

Melanoma image synthesis: a review using generative adversarial networks

Mohammed Altaf Ahmed¹, Mohammad Naved Qureshi², Mohammad Sarosh Umar³, Mouna Bedoui⁴

¹Department of Computer Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia

²Electrical Engineering Section, University Polytechnic, Aligarh Muslim University, Aligarh, India

³Department of Computer Engineering, Zakir Hussain College of Engineering and Technology, Aligarh Muslim University, Aligarh, India

⁴Electronics and Micro-Electronics Laboratory (E. µ. E. L), Faculty of Sciences of Monastir, University of Monastir, Monastir, Tunisia

Article Info

Article history:

Received Jan 15, 2024

Revised Mar 3, 2024

Accepted Mar 25, 2024

Keywords:

Deep learning

GAN

HAM10000

ISIC

Melanoma

SDG

Sustainability

ABSTRACT

Melanoma is a highly malignant skin cancer that may be fatal if not promptly detected and treated. The limited availability of high-quality melanoma images, which are needed for training machine learning models, is one of the obstacles to detecting melanoma. Generative adversarial networks (GANs) have grown in popularity as a strong technique for image synthesis. This research is also targeted at the sustainable development goal (SDG) for health care. In this study, we survey existing GAN-based melanoma image synthesis methods. In this work, we briefly introduce GANs and how they may be used for generating synthetic images. Ensuring healthy lifestyles and promoting well-being for everyone, regardless of age, is the main aim. A comparative study is carried out on how GANs are used in current research to generate melanoma images and how they improve the classification performance of neural networks. Various public and proprietary datasets for training GANs in melanoma image synthesis are also discussed. Lastly, we assess the examined studies' performance using measures like the Frechet Inception distance (FID), Inception score, structural similarity index (SSIM), and various classification performance metrics. We compare the evaluated findings and suggest further GAN-based melanoma image-creation research.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Mohammed Altaf Ahmed

Department of Computer Engineering, College of Computer Engineering and Sciences

Prince Sattam bin Abdulaziz University

Al-Kharj, Saudi Arabia

Email: m.altaf@psau.edu.sa

1. INTRODUCTION

Skin cancer is a prevalent form of cancer, representing 1 to 2% of all new malignancies and over 90% of all identified cases of skin tumours [1], [2]. Melanoma, the least common and most fatal type of skin cancer, has a global mortality rate of 25% [3]. Early detection can significantly improve survival rates in numerous cases [4], whereas without an accurate diagnosis, the 5-year survival rate drops to 19.9%. Diagnosing melanoma early is challenging because of variations in skin tone and lesion types. Even skilled dermatologists can only identify it with 60% accuracy through visual examinations [5]. Dermoscopy, a precise technique [6], necessitates sophisticated equipment and skilled personnel, highlighting the importance of enhancing computer-aided design (CAD) systems to support dermatologists and alleviate strain on healthcare resources [7].

Image processing filters such as the Harris corner detector are used to analyse specific image features in traditional CAD and machine-based systems. Nevertheless, these techniques were labor-intensive and restricted to small datasets. In recent decades, convolutional neural networks (CNNs) and other image processing methods have attracted attention in the field of medical imaging because of their diagnostic accuracy [8]-[12]. Deep learning models with intricate structures necessitate extensive datasets to prevent overfitting.

Insufficient data in medical imaging, especially for melanoma, presents a major challenge. Biopsies provide a limited amount of training data due to the fact that they only represent a portion of cases for pathologic diagnoses [13]. Data augmentation and transfer learning are used to address the scarcity of data in medical datasets, despite facing challenges due to their limited availability [14]. Generative adversarial networks (GANs) have proven to be effective in tasks such as medical image generation, alteration, segmentation, and categorization [15]-[19].

Some of the review consolidates important contributions and findings from pivotal research in the field. Such as GANs in unsupervised and semi-supervised settings and uncovering their capabilities and constraints. A comparison of various GAN architectures and demonstrations that personalize GAN (PGAN) outperform others in producing lifelike images, and GAN-based data augmentation techniques like demonstrating superior performance compared to conventional models. Section 3 discusses these studies in detail. Yet, unresolved challenges include generalizability across datasets, computational efficiency, and detailed evaluation of synthetic image quality and biases.

This review seeks to address these gaps by thoroughly examining the methodologies, computational requirements, and reliability of synthetic images in different studies. Thus, our review aims to provide a systematic and comprehensive overview of the use of GANs in melanoma image synthesis. By highlighting current advancements and identifying areas for future development, this study contributes towards the healthcare sector's sustainable goals, particularly enhancing early detection and treatment of skin cancers. Following this introduction, the paper is organised as outlined below: section 2 outlines the research methodology for the proposed systematic review, while section 3 delves into the operations of GANs and their architectures. Section 4 provides a literature review on GANs in melanoma image synthesis. Section 5 examines standard skin lesion datasets accessible online. Section 6 investigates the current research challenges in utilising GANs for image synthesis. Section 7 wraps up the study with recommendations for future research.

2. METHOD

We conducted this systematic review to select and classify the most effective methods currently available for generating synthetic melanoma images to overcome class imbalance problems using GAN and the classification of the same at an early stage with good accuracy. Systematic literature reviews involve collecting and evaluating previously published research as per a set of predetermined evaluation criteria. These kinds of reviews help determine what previous research in the relevant field of study has uncovered [20].

Every piece of information obtained from primary sources is sorted and then analyzed. When the literature review is finished, it will provide a more reasonable, logical, and solid answer to the fundamental question that the research was attempting to answer [21]. The research papers that were relevant to melanoma image synthesis and classification using GANs made up the population of this study that was considered during the systematic literature review.

2.1. Research framework

We began our systematic review by creating a detailed conceptual framework to guide us through three phases. The planning phase consisted of creating research questions and determining the scope. We moved on to the data collection and analysis stage, during which we carried out a focused literature search using specific databases and established criteria. The last stage focused on consolidating our discoveries and reaching conclusions that effectively answered our research inquiries.

- Research questions: research question formulation is important for an effective systematic literature review. The following questions can be addressed in this study:
 - Question 1: which are the primary features of the datasets that are accessible for lesions of melanoma skin cancer?
 - Question 2: what are the major GAN methods for synthesizing images of melanoma skin cancer, and how have they improved the classification performance?
- Search criteria: we conducted a systematic review by strategically selecting search keywords related to our study's theme, including melanoma, skin cancer, image synthesis, image classification, deep learning, GAN, and generative adversarial networks. We conducted a comprehensive search across

major databases such as PubMed, Scopus, and Web of Science by using a combination of keywords and Boolean operators. This systematic search strategy was created to thoroughly gather pertinent literature, guaranteeing a strong and comprehensive dataset for our research. The keywords used to search for information relevant to skin cancer are listed in Table 1.

Criteria for selecting articles:

- Articles published between 2015 and 2023.
- Concentrate on generating melanoma images using GAN.
- Include studies on image synthesis and classification utilising GAN or alternative classifiers.
- Only articles that have undergone peer review were taken into account.
- Articles are screened initially using titles, abstracts, and keywords.
- Thorough examination of complete texts to gather data on study design, methodologies, findings, and conclusions.

Resources explored: performed literature searches in reputable databases such as PubMed, Arxiv.org, IEEE Xplore, ACM Digital Library, Springer, Science Direct, and selected open-access sources.

Table 1. Search terms for literature collection

Search keywords/terms
melanoma AND image synthesis
(melanoma OR skin cancer) AND image synthesis
(melanoma or skin cancer) AND (image synthesis OR image generation)
(melanoma or skin cancer) AND (image synthesis OR image generation) AND GAN
(melanoma or skin cancer) AND (image synthesis OR image generation) AND GAN AND classification

2.2. Selection criteria and evaluation methods

There were 26 research papers and conference reports found using the search criteria. Of the identified papers, 18 were chosen because their titles and abstract were most relevant to our research. Table 2 displays the results of the search.

Some quality control questions were attempted to be answered by carefully reading the full texts of the selected research papers. The current systematic research asks the following questions to evaluate the quality.

- i) Did the article that was reviewed include all relevant information?
- ii) Has the paper's quality been confirmed?
- iii) Does the chosen study provide satisfactory responses to the research questions?

The first quality assessment item inquired whether data augmentations using GAN methods for generating skin lesion images were discussed at length. Secondly, the standing of the journal/publisher in which a given article appeared was to confirm the article's quality. The third inquiry checked whether or not the study addressed the research questions posed in section 2. We selected only the most relevant articles from the research literature that pertained to our field. These papers have been selected because they answer the aforementioned research questions. Non-relevant text or papers that did not fully respond to the research or quality control questions were also disqualified.

The above-mentioned research questions were answered with 'true/false' responses. We assigned each true response a value of 1, and for each false response, a value of 0 was assigned. The 18 research papers chosen for the first quality control question had a topic coverage evaluation of 95%, which is very good. The chosen papers were found to be of higher quality, so a score of 90% is achieved for the second question; it is beyond average. The third question was crucial to address the primary research questions posed by the review. An indicator of the studies' ability to address the review's research questions yielded 85% results. The aggregate response rate to these high-standard questions was encouraging.

Table 2. Search results for literature review

Resource	Related articles
Springer	2
arXiv	4
IEEE Xplore	4
PubMed	3
Science direct	3
SBC open lib	1

2.2.1. Methodology justification

The chosen databases are well-known for their scientific rigour, guaranteeing access to top-quality research related to GANs and melanoma. The keywords are directly connected to the main focus of our study, guaranteeing a thorough review of relevant literature. The importance of a systematic approach lies in its necessity for an impartial evaluation, which supports our goal of addressing knowledge deficiencies regarding GAN applications in melanoma image synthesis, as outlined in the introduction.

