

Original Research

Received 11 June 2026

Revised 20 July 2026

Accepted 03 Aug 2026

Natural Resources for Human
Health 2026, Vol.6(15s):1621–1638

KEYWORDS:

Applied Artificial Intelligence;
Medical Image Processing, Deep
Learning, Generative Adversarial
Networks (GANs), Skin Cancer
Analysis, Healthcare AI.

Performance Analysis of Deep Learning Models for Hair Detection and Removal in Dermoscopic Images

Nutan Pathania^{1*}, Aman Kumar Sharma¹

^{1*}Department of Computer Science, Himachal Pradesh University, Shimla, India, Orcid
Id: 0009-0002-4507-6458

¹Department of Computer Science, Himachal Pradesh University, Shimla, India, Orcid
Id: 0009-0009-6815-2602

Corresponding Author's email id: pathaniatanu3@gmail.com

ABSTRACT: Removing hair from dermoscopic images is a key pre-processing step, because it directly affects the accuracy of later image analysis and diagnosis. This study implements, optimises, and evaluates three deep learning models for automated dermoscopic hair removal: a Multilayer Perceptron (MLP), a U-Net, and an Enhanced Attention Generative Adversarial Network (EA-GAN). Each model was tuned independently to its own best configuration, trained on a common dataset, and compared using six complementary image-quality metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Multiscale SSIM (MS-SSIM), and Visual Information Fidelity (VIF). U-Net performed best overall, with the highest PSNR, SSIM, MS-SSIM and VIF, which indicates superior structural preservation and perceptual fidelity. EA-GAN gave the lowest pixel-level reconstruction error but a comparatively low VIF, suggesting some loss of perceptual detail despite accurate pixel-level reconstruction. The fully connected MLP consistently trailed both encoder–decoder models, confirming its limited capacity to capture the complex spatial structure of hair artefacts. These findings indicate that convolutional encoder–decoder architectures with skip connections are best suited to dermoscopic hair removal and can meaningfully improve the reliability of fully automated skin-lesion analysis pipelines

1. INTRODUCTION

Skin cancer ranks among the most common and dangerous cancers worldwide, and melanoma is its deadliest form (Akilandasowmya et al., 2024), accounting for a disproportionate share of skin-cancer deaths, which makes early detection central to patient survival. Outcomes are good when the disease is caught early: the five-year survival rate for melanoma approaches 95%, which shows how much accurate and timely diagnosis matters. Standard diagnostic routes such as visual inspection and biopsy (Mahmoud & Soliman, 2024), rely heavily on the clinician's experience and are error-prone, especially for lesions that look very much like benign growths. Dermoscopy offers a non-invasive alternative that exposes structural features of skin lesions invisible to the naked eye, and so improves diagnostic accuracy (Kang et al., 2018). Dermoscopic images, however, are often degraded by body hair. Thin, overlapping, or skin-coloured hairs hide lesion borders and texture, and this interference is carried into automated pipelines, lowering the accuracy of later segmentation and classification (Li et al., 2021), (Lama et al., 2022). The large variation in hair density, colour, and distribution makes it harder still to design one removal technique that generalises well.

Digital hair removal (DHR) has therefore become an essential pre-processing step for dependable dermoscopic image analysis. The first approaches relied on thresholding, morphological operations, and edge detection. They are cheap to compute, but they often leave hair behind or delete clinically relevant lesion detail along with it, which distorts the image and loses information

(Nancy et al., 2023). Convolutional Neural Networks (CNNs), in contrast, can learn discriminative hair features directly from data, which makes deep-learning-based segmentation and inpainting a credible replacement for rule-based DHR.

Classical methods that do not learn from data remain bound to hand-crafted rules and fixed parameters. In practice they either over-remove, erasing fine lesion detail together with the hair, or under-remove, leaving residual hair artefacts that corrupt later analysis (Bardou et al., 2026; Kim & Hong, 2021; Shahzaib et al., 2025). Either failure model passes error on to the downstream diagnostic stages.

To make automated skin-cancer detection pipelines more reliable, this study systematically compares three deep learning approaches to dermoscopic hair removal: an MLP, a U-Net, and an EA-GAN. Each model is optimised independently to its own best configuration and evaluated on the same dataset with a consistent set of quantitative metrics that together capture reconstruction error, structural similarity, and perceptual quality. The comparison then identifies the most effective architecture for the task, with the aim of guiding the design of future automated hair-removal pipelines for dermoscopic image analysis.

The key contributions of this study are as follows:

- A like-for-like comparison of three architecturally different deep learning models (fully connected, convolutional encoder–decoder, and adversarial) for dermoscopic hair removal on one shared dataset, which fills the gap in earlier comparisons identified in the literature (Section 2.4).
- A depth-optimisation procedure specific to each architecture (Sections 3–5), so that the final comparison reflects what each model can achieve at its best.
- A six-metric evaluation protocol (Section 2.2) covering pixel-level error, structural similarity, and perceptual fidelity, which gives a fuller picture of hair-removal quality than the error-only or similarity-only evaluations used in earlier work.

2. RELATED WORK

2.1 Traditional (Non-Learning) Approaches

DullRazor (Lee et al., 1997) was the first DHR method to be widely adopted. It detected dark, coarse hairs with generalised morphological closing and filled them by bilinear interpolation; this set the template for later DHR pipelines, although it failed on thin or light-coloured hair and was validated on just five images. Later studies refined detection through other colour spaces and filters, for example CIE L^*u^*v morphological closing with median-filter inpainting (Schmid-Saugeon et al., 2003), a Radon-transform-based orientation detector called E-Shaver (Kiani & Sharafat, 2011), top-hat segmentation with PDE-based inpainting in VirtualShave (Fiorese et al., 2011), matched-filter and coherence-transport approaches (Abbas et al., 2011; Zhou et al., 2008), adaptive Canny edge detection (Toossi et al., 2013a), multiscale filtering with hysteresis thresholding (Huang et al., 2013), threshold-set decomposition combined with fast-marching inpainting (Koehoorn et al., 2015), unsupervised repair based on PDEs (Xie et al., 2009), and soft-colour morphological operators (Bibiloni et al., 2016). Taken together, these methods gradually became better at finding thinner and lighter hair and adopted more texture-aware inpainting. Even so, all of them were sensitive to manual parameter tuning, were validated on small datasets (2–460 images), and did not generalise well across the wide variability of dermoscopic hair. Table 1 traces how these techniques evolved from 1997 to 2023.

Table 1. Evolution of digital hair removal (DHR) techniques in dermoscopy, 1997–2023.

Ref.	Year	Hair Segmentation	Inpainting Method	Test Images	Key Limitation
(Lee et al., 1997)	1997	Generalised morphological closing	Bilinear interpolation	5	Fails on thin/light hair; validated on 5 images only
(Schmid-Saugeon et al., 2003)	2003	Morphological closing (CIE L^*u^*v space)	Median filtering	200	Median filter blurs lesion texture; still threshold-based
(Zhou et al., 2008)	2008	Steger line detection + curve fitting	Exemplar-based inpainting	460	Many difficult-to-tune parameters; high computational cost

(Xie et al., 2009)	2009	Morphological top-hat + thresholding	PDE-based inpainting	40	Hard thresholding fails on varying-colour hair
(Kiani & Sharafat, 2011)	2011	Prewitt edge detector + Radon transform	Colour averaging	5	Very small validation set; visible repair traces
(Fiorese et al., 2011)	2011	Top-hat operator	PDE-based inpainting	20	PDE convergence not guaranteed; limited to dark hair
(Abbas et al., 2011)	2011	Derivative-of-Gaussian matched filter	Coherence transport	100	Requires extensive per-image parameter tuning
(Toossi et al., 2013b)	2013	Adaptive Canny edge detector	Coherence transport	50	Prone to over-segmentation of high-contrast edges
(Huang et al., 2013)	2013	Multiscale matched filters + hysteresis threshold	Median filtering	20	High false-positive rate (Jaccard 31.64)
(Koehoorn et al., 2015)	2015	Multiscale skeleton + morphological operators	Fast marching	>300	Fast marching causes blurring on thick hair
(Bibiloni et al., 2016)	2017	Soft-colour top-hat morphology	Soft colour morphology	2	Only 2 test images; no inpainting quality metrics
(Attia et al., 2019)	2019	Hybrid CNN + RNN	Not proposed	>300	Weakly labelled data; no ground-truth hair masks
(Li et al., 2021)	2021	U-Net segmentation	GAN-based (SN-PatchGAN)	206	Small training set (80 images); domain gap from face pretraining
(Lama et al., 2022)	2022	U-Net variant (ChimeraNet)	Not evaluated	25	Very small test set; no inpainting component
(El-Shafai et al., 2023)	2023	Black top-hat + CNN classification	N/A	206	No novel segmentation/inpainting method proposed

2.2 CNN-Based Approaches

Attia (Attia et al., 2019) was the first to use a hybrid CNN–RNN architecture for hair segmentation; however, the work relied on weakly labelled data with no ground-truth hair masks and offered no inpainting component. Li et al. (Li et al., 2021) paired a U-Net segmentation network with GAN-based inpainting trained on a newly built Digital Hair Dataset and clearly outperformed DullRazor (Jaccard 95.42 vs. 61.88). Its limits were a training set of only 80 images and an inpainting network pre-trained on celebrity faces, which introduces a domain gap. Lama et al. (Lama et al., 2022) proposed ChimeraNet, a U-Net variant trained with a combined loss function that sharpened the detection of fine hair boundaries, although it was evaluated on only 25 images and had no inpainting stage.

