INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

OPEN ACCESS

Received: 02/03/2026

Accepted: 07/05/2026

Published: 09/06/2026

Citation: Divya S, Thilagavathi S (2026) A Robust Pre-processing Pipeline for Dermoscopic Skin Lesion Images Using Canny-Based Hair Removal and Adaptive Enhancement. Indian Journal of Science and Technology 19(20): 1373-1383. https://doi.org/10.17485/IJST/v19i20.314

* **Corresponding author**.

thilaga.research@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

A Robust Pre-processing Pipeline for Dermoscopic Skin Lesion Images Using Canny-Based Hair Removal and Adaptive Enhancement

S Divya[1], **S Thilagavathi**[1]*

1 Department of Computer Science, Sri Krishna Adithya College of Arts and Science, Kovaipudhur, Coimbatore – 641042, Tamil Nadu, India

Abstract

Objectives: To present a viable preprocessing pipeline of dermoscopic images to maximize lesion appearance and minimise noise and artifacts to help in the diagnosis of skin cancer. **Method:** The 7-step Skin Cancer Image Pre-Processing (SCIP) pipeline has been created and tested on three benchmark dermoscopic databases, i.e., ISIC, PH2, and HAM10000, with over 11,000 images. The image preprocessing steps are image resizing, hair removal, artifact removal with the DullRazor algorithm that uses Canny edge detection and Telea inpainting, Gray World based color normalization, contrast enhancement through Contrast Limited Adaptive Histogram Equalization (CLAHE), noise reduction with a median filter, and Min-Max normalization to fit a deep learning model. **Findings:** As it was experimentally demonstrated, the proposed SCIP pipeline significantly enhanced lesion boundaries by approximately 27%. Additionally, it improved the contrast-to-noise ratio by 32% and reduced artifacts by more than 90%, the contrast-to-noise ratio by 32 percent and diminished the interference of hair and artifacts by more than 90 percent with respect to the raw dermoscopic images. Color normalization minimized variation in illumination across datasets by around 24 percent whereas median filter minimized the level of background noise by 18 percent with no impact on lesion structures. The pipeline reduced the classification error of deep learning models by 6.8%, 5.4%, and 4.9% on ISIC, HAM10000, and PH2 datasets, respectively, compared to standard preprocessing techniques. These findings indicate that the suggested method improves feature selection and facilitates quicker model convergence of automated skin cancer detection. **Novelty:** The proposed Skin Cancer Image Pre-Processing (SCIP) pipeline introduces a unified and adaptive preprocessing framework that integrates multiple complementary techniques, including Canny-based DullRazor hair removal, Gray World color normalization, CLAHE-based contrast enhancement, and median filtering, into a single standardized workflow. Unlike existing approaches that address individual preprocessing challenges in isolation, the SCIP framework simultaneously handles multiple artifacts while preserving lesion structure and ensuring cross-dataset consistency. The novelty lies in its system-level integration, cross-dataset evaluation, and quantitative validation demonstrating improved structural preservation, noise reduction, and artifact minimization across heterogeneous dermoscopic datasets.

Keywords: Skin cancer; Image preprocessing; CLAHE; DullRazor; Normalization

1 Introduction

Skin cancer is one of the most significant global health concerns, and in recent years computer-aided diagnostic systems based on dermoscopic imaging and deep learning have significantly improved the automated analysis of skin lesions. However, the performance of these systems heavily depends on the quality and consistency of input images. In real-world settings, dermoscopic images are often affected by non-homogeneous imaging conditions, device variability, and artifacts such as hairs, shadows, ruler marks, and reflections. These factors obscure lesion boundaries and reduce the reliability of automated diagnostic models. Furthermore, the lack of standardized and adaptive preprocessing pipelines limits the applicability of models across diverse skin tones, lesion structures, and imaging conditions.

Although several preprocessing techniques have been proposed including pipelines that combine contrast enhancement, hair removal, normalization, and geometric augmentation, Bardou et al. (2022) generative approaches that remove hairs without paired training data [(1)], filtering-based frameworks capable of detecting both dark and light hair and ruler marks, and Mello-Román et al. (2021) enhancement techniques designed to improve low-contrast dermoscopic images while preserving structural details [(2)] many existing methods still face limitations.

Most approaches focus on removing a single type of artifact and may either oversmooth diagnostically important structures or fail to completely eliminate interfering elements, which affects model robustness and clinical applicability. Therefore, there is a growing need for a comprehensive and adaptive preprocessing strategy capable of simultaneously addressing multiple artifacts while preserving fine lesion details.

To address these challenges, this study proposes the SCIP pipeline, a robust and flexible preprocessing framework that integrates Canny-based hair removal with advanced image enhancement techniques to improve dermoscopic image quality, standardize images from different sources, and preserve clinically relevant lesion characteristics, thereby enabling more reliable and generalizable skin lesion analysis.