- Technical details included: we used a binary true/false system to evaluate the relevance and quality of each paper consistently, focusing on criteria directly related to our study's scope. The true/false evaluation system quantifies 'true' as 1 and 'false' as 0, offering an objective method to assess each article's suitability for review.
- Connecting methodology to research questions and gaps in knowledge: we address our research questions on GANs in melanoma image synthesis using a systematic approach. Through analysing recent literature, we pinpoint current trends and shortcomings, successfully addressing the gaps outlined in the introduction.

2.3. Generative adversarial networks for image synthesis

GAN, a popular deep learning framework, is often used to make new images [22]. It has completely changed the field of image synthesis, making it possible to create images from scratch that look real and are of high quality [23]. It has two main components, a generator for generating synthetic images and a discriminator to check whether image is fake or real using a neural network trained jointly in a zero-sum game. During the training process, the generator and the discriminator play a game together in which the random noise is given to generator which create a image from it that convince the discriminator that it is real. In contrast, the discriminator classifies whether the image is real or fake. Because the two networks are always competing, the generator learns from discriminator feedback and further make images look more real so the discriminator is not able to distinguish it from real one.

After the training process is completed for the GAN, the generator is used to generate new images that share characteristics with the training data. The GAN can generate these new images, and because of this, GANs are applicable in image editing, style transfer, and the synthesis of realistic images of people, objects, and landscapes. The popular GANs architecture for image synthesis is as discussed:

2.3.1. Vanilla GAN

Vanilla GAN comprises a generator and discriminator as given in Figure 1, which were first proposed by Goodfellow *et al.* 2014 as a classic GAN [24]. The generator generates images with the random noise and the discriminator classifies whether the image is synthetic or real. Iteratively, the generator tries to create images that are indistinguishable from real ones, approximating Nash equilibrium-the point where the generator consistently conned or deceived the discriminator. The training process of Vanilla GAN includes unique loss functions.

- Discriminator loss:

$$L_D = -E[\log(D(x))] - E[\log(1 - D(G(z)))] \quad (1)$$

where $D(x)$ is the real data output from the discriminator, $G(z)$ is the noise output from the generator, and E is the expected value.

- Generator loss:

$$L_G = -E[\log(D(G(z)))] \quad (2)$$

where $D(G(z))$ is the discriminator output for the data $G(z)$.

- Update rules: when training the discriminator, L_D is made as small as possible in relation to its parameters.

$$\theta_D = \theta_D - \alpha * \frac{\partial L_D}{\partial \theta_D} \quad (3)$$

Where θ_D , are the parameters of the discriminator and is α the learning rate. In order to train the generator, L_G is minimized in relation to its parameters:

$$\theta_G = \theta_G - \alpha * \frac{\partial L_G}{\partial \theta_G} \tag{4}$$

where θ_G represents the generator's parameters and α represents the learning rate.

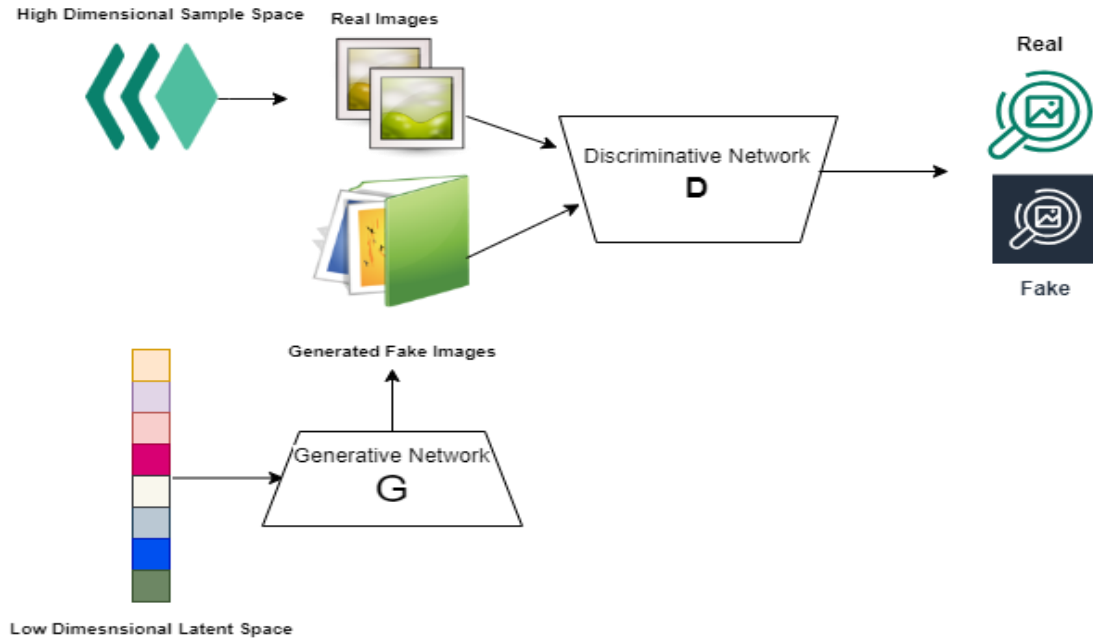


Figure 1. Vanilla GAN architecture

2.3.2. Conditional GAN

Conditional GAN receives class labels or other pertinent data as an additional conditioning input [25]. The generator aims to produce realistic samples corresponding to the provided conditioning information as shown in Figure 2. The discriminator role is to find whether the image is real or fake sample which is conditioned on the same information. The CGAN equations can be expressed as follows:

- The generator G accepts a noise vector z and conditioning information y as inputs and generates a fake sample \hat{x} :

$$\hat{x} = G.(z, y) \tag{5}$$

- The discriminator D receives as input a real sample x and conditioning information y, and generates a probability score $D(x, y)$, indicating the likelihood that x is real:

$$D(x, y) \tag{6}$$

- The discriminator also accepts a fake sample \hat{x} generated by the generator G and conditioning information y as inputs and generates a probability score $D(\hat{x}, y)$, indicating the probability that \hat{x} is fake:

$$D(\hat{x}, y) \tag{7}$$

- Generator G seeks to minimize the objective function listed below:

$$\min_G \max_D E[\log D(x, y)] + E[\log(1 - D(G(z, y), y))] \tag{8}$$

where E represents the expected value, log represents the natural logarithm, and z is a noise vector sampled from a prior distribution such as the normal or uniform distribution.

- The discriminator D seeks to maximize the same objective function as the generator, but in reverse:

$$\max_D \min_G E[\log D(x, y)] + E[\log(1 - D(G(z, y), y))] \tag{9}$$

Minimizing this objective function enables the generator and discriminator to play a minimax game in which the generator learns to generate realistic samples that match the conditioning information. The discriminator learns to distinguish between real and fake samples conditioned on the same information.

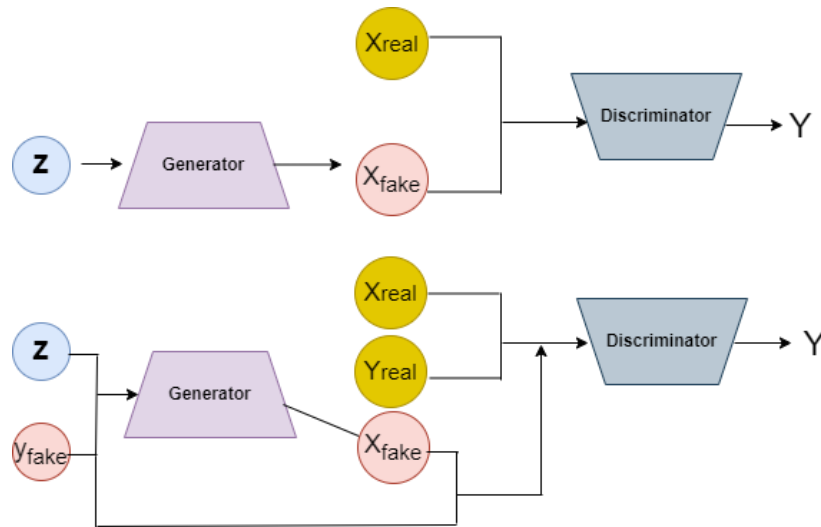


Figure 2. Conditional GAN architecture

2.3.3. Deep convolutional GAN

Deep convolutional GAN (DCGAN) use CNNs in both the generator and discriminator. DCGAN was introduced by Radford *et al.* [26] in 2016 and has since become one of the most popular image synthesis architectures. Figure 3 shows DCGAN generates an image by passing a normal distribution noise vector through several layers of transposed CNNs (deconvolutional layers). Each generator layer increases image spatial resolution while decreasing channels. The most important equations in DCGAN are:

A. Generator

The noise vector z is used as an input, and the output is an image $G(z)$. Most of the time, the generator is made with deconvolutional layers, also called transposed convolutional layers. These layers learn to upsample the noise vector to create an image. The following is the way the generator can be described mathematically:

$$G(z) = f(Wz + b) \quad (10)$$

where b is a bias vector, W is a weight matrix, and f is a non-linear activation function like rectified linear unit (ReLU).

