Received: March 15, 2023. Revised: April 21, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

A Novel Skin and Mole Pattern Identification Using Deep Residual Pooling Network (DRPN)

Rohan Don Salins[1]* **G Ananth Prabhu[1]** **Jason Elroy Martis[2]** **Sannidhan M S[2]**

[1]*Sahyadri College of Engineering and Management, Mangaluru, Karnataka-575007, India*
[2]*NMAM Institute of Technology (Nitte Deemed to be University), Udupi, Karnataka-574110, India*
* Corresponding author's Email: rohan25nov@gmail.com

Abstract: The process of identifying a criminal is extremely challenging if no eyewitnesses are available or closed-circuit television (CCTV) footage does not show the criminal's face. With the use of high-definition cameras and soft biometrics, forensic science has made great strides in identifying criminals based on their bodily traits. Due to the use of one type of soft biometric for identification, many researchers have focused on this area and have failed to achieve better matching. Despite the fact that similar soft biometrics share many of the same features, one type of soft biometric matching has been insufficient and resulted in inaccurate results. As a result, it becomes difficult to make distinct classifications. Therefore, this study proposes the implementation of a suspect identification system that uses mole pattern analysis and skin tone matching as soft biometrics. The deep residual pooling network is employed to analyze mole patterns, while the histogram equalization method is employed to analyze skin tones. We have also improved the performance of the deep seg-net and the histogram equalization technique by incorporating an aggregator operator and an adaptive pixel-wise noise cancellation filter. The experimental results show that the proposed method attains higher accuracy when the loss rate is gradually decreased. The proposed method attains 89.21% accuracy which is higher than the existing approaches and the efficiency of the proposed method is compared with various texture encoders.

Keywords: Deep residual pooling network, Histogram equalization, Mole identification, Semantic segmentation, Skin tone analysis, Soft biometrics.

1. Introduction

Identification of the suspect through valid evidence is still challenging for any crime investigating agency when there is no eyewitness involved in the criminal activity. However, with the advancements in science and technology, many CCTV cameras with very high-resolution are being mounted everywhere to track various illicit activities in the surroundings [1]. With the introduction of soft biometric technology, it is now possible to identify the offender even if one's face is not visible. Soft biometrics usually involves matching hair patterns, an individual's age, sex, skin texture, mole, etc. [2].

The application of artificial intelligence with better resolution cameras further boosted the soft biometric identification process using computer vision techniques. They are good enough to capture significantly finer details of soft biometrics such as vein patterns, skin texture, moles, and androgenic hairs. [3] These captured details can be further classified using deep learning techniques to match them against a set of available soft biometric data for the possible identification of any individual who might even be a preparator of any incident. [4] Few current research works utilized the application of soft biometric matching using deep learning techniques that focused on matching either hair patterns or moles [5, 6]. Hence to gain better accuracy, the system demands an identification technique that combines the matching of multiple soft biometrics to accomplish the identification process [7, 8]. With the next level of advancements in deep learning techniques, mole detection algorithms have gained better prospects in achieving better identification accuracy [9, 10]. Hence, to overcome the existing system's drawback and to incorporate the matching of multiple soft biometrics, we propose a novel criminal

identification system that relies on a mole identification technique using a combination of mole texture analysis and skin tone matching [11-13].

The proposed system is designed in such a way that it is well-suited to scenarios when the face is completely hidden by exposing other bodily parts of an individual. This work targets separating moles from the skin area to improve detection accuracy to analyze the mole's texture using a dedicated deep learning encoder network [14, 15]. The scores obtained from mole texture analysis and skin tone analysis are combined with mapping it against a set of mole samples existing in the adopted dataset.

To design this proposed system, the following are the key contributions of this research article.

- Incorporating a pre-processor with standard noise filtering techniques to filter out different types of noise in the mole image sample.
- Implement a deep learning-based semantic segmentation approach to differentiate and isolate mole from the skin area.
- Design of a deep residual pooling network to perform texture analysis.
- Application of skin histogram equalization technique for skin tone analysis.

The rest of this research is organized as follows: section 2 discusses the details of a related study and its results that motivate the design of the proposed system. Section 3 provides a detailed explanation of the proposed system, including the materials, methods, and formulations used. A comparative analysis of the proposed system is presented in section 4, along with the results of different experimental investigations. In section 5, we present the conclusions of our research work portraying our accomplishments.