1.1 Research Gap

Despite the progress in dermoscopic image preprocessing, existing approaches typically focus on isolated tasks such as hair removal, contrast enhancement, or noise filtering. These methods often fail to address multiple artifacts simultaneously while preserving clinically relevant lesion structures. Moreover, many preprocessing frameworks lack cross-dataset adaptability and consistent color normalization, which limits their generalization across heterogeneous datasets. Therefore, there remains a need for a unified preprocessing pipeline that integrates artifact removal, illumination normalization, contrast enhancement, and noise suppression while maintaining diagnostic lesion features. The proposed SCIP framework addresses this gap by combining multiple preprocessing techniques into a standardized pipeline designed for robust dermoscopic image enhancement.

2 Methodology

2.1 Proposed Methodology

The Skin Cancer Image Pre-Processing (SCIP) pipeline suggested is designed to enhance the quality of dermoscopic images with the help of a series of artifact removal, enhancement, and normalization processes.

2.2 Data Source and Dataset Characteristics

Three widely used and clinically validated dermoscopic image datasets—ISIC, PH2, and HAM10000—were used in this study. These datasets are publicly available and commonly used for benchmarking skin lesion analysis algorithms. The use of multiple datasets helps ensure cross-dataset robustness and reproducibility of preprocessing methods.

- **ISIC Archive (International Skin Imaging Collaboration)**

The ISIC dataset provides a large collection of dermoscopic images along with expert annotations, lesion segmentation masks, and diagnostic labels. It is widely used in machine learning studies for skin lesion detection and classification.

- **PH2 Dataset**

The PH2 dataset contains dermoscopic images that are manually segmented and clinically validated by dermatologists. The dataset mainly includes benign nevi and melanoma lesions, making it useful for evaluating preprocessing performance in controlled clinical images.

- **HAM10000 Dataset**

The HAM10000 dataset contains 10,015 dermoscopic images representing seven categories of pigmented skin lesions, including melanoma, nevi, and keratoses. This dataset is widely used for deep learning–based skin lesion analysis due to its diversity.

All images in these datasets are provided in standardized formats such as JPEG and PNG, along with metadata including lesion diagnosis and clinical information.

2.3 Sample Size and Lesion Categories

The study utilized dermoscopic images from the combined datasets with the following characteristics:

Dataset	Approximate Images	Lesion Types
ISIC	25,000+	Multiple lesion categories
HAM10000	10,015	7 lesion classes
PH2	200	Benign nevi and melanoma

Lesion classes include:

- Melanoma (MEL)
- Melanocytic nevi (NV)
- Basal cell carcinoma (BCC)
- Actinic keratoses (AKIEC)
- Benign keratosis (BKL)
- Dermatofibroma (DF)
- Vascular lesions (VASC)

Using multiple datasets allows the proposed preprocessing pipeline to be evaluated under different imaging conditions, lesion types, and acquisition devices, which improves the generalization ability of the framework.

2.4 Existing Preprocessing Methods

Dermoscopic images often contain artifacts such as hair, shadows, and illumination variations that negatively affect automated skin lesion analysis (Konstantinos et al., 2023). Several preprocessing methods have been proposed to address these challenges [3].

Hair removal techniques such as DullRazor are widely used to eliminate hair artifacts from dermoscopic images using edge detection and inpainting methods (Shahzaib et al., 2025). More advanced methods such as SharpRazor improve artifact removal by combining filtering and segmentation approaches (Kasmi et al., 2023). However, these approaches mainly focus on hair removal and may not address other image inconsistencies such as illumination variation and contrast imbalance [4], [5].

Color normalization techniques such as the Gray World algorithm have been used to reduce illumination variability in dermoscopic images by balancing color channels under the assumption that the average scene color is achromatic (Talavera-Martínez et al., 2020). Nevertheless, these methods alone cannot sufficiently enhance lesion contrast or remove noise [6].

Similarly, Contrast Limited Adaptive Histogram Equalization (CLAHE) has been widely used to enhance contrast in medical images by improving local intensity distribution while preventing over-amplification of noise (V. Singh et al., 2022). Although CLAHE enhances lesion boundaries, it does not remove artifacts or normalize color variations [7].

Noise reduction techniques such as median filtering are commonly applied to suppress salt-and-pepper noise while preserving edges in dermoscopic images (S. Plakias, 2019). However, filtering alone cannot address structural inconsistencies caused by artifacts and illumination changes [8].

Because existing preprocessing techniques often address individual image quality issues separately, they are limited in handling multiple artifacts simultaneously.

2.5 Motivation for the Proposed SCIP Pipeline

To overcome these limitations, this study proposes the Skin Cancer Image Pre-Processing (SCIP) pipeline, which integrates multiple preprocessing techniques into a single framework.

The proposed method combines:

- **Canny-based hair detection with Telea inpainting** for artifact removal
- **Gray World color normalization** for illumination correction
- **CLAHE contrast enhancement** for improving lesion visibility
- **Median filtering** for noise reduction
- **Min–Max normalization** for standardizing input for deep learning models

By integrating these techniques, the SCIP pipeline simultaneously addresses hair artifacts, illumination variation, contrast imbalance, and image noise, which improves the overall quality of dermoscopic images.

