

*cancers*



Article

---

# *DM–AHR*: A Self-Supervised Conditional Diffusion Model for AI-Generated Hairless Imaging for Enhanced Skin Diagnosis Applications

---

Bilel Benjdira, Anas M. Ali, Anis Koubaa, Adel Ammar and Wadii Boulila

Special Issue

Skin Cancer: Imaging and Radiotherapy

Edited by






Dr. Fabio Sandomenico



<https://doi.org/10.3390/cancers16172947>

## Article

# DM-AHR: A Self-Supervised Conditional Diffusion Model for AI-Generated Hairless Imaging for Enhanced Skin Diagnosis Applications

Bilel Benjdira <sup>1,2,\*</sup> , Anas M. Ali <sup>1,3</sup> , Anis Koubaa <sup>1</sup> , Adel Ammar <sup>1</sup>  and Wadii Boulila <sup>1,4</sup> 

<sup>1</sup> Robotics and Internet-of-Things Laboratory, Prince Sultan University, Riyadh 11586, Saudi Arabia; aaboessa@psu.edu.sa (A.M.A.); akoubaa@psu.edu.sa (A.K.); aammar@psu.edu.sa (A.A.); wboulila@psu.edu.sa (W.B.)

<sup>2</sup> SE & ICT Laboratory, LR18ES44, ENICarthage, University of Carthage, Tunis 1054, Tunisia

<sup>3</sup> Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

<sup>4</sup> RIADI Laboratory, University of Manouba, Manouba 2010, Tunisia

\* Correspondence: bbenjdira@psu.edu.sa

**Simple Summary:** Skin diseases can be serious, and early detection is key to effective treatment. Unfortunately, the quality of images used to diagnose these diseases often suffers due to interference from hair, making accurate diagnosis challenging. This research introduces a novel technology, the *DM-AHR*, a self-supervised conditional diffusion model designed specifically to generate clear, hairless images for better skin disease diagnosis. Our work not only presents a new, advanced model that expertly identifies and removes hair from dermoscopic images but also introduces a specialized dataset, *DERMAHAIR*, to further research and improve diagnostic processes. The enhancements in image quality provided by *DM-AHR* significantly improve the accuracy of skin disease diagnoses, and it promises to be a valuable tool in medical imaging.



**Citation:** Benjdira, B.; M. Ali, A.; Koubaa, A.; Ammar, A.; Boulila, W. *DM-AHR: A Self-Supervised Conditional Diffusion Model for AI-Generated Hairless Imaging for Enhanced Skin Diagnosis Applications.* *Cancers* **2024**, *16*, 2947. <https://doi.org/10.3390/cancers16172947>

Academic Editor: Brigitta G. Baumert

Received: 10 June 2024

Revised: 8 August 2024

Accepted: 15 August 2024

Published: 23 August 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Accurate skin diagnosis through end-user applications is important for early detection and cure of severe skin diseases. However, the low quality of dermoscopic images hampers this mission, especially with the presence of hair on these kinds of images. This paper introduces *DM-AHR*, a novel, self-supervised conditional diffusion model designed specifically for the automatic generation of hairless dermoscopic images to improve the quality of skin diagnosis applications. The current research contributes in three significant ways to the field of dermatologic imaging. First, we develop a customized diffusion model that adeptly differentiates between hair and skin features. Second, we pioneer a novel self-supervised learning strategy that is specifically tailored to optimize performance for hairless imaging. Third, we introduce a new dataset, named *DERMAHAIR* (DERMatologic Automatic HAIR Removal Dataset), that is designed to advance and benchmark research in this specialized domain. These contributions significantly enhance the clarity of dermoscopic images, improving the accuracy of skin diagnosis procedures. We elaborate on the architecture of *DM-AHR* and demonstrate its effective performance in removing hair while preserving critical details of skin lesions. Our results show an enhancement in the accuracy of skin lesion analysis when compared to existing techniques. Given its robust performance, *DM-AHR* holds considerable promise for broader application in medical image enhancement.

**Keywords:** skin diagnosis consumer applications; healthcare image generation; healthcare image enhancement; melanoma detection; diffusion models; skin medical images; automatic skin hair removal; self-supervised techniques

## 1. Introduction

Global skin cancer rates are escalating, highlighting the urgent need for accurate and efficient diagnostic methods. The World Health Organization (WHO) reports that skin

cancers comprise a significant proportion of all cancer cases worldwide, with melanoma recognized as the most lethal form. The incidence of melanoma has been steadily increasing over the past few decades, marking it as a critical public health concern (WHO, 2021) [1]. Early detection is paramount in the fight against skin cancer, as it significantly improves prognosis. As part of this fight, there has been an increase in end user applications designed to diagnose skin health conditions. However, these applications are still hampered by the presence of many conditions that decrease the detection accuracy.

The integration of artificial intelligence (AI) [2–6] in medical imaging, particularly in dermatology, has greatly benefited the early detection and diagnosis of skin cancers, including melanoma [7–9]. Dermoscopy, a noninvasive skin imaging technique, is crucial in this process. However, the presence of hair in dermoscopic images often poses significant challenges [10], as it can obscure critical details of skin lesions and may lead to potential diagnostic errors. Physical removal of hair is a solution, but it can be uncomfortable for patients and may not always be feasible. To address this issue, we introduce *DM-AHR*, a novel conditional diffusion model specifically designed to generate hairless dermoscopic images that are more suitable for diagnosis and detection. Diffusion models have emerged as powerful tools in the field of machine learning, particularly for image-generation tasks. These models create high-quality, detailed images by iteratively refining a signal from a random noise distribution based on stochastic process principles. This approach allows for the generation of images with remarkable fidelity and diversity. Our research focuses on improving the quality of dermoscopic images, which is vital for early skin cancer detection and diagnosis. We believe *DM-AHR* can significantly contribute to this field.

The flexibility and robustness of diffusion models make them especially effective for tasks requiring nuanced understanding and manipulation of visual data. Our study repurposes diffusion models to enhance the clarity of medical skin images. The *DM-AHR* model is designed to discriminate between hair and underlying skin features, gradually generating a hair-free image from the original hairy image. This approach not only augments the visibility of critical dermatological details but also streamlines the pre-diagnostic process by eliminating the need for physical hair removal. The main contributions of this paper are as follows:

- We introduce *DM-AHR*, a customized self-supervised diffusion model for automatic hair removal in dermoscopic images. This model enhances the clarity and diagnostic quality of skin images by effectively differentiating hair from skin features.
- We demonstrate significant improvement in skin lesion detection accuracy after applying *DM-AHR* as a preprocessing step before the classification model.
- We introduce the novel DERMAHAIR dataset, advancing research in the task of automatic hair removal from skin images.
- We introduce a new self-supervised technique specifically tailored for the task of automatic hair removal from skin images.
- We demonstrate *DM-AHR*'s potential in broader skin diagnosis applications and automated dermatological systems.

This study details the architecture and optimization of *DM-AHR*. The structure of the paper is as follows: Section 2 discusses related works in the field, providing background and context for the current research. Section 3 describes the proposed methodology, detailing the design, implementation, and theoretical foundation of the *DM-AHR* model. Section 4 presents the experimental results, offering a thorough analysis of the model's performance, comparisons with existing methods, and insights gained from the experiments. Finally, Section 5 concludes the paper, summarizing the key findings, discussing implications, and suggesting directions for future research.

## 2. Related Works

The domain of medical image analysis has seen substantial advancements due to the integration of AI and machine learning techniques. Particularly, in dermatology, the automatic analysis of dermoscopic images plays a crucial role in the early detection of skin diseases, including melanoma. This section reviews the existing literature in three primary areas: diffusion models in image processing, automatic hair removal techniques in dermoscopic images, and their implications for skin disease diagnosis.

First, diffusion models have gained prominence in image processing due to their exceptional ability to generate and reconstruct high-quality images. Recent studies, such as those by [11,12], have demonstrated the potential of these models in tasks like image denoising, super-resolution, and inpainting. Their application in medical image analysis, however, remains relatively unexplored. The adaptability of diffusion models to different contexts, as shown by [13], provides a foundation for their application in enhancing the quality of medical images, including dermoscopic images. Another notable development in the domain of dermatoscopic image enhancement is the Derm-T2IM approach, as explored by Farooq et al [14]. Their work utilizes stable diffusion models to generate large-scale, high-quality synthetic skin lesion datasets, enhancing machine learning model training by providing a rich variety of realistic dermatoscopic images that improve the robustness and accuracy of skin disease classification models when integrated into training pipelines.