2.3 GAN-Based Approaches

Kim and Hong (Kim & Hong, 2021) trained a GAN inpainting network together with a U-Net segmentation network on a curated ground-truth dataset; removal quality improved, but scalability suffered because manual curation limits dataset size. El-Shafai et al. (El-Shafai et al., 2023) showed that black top-hat segmentation followed by CNN classification, used as a pre-processing step, raised downstream diagnostic accuracy on the ISIC 2018 dataset. This gives direct clinical support for DHR, although the study did not propose a new removal algorithm.

2.4 Research Gaps

Even after more than twenty years of research, no dermoscopic hair-removal method performs reliably across all hair types. Classical techniques have trouble with fine or light hair, over-segmentation artefacts, and texture blurring, and they are validated on very small datasets. Learning-based methods do better, but each has its own weakness: weakly labelled training data (Attia et al., 2019), a domain gap caused by unrelated pre-training data (Li et al., 2021), very small validation sets (Lama et al., 2022), and, for GAN-based methods, too little supervision from binary hair labels. Most importantly, no earlier study has compared several DHR architectures on an identical dataset with a consistent set of quantitative metrics, so it is hard to conclude reliably which architectural family suits the task best.

To address this gap, this study implements and compares three representative architectures (MLP, U-Net, and EA-GAN). They were chosen because together they cover the main architectural families in the DHR literature (pixel-wise classification, encoder-decoder segmentation, and adversarial image-to-image translation), and each has performed well in related image-restoration tasks (El-Shafai et al., 2023; Jütte et al., 2024; Lama et al., 2022; Li et al., 2021). All three models are trained and evaluated on the same dermoscopy dataset with an identical set of quantitative criteria, so dataset variability cannot act as a confounding factor. Architectural depth governs the trade-off between representational capacity and generalisation, so each model was first tuned over a range of depths to find its own optimal configuration before the final comparison. The MLP was tested at 3, 5, 8, and 12 layers; for U-Net, three variants differing in depth, normalisation, and up-sampling strategy were compared; and for EA-GAN, generator depth and attention configuration were tuned to balance reconstruction quality against computational cost. This depth search means each architecture is compared at its best-performing configuration and not under an artificially uniform constraint.

2. MATERIALS AND METHODS

2.1 Dataset

The experiments used the DERMAHAIR dataset, a public benchmark for automatic hair removal in dermoscopic images, available on Kaggle [43]. It is derived from HAM100, a large collection of dermoscopic images of common pigmented skin lesions. Artificial hair patterns from the Digital Hair Dataset [44] were superimposed on hair-free images using binary masks (white pixels for hair, black for background). The masks were normalized and merged with the images, and pixel values were then adjusted in post-processing. This produced matched pairs of a hair-occluded image and its hair-free ground truth.

The dataset contains 707 hair-occluded and 707 hair-free images (1,414 in total), with 101 pairs for each of seven lesion categories: basal cell carcinoma, dermatofibroma, melanocytic nevus, vascular lesions, benign keratosis, melanoma and actinic keratosis. The images are stored in two directories, *Data_Skin_with_Hair* and *Data_Skin_without_Hair*, each with one subfolder per category. The dataset comprised images with both low- and high-density hair. The hairs superimposed on the skin images exhibited varying orientations and angles, as illustrated in Fig. 1. In addition, the dataset contained lesions with varying sizes, shapes, colors, and visual characteristics, representing differences in lesion magnitude and appearance across the images. Representative samples are shown in Figure 1.



Figure 1. Representative samples from the dermoscopic dataset: (a) with hair, ‘Data_Skin_with_Hair’; (b) hair-free, ‘Data_Skin_without_Hair’.

2.2 Performance Evaluation Metrics

Performance was measured with six complementary metrics that cover pixel-level error, structural similarity, and perceptual quality: MSE (Zhou Wang & Bovik, 2009), RMSE (Barnston, 1992), PSNR (Almohammad & Ghinea, 2010), SSIM (Zhou Wang et al., 2004), MS-SSIM (Wang et al., 2003), and VIF (Han et al., 2013), as defined in Equations (1)–(6).

$$MSE = \left(\frac{1}{n} \right) \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (1)$$

$$RMSE = \sqrt{\left[\left(\frac{1}{n} \right) \sum_{i=1}^n (\hat{y}_i - y_i)^2 \right]} \quad (2)$$

$$PSNR = 10 \log_{10} (MAX^2 / MSE) \quad (3)$$

$$SSIM(x, y) = [(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)] / [(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)] \quad (4)$$

$$MS-SSIM(x, y) = (1/L) \sum_{i=1}^L SSIM_i(x, y) \quad (5)$$

$$VIF = 1 / (1 - R_i^2) \quad (6)$$

Here, $\hat{y}_i$ and y_i denote the predicted and reference pixel values, respectively, and n is the number of observations. MSE and RMSE quantify the average pixel-wise reconstruction error, and RMSE is expressed in the same units as the original pixel intensities. PSNR (Eq. 3) puts this error on a logarithmic decibel scale relative to MAX, the maximum possible pixel value (255 for 8-bit images); a higher PSNR means better reconstruction. SSIM (Eq. 4) compares local luminance (means μ_x , μ_y), contrast (variances σ_x^2 , σ_y^2), and structure (covariance σ_{xy}) between the reconstructed and reference images, with C_1 and C_2 as stabilising constants, and MS-SSIM (Eq. 5) averages SSIM over L spatial scales to capture both coarse and fine structural agreement. VIF (Eq. 6) estimates the share of reference-image information preserved in the reconstruction, where R_i^2 is the coefficient of determination for image component i ; a higher VIF means greater perceptual fidelity. For every metric except MSE and RMSE, higher values indicate better performance.

The three models (MLP, U-Net, and EA-GAN) were each designed, optimised, and evaluated separately for hair removal on this dataset. Their architectures, training protocols, and loss functions are described in Sections 3–5, and the comparison of their performance is given in Section 6.

3. MLP-BASED HAIR DETECTION AND REMOVAL

A Multilayer Perceptron (MLP) is a fully connected feed-forward neural network. Here, MLP-based hair removal runs in five consecutive stages, shown in Figure 2: (1) image acquisition and pre-processing, (2) multichannel feature extraction, (3) building a class-balanced training set, (4) pixel-wise MLP classification, and (5) reconstruction by inpainting.

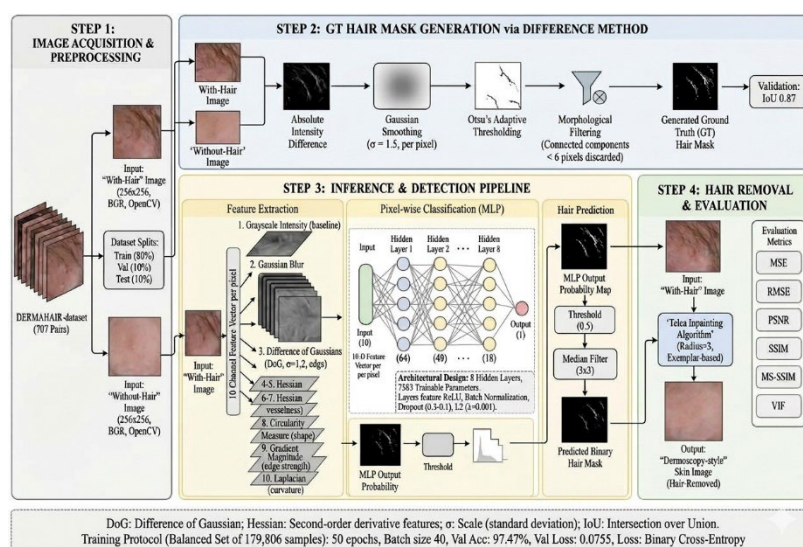


Figure 2. Hair-removal pipeline using the MLP classifier and Telea inpainting.

