Hair removal in dermoscopic images using SegFormer and inpainting*

Fayzulla Ollamberganov[1,*,†], Rustam Utemuratov[2,†], Rakhim Shikhiyev[2,†] and Bayram Absametov[2,†]

[1] *Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan*
[2] *Karakalpak State University named after Berdakh, Nukus, Uzbekistan*

Abstract

Dermoscopic images are widely used in the early detection and differential diagnosis of skin tumors. However, hair artifacts found in these images cause significant problems in the diagnostic process. As a result of hair covering pigment networks, lesion boundaries, and color gradients, both AI-based models and physicians visual assessment processes may lose reliability. In this study, a two-stage deep learning-based approach was proposed for hair removal from dermoscopic images. In the first stage, hair and shadows are segmented using the SegFormerWithDropout architecture. The dropout mechanism enhances the models stability, while the shadow-aware extended channel allows for masking of shadows as well. In the second stage, a hybrid MAT+LaMa inpainting generator is employed to restore masked areas based on global and local context, reconstructing clinical structures without artifacts. The proposed approach was tested on the HAM10000 dataset and achieved high performance metrics, including Dice (0.967), IoU (0.935), and SSIM (0.915). The obtained results demonstrated that the proposed method can potentially improve the overall effectiveness of diagnostic systems by enhancing the quality of dermoscopic images.

Keywords

Dermoscopic images, hair artifacts, SegFormer, inpainting, deep learning.

1. Introduction

The dermoscopic images play a crucial role in medical diagnostics, particularly in the early detection and differential diagnosis of skin tumors. The global increase in melanoma and other skin diseases further emphasizes the relevance of this process [1,2]. The dermoscopy method enables physicians to identify differences in color, texture, and structure that are invisible to the naked eye. However, image quality can be significantly compromised due to various artifacts such as hair, shadows, surgical marks, and measurement lines [3]. Hair artifacts are one of the most common problems in dermoscopic images, as they obscure lesion boundaries, pigment networks, or color variations. This situation leads to a decrease in accuracy not only during medical examinations but also in automated segmentation and classification systems. Recent studies have shown that the presence of hair significantly reduces the accuracy of AI models [4]. It has also been found that in some cases, sparse hair can help make the model more stable, while dense hair exacerbates diagnostic errors. For this reason, effective removal of hair in dermoscopic images is considered a pressing task today.

* *IVUS 2025: Information Society and University Studies 2025, May 15, Kaunas, Lithuania*
* Corresponding author.
† These authors contributed equally.
fayzulla0804@gmail.com (F. Ollamberganov); u_rustam@karsu.uz (R. Utemuratov); raximm82@mail.ru (R. Shikhiyev); absametovbayram@gmail.com (B. Absametov)
0000-0001-6450-0467 (F. Ollamberganov); 0009-0008-9409-3020 (R. Utemuratov); 0000-0003-2207-9357 (R. Shikhiyev)

In the past decade, various approaches have been proposed. Classical methods include morphological operators, filtration, interpolation, and inpainting techniques based on PDE (partial differential equation) [5]. These methods are simple and fast, but not robust for hair of different colors and thicknesses [6]. CNN approaches based on deep learning, such as U-Net and its modifications, have been widely used in recent years for hair removal from dermoscopic images [7]. For instance, DullRazor and subsequent algorithms operate through a segmentation + inpainting pipeline, but are limited in preserving complex shadows and boundary structures. Transformer approaches: Recently proposed models based on the SegFormer architecture stand out for their ability to account for long-range connections and demonstrate high accuracy even on small datasets [8]. However, this approach only provides a segmentation mask and does not reconstruct the texture and color of the lesion beneath the hair. The aim of this study is to develop a new two-stage deep learning algorithm for hair removal on dermoscopic images. The proposed method is distinguished by the following features:

1. Clear masking of hair and shadows through transformer-based segmentation (SegFormer + attention mechanisms).
2. Reproduction of clinical signs under the hair using diffusion or GAN-based inpainting, depending on the context, without artifacts.
3. Shadow-aware extended masking, that is, taking into account not only the hair, but also its shadow.

As a result, the proposed algorithm not only effectively removes hair, but also preserves the diagnostic features of the lesions, the pigment network, border, texture, and color gradients.