B. Discriminator

The discriminator receives an input image, x , and returns a scalar, $D(x)$, which represents the likelihood that x is an actual image (as opposed to generated). Convolutional layers, which learn to downsample the input image to a smaller feature representation, are commonly used to implement the discriminator. Mathematical expressions for the discriminator are as (11).

$$D(x) = g(Wx + b) \quad (11)$$

Non-linear activation functions like sigmoid are used in the formula, where W is a weight matrix, and b is a bias vector.

The adversarial loss is used while training a discriminator and a generator. The discriminator aims to perform the exact opposite of what the generator does, which is to produce images that the discriminator is unable to distinguish from genuine ones. The adversarial loss is represented mathematically as (12).

$$L_{adv} = -E[\log(D(x))] - E[\log(1 - D(G(z)))] \quad (12)$$

Where x is a real image, z is a noise vector, and $G(z)$ is the generated image; E is the expectation over the real data distribution and the noise distribution; the \log is the natural logarithm; and z is the generated image.

The image quality of generated images can be improved through training with the generator loss. The core concept is that the generator should produce fake images that the discriminator will mistake for the real thing. The mathematical expression for the generator loss is as (13).

$$L_{gen} = -E[\log(D(G(z)))] \tag{13}$$

Where E stands for the expectation over the noise distribution, \log stands for the natural logarithm, z is a noise vector, and $G(z)$ is the generated image.

It is trained with the discriminator loss to improve the discriminator's ability to tell fake from real images. The theory behind it is that the discriminator should return a high probability for real images and a low probability for fake ones. Here is a mathematical representation of the discriminator loss:

$$L_{dis} = -E[\log(D(x))] - E[\log(1 - D(G(z)))] \tag{14}$$

where x is a real image, z is a noise vector, $G(z)$ is the generated image, E is the expectation over the real data distribution and the noise distribution, the \log is the natural logarithm, and x is the natural logarithm.

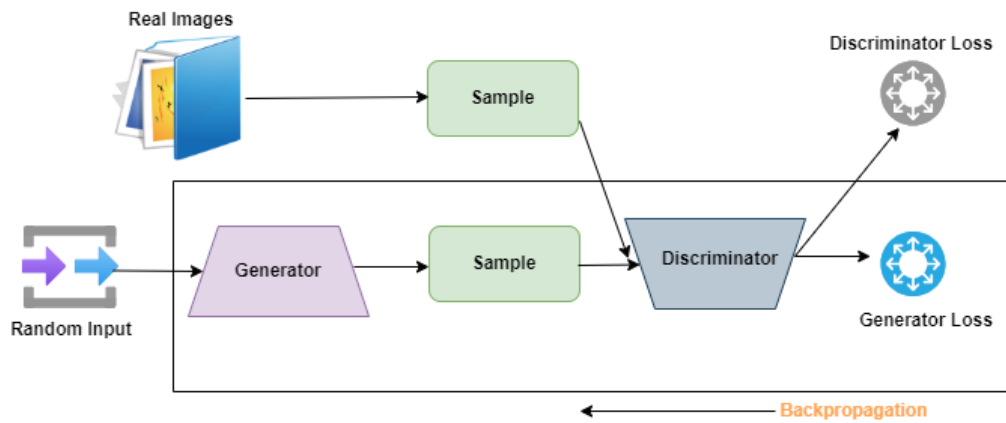


Figure 3. DCGAN architecture

2.3.4. StyleGAN

Karras *et al.* [27] introduced a new type of GAN called StyleGAN. The novel feature of StyleGAN is the use of a mapping network for first converting the input noise into a latent space before applying to the image generation process. The mapping network is composed of several fully connected layers; it turns the input noise vector into a high-dimensional latent space Figure 4. The latent space is then used as an input to a generator network, which comprises several convolutional layers that generate the final image. The core component of StyleGAN is a mapping network, which receives a noise vector z as input and produces a latent code w as output.

$$w = F(z) \tag{15}$$

Where F is a fully connected neural network that transforms z , a vector representing noise, into a vector representing the latent space.

- Noise injection: StyleGAN randomizes noise at various levels to ensure diversity in the generated images using this method. To the intermediate feature maps of the generator, nonzero noise is added via a sample from $N(0, 1)$.

$$\hat{X} = x + \alpha * \epsilon \tag{16}$$

Where x represents the intermediate feature map, α is a scalable factor that can be trained, and ϵ is the noise vector.

- Adaptive instance normalization (AdaIN): it is used to alter the appearance and sensation of the generated images. AdaIN normalizes the feature maps x by the style vectors s_i according to mean and standard deviation of the style vector.

Here $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation of the feature map x , and are learnable scale and shift parameters, and s_i is the i^{th} element of the style vector, then.

$$Y = \sin\left(\gamma + \left(\frac{x - \mu(x)}{\sigma(x)}\right) * \beta\right) \tag{17}$$

The generator takes a latent code w as input and outputs an image x .

$$x = G(w) \tag{18}$$

Where G is a convolutional neural network that converts the latent code w into an image.

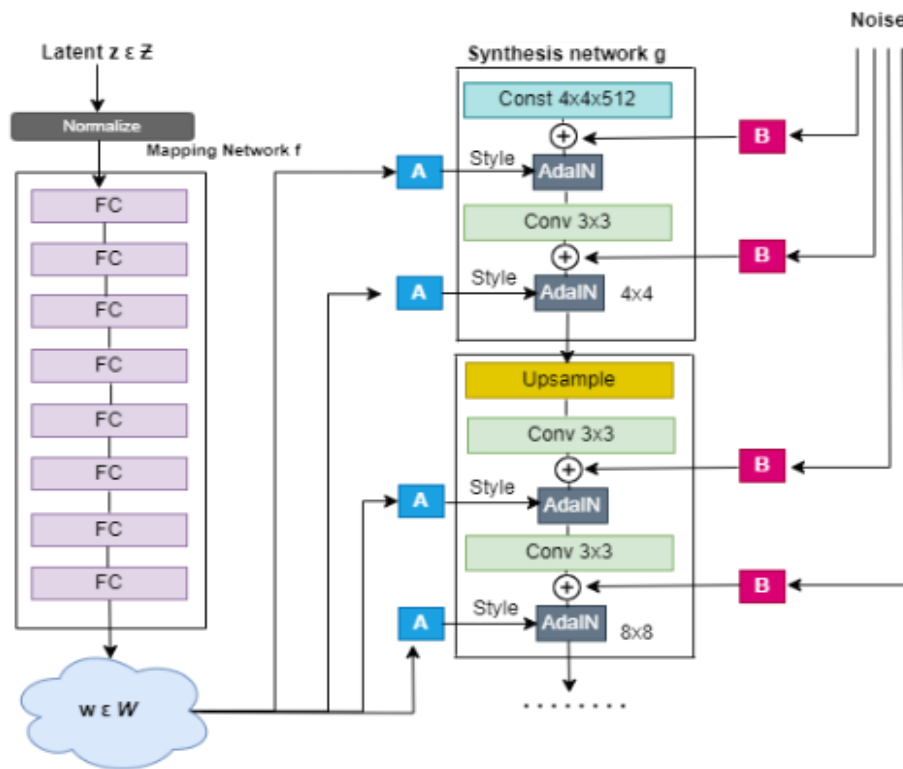


Figure 4. StyleGAN architecture

2.3.5. Progressive GAN

As opposed to conventional GANs that train the generator and discriminator on full-resolution images simultaneously, the progressive GANs [28] start with low-resolution image input as shown in Figure 5. This method involves slowly turning from basic model (generator and discriminator) towards a larger and more complex one while training, which leads to the generation of images with growing quality, resolution and detail. This approach helps to stabilize and standardize training formation in a step-by-step manner, starting from the simplest and ending with complicated ones. In progressive GANs, multiple generators and discriminators are used for various resolutions, which have equations of the same format that is used in a standard GAN but adjusted to the corresponding progressive architecture.

A. Generator

A random noise vector z with dimensions (batch size, latent size) is provided as input.

$$G(z) = x \tag{19}$$

Where x is the resultant image, the adversarial loss is one of several losses that the generator is trained to reduce; others include the feature matching loss and the pixel-wise reconstruction loss.

B. Discriminator

Input: an image of particular channels, height, and width, output: a scalar between 0 and 1 representing the likelihood that the input image is real.

$$D(x) = y \tag{20}$$

Where y is a scalar between 0 and 1, the discriminator is optimized through training to maximize the adversarial loss, defined as the discrepancy between the predicted probability and the correct label (0 or 1), and minimize other losses, such as the feature matching loss.

Each of the GANs discussed in this space has its benefits and drawbacks; the choice of model required for a given scenario will depend on other, context-specific factors. For instance, Vanilla GANs are easy to implement and deliver high-quality image but not of more variety because mode collapse stops the generator from producing a lot of images. Thanks to using conditional GANs it is possible to generate images by observing certain standards. The limitation of conditional GANs is that their training becomes more complex and furthermore requires a higher amount of data. DCGAN uses deeper convolution networks, and in comparison, with vanilla GAN, it produces images that are more realistic due to the higher quality of the textures and details. Yet they still may suffer from the impact of mode collapse. Concerning progressive GANs, these GAN models result in higher-resolution and more realistic images than any other.

GAN model. Unfortunately, they require far more time and computational power to train. The most important achievement of StyleGAN is the ability to generate images with the highest degree of realism and characteristics that make each of them stand out from others. However, they require vast amounts of training data and processing resources.

