DermaGAN: Enhancing Skin Lesion Classification with Generative Adversarial Networks

Sri Sai Teja M S
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
srisaitejms2002@gmail.com

Ullas P
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
ullasputtaraju@gmail.com

Vikas K R
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
vikaskr8055@gmail.com

R C Ravindranath
Assistant Professor, Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India

***Abstract*—In the domain of dermatology diagnosis, the convergence of cutting-edge image processing techniques and deep learning methodologies stands as a pivotal frontier. This paper introduces "DermaGAN," a pioneering methodology harnessing Generative Adversarial Networks (GANs) to synthesize a comprehensive array of dermatology images, covering four essential classes—melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma. The synthesized dataset addresses the critical need for diverse and expansive training data, mitigating the inherent constraints of data scarcity in medical image analysis. Convolutional Neural Network (CNN) models are then used to classify skin lesions, demonstrating via rigorous comparative analysis the extraordinary effectiveness of DermaGAN-generated images in improving classification performance across a range of skin cancer types. Interestingly, our customized CNN architecture has performed better than pre-trained CNN models like VGG16, ResNet50, VGG19, MobileNet, and MobileNetV2. This customized CNN uses GAN-generated images for the categorization of skin lesion classifications. Our results highlight the critical role that synthetic picture augmentation plays in strengthening CNN models' diagnostic skills, especially in situations where annotated data is sparse. DermaGAN represents a transformative paradigm in dermatology diagnosis, promising to revolutionize patient care by enabling earlier and more precise detection of skin cancers, thereby facilitating timely interventions and improved patient outcomes while opening novel avenues in dermatological research and clinical practice.**

Keywords*—Generative Adversarial Network, Convolutional Neural Network, Image Augmentation.

I. INTRODUCTION

Dermatological disorders, particularly skin cancers, represent a significant burden on healthcare systems globally due to their widespread prevalence and potential for severe morbidity and mortality. Among these conditions, melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma emerge as the most prevalent and clinically significant. Timely and accurate diagnosis of these diseases is important for correct treatment planning and good patient outcomes. However, dermatology diagnostic procedures often rely on observations, which can rely on doctors' opinions and lead to bias. In recent years, the intersection of artificial intelligence (AI) and deep learning technologies with advances in medical imaging has ushered in a new era of innovation in diagnostic dermatology.

"DermaGAN" is a GAN application designed for dermatology diagnosis. DermaGAN combines digital imaging and dermatology, providing rich and diverse training for deep learning models. There are four main types of these synthetic data: actinic keratosis, basal cell carcinoma, benign keratosis, and melanoma, which represent skin diseases. Potential in convolution performance. CNNs have been highly successful in many image classification applications, including dermatology. We aim to improve the accuracy and robustness of skin cancer classification by training CNN on the augmented dataset generated by DermaGAN. A comparison of different CNN architectures was also made to evaluate their effectiveness in classifying different types of cancer.

This article aims to investigate the potential of DermaGAN-generated images to improve the performance of convolutional neural network (CNN) models in skin classification. CNNs have been highly successful in many image classification applications, including dermatology. We aim to improve the accuracy and robustness of skin cancer classification by training CNN on the augmented dataset generated by DermaGAN. Additionally, we conducted a comparison of various CNN architectures to evaluate their effectiveness in classifying different types of cancer. The synergy of dermatology diagnoses increases the field. The ultimate goal is to create a broad and effective capacity to improve the efficiency and accuracy of skin cancer diagnosis, thus improving the outcome of individual pain and reducing the pain of treatment.

With this research project, we hope to improve dermatology diagnosis by utilizing the complementary strengths of deep

learning methodologies and cutting-edge picture augmentation techniques. The ultimate goal is to create an efficient and scalable framework that will improve patient outcomes and lessen the burden on healthcare systems by increasing the efficiency and accuracy of skin cancer diagnosis. This work intends to advance dermatology by offering a thorough evaluation of CNN models for skin cancer classification.

Clinical practice could be completely transformed by combining dermatological diagnostics with AI-driven technology. Our goal is to develop a strong and dependable system for the accurate and timely diagnosis of skin cancer lesions by utilizing the strengths of DermaGAN and CNNs. We want to develop a framework that can generate consistently correct results in a variety of clinical contexts by optimizing the blend of deep learning and picture enhancement approaches.

Additionally, this study aims to tackle the current obstacles and constraints in dermatological diagnosis, such as a lack of available data, arbitrary methods of evaluation, and variations in diagnostic precision. We hope to provide light on the possible uses and constraints of CNN models and DermaGAN-generated images in clinical settings by presenting a methodical assessment of both technologies. We work together with researchers, physicians, and industry partners to convert these developments into real advantages for both patients and healthcare providers.

To sum up, this study is a big step toward using deep learning and artificial intelligence to diagnose dermatological conditions. Our goal is to enhance the precision, effectiveness, and ease of use of skin cancer diagnosis through the use of cutting-edge technologies like DermaGAN and CNNs. This will ultimately result in better patient outcomes and lower healthcare expenses. We foresee a future when AI-driven dermatology diagnostics play a major role in the early diagnosis and management of skin malignancies, contributing to general gains in public health and well-being. This will be achieved through ongoing research and development activities.

A. Motivation

Skin cancer is one of the most prevalent forms of cancer globally, affecting millions annually. Over 100,000 new cases of melanoma, the deadliest type of skin cancer, are detected each year in the United States alone, out of an estimated 5.4 million cases of basal and squamous cell skin cancer. Nearly 20 Americans die from melanoma daily. Factors like excessive UV exposure, genetic predispositions, and environmental elements contribute to the high incidence of skin cancer. Early detection is crucial, as melanoma's survival rate dramatically increases to 99% when diagnosed early, compared to just 27% for advanced stages.

This project's motivation stems from the urgent need to enhance early detection and diagnosis of skin cancers, which can significantly improve patient outcomes and reduce mortality rates. Traditional diagnostic methods, reliant on visual inspections and biopsies, are time-consuming and subjective.Our goal is to create an automatic and precise skin lesion categorization system by utilizing cutting-edge deep learning techniques, particularly Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs). This system aids dermatologists in making quicker and more precise diagnoses and addresses data scarcity in medical image analysis by generating synthetic images, thus improving model performance and reliability. The potential for this project to greatly improve patient care and public health by offering a trustworthy tool for early skin cancer diagnosis is what motivated us to take it on.