2. Related works

In recent years, three trends have emerged in the field of hair artifact removal in dermoscopic images: (i) deep learning-based segmentation + inpainting pipelines, (ii) studies focused on measuring the direct impact of hair presence on AI results, and (iii) a shift from CNNs to transformers, particularly the SegFormer family and transformer-inpainting solutions. For instance, a combination of U-Net-based hair segmentation and inpainting on free masks has been proposed, showing improved lesion segmentation results; their approach is suitable for real dermoscopy, but texture consistency may not always be perfectly maintained in thin-boundary structures [7]. Additionally, the encoder-decoder CNN stabilizes the restoration stage by separating the hair mask and using a combined loss of L1+TV+SSIM. While the solution is practical, its generalization is noted to be limited for wide shadow ranges and various hair morphologies (white/thin) [9]. In study [10], a “hairless” reconstruction is performed using an unpaired VAE, reducing the need for extensive annotation. However, the VAE decoder may sometimes cause oversmoothing, potentially erasing fine structures. Concurrently, hybrid approaches combining classical and deep learning methods are emerging: they integrate surface and deep models, improving cleaning through morphological post-processing after segmentation. This approach is easily integrated into the pipeline, but the issue of shadow-awareness remains unresolved [11]. White and black hair, which lift the artifact in a complex form, even markers, are identified and inpainted; high indicators are noted in comparison with many methods, but the risk of over-inpainting is mentioned in complex texture-boundaries [12].

One of the important directions in recent years is to more accurately measure the impact of hair on AI systems. [4] experimentally demonstrates that the presence of hair on more than 10,000 images reduces classification/segmentation results and offers metrics such as Inf1/Inf2; this reinforces the need for evaluation not only by “Dice/IoU” but also by downstream tasks. In a scenario closely resembling clinical symptoms, the study [13] demonstrates that the pipeline proposed for Mpox lesions, enhanced with a hair-cleaning stage, improved the diagnostic stability of the U-Net. This integrative approach emphasizes that hair should be viewed as a task-driven component rather than simply as pre-processing.

In model architectures, there is a noticeable shift from CNNs to transformers. SegFormer achieves a balance between efficiency and accuracy in segmentation with its hierarchical MiT encoder + lightweight MLP decoder design. The absence of a large receptive field and positional encoding requirements improves adaptability to various resolutions [14]. In medical segmentation, evaluating SegFormer against U-Net across different modalities demonstrates that the transformer is indeed competitive, providing strong empirical evidence for its application in dermoscopy [8]. A promising approach to addressing the transformer-specific edge/boundary problem is improved edge segmentation using neighborhood attention and patch-merging in colonic polyps. These principles could potentially be applied to hair segmentation masks as well [15].

The second half of the cleaning process is inpainting. In addition to CNN-based approaches, two directions have been leading in recent years: LaMa [16] with a high receptive field and transformer-based inpainting. LaMa, using Fast Fourier Convolutions (FFC), reliably fills even large masks based on global context; in dermoscopy, this is ideal for broad hair shadows or hair attached to lesion edges, as the global spectral focus maintains structural coherence. Among transformers, MAT [17] is recognized as the first high-resolution mask-aware transformer inpainting system: its ability to generate realistic, diverse content in large holes while maintaining stylistic consistency provides an advantage in restoring dermoscopic textures. Additionally, between 2023 and 2025, various versions of transformer-inpainting emerged, demonstrating a focus on preserving multi-dimensional structure and restoring fine details; this aligns with the "shadow-aware" inpainting design trend in dermoscopy.

Another interesting trend is synthetic hair benchmarks. In [18], an artificial hair dermoscopic dataset and DL model are proposed, serving to stabilize evaluation by multiplying ground truth pairs; this creates a solid foundation for self-supervised/synthetic overlay strategies. However, this work also does not ultimately include a complete artifact removal stage, meaning the "mask-only" limitation remains; it is precisely here that the need for integrated diffusion/transformer-inpainting arises.

3. Methodology

This study proposes a two-stage algorithm for hair removal in dermoscopic images. The approach consists of the following main components: (1) a transformer-based segmentation module for detecting hair and their shadows, (2) an inpainting generator for contextually reconstructing masked areas without artifacts.