3.1 Image Acquisition and Pre-processing

Image pairs from the DERMAHAIR dataset were resized to 256×256 pixels and converted to BGR format with OpenCV. The data were divided into 80% training, 10% validation, and 10% test sets, a ratio commonly used for small biomedical image datasets.

3.2 Ground-Truth Hair-Mask Generation

Ground-truth hair masks were produced automatically (Eq. 7). The absolute intensity difference between each with-hair and without-hair pair was computed, smoothed with a Gaussian filter ($\sigma = 1.5$), and thresholded using Otsu's method; connected components smaller than 6 pixels were discarded as noise. Because of this automated procedure, no manual mask annotation was needed.

$$M_{hair} = T(G\sigma = 1.5 (|I_{with} - I_{without}|)) \in \{0,1\}H \times W \quad (7)$$

3.3 Multichannel Feature Extraction

Every pixel was represented by a 10-dimensional feature vector that mixes intensity, edge, and structural information (Table 2). The features are grayscale intensity, Gaussian-blurred intensity, a difference-of-Gaussians edge response, Hessian eigenvalues at two scales (which capture tubular, hair-like structures), an eigenvalue-ratio circularity measure, gradient magnitude, and the Laplacian (Kusnik & Smolka, 2022; Pauwelyn et al., 2025). Using appearance and geometric cues at several scales helps the classifier separate thin, curvilinear hairs from the surrounding skin texture.

Table 2. Handcrafted multi-channel features used for pixel-level hair detection.

Ch.	Feature	Formulation	Purpose
1	Original intensity	$f_1 = I_{gray}$	Baseline intensity
2	Gaussian blur	$f_2 = I * G(\sigma = 1)$	Noise reduction
3	Difference of Gaussians	$f_3 = I * G(\sigma = 1) - I * G(\sigma = 2)$	Edge enhancement
4–5	Hessian eigenvalues ($\sigma=1$)	λ_1, λ_2 of $H(\sigma = 1)$	Tubular structure detection
6–7	Hessian eigenvalues ($\sigma=2$)	λ_1, λ_2 of $H(\sigma = 2)$	Multi-scale vesselness
8	Circularity measure	$f_8 = 1 - (\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2 + \epsilon)$	Shape characterisation
9	Gradient magnitude	$f_9 = \sqrt{(\partial I/\partial x)^2 + (\partial I/\partial y)^2}$	Edge strength
10	Laplacian	$f_{10} = \nabla^2 I$	Curvature information

3.4 Dataset Construction and Class Balancing

The resulting pixel-level dataset is given in Eq. 8; each sample couples a 10-dimensional feature vector with a binary label (1 = hair, 0 = non-hair).

$$D = \{ (x_i, y_i) \mid x_i \in \mathbb{R}^{10}, y_i \in \{0,1\} \} \quad (8)$$

Hair pixels make up only a small minority of a dermoscopic image, so an adaptive sampling strategy kept every hair pixel and randomly sub-sampled the non-hair pixels according to $N(\text{non-hair}) = \min(|\text{Hair}| \times 5 + 500, |\text{Non-Hair}|)$. This stopped the majority class from dominating training and gave a final training set of 179,806 samples, split equally between the two classes (89,903 each).

3.5 Architecture Design

Network depth was varied over 3, 5, 8, and 12 layers under identical training conditions. The three- and five-layer networks under-fit and could not learn discriminative pixel-level features, while the 12-layer network overfit because the dataset is small. The eight-layer configuration struck the best balance between capacity and generalisation and was adopted as the final

architecture (Table 3). The hidden-layer width shrinks gradually from 64 to 8 neurons ($64 \rightarrow 48 \rightarrow 32 \rightarrow 24 \rightarrow 18 \rightarrow 14 \rightarrow 10 \rightarrow 8$). Each layer uses ReLU activation, Batch Normalisation, Dropout (0.3 for the first two layers, 0.2 for the middle layers, 0.1 for the last two), and L2 regularisation ($\lambda = 0.001$) (Sayin et al., 2025; Yahia et al., 2025). A sigmoid output layer gives the pixel-wise probability of hair. The final model has 7,583 trainable parameters (Table 3) and is trained with binary cross-entropy loss (Eq. 9).

$$L = -(1/N) \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (9)$$

3.6 Training Protocol

The balanced set of 179,806 samples was divided 80/20 into training and validation subsets, and the model was trained for 50 epochs with a batch size of 40. Validation accuracy was already 96.25% after the first epoch and rose steadily to 97.47% (validation loss 0.0755) by epoch 50. Over the same period, training and validation loss dropped from 0.8139 to 0.0808. An exponential learning-rate schedule halved the learning rate from 0.001 to 0.0005 in the later epochs (from epoch 46 onward), which contributed to the last small accuracy gains (Figure 3).

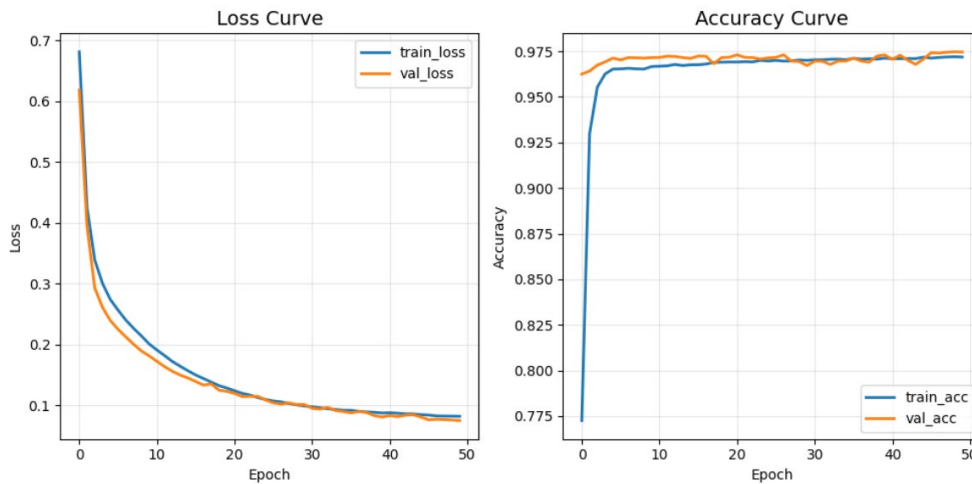


Figure 3. Training and validation loss and accuracy curves of the MLP model over 50 epochs.

3.7 Hair Removal via Inpainting

The detected hair regions were removed with Telea exemplar-based inpainting (radius = 3), which rebuilds occluded pixels as a weighted combination of nearby texture and structure (Eq. 10).

$$I_{\text{clean}}(p) = [\sum_{q \in B_{\varepsilon}(p)} w(p, q) (I_{\text{orig}}(q) + \nabla I(q) \cdot (p - q))] / \sum_{q \in B_{\varepsilon}(p)} w(p, q) \quad (10)$$

where $w(p, q)$ is a weighting function based on distance and directional coherence; the algorithm propagates texture and structure from the surrounding area within an inpainting radius of 3 pixels.

Table 3. MLP architecture used for pixel-level hair detection.

Layer	Neurons	Activation	Regularisation	Trainable Params
Input	10	—	—	—
Hidden 1	64	ReLU	BN + Dropout(0.3) + L2	832
Hidden 2	48	ReLU	BN + Dropout(0.3) + L2	3,216
Hidden 3	32	ReLU	BN + Dropout(0.2) + L2	1,632
Hidden 4	24	ReLU	BN + Dropout(0.2) + L2	840
Hidden 5	18	ReLU	BN + Dropout(0.2) + L2	486

Hidden 6	14	ReLU	BN + Dropout(0.2) + L2	294
Hidden 7	10	ReLU	BN + Dropout(0.1) + L2	170
Hidden 8	8	ReLU	BN + Dropout(0.1) + L2	104
Output	1	Sigmoid	—	9
Total				7,583

BN = Batch Normalisation; L2 regularisation coefficient $\lambda = 0.001$ applied to all Dense layers. Non-trainable Batch Normalisation parameters are excluded from the trainable count.

4. U-NET-BASED HAIR SEGMENTATION AND REMOVAL

U-Net was first proposed for biomedical image segmentation. It is a symmetric encoder–decoder CNN in which skip connections concatenate encoder feature maps directly with the matching decoder stages, retaining fine spatial detail that downsampling would otherwise discard. This makes it well suited to segmenting thin, elongated structures such as dermoscopic hair. The U-Net pipeline has three stages: (1) pre-processing and automatic ground-truth mask generation, (2) training with a combined BCE–Dice loss, and (3) Telea-based inpainting of the segmented hair regions.