2. Literature review

Mole detection on the skin is a common method for detecting skin lesions of various lengths. A limited amount of research has concentrated on identifying skin lesions using mole detection methods. A study [16] used fuzzy k-means clustering segmentation techniques to identify moles on the skin. The segmentation technique generates a score that categorizes skin lesions from normal moles. In the study, a classification accuracy of more than 90% was achieved in the identification of lesions. The work, however, revealed limitations concerning asymmetrical boundaries in terms of identification. The article [17] suggested an approach for a skin lesion classification system based on deep learning techniques and collective intelligence system (CIS), which involves multiple convolutional neural networks (CNNs), trained on the HAM10000 dataset, which is able to predict seven skin lesions including melanoma. Article [18] suggests an approach for multi class skin lesion classification and detection by teledermatology. In this approach four datasets were used for evaluating the segmentation and the HAM10000 dataset was used for classification.

The article [19] presents an approach for localization of melanoma and classification. Faster-region-based CNN with support vector machine classifier is applied for deep feature extraction and detection of melanoma. The article [20], presented an approach for detecting skin disease and classification. This approach helps for detecting melanoma a severe skin cancer. Melanoma can occur anywhere on the skin. Unusual moles, exposure to sunlight, and health history can affect the risk of melanoma. The article [21], presents an approach for skin disease diagnosis by mask region-based CNN and super resolution generative adversarial network. The technical challenges in diagnosis of skin problem were addressed first. The article [22] suggests an approach for detection of facial spots with the help of CNN. This method helps to detect the face in various lighting condition. This method helps to detects several facial spots like freckles, scars, mole, etc.

A few factors that motivate our proposed system's design have been conclusively identified through a comprehensive literature review. Currently, very few research efforts have focused on the difference between moles and lesions. This has opened the scope for identifying moles by incorporating some techniques that have the potential to extend previous research to develop a mole identification and classification capability. A substantial amount of research has been done on skin tone analysis in multiple domains, from identifying individuals to analyzing skin lesions. An in-depth literature review confirms that the histogram equalization technique can be used to extract color features for further analysis. Hence, we adopted histogram equalization for skin tone patch color analysis. A significant problem with our proposed system is the extraction of texture information, and we found that a deep residual pooling network is the most appropriate one for this purpose. To isolate moles from the skin area, the research presented in this review recommended the use of semantic segmentation via deep seg-net architecture.

3. Proposed system

Considering the earlier works' drawbacks and

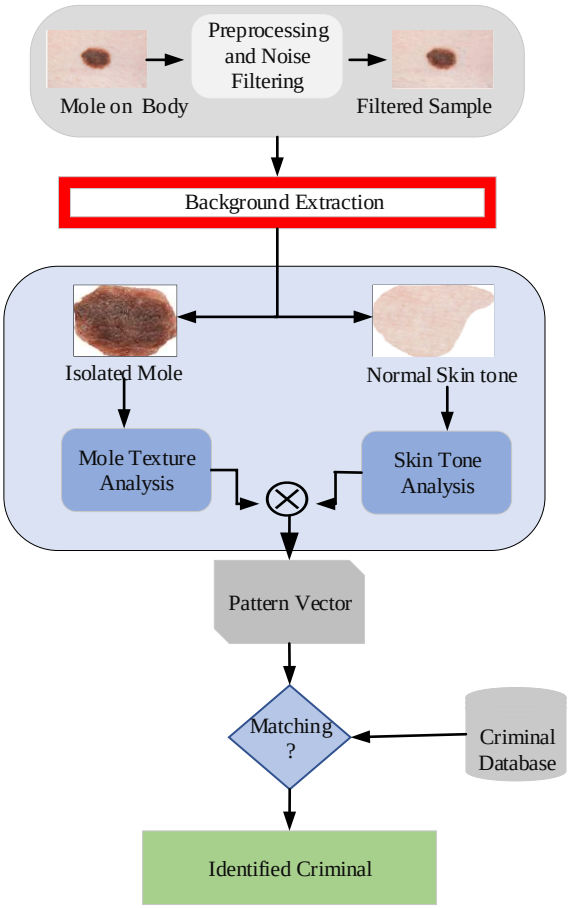


Figure 1. A simple architecture shows our proposed system for criminal identification workflow

upholding the literature, we propose an architecture for our criminal identification system, as shown in Fig. 1. As portrayed in Fig. 1, our proposed system is divided into five main modules: pre-processing, background elimination, mole texture analysis, skin tone analysis, and pattern matching. The detailed structure of this module will be explained in the subsequent sections.