3.1. Datasets

The experiments were conducted on the HAM10000 (Human Against Machine with 10,000 training images) dataset. This collection consists of 10,015 dermoscopic images classified according to 7 different clinical diagnoses: melanocytic nevi (NV), melanoma (MEL), benign keratosis-like lesions (BKL), basal cell carcinoma (BCC), actinic keratosis and intraepithelial carcinoma (AKIEC), dermatofibroma (DF), and vascular lesions (VASC).

The dataset contains artifacts of various sizes and quality. Table 1 shows the distribution of the HAM10000 dataset based on the presence or absence of hair. Statistical analysis reveals that hair artifacts are present in all classes, with the majority (≈80%) of images in the nevus (NV) class being covered with hair. This justifies the necessity of removing artifacts during the diagnostic process.

Table 1
Distribution of the HAM10000 dataset by hair artifacts

Target	**Train Validation (n=8008)**			**Test (n=1998)**		
	Non-hair occlusions	Hair occlusions	Hair occlusions	Non-hair occlusions	Hair occlusions	Hair occlusions

	(n=4916)	Few (n=1437)	Many (n=1655)	(n=1311)	Few (n=416)	Many (n=271)
AKIEC	173 (3.5%)	26 (1.8%)	53 (3.2%)	53 (4.0%)	14 (3.3%)	8 (3.0%)
BCC	300 (6.1%)	31 (2.2%)	85 (5.2%)	69 (5.3%)	15 (3.6%)	14 (5.2%)
BKL	657 (13.4%)	66 (4.6%)	146 (8.8%)	163 (12.4%)	34 (8.2%)	31 (11.4%)
DF	54 (1.1%)	18 (1.3%)	20 (1.2%)	14 (1.1%)	5 (1.2%)	4 (1.5%)
MEL	641 (13.0%)	133 (9.2%)	166 (10.0%)	133 (10.1%)	29 (7.0%)	10 (3.7%)
NV	2998 (61.0%)	1158 (80.6%)	1173 (70.9%)	853 (65.1%)	317 (76.2%)	200 (73.8%)
VASC	93 (1.9%)	30 (2.0%)	50 (3.0%)	26 (2.0%)	20 (5.0%)	4 (1.4%)

3.2. Segmentation method architecture

The SegFormerWithDropout architecture was employed to identify hair and its shadows. This model comprises a Mix Transformer (MiT) encoder and a multilayer perceptron (MLP) decoder, enabling the simultaneous learning of local and global features in dermoscopic images. The input image is first divided into small, partially overlapping patches through overlap patch embedding, which helps preserve local context similar to classical convolutional filters. Each patch then passes through four sequential transformer blocks, each consisting of an efficient self-attention mechanism and a Mix-Feed Forward Network (Mix-FFN) combination. The attention mechanism effectively captures global structures in the dermoscopic image by considering long-range connections, while the Mix-FFN represents deeper features through nonlinear transformations. After each stage, spatial dimensions are reduced and the number of channels is increased using overlap patch merging, which allows the model to generate representations of the image at different levels. At the encoder output, feature maps of various resolutions are obtained and transmitted to the decoder section.

The decoder has a very lightweight architecture, consolidating feature maps from all stages of the encoder into a single dimension through MLP. These maps are then restored to their original spatial resolution using an upsampling process and combined. Finally, a segmentation map is generated, in which hair and its shadows are represented as precise masks. Additionally, dropout ($p=0.3$) was applied to the model to prevent overfitting on small datasets. Furthermore, the inclusion of a shadow-aware channel resulted in masking not only the hair itself but also its shadows, thereby improving accuracy in real clinical settings.

The main advantage of the SegFormer architecture is that, unlike classical U-Net models, it does not require positional encoding and can therefore operate stably on images of various resolutions. The efficient design of the encoder allows for feature extraction while fully considering the global context, while the decoder, being simple and computationally light, enables near real-time operation. These aspects make the SegFormerWithDropout architecture particularly effective for hair segmentation in dermoscopic images.