B. Contribution

In order to solve the crucial problem of data scarcity in medical image analysis, this research presents "DermaGAN," a unique approach that uses Generative Adversarial Networks (GANs) to produce synthetic images for skin lesion categorization. Our primary contribution is the development of an enhanced dataset comprising both real and GAN-generated images, ensuring a balanced representation across four skin lesion categories: melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma.We show that, in terms of accuracy and loss measures, our customized Convolutional Neural Network (CNN) architecture, trained on this enhanced dataset, performs much better than pre-trained models like VGG16, ResNet50, and MobileNet versions. Additionally, we conduct a thorough dropout analysis to optimize model performance and mitigate overfitting. Our work aids physicians in the early and precise detection of skin malignancies by offering a reliable and accurate categorization system, thereby improving patient outcomes and furthering the field of dermatological research.

C. Paper outline

The structure of this document is as follows: Section 2 examines relevant studies and identifies gaps in the body of knowledge. Section 3 details the methodology, including data collection, preprocessing, GAN implementation, and CNN model development. Section 4 presents the experiments and results, including performance comparisons and dropout analysis. Section 5 outlines potential future improvements, suggesting ways to enhance the model's robustness and performance further. Section 6 wraps up by summarizing contributions and outlining potential future research topics.

II. EXISTING WORK

This part reviews previous research in the topic, with an emphasis on pertinent studies and developments that serve as the framework for our investigation. The approach used in [1] entails creating a framework for classifying skin lesions in images by utilizing the DenseNet201 architecture and a skin lesion augmentation style-based GAN (SLA-StyleGAN).This framework improves balanced multiclass accuracy (BMA) by restructuring the generator structure and adding a new loss function, resulting in the creation of high-quality skin lesion images. Improved classification accuracy, help with correct diagnosis of different types of skin lesions, and study of lesions at different stages and those that are hard to differentiate are some of the benefits of this technique. The necessity for additional validation against current techniques and the suitability of the suggested framework in actual clinical settings are possible drawbacks, though.

[2]suggests using a Deep Convolutional Generative Adversarial Network (DC-GAN) for image classification in order to diagnose skin cancer. Its methodology integrates GANs with deep convolutional neural networks. The approach offers automated skin cancer diagnosis, leveraging

advanced image classification techniques. However, potential drawbacks may include challenges in accurately classifying complex lesions and the need for validation across diverse datasets for robustness.

Tsai et. al. [3] introduce a novel approach to analyze facial pigmentation using Conditional Generative Adversarial Networks (CGANs). Their method enables precise evaluation of pigmented areas on the face, offering valuable insights for dermatological diagnosis and treatment planning. However, potential drawbacks include the requirement for extensive and diverse datasets to effectively train and validate the model. Zhang et. al. [4] provide a segmentation strategy that uses dual-stream patch-based discriminators in conjunction with Superpixel Guided Generative Adversarial Networks (GANs) to identify skin lesions. By increasing the precision of skin lesion identification, this method facilitates better dermatological condition diagnosis and treatment planning. Nevertheless, challenges may arise in optimizing the GAN architecture to perform optimally across various skin lesion types and image qualities.

Nugroho et. al. [5] conducted a systematic review on deep learning-based methods for classifying dermoscopy skin lesions from images, highlighting their potential in dermatological diagnostics. However, challenges may arise in selecting appropriate deep learning architectures and optimizing model parameters to achieve robust and accurate classification performance. Metta et. al. [6] presented interpretable AI methods to improve the classification of skin lesions, providing important information about how AI models make decisions. Nevertheless, complexities may arise in developing interpretable AI algorithms that balance transparency with classification accuracy. Bansal and Sridhar [7] proposed HEXA-GAN, a novel approach for inpainting skin lesion images using hexagonal sampling-based generative adversarial networks. However, limitations may include the need for further refinement to address challenges in preserving fine details and textures during image inpainting. Farady et al.[8] introduced PSIG-Net, a pseudo skin image generator intended to control outliers and generate samples free of ambiguity in tasks involving the categorization of skin lesions. Despite its potential benefits, challenges may arise in fine-tuning the generator architecture and optimizing training procedures to ensure consistent and reliable performance across diverse datasets.

Su et. al. [9]suggested a GAN-based data augmentation method to improve the classification performance of deep learning models by resolving class imbalance in multi-class skin lesion classification. However, challenges may arise in optimizing the GAN architecture and training process to effectively generate diverse and realistic synthetic data samples across all classes.

Nirmala and Premaladha [10] presented a GAN-based approach for augmenting medical images, focusing on melanoma skin lesion classification with deep learning models. Despite its potential benefits, limitations may include the need for further validation and optimization to ensure the reliability and generalizability of the augmented images for classification tasks.

Mutepfe et. al. [11] presented a GAN image synthesis technique to create artificial skin lesion images with the goal of enhancing the efficiency of classification systems. Still, there might be difficulties in adjusting the GAN's parameters and streamlining the training procedure to generate high-quality artificial images that closely mimic genuine skin lesions.

In order to improve the discriminative qualities of the input photos, Teodoro et al. [12] suggested a skin cancer classification method utilizing GANs and a region-of-interest (ROI)-based attention mechanism. The intricacy of combining the attention mechanism with the GAN architecture and enhancing the classification performance overall, however, can be a drawback.

Bissoto et. al. [13] presented a GAN-based method for synthesizing skin lesion images, aiming to create realistic and diverse datasets for training classification models. Despite its potential advantages, challenges may include the need for rigorous evaluation and validation to ensure the authenticity and utility of the synthetic images for dermatological diagnostics and research.

Xiang and Wang[14] investigated the use of deep learning methods to create interpretable skin lesion classification models with the goal of improving the clarity and interpretability of classification results. However, challenges may arise in balancing model complexity with interpretability and ensuring the reliability of the classification outcomes in clinical settings.

Bissoto et al.[15] analyzed and emphasized the possible advantages and drawbacks of GAN-based data augmentation and anonymization approaches for skin lesion analysis. Nevertheless, challenges may include the need for standardized evaluation metrics and protocols to assess the effectiveness and privacy implications of GAN-based augmentation methods.