3.1 Pre-processing stage

Pre-processing is the first stage of our criminal identification system, which includes noise-filtering strategies. In this approach, we eliminate all background noise from the system while capturing the image. The image consists of three types of noise: gaussian, salt-and-pepper, and poisson. To effectively filter out all of these types of noise, our filters must instigate and clean them. This is achieved by using a fusion-based filter that eliminates these noises [23]. Eqs. (1) and (2) are used for different types of noise structures.

$$P_{noise} \approx I_{RGB} \Leftarrow Camera \tag{1}$$

Here, P_{noise} is the transformation function that transforms the coordinate geometrical image into a noise model. I_{RGB} indicates the image taken from the camera in spatial form. Eq. (2) represents the type of filter used for a specific type of noise. σ^2 represents the global variance from the image. μ indicates the global mean. P_a and P_b indicates the extreme intensity of the salt and pepper noise pixels.

$$P_{noise} = \begin{cases} \sqrt{\frac{1}{2\pi\sigma^2}} \times e^{\frac{-(x-\mu)^2}{2\sigma^2}}, for\ Gaussian\ noise \\ P_a, for\ noise = a \\ P_b, for\ noise = b, for\ salt\ and\ pepper\ noise \\ \frac{g^k ie^{-noise}}{k!}, for\ poisson\ noise \end{cases} \tag{2}$$

g^k represents the mean value of the pixels under k intensity levels. The total value of k is 256. Since our P_{noise} function stands for a fused set of noises, and we used a sequential set of filters for noise filtration. To evaluate the quality of the filtered image obtained through the pre-processing stage, we have used three standard no-referential quality evaluators which are naturalness image quality evaluator (NIQE), blind/reference less image spatial quality evaluator (BRISQUE), perceptional image quality evaluator (PIQE) [24-26].

1) Naturalness image quality evaluator (NIQE)

The naturalness factor of an image is determined by this metric. The network generates score values ranging from 1 to 20 for each test image.

2) Blind/reference less image spatial quality evaluator (BRISQUE)

The quality of an image can also be evaluated by its noise level and other distortion parameters using this non-referential evaluator.

3) Perceptional image quality evaluator (PIQE)

PIQE generates scores between 0 and 100, with 0 marking highly exceptional and 100 marking the poorest performance.

3.2 Segmentation network

A deep seg-net is used to eliminate the contextual background and maintain the mole's integrity. To create such a network, we must collect an enormous number of images and manually isolate the background, creating a dataset of moles with and without backgrounds. The deep seg-net consists of an encoder network (EN) and decoder network (DN). There is a total of 13 convolutional layers in the EN. The encoder layer had a corresponding decoder layer,

so the DN also consists of 13 convolutional layers. For the final background-separated output, the output of the final decoder is given to a multi-class SoftMax classifier.

1) Encoder network

Every encoder performs convolution with a filter bank to generate a set of feature maps then these are batch normalized. Layers of max-pooling and sub-sampling may attain additional translation invariance for background separation correspondingly there is a feature maps spatial resolution loss.

2) Decoder network

The suitable decoder in the DN up samples, its input feature maps and utilizes the remembered max-pooling indices from the consistent encoder feature maps. The high-dimensional feature illustration at the outcome of the final decoder is fed to a soft-max classifier. This soft-max classifies every pixel independently.

3) Attention Layer (AL)

The attention layer is added to the deep seg-net for better segmentation. The AL only focuses on the important feature in the image, not the entire feature map. While adding the attention layer with deep seg-net we can attain better background removed the image as output.

4) Assessment of segmentation network

Our proposed segmentation network is evaluated by two standard statistical assessment parameters: the jaccard index and the dice coefficient [27]. These parameters are generally used for comparing two different sets, and we used them here to compare the similarity between the ground truth image and the segmented mole image.

3.3 Texture analysis

It is imperative to analyze the texture of the image after we have isolated the mole. Texture analysis is a complex process used to identify a specific vector graph that outliers the entire image. The number of texture score identifiers is innumerable, but we need our system to isolate the texture so that other similar texture forms can be separated from the original. In light of this, we constructed a modified deep residual pooling network structure to analyze mole textures.