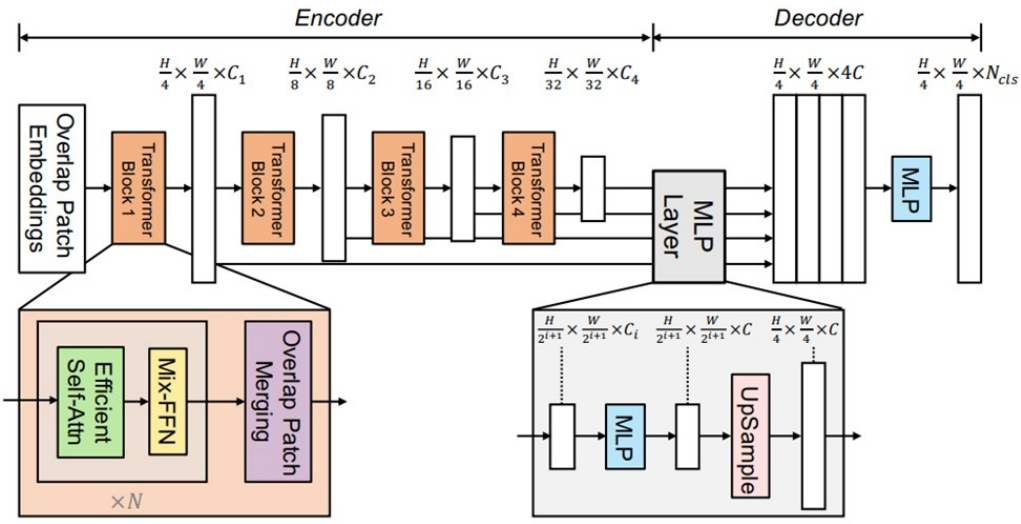

Figure 1: Block diagram of the SegformerWithDropout architecture, highlighting the MiT-B2 encoder, decoder, and the added dropout layer in the segmentation head

3.3. Proposed Hybrid inpainting model

After masking the hair through segmentation, the main objective is to reconstruct the masked areas without disrupting the clinical structures. For this task, a hybrid inpainting generator was employed, based on Mask-Aware Transformer (MAT) blocks and LaMa (Large Mask Inpainting with Fourier Convolutions). The MAT blocks adapt the self-attention mechanism in masked areas, taking into account both global and local context in the image. This allows the reconstructed area to be harmonized with the surrounding color, texture, and boundary features. The LaMa module enables high-precision restoration even in extensively masked areas using Fourier convolutions, which is particularly crucial for preserving the pigment network and lesion boundaries in dermoscopic images.

Formally, the input image is denoted by $I \in R^{H \times W \times 3}$ and the mask obtained at the segmentation stage by $M \in \{0,1\}^{H \times W}$. The masked image is generated as follows:

$$I_M = I \odot (1 - M), \tag{1}$$

where ⊙ denotes element-wise multiplication. The inpainting generator G_θ with parameter set θ takes the masked image and mask as input to produce a restored image:

$$\hat{I} = G_\theta(I_M, M). \tag{2}$$

The generator output is evaluated by the discriminator D_ϕ, which is used to determine the difference between the real image I and the reconstructed image $\hat{I}$.

The overall loss function consists of a combination of several components:

1. Reconstruction loss (L1):

$$L_{rec} = ||M \odot (\hat{I} - I)||_1, \tag{3}$$

it minimizes pixel-level errors in masked areas.

2. Perceptual loss:

$$L_{perc} = \sum_l ||\phi_l(\hat{I}) - \phi_l(I)||_2^2, \tag{4}$$

where ϕ_l – represents the feature maps in the l-th layer of VGG19.

3. Style loss:

$$L_{style} = \sum_{l} |G_l(\hat{I}) - G_l(I)|_F^2, \quad (5)$$

where $G_l(\cdot)$ – Gram matrix, $||\cdot||_F$ and Frobenius is the norm.

4. Adversarial loss:

$$L_{adv} = E\left[\log D_\phi(I)\right] + E\left[\log\left(1 - D_\phi(\hat{I})\right)\right]. \quad (6)$$

5. Total loss:

$$L = \lambda_1 L_{rec} + \lambda_2 L_{perc} + \lambda_3 L_{style} + \lambda_4 L_{adv}. \quad (7)$$

This combination improves texture and color consistency while minimizing artifacts. As a result, after removing hair and shadows from dermoscopic images, the inpainting generator reconstructs the underlying clinical structures in a comprehensive and diagnostically significant manner.