Hair removal in dermoscopic images is a critical preprocessing step in skin lesion analysis. Traditional methods, as reviewed by [15,16], have relied on techniques such as morphological operations and edge detection. However, these methods often struggle with maintaining lesion integrity. The advent of deep learning has introduced new possibilities [17–19]. More recent studies have started exploring convolutional neural networks (CNNs) [20–22] and Generative Adversarial Networks (GAN) [23–25] for this purpose, achieving better accuracy while preserving important dermatological features.

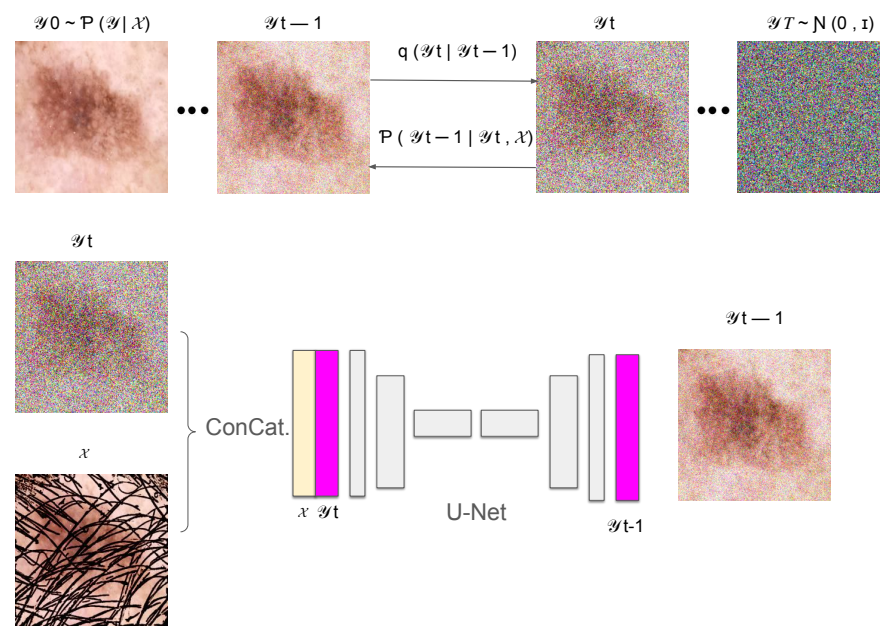
The accuracy of skin disease diagnosis, particularly when diagnosing melanoma, is heavily dependent on the quality of dermoscopic images. As highlighted by [26], combining AI with high-quality images can outperform traditional diagnostic methods. The removal of occluding elements like hair is therefore not just a technical challenge, but a clinical necessity, as emphasized by [27,28]. The integration of recent advanced image processing techniques, such as diffusion models, in this domain holds significant promise for improving diagnostic accuracy and patient outcomes. Moreover, leveraging self-supervised techniques and constructing new datasets could advance the state of the art in this domain. This study addresses these improvement gaps and introduces the *DM-AHR* model that will be described in the next section.

## 3. Proposed Method

In this section, we introduce *DM-AHR* (Diffusion Model for Automatic Hair Removal), based on modifying the standard architecture of Denoising Diffusion Probabilistic Models (DDPM) [11] for the purpose of automatic hair removal in dermoscopic images. The global architecture of *DM-AHR* is illustrated in Figure 1.

### 3.1. Dataset and Problem Formulation

Starting from a dataset  $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$ , where each  $\mathbf{x}_i$  is a dermoscopic image exhibiting hair artifacts, and  $\mathbf{y}_i$  represents the corresponding hair-free image. The dataset is crucial for training the *DM-AHR* model to understand and learn the complex patterns associated with hair in such medical images. Formally, we define our task as learning a mapping function  $F$ , such that  $F(\mathbf{x}_i) \approx \mathbf{y}_i$ , where  $F$  represents the *DM-AHR* model's transformation capability. Due to the difficulty of constructing real pairs of hairy and hair-free images, the hair is added artificially to real hair-free skin images, as explained in the Experimental Section later.



**Figure 1.** Illustration of the *DM-AHR* Diffusion process.

### 3.2. Conditional Denoising Diffusion Process

The *DM-AHR* model incorporates a sophisticated conditional denoising diffusion process inspired by prominent diffusion model advancements [11,13,29]. The architecture central to this model is a U-Net [30], which is pivotal for denoising and feature differentiation, specifically between hair and skin features in images. This U-Net employs a multi-scale strategy with channels varying in dimensionality corresponding to the resolution of input images, ensuring a tailored approach for different scales, as described in Figure 1.

Although the *DM-AHR* model utilizes the U-Net structure, it is distinctly modified to include a conditioning branch at every denoising step. The inputs to the model are images of skin cancer with hair, which are instrumental in guiding the model to align more accurately with the content of the skin cancer image. This conditioning mechanism significantly enhances the model's ability to adhere to the relevant features in the images, as visually represented in Figure 1.

The U-Net comprises several layers with depth multipliers at each resolution stage to enhance its capacity and adaptability to the input's complexity. It utilizes a series of ResNet blocks [31] at each stage, and these blocks help it to learn residual behaviors between noisy and denoised states, thus improving the model's efficiency and accuracy. Activation functions within these blocks are primarily ReLU, ensuring non-linearity preservation in the denoising process.

For the diffusion process, the model starts with a forward diffusion to generate noisy versions of the original image, which are subsequently refined by the conditional U-Net. This U-Net takes both the noisy image and a low-resolution, interpolated, and up-sampled version of the original image as inputs. It then progressively enhances the image resolution through its depth stages, adapting its parameters dynamically as dictated by the underlying complexity of different image features. The loss function employed is tailored to minimize the difference between the clean and generated images, ensuring the preservation and accurate reconstruction of critical features such as hair and skin textures.

### 3.2.1. Forward Diffusion Process

The forward diffusion process is modeled as a Markov chain that incrementally adds Gaussian noise to an image. Formally, the process is defined as follows:

$$q(\mathbf{y}_{1:T} | \mathbf{y}_0) = \prod_{t=1}^T q(\mathbf{y}_t | \mathbf{y}_{t-1}), \quad (1)$$

$$q(\mathbf{y}_t | \mathbf{y}_{t-1}) = \mathcal{N}(\mathbf{y}_t | \sqrt{\alpha_t} \mathbf{y}_{t-1}, (1 - \alpha_t) \mathbf{I}) \quad (2)$$

where  $\mathbf{y}_0$  is the original hair-free image,  $\mathbf{y}_{1:T}$  represents the sequence of images with incrementally added noise, and  $\alpha_{1:T}$  are the noise levels at each step. This part is illustrated in the upper part of Figure 1.

### 3.2.2. Reverse Diffusion Process

In the reverse diffusion process, the aim is to recover the original hair-free image from the noisy image. This is achieved by a learned neural network  $f_\theta$ , which performs denoising at each step. Here, we added the conditioning based on the original image with hair, represented as  $\mathbf{x}$ , to guide the reconstruction process to remove hair and noise jointly. This is shown in the bottom part of the Figure 1.

The reverse process is defined as follows:

$$\begin{aligned} p_\theta(\mathbf{y}_{0:T} | \mathbf{x}) &= p(\mathbf{y}_T) \prod_{t=1}^T p_\theta(\mathbf{y}_{t-1} | \mathbf{y}_t, \mathbf{x}), \\ p(\mathbf{y}_T) &= \mathcal{N}(\mathbf{y}_T | \mathbf{0}, \mathbf{I}), \\ p_\theta(\mathbf{y}_{t-1} | \mathbf{y}_t, \mathbf{x}) &= \mathcal{N}(\mathbf{y}_{t-1} | \mu_\theta(\mathbf{x}, \mathbf{y}_t, \alpha_t), \sigma_t^2 \mathbf{I}). \end{aligned} \quad (3)$$

where  $\mu_\theta$  and  $\sigma_t^2$  are learned parameters of the model and  $\mathbf{x}$  is the input hair-containing image. The reverse diffusion process is iterative, starting from a purely noisy image and progressively recovering the hair-free image, guided by the input image characteristics.