- *Deep residual pooling network (DRPN)*

The DRPN can integrate with any CNN architecture, which helps to form a learning framework (end-to-end). The input given to the DRPN is the segmented output. The DRPN consists of two components such as; convolutional transfer module (CTM) and the aggregation module (AM). The aggregation module in the DRPN can generate features that have similar dimensions compared to the number of feature channels. Then the feature is given as classifier's input.

1) Convolutional transfer module

The CTM is utilized for learning the feature. The CTM had a convolutional layer that consists of 2048 kernels, the size of each kernel is 1×1, 1 stride with the activation function (sigmoid). To avoid over-fitting batch normalization (BN) and dropout are applied. In the residual maps for computing the spatial sensitive hard assignment outcomes Eq. (3) is applied,

$$A_{I,J} = S_{I,J} - \sigma\left(E_{I,J}\right), A_{I,J}, S_{I,J}\ and E_{I,J}\ \epsilon \mathbb{R}^{2048} \quad (3)$$

where the sigmoid function (SF) is mentioned as σ. The SF can attenuate the differences between learned and pre-trained features. This made CTM learn a novel feature that is appropriate for texture recognition.

2) Aggregation module

This is the second part of DRPN. The AM is utilized for residual aggregation for attaining an feature that is order less. The AM contains preceded BN and is followed by a rectified linear unit (ReLU) layer, that acts as a feature selector. The global average (GA) of AM is computed using Eq. (4),

$$GA = Agg(a)\ , GA \epsilon \mathbb{R}^{2048} \quad (4)$$

where $Agg(a)$ is the ReLU's composition, ReLU, BN and GA pooling respectively.

3) Residual encoding module

The final convolutional layer of pre-trained CNN is utilized as the dictionary for computing the residual [28].

3.4 Skin tone analysis

The previous sections explained how to read mole textures and outfit a score value, but there's a 30% chance that one mole will match another. Our system performs skin tone analysis in a two-fold technique which involves skin filtering and histogram analysis. The detailed structure is explained in the subsequent sections.

- *Skin filtering*

A skin filter works as a pre-processor to remove small hairs and black marks, or blemishes [29]. A pixel dominance method is used to extract blemishes. Using pixel summation, we filter the pixel count of blemishes based on a histogram plotting the overall

Table 1. Dataset division under different mole categories

Dataset	Resolution	Ratio Division	Set Name	Mole Category	Training Images	Testing Images	Validation Images	Total Images
HAM10000	1872×1053 px cropped to 800×600 px resized to 224×224 with 72 DPI	27.04%	Set 1	Melanoma	2163	270	271	2704
		48.75%	Set 2	Vascular Lesions	3900	487	488	4875
		7.15%	Set 3	Nevus	572	72	71	715
		17.06%	Set 4	Other Categories	1365	170	171	1706
	Total Images (Entire Dataset)				8000	999	1001	10000

pixel count across all pixels within the skin range. The immediate succeeding step is to perform smoothening of the image. We perform this step to avoid sharp hair and other matter texturing interrupting the skin tone analysis.

- *Histogram equalization*

The image is extracted using an RGB camera, which has limitations regarding lighting that can create varying shades of skin tone. The issue of skin tone can be vital when comparing criminals from different backgrounds. Furthermore, environmental, and physical conditions can also cause sunburn and skin darkening. To avoid these issues, we apply histogram equalization only across the entire low-resolution filtered image [30].

- *Score calculation*

At this stage, the filtered skin patch is compared with an existing skin patch stored in the criminal database. This comprehensive (bit-by-bit) comparison of the entire patch is performed by our process. To perform individual comparisons, it is mainly necessary to compare the same patches of skin in the database with the filtered patch. Because the photo/image/picture is extracted manually, there will be a degree of error that can be approximated by invariant rotation scaling.

- *Evaluation of skin tone analyzer*

To evaluate the operational capability of our proposed skin tone analyzer module, we have incorporated three standard evaluation procedures [31] explained below:

1) Peak signal-to-noise ratio (PSNR)

The evaluator is a statistical component that evaluates signal strength of any image under test in comparison with a reference image and produces a value ranging between 0 to 100 with 100 representing the maximum strength and 0 as minimum strength.