The goal of Rashid et al.'s [16] GAN-based data augmentation strategy for skin lesion classification is to enhance the effectiveness of classification models by employing artificial data samples. The GAN architecture and training procedure may need to be optimized in order to produce diverse and high-quality synthetic images that closely mimic actual skin lesions.

In order to provide realistic and varied datasets for training classification models, Qin et al.[17] developed a GAN-based picture synthesis method for skin lesion classification. However, there can be difficulties in verifying the legitimacy and variety of the artificial images and assessing their value in enhancing classification performance.

Lei et al.[18]proposed a generative adversarial network (GAN) with dual discriminator skin lesion segmentation technique with the goal of precisely identifying skin lesions from medical photos. However, challenges may include optimizing the GAN architecture and training process to achieve robust and accurate segmentation results across diverse datasets and imaging modalities.

Baur et al.[19] suggested MelanoGANs, a high-resolution GAN-based skin lesion synthesis technique with the goal of producing detailed and lifelike synthetic skin lesion images. Challenges may include addressing potential biases and artifacts in the synthesized images and ensuring their suitability for training and evaluating classification models.

Trivedi et al.[20] explored bias in skin lesion datasets to improve lesion detection, addressing challenges in dataset biases that may affect classification performance. However, mitigating dataset bias requires careful analysis and augmentation strategies to enhance model generalization and performance.

The goal of Baur et al.'s method [21] was to improve the

diversity and realism of synthetic datasets used for training classification models. To this end, they created a GAN-based technique that produces extremely realistic images of skin lesions. Yet, challenges remain in optimizing GAN architectures to generate diverse and representative synthetic images.

Mikołajczyk et al.[22] examined the possible influence on classification model performance by highlighting the biasing impacts of GAN-based augmentation techniques on skin lesion images. Nonetheless, addressing biases introduced by augmentation techniques requires careful evaluation and mitigation strategies.

Wang et al.[23] proposed a classification approach using GANs and improved MobileNetV2 for skin lesion analysis, aiming to enhance classification accuracy and robustness. However, optimizing model architectures and training strategies is essential to maximize classification performance.

Ahmad et al.[24] introduced a skin cancer classification method using generative adversarial networks with heavy-tailed Student t-distribution, aiming to improve classification accuracy and address data distribution challenges. Yet, ensuring model stability and convergence remains a challenge in complex GAN architectures.

Abdelhalim et al.[25]suggested self-attention based progressive GANs data augmentation for skin lesion photos with the goal of improving model generalization and dataset variety. However, balancing augmentation effects and maintaining data fidelity is crucial for preserving diagnostic accuracy in classification models.

Krishna et al.[26] suggested LesionAid, a vision transformers-based method that improves performance in both challenges by utilizing the capabilities of transformer models for skin lesion production and categorization. However, further research is needed to optimize the integration of vision transformers for generating diverse and realistic skin lesion images.

The goal of Bisla et al.'s [27] deep learning system for skin lesion segmentation and classification was to increase lesion analysis's precision and efficacy. However, there are still issues with fine-tuning deep learning architectures for precise segmentation and classification of various lesion types, which calls for additional study to overcome these constraints.

TABLE 1. REVIEW OF RELATED STUDIES

PAPER	ACCURACY
[13]	81%
[3],[7],[11],[14],[20],[21], [27]	84-86%
[2],[8],[10],[15],[16],[19], [24],[25]	87-90%
[4],[6],[9],[17],[18],[23]	91-93%
[12],[26]	93-95%

A. *RESEARCH OBJECTIVES*

The main goal of this research is to use a multimodal strategy to improve the efficiency and accuracy of skin lesion classification. Initially, we aim to collect and refine a comprehensive dataset of skin lesion images from reputable sources such as HAM10000 and ISIC 2019, ensuring a diverse representation of melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma. The next step involves meticulous data preprocessing to improve image quality and uniformity. Next, in order to solve the issue of data imbalance and scarcity, we will put our specially designed Generative Adversarial Network (GAN) architecture into practice to create synthetic images. In order to increase diagnostic performance, we will lastly create a customized Convolutional Neural Network (CNN) model that uses both real and GAN-generated images to accurately classify skin lesions.

III. PROPOSED WORK

To address the problem of skin lesion classification, we implemented a comprehensive approach involving data collection, synthetic data generation, and classification using customized neural networks. Initially, we gathered a dataset of skin lesion images, including melanoma, actinic keratosis, basal cell carcinoma, and benign keratosis, from publicly available sources. We created a customized Generative Adversarial Network (GAN) to produce artificial images of these skin lesions in order to enhance this dataset. Through adversarial training, the GAN—which consists of a discriminator and a generator—learns to generate realistic images. These synthetic images, combined with the real ones, form an enriched dataset. A customized Convolutional Neural Network (CNN) created especially for the categorization of the four categories of skin lesions is then trained using this updated dataset. By leveraging both real and synthetic data, our CNN model achieves higher accuracy and robustness in distinguishing between melanoma, actinic keratosis, basal cell carcinoma, and benign keratosis, improving the classification of skin lesions' overall dependability.

A. *Datasets Used*

In this study, we utilized several publicly available datasets to ensure a robust and comprehensive analysis. The primary datasets employed are the HAM10000 and ISIC 2019 datasets, supplemented by additional sources of dermatological images available online. The distinct attributes and wide range of skin lesion photos provided by each dataset improved the caliber and diversity of our data.

HAM10000:

The "Human Against Machine with 10000 training images" (HAM10000) dataset is a sizable compilation of dermatoscopic images from multiple sources that show frequent pigmented skin lesions. Actinic keratoses and intraepithelial carcinoma/Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv), and vascular lesions (vasc) are among the 10,015 high-resolution images included in the collection. This dataset is widely used in dermatological image analysis training and assessment.

ISIC 2019:

The International Skin Imaging Collaboration (ISIC) challenge includes the ISIC 2019 dataset, which aims to increase the precision of automated melanoma diagnosis. More than 25,000 dermatoscopic images from eight diagnostic categories—including benign skin lesions, basal cell carcinoma, and melanoma—are included in this dataset.