## 3.3. Training of the Denoising Model

The denoising model  $f_\theta$ , a pivotal component of the DM-AHR architecture, is tasked with estimating and eliminating the noise introduced at each step of the forward diffusion process. This process is crucial for reconstructing the hair-free image from its noisy counterpart. The training of  $f_\theta$  is conducted using pairs of noisy and original hair-free images, with the aim of fine-tuning the model parameters to accurately estimate the noise vector.

### 3.3.1. Loss Function

The training objective is encapsulated in a loss function that measures the difference between the noise estimated by the model and the actual noise applied during the forward process. Formally, the loss function is as follows:

$$\mathcal{L}(\theta) = \mathbb{E}_{(\mathbf{x}, \mathbf{y}), \epsilon, t} [\|f_\theta(\mathbf{x}, \mathbf{y}_t, t) - \epsilon\|^2], \quad (4)$$

where  $\mathbf{y}_t$  represents the noisy image at time step  $t$  and  $\epsilon$  denotes the noise vector. This function aims to minimize the mean squared error between the noise predicted by  $f_\theta$  and the actual noise.

### 3.3.2. Training Algorithm

The training algorithm for  $f_\theta$  is carefully designed to iteratively refine the model parameters, thereby optimizing its ability to estimate and remove noise from images. The process can be dissected into several key steps, each with a mathematical basis, as demonstrated in Algorithm 1:



**Algorithm 1:** Training algorithm for the *DM-AHR* Model  $f_\theta$ 


---

```

while not converged do
  for each batch in the training dataset do
    Sample a pair  $(\mathbf{x}, \mathbf{y}_0)$  from the batch
    for  $t = 1$  to  $T$  do
      Sample noise level  $\gamma_t$  from a predefined noise schedule
      Generate noise  $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
      Create a noisy image  $\mathbf{y}_t$  using the formula:
         $\mathbf{y}_t = \sqrt{1 - \gamma_t} \mathbf{y}_0 + \sqrt{\gamma_t} \epsilon_t$ 
      Compute the loss function for backpropagation:
         $\mathcal{L}(\theta) = \|f_\theta(\mathbf{x}, \mathbf{y}_t, t) - \epsilon_t\|^2$ 
      Update  $f_\theta$  by applying the gradient descent step:
         $\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}(\theta)$ 
    end for
  end for
end while

```

---

In the training algorithm for the *DM-AHR* model, each iteration commences with sampling a pair of images  $(\mathbf{x}, \mathbf{y}_0)$ , where  $\mathbf{x}$  represents the original image with hair and  $\mathbf{y}_0$  its hair-free counterpart. At each time step  $t$ , noise is added to  $\mathbf{y}_0$  based on a sampled noise level  $\gamma_t$ , creating a noisy version  $\mathbf{y}_t$ . The loss function, which is critical for guiding the model, measures the difference between the model's estimated noise and the actual noise added. Model parameters are updated via gradient descent, where the learning rate ( $\eta$ ) and the gradient of the loss function ( $\nabla_\theta \mathcal{L}(\theta)$ ) are key factors. This approach ensures effective learning and robust noise estimation capabilities, enabling the model to accurately map noisy images to their hair-free versions under a variety of noise levels.

### 3.4. Inference Process of the *DM-AHR* Model

During the inference process of *DM-AHR* model, the Iterative Refinement Algorithm [29] is adapted to the current problem. The inference is performed for multiple iterations and guided by the input image containing hair and it leverages the denoising model  $f_\theta$  (the U-Net model shown in Figure 1) at each iteration.

The algorithm initializes with a noisy image  $\mathbf{y}_T$  and progressively applies denoising to reach the final hair-free image  $\mathbf{y}_0$ . The refinement at each step is mathematically governed by the following equation:

$$\mathbf{y}_{t-1} = f_\theta(\mathbf{y}_t, \mathbf{x}, t) + \mathbf{y}_t, \quad (5)$$

where  $\mathbf{y}_t$  is the noisy image at step  $t$ , and  $\mathbf{x}$  is the original hair-containing image. The function  $f_\theta$  is responsible for estimating the difference between  $\mathbf{y}_t$  and the hair-free image, effectively removing noise and hair artifacts.

The effectiveness of this algorithm in hair removal can be attributed to its iterative nature, allowing gradual refinement of the image. Each iteration removes a portion of the noise and hair, refining the image closer to its hair-free version. The process is described in Algorithm 2, where  $T$  is the number of iterations fixed by the user to regulate the inference.

Each step of the algorithm incrementally removes noise and hair, leveraging the learned model  $f_\theta$  to guide the transformation towards a clear, hair-free image. An example of the inference process applied on a one sample image is illustrated in Figure 2, and the inference number is 10 ( $T = 10$ ).

The noise schedule of the *DM-AHR* model is defined by a piecewise uniform distribution, which is crucial for controlling the noise variance at each diffusion step:

$$p(\gamma) = \sum_{t=1}^T \frac{1}{T} U(\gamma_{t-1}, \gamma_t), \quad (6)$$

where  $U(\gamma_{t-1}, \gamma_t)$  is a uniform distribution between noise levels  $\gamma_{t-1}$  and  $\gamma_t$ , and  $T$  is the total number of diffusion steps. This noise schedule is vital for the model's ability to handle various noise levels, enhancing its robustness and adaptability for effective hair removal in medical imaging applications.

---

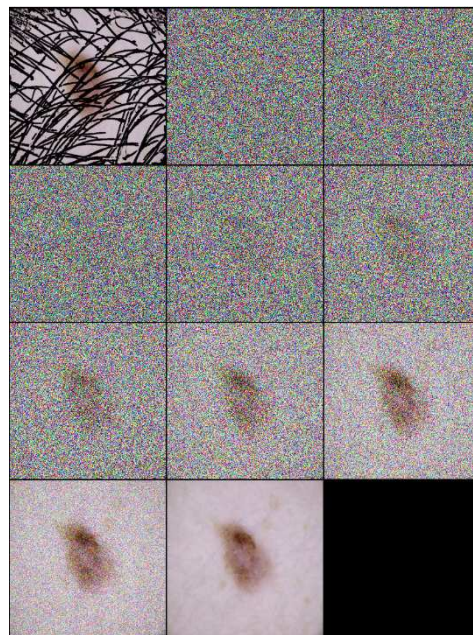
**Algorithm 2:** Inference process for hair removal in *DM-AHR*

---

```

Initialize  $\mathbf{y}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
for  $t = T$  down to 1 do
    Sample noise  $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
    Apply the denoising model:
         $\mathbf{y}_{t-1} = f_{\theta}(\mathbf{y}_t, \mathbf{x}, t) + \mathbf{y}_t$ 
end for
return  $\mathbf{y}_0$ , the final hair-free image
  
```

---



**Figure 2.** Inference steps of *DM-AHR* applied on one melanoma skin case performed in 10 iterations ( $T = 10$ ).

### 3.5. Self-Supervised Technique

The training of the *DM-AHR* model employs a novel self-supervised technique, which is specifically tailored for the automatic removal of hair from dermoscopic skin images. This approach leverages the algorithm as shown in Algorithm 3. The key idea is to artificially synthesize hair-like patterns on hair-free dermoscopic images, creating a modified dataset that simulates real-world scenarios in which hair obscures skin lesions.

The process begins with an original, hair-free dermoscopic image, denoted as  $I_{original}$ . A pre-set number of lines,  $N_{lines}$ , for example, 1000, are algorithmically drawn on this image to imitate hair. These lines are generated with randomized starting points, lengths, and directions, ensuring a diverse and realistic representation of hair patterns across the dataset.