4.1 Data Preparation and Ground-Truth Generation

As in the MLP pipeline, ground-truth hair masks came from the absolute grayscale difference between paired images, which was smoothed with a 3×3 Gaussian kernel and thresholded with Otsu's method; connected components under 6 pixels were removed (Ashour et al., 2022). Images were converted from BGR to RGB, resized to 256×256 with area interpolation, and normalised to $[0, 1]$; masks were resized with nearest-neighbour interpolation so that they stay binary (Guo et al., 2025). The training data were augmented by random horizontal and vertical flips ($p = 0.5$, applied identically to an image and its mask), whereas the validation set was not augmented. The data were split 80/10/10 (training/validation/test) with a fixed random seed (42).

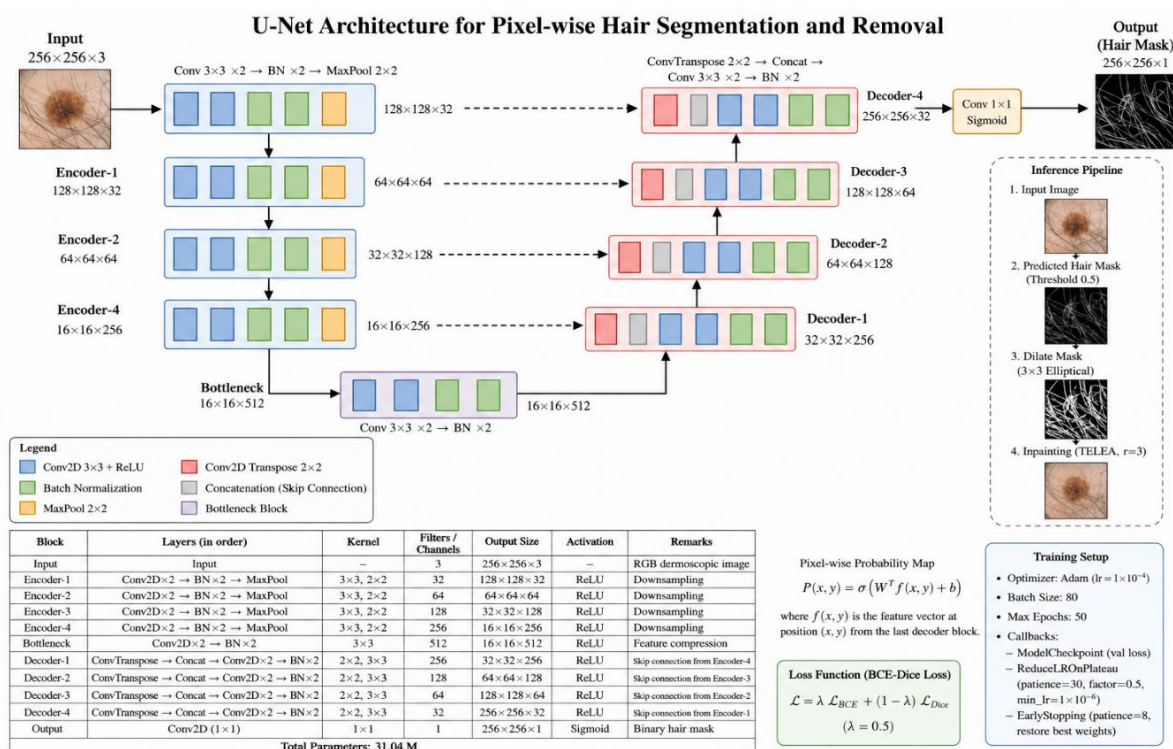


Figure 4. U-Net architecture used for pixel-wise hair segmentation and removal.

4.2 Architecture

The network accepts a $256 \times 256 \times 3$ RGB input and returns a single-channel binary segmentation mask. Its basic building block is a double-convolution unit made of two 3×3 convolutions with He-normal initialisation, Batch Normalisation, and ReLU activation (Zollanvari, 2023). Across four encoder stages the filter count doubles while the spatial resolution is halved, and a 512-filter bottleneck at 16×16 resolution follows. The decoder mirrors the encoder with transposed convolutions and receives skip connections from each matching encoder stage (Gowda et al., 2024). Depth was tuned at approximately 54, 67, and 85 layers. The 54-layer network performed worse in hair removal and lesion retention, and the 85-layer network was unstable and gave lower image quality without any gain in accuracy. The 67-layer configuration (Table 4; 31.04M parameters) performed best and became the final architecture.

A last 1×1 convolution with sigmoid activation gives a pixel-wise hair-probability map (Eq. 11), which is thresholded at 0.5 to obtain the binary mask.

$$M(x, y) = \sigma(WT \cdot f(x, y) + b), \quad \sigma(z) = 1 / (1 + e^{-z}) \quad (11)$$

A combined Binary Cross-Entropy and Dice loss (BCE–Dice) was used for training to deal with the strong class imbalance between hair and skin pixels. Dice loss maximises the spatial overlap between predicted and ground-truth masks, and BCE penalises the misclassification of individual pixels (Hosen & Zhang, 2025)

4.3 Training Protocol

The model was trained with the Adam optimiser (learning rate 1×10^{-4} , batch size 8, at most 50 epochs) together with three callbacks: ModelCheckpoint (best validation loss), ReduceLROnPlateau (halving the learning rate after 30 stagnant epochs, with a floor of 1×10^{-6}), and EarlyStopping (patience of 8 epochs, restoring the best weights). Training and validation loss fell sharply from 0.27/0.13 at epoch 1 to a stable value of approximately 0.03 by epoch 20, while accuracy rose from 0.90 to approximately 0.99 over the same span. Convergence was therefore faster than for the MLP, which needed all 50 epochs to reach 0.97 (Figure-5). The training and validation curves stayed in close agreement throughout, which confirms stable convergence without overfitting.

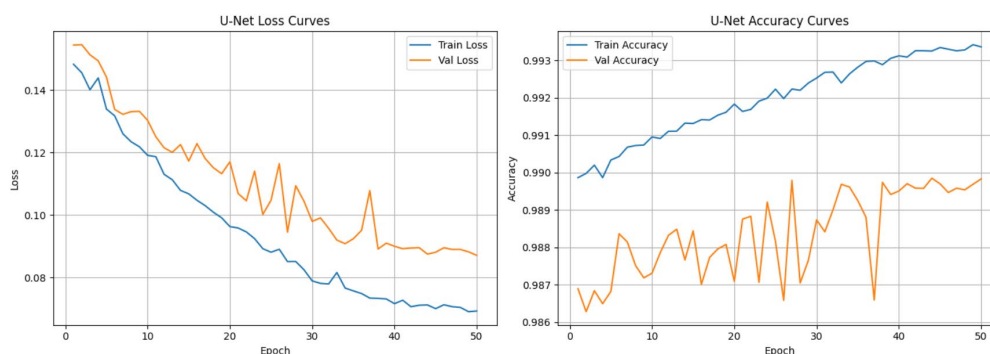


Figure 5. Training and validation loss and accuracy curves of the U-Net model over 50 epochs.

Table 4. U-Net architecture specifications (final 67-layer configuration).

Block	Layers (in order)	Kernel	Filters	Output Size	Remarks
Input	—	—	3	$256 \times 256 \times 3$	RGB dermoscopic image
Encoder-1	Conv2D $\times 2 \rightarrow$ BN $\times 2 \rightarrow$ MaxPool	$3 \times 3, 2 \times 2$	32	$128 \times 128 \times 32$	Downsampling
Encoder-2	Conv2D $\times 2 \rightarrow$ BN $\times 2 \rightarrow$ MaxPool	$3 \times 3, 2 \times 2$	64	$64 \times 64 \times 64$	Downsampling
Encoder-3	Conv2D $\times 2 \rightarrow$ BN $\times 2 \rightarrow$ MaxPool	$3 \times 3, 2 \times 2$	128	$32 \times 32 \times 128$	Downsampling
Encoder-4	Conv2D $\times 2 \rightarrow$ BN $\times 2 \rightarrow$ MaxPool	$3 \times 3, 2 \times 2$	256	$16 \times 16 \times 256$	Downsampling
Bottleneck	Conv2D $\times 2 \rightarrow$ BN $\times 2$	3×3	512	$16 \times 16 \times 512$	Feature compression
Decoder-1	ConvT $\rightarrow$ Concat $\rightarrow$ Conv2D $\times 2 \rightarrow$ BN $\times 2$	$2 \times 2, 3 \times 3$	256	$32 \times 32 \times 256$	Skip from Encoder-4
Decoder-2	ConvT $\rightarrow$ Concat $\rightarrow$ Conv2D $\times 2 \rightarrow$ BN $\times 2$	$2 \times 2, 3 \times 3$	128	$64 \times 64 \times 128$	Skip from Encoder-3
Decoder-3	ConvT $\rightarrow$ Concat $\rightarrow$ Conv2D $\times 2 \rightarrow$ BN $\times 2$	$2 \times 2, 3 \times 3$	64	$128 \times 128 \times 64$	Skip from Encoder-2

Decoder-4	ConvT $\rightarrow$ Concat $\rightarrow$ Conv2D $\times$ 2 $\rightarrow$ BN $\times$ 2	2 $\times$ 2, 3 $\times$ 3	32	256 $\times$ 256 $\times$ 32	Skip from Encoder-1
Output	Conv2D (Sigmoid)	1 $\times$ 1	1	256 $\times$ 256 $\times$ 1	Binary hair mask
Total					31.04 M parameters

4.4 Hair Removal via Inpainting

At inference time, the predicted probability map is resized back to the original resolution and thresholded at 0.5. The binary mask obtained is dilated with a 3 $\times$ 3 elliptical kernel to cover edge artefacts, and the hair regions are then removed by Telea inpainting with radius = 3 (Gertsolf et al., 2024).