3.4. Metrics

The proposed algorithm used several different categories of metrics to evaluate its effectiveness. They are aimed at assessing the accuracy of segmentation, the quality of inpainting, and the results in downstream clinical tasks.

Segmentation metrics were used to measure the correctly separated level of the mask.

The Dice coefficient (DSC) describes how well two sets match and is calculated as follows:

$$Dice = \frac{2|P \cap G|}{|P| + |G|}, \quad (8)$$

where P is the predicted mask by the model, and G is the actual mask.

Intersection over Union (IoU) represents the predicted and actual coverage of masks:

$$IoU = \frac{|P \cap G|}{|P \cup G|}. \quad (9)$$

The Boundary F-score is used to assess the accuracy of mask boundaries. It is defined based on precision and recall as follows:

$$BF = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}.$$

Inpainting metrics are used to assess the visual and structural quality of the restored image.

PSNR (Peak Signal-to-Noise Ratio) is a criterion for determining the noise level and is calculated according to the following formula:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right), \quad (10)$$

where MAX_I is the maximum pixel value in the image, MSE is the root mean square error.

SSIM (Structural Similarity Index) calculates the combination of structure, light, and contrast between images:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (11)$$

where μ_x, μ_y - mean values, σ_x^2, σ_y^2 - variances, σ_{xy} - covariance, C_1, C_2 -stabilizing constants.

LPIPS (Learned Perceptual Image Patch Similarity) is a perceptual similarity criterion based on deep neural networks, with lower values indicating better results.

FID (Fréchet Inception Distance) measures the distance between the distributions of real and recovered images:

$$FID = \left\| \mu_r - \mu_g \right\|^2 + Tr\left(\Sigma_r + \Sigma_g - 2\left(\Sigma_r \Sigma_g \right)^{\frac{1}{2}} \right), \quad (12)$$

where μ_r, Σ_r are the average and covariance values of real images, and μ_g, Σ_g are statistics related to reconstructed images.

4. Results and Discussion

The effectiveness of the two-stage algorithm proposed in this section is evaluated based on experiments conducted on the HAM10000 dataset. During the experiments, the model's segmentation accuracy, inpainting quality metrics, and impact on downstream tasks were analyzed. The experiments were performed using the PyTorch 2.0 framework on an NVIDIA RTX 3060 GPU (24GB VRAM) computing environment. For the training process, the AdamW optimizer $\left(lr = 1 \times 10^{-4} \right)$ was employed, with a batch size of 32, and training was carried out for 50 epochs. To prevent overfitting, early stopping and a learning rate scheduler were implemented. The dataset was split into 80% for training and validation, and 20% for testing. Various augmentation techniques were applied to the training set, including adjustments to brightness and contrast, addition of Gaussian noise, and introduction of additional hair artifacts through synthetic hair overlay.

This approach contributed to the stability of the model to various artifacts in real clinical conditions. Several indicators were used for the evaluation. At the segmentation stage, the Dice coefficient, IoU, and the Boundary F-score were calculated. To assess the quality of inpainting, the PSNR, SSIM, LPIPS, and FID criteria were used. In addition, the results of lesion segmentation (Dice, IoU) and classification (AUC, Accuracy) were also compared to measure the influence of hair removal on clinical tasks. The effectiveness of the proposed approach was also compared with other existing methods: DullRazor, U-Net+Inpainting, DeepLabV3+, SegFormerWithDropout (mask-only). With the help of these comparisons, the advantages of our algorithm were determined.

4.1. Hair segmentation results

The proposed SegFormerWithDropout (shadow-aware) model achieved high accuracy in the segmentation of hair and shadows in the HAM10000 dataset. Assessment was carried out based on such criteria as the Dice coefficient, IoU, and the Boundary F-score. Several basic approaches were selected for comparison: the classic DullRazor algorithm, U-Net+Inpainting, DeepLabV3+, and SegFormer (mask-only). Among these methods, the highest result was observed in this approach.