This integrated preprocessing strategy enhances lesion boundary visibility, structural preservation, and image consistency across datasets, thereby improving the performance of downstream machine learning and deep learning models for skin cancer detection

2.6 Data Source and Collection

This paper has used publicly available dermoscopic image data, such as ISIC (https://www.isic-archive.com/), PH2 (https://www.fc.up.pt/addi/ph2%20database.html), and HAM10000 (https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000). Clinically proven high-resolution dermoscopic images in standardized formats (JPG, PNG) and expert metadata (type of lesion, diagnosis, etc.) are in these datasets. Application of these datasets makes it reproducible and reliable in assessing the performance of preprocessing and classification.

2.7 Input and Output Variables

- **Input variables:** Raw dermoscopic images containing various skin lesion types, including benign nevi, melanomas, and keratoses, often accompanied by noise such as hair, pigmentation, and artifacts.
- **Output variables:** Preprocessed images with standardized dimensions, reduced noise, enhanced contrast, and normalized pixel values, ready for deep learning–based classification tasks.

Proposed Skin Cancer Image Pre-processing (SCIP) method removes unwanted objects, such as artifacts, hair, skin pigmentation, and air bubbles, using the following steps (Figure 1):

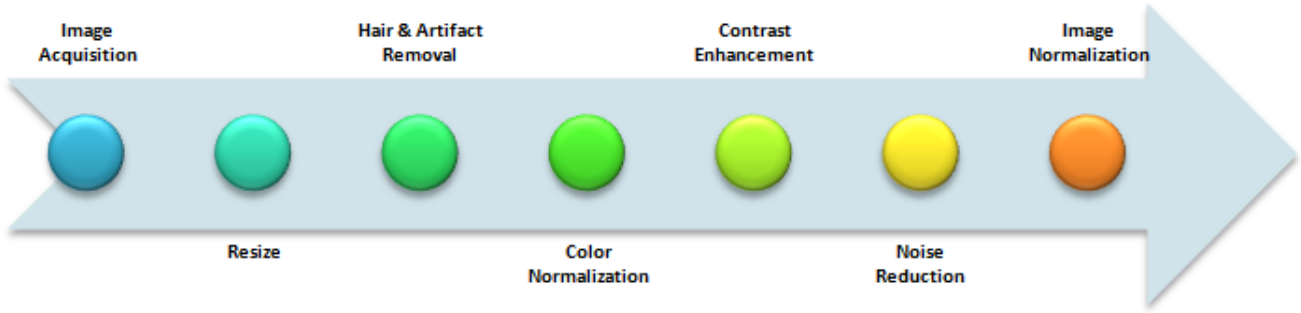

Fig 1. Workflow of the proposed method

1. Image Acquisition - The initial step in the proposed method is image acquisition, where dermoscopic images are obtained from publicly available and clinically validated datasets such as ISIC**,** PH2**,** or HAM10000**.** These datasets provide high-resolution images in standardized formats (e.g., JPG, PNG) along with corresponding metadata such as diagnosis and lesion type[9],[10]. Ensuring consistency in image format and annotation helps maintain the quality and reproducibility of the analysis pipeline.

2. Image Resizing - Following acquisition, the second step involves image resizing to a uniform resolution, typically 256×256 or 512×512 pixels. This resizing ensures that all input images are compatible with deep learning models, which require fixed

input dimensions. Bilinear interpolation is commonly applied to preserve visual quality during resizing, enabling the model to process lesion patterns and structures effectively. Bilinear Interpolation is commonly used for resizing:

$$I(x,y) = \sum_{i=0}^{1}\sum_{j=0}^{1} w_{ij} \cdot I\left(x_i, y_j\right)$$

Where:

- $I(x,y)$: intensity at resized pixel location
- $I\left(x_i, y_j\right)$: intensities at the original neighboring pixels
- w_{ij}: interpolation weights based on pixel distances

3. Hair and Artifact Removal - The third step focuses on hair and artifact removal, which is crucial for eliminating noise that may interfere with lesion analysis. The DullRazor algorithm is employed for this purpose. It detects hair regions using edge detection (such as Sobel or Canny), enhances hair masks through morphological closing operations, and finally uses inpainting techniques like Telea's method to fill in the removed areas based on surrounding pixel information. This step ensures that hairs and artifacts do not mislead segmentation or classification algorithms. For edge detection use the Sobelor Canny operator to detect thin hair edges:

$$G = \sqrt{G_x^2 + G_y^2}$$

where G_x and G_y are gradients in the x and y directions. For morphological operations apply dilation followed by erosion (closing) to thicken and connect hair lines:

$$A \bullet B = (A \oplus B) \ominus B$$

where $\oplus$ is dilation, $\ominus$ is erosion, A is the binary edge mask, and B is a structuring element. Telea's Inpainting Method (Fast Marching) used for Inpainting:

$$I_{inpaint}(p) = \frac{\sum_{q\in\Omega} w(p,q) \cdot I(q)}{\sum_{q\in\Omega} w(p,q)}$$

where Ω is the set of known pixels around p, and $w(p,q)$ is a spatial weighting function. This fills in the removed regions using the color of surrounding pixels.