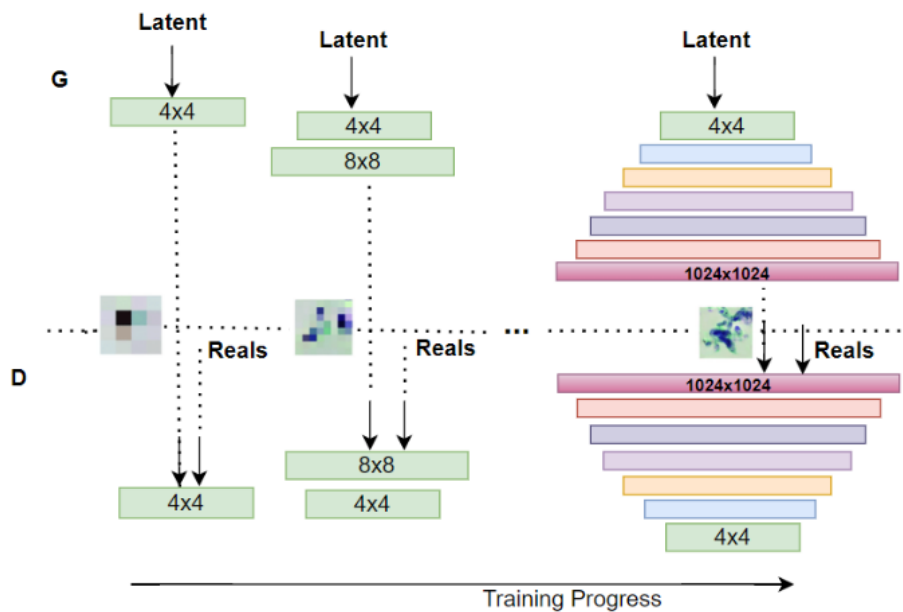


Figure 5. Progressive GAN operation

3. MELANOMA IMAGE SYNTHESIS USING GANS

This section provides in-depth analysis of the various methods selected from the literature search as per the strategy discussed in section 2 for synthesizing melanoma images using GAN to improve classification accuracy. These days, GANs are being used effectively in skin cancer diagnostic systems [29]. A method called catWGAN was proposed that use categorical generative adversarial networks in an unsupervised and semi-supervised manner [30]. To eliminate background interference and any segmentation algorithm confusion, ground truth segmentation maps were used to segment the skin lesions. On ISIC 2016 dataset an average precision score of 0.424 with only 140 labeled images is achieved. The proposed method was only tested on single dataset in this paper, which may reduce the generalization of the findings to other

datasets. The proposed method has high computational requirements, which may restrict its use in real-world scenarios.

Baur *et al.* [31], compared the performance of DCGAN, the LAPGAN, and the PGAN, in which PGAN with progressive growth outperformed the other two. The authors utilized the ISIC2018 dataset of benign and malignant skin lesions for their experiments. They successfully synthesized highly realistic dermoscopic images of skin lesions using GANs at a high resolution of 256×256 pixels. The results demonstrated images of high quality and realism, and even expert dermatologists find it hard to distinguish between real and fake images. Here, the user study conducted to evaluate the realism of the synthesized images was limited to a small number of participants and could benefit from a larger and more diverse sample size.

Another method for skin lesion classification using GAN based data augmentation is proposed by Rashid *et al.* [32]. They compared the performance of a deep learning model that uses GAN for data augmentation with two other popular deep learning models, ResNet and DenseNet, for skin lesion classification. Using the identical training data as the GAN-based model, the ResNet and DenseNet models were fine-tuned; however, synthetic GAN-generated images were left out for augmentation. Conventional augmentation techniques such as rotation, flipping, and adding noise were used for all models during training. The balance accuracy score was used as the metric for performance comparison, which considers class imbalance in the data. The results showed that the GAN-based model outperformed both ResNet and DenseNet, even though these models are deeper than the GAN-based classifier.

Pollastri *et al.* [33], presented an innovative approach using GAN to augment data for skin lesion segmentation, which is a key initial phase in the automated process of detecting melanoma. The effect of the diversity and quality of synthetic images on the method's performance was examined by modifying two well-known GANs: aDCGAN and a Laplacian GAN. The baseline CDNN architecture was implemented that has the highest score in the ISIC2017 challenge in order to map the input image to a map of posterior probability. The images were resized to 192×256, and with the HSV and L* channels the original RGB channels were augmented. The proposed method is evaluated by populating training data with synthetic data to train CNN and overall accuracy was measured for CDNN for melanoma lesion segmentation. The proposed methodology outperforms the DCGAN in terms of diversity and quality of synthetic images.

While in the work proposed by [34], DermGAN is introduced, which is an adaptation of the popular Pix2Pix architecture. Opinion of several board-certified dermatologists was taken to find the actual skin condition in each case. This was done to differentiate between 26 common skin conditions and "other" category. Human turing test and objective GAN evaluation metrics were used to evaluate the generated images. The synthetic images were used along with original ones for training a classifier, and the model's performance was compared to the baseline model. Using the synthetic images, the model worked about the same overall as the baseline model, but it did better in some rare but malignant conditions.

The enhancements in the pix2pixHD GAN is proposed by [35] that combines the coarse-to-fine generator, a multi-scale discriminator architecture, and a robust adversarial learning objective function. These improvements allow the network to work with high-resolution samples. The authors only use the global generator from pix2pixHD to generate 1,024×512 resolution images. The synthetic images have features that aren't in the real images. This makes the classification network better by an average of 1.3% percentage points and keeps it more stable. All of the differences were important (95% confidence level) (p-value <0.05). The p-value of a paired samples t-test is given to make sure that the comparison between "real+instance+PGAN" and other data is statistically significant. At a 95% confidence level (p-value 0.05), all of the differences were statistically important.

The work proposed by [36] modifies the structure of style control and noise input in the original generator. Various metrics such as accuracy, specificity, sensitivity, average precision, and balanced multiclass accuracy achieved were 95.2%, 83.2%, 74.3%, 96.6%, and 83.1% respectively on ISIC 2018 challenge dataset. However, the paper does not provide a detailed analysis of the generated images by the proposed skin lesion style-based GANs model. It would be interesting in future work to see how the generated images compare to real skin lesion images and whether they have any artifacts or biases.

Another work proposed by [37] used PGAN for melanoma image synthesis and populated the ISIC archive dataset with the generated images to build a classifier model for melanoma classification. The proposed model achieved the highest area under the curve (AUC) of 84.7 on the populated data set, while the AUC on the real dataset is 82.8. Zhao *et al.* [38] suggested a new skin lesion augmentation style-based GAN called SLA-StyleGAN that changes how style control and noise input work in the original generator network. It reconstructs the discriminator so that high-quality skin lesion images can be generated quickly. The proposed framework uses DenseNet201 as the backbone network for skin lesion classification. It also used an enhanced loss function that mixes focal loss with weighted cross-entropy and A-Softmax loss to fix the problems caused by sample imbalance in the data and make the weights of the different types of samples

more equal. This can help the classification model work better. On the ISIC2019 dataset, the suggested method gets a balanced multiclass accuracy (BMA) of 93.64%, which is better than other cutting-edge methods.

A new framework is implemented called TED-GAN by [39] that incorporated an encoder-decoder that was trained obtaining the newly noise vector with the image manifold's data, and the GAN sampled the intake from this noise vector which is full informative and generate the skin lesion images. The variety of the images made was better with another GAN that had an extra classifier that took samples from a heavy-tailed student t-distribution as a replacement of random-noise distribution of Gaussian. The authors improved classification performance rising from 66% average accuracy to 92.5% on the skin lesion classification task. To recreate super-resolution skin lesion images from low-resolution ones, [40] also suggests a brand-new cascade ensemble super-resolution GAN (CESR-GAN) technique. The authors designed a novel feature-based measurement loss function to obtain more details and generate higher-quality images. It achieved a peak signal-to-noise ratio (PSNR) of 34.23 dB and a structural similarity index (SSIM) of 0.94 on the ISIC dataset, a PSNR of 33.12 dB, and an SSIM of 0.93 on the PH 2 dataset.

A self-attention PGGAN (SPGGAN) to generate fine-grained 256×256 skin lesion images for CNN-based melanoma diagnosis is proposed by [41]. They used the two-timescale update rule (TTUR) to improve stability of SPGGAN while generating melanoma images. The results of the paper show that the proposed SPGGAN and TTUR can lead to statistically significant improvements in the sensitivity (recall) over non-augmented and augmented counterparts, with classical data augmentation, for all classes and specifically for melanoma class. Further, a novel method for melanoma skin cancer using an advanced deep neural network and adversarial training to achieve better accuracy even with a small amount of data [42]. The input image's depth, gradient, and shade are amplified to extract useful information. The the gradients of the loss function were used for the input image to generate a new adversarial image that maximizes the loss for the input image. and generated images were used in the training of classifier model. Training accuracy of 74.65% with 0.75 loss value is achieved and during validation, accuracy of 74.76% at a loss of 0.5865 is achieved. A state-of-the-art performance using ResNet101 architecture with adversarial training was achieved with accuracy of 84.77% for melanoma diagnosis.

In another novel work [43], StyleGAN2 was used to generate synthetic melanoma images and additionally, a U-Net model was utilised for creating masks for regions of interest in the images, to focus on the image's relevant features. Finally, a CNN called EfficientAttentionNet was trained with the mask-based attention mechanism for classification of lesions. A classification accuracy of 96% with GAN-generated synthetic images was achieved using EfficientAttentionNet. Al-Rasheed *et al.* [44] fine-tuned VGG16, ResNet50, and ResNet101 for multi class classification of skin cancers. In order to address class imbalance problems that could result in model overfitting, extra photos were added to the training set using conditional GAN and traditional data augmentation. The model's output is contrasted with models trained on the imbalanced dataset. By fine-tuning and training it on both balanced and unbalanced datasets, an ensemble of transfer learning models is created. With suitable data augmentation, an accuracy of 9.35% for the ensemble one and 92% for VGG16, 92% for ResNet50, and 92.25% for ResNet101 is obtained.