5. EA-GAN-BASED HAIR REMOVAL

The Enhanced Attention Generative Adversarial Network (EA-GAN) restores hair-free dermoscopic images with an adversarial generator–discriminator pair. The generator learns to synthesise realistic hair-free images from hair-occluded input, and the discriminator learns to tell generated images from real hair-free ones. The pipeline has four stages: (1) data pre-processing, (2) design of the Attention U-Net generator and the PatchGAN discriminator, (3) adversarial training with a combined adversarial and L1 reconstruction loss, and (4) performance evaluation.

The first baseline used a deeper U-Net generator with eight encoder–decoder stages and was trained with adversarial, L1, and edge-consistency losses. It removed hair artefacts but left blurred regions, retained lesion detail poorly, and was computationally expensive. Adding attention gates to the skip connections and cutting the generator to four encoder–decoder stages preserved features and lesion structure better. Since the attention mechanism already enforced structural fidelity, the edge-consistency loss became unnecessary. Making the network deeper or more complex gave marginal or even negative returns, which is consistent with overfitting, so the four-stage, attention-gated configuration was adopted as the final architecture. It balances reconstruction quality, training efficiency, and computational cost.

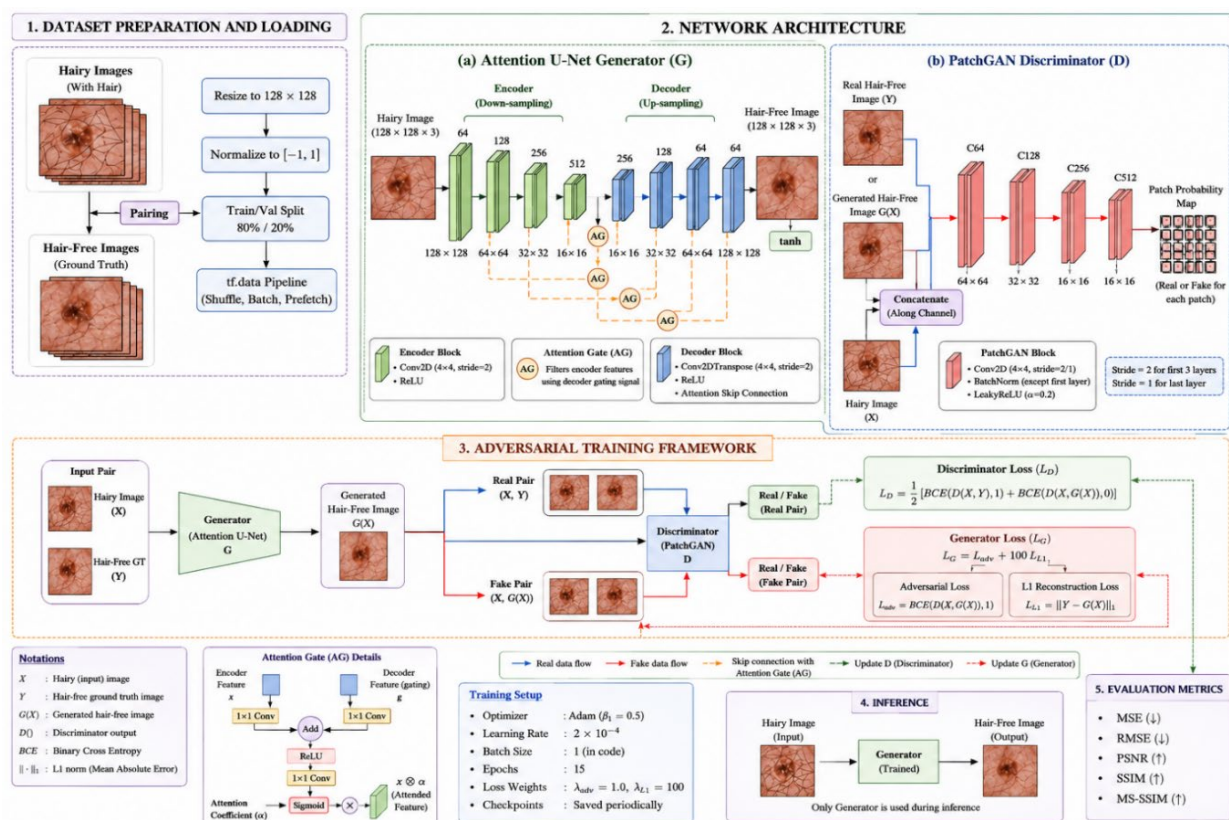


Figure 6. Architecture of the EA-GAN framework for dermoscopic hair removal, comprising an Attention U-Net generator and a PatchGAN discriminator.

5.1 Dataset Preparation

Hair and hair-free images were matched into pairs, resized to 128×128 pixels, and normalised to $[-1, 1]$ by min–max scaling. The data were split 80:20 into training and validation subsets and fed through a TensorFlow tf.data pipeline with shuffling, batching, and prefetching.

5.2 Network Architecture

Attention U-Net generator The generator follows a U-Net design with attention gates placed on the skip connections. Its four encoder stages (filter sizes 64, 128, 256, 512) each downsample with a 4×4 strided convolution and ReLU activation, and the bottleneck holds high-level semantic information about skin texture, lesion borders, and hair structure. The decoder restores spatial resolution with transposed convolutions. Instead of concatenating raw encoder features, an attention gate first computes coefficients that emphasise lesion-relevant structure and suppress background and hair, and only then are the features merged with the decoder path. A final 1×1 convolution with Tanh activation outputs a 128×128×3 reconstructed hair-free image, where the three channels are the RGB colour components of the dermoscopic image.

Attention gate Encoder features x and decoder gating features g are each projected into a lower-dimensional space by a 1×1 convolution, added together, and passed through ReLU. Another 1×1 convolution followed by a sigmoid yields an attention coefficient map, which is multiplied element-wise with the encoder features before they reach the decoder. In this way redundant information is suppressed and diagnostically relevant lesion structure is emphasised during reconstruction.

PatchGAN discriminator The discriminator takes the hair-occluded input, channel-concatenated with either the real or the generated hair-free image, and passes it through four convolutional stages (64, 128, 256, 512 filters; 4×4 kernels). Every stage after the first is followed by Batch Normalisation and LeakyReLU ($\alpha = 0.2$). PatchGAN does not output a single real/fake score for the whole image; it returns a grid of patch-level real/fake probabilities, which pushes the generator to reproduce realistic local texture and structure that matter for dermoscopic detail (Table 5).

Table 5. Architectural configuration of the Attention-Based EA-GAN.

Component	Architecture	Key Parameters	Output Shape
Generator (Attention U-Net)	Encoder: 4× [Conv 4×4, stride 2, ReLU]; Decoder: 4× [ConvT 4×4, stride 2, ReLU, attention skip]	Filters 64→512; attention gates; final activation Tanh	128×128×3
Discriminator (PatchGAN)	C64–C128–C256–C512	Kernel 4×4; stride 2/1; LeakyReLU $\alpha = 0.2$	Patch probability map

5.3 Loss Functions

The generator loss is the sum of an adversarial term and an L1 reconstruction term. The adversarial loss (Eq. 12) uses binary cross-entropy to push the generator towards images that the discriminator classifies as real; here $G(x)$ is the generated hair-free image.

$$L_{adv} = BCE(D(x, G(x)), 1) \quad (12)$$

The L1 loss (Eq. 13) is the mean absolute difference between the generated image and the ground-truth hair-free image. It helps preserve the overall lesion structure and avoids unnecessary changes to healthy skin.

$$LL1 = ||y - G(x)||_1 \quad (13)$$

In the total generator loss (Eq. 14), the L1 term is weighted 100 times relative to the adversarial term, which favours accurate reconstruction while still keeping the output visually realistic.

$$LG = L_{adv} + 100 \cdot LL1 \quad (14)$$

The discriminator loss (Eq. 15) averages the binary cross-entropy over real pairs (label 1) and generated pairs (label 0), so the discriminator learns features that separate authentic hair-free images from synthesised ones.

$$LD = \frac{1}{2} [BCE(D(x, y), 1) + BCE(D(x, G(x)), 0)] \quad (15)$$

5.4 Training Procedure

The generator and discriminator were trained in alternation for 15 epochs with the Adam optimiser ($\beta_1 = 0.5$, learning rate 0.0002). In each iteration the generator produced a synthetic hair-free image from the hair-occluded input. The discriminator was evaluated on real and generated pairs to obtain the adversarial loss, which was combined with the L1 loss to update the generator, while the discriminator itself was updated with its binary cross-entropy classification loss. Checkpoints were saved periodically to follow intermediate performance (Figure 7).