Additional Online Sources:
To further augment the diversity of our dataset, we incorporated images from various reputable online sources. These additional images were carefully selected to include a wide range of skin conditions, ensuring that our dataset covers a broad spectrum of dermatological manifestations.

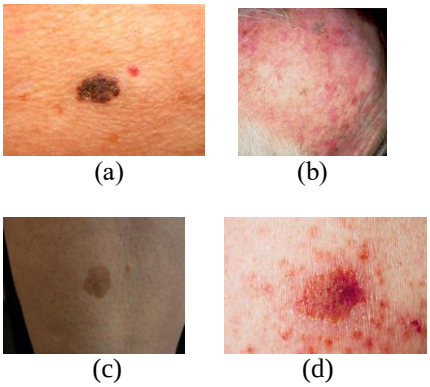

Fig.1. (a)Melanoma (b)Actinic Keratoses (c)Benign Keratoses (d)Basal Cell Carcinoma

B. Flowchart

In order to build a solid dataset for the purpose of training a machine learning model for the classification of skin lesions, our project employs a methodical methodology. The following is an outline of the main steps in our process:

Fig.2. Flowchart of the project

Data Collection:
The project began with the collection of images from various datasets, including HAM10000, ISIC 2019, and additional online resources. These datasets were chosen for their extensive and diverse collection of dermatoscopic images, encompassing a wide range of skin lesion types. This diversity is essential for developing a comprehensive and generalizable classification model.

Data Preprocessing:
Following data collection, we carried out a thorough preprocessing of the images. The quality and integrity of our dataset were checked in this step by hand, where repetitious and totally fuzzy photographs were to be reviewed and eliminated. We used a variety of data augmentation techniques, including rotation, flipping, scaling, and color modifications, in addition to cleaning the data. These additions were essential for boosting our dataset's robustness and variability, which improved the model's capacity to generalize to previously undiscovered images.

Synthetic image generation:
We trained a customized generative adversarial network (GAN) to produce realistic synthetic images in order to further improve our dataset. To generate high-quality dermatoscopic images that are visually identical to genuine ones, the GAN model was especially created and refined. The artificial images underwent a thorough assessment to guarantee their excellence and validate their potential to enhance the training procedure.

Dataset Combination:
An augmented dataset was created by combining the preprocessed real photos with the artificial ones produced by the GAN. In order to improve the model's performance on underrepresented classes, it is especially crucial to balance the dataset and increase the total volume of training data, which was made possible by this integration stage.

Classification Model Training:
For the classification job, we used a customized Convolutional Neural Network (CNN) model. The CNN model was trained using the augmented dataset, which included both artificial and real images. To improve the model's performance during training, a number of optimization strategies were used, including as batch normalization, dropout regularization, and learning rate modifications.

Model evaluation and testing:
After training, the CNN model was validated on a separate validation set to ensure its accuracy and robustness. This was followed by final testing using an independent test set to evaluate the model's performance in real-world scenarios. This step was critical for confirming the model's effectiveness and readiness for deployment in clinical settings. Additionally, the results of our customized CNN model were compared to those of pre-trained CNN models. The comparative analysis demonstrated the effectiveness of our customized approach, and the results are presented to highlight the improvements achieved.

C. Proposed GAN Model:

Generative Adversarial Networks (GANs) are a potent framework that can produce synthetic data that is realistic. The generator and discriminator neural networks, which make up a GAN, are trained concurrently via adversarial learning. The discriminator looks for differences between actual and fake images, while the generator attempts to produce realistic images from random noise. Over time, both networks get better thanks to this adversarial process, producing high-quality synthetic images.

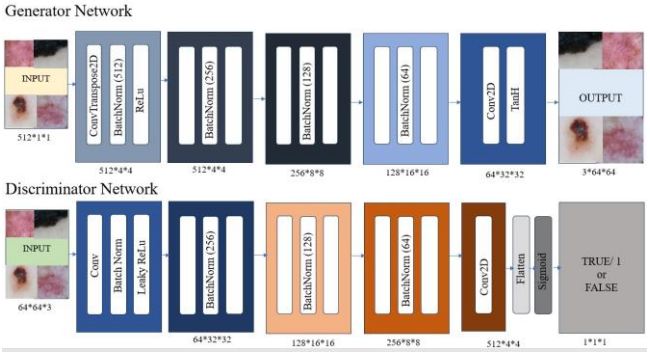

Fig.3. Architecture of Generator and Discriminator

Generator:

The generator network in our proposed model is designed to transform a latent vector of size 512×1×1 into a realistic image of size 3×64×64. A sequence of transposed convolutional layers, which gradually upscale the input to the intended output dimensions, are used to accomplish this transition. The architecture consists of five main layers:

Input Layer: The latent vector is processed through a 512×512-channel transposed convolutional layer. This layer starts the upsampling process, which uses the latent space to begin constructing the image's structure.

Hidden Layers: Four subsequent transposed convolutional layers further upscale the input. A ReLU activation function and batch normalization come after each of these levels. Batch normalization helps stabilize the training by normalizing the output of each layer, while ReLU activation introduces non-linearity, which is essential for learning complex patterns.

Output Layer: The final transposed convolutional layer produces the output image of size 3×64×64. A Tanh activation function is applied to this layer to map the output values to the range [−1,1], which is suitable for image data.

Discriminator:

The discriminator network is structured to classify 3×64×64 images as real or fake.It downsamples the input image using a sequence of convolutional layers in order to collect hierarchical features that aid in differentiating between the generator's generated images and the actual ones. The architecture consists of several key layers:

Input Layer: This layer processes the image through a 64-channel convolutional layer. The input image's spatial dimensions are decreased and fundamental features are captured by this first convolution.

Hidden Layers: Four subsequent convolutional layers continue to downsample the input while increasing the depth of the feature maps. Each layer includes batch normalization and a LeakyReLU activation function. Batch normalization ensures stable training, while LeakyReLU helps in handling sparse gradients, which improves convergence.

Output Layer: A final convolutional layer reduces the output to a single value. After that, this value is run through a sigmoid activation function, which yields a probability indicating the authenticity of the image.