A novel GAN model with four transposed convolutional layers and a leaky rectified linear unit (LReLU) as an activation function in generator was proposed [45] to avoid vanishing gradients and dying neurons problems in ReLU. The generator network also includes a convolutional layer with the sigmoid activation function and the discriminant consists of five convolutional layers with LReLU as an activation function, followed by a dense layer for classification with sigmoid function. The U-Net architecture, a type of CNN, is used with an attention mechanism to improve image segmentation performance. For classification, the pre-trained EfficientNetV2S and Efficient-NetV2B2 networks were used. On ISIC dataset it achieved an accuracy of 96%, recall of 0.95, precision of 0.88, and F1-score of 0.91, outperforming the state-of-the-art models.

Table 3 (in Appendix) summarizes all GAN-based melanoma image synthesis and detection systems discussed so far, along with the diagnosed skin cancer type, classifier, dataset, and the result obtained. In the table AUC is area under curve, AC: accuracy, SWD: sliced wasserstein distances, FID: Fréchet Inception distance, IS: Inception score, SS: sensitivity, SP: specificity, RL: recall, FS: F1-score, PR: precision, SSIM: structural similarity index measure, FSIM: feature similarity index measure, and PSNR: peak signal-to-noise ratio. The classification and image synthesis results using GAN of different studies are shown below in Figures 6 and 7.

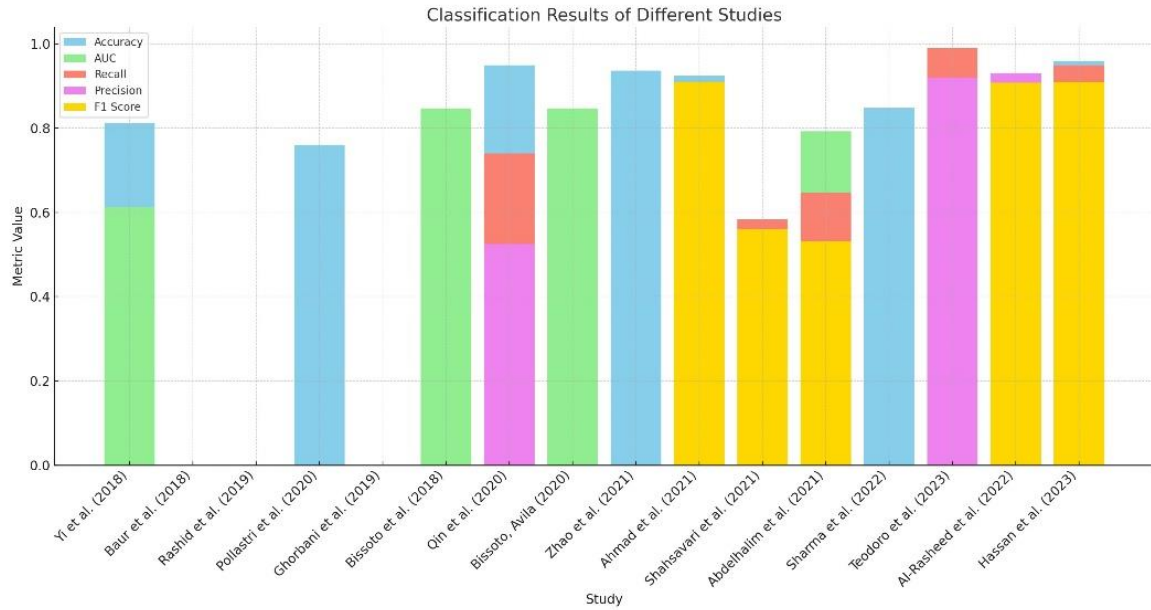


Figure 6. Combined classification results of different studies

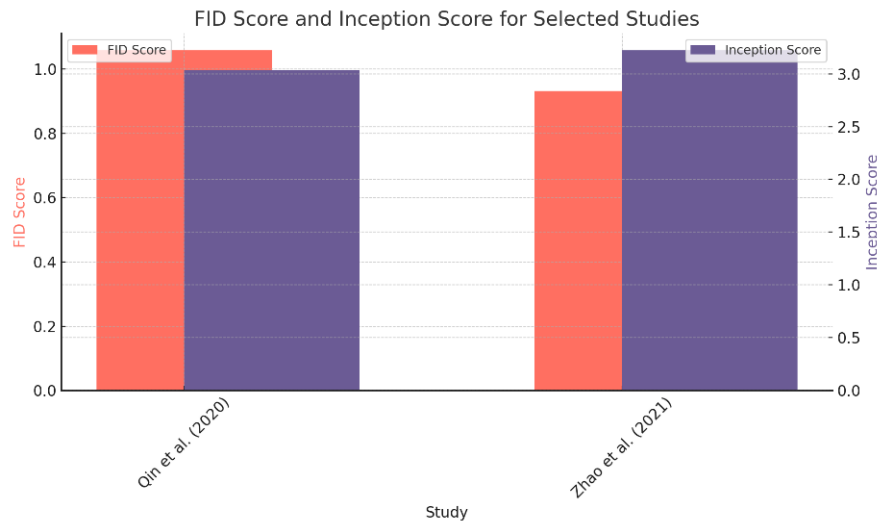


Figure 7. FID and IS score for selected studies

4. SKIN LESION DATASETS

One of the challenges in developing and evaluating a classification system is the availability of a solid and reliable collection of dermoscopic images [46]. In the case of melanoma, the most deadly form of skin cancer, there is a need for diverse and comprehensive datasets for training and validating artificial neural networks. Traditionally, the datasets available for automated skin cancer diagnosis have primarily focused on melanocytic lesions, such as nevi or melanoma [47].

Furthermore, the datasets have not adequately represented the wide range of patients who commonly experience non-melanocytic lesions. To address these limitations and improve the performance of melanoma detection systems, researchers have turned to GANs for melanoma image synthesis. Using a provided dataset as a source, GANs are a type of machine learning model that can produce artificial data. For training and testing GANs for melanoma image synthesis, a standard, trustworthy collection of dermoscopic pictures is essential. This section discusses about various real-world datasets of skin lesions.

4.1. HAM10000

HAM10000 [48] is a human-vs.-machine dataset consisting of 10,000 training images. It's among the most up-to-date skin lesions datasets we can find online, and it fixes the issue of data homogeneity. Cliff Rosendahl's skin cancer practice in Queensland, Australia, and the Dermatology Department at the Medical University of Vienna, Austria, contributed 10,015 dermoscopic images to the final HAM10000 dataset. This dataset has 6,705 images of melanomas, 1,113 images of melanocytic nevi, 115 images of dermatofibromas, and 327 images of AK. There are also 1,099 images of benign keratoses, 115 dermatofibromas, and 115 images of AK.

4.2. ISIC archive

The ISIC repository [49] houses a variety of skin lesion datasets, comprising of ISIC2016 released as part of the round 3 contest through the challenge organized at ISIC-ISBI 2016. It is composed from 900 training and 379 test images, whereabout the proportion of melanoma stands in approximate completeness of 30%. In an attempt to accelerate automated skin cancer diagnosis, the perpetual challenges organized by ISIC each year. The ISIC2017 dataset contains melanomas, seborrheic keratoses and benign nevi and in all 2,500 images. ISIC2018 and ISIC2019 enlarged the image library to larger extent including an outlier class for testing systems of diagnosis which had over 25,000 images of eight lesion types.

4.3. PH2

The dermoscopy images in the PH2 dataset were taken at the Pedro Hispano Hospital Dermatology Center in Portugal [50]. The dataset has 200 dermoscopy images. There are 80 common nevi images, 80 atypical nevi images, and 40 melanoma skin cancer images. The lesion images in this dataset have medical annotations, including evaluation of several dermoscopic criteria, histological and clinical diagnosis, and medical segmentation of pigmented skin lesions. The evaluation was done using dermoscopy criteria like streaks, colors, regression areas, pigment network, and blue-white veil globules.

4.4. Dermofit

The dermofit [51] provides a variety of images of skin lesions along with their metadata, which can help advance the field of computer-assisted diagnosis. Various types of skin cancer and other lesions are represented in the dataset, including basal cell carcinoma, malignant melanoma, melanocytic nevus, seborrheic keratosis, and squamous cell carcinoma. The dermofit dataset gives researchers a sizable sample size for training and evaluating their models with 1,000 dermoscopic images. This dataset is publicly available for academic use.

4.5. DermIS

The acronym "DermIS" stands for "dermatology information system". The University of Erlangen's Dermatology Department and the University of Heidelberg's Clinical Social Medicine Department worked together to compile this dataset. There are 6,588 images in total. Recently, this data set was split in half to create a DOIA (dermatology online image atlas) and a PeDOIA (pediatric dermatology online image atlas). About 600 skin lesions are represented in the DOIA's 3000 images. Case reports, differential diagnoses, and other details on nearly every conceivable skin disorder are included alongside dermoscopy images.

5. RESULTS AND DISCUSSION

This section outlines and examines the main discoveries resulting from the comparative investigation of different GAN approaches in the realm of melanoma picture creation.