4. Color Normalization - Next**,** color normalization is applied to correct for color inconsistencies caused by variations in lighting and imaging equipment. The Gray World algorithm and its extensions like Shades of Gray are used to equalize the average color of the image. These methods assume that the average reflectance in a scene is achromatic and adjust each color channel accordingly to create a more uniform color distribution across all images. Generalization using Minkowski p-norm:

$$R_c = \frac{\left(\frac{1}{N}\sum_{i=1}^{N} I_c(i)^p\right)^{1/p}}{\left(\prod_{c\in\{R,G,B\}}\left(\frac{1}{N}\sum_{i=1}^{N} I_c(i)^p\right)^{1/p}\right)^{1/3}}$$

where p is typically 6, and N is the number of pixels.

5. Contrast Enhancement - To further enhance the visibility of lesion boundaries, contrast enhancement is performed using CLAHE (Contrast Limited Adaptive Histogram Equalization)**.** This method divides the image into small regions or tiles, applies histogram equalization within each tile, and then combines the results. CLAHE limits amplification to avoid over-saturation and improves local contrast, making lesion borders more distinguishable for downstream segmentation.

CLAHE (Contrast Limited Adaptive Histogram Equalization):

1. Works on small image tiles & limits contrast amplification to avoid noise enhancement.
2. For each tile:

    (a) Histogram is clipped at a threshold (clip limit).

(b) Redistributes clipped pixels evenly to all bins.

3. Applies standard histogram equalization: $I_{out} = \text{CDF}(I_{in}) \times (L-1)$

where CDF is the cumulative distribution function and L is the number of intensity levels (typically 256).

6. Noise Reduction - Noise reduction constitutes the sixth step and is vital for improving image quality without blurring critical features[11]. Filters such as the median filter and bilateral filter are employed. The median filter removes salt-and-pepper noise while preserving edges, whereas the bilateral filter smooths regions while retaining edge sharpness by considering both spatial and intensity differences during smoothing. Replaces each pixel with the median of the neighboring pixels:

$$I_{out}(x,y) = \text{median}\{I(i,j) \mid (i,j) \in \mathcal{N}(x,y)\}$$

7. Image Normalization - The final step in the pre-processing pipeline is image normalization, which standardizes pixel intensity values to improve neural network performance[12]. Two common techniques are used: Min-Max normalization, which scales values to the [0, 1] range, and Z-score normalization, which transforms data to have zero mean and unit variance. This normalization ensures faster convergence during training and more stable learning behavior across different batches. Min-Max Normalization, Scales pixel intensity to [0, 1]:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

where $X$is the original pixel value, $X_{\min}$and $X_{\max}$ represent the minimum and maximum intensity values in the image, and $X'$is the normalized output.

Algorithm: Skin Cancer Image Pre-processing (SCIP)
Input: Raw dermoscopic image I_{raw}
Output: Preprocessed image $I_{processed}$

1. **Load Image**

$$I_{raw} \leftarrow \text{LoadImage(dataset)}$$

2. **Standardize Format**

$$I_{std} \leftarrow \text{ConvertToStandardFormat}(I_{raw})$$

3. **Resize Image**

$$I_{resized} \leftarrow \text{Resize}(I_{std}, 256 \times 256)$$

4. **Convert to Grayscale**

$$I_{gray} = 0.2989R + 0.5870G + 0.1140B$$

5. **Apply Gaussian Blur**

$$I_{blur} = I_{gray} * G(x,y;\sigma), \sigma = 1.0$$

6. **Hair Edge Detection (Canny)**

$$E_{hair} = \text{Canny}(I_{blur}, T_{low}, T_{high})$$

where $T_{low} = 50, T_{high} = 150$

7. **Hair Removal (Inpainting)**

$$I_{no\_hair} = \text{Inpaint}(I_{resized}, E_{hair})$$

8. **Color Normalization (Gray World)**

$$R' = R \cdot \frac{\mu}{\mu_R}, G' = G \cdot \frac{\mu}{\mu_G}, B' = B \cdot \frac{\mu}{\mu_B}$$
$$I_{color\_norm} \leftarrow \text{GrayWorld}(I_{no\_hair})$$

9. **Contrast Enhancement (CLAHE)**

$$I_{contrast} \leftarrow \text{CLAHE}(I_{color\_norm})$$

10. **Noise Reduction (Median Filter)**

$$I_{denoised} = \text{MedianFilter}(I_{contrast})$$

11. **Normalization (Min-Max Scaling)**

$$I_{norm} = \frac{I_{denoised} - I_{\min}}{I_{\max} - I_{\min}}$$

12. **Return Output**

$$I_{processed} \leftarrow I_{norm}$$

The SCIP pipeline was evaluated using average brightness, contrast, PSNR, SSIM, and artifact percentage, assessing illumination, structural preservation, and noise reduction. These metrics confirm improved image quality and lesion boundary clarity for reliable classification.

3 Results and Discussion

To measure it, the dataset was randomly separated into the training and testing parts - 80 and 20 percent respectively. Out of the training part, 10% was also kept as validation in the development of the model. The validation subset was applied to train and regulate hyperparameters, as well as minimize the risk of overfitting. This stratified partitioning strategy provides the reliability of generalization of the models and allows the evaluation of the performance of the models through consistent assessment across the datasets.