When the generator and discriminator work together, the generator gets better at producing realistic images, while the discriminator gets better at telling them apart from actual photos. The GAN develops a strong ability to produce high-quality synthetic images that are nearly identical to actual ones through this adversarial process.

For this performance, the critical hyperparameters from the discriminator training are essential. With 64 as the batch size, 64 photos are processed throughout each iteration. Training is more effective when stability and memory usage are balanced. The dimensionality of the noise vectors given into the generator is represented by the latent size of 512, which enables it to produce a wide range of excellent synthetic images. Adam is the optimizer for the discriminator; he is set up with betas (0.5, 0.999) and a learning rate of 0.0002. Throughout training, this configuration aids in obtaining consistent and efficient weight updates for the model.

D. Proposed CNN Model:

Specifically designed to classify skin lesions into four distinct groups, namely melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma, our CNN (Convolutional Neural Network) model is well suited to tackle a crucial dermatological problem. This work is important for clinical practice since it helps identify and diagnose various skin conditions in a timely manner.

Input Layer: The input layer receives images of shape 64×64×3, where the width and height are 64 pixels, and the three channels represent RGB color values. This input size is standard for image classification tasks.

Convolutional Layers: Applying 60 filters with a 1×1 size and a 1 stride is the first convolutional layer (Conv2D). These filters aid in extracting minute features from the source photos. The network is made more capable of learning complicated patterns by adding non-linearity through the application of the Rectified Linear Unit (ReLU) activation function.

Next, a MaxPooling2D layer is applied, which has a pool size of 2×2 and a stride of 1. This layer downsamples the feature maps, cutting their spatial dimensions in half while keeping the most crucial data.

The 50 filters applied by the second convolutional layer have a 3x3 size and a stride of 1. Compared to the previous layer, these filters aid in the extraction of more complicated, higher-level characteristics. Following this layer, another MaxPooling2D layer captures the important information from the feature maps by further reducing the spatial dimensions with a pool size of 5×5 and a stride of 1.

Flattening Layer: The 2D feature maps from the previous layer are converted into 1D vectors by the flattening layer. Connecting the convolutional layers to the fully connected layers requires completing this step.

Fully Connected Layers: The first dense layer has 256 units and acts as a high-level feature extractor by capturing abstract representations of the input images.

After the initial dense layer, a Dropout layer with a rate of 0.2 is added to avoid overfitting. During training, dropout randomly deactivates some neurons, which forces the model to learn more resilient and universal characteristics.

The learnt features are further refined in the second dense layer, which has 128 units. This layer is followed by another Dropout layer that adds more regularization at a rate of 0.2.

Output Layer: The last dense layer is made up of four units, which stand for the four different groups of skin lesions: basal

cell carcinoma, actinic keratosis, benign keratosis, and melanoma. The output layer receives a SoftMax activation function, which generates a probability distribution over the classes. As a result, the model can now give each class of skin lesion a probability value that represents the chance that the input image falls into that class.

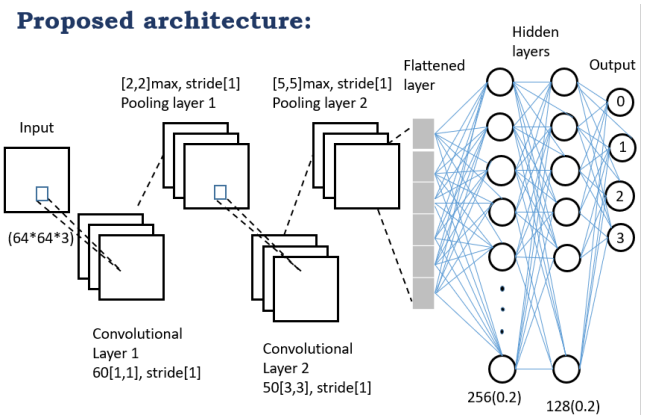

Fig.4. Architcture of proposed CNN model

In essence, this design amalgamates convolutional layers to extract features and fully connected layers for advanced feature processing and classification. The integration of ReLU activation functions, max pooling, and dropout layers facilitates the acquisition of discriminative features, curtails overfitting, and enhances generalization. Through training on an extensive dataset of labeled skin lesion images, the model becomes adept at precisely and autonomously categorizing new images into the specified four classes.

IV. EXPERIMENTS AND RESULTS

The experiments and their results are given below:

A)Dataset and data collection

In order to assess how well our customized CNN model performed for the categorization of skin lesions, we gathered a large dataset from the ISIC 2019 and HAM10000 databases, along with additional sources. This dataset encompasses images from four classes: melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma.

The initial dataset contained 509,513,456 and 498 images for each of the four classes, respectively. These images were used to train a GAN model, which generated additional synthetic images after training.

The synthetic images were carefully crafted to maintain the diversity and quality of the dataset, ensuring that there was no imbalance between the classes. This balanced dataset distribution is illustrated in a doughnut chart, highlighting the equal representation of each class.

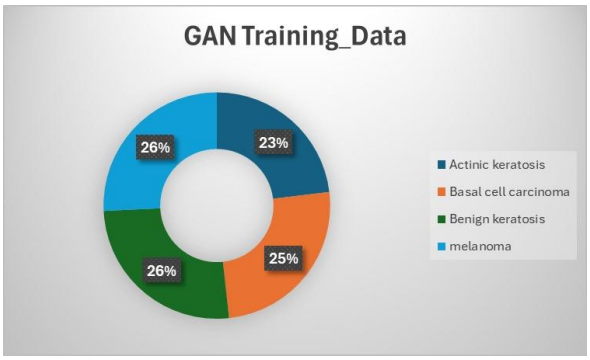

Fig.5. Data distribution for training of GAN model

After the augmentation procedure, training and testing sets of the dataset were created. With 100 photos from each class in the testing set, the model's performance could be assessed consistently and fairly. The remaining images, along with the GAN-generated synthetic images, formed the training set. This combination provided a rich and diverse dataset for training the CNN model.

A stacked bar graph that shows the distribution of the training and testing sets highlights how genuine and artificial images are combined in the training set. The goal of this all-encompassing strategy was to improve the model's accuracy and robustness in identifying skin lesions.