5.1. Comparative analysis of GAN strategies

We thoroughly examined various GAN models utilized in generating melanoma images in our review. The table provides a concise overview of the main features, approaches, and results of these studies. The work conducted by Yi, Walia, and Babyn (2018) [30] showcased catWGAN's favourable outcomes in unsupervised and semi-supervised scenarios, with AUC values of 0.613 and 0.690, respectively. According to Baur, Albarqouni, and Navab (2018) [31], PGAN produced more realistic melanoma images than DCGAN and LAPGAN, as seen by reduced SWDs when compared to genuine images. Rashid, Tanveer, and Aqeel Khan (2019) [32] discovered that traditional GAN performed better than DenseNet and ResNet-50 in categorizing different skin lesions. The experiments demonstrate improvements in realism and diagnostic precision with the use of GANs. Nevertheless, issues like as data quality, generalization across various skin types, and computational efficiency continue to exist.

5.2. Data quality and annotation

The importance of precise and well-annotated data for training GANs to create synthetic melanoma images is emphasised by our results. The difficulties in acquiring a varied and thoroughly annotated dataset are clear, especially due to the subjective and time-intensive process of manual annotation. This emphasises the necessity for research in creating standardised annotation protocols and improving annotation methods for greater efficiency.

5.3. Addressing class imbalance and rare subtypes

The study uncovers substantial class imbalance problems in existing datasets, where rare and aggressive melanoma subtypes are not adequately represented. GANs face a challenge in creating synthetic images that accurately represent all types of melanoma.

5.4. Authenticity and variety of produced images

Our analysis shows that although GANs can generate images resembling real melanoma lesions, there is a requirement for enhancing the authenticity and variety of these images. Improving GAN structures and training techniques, while integrating sophisticated loss functions, may elevate the quality of generated images.

5.5. Generalisation and transferability

Generalising and transferring GAN models between various populations, skin tones, and imaging conditions is still difficult. Using these models in real clinical settings can be challenging, indicating a necessity for further research to improve GANs' flexibility.

5.6. Interpretability and explainability

Interpretability and explainability are essential for the clinical approval of GAN-generated images. Our research shows that existing GAN models do not have enough interpretability, highlighting the need to create methods for improving the understanding of GAN features and decisions.

5.7. Ethical considerations

Ethical considerations, especially regarding patient privacy and data protection, are of utmost importance. Our review indicates an increasing demand for research that concentrates on privacy-preserving methods in the clinical utilisation of GAN-generated images.

5.8. Data representation

Most images in current datasets mainly show individuals with light skin tones, indicating a notable lack of diversity in representation. It is necessary to train GAN models on a wider range of datasets to guarantee accurate skin cancer detection on various skin tones.

5.9. Diversity of skin lesions

We observed little variation among various categories of skin lesions in medical imaging. The minimal variation between different classes makes it difficult to distinguish melanoma from other skin conditions, which poses a challenge for creating synthetic images.

5.10. Computing needs

Generating synthetic images using GANs requires powerful GPUs and extensive RAM, which can be challenging due to resource constraints in research settings.

6. CONCLUSION

The review paper thoroughly analyzed GANs in melanoma image generation, emphasizing its visual quality, diversity, and clinical significance. We have analyzed many GAN models such as DCGAN, WGAN-GP, and CycleGAN to determine their specific advantages and drawbacks, especially in duplicating the intricate structure of melanoma lesions accurately. The findings show substantial development in the sector but also highlight the need for further advancements, aligning with our objective. Although models such as PGAN demonstrate potential in creating high-quality artificial pictures, there are still obstacles in guaranteeing data variety and handling the computational requirements for real-world clinical use. The constraints in the generalizability and applicability of existing GAN models highlight a crucial topic for further investigation.

Our evaluation, which includes assessing visual quality and diversity, indicates that GAN-generated melanoma images have the potential to greatly improve the precision and dependability of melanoma detection systems. Synthetic images have been successfully incorporated into current classification methods, resulting in significant progress in dermatological imaging. Nevertheless, this advancement highlights the need for ongoing research, specifically emphasizing transfer learning methods and the utilization of pre-trained GAN models on large datasets from relevant areas such as general medical imaging. These methods could significantly enhance the generalization capacities of GAN-generated melanoma images, making them more adaptable and successful in different therapeutic situations.

Our findings have significant consequences for dermatological research and the medical community. They highlight the significant impact GANs can have on enhancing early melanoma identification and therapy, potentially resulting in improved patient outcomes. Advancements in GAN technology can be utilized in dermatology and other medical imaging fields, sparking a new era of innovation in healthcare. Our study provides opportunities for the scientific community to investigate sophisticated machine learning methods in medical diagnostics. It highlights the importance of interdisciplinary collaboration to connect technology with clinical practice. Our analysis highlights the significant progress in GANs for generating melanoma images and suggests future research to address current obstacles. The technology has promising possibilities for improving healthcare delivery and patient care in clinical settings and beyond.

7. FUTURE WORK

Future research should prioritize the development of novel structures, loss functions, and training techniques for GANs to enhance the quality and diversity of generated images. These developments are essential to ensure that GAN-generated images can accurately replicate the wide variety of melanoma presentations observed in real-world clinical environments. Future research also relies significantly on collaborating with dermatologists and medical specialists. Their clinical feedback on the assessment of GAN-generated melanoma images is crucial to assure the technical quality and clinical significance of the images. As we advance in using GAN technology in clinical settings, it is crucial to focus on ethical issues, especially regarding patient data privacy and permission. Creating frameworks to ensure the responsible and ethical utilization of GAN-generated images in clinical environments is crucial for their effective incorporation into medical procedures.

Ultimately, our research sets the stage for future progress in GAN-generated melanoma picture creation. This discovery has the potential to improve patient outcomes in dermatology by boosting melanoma identification and therapy. By focusing on the specified future research areas, we can expect substantial advancements in the field, resulting in more precise, varied, and clinically significant tools for diagnosing and treating melanoma.

ACKNOWLEDGEMENTS

This project is sponsored by Prince Sattam Bin Abdulaziz University (PSAU) as part of funding for its SDG Roadmap Research Funding Programme project number PSAU-2023/SDG/70.

APPENDIX

Table 3. Summary of GAN-based melanoma image synthesis and classification methods

Reference	Diagnostic framework	Description	Results
(Yi, Walia, and Babyn 2018 [30])	Classes: benign and melanoma, model(s): catWGAN dataset(s): ISIC2016, PH2	The proposed model can learn a feature representation that outperforms the denoising autoencoder and simple hand-crafted features.	catWGAN-unsupervised AUC→0.613, AC→0.812 catWGAN-semisupervised AUC→0.69, AC→0.810.
(Baur, Albarqouni, and Navab 2018 [31])	Classes: benign and melanoma, model(s): PGAN, DCGAN, LAPGAN dataset(s): ISIC2018	PGAN generated realistic-looking melanoma images and outperformed DCGAN and LAPGAN	SWDs: PGAN vs real→20.0197, DCGAN vs real→94.71508 LAPGAN vs real→96.68380.
(Rashid, Tanveer, and Aqeel Khan 2019 [32])	Classes actinic keratosis/basal cell carcinoma/benign eratosis/dermatofibroma/melano ma/melanocytic nevus/vascular lesion, model(s): traditional GAN, DenseNet, ResNet-50 Dataset(s): ISIC2018	Traditional GAN generated realistic-looking images and was also used as a classifier for various skin lesion images. It outperformed DenseNet and ResNet-50	AC: GANs 0.861 DenseNet→0.815 ResNet→50 0.792.

Table 3. Summary of GAN-based melanoma image synthesis and classification methods (*Continue...*)