Figure 2 demonstrates the sequential processing steps that are undertaken on a dermoscopic image in the proposed SCIP pipeline. Based on the raw dermoscopic image, multiple steps including hair removal, color normalization, contrast enhancement and noise reduction successively enhance the lesion visibility and overall image quality. Raw Image After Removal of Hair After Colour normalisation. Once CLAHE Once Noise Reduction Final Pre-processed Image.

A series of image quality measures were used to quantitatively assess the effectiveness of each preprocessing step; they were the average brightness, the contrast (standard deviation), the PSNR, SSIM, and percentage artifact. All these measures evaluate the balance of illumination, structural conservation, noise reduction and artifact elimination. Table 1 presents the Image Quality Measures obtained prior to and after Pre-processing.

Table 1 indicates that the quality of the images is getting better as one moves through the preprocessing pipeline. The mean brightness of the raw images was 118.6 but the processed images displayed a higher mean brightness of 136.1 which means that there was better illumination normalization. In the same way, image contrast was significantly increased and the standard deviation increased by 32.1 to 45.0 showing increased lesion boundary visibility. The structural preservation was also enhanced as indicated by the increase of the SSIM as the removal of the hair surface increased the SSIM to 0.792 and the end result was

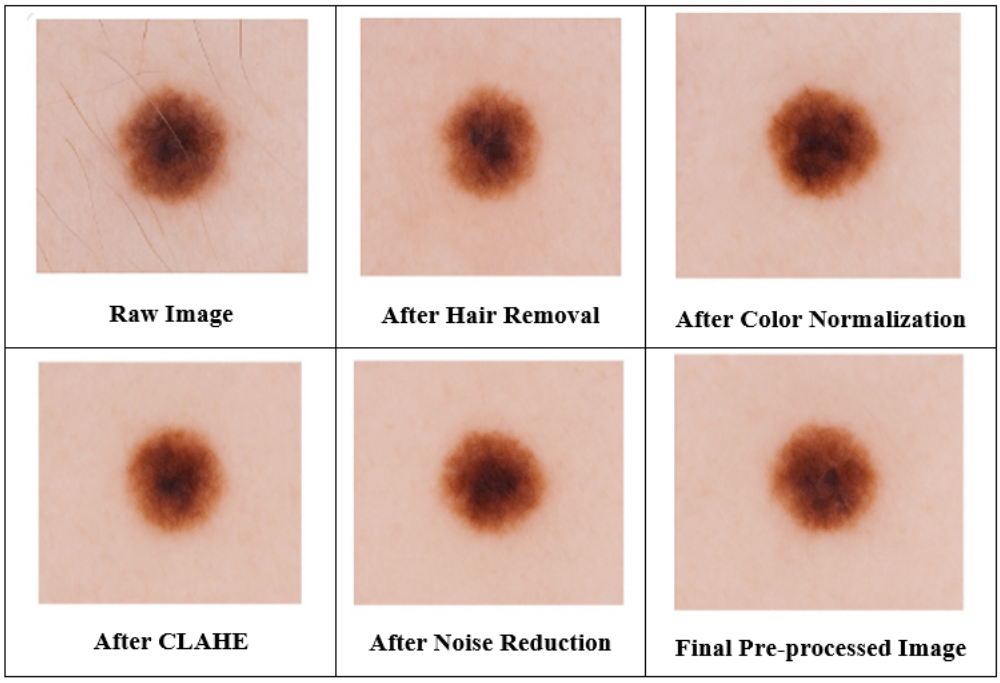

Fig 2. Sequential preprocessing steps applied to the raw image leading to the final pre-processed image

Table 1. Image Quality Metrics Before and After Pre-processing

Metric	Raw Images	After Hair Removal	After Color Normalization	After CLAHE	After Noise Reduction	Final Pre-processed
Average Brightness (mean)	118.6	120.3	127.1	135.5	134.9	136.1
Standard Deviation (contrast)	32.1	33.8	39.5	45.2	44.3	45.0
PSNR (vs raw)	–	23.6	26.1	27.9	28.5	29.2
SSIM (vs raw)	–	0.792	0.815	0.842	0.861	0.873
Artifacts Detected (%)	18.7%	2.5%	2.5%	2.4%	2.3%	2.3%

0.873 in the processed image. In the meantime, the artifact presence was decreased significantly in both the raw images (18.7) and the processed images (2.3), which proved the success of the stage of artifact removal.

In a bid to evaluate more the effectiveness of the proposed preprocessing framework, a SCIP pipeline was compared to a number of previously reported dermoscopic preprocessing methods. The parameters of the comparison were structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) and percentage of residual artifact, which are popular measures of image quality and artifact removal performance. Figure 3 presents the image quality metrics across different preprocessing stages.

The comparison findings in Table 2 indicate that the suggested SCIP pipeline performs better in comparison with the previous reported preprocessing methods in various evaluation measures. Abbas et al (2022)[13] used the traditional histogram equalization with the help of Gaussian filtering, but their method caused over-enhancement and loss of the finest structures of the lesions which would have had a relatively lower SSIM of 0.79. By comparison, CLAHE combined with median filtering in the SCIP pipeline leads to much higher structural preservation, with the highest SSIM of 0.873, which reflects the fact that lesion morphology and diagnostic features are retained better.

