tasks, where the output is a probability between 0 and 1.

E. Ensemble Model

Convolutional Neural Networks (CNNs) are trained on the expanded dataset, utilizing architectures such as ResNet50, InceptionResNetV2, and NASNetLarge. The training focuses on optimizing hyperparameters and ensuring robust performance across diverse skin tones. The final ensemble model combines the predictions generated by NASNetLarge, ResNet50, and InceptionResNetV2, averaging their outputs to improve classification accuracy and enhance robustness against overfitting.

V. RESULTS

The results focus on the augmentation of images using Generative Adversarial Networks (GANs), hair removal from dermatological images, and classification through a CNN ensemble. Hair removal is conducted using specialized image processing techniques to improve feature visibility, critical for accurate analysis. The effectiveness of this process is assessed both visually and quantitatively, with Fig.3. illustrating sample images before and after hair removal, showing improved clarity and feature detail.

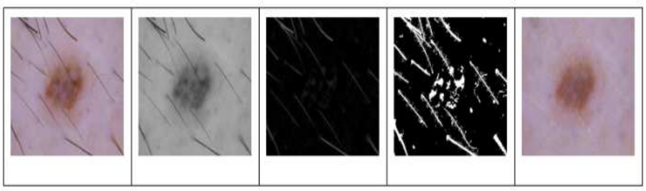

Fig. 3. The stages involved in hair removal using the Blackhat method

The dataset is further expanded by generating synthetic images using GANs, trained specifically for dermatological images. The quality and diversity of these GAN-generated images are compared to the original dataset, as shown in Fig. 4., to ensure they provided valuable variability.

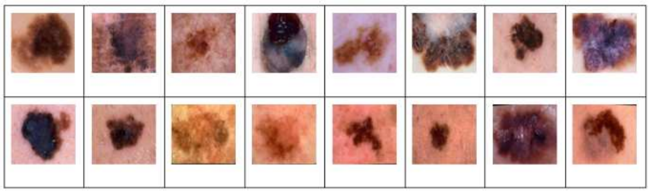

Fig. 4. Images from the dataset

Fig. 5. showcases examples of GAN-generated images alongside corresponding dataset images. The generated images are added into the dataset to improve the variations in representation of skin ethnicities. The classification is carried out through a CNN ensemble, integrating ResNet50, NASNetLarge, and InceptionRes- NetV2 architectures. This ensemble approach leverages the individual strengths of each model to improve classification accuracy and robustness. The overall impact of CNNs, GAN augmentation, and preprocessing steps effectively enhanced the accuracy of skin lesion classification, contributing to better model performance and generalizability.

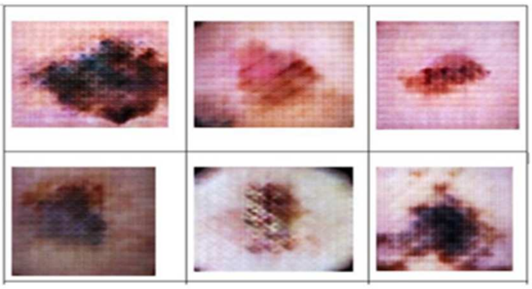

Fig. 5. Images generated by the Generative Adversarial Network.

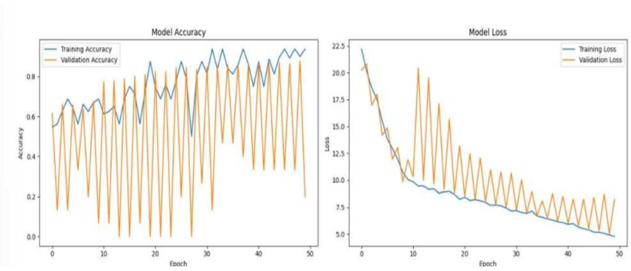

Fig. 6. Training history pre-augmentation.

As seen in Fig. 6 and Fig. 7., the initial model demonstrates moderate performance with notable class imbalance issues. The confusion matrix reveals strong prediction for the NV class (1276 correct), but significant misclassifications among other classes, particularly for SCC. The model accuracy plots show high volatility in validation accuracy, ranging from 20-80%, while training accuracy reaches approximately 90%. The loss curves indicate potential overfitting, with a consistent gap between training and validation loss. These results suggest the need for improved generalization and balanced class representation.