Table 2
Comparison of hair segmentation results (in the HAM10000 dataset)

Model	Dice	IoU	Boundary F-score
DullRazor (classic)[4]	0.781	0.652	0.610
U-Net + Inpainting[8]	0.871	0.794	0.763
DeepLabV3+[4]	0.884	0.808	0.779
SegFormer (mask-only)[14]	0.932	0.891	0.853
Proposed method	0.967	0.935	0.908

As can be seen from the results, classical algorithms such as DullRazor are not stable with respect to the influence of different hair and shadows. While CNN-based U-Net and DeepLabV3+ have provided significant improvements, they do not adequately consider the global context. A clear improvement was observed in the SegFormer model based on the Transformer architecture, but limited to mask-level segmentation. Our shadow-aware SegFormerWithDropout model achieved the highest results thanks to an additional shadow channel and a stabilized drop-off mechanism.

4.2. Inpainting results

After masking the hair and shadows at the segmentation stage, the inpainting generator performs the task of reconstructing the masked areas based on the context. The results of the proposed hybrid MAT+LaMa inpainting generator were evaluated according to several criteria: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), LPIPS (Learned Perceptual Image Patch Similarity), and FID (Fréchet Inception Distance).

Table 3
Comparison of inpainting results (on the HAM10000 dataset)

Model	PSNR	SSIM	LPIPS	FID
DullRazor (classic)[4]	24.12	0.811	0.321	89.4
U-Net + PDE inpainting[8]	26.54	0.843	0.280	73.1
DeepLabV3+ + CNN inpainting [4]	27.02	0.857	0.251	66.7
SegFormer (mask-only) + LaMa [14]	28.93	0.889	0.212	51.6
Proposed method (MAT+LaMa)	**30.41**	**0.915**	**0.184**	**42.3**

As can be seen from the results, methods based on classical interpolation were unable to restore the basic structures during hair removal, resulting in low SSIM and PSNR values. Although CNN-based inpainting methods (U-Net, DeepLabV3+) yield better results, they are not capable of fully preserving textural harmony. The combination of SegFormer segmentation and LaMa inpainting yielded strong results, but focused only on global structures. The proposed MAT+LaMa hybrid approach made it possible to restore not only global, but also local texture harmony, significantly increasing the PSNR and SSIM indicators and reducing the LPIPS and FID values.

4.3. Discussion

The proposed approach also has some limitations. For example, in images with very dense and thick hair, the inpainting stage can sometimes lead to a slight smoothing of textures. Additionally, since the model training primarily used the HAM10000 dataset, further testing in a broader database is required for application in wider clinical settings. From this perspective, future work plans include utilizing more multimodal data, further stabilizing the model based on images

obtained under various lighting conditions, and applying more powerful inpainting mechanisms based on diffusion models.

5. Conclusion

This study proposes a novel two-stage deep learning-based algorithm for removing hair artifacts in dermoscopic images. In the first stage of the approach, hair and its shadows were segmented with high accuracy using the SegFormerWithDropout architecture. The dropout mechanism enhanced the model's stability, while the shadow-aware extended channel enabled correct masking of shadows encountered in real clinical settings. In the second stage, the masked areas were reconstructed based on global and local context using the MAT+LaMa hybrid inpainting generator, restoring the pigment network, lesion boundaries, and color gradients without artifacts.

At the same time, the method has certain limitations. In images with very dense hair, the inpainting stage can sometimes lead to a slight smoothing of clinical structures. Future research may focus on applying more sophisticated inpainting mechanisms based on the diffusion model, utilizing multimodal data, and testing the model under various lighting conditions.

Overall, the proposed approach significantly improves upon previously existing methods for hair removal in dermoscopic images, providing high-quality visual data for clinical diagnostics. The results of this study are of significant importance not only scientifically but also practically, and will serve as a solid foundation for making future artificial intelligence-assisted systems for early detection of skin diseases even more effective.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] S. Mocellin and C. R. Rossi, “The Melanoma Molecular Map Project,” Melanoma Res., vol. 18, no. 3, 2008, doi: 10.1097/CMR.0b013e328300c50b.

[2] G. Randi *et al.*, “The European Cancer Information System: exploring linkages between indoor radon concentrations and data on cancer burden,” *J. Eur. Radon Assoc.*, pp. 1–8, 2022, doi: 10.35815/radon.v3.7607.