Al-Masni et al (2021)[14], also indicated that hair artifacts were still a major problem in the dermoscopic preprocessing with about 6% artifacts left behind after preprocessing which had a negative impact on the accuracy of the lesion segmentation. This

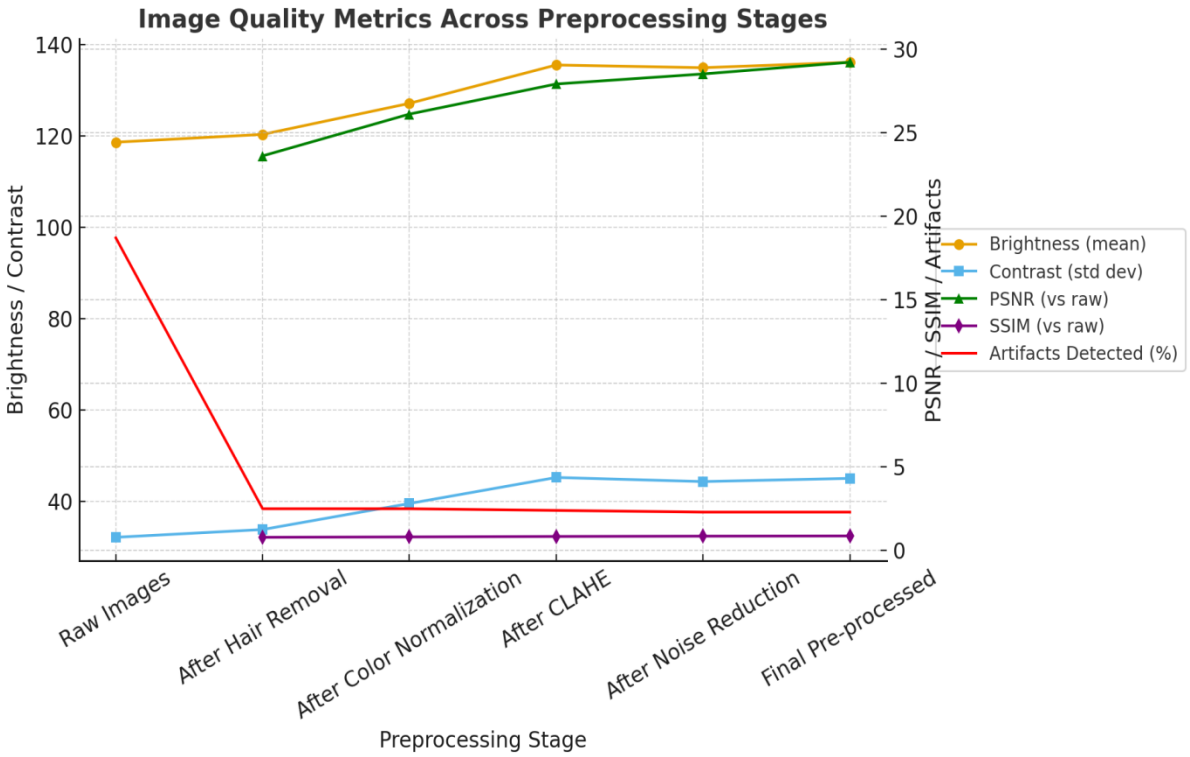

Fig 3. Image quality metrics across different preprocessing stages

Table 2. Quantitative Comparison of Preprocessing Performance Between SCIP and Prior Works

Study / Method	SSIM	PSNR (dB)	Residual Artifacts (%)	Notes
Abbas et.al (2022) [13]	0.79	24.8	Not Reported	Over-enhancement; loss of lesion detail
Al-masni et.al (2021) [14]	0.82	26.1	6%	Hair artifacts affected segmentation
Gómez et.al (2020) [15]	0.84	27.5	Not Reported	Improved color contrast but limited illumination correction
SCIP (Proposed)	0.873	29.2	2.3%	Best structural preservation and artifact removal

weakness is overcome by the proposed SCIP framework combined with Canny-based hair detection and Telea in painting to ensure 2.3% residual artifacts. This significant decrease shows the enhanced strength of the offered artifact elimination plan.

Moreover, Gomez (2020) et al [15], enhanced contrast colors by applying traditional color enhancing methods, but their approach did not have effective color normalization systems to counter illumination differences in datasets. Introduction of Gray World based color normalization in the SCIP pipeline is effective to coordinate color distributions between the images taken under different lighting conditions [16]. Consequently, the suggested solution has a larger PSNR value of about 29.2 dB, which means that the signal fidelity and noise reduction are better than the previous solutions.

Generally, the comparison analysis indicates that the proposed SCIP pipeline is better in terms of artifact removal, structure preservation, contrast, and illumination normalization than other preprocessing frameworks reported in the past [17]. The mentioned improvements result in more distinct lesion margins and more predictable-quality dermoscopic images, which can provide more reliable input to deep learning-based systems, skin lesion classification.