---

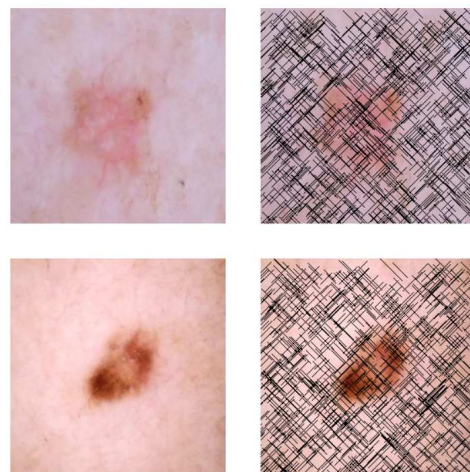
**Algorithm 3:** Self-supervised technique customized for training skin hair removal models
 

---

**Input:**  $I_{original}$  {the original hair-free dermoscopic image}  
**Output:**  $I_{modified}$  {Image with the small artificial lines simulating hair-like patterns}  
 $N_{lines} \leftarrow 1000$  {Fix the number of lines to draw}  
 $(W, H) \leftarrow \text{Size of } I_{original}$  {Get the input image dimensions}  
 $I_{modified} = I_{original}$   
**for**  $i = 1$  to  $N_{lines}$  **do**  
      $x_{start} \sim \mathcal{U}(0, W)$  {pick a random integer in  $[0, W]$ }  
      $y_{start} \sim \mathcal{U}(0, H)$  {pick a random integer in  $[0, H]$ }  
      $L_{length} \sim \mathcal{U}(10, 30)$  {pick a random integer in  $[10, 30]$ }  
      $d \sim \{-1, 1\}$  {pick a random direction for diagonal orientation}  
      $x_{end} \leftarrow x_{start} + d \cdot L_{length}$  {End x-coordinate}  
      $y_{end} \leftarrow y_{start} + L_{length}$  {End y-coordinate}  
     Draw line on  $I$  from  $(x_{start}, y_{start})$  to  $(x_{end}, y_{end})$  on  $I_{modified}$   
**end for**  
**return**  $I_{modified}$

---

This self-supervised approach significantly augments the model's proficiency in processing real-world dermoscopic images while obviating the necessity for a large corpus of manually annotated images obscured by hair. Figure 3 presents two exemplar images produced by this technique, which effectively mimic natural hair patterns. This visual mimicry is instrumental in enabling the *DM-AHR* model to autonomously discern and segregate hair from skin features. Notably, the implementation of this technique has led to a substantial enhancement in the performance of the *DM-AHR* model, which complements its prior training on the meticulously constructed DERMAHAIR dataset, as detailed in the subsequent section.



**Figure 3.** Generated samples using the customized self-supervised technique.

## 4. Experiments

### 4.1. Experimental Settings

The *DM-AHR* model's training and evaluation process was conducted using a high-performance NVIDIA Quadro RTX 8000 GPU with 48 GB VRAM, known for its substantial computational capabilities. The *DM-AHR* model underwent an intensive training protocol consisting of 2000 time steps for detailed refinement and 20,000 iterations for comprehensive learning across diverse imaging conditions. This extensive training regime was crucial for iterative improvement in image quality and ensuring the model's robustness. The Adam optimizer was employed with a fixed rate of 0.0001, facilitating stable and

consistent training. The model and resources are made available to the community at: <https://github.com/AnasHXXH/DM-AHR> (accessed on 7 August 2024).

#### 4.2. Dataset Description

This study introduces the dataset DERMAHAIR (DERMatologic Automatic HAIR Removal Dataset). It is created to benchmark the task of Automatic Skin Hair Removal to enhance skin diagnosis and analysis. DERMAHAIR was created by sampling 100 images from each of the seven lesion categories within the HAM10000 dataset [32], resulting in 700 unique images. Later, for every image, an artificial hair pattern was picked randomly from the Digital Hair Dataset [33] and added artificially to the image. In the following paragraphs, we will provide a description of the HAM10000 dataset, and we will explain the artificial hair addition process. The DERMAHAIR dataset is open sourced at the following link: <https://www.kaggle.com/datasets/riotulab/skin-cancer-hair-removal> (accessed on 7 August 2024).

##### 4.2.1. Description of the HAM10000 Dataset

The HAM10000 dataset [32] is a cornerstone in dermatological research, encompassing 10,000 dermoscopic images that represent a broad spectrum of skin lesions, including melanomas and nevi. This dataset is distinguished by its comprehensive diversity in lesion characteristics and patient demographics. Each image is meticulously annotated with essential metadata such as diagnosis, lesion location, and patient specifics, thereby providing a robust foundation for the development of machine learning models aimed at skin lesion classification and melanoma detection. The HAM10000 dataset includes seven lesion categories: Melanocytic nevi (benign melanocyte tumors, nv), Melanoma (the deadliest form of skin cancer, mel), Benign keratosis (comprising seborrheic keratoses and lentigines, non-malignant growths from keratinocytes, bkl), Basal cell carcinoma (the most common but least aggressive skin cancer, bcc), Actinic keratoses (sun-related precancerous lesions, akiec), Vascular lesions (examples include angiomas, benign vascular proliferations, vasc), and Dermatofibroma (non-malignant fibrous skin tumors, df). The diverse and detailed classification of skin conditions in this dataset renders it invaluable for training advanced AI models in dermatological diagnostics.

We reserved 100 images from each category to form a test set of 700 images, ensuring each class is equally represented in the evaluation process. This stratified sampling minimizes bias and robustly measures the model's performance across varied conditions. The remaining images were augmented with synthetic hair, forming a diverse training set. Given the specific challenge of hair removal in skin cancer images, the size of the dataset has proven sufficient for achieving high accuracy and reliability in this niche application, as evidenced by our experimental results and the specific adaptations we have made to mitigate the limitations of a smaller sample size.

##### 4.2.2. Description of the Artificial Hair Addition Process

Let us consider a simplified 2D matrix representation of a single color channel from an RGB image:

$$\text{Dermatoscopic Image (img)} = \begin{bmatrix} 0.3 & 0.6 \\ 0.5 & 0.8 \end{bmatrix} \quad (7)$$

These matrices symbolize pixel intensity values within a specific color channel, normalized between 0 and 1.

**Hair Mask Images:** To simulate hair artifacts, we employed binary masks sourced from a digital hair dataset [33]. In these masks, hair is represented by white pixels (value = 1), contrasting with the black background pixels (value = 0). Let us consider a simplified representation:

$$\text{Hair Mask Image (img\_h)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (8)$$

**Image Processing and Synthesis:** The process of artificial hair addition encompassed several key steps:

*Normalization of Hair Mask:* Initially, the hair mask image underwent normalization. This process involved inverting the pixel values, effectively converting white hair pixels to black, thus mirroring the appearance of hair in authentic dermatoscopic images:

$$\text{Normalized Hair Mask} = -1 \times (\text{img\_h} - 1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (9)$$

*Merging Images:* Subsequently, the normalized hair mask was merged with the dermatoscopic image. This crucial step superimposed the artificial hair structure onto the skin lesion image:

$$\text{Merged Image} = \text{img} + \text{Normalized Hair Mask} = \begin{bmatrix} 0.3 & 1.6 \\ 1.5 & 0.8 \end{bmatrix} \quad (10)$$

*Post-Processing:* The final stage entailed a post-processing routine. By subtracting one from the merged image's pixel values, the hair regions were effectively restored to their original skin color hues, while normalizing the background:

$$\text{Final Image} = \text{Merged Image} - 1 = \begin{bmatrix} -0.7 & 0.6 \\ 0.5 & -0.2 \end{bmatrix} \quad (11)$$

The DERMAHAIR dataset is constructed by picking 100 images from every class in the HAM10000 [32] and constructing a hairy version using this described hair addition process. The generated pair will be used to train any hair removal model, including the DM-AHR model described in this study.

In terms of its organization, the dataset is methodically divided into two principal directories: Data\_Skin\_with\_Hair and Data\_Skin\_without\_Hair. Each directory encompasses seven subdirectories corresponding to the seven diagnostic categories, namely akiec, bcc, bkl, df, mel, nv, and vasc. This structure ensures streamlined access and efficient categorization of the dataset.

#### 4.3. Evaluation Metrics

The evaluation metrics are divided into two groups. First, there are the metrics used to measure the performance of the DM-AHR model itself. Second, there are the metrics used to measure the performance of the skin lesion classification before and after using the DM-AHR model to improve the quality of the skin images.

##### 4.3.1. Evaluation Metrics of the DM-AHR Performance

In the evaluation of the DM-AHR model for Skin Image Dehairing, three key metrics are employed: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS).