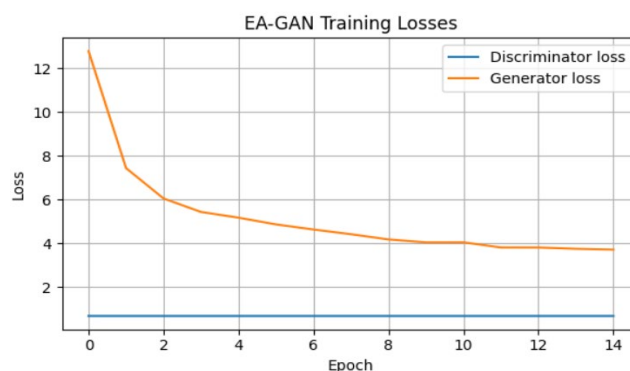


Figure 7. Generator and discriminator loss curves of the EA-GAN model during training.

6. RESULTS AND DISCUSSION

6.1 Qualitative Comparison

The three models were compared visually on randomly chosen test images that vary in hair density, lesion shape, and hair colour, with all outputs resized to a common resolution and set against the ground-truth hair-free images (Figure 8). The MLP was inconsistent and often left residual hair in dense or curved-hair regions. EA-GAN produced realistic, well-reconstructed images but sometimes introduced smoothing artefacts in complex regions. U-Net did best overall: it removed hair effectively while keeping lesion morphology, texture, and boundary detail, and its outputs closely matched the ground truth apart from a little residual hair in a few difficult cases.

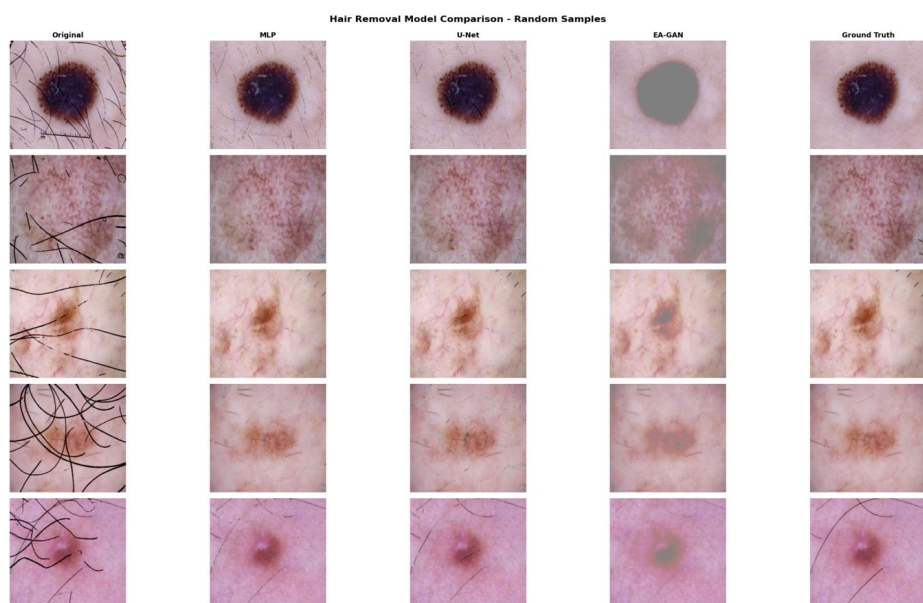


Figure 8. Qualitative comparison of hair-removal outputs from MLP, U-Net, and EA-GAN on a randomised sample of dermoscopic images.

6.2 Quantitative Evaluation

Table 6 gives the average performance over all test images, with each metric computed per image and then averaged across the test set. An $\uparrow$ means higher values are better and a $\downarrow$ means lower values are better.

Table 6. Average quantitative performance of the hair-removal models.

Model	MSE $\downarrow$	RMSE $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	MS-SSIM $\uparrow$	VIF $\uparrow$
MLP	61.36	7.83	33.31	0.836	0.934	1.012
U-Net	47.56	6.90	34.20	0.892	0.954	1.024
EA-GAN	42.66	6.53	33.77	0.888	0.945	0.587

The MLP was worst on every metric, which reflects its limited ability to model the complex spatial structure of hair artefacts. EA-GAN reached the lowest MSE (42.66) and RMSE (6.53), meaning the most accurate pixel-level reconstruction, but its low VIF (0.587) points to weaker preservation of perceptual detail. U-Net obtained the highest PSNR (34.20 dB), SSIM (0.892), MS-SSIM (0.954), and VIF (1.024), showing the best overall structural preservation and perceptual fidelity. Overall, U-Net gives the most balanced performance for dermoscopic hair removal among the three architectures tested.

6.3 Visualisation of Metric Comparisons

Figure 9 (bar chart) and Figure 10 (heatmap) compare all six metrics across the three models and confirm the pattern described in Section 7.2. U-Net leads on four of the six metrics (PSNR, SSIM, MS-SSIM, VIF), EA-GAN leads on the two pixel-error metrics (MSE, RMSE), and the MLP is last on all six.



Figure 9. Bar-chart comparison of MLP, U-Net, and EA-GAN across MSE, RMSE, PSNR, SSIM, MS-SSIM, and VIF.

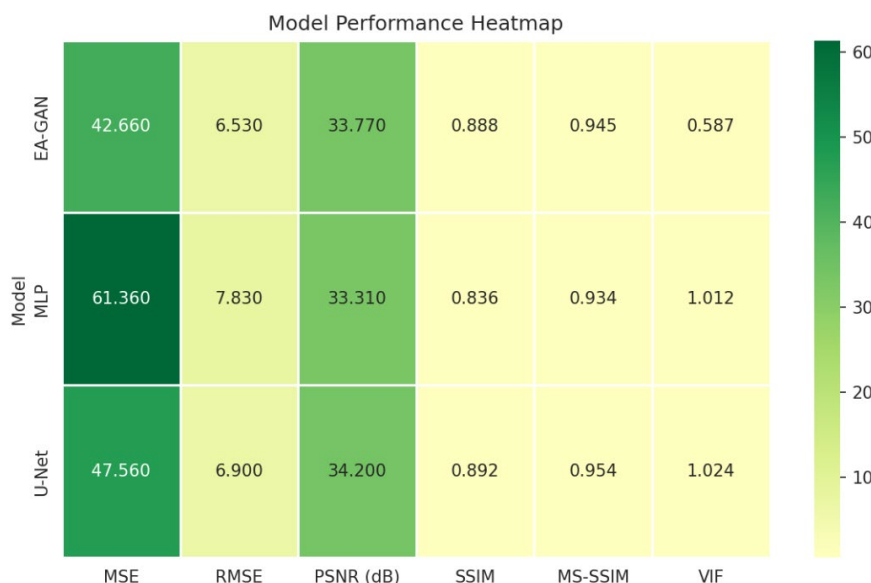


Figure 10. Heatmap visualisation of the quantitative performance of the MLP, U-Net, and EA-GAN hair-removal models across all six metrics.

6.4 Discussion

The MLP, U-Net, and EA-GAN stand for three different architectural families: a fully connected network, a convolutional encoder–decoder, and an adversarial image-to-image framework. For a fair comparison, each was optimised to its own best configuration and not forced into a shared depth or parameter budget. The MLP performed best with eight hidden layers; beyond that its performance levelled off or dropped because of overfitting. U-Net gained from a deeper 67-layer configuration, which supports richer hierarchical feature extraction, and EA-GAN was trained at the depth and attention configuration that gave the most stable adversarial convergence. These results suggest that the encoder–decoder inductive bias, meaning spatial downsampling combined with skip-connection recovery of fine detail, is much better suited to hair removal than a purely pixel-wise, fully connected classifier. They also suggest that adversarial training can reach very low pixel-level error, but at some cost in perceptual fidelity compared with a direct segmentation-and-inpainting approach based on an encoder–decoder.