Reference	Diagnostic framework	Description	Results
(Pollastri <i>et al.</i> 2020 [33])	Classes: benign and melanoma, model(s): modified DCGAN and LAPGAN, CDNN for classification, dataset(s): ISIC2017	Data augmentation with DCGAN and LAPGAN boosts the classification performance of various CDNNs	AC→0.76 on augmented data using LAPGAN.
(Ghorbani <i>et al.</i> 2019 [34])	Classes: actinic keratosis/basal cell carcinoma/benign eratosis/dermatofibroma/melano ma/melanocytic nevus/vascular lesion, model(s): DermGAN and MobileNet, dataset(s): teledermatology service US 2010-2018	Data augmentation with DermGAN and classification of eught skin lesions with MobileNet	FID: real data→83.60, DermGAN→122.40, AC→0.75.
(Bissoto <i>et al.</i> 2018[35])	Classes :benign and melanoma, model(s): DCGAN, conditional version of PGAN and version of pix2pixHD GAN, Inception-v4, dataset(s): ISIC 2017, ISIC archive, dermot image library, PH2 dataset, interactive atlas of dermoscopy	Using a multi-scale discriminator architecture, a coarse-to-fine generator, and resilient adversarial learning goal function resolution samples, a pix2pixHD GAN is employed.	AUC→84.7 using PGAN.
(Qin <i>et al.</i> 2020 [36])	Classes: actinic keratosis/basal cell carcinoma/benign eratosis/dermatofibroma/melano ma/melanocytic nevus/vascular lesion, model(s): skin lesion style-based GAN, ResNet-50, dataset(s): ISIC2018	By modifying the style control and noise input structure of the generator, the suggested model effectively synthesizes high-quality skin lesion images by adjusting the generator and discriminator.	SL-StyleGAN IS→3.037, FID→1.059 PR→0.525, RL→0.220 Classification result-AC→0.95, SS→0.74, SP→0.96.
(Bissoto and Avila 2020 [37])	Classes: benign and melanoma, model(s): PGAN, dataset(s): ISIC archive	The proposed PGAN generates realistic looking melanoma images and the classifier is trained with and without synthetic images to analyse its performance.	AUC→84.7 achieved with orginal plus synthetic images.
(Zhao <i>et al.</i> 2021[38])	Classes: benign and melanoma, model(s): SLA-StyleGAN and DenseNet201, dataset(s): ISIC 2019	(SLA-StyleGAN) is proposed to generate melanoma skin lesion images and give as input to DenseNet201 pre-trained network to do transfer learning for melanoma classification. A novel loss function with focal loss and weighted cross-entropy and A-Softmax loss to solve data imbalance problem	DenseNet201: AC→0.9364 SLA-StyleGAN: IS→3.224, FID→0.932.
(Ahmad <i>et al.</i> 2021[39])	Classes: melanoma, melanomic neves, basal cell carcinoma, benign keratosis, model(s): VAE, GAN1, GAN2, dataset(s): HAM10000	The decoder-encoder network give noise as input to GAN 1 which generate sythetic images of given classes and these images along with real images are given to GAN 2 which take student t-distribution as input instead of gaussian, for classification purpose.	Classification results: SS→89%, SP→94%, FS→0.91 and AC→92.5%.
(Shahsavari, Ranjbari, and Khatibi 2021[40])	Classes: basal cell carcinoma, malignant melanoma, nevus, and seborrheic keratoses, model(s): CESR-GAN, U-Net, dataset(s): ISIC archive and PH2 dataset	The low-resolution images are converted to high resolution using the proposed CESR-GAN and then they are classified using U-Net architecture.	Image conversion results: SSIM→0.94, FSIM→0.98, PSNR→42 Classification results: AC→0.55, PR→0.538, RL→0.583, FS→0.560.
(Abdelhalim, Mohamed, and Mahdy 2021[41])	Classes: benign and melanoma, model(s): SPGGANs and ResNet 18, dataset(s): HAM10000	SPGGAN generated fine-grained 256×256 skin lesion images . TTUR is applied to SPGGAN to improve stability while generating images. ResNet18 is used for finally classifying melanoma.	Qualitative results: P-value of the T-test. 68.1 (GAN-train) and 60.8 (GAN-test) quantitative results: AC→66.1, AUC→79.3, precision→50.1, RL→64.7, FS→53.1
(Sharma <i>et al.</i> 2022 [42])	Classes: benign and melanoma, model(s): GAN, VGG16, VGG19, DenseNet121 and ResNet101, dataset(s): HAM10000	Adversarial images are generated with an FGSM attack using GAN, and then classification is done using various CNN mentioned.	Classification: AC→84.77%.

Table 3. Summary of GAN-based melanoma image synthesis and classification methods (Continue...)





Reference	Diagnostic framework	Description	Results
(Teodoro <i>et al.</i> 2023 [43])	Classes: benign and melanoma, model(s): StyleGAN, U-Net and EfficientNet, dataset(s): HAM10000, ISIC2020, MKS and BCN20000	Melanoma images are input to GAN, and synthetic images are generated. U-Net is used to generate clipping masks from the balanced dataset. EfficientNet was trained using these masks as a RoI-based attention mechanism that improved the classification performance	Segmentation results by U-Net: AC→0.89, RL→0.83, PR→0.89, IoU→0.76, DC→0.85, classification results: with GAN images AC→0.96, RL→0.99, PR→0.92, AUC→0.96 EfficientAttentionNet: AC→0.97, RL→0.99, PR→0.94, AUC→0.97
(Al-Rasheed <i>et al.</i> 2022 [44])	Classes: dermatofibroma, melanoma, nevus, and vascular cancer, model(s): CGAN, VGG16, ResNet50, and ResNet101, dataset(s): HAM10000	First, data augmentation is done using traditional techniques and with GAN also. The various CNN models, as mentioned, are used to analyze the classification performance. Also, the ensemble of all models is used for classification.	classification results: ResNet101 AC→92.25, RL→85.40 R→90.63 FS→87.79 Ensemble model AC→93.5, RL→88.98 PR→93.20, FS→90.82.
(Hassan, Mahar, and Fouad 2023 [45])	Classes: benign and melanoma, model(s): novel GAN, U-Net and EfficientNetV2S, Efficient-NetV2B2, dataset(s): ISIC archive	The proposed GAN generates the synthetic images which are added to training data. Segmentation masks are generated with U-Net and are given as input to EfficientNetV2S, Efficient-NetV2B2 for classification.	Classification results: AC→96, FS→0.91, RL→0.95, PR→0.88.

REFERENCES





- [1] R. Ashraf *et al.*, “Region-of-interest based transfer learning assisted framework for skin cancer detection,” *IEEE Access*, vol. 8, pp. 147858–147871, 2020, doi: 10.1109/ACCESS.2020.3014701.
- [2] A. L. Byrd, Y. Belkaid, and J. A. Segre, “The human skin microbiome,” *Nature Reviews Microbiology*, vol. 16, no. 3, pp. 143–155, Jan. 2018, doi: 10.1038/nrmicro.2017.157.
- [3] Skin Cancer Foundation, A 501(c)(3) organization [EIN: 13-2948778] “What is melanoma?., <https://www.skincancer.org/skin-cancer-information/melanoma/> Accessed on (29-03-2024)
- [4] “Cancer facts and statistics,” *American Cancer Society*, 2023. <https://www.cancer.org/research/cancer-facts-statistics.html> (accessed Jun. 04, 2023).
- [5] A. Blundo, A. Cignoni, T. Banfi, and G. Ciuti, “Comparative analysis of diagnostic techniques for melanoma detection: a systematic review of diagnostic test accuracy studies and meta-analysis,” *Frontiers in Medicine*, vol. 8, Apr. 2021, doi: 10.3389/fmed.2021.637069.
- [6] V. Dick, C. Sinz, M. Mittlböck, H. Kittler, and P. Tschandl, “Accuracy of computer-aided diagnosis of melanoma: a meta-analysis,” *JAMA Dermatology*, vol. 155, no. 11, pp. 1291–1299, Nov. 2019, doi: 10.1001/jamadermatol.2019.1375.
- [7] C. Cantisani *et al.*, “Melanoma detection by non-specialists: an untapped potential for triage?,” *Diagnostics*, vol. 12, no. 11, p. 2821, Nov. 2022, doi: 10.3390/diagnostics12112821.
- [8] S. Duma, “Dermoscopy of pigmented skin lesions,” in *European Handbook of Dermatological Treatments*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1167–1177.
- [9] O. Yélamos, L. Mary Diem, R. P. Braun, K. K. French, and A. A. Marghoob, “Dermoscopy for dermatopathologists,” in *Pathology of Melanocytic Tumors*, Elsevier, 2019, pp. 331–347.
- [10] S. W. Menzies *et al.*, “The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma,” *Archives of Dermatology*, vol. 141, no. 11, pp. 1388–1396, Nov. 2005, doi: 10.1001/archderm.141.11.1388.
- [11] M. J. Lin, V. Mar, C. McLean, R. Wolfe, and J. W. Kelly, “Diagnostic accuracy of malignant melanoma according to subtype,” *Australasian Journal of Dermatology*, vol. 55, no. 1, pp. 35–42, Nov. 2014, doi: 10.1111/ajd.12121.
- [12] A. C. Halpern and J. A. Lieb, “Early melanoma diagnosis: a success story that leaves room for improvement,” *Current Opinion in Oncology*, vol. 19, no. 2, pp. 109–115, Mar. 2007, doi: 10.1097/CCO.0b013e32801497b2.
- [13] C. A. Bauman, P. Emary, T. Damen, and H. Dixon, “Melanoma in situ: a case report from the patient’s perspective,” *Journal of the Canadian Chiropractic Association*, vol. 62, no. 1, pp. 56–61, 2018.
- [14] L. Alzubaidi *et al.*, “Towards a better understanding of transfer learning for medical imaging: a case study,” *Applied Sciences (Switzerland)*, vol. 10, no. 13, p. 4523, Jun. 2020, doi: 10.3390/app10134523.
- [15] X. Li, Y. Jiang, J. J. Rodriguez-Andina, H. Luo, S. Yin, and O. Kaynak, “When medical images meet generative adversarial network: recent development and research opportunities,” *Discover Artificial Intelligence*, vol. 1, no. 1, Sep. 2021, doi: 10.1007/s44163-021-00006-0.
- [16] Y. Skandarani, P. M. Jodoin, and A. Lalonde, “GANs for medical image synthesis: an empirical study,” *Journal of Imaging*, vol. 9, no. 3, p. 69, Mar. 2023, doi: 10.3390/jimaging9030069.
- [17] L. Fetty *et al.*, “Latent space manipulation for high-resolution medical image synthesis via the StyleGAN,” *Zeitschrift für Medizinische Physik*, vol. 30, no. 4, pp. 305–314, Nov. 2020, doi: 10.1016/j.zemedi.2020.05.001.
- [18] S. Xun *et al.*, “Generative adversarial networks in medical image segmentation: a review,” *Computers in Biology and Medicine*, vol. 140, p. 105063, Jan. 2022, doi: 10.1016/j.combiomed.2021.105063.
- [19] M. Gong, S. Chen, Q. Chen, Y. Zeng, and Y. Zhang, “Generative adversarial networks in medical image processing,” *Current Pharmaceutical Design*, vol. 27, no. 15, pp. 1856–1868, Apr. 2020, doi: 10.2174/1381612826666201125110710.
- [20] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, and G. Lasa, “How-to conduct a systematic literature review: a quick guide for computer science research,” *MethodsX*, vol. 9, p. 101895, 2022, doi: 10.1016/j.mex.2022.101895.
- [21] H. Snyder, “Literature review as a research methodology: an overview and guidelines,” *Journal of Business Research*, vol. 104, pp. 333–339, Nov. 2019, doi: 10.1016/j.jbusres.2019.07.039.
- [22] J. Brownlee, “A gentle introduction to generative adversarial networks (GANs),” *Machine Learning Mastery*, vol. 17, pp. 1–31, 2019.