**Peak Signal-to-Noise Ratio (PSNR):** PSNR is widely used in image processing to assess the quality of reconstructed images. It is defined as follows:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right), \quad (12)$$

where  $\text{MAX}_I$  is the maximum possible pixel value of the image and MSE is the Mean Squared Error between the original and processed images. Higher PSNR values indicate better quality, with values above 30 dB generally considered good in medical imaging.

**Structural Similarity Index Measure (SSIM):** SSIM measures the perceived quality of an image by comparing its structural information with that of the original image. It considers changes in texture, luminance, and contrast, and is calculated as follows:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (13)$$

where  $x, y$  are the original and processed images,  $\mu$  and  $\sigma$  are the mean and variance,  $\sigma_{xy}$  is the covariance, and  $C_1, C_2$  are constants to stabilize the division. The SSIM values range from  $-1$  to  $1$ , with higher values indicating better similarity.

**Learned Perceptual Image Patch Similarity (LPIPS):** LPIPS [34] is a metric that evaluates the perceptual similarity between two images, using deep learning models to capture the human visual system's response. It is defined as follows:

$$\text{LPIPS}(I_1, I_2) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \omega_l \|F_l(I_1)_{h,w} - F_l(I_2)_{h,w}\|_2^2, \quad (14)$$

where  $I_1, I_2$  are the images being compared,  $F_l$  denotes the feature map at layer  $l$ ,  $H_l, W_l$  are the dimensions of the feature map, and  $\omega_l$  are layer-wise weights. LPIPS quantifies perceptual differences more aligned with human visual perception.

These metrics collectively provide a multi-faceted evaluation of image quality, which is crucial for accurate diagnosis and analysis in medical imaging.

#### 4.3.2. Evaluation Metrics of the Skin Lesion Classification

After assessing the performance of the *DM-AHR* itself, we move on to the evaluation of the skin diagnosis task before and after using the *DM-AHR*. The classification of skin classes used six key performance metrics: the accuracy, the loss value, the precision, the recall, the specificity, and the F1-score. Each metric provides a unique perspective on the model's performance before and after using *DM-AHR* to enhance the skin image quality for skin diagnosis purposes.

**Accuracy** is the most intuitive performance measure and it is simply a ratio of correctly predicted observations to the total observations. It is defined as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \quad (15)$$

where accuracy is particularly useful when the classes in the dataset are nearly balanced.

**Loss**, specifically in the context of neural networks, is a measure of how well the model's predictions match the actual labels during the training. For the current study, the cross-entropy loss is used in all the experiments:

$$\text{Loss} = - \sum_i y_i \log(p_i) \quad (16)$$

where  $y_i$  is the actual label and  $p_i$  is the model's predicted probability for that class.

**Precision** is the ratio of correctly predicted positive observations to the total predicted positive observations. High precision relates to the low false-positive rate. It is defined as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (17)$$

**Recall** (sensitivity) measures the proportion of actual positives that were correctly identified. It is defined as follows:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (18)$$

**Specificity** is a measure that tells us what proportion of negatives was identified correctly. It complements recall and is defined as follows:

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} \quad (19)$$

**F1-score** is the weighted average of precision and recall. Therefore, this score takes both false-positives and false-negatives into account. It is a good way to show that a classifier has a good value for both recall and precision. The F1-score is defined as follows:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (20)$$

Understanding these metrics in tandem allows for a more nuanced interpretation of a model's performance, particularly in applications like medical image analysis, where the cost of false negatives or false positives can be high.

#### 4.4. Performance of the Skin Image Enhancement Using the DM-AHR Model

The *DM-AHR* model exhibits high efficacy in enhancing and reconstructing images across various cancer types, as evidenced by its PSNR, SSIM, and LPIPS metrics in Table 1 (PSNR), Table 2 (SSIM), and Table 3 (LPIPS). The *DM-AHR* model performance is compared with the performance of the GAN architecture [35], the SwinIR architecture [36], the standard Stable Diffusion Model (SDM) [37], the *DM-AHR* model without self-supervision (*DM-AHR*), and the full *DM-AHR(SS)* model with self-supervision. The best results are highlighted in bold.

**Table 1.** Comparison of the PSNR performance of *DM-AHR* with other models.

Skin Cancer Type	SRGAN	SwinIR	SDM	<i>DM-AHR</i>	<i>DM-AHR(SS)</i>
Actinic keratosis	32.401835	39.401370	33.731562	39.095776	<b>41.280412</b>
Basal cell carcinoma	31.700830	39.625874	34.317885	38.633978	<b>40.806729</b>
Benign Keratosis	31.644829	39.894135	34.109435	39.228128	<b>41.026910</b>
Dermatofibroma	31.807330	39.207628	34.196231	39.148305	<b>40.910219</b>
Melanoma	30.917425	38.493721	33.536934	38.728289	<b>39.883010</b>
Melanocytic nevi	32.242417	39.974818	34.960045	39.091819	<b>41.212679</b>
Vascular lesions	32.633316	41.737362	35.790086	39.963068	<b>42.170609</b>
<b>Average</b>	<u>31.906854</u>	<u>39.762129</u>	<u>34.377454</u>	<u>39.12705</u>	<b><u>41.041512</u></b>

The PSNR values, as shown in Table 1, reveal a consistent superiority of *DM-AHR(SS)*, the self-supervised version of *DM-AHR*, over other models. Such high PSNR values indicate that *DM-AHR(SS)* retains maximal fidelity to the original image. This is crucial in medical imaging, where detail preservation is paramount.

**Table 2.** Comparison of the SSIM performance of *DM-AHR* with other model.

Skin Cancer Type	SRGAN	SwinIR	SDM	<i>DM-AHR</i>	<i>DM-AHR(SS)</i>
Actinic keratosis	0.881940	0.959481	0.865522	0.960036	<b>0.967910</b>
Basal cell carcinoma	0.882850	0.957559	0.872703	0.955464	<b>0.964175</b>
Benign Keratosis	0.879369	0.959514	0.871176	0.956925	<b>0.965370</b>
Dermatofibroma	0.869542	0.957726	0.865824	0.956125	<b>0.964412</b>
Melanoma	0.868869	0.953418	0.860095	0.951718	<b>0.961027</b>
Melanocytic nevi	0.879389	0.955402	0.892593	0.954957	<b>0.962382</b>
Vascular lesions	0.899865	0.968740	0.917110	0.967328	<b>0.973995</b>
<b>Average</b>	<u>0.880261</u>	<u>0.958834</u>	<u>0.87786</u>	<u>0.957508</u>	<b><u>0.965613</u></b>



Table 2 presents the SSIM scores, underscoring *DM-AHR(SS)*'s ability to maintain structural integrity. The model exhibits exceptional performance in preserving textural and contrast details, as evidenced by it achieving the highest SSIM in every class.

The LPIPS metric, detailed in Table 3, further cements the superiority of *DM-AHR(SS)*. LPIPS, as a perception-based measure, highlights the model's proficiency in maintaining perceptual image quality, a critical aspect that is often overlooked by conventional metrics. The *DM-AHR(SS)* model consistently exhibits the lowest (best) LPIPS scores across all cancer types, confirming its excellence in perceptual image fidelity.

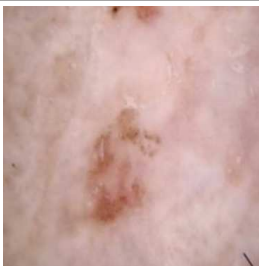
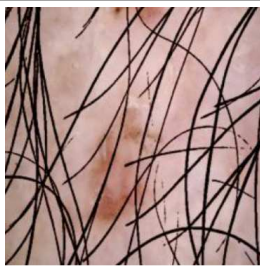

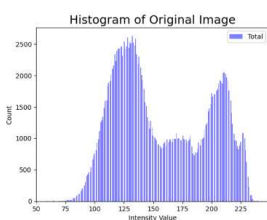
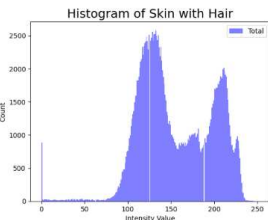
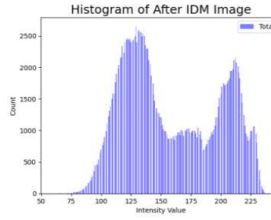
**Table 3.** Comparison of the LPIPS performance of *DM-AHR* with other models.