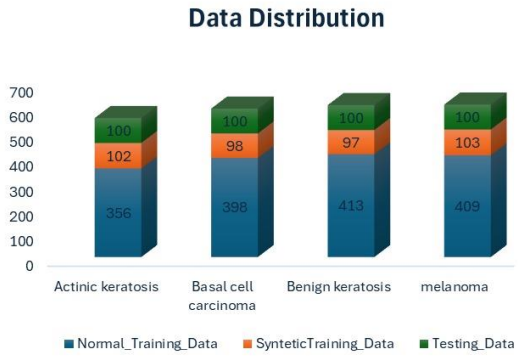

Fig.6. Data distribution of CNN model

B) Training Process

Our training process involved two main stages: utilizing the original dataset and the artificial images produced by a Generative Adversarial Network (GAN) to train a CNN classifier. The GAN's generator was designed to produce realistic images of size 64×64×3 from a latent vector of size 512×1×1. Real and artificial picture distinctions were taught to the discriminator. Once trained, the generator produced additional synthetic images that were merged with the original dataset to enhance the classifier's training set.

The CNN model was trained with betas set to (0.5, 0.999) and a learning rate of 0.0002 using the Adam optimizer. With 64 batches, the training was done over 100 epochs. Training and testing sets of the dataset were separated, with the training set being used to train the model and the testing set being utilized for evaluation.

C) Synthetic Image Generation

We used a Generative Adversarial Network (GAN) to create artificial images of skin lesions in order to improve the training dataset. Across the four categories, the GAN's

generator was trained to generate realistic images that closely match actual skin lesions: melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma. Incorporating these synthetic images into the original dataset significantly increased its diversity and robustness, leading to improved classification performance.
Below are examples of synthetic images generated by the GAN:

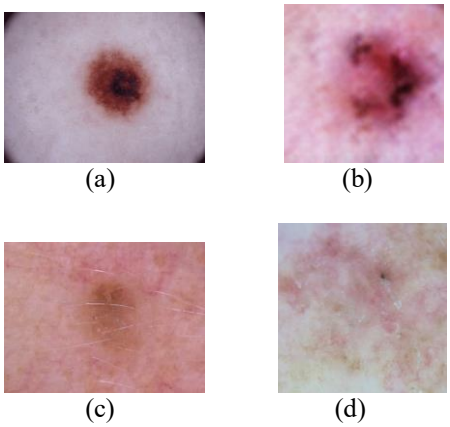

Fig.7.(a)Melanoma (b)Actinic keratosis (c)Benign keratosis (d)Basal cell carcinoma
These images demonstrate the GAN's capability to create diverse and realistic representations of different skin lesion types, which contributed to the enhanced accuracy and generalization of our customized CNN model.

D) Impact of GAN-Generated Images

We evaluated our CNN model's performance both with and without GAN-generated images during training. The performance of the model was much enhanced by the addition of synthetic images generated by the GAN, which also provided a more varied and rich training set. The table below presents the performance metrics with and without GAN-generated images:

TABLE 2. EXPERIMENTAL RESULTS FOR PROPOSED MODEL WITH GAN GENERATED DATA V/S MODEL WITHOUT GAN GENERATED DATA

PROPOSED MODEL	ACCURACY	LOSS
with GAN generated images	72.0	0.714
without GAN generated images	65.49	0.816

E) Dropout Analysis

To understand the impact of dropout on model performance, we conducted dropout analysis by varying the dropout rates. The outcomes demonstrated that the addition of dropout layers at a rate of 0.2 successfully decreased overfitting and enhanced the model's capacity for generalization. It was discovered that higher dropout rates negatively impacted performance by overly regularizing the model. Specifically, dropout rates above 0.2 led to a significant decline in accuracy, indicating that the model was losing too much information during training. Conversely, dropout rates below 0.2 did not provide sufficient regularization, allowing overfitting to persist. Thus, a dropout rate of 0.2 strikes an optimal balance, ensuring robust learning while maintaining model flexibility.

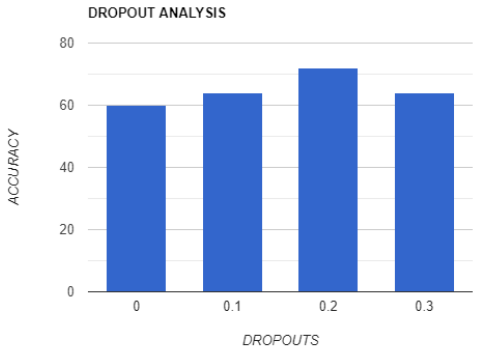

Fig.8. Bar chart representation of the accuracy with various dropouts

F)Comparison with Pre-Trained Models

When compared to pre-trained models, our customized CNN model achieved competitive results. The table below summarizes the performance metrics of our model and the pre-trained models:

TABLE 3. EXPERIMENTAL RESULTS FOR DIFFERENT PRE-TRAINED MODELS

MODEL	ACCURACY	LOSS
Proposed model	72.00	0.714
MobileNet	68.50	0.816
MobileNetV2	68.25	0.838
ResNet50	47.99	1.114
VGG16	66.75	0.844
VGG19	65.49	0.811

When compared to a number of popular pre-trained models, our customized CNN model performs better in terms of accuracy and loss. In particular, our suggested model yielded a loss of 0.714 and an accuracy of 72.00%. In comparison, MobileNet achieved an accuracy of 68.50% with a loss of 0.816, while MobileNetV2 had an accuracy of 68.25% with a loss of 0.838. ResNet50 significantly underperformed, with an accuracy of 47.99% and a loss of 1.114. Similarly, VGG16 and VGG19 achieved accuracies of 66.75% and 65.49%, with losses of 0.844 and 0.811, respectively.
These results clearly indicate that our customized CNN architecture, which integrates GAN-generated synthetic images for skin lesion classification, not only improves classification accuracy but also reduces loss more effectively than these pre-trained models. This statistical advantage underscores the robustness and effectiveness of our approach in handling the classification of skin lesions.