Table 7 brings together the architectural and practical trade-offs of the three models to help choose among them in future deployments.

Table 7. Qualitative comparison of the three hair-removal architectures evaluated.

Criterion	MLP	U-Net	EA-GAN
Architectural family	Fully connected (pixel-wise)	Convolutional encoder–decoder	Adversarial (generator–discriminator)
Best configuration	8 hidden layers	67 layers	4 encoder–decoder stages + attention gates
Trainable parameters	7,583	31.04 M	13.82 M
Training length	50 epochs	50 epochs	15 epochs
Strongest metric	—	PSNR, SSIM, MS-SSIM, VIF	MSE, RMSE
Key strength	Simple, lightweight, fast inference	Best structural & perceptual fidelity	Lowest pixel-level error; realistic textures

6.5 Limitations

All models were evaluated on one dataset (DERMAHAIR) in which hair patterns were superimposed synthetically; performance on natural, unconstrained clinical hair occlusion and on other dermoscopy datasets was not examined. No statistical significance testing (for example paired t-tests or Wilcoxon signed-rank tests on per-image metric distributions) was carried out, so the reported differences are averages and not confidence-bounded comparisons. Downstream clinical impact, such as changes in lesion segmentation or classification accuracy after hair removal, was not measured directly and is left for future work.

7. CONCLUSION AND FUTURE SCOPE

In this study, three deep learning approaches (MLP, U-Net, and EA-GAN) were implemented and compared for hair removal in dermoscopic images, and each was optimised to its own best configuration so that the evaluation would be fair. All three reduced hair artefacts, but their performance profiles differed clearly. U-Net gave the best overall balance of structural and perceptual quality (PSNR 34.20 dB, SSIM 0.892, MS-SSIM 0.954, VIF 1.024). EA-GAN reached the lowest pixel-level reconstruction error (MSE 42.66, RMSE 6.53) but had weaker perceptual fidelity (VIF 0.587). The fully connected MLP was consistently behind both encoder–decoder-based approaches, reflecting its limited capacity to capture the complex spatial structure of hair artefacts. These findings suggest that convolutional encoder–decoder architectures are the best choice for dermoscopic hair removal and can make fully automated skin-lesion analysis pipelines more reliable. The primary clinical benefit of hair removal from the lesion surface in dermoscopic images is to facilitate clearer lesion visualization, which may support more accurate lesion identification and improve the quality of the predicted lesion image.

Future work could look at hybrid architectures that pair U-Net's structural preservation with adversarial reconstruction, add attention mechanisms to localise hair better, and use perceptual or edge-aware loss functions. Training on larger and more varied datasets and building more computationally efficient models would further help bring these methods into real clinical workflows.

DECLARATIONS

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability: The dataset analysed in this study (DERMAHAIR) is publicly available on Kaggle (Riotulab, n.d.). The code and trained model weights are available from the corresponding author upon reasonable request.

Ethical Approval: This study used a publicly available, de-identified image dataset; it did not involve direct data collection from human or animal subjects, and no additional ethical approval was required.

ACKNOWLEDGEMENTS

Acknowledgements: The authors thank the creators and maintainers of the DERMAHAIR and Digital Hair Dataset resources for making their data publicly available for research.

AUTHOR CONTRIBUTIONS

Nutan Pathania (corresponding author, research scholar): conceptualisation of the research problem; data curation of dermoscopic skin-cancer datasets; development and implementation of machine learning and deep learning models; software; investigation; formal analysis; visualisation; validation; writing original draft; writing review and editing.

Aman Kumar Sharma (supervisor): conceptualisation; supervision; methodology design and refinement; resources; project administration; validation; writing review and editing.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ETHICS STATEMENT

This study used a publicly available, de-identified image dataset; it did not involve direct data collection from human or animal subjects, and no additional ethical approval was required.

DATA AVAILABILITY STATEMENT

The dataset analysed in this study (DERMAHAIR) is publicly available on Kaggle (Riotulab, n.d.). The code and trained model weights are available from the corresponding author upon reasonable request.

REFERENCES

- [1] Abbas, Q., Celebi, M. E., And García, I. F. (2011). Hair Removal Methods: A Comparative Study For Dermoscopy Images. *Biomedical Signal Processing And Control*, 6, (4), 395–404. <https://doi.org/10.1016/j.bspc.2011.01.003>
- [2] Akilandasowmya, G., Nirmaladevi, G., Suganthi, Su., And Aishwariya, A. (2024). Skin Cancer Diagnosis: Leveraging Deep Hidden Features And Ensemble Classifiers For Early Detection And Classification. *Biomedical Signal Processing And Control*, 88, 105306. <https://doi.org/10.1016/j.bspc.2023.105306>
- [3] Almohammad, A., And Ghinea, G. (2010). Stego Image Quality And The Reliability Of Psnr. 2010 2nd International Conference On Image Processing Theory, Tools And Applications, Ieee. <https://doi.org/10.1109/Ipta.2010.5586786>
- [4] Ashour, A. S., El-Wahab, B. S. A., Wahba, M. A., Mansour, D.-E. A., Hodeib, A. A. E., Khedr, R. A. E.-G., And Hassan, G. F. R. (2022). Cascaded Hough Transform-Based Hair Mask Generation And Harmonic Inpainting For Automated Hair Removal From Dermoscopy Images. *Diagnostics*, 12, (12), 3040. <https://doi.org/10.3390/Diagnostics12123040>
- [5] Attia, M., Hossny, M., Zhou, H., Nahavandi, S., Asadi, H., And Yazdabadi, A. (2019). Digital Hair Segmentation Using Hybrid Convolutional And Recurrent Neural Networks Architecture. *Computer Methods And Programs In Biomedicine*, 177, 17–30. <https://doi.org/10.1016/j.cmpb.2019.05.010>
- [6] Bardou, D., Bouaziz, H., Lv, L., Bounezra, M., Vajdi, A., Zhang, T., Abbas, F., And Malah, M. (2026). A Review Of Automatic Hair Removal In Dermoscopy Images: From Image Processing To Deep Learning. *Journal Of Imaging Informatics In Medicine*. <https://doi.org/10.1007/S10278-026-01950-Y>
- [7] Barnston, A. G. (1992). Correspondence Among The Correlation, Rmse, And Heidke Forecast Verification Measures; Rsefinement Of The Heidke Score. *Weather And Forecasting*, 7, (4), 699–709. [https://doi.org/10.1175/1520-0434\(1992\)007%3c0699:Catcra%3e2.0.Co;2](https://doi.org/10.1175/1520-0434(1992)007%3c0699:Catcra%3e2.0.Co;2)
- [8] Bibiloni, P., González-Hidalgo, M., And Massanet, S. (2016). A Survey On Curvilinear Object Segmentation In Multiple Applications. *Pattern Recognition*, 60, 949–970. <https://doi.org/10.1016/j.patcog.2016.07.023>
- [9] El-Shafai, W., El-Fattah, I. A., And Taha, T. E. (2023). Deep Learning-Based Hair Removal For Improved Diagnostics Of Skin Diseases. *Multimedia Tools And Applications*, 83, (9), 27331–27355. <https://doi.org/10.1007/S11042-023-16646-6>
- [10] Fiorese, M., Peserico, E., And Silletti, A. (2011). Virtualshave: Automated Hair Removal From Digital Dermatoscopic Images. 2011 Annual International Conference Of The Ieee Engineering In Medicine And Biology Society, Ieee. <https://doi.org/10.1109/Iembs.2011.6091274>
- [11] Gertsvolf, D., Horvat, M., Aslam, D., Khademi, A., And Berardi, U. (2024). A U-Net Convolutional Neural Network Deep Learning Model Application For Identification Of Energy Loss In Infrared Thermographic Images. *Applied Energy*, 360, 122696. <https://doi.org/10.1016/j.apenergy.2024.122696>
- [12] Gouda, W., Sama, N. U., Al-Waakid, G., Humayun, M., And Jhanjhi, N. Z. (2022). Detection Of Skin Cancer Based On Skin Lesion Images Using Deep Learning. *Healthcare*, 10, (7), 1183. <https://doi.org/10.3390/Healthcare10071183>
- [13] Gowda, V. B., Gopalakrishna, M. T., Megha, J., And Mohankumar, S. (2024). Foreground Segmentation Network Using Transposed Convolutional Neural Networks And Up Sampling For Multiscale Feature Encoding. *Neural Networks*, 170, 167–175. <https://doi.org/10.1016/j.neunet.2023.11.015>
- [14] Guo, Z., Zheng, L., And Xu, W. (2025). Bgir: A Low-Illumination Remote Sensing Image Restoration Algorithm With Zynq-Based Implementation. *Sensors*, 25, (14), 4433. <https://doi.org/10.3390/S25144433>
- [15] Han, Y., Cai, Y., Cao, Y., And Xu, X. (2013). A New Image Fusion Performance Metric Based On Visual Information Fidelity. *Information Fusion*, 14, (2), 127–135. <https://doi.org/10.1016/j.inffus.2011.08.002>
- [16] Hosen, M. S., And Zhang, H. (2025). Bas-Segnet: A Boundary-Aware Sobel-Enhanced Deep Learning Framework For Breast Cancer And Skin Cancer Segmentation. *Electronics*, 15, (1), 75. <https://doi.org/10.3390/Electronics15010075>