Fig. 7. Confusion matrix pre-augmentation

As seen in Fig. 8. and Fig. 9., following augmentation, the model exhibits substantial improvements across all metrics. The confusion matrix shows enhanced performance for all classes, with BCC predictions dramatically increasing from 91 to 6522 correct classifications. Overall class balance improved significantly. The model accuracy plot demonstrates higher and more stable validation accuracy, peaking around 90%, with training accuracy approaching 100%. Loss curves for both training and validation sets decreases more consistently, with training loss nearing zero. These outcomes indicate that the augmentation techniques effectively addresses class imbalance, reduced overfitting, and improved the model's generalization capabilities across various skin lesion types.

Fig. 8. Confusion matrix post-augmentation.

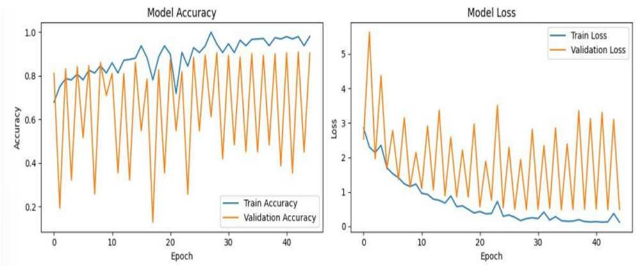

Fig. 9. Training history post-augmentation.

VI. CONCLUSION AND FUTURE SCOPE

This study makes the effective utilization of Generative Adversarial Networks (GANs) for data augmentation and hair removal techniques, alongside Convolutional Neural Networks (CNNs) for classification, to improve the quality and diversity of dermatological datasets. The experiments demonstrate the effectiveness of GAN-based augmentation in generating synthetic images that closely resemble real dermatological lesions, addressing data insufficiency and enhancing dataset representation. Moreover, the successful removal of hair from dermatological images results in clearer, more visually informative images, facilitating more accurate diagnosis and analysis. The integration of CNNs further improve the classification accuracy by leveraging the enhanced datasets, showcasing their capability in capturing intricate features essential for skin lesion classification. These findings underscore the importance of both preprocessing techniques and advanced classification models in enhancing the quality of dermatological datasets and improving the efficacy of computer-aided diagnosis systems. By mitigating biases arising from data insufficiency and improving dataset diversity, this combined approach with GAN-based augmentation and CNN-based classification has the potential to significantly enhance the accuracy and fairness of dermatological diagnoses, ultimately benefiting healthcare outcomes for individuals across all skin tones.

The future work may extend further efficiency with GAN-augmented data in other fields of medical images, find architectures of GANs specifically for classification tasks, and incorporate domain-specific preprocessing techniques to better improve results. The key issues are imbalance and fair representation of underrepresented demographics; GANs can be designed to produce extra samples for such classes, thus promoting more accurate and equitable diagnosis. Further, these models can be deployed on mobile or cloud platforms that would give the basis for real dermatological assessments in real time, most importantly in remote areas, thus improving accessibility. Regulation in Health Compliance and pre-processing techniques such as hair removal or shaving may further an ethically sound and reliable AI-powered healthcare solution.

REFERENCES

[1] Joseph, Supriya, and Janu R. Panicker, "Skin lesion analysis system for melanoma detection with an effective hair segmentation method", 2016 International Conference on Information Science (ICIS), IEEE, 2016.

[2] ur Rehman, Mobeen, Sharzil Haris Khan, SM Danish Rizvi, Zeeshan Abbas, and Adil Zafar, "Classification of skin lesion by interference of segmentation and convolotion neural network", In 2018 2nd International Conference on Engineering Innovation (ICEI), pp. 81-85. IEEE, 2018.

[3] Qin, Zhiwei, Zhao Liu, Ping Zhu, and Yongbo Xue, "A GAN-based image synthesis method for skin lesion classification", Computer Methods and Programs in Biomedicine 195 (2020): 105568.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David WardeFarley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, “Generative adversarial nets”, Advances in neural information processing systems 27, 2014.