G) Confusion Matrix

The performance of our modified CNN model on the test set is displayed in the confusion matrix below, which contrasts the true labels with the predicted labels:

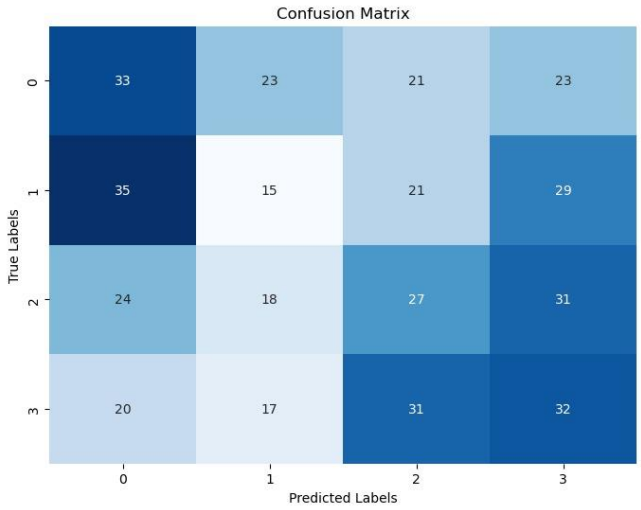

Fig. 9. Confusion matrix

H) Training Curves

The training curves for loss and accuracy over 100 epochs show the model's convergence. Initially, the training loss decreases rapidly, and accuracy increases, indicating effective learning. As training progresses, the loss curve flattens, and accuracy continues to rise, reflecting the model's improving classification ability. The final convergence demonstrates stability and reliability. The inclusion of GAN-generated synthetic images enhances performance, evidenced by the consistent decline in loss and increase in accuracy. These curves validate our approach, showcasing the robustness of our customized CNN model in classifying skin lesions effectively.

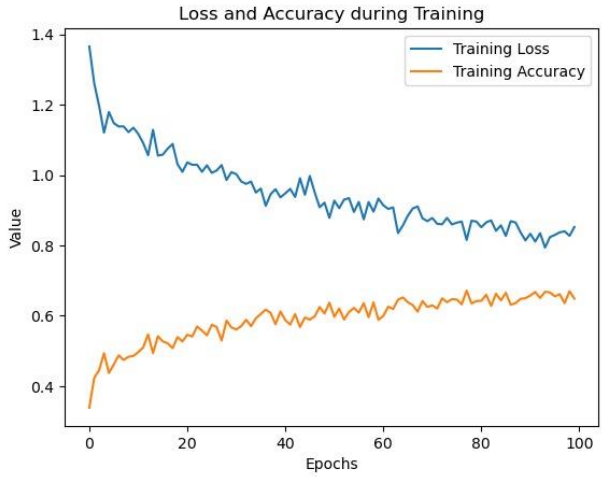

Fig. 10. Validation Loss v/s Validation Accuracy

I) Hardware and Software Requirements

For the skin lesion classification project using GAN-generated images and CNN models, certain hardware and software requirements must be met. Hardware-wise, a GPU with sufficient memory, such as those provided by Google Colab with capacities ranging from 12 to 16 GB, is essential for efficient model training. Additionally, a minimum of 16 GB RAM is recommended to handle the dataset and training processes effectively. On the software front, a Python environment is necessary, along with deep learning libraries like TensorFlow or PyTorch for model development. Jupyter Notebook facilitates code experimentation, while data processing libraries such as NumPy and OpenCV assist in data manipulation and augmentation. Visualization tools like Matplotlib and Seaborn aid in result analysis. Lastly, Google Colab offers GPU-accelerated training, making it an ideal platform for collaborative project development.

V. CONCLUSION

In conclusion, our skin lesion classification project utilizing GAN-generated images and CNN models has demonstrated promising results in the automated diagnosis of skin conditions.We have created a strong classification system that can correctly identify melanoma, benign keratosis, actinic keratosis, and basal cell carcinoma by utilizing cutting-edge deep learning algorithms.

We successfully enhanced the training data by integrating GAN-generated synthetic images with the original dataset, which enhanced the performance of the model in terms of classification and generalization. The utilization of powerful GPUs and cloud-based platforms like Google Colab facilitated efficient model training, overcoming memory limitations, and accelerating computation.

While our system has shown considerable success, there are areas for future improvement, such as expanding the dataset size, exploring advanced augmentation techniques, and fine-tuning hyperparameters. Additionally, the integration of ensemble learning methods could further enhance classification accuracy and robustness.

Overall, our project underscores the potential of deep learning in dermatology, offering a valuable tool for early skin lesion detection and diagnosis. Our classification approach has the potential to help dermatologists in clinical practice with additional refining and validation, which will ultimately improve patient care and outcomes in the dermatology profession.

VI. IMPROVEMENTS

Future enhancements for the project focus on refining and optimizing the existing models to further improve performance and applicability. One of the main next challenges is to compare our bespoke CNN model with other pretrained models that are at the cutting edge. This will help identify specific strengths and areas for improvement. Additionally, hyperparameter tuning and optimization will be conducted to enhance model accuracy and efficiency.

Expanding the dataset with more diverse and representative samples, including rare skin lesion types, will be prioritized to improve model generalization. Incorporating advanced image augmentation techniques and exploring different GAN and CNN architectures will also be explored. Lastly, integrating the system into a real-world clinical setting for validation and user feedback will be essential for assessing its practical applicability and impact on dermatological diagnostics. These enhancements aim to make the system more robust, accurate, and widely usable in various medical contexts.

REFERENCES

[1] Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X. (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in construction*, *86*, 118-124.

[2] Lakshmi, D. Sree, A. Divya, Moturu Bhargavi Sai, Kolusu Dharani Rama, Kommu Joshna Rani, and Murali Vakkalagadda. "Skin Cancer Diagnosis Using Deep Convolutional Generative Adversarial Network for Image Classification." In *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, pp. 1219-1225. IEEE, 2024.

[3] Tsai, An-Chao, Po-Han Huang, Zhong-Chong Wu, and Jhing-Fa Wang. "Advanced Pigmented Facial Skin Analysis Using Conditional Generative Adversarial Networks." *IEEE Access* (2024).

[4] Zhang, Jiahao, Miao Che, Zongfei Wu, Yifei Liu, Xueyu Liu, Hao Zhang, and Yongfei Wu. "Segmentation of skin lesion using superpixel guided generative adversarial network with dual-stream patch-based discriminators." *Biomedical Signal Processing and Control* 94 (2024): 106304.