Skin Cancer Type	SRGAN	SwinIR	SDM	<i>DM-AHR</i>	<i>DM-AHR(SS)</i>
Actinic keratosis	0.171810	0.042776	0.050473	0.018420	<b>0.012620</b>
Basal cell carcinoma	0.175114	0.038973	0.046981	0.020546	<b>0.013838</b>
Benign Keratosis	0.185645	0.036158	0.047129	0.020156	<b>0.013638</b>
Dermatofibroma	0.182939	0.042536	0.045433	0.020481	<b>0.013376</b>
Melanoma	0.192164	0.045766	0.049873	0.023361	<b>0.015010</b>
Melanocytic nevi	0.164768	0.040105	0.036954	0.019475	<b>0.012884</b>
Vascular lesions	0.131372	0.025131	0.027423	0.013758	<b>0.008439</b>
<b>Average</b>	<u>0.171973</u>	<u>0.038777</u>	<u>0.043467</u>	<u>0.019457</u>	<b><u>0.012829</u></b>

When juxtaposed with models like SRGAN and SwinIR, *DM-AHR(SS)* not only surpasses them in terms of numerical metrics but also in qualitative aspects such as noise reduction, detail enhancement, and overall visual quality. The self-supervised learning approach ingrained in *DM-AHR (SS)* evidently contributes to its nuanced understanding and handling of complex dermatoscopic images.

Table 4 shows an example of a skin image without hair, the artificially added hair, and the reconstructed image using the *DM-AHR* model. We can see visually that the image is reconstructed at a high fidelity scale, alongside its histogram.

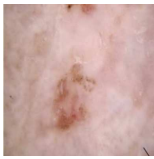



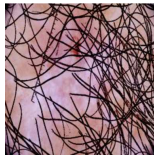











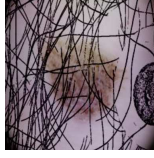


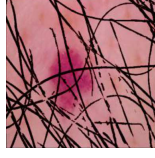

**Table 4.** Visual results of *DM-AHR* applied on one set of skin images with the corresponding histograms.

	Original Skin Image	Skin Image with Hair	Skin Image after <i>DM-AHR</i>
Result			
Histogram			

In Table 5, we show samples from different skin lesion classes with the reconstructed images using the *DM-AHR* model. We can also see the high reconstruction quality of the model and its successful discrimination between the skin features and the hair degradation patterns.



**Table 5.** Comparison of skin images before and after application of *DM-AHR* dehairing

Cancer Type	Original	After Hair	After <i>DM-AHR</i>
Actinic keratosis			
Basal cell carcinoma			
Benign keratosis			
Dermatofibroma			
Melanoma			
Melanocytic nevi			
Vascular lesions			

#### 4.5. Improvement of the Medical Skin Diagnosis after Using *DM-AHR*

In the evolving landscape of medical skin diagnosis, the robustness and accuracy of skin cancer classification models are paramount. This study presents a comprehensive evaluation of three state-of-the-art deep learning architectures: resnext101\_32x8d [38] (in Table 6), maxvit\_t [39] (in Table 7), and Swin-Transformer [40] (in Table 8). The performance of these models is assessed based on the original images, on the hairy version of the images, and finally on the dehaired version of the images using *DM-AHR* model. The ensuing tables encapsulate key performance metrics, including accuracy, loss, precision, recall, specificity, and F1-score for each model under the specified conditions.

For the resnext101\_32x8d model, the introduction of hair to the images led to a substantial decrease in accuracy, from an impressive 98.5% in the original state to 87.9% post-modification. This significant decline highlights the model's sensitivity to visual perturbations. However, the subsequent application of the IDM hair removal method effectively mitigated this impact, elevating the accuracy to 97.3%. The precision and recall metrics mirrored this trend, suggesting that *DM-AHR* preprocessing can significantly recover the model's discriminatory power.

**Table 6.** Performance metrics for the resnext101\_32x8d model.

Metric	Original	After Adding Hair	After <i>DM-AHR</i> Hair Removal
Accuracy	0.985242	0.879514	0.972731
Loss	0.050478	0.565113	0.099951
Precision	0.9856	0.8624	0.9825
Recall	0.9852	0.8795	0.9727
Specificity	0.9973	0.9778	0.9964
F1-score	0.9853	0.8660	0.9766

**Table 7.** Performance metrics for the maxvit\_t model.

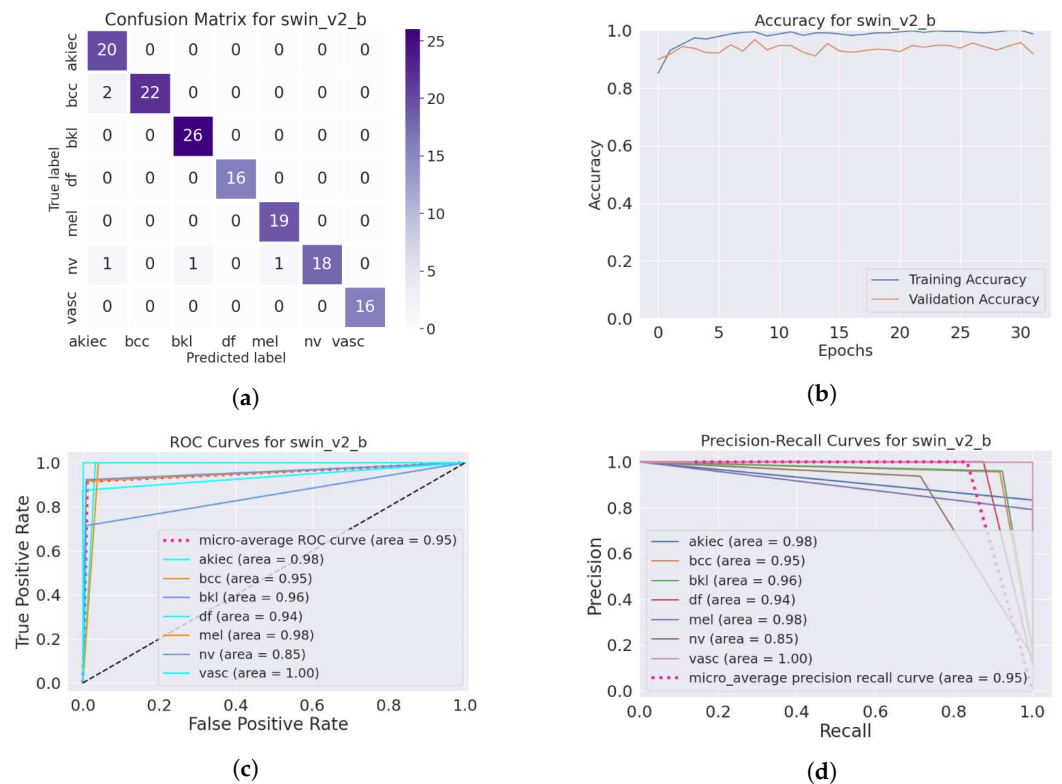
Metric	Original	After Adding Hair	After <i>DM-AHR</i> Hair Removal
Accuracy	0.988786	0.950656	0.992857
Loss	0.042649	0.168300	0.024280
Precision	0.9889	0.9436	0.9940
Recall	0.9888	0.9507	0.9929
Specificity	0.9980	0.9906	0.9988
F1-score	0.9888	0.9462	0.9933

Similarly, the maxvit\_t model, while demonstrating remarkable resilience with an original accuracy of 98.9%, also experienced a decline in performance after hair was added to the image, though to a lesser extent (95.1% accuracy). Post-*DM-AHR* application, the model not only recovered but exceeded its original performance, achieving an accuracy of 99.3%. This result suggests that the IDM method may not only compensate for the added noise but also enhances the model's overall interpretive accuracy. This is a notable property of the *DM-AHR* model that is worthy of investigation in future research studies.