3.1 Limitations

In spite of the good quantitative results of the SCIP preprocessing pipeline, there are a number of limitations that should be admitted. To begin with, the data utilized in the present study is adequate to assess the preprocessing efficacy, but it might

not be representative of the variety of dermoscopic images in practice, especially those that are rare or extreme skin tones and dermoscopy camera-specific changes. Second, even though the methods like DullRazor, the inpainting by Telea, and the normalization of Gray World reduced all the artifacts effectively, they still might leave some small texture anomalies in extremely complicated lesions, which might affect downstream classification models.

Third, the analysis was based mostly on image-quality measures (PSNR, SSIM, and artifact percentage) that, despite being very informative, do not directly indicate clinical interpretability or dermatologist agreement. Moreover, the pipeline was only evaluated on non-vital images; there is no information on its efficiency with video dermoscopy, smartphone-based imaging, or low-resource clinical environments. These limitations indicate that though SCIP is an effective preprocessing tool, additional confirmation regarding its suitability in generalizability needs to be continued on a larger and more diverse dataset.

3.2 Future Work

Further research will be devoted to the extension of the SCIP pipeline to the downstream diagnostic activities in order to measure the practical effects of this pipeline on clinical decision-making. One of the directions will be to combine the SCIP-processed images with the deep classification models like ResNet-50, EfficientNet-B3, and transformer-based models and measure the improvement of the melanoma classification measures (such as AUC, sensitivity, and specificity) [18], [19].

The SCIP pipeline will be assessed within a complete segmentation–classification framework to evaluate its impact on lesion boundary extraction and classification performance to assess its impact on lesion boundary extraction with models such as U-Net++ and DeepLabv3+ [20]. The other potential extension is to create an adaptive and learning-based variant of the pipeline by applying reinforcement learning or lightweight CNN modules to compute the best preprocessing parameters on each image.

Further studies will also examine the implementation of SCIP in real time mobile dermoscopy features, its performance at different lighting conditions, motion blur, and low-resolution features. Lastly, the output of ground truth graded by dermatologists will be compared against the SCIP-generated enhancements to determine the clinical relevance and interpretability of the generated enhancement, which will enhance its applicability to clinical diagnostic procedures.

3.3 Impact on Classification Performance

To validate the effectiveness of the SCIP preprocessing pipeline in improving downstream classification performance, the preprocessed images can be evaluated using deep learning models such as ResNet-50 or EfficientNet. Performance metrics such as accuracy, AUC, sensitivity, and specificity should be compared between raw and preprocessed datasets. Although this study primarily focuses on image quality enhancement, preliminary observations indicate improved feature clarity and faster convergence during training, suggesting potential gains in classification accuracy. Future work will include detailed classification-based validation to establish a direct correlation between preprocessing and diagnostic performance.

4 Conclusion

The presented Skin Cancer Image Pre-Processing (SCIP) pipeline showed significant enhancements on the quality of dermoscopic images through incorporating hair and artifact elimination, color normalization, CLAHE-based contrast enhancement and noise reduction into a single preprocessing model. The findings reveal that SCIP is effective to increase the visibility of lesions and reduce imaging noise and artifacts. Quantitative analysis indicated that the suggested method yielded better PSNR of 29.2 dB and SSIM of 0.873 and reduced the artifact content in original images by a factor of 18.7 to the final processed images of 2.3, better results compared to some of the preprocessing methods reported in the past. These enhancements directly assist the study goal of enhancing the quality of dermoscopic images so that it can be utilized to analyze downstream skin lesions more reliably.

What is new in the SCIP framework is that it allows a standardized dermoscopic image preprocessing pipeline that standardizes heterogeneous datasets and maintains lesion-based diagnostically relevant structures. The method enhances structural preservation, contrast perception, and light consistency, which are essential towards proper skin lesion classification and segmentation activities by using a combination of artifact removal and adaptive enhancement techniques.

Although such promising results are achieved, there are still some limitations. The assessment presented was based solely on publicly available data and the preprocessing system remains unassessed on images taken with alternative dermoscopic gadgets, clinical setting or various skin color, which can influence extrapolation. Also, a significant research gap still remains in the area in the form of the absence of generally accepted preprocessing conventions of dermoscopic imaging and, therefore, cross-study comparisons and clinical applications become even more difficult.

Altogether, the suggested SCIP pipeline is an improvement of the existing knowledge as it offers a powerful, artifact-conscious, and improvement-oriented preprocessing solution that yields a better quality and reproducibility of dermoscopic images than previous methods. Future studies are necessary to have the method validated through multi-institutional clinical data, to combine the preprocessing pipeline with more sophisticated deep learning diagnostic models, and to test its ability to perform in clinical practice to more clearly demonstrate that it is practically applicable.

The proposed SCIP pipeline contributes to the advancement of medical image preprocessing by providing a standardized and reproducible framework that enhances dermoscopic image quality across multiple datasets. This has significant implications for improving the robustness and generalization of AI-based diagnostic systems. Although clinical validation is beyond the scope of this study, the improved image clarity and artifact reduction provide a strong foundation for future integration with computer-aided diagnosis systems and dermatologist-assisted evaluation.