[3] J. K. Winkler et al., “Association between different scale bars in dermoscopic images and diagnostic performance of a market-approved deep learning convolutional neural network for melanoma recognition,” Eur. J. Cancer, vol. 145, pp. 146–154, 2021, doi: 10.1016/j.ejca.2020.12.010.

[4] Z. Wang et al., “Influence of hair presence on dermoscopic image analysis by AI in skin lesion diagnosis,” Comput. Biol. Med., vol. 183, Dec. 2024, doi: 10.1016/j.compbiomed.2024.109335.

[5] J. A. A. Salido and C. Ruiz, “Using morphological operators and inpainting for hair removal in dermoscopic images,” in Proceedings of the Computer Graphics International Conference, in CGI ’17. New York, NY, USA: Association for Computing Machinery, 2017. doi: 10.1145/3095140.3095142.

[6] P. A. Somnathe and P. P. Gumaste, “A Review of Existing Hair Removal Methods in Dermoscopic Images,” IOSR J. Electron. Commun. Eng., pp. 73-76.

[7] W. Li, A. N. Joseph Raj, T. Tjahjadi, and Z. Zhuang, “Digital hair removal by deep learning for skin lesion segmentation,” Pattern Recognit., vol. 117, p. 107994, 2021, doi: 10.1016/j.patcog.2021.107994.

[8] T. Sourget, S. N. Hasany, F. Mériaudeau, and C. Petitjean, “Can SegFormer be a True Competitor to U-Net for Medical Image Segmentation?,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 14122 LNCS, pp. 111–118, 2024, doi: 10.1007/978-3-031-48593-0_8.

[9] L. Talavera-Martinez, P. Bibiloni, and M. Gonzalez-Hidalgo, “Hair Segmentation and Removal in Dermoscopic Images Using Deep Learning,” IEEE Access, vol. 9, pp. 2694–2704,

2021, doi: 10.1109/ACCESS.2020.3047258.
[10] D. Bardou, H. Bouaziz, L. Lv, and T. Zhang, “Hair removal in dermoscopy images using variational autoencoders,” Ski. Res. Technol., vol. 28, no. 3, pp. 445–454, 2022, doi: 10.1111/srt.13145.
[11] K. Delibasis, K. Moutselos, E. Vorgiazidou, and I. Maglogiannis, “Automated hair removal in dermoscopy images using shallow and deep learning neural architectures,” Comput. Methods Programs Biomed. Updat., vol. 4, no. June, p. 100109, 2023, doi: 10.1016/j.cmpbup.2023.100109.
[12] R. Kasmi et al., “SharpRazor: Automatic removal of hair and ruler marks from dermoscopy images,” Ski. Res. Technol., vol. 29, no. 4, pp. 1–12, 2023, doi: 10.1111/srt.13203.
[13] E. M. Onyema et al., “Deep learning model for hair artifact removal and Mpox skin lesion analysis and detection,” Sci. Rep., vol. 15, no. 1, pp. 1–18, 2025, doi: 10.1038/s41598-025-05324-2.
[14] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, “SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers,” Adv. Neural Inf. Process. Syst., vol. 15, no. NeurIPS, pp. 12077–12090, 2021.
[15] D. Liu, C. Lu, H. Sun, and S. Gao, “NA-segformer: A multi-level transformer model based on neighborhood attention for colonoscopic polyp segmentation,” Sci. Rep., vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-74123-y.
[16] R. Suvorov et al., “Resolution-robust Large Mask Inpainting with Fourier Convolutions,” Proc. - 2022 IEEE/CVF Winter Conf. Appl. Comput. Vision, WACV 2022, pp. 3172–3182, 2022, doi: 10.1109/WACV51458.2022.00323.
[17] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang, and J. Jia, “MAT: Mask-Aware Transformer for Large Hole Image Inpainting,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2022-June, pp. 10748–10758, 2022, doi: 10.1109/CVPR52688.2022.01049.
[18] L. Jütte, H. Patel, and B. Roth, “Advancing dermoscopy through a synthetic hair benchmark dataset and deep learning-based hair removal,” J. Biomed. Opt., vol. 29, no. 11, pp. 1–17, 2024, doi: 10.1117/1.jbo.29.11.116003.