[5] Al-Rasheed, Ksibi, Ayadi, Alzahrani and Mamun Elahi, “An Ensemble of Transfer Learning Models for the Prediction of Skin Lesions with Conditional Generative Adversarial Networks”, Contrast Media & Molecular Imaging, pp. 1-15, 2023.

[6] H. Rashid, M. A. Tanveer and H. Aqeel Khan, “Skin Lesion Classification Using GAN based Data Augmentation”, 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, pp. 916-919, 2019.

[7] Alsaidi, Mostapha & Jan, Muhammad & Altaher, Ahmed & Zhuang, Hanqi & Zhu, Xingquan, “Tackling the class imbalanced dermoscopic image classification using data augmentation and GAN”, Multimed Tools Appl 83, pp. 49121–49147, 2024.

[8] Qichen Su, Haza Nuzly Abdull Hamed, Mohd Adham Isa, Xue Hao, and Xin Dai. Q. Su, H. N. A. Hamed, M. A. Isa, X. Hao and X. Dai, “A GAN-Based Data Augmentation Method for Imbalanced Multi-Class Skin Lesion Classification”, IEEE Access, vol. 12, pp. 16498-16513, 2024.

[9] E. Gören And G. Çınarer, “Cancer Lesion Classification with GAN-Based Image Augmentation Method from Skin Images”, 2nd International Conference on Engineering, Natural and Social Sciences, vol.1, no.1, Konya, Turkey, pp.658-666, 2023.

[10] Shubham Innani, Prasad Dutande, Ujjwal Baid, Venu Pokuri, Spyridon Bakas, Sanjay Talbar, Bhakti Baheti and Sharath Chandra Guntuku, “Generative adversarial networks-based skin lesion segmentation”, Sci Rep 13, 13467, 2023.

[11] Bevan, P.J. and Atapour-Abarghouei, A, “Detecting Melanoma Fairly: Skin Tone Detection and Debiasing for Skin Lesion Classification”, Lecture Notes in Computer Science, vol 13542. Springer, Cham, 2022.

[12] Mikołajczyk, A, Majchrowska, S. Carrasco Limeros, S, “The (de)biasing Effect of GAN-Based Augmentation Methods on Skin Lesion Images”, Medical Image Computing and Computer Assisted Intervention – MICCAI, vol 13438, Springer, Cham, 2022.

[13] Alankrita Aggarwal, Mamta Mittal and Gopi Battineni, "Generative adversarial network: An overview of theory and applications", International Journal of Information Management Data Insights, Volume 1, Issue 1,2021.

[14] Sara Atito Ali Ahmed, Berrin Yanıkoglu, Özgü Göksu, and Erchan Aptoula, "Skin Lesion Classification With Deep CNN Ensembles", 2 2020 28th Signal Processing and Communications Applications Conference (SIU), October 2020, pp. 1-5. DOI: 10.1109/SIU49456.2020.9302125.

[15] Gouda W, Sama NU, Al-Waakid G, Humayun M and Jhanjhi NZ, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning", Healthcare, vol. 10, no. 11, 2022, p. 1183.

[16] Albahar, Marwan, “Skin Lesion Classification Using Convolutional Neural Network With Novel Regularizer”, IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2906241, 2019.

[17] T. Alkarakatly, S. Eidhah, M. Al-Sarawani, A. Al-Sobhi and M. Bilal, "Skin Lesions Identification Using Deep Convolutional Neural Network," 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), Al Madinah Al Munawwarah, Saudi Arabia, 2020.

[18] M. Elgamal, "Automatic skin cancer images classification", International Journal of Advanced Computer Science and Applications, vol. 4, no. 3, 2013.

[19] S. Bhosale, "Comparison of deep learning and machine learning models and frameworks for skin lesion classification," arXiv:2207.12715, Jul. 2022.

[20] Maron RC, Weichenthal M, Utikal JS, Hekler A, Berking C, Hauschild A, Enk AH, Haferkamp S, Klode J, Schadendorf D, Jansen P, Holland-Letz T, Schilling B, von Kalle C, Fröhling S, Gaiser MR, Hartmann D, Gesierich A, Kähler KC, Wehkamp U, Karoglan A, Bär C, Brinker TJ, “Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks”, Eur J Cancer. 2019 Sep; 119:57-65, 2019.