- [17] Huang, A., Shun-Yuen Kwan, Wen-Yu Chang, Min-Yin Liu, Min-Hsiu Chi, And Gwo-Shing Chen. (2013). A Robust Hair Segmentation And Removal Approach For Clinical Images Of Skin Lesions. 2013 35th Annual International Conference Of The Ieee Engineering In Medicine And Biology Society (Embc), Ieee. <https://doi.org/10.1109/Embc.2013.6610250>
- [18] Kang, D., Kim, S., And Park, S. (2018). Flow-Guided Hair Removal For Automated Skin Lesion Identification. *Multimedia Tools And Applications*, 77, (8), 9897–9908. <https://doi.org/10.1007/S11042-018-5672-7>
- [19] Keerthana, D., Venugopal, V., Nath, M. K., And Mishra, M. (2023). Hybrid Convolutional Neural Networks With Svm Classifier For Classification Of Skin Cancer. *Biomedical Engineering Advances*, 5, 100069. <https://doi.org/10.1016/J.Bea.2022.100069>
- [20] Kiani, K., And Sharafat, A. R. (2011). E-Shaver: An Improved Dullrazor® For Digitally Removing Dark And Light-Colored Hairs In Dermoscopic Images. *Computers In Biology And Medicine*, 41, (3), 139–145. <https://doi.org/10.1016/J.Compbimed.2011.01.003>
- [21] Kim, D., And Hong, B.-W. (2021). Unsupervised Feature Elimination Via Generative Adversarial Networks: Application To Hair Removal In Melanoma Classification. *Ieee Access*, 9, 42610–42620. <https://doi.org/10.1109/Access.2021.3065701>
- [22] Koehoorn, J., Sobiecki, A. C., Boda, D., Diaconeasa, A., Doshi, S., Paisey, S., Jalba, A., And Telea, A. (2015). Automated Digital Hair Removal By Threshold Decomposition And Morphological Analysis. *Mathematical Morphology And Its Applications To Signal And Image Processing*, Springer International Publishing. https://doi.org/10.1007/978-3-319-18720-4_2
- [23] Kusnik, D., And Smolka, B. (2022). Robust Mean Shift Filter For Mixed Gaussian And Impulsive Noise Reduction In Color Digital Images. *Scientific Reports*, 12, (1), 14951. <https://doi.org/10.1038/S41598-022-19161-0>
- [24] Lama, N., Kasmir, R., Hagerty, J. R., Stanley, R. J., Young, R., Miinch, J., Nepal, J., Nambisan, A., And Stoecker, W. V. (2022). Chimeranet: U-Net For Hair Detection In Dermoscopic Skin Lesion Images. *Journal Of Digital Imaging*, 36, (2), 526–535. <https://doi.org/10.1007/S10278-022-00740-6>
- [25] Lee, T., Ng, V., Gallagher, R., Coldman, A., And Mclean, D. (1997). Dullrazor: A Software Approach To Hair Removal From Images. *Comput Biol Med*, 27. [https://doi.org/10.1016/S0010-4825\(97\)00020-6](https://doi.org/10.1016/S0010-4825(97)00020-6)
- [26] Li, W., Joseph Raj, A. N., Tjahjedi, T., And Zhuang, Z. (2021). Digital Hair Removal By Deep Learning For Skin Lesion Segmentation. *Pattern Recognition*, 117, 107994. <https://doi.org/10.1016/J.Patcog.2021.107994>
- [27] Mahmoud, N. M., And Soliman, A. M. (2024). Early Automated Detection System For Skin Cancer Diagnosis Using Artificial Intelligent Techniques. *Scientific Reports*, 14, (1), 9749. <https://doi.org/10.1038/S41598-024-59783-0>
- [28] Nancy, V. A. O., Prabhavathy, P., Arya, M. S., And Ahamed, B. S. (2023). Comparative Study And Analysis On Skin Cancer Detection Using Machine Learning And Deep Learning Algorithms. *Multimedia Tools And Applications*, 82, (29), 45913–45957. <https://doi.org/10.1007/S11042-023-16422-6>
- [29] Pauwelyn, A., Carré, M., Jourlin, M., Ginhaç, D., And Meriaudeau, F. (2025). Image Visual Quality: Sharpness Evaluation In The Logarithmic Image Processing Framework. *Big Data And Cognitive Computing*, 9, (6), 154. <https://doi.org/10.3390/Bdcc9060154>
- [30] Sayin, K. A., Gürsoy, N. K., Yolcu, T., And Gürsoy, A. (2025). On The Synergy Of Optimizers And Activation Functions: A Cnn Benchmarking Study. *Mathematics*, 13, (13), 2088. <https://doi.org/10.3390/Math13132088>
- [31] Schmid-Saugeon, P., Guillo, J., And Thiran, J. P. (2003). Towards A Computer-Aided Diagnosis System For Pigmented Skin Lesions. *Comput Med Imaging Graph*, 27. [https://doi.org/10.1016/S0895-6111\(02\)00048-4](https://doi.org/10.1016/S0895-6111(02)00048-4)
- [32] Shahzaib, A., Siddiqui, A. B., Anjum, N., Rehman, M. U., And Ramzan, N. (2025). Hair Removal And Lesion Segmentation Of Dermoscopic Images For Classification Of Skin Cancer Using Deep Neural Networks. *Egyptian Informatics Journal*, 32, 100844. <https://doi.org/10.1016/J.Eij.2025.100844>
- [33] Toossi, M. T. B., Pourreza, H. R., Zare, H., Sigari, M., Layegh, P., And Azimi, A. (2013a). An Effective Hair Removal Algorithm For Dermoscopy Images. *Skin Research And Technology*, 19, (3), 230–235. <https://doi.org/10.1111/Srt.12015>
- [34] Wang, Z., Simoncelli, E. P., And Bovik, A. C. (2003). Multiscale Structural Similarity For Image Quality Assessment. The Thirty-Seventh Asilomar Conference On Signals, Systems & Computers, 2003, Ieee. <https://doi.org/10.1109/Acssc.2003.1292216>

- [35] Xie, F.-Y., Qin, S.-Y., Jiang, Z.-G., And Meng, R.-S. (2009). Pde-Based Unsupervised Repair Of Hair-Occluded Information In Dermoscopy Images Of Melanoma. *Computerized Medical Imaging And Graphics*, 33, (4), 275–282. <https://doi.org/10.1016/J.Compmmedimag.2009.01.003>
- [36] Yahia, A. B., Kadir, I., Abdallaoui, A., And El-Hmaidi, A. (2025). Enhancing Relative Humidity Modelling Using L2 Regularization Updates. *Scientific Reports*, 15, (1), 14796. <https://doi.org/10.1038/S41598-025-94356-9>
- [37] Zhou, H., Chen, M., Gass, R., Rehg, J. M., Ferris, L., Ho, J., And Drogowski, L. (2008). Feature-Preserving Artifact Removal From Dermoscopy Images. <https://doi.org/10.1117/12.770824>
- [38] Zhou Wang, And Bovik, A. C. (2009). Mean Squared Error: Love It Or Leave It? A New Look At Signal Fidelity Measures. *Ieee Signal Processing Magazine*, 26, (1), 98–117. <https://doi.org/10.1109/Msp.2008.930649>
- [39] Zhou Wang, Bovik, A. C., Sheikh, H. R., And Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility To Structural Similarity. *Ieee Transactions On Image Processing*, 13, (4), 600–612. <https://doi.org/10.1109/Tip.2003.819861>
- [40] Zollanvari, A. (2023). *Convolutional Neural Networks. Machine Learning With Python*, Springer International Publishing. https://doi.org/10.1007/978-3-031-33342-2_14
- [41] Riotulab. (N.D.). Dermahair-Dataset [Data Set]. Kaggle. <https://www.kaggle.com/Datasets/Riotulab/Skin-Cancer-Hair-Removal>
- [42] Digital Hair Dataset. Kaggle. <https://www.kaggle.com/Datasets/Weilizai/Digital-Hair-Dataset>