- [23] H. Huang, P. S. Yu, and C. Wang, "An introduction to image synthesis with generative adversarial nets," *arXiv preprint*, p. 17, 2018, doi: 10.48550/arXiv.1803.04469.
- [24] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.
- [25] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint*, p. 7, 2014.
- [26] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 2016, p. 16.
- [27] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2019, vol. 2019-June, pp. 4396–4405, doi: 10.1109/CVPR.2019.00453.
- [28] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," in *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, 2018, p. 26.
- [29] M. Dildar *et al.*, "Skin cancer detection: a review using deep learning techniques," *International Journal of Environmental Research and Public Health*, vol. 18, no. 10, p. 5479, May 2021, doi: 10.3390/ijerph18105479.
- [30] X. Yi, E. Walia, and P. Babyn, "Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by wasserstein distance for dermoscopy image classification," *arXiv preprint*, p. 10, 2018, [Online]. Available: <http://arxiv.org/abs/1804.03700>.
- [31] C. Baur, S. Albarqouni, and N. Navab, "Generating highly realistic images of skin lesions with GANs," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11041 LNCS, Springer International Publishing, 2018, pp. 260–267.
- [32] H. Rashid, M. A. Tanveer, and H. Aqeel Khan, "Skin lesion classification using GAN based data augmentation," in *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, Jul. 2019, vol. 2019, pp. 916–919, doi: 10.1109/EMBC.2019.8857905.
- [33] F. Pollastro, F. Bolelli, R. Paredes, and C. Grana, "Augmenting data with GANs to segment melanoma skin lesions," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 15575–15592, May 2020, doi: 10.1007/s11042-019-7717-y.
- [34] A. Ghorbani, V. Natarajan, D. Coz, and Y. Liu, "DermGAN: synthetic generation of clinical skin images with pathology," in *Proceedings of Machine Learning Research*, 2019, vol. 116, pp. 155–170.
- [35] A. Bissoto, F. Perez, E. Valle, and S. Avila, "Skin lesion synthesis with generative adversarial networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11041 LNCS, Springer International Publishing, 2018, pp. 294–302.
- [36] Z. Qin, Z. Liu, P. Zhu, and Y. Xue, "A GAN-based image synthesis method for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105568, Oct. 2020, doi: 10.1016/j.cmpb.2020.105568.
- [37] A. Bissoto and S. Avila, "Improving skin lesion analysis with generative adversarial networks," in *Anais Estendidos da Conference on Graphics, Patterns and Images (SIBRAPI Estendido 2020)*, Nov. 2021, pp. 70–76, doi: 10.5753/sibgrapi.est.2020.12986.
- [38] C. Zhao, R. Shuai, L. Ma, W. Liu, Di. Hu, and M. Wu, "Dermoscopy image classification based on StyleGAN and DenseNet201," *IEEE Access*, vol. 9, pp. 8659–8679, 2021, doi: 10.1109/ACCESS.2021.3049600.
- [39] B. Ahmad, S. Jun, V. Palade, Q. You, L. Mao, and M. Zhongjie, "Improving skin cancer classification using heavy-tailed student t-distribution in generative adversarial networks (Ted-gan)," *Diagnostics*, vol. 11, no. 11, p. 2147, Nov. 2021, doi: 10.3390/diagnostics11112147.
- [40] A. Shahsavari, S. Ranjbari, and T. Khatibi, "Proposing a novel cascade ensemble super resolution generative adversarial network (CESR-GAN) method for the reconstruction of super-resolution skin lesion images," *Informatics in Medicine Unlocked*, vol. 24, p. 100628, 2021, doi: 10.1016/j.imu.2021.100628.
- [41] I. S. A. Abdelhalim, M. F. Mohamed, and Y. B. Mahdy, "Data augmentation for skin lesion using self-attention based progressive generative adversarial network," *Expert Systems with Applications*, vol. 165, p. 113922, Mar. 2021, doi: 10.1016/j.eswa.2020.113922.
- [42] P. Sharma, A. Gautam, R. Nayak, and B. K. Balabantaray, "Melanoma detection using advanced deep neural network," in *4th International Conference on Energy, Power, and Environment, ICEPE 2022*, Apr. 2022, pp. 1–5, doi: 10.1109/ICEPE55035.2022.9798123.
- [43] A. A. M. Teodoro *et al.*, "A skin cancer classification approach using GAN and RoI-based attention mechanism," *Journal of Signal Processing Systems*, vol. 95, no. 2–3, pp. 211–224, Apr. 2023, doi: 10.1007/s11265-022-01757-4.
- [44] A. Al-Rasheed, A. Ksibi, M. Ayadi, A. I. A. Alzahrani, M. Zakariah, and N. Ali Hakami, "An ensemble of transfer learning models for the prediction of skin cancers with conditional generative adversarial networks," *Diagnostics*, vol. 12, no. 12, p. 3145, Dec. 2022, doi: 10.3390/diagnostics12123145.
- [45] A. M. Hassan, K. Mahar, and M. M. Fouad, "MSDSC: a multistage deep learning-based skin cancer classifier," *Research Article preprint*, Apr. 2023, doi: 10.21203/rs.3.rs-2792248/v1.
- [46] S. Saravanan, B. Heshma, A. V. Ashma Shanofar, and R. Vanithamani, "Skin cancer detection using dermoscope images," in *Materials Today: Proceedings*, 2020, vol. 33, pp. 4823–4827, doi: 10.1016/j.matpr.2020.08.388.
- [47] A. Marka, J. B. Carter, E. Toto, and S. Hassanpour, "Automated detection of nonmelanoma skin cancer using digital images: a systematic review," *BMC Medical Imaging*, vol. 19, no. 1, p. 21, Feb. 2019, doi: 10.1186/s12880-019-0307-7.
- [48] P. Tschandl, C. Rosendahl, and K. Hararid, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci Data* 5, vol. 5, p. 180161, 2018, doi: 10.7910/DVN/DBW86T.
- [49] J. Malvehy, K. Nelson, and M. Nicholas, "ISIC: international skin imaging collaboration." <https://www.isic-archive.com/technology-working-group>, accessed 29-03-2024.
- [50] T. Mendonca, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "PH2 - a dermoscopic image database for research and benchmarking," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Jul. 2013, pp. 5437–5440, doi: 10.1109/EMBS.2013.6610779.
- [51] DermIS, "Dermofit image library." <https://licensing.edinburgh-innovations.ed.ac.uk/product/dermofit-image-library>, accessed on 29-03-2024.





BIOGRAPHIES OF AUTHORS

Dr. Mohammed Altaf Ahmed     received his Ph.D. in Engineering from GITAM University, Vishakhapatnam, India, for the thesis entitled "Design of a built-in self-test controller using the march algorithm for fault diagnosis in embedded memories" in the year 2018. He received the Best Thesis Award and the Gold Medal for his Ph.D. He completed a master's in engineering and technology with a specialization in embedded systems in 2008 from Jawaharlal Nehru Technological University, Hyderabad, India. He received his bachelor's in engineering (Electronics and Communications) in 2000 from Swami Ramanand Teerth Marathwada University, Nanded, India. He is presently focusing his research on VLSI design, embedded systems, and IoT devices. He has successfully implemented several research grants projects in the same area and published research articles in leading academic journals indexed in Scopus and Science. He can be contacted at email: m.altaf@psau.edu.sa.







Mohammad Naved Qureshi Ahmed     is an Assistant Professor at the Electrical Engineering Section, University Polytechnic, Aligarh Muslim University, since 2015. He holds an M.Tech degree in Software Engineering from the same university, acquired in 2014. A prolific author, he has contributed extensively to various technical publications, proceedings, and book chapters. His primary research interest lies in artificial intelligence and its applications, a field in which he has demonstrated considerable expertise. He can be contacted at email: naved.ubp@amu.ac.in.



Mohammad Sarosh Umar     is working as Professor in the Department of Computer Engineering at Aligarh Muslim University, India. He has teaching/research and industrial experience of more than 32 years in India as well as abroad. He has served as Chairman, Department of Computer Engineering, Aligarh Muslim University for 8 and 1/2 years. Sarosh Umar has delivered more than 60 invited talks in the areas of E-governance, E-commerce, user authentication, soft-ware engineering and other related topics. His research interests are user authentication using graphical methods, computer security and software engineering, and he has published 3 patents. He can be contacted at email: saroshumar@zhcet.ac.in.



Mouna Bedoui     received Mouna earned her M.S. in Microelectronic Systems from Tunisia's Faculty of Sciences in 2016. She completed her doctorate in the year 2022. Design and implementation of a reconfigurable crypto processor, as well as embedded system security on FPGA, are among her research interests. He can be contacted at email: mouna.bedoui1991@gmail.com.