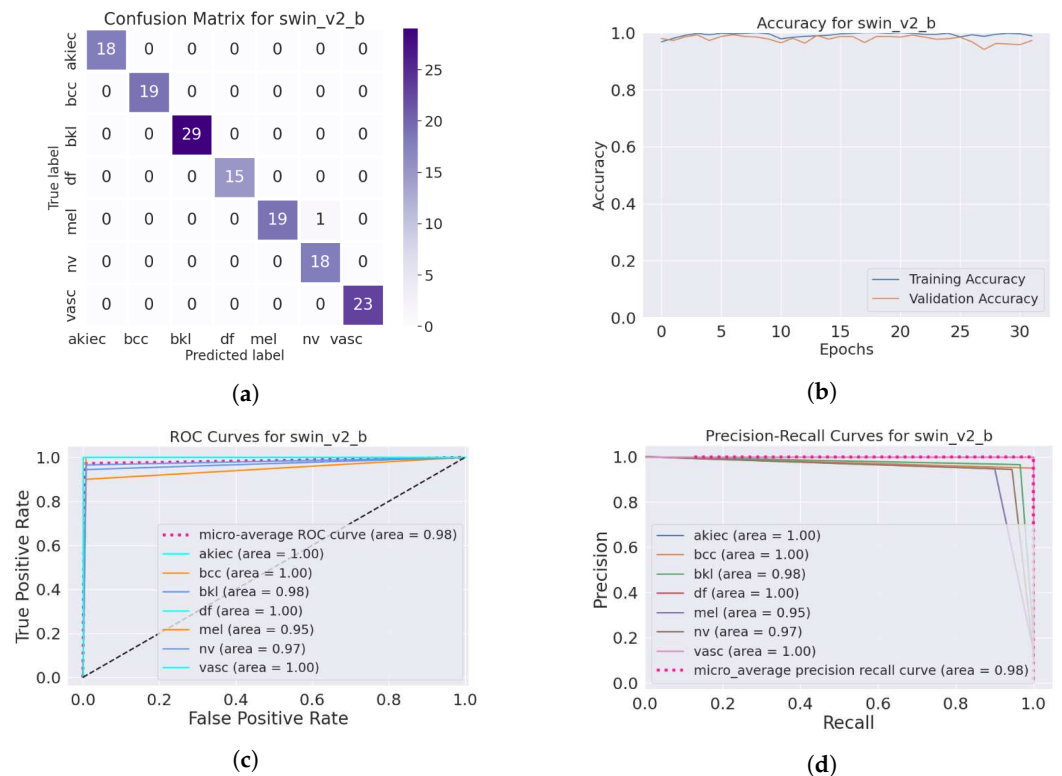
**Table 8.** Performance Metrics for the SwinTransformer Model.

Metric	Original	After Adding Hair	After IDM Hair Removal
Accuracy	0.983420	0.967687	0.992957
Loss	0.066955	0.114440	0.027340
Precision	0.9837	0.9689	0.9925
Recall	0.9834	0.9677	0.9929
Specificity	0.9970	0.9941	0.9988
F1-score	0.9834	0.9665	0.9925

The SwinTransformer model showed a consistent pattern of performance reduction upon hair addition, followed by a notable recovery with IDM processing. The original accuracy of 98.3% reduced to 96.8% after hair was added and then increased to 99.3% post-*DM-AHR*, paralleling the trends seen in the other models. This consistency across different architectures underscores the general applicability and effectiveness of the *DM-AHR* preprocessing technique in enhancing model robustness. Figures 4 and 5 illustrate the performance of the skin classification and diagnosis using the Swin Transformer before and after the enhancement step using the *DM-AHR* model. The analysis of the confusion matrices, the accuracy curves, the ROC curves, and the precision–recall curves shows the significant improvement of the skin analysis after the preprocessing using the *DM-AHR* model.



**Figure 4.** Performance of the skin classification using Swin Transformer before using *DM-AHR* (hairy occluded images): (a) Confusion matrix, (b) accuracy curve during training, (c) ROC curve, and (d) precision-recall curve.



**Figure 5.** Performance of the skin classification using Swin Transformer after using *DM-AHR* (de-haired images): (a) Confusion matrix, (b) accuracy curve during training, (c) ROC curve, (d) precision-recall curve.

Concerning the comparative study between these three skin classification models, this study suggests distinct applications for each model based on their performance. The resnext101 model is ideal for environments with clean, structured data due to its high precision and accuracy in such conditions. The maxvit\_t model, which demonstrates remarkable robustness to occlusions, is recommended for more challenging clinical settings where data variability and robustness are critical. SwinTransformer, with its consistent performance across varying conditions, serves as a versatile option for diverse medical imaging tasks. In scenarios requiring robust preprocessing, both maxvit\_t and SwinTransformer show significant improvements, making them suitable choices. A combination of these models or ensemble techniques could be advantageous in environments with diverse and unpredictable image qualities, leveraging the individual strengths of each model. Table 9 shows the training time and the model size for every model apart.

In the deployment of the *DM-AHR* model, its primary function is to preprocess dermoscopic images by removing hair before classifying the type of skin cancer. While the model achieves high accuracy, there are instances of misidentification, albeit at a very low rate. These rare errors typically involve the misclassification between hair artifacts and certain types of skin lesions. However, it is important to highlight that in a clinical setting, the final diagnosis is corroborated by expert dermatologists. The *DM-AHR* model significantly aids this process by enhancing image clarity, thus making the detection of skin anomalies more straightforward and rapid. This not only assists dermatologists in making more informed decisions but also contributes to quicker diagnostic processes, potentially speeding up the initiation of appropriate treatment plans and improving patient outcomes. The integration of the *DM-AHR* model into clinical workflows is therefore seen not just as a technological enhancement but as a vital tool in the fight against skin cancer, offering a blend of precision and speed that is crucial for effective medical care.

**Table 9.** Model size and training time for different models.

Model	Model Size (MB)	Training Time
resnext101_32x8d	1340.72	403 m 32 s
maxvit_t	834.43	196 m 23 s
SwinTransformer	911.96	418 m 49 s

The collective findings from this study highlight the critical role of preprocessing in medical image analysis. Introduced herein, the *DM-AHR* model has demonstrated notable efficiency in mitigating a common form of medical image degradation, underscoring its potential utility as a preprocessing tool not only for dermatological assessments but also across various medical imaging applications. However, it is important to note that the computational demands of the *DM-AHR* model are considerable. Specifically, processing an image of size  $256 \times 256$  pixels through 2000 steps requires approximately 2 min on a single GPU RTX 8000 with 48 GB of VRAM. This substantial resource requirement is justified by the model's capability to effectively remove hair from dermoscopic images, which is a crucial step in enhancing the visibility and detectability of skin cancer lesions. Skin cancer, as a potentially lethal disease, necessitates the highest accuracy in diagnostic imaging processes.

The *DM-AHR* model, therefore, presents a trade-off between computational expense and clinical value. The increased processing time and resource use are offset by the model's ability to deliver clearer, more interpretable images that are essential for accurate diagnosis and subsequent treatment planning. This balance underscores the model's practical relevance, particularly in settings where diagnostic precision is paramount.

## 5. Conclusions

In conclusion, the *DM-AHR* model presents a significant contribution to the topic of dermatological image processing. The model, with its self-supervised learning approach,

has demonstrated remarkable proficiency in enhancing skin lesion images, thereby facilitating more accurate and efficient skin diagnostics. With skin cancer rates increasing globally, tools like *DM-AHR* can play a major role in early detection and treatment, possibly saving lives. The model's ability to effectively remove hair from dermatoscopic images addresses a significant challenge in the field, paving the way for clearer and more reliable lesion analysis. However, the high processing cost of *DM-AHR* remains its main limitation that should be addressed in future research works.

Future research should focus on integrating the *DM-AHR* model into clinical workflows to harness its diagnostic capabilities in real-world settings and extend its application to other medical imaging modalities such as ultrasound, CT, and MRI. Exploring the adaptation of diffusion-based models to address various imaging artifacts could lead to significant development in medical image analysis by improving diagnostic accuracy, expediting disease detection, and enhancing patient care. This integration promises comprehensive pre-processing tools that support a wide range of diagnostic procedures, potentially transforming patient outcomes by facilitating early disease detection and treatment through advanced technological applications.