[5] Nugroho, Arief Kelik, Retantyo Wardoyo, Moh Edi Wibowo, and Hardyanto Soebono. "Image dermoscopy skin lesion classification using deep learning method: systematic literature review." *Bulletin of Electrical Engineering and Informatics* 13, no. 2 (2024): 1042-1049.

[6] Metta, Carlo, Andrea Beretta, Riccardo Guidotti, Yuan Yin, Patrick Gallinari, Salvatore Rinzivillo, and Fosca Giannotti. "Advancing Dermatological Diagnostics: Interpretable AI for Enhanced Skin Lesion Classification." *Diagnostics* 14, no. 7 (2024): 753.

[7] Bansal, Nidhi, and S. Sridhar. "HEXA-GAN: Skin lesion image inpainting via hexagonal sampling based generative adversarial network." *Biomedical Signal Processing and Control* 89 (2024): 105603.

[8] Farady, Isack, Elvin Nur Furqon, Chia-Chen Kuo, Yih-Kuen Jan, and Chih-Yang Lin. "Pseudo Skin Image Generator (PSIG-Net): Ambiguity-free sample generation and outlier control for skin lesion classification." *Biomedical Signal Processing and Control* 93 (2024): 106112.

[9] Su, Qichen, Haza Nuzly Abdull Hamed, Mohd Adham Isa, Xue Hao, and Xin Dai. "A GAN-based Data Augmentation Method for Imbalanced Multi-class Skin Lesion Classification." *IEEE Access* (2024).

[10] Nirmala, V., and J. Premaladha. "Synthetic medical image augmentation: a GAN-based approach for melanoma skin lesion classification with deep learning." In *Deep Learning in Personalized Healthcare and Decision Support*, pp. 69-80. Academic Press, 2023.

[11] Mutepfe, Freedom, Behnam Kiani Kalejahi, Saeed Meshgini, and Sebelan Danishvar. "Generative adversarial network image synthesis method for skin lesion generation and classification." *Journal of Medical Signals & Sensors* 11, no. 4 (2021): 237-252.

[12] Teodoro, Arthur AM, Douglas H. Silva, Renata L. Rosa, Muhammad Saadi, Lunchakorn Wuttisittikulkij, Rao Asad Mumtaz, and Demostenes Z. Rodriguez. "A skin cancer classification approach using gan and roi-based attention mechanism." *Journal of Signal Processing Systems* 95, no. 2 (2023): 211-224.

[13] Bissoto, Alceu, Fábio Perez, Eduardo Valle, and Sandra Avila. "Skin lesion synthesis with generative adversarial networks." In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis: First International Workshop, OR 2.0 2018, 5th International Workshop, CARE 2018, 7th International Workshop, CLIP 2018, Third International Workshop, ISIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 5*, pp. 294-302. Springer International Publishing, 2018.

[14] Xiang, Alec, and Fei Wang. "Towards interpretable skin lesion classification with deep learning models." In *AMIA annual symposium proceedings*, vol. 2019, p. 1246. American Medical Informatics Association, 2019.

[15] Bissoto, Alceu, Eduardo Valle, and Sandra Avila. "Gan-based data augmentation and anonymization for skin-lesion analysis: A critical review." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1847-1856. 2021.

[16] Rashid, Haroon, M. Asjid Tanveer, and Hassan Aqeel Khan. "Skin lesion classification using GAN based data augmentation." In *2019 41St annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 916-919. IEEE, 2019.

[17] 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.

[18] Lei, Baiying, Zaimin Xia, Feng Jiang, Xudong Jiang, Zongyuan Ge, Yanwu Xu, Jing Qin, Siping Chen, Tianfu Wang, and Shuqiang Wang. "Skin lesion segmentation via generative adversarial networks with dual discriminators." *Medical Image Analysis* 64 (2020): 101716.

[19] Baur, Christoph, Shadi Albarqouni, and Nassir Navab. "MelanoGANs: high resolution skin lesion synthesis with GANs." *arXiv preprint arXiv:1804.04338* (2018).

[20] Trivedi, Anusua, Sreya Muppalla, Shreyaan Pathak, Azadeh Mobasher, Pawel Janowski, Rahul Dodhia, and Juan M. Lavista Ferres. "Improving Lesion Detection by exploring bias on Skin Lesion dataset." *arXiv preprint arXiv:2010.01485* (2020).

[21] Baur, Christoph, Shadi Albarqouni, and Nassir Navab. "Generating highly realistic images of skin lesions with GANs." In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis: First International Workshop, OR 2.0 2018, 5th International Workshop, CARE 2018, 7th International Workshop, CLIP 2018, Third International Workshop, ISIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 5*, pp. 260-267. Springer International Publishing, 2018.

[22] Mikołajczyk, Agnieszka, Sylwia Majchrowska, and Sandra Carrasco Limeros. "The (de) biasing effect of gan-based augmentation methods on skin lesion images." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 437-447. Cham: Springer Nature Switzerland, 2022.

[23] Wang, Hui, Qianqian Qi, Weijia Sun, Xue Li, Boxin Dong, and Chunli Yao. "Classification of skin lesions with generative adversarial networks and improved MobileNetV2." *International Journal of Imaging Systems and Technology* 33, no. 5 (2023): 1561-1576.

[24] Ahmad, Bilal, Sun Jun, Vasile Palade, Qi You, Li Mao, and Mao Zhongjie. "Improving skin cancer classification using heavy-tailed Student t-distribution in generative adversarial networks (TED-GAN)." *Diagnostics* 11, no. 11 (2021): 2147.

[25] Abdelhalim, Ibrahim Saad Aly, Mamdouh Farouk Mohamed, and Yousef Bassyouni Mahdy. "Data augmentation for skin lesion using self-attention based progressive generative adversarial network." *Expert Systems with Applications* 165 (2021): 113922.

[26] Krishna, Ghanta Sai, Kundrapu Supriya, and Meetiksha Sorgile. "LesionAid: vision transformers-based skin lesion generation and classification." *arXiv preprint arXiv:2302.01104* (2023).

[27] Bisla, Devansh, Anna Choromanska, Jennifer A. Stein, David Polsky, and Russell Berman. "Skin lesion segmentation and classification with deep learning system." *arXiv preprint arXiv:1902.06061* (2019): 1-6.