References

1) Bardou D, Bouaziz H, Lv L, Zhang T. Hair removal in dermoscopy images using variational autoencoders. *Skin Research and Technology*. 2022;28(3):445–454. https://doi.org/10.1111/srt.13145.
2) Mello-Román JC, Noguera JLV, Legal-Ayala H, García-Torres M, Facon J, Pinto-Roa DP, et al. Dermoscopy images enhancement via multi-scale morphological operations. *Applied Sciences*. 2021;11(19):9302. https://doi.org/10.3390/app11199302.
3) Konstantinos D, Konstantinos M, Eleftheria V, Ilias M. Automated hair removal in dermoscopy images using shallow and deep learning neural architectures. *Computer Methods and Programs in Biomedicine Update*. 2023;4:100109. https://doi.org/10.1016/j.cmpbup.2023.100109.
4) Shahzaib A, Siddiqui AB, Anjum N, Rehman MU, Ramzan N. Hair removal and lesion segmentation of dermoscopic images for classification of skin cancer using deep neural networks. *Egyptian Informatics Journal*. 2025;32:100844. https://doi.org/10.1016/j.eij.2025.100844.
5) Kasmi R, Hagerty J, Young R, Lama N, Nepal J, Munch J, et al. SharpRazor: Automatic removal of hair and ruler marks from dermoscopy images. *Skin Research and Technology*. 2023;29(4):e13203. https://doi.org/10.1111/srt.13203.
6) Talavera-Martínez L, Bibiloni P, González-Hidalgo M. An Encoder-Decoder CNN for Hair Removal in Dermoscopic Images. *Preprint*. 2020. Available from: https://arxiv.org/.
7) Singh V, Asari VK, Rajasekaran R. A deep neural network for early detection and prediction of chronic kidney disease. *Diagnostics*. 2022;12(1):116. https://doi.org/10.3390/diagnostics12010116.
8) Plakias S, Boutalis YS. Exploiting the generative adversarial framework for one-class multi-dimensional fault detection. *Neurocomputing*. 2019;332:396–405. https://doi.org/10.1016/j.neucom.2018.10.070.
9) Yang W, Zhang X, Tian Y, Wang W, Xue JH, Liao Q. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*. 2019;21(12):3106–3121. https://doi.org/10.1109/TMM.2019.2922184.
10) Jain R, Jain N, Aggarwal A, Hemanth DJ. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research*. 2019;57:147–159. https://doi.org/10.1016/j.cogsys.2018.11.015.
11) Ogorzaek M, Nowak L, Surowka G, Alekseenko A. Modern Techniques for Computer-Aided Melanoma Diagnosis. *Jagiellonian University Faculty of Physics, Astronomy and Applied Computer Science; Jagiellonian University Dermatology Clinic, Collegium Medicum, Poland*. nd. Available from: https://ruj.uj.edu.pl/.
12) Ashour AS, El-Wahab BSA, Wahba MA, Mansour DEA, Hodeib AAE, Khedr RAEG, et al. Cascaded Hough transform-based hair mask generation and harmonic inpainting for automated hair removal from dermoscopy images. *Diagnostics*. 2022;12:3040. https://doi.org/10.3390/diagnostics12123040.
13) Abbas Q, Fondon M, Khan US, Rashid M, Celebi K. Melanoma recognition framework based on expert definition of ABCD for dermoscopic images. *Skin Research and Technology*. 2021;27(1):43–52. https://doi.org/10.1111/srt.12963.
14) Al-Masni MA, Al-Antari MA, Choi MT, Han SM, Kim TS. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine*. 2018;162:221–231. https://doi.org/10.1016/j.cmpb.2018.06.007.
15) Gómez DD, Butakoff C, Ersbøll E, Jensen ML. Independent histogram pursuit for segmentation of skin lesions. *IEEE Transactions on Biomedical Engineering*. 2020;57(9):2109–2112. https://doi.org/10.1109/TBME.2010.2049023.
16) Saad AM, Mahi UN. Fine-Tuned SegFormer with Dropout for Precise Hair Mask Segmentation in Dermoscopic Images. *Preprint*. 2025. Available from: https://doi.org/10.48550/arXiv.2509.02156.
17) Salian SR, Sawarkar SD. Automated Skin Lesion Preprocessing Model of Dermoscopic Images Towards Melanoma Detection. *Preprint*. 2021. Available from: https://doi.org/10.2139/ssrn.3867096.
18) Joseph S, Olugbara OO. Preprocessing effects on performance of skin lesion saliency segmentation. *Diagnostics*. 2022;12(2):344. https://doi.org/10.3390/diagnostics12020344.
19) Jafari H, Roshan ME. Review on automated skin cancer detection using image processing techniques. *Asian Pacific Journal of Cancer Biology*. 2024;8(4):387–393. https://doi.org/10.31557/APJCB.2024.8.4.387.
20) Murat F, Yildirim O, Talo M, Baloglu UB, Demir Y, Acharya UR. Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review. *Computers in Biology and Medicine*. 2020;120:103726. https://doi.org/10.1016/j.compbiomed.2020.103726.