**Author Contributions:** Conceptualization, B.B. and A.M.A.; formal analysis, A.M.A. and B.B.; investigation, A.M.A. and B.B.; methodology, A.M.A. and B.B.; project administration, B.B. and A.K.; software, A.M.A.; supervision, B.B. and A.K.; validation, A.M.A.; visualization, A.M.A.; writing—original draft, A.M.A. and B.B.; writing—review and editing, B.B., A.K., W.B. and A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Prince Sultan University.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data Availability Statements at <https://www.kaggle.com/datasets/riotulab/skin-cancer-hair-removal> (accessed on 7 August 2024).

**Acknowledgments:** The authors thank Prince Sultan University for the support and funding in conducting this research.

**Conflicts of Interest:** The authors declare no conflicts of interest associated with this publication.

## References

1. World Health Organization. Ultraviolet (UV) Radiation and Skin Cancer. 2023. Available online: [https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-\(uv\)-radiation-and-skin-cancer](https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv)-radiation-and-skin-cancer) (accessed on 25 December 2023).
2. Mahmood, T.; Saba, T.; Rehman, A.; Alamri, F.S. Harnessing the power of radiomics and deep learning for improved breast cancer diagnosis with multiparametric breast mammography. *Expert Syst. Appl.* **2024**, *249*, 123747. [CrossRef]
3. Omar Bappi, J.; Rony, M.A.T.; Shariful Islam, M.; Alshathri, S.; El-Shafai, W. A Novel Deep Learning Approach for Accurate Cancer Type and Subtype Identification. *IEEE Access* **2024**, *12*, 94116–94134. [CrossRef]
4. Abu Tareq Rony, M.; Shariful Islam, M.; Sultan, T.; Alshathri, S.; El-Shafai, W. MediGPT: Exploring Potentials of Conventional and Large Language Models on Medical Data. *IEEE Access* **2024**, *12*, 103473–103487. [CrossRef]
5. Soleimani, M.; Harooni, A.; Erfani, N.; Khan, A.R.; Saba, T.; Bahaj, S.A. Classification of cancer types based on microRNA expression using a hybrid radial basis function and particle swarm optimization algorithm. *Microsc. Res. Tech.* **2024**, *87*, 1052–1062. [CrossRef] [PubMed]
6. Emara, H.M.; El-Shafai, W.; Algarni, A.D.; Soliman, N.F.; El-Samie, F.E.A. A Hybrid Compressive Sensing and Classification Approach for Dynamic Storage Management of Vital Biomedical Signals. *IEEE Access* **2023**, *11*, 108126–108151. [CrossRef]
7. Vestergaard, M.; Macaskill, P.; Holt, P.; Menzies, S. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting. *Br. J. Dermatol.* **2008**, *159*, 669–676. [CrossRef] [PubMed]
8. Hammad, M.; Pławiak, P.; ElAffendi, M.; El-Latif, A.A.A.; Latif, A.A.A. Enhanced deep learning approach for accurate eczema and psoriasis skin detection. *Sensors* **2023**, *23*, 7295. [CrossRef]
9. Alyami, J.; Rehman, A.; Sadad, T.; Alruwaythi, M.; Saba, T.; Bahaj, S.A. Automatic skin lesions detection from images through microscopic hybrid features set and machine learning classifiers. *Microsc. Res. Tech.* **2022**, *85*, 3600–3607. [CrossRef]
10. Nawaz, M.; Mehmood, Z.; Nazir, T.; Naqvi, R.A.; Rehman, A.; Iqbal, M.; Saba, T. Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microsc. Res. Tech.* **2022**, *85*, 339–351. [CrossRef]



11. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.
12. Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 2256–2265.
13. Song, Y.; Sohl-Dickstein, J.; Kingma, D.P.; Kumar, A.; Ermon, S.; Poole, B. Score-based generative modeling through stochastic differential equations. *arXiv* **2020**, arXiv:2011.13456.
14. Farooq, M.A.; Yao, W.; Schukat, M.; Little, M.A.; Corcoran, P. Derm-t2im: Harnessing synthetic skin lesion data via stable diffusion models for enhanced skin disease classification using vit and cnn. *arXiv* **2024**, arXiv:2401.05159.
15. Abuzaghlh, O.; Barkana, B.D.; Faezipour, M. Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention. In Proceedings of the IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014, Farmingdale, NY, USA, 2 May 2014; pp. 1–6.
16. Maglogiannis, I.; Doukas, C.N. Overview of advanced computer vision systems for skin lesions characterization. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 721–733. [\[CrossRef\]](#)
17. Jing, W.; Wang, S.; Zhang, W.; Li, C. Reconstruction of Neural Radiance Fields With Vivid Scenes in the Metaverse. *IEEE Trans. Consum. Electron.* **2023**, *70*, 3222–3231. [\[CrossRef\]](#)
18. Bao, Q.; Liu, Y.; Gang, B.; Yang, W.; Liao, Q. S2Net: Shadow Mask-Based Semantic-Aware Network for Single-Image Shadow Removal. *IEEE Trans. Consum. Electron.* **2022**, *68*, 209–220. [\[CrossRef\]](#)
19. Ji, Z.; Zheng, H.; Zhang, Z.; Ye, Q.; Zhao, Y.; Xu, M. Multi-Scale Interaction Network for Low-Light Stereo Image Enhancement. *IEEE Trans. Consum. Electron.* **2023**, *70*, 3626–3634. [\[CrossRef\]](#)
20. Anand, V.; Gupta, S.; Nayak, S.R.; Koundal, D.; Prakash, D.; Verma, K. An automated deep learning models for classification of skin disease using Dermoscopy images: A comprehensive study. *Multimed. Tools Appl.* **2022**, *81*, 37379–37401. [\[CrossRef\]](#)
21. Li, W.; Raj, A.N.J.; Tjahjadi, T.; Zhuang, Z. Digital hair removal by deep learning for skin lesion segmentation. *Pattern Recognit.* **2021**, *117*, 107994. [\[CrossRef\]](#)
22. Guo, Z.; Chen, J.; He, T.; Wang, W.; Abbas, H.; Lv, Z. DS-CNN: Dual-Stream Convolutional Neural Networks-Based Heart Sound Classification for Wearable Devices. *IEEE Trans. Consum. Electron.* **2023**, *69*, 1186–1194. [\[CrossRef\]](#)
23. Kim, D.; Hong, B.W. Unsupervised feature elimination via generative adversarial networks: application to hair removal in melanoma classification. *IEEE Access* **2021**, *9*, 42610–42620. [\[CrossRef\]](#)
24. Dong, L.; Zhou, Y.; Jiang, J. Dual-Clustered Conditioning towards GAN-based Diverse Image Generation. *IEEE Trans. Consum. Electron.* **2024**, *70*, 2817–2825. [\[CrossRef\]](#)
25. He, Y.; Jin, X.; Jiang, Q.; Cheng, Z.; Wang, P.; Zhou, W. LKAT-GAN: A GAN for Thermal Infrared Image Colorization Based on Large Kernel and AttentionUNet-Transformer. *IEEE Trans. Consum. Electron.* **2023**, *69*, 478–489. [\[CrossRef\]](#)
26. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [\[CrossRef\]](#)
27. Delibasis, K.; Moutselos, K.; Vorigiazidou, E.; Maglogiannis, I. Automated hair removal in dermoscopy images using shallow and deep learning neural architectures. *Comput. Methods Programs Biomed. Update* **2023**, *4*, 100109. [\[CrossRef\]](#)
28. Abbas, Q.; Celebi, M.E.; García, I.F. Hair removal methods: A comparative study for dermoscopy images. *Biomed. Signal Process. Control* **2011**, *6*, 395–404. [\[CrossRef\]](#)
29. Saharia, C.; Ho, J.; Chan, W.; Salimans, T.; Fleet, D.J.; Norouzi, M. Image super-resolution via iterative refinement. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 4713–4726. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015, Proceedings, Part III* 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
32. Tschandl, P.; Rosendahl, C.; Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **2018**, *5*, 180161. [\[CrossRef\]](#)
33. Digital Hair Dataset. Available online: <https://www.kaggle.com/datasets/weilizai/digital-hair-dataset> (accessed on 17 December 2023).
34. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 586–595.
35. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.
36. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1833–1844.
37. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10684–10695.
38. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.



39. Tu, Z.; Talebi, H.; Zhang, H.; Yang, F.; Milanfar, P.; Bovik, A.; Li, Y. Maxvit: Multi-axis vision transformer. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 459–479.
40. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.