2) Mean squared error (MSE)

This evaluator basically reveals the amount of difference appearing in a test image in comparison to the reference image and generates value between 0 to 100.

3) Pixel error rate

The percentage of error is computed for each of the pixel in the filtered image and the value ranges between 0 to 100 with lowest value considered to be the best in terms of error rate.

3.5 Score level fusion

In the final stage of our proposed system, the texture encoder scores are fused with the scores generated in section 3. Our proposed system will produce a pattern vector using this fusion score. The pattern vector is then compared with a pattern vector previously stored in the criminal database. Pattern vectors are associative tagged arrays, as shown in Eq. (5).

$$P_v \leftarrow [[\{'texture'\},0.83],[\{'patch'\},0.34]] \quad (5)$$

$$C_{id} \leftarrow SearchDB(\langle P_v \rangle_{5\%}) \quad (6)$$

Here, P_v determines the fused pattern vector associated with texture and skin patch scores. Eq. (6) uses a function to search the criminal database for texture and skin patch scores within a certain range. The five percent is prefixed using a heuristic approach. Finally, the system outputs a matching criminal id C_{id} which is closely associated with the pre-existing pattern.

4. Results

A dual-card Tesla P100 1024-core central processing unit with 128GB RAM and a maximum clock speed of 4.5 TFLOPS is used for the experimental setup. During the installation process, ubuntu 18.04 long-term support was chosen as the operating system This section presents the details of all the experiments we performed to check the performance of our proposed system. Practical investigations are conducted using standard open-source datasets available to the research community. Furthermore, we have evaluated all the standard evaluation metrics as discussed in section 3. A detailed description of the adopted data set and a range of investigations and depth analyses have been

Table 2. Image quality assessment of our proposed system with and without pre-processor

Performance Measures	With Pre-processor			Without Pre-processor		
	High	Low	Average	High	Low	Average
NIQE↓	16.97	6.8	11.97	19.163	10.17	14.1635
BRISQUE↑	16.643	5.734	7.743	13.56	3.563	10.8546
PIQE↓	7.2	3.2	5.321	17.5	13.3	15.314

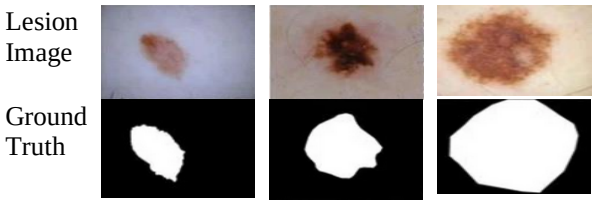


Figure. 2 Dataset samples of some lesions and their corresponding ground truths

conducted in the following subsections to maintain the proposed system's performance.

4.1 Dataset description

The set of images from the human against machine with 10000 training image datasets (HAM10000) [32] has been adopted in our studies. These datasets were divided into four sets containing images of melanoma (Set 1), vascular lesions (Set 2), nevus (Set 3), and other categories (Set 4). To improve training accuracy, we also used augmented samples to increase the dataset size. A complete representation of data samples based on their lesion categories can be found in Table 1. This dataset was collected over twenty years from Vienna and Austria universities. All the images were extracted from DermLiteTM FOTO (3GenTM) camera.

- *Visual presentation of dataset samples*

A selection of images from different mole categories used in our research is presented in this subsection. These samples are just a few of the hundreds that were randomly selected. Sample ground truth images for moles and lesions. As discussed in the earlier section of the research article, the assessment of the segmentation network of the proposed system relies on the lesion image and its corresponding ground truth image.

Hence, for the lesions in the dataset, we generated ground truth images using the ImageJ tool by selecting all the lesion images manually and masking them individually on the utilized HAM10000 lesion dataset. Fig. 2 presents a few handpicked samples of ground truths corresponding to the lesion images.

- *Evaluation of data pre-processing*

This section of the experiment evaluates the effect of pre-processors on the quality of the original image samples. Therefore, we have presented a visual and parametric analysis. Detailed outcomes are presented in the following subsections.

- *Visual analysis of pre-processor influence*

In this section, the quality of selected pre-processed image samples is compared with that of non-pre-processed samples. The visual analysis depicted indicates that the image quality of the pre-processed samples appears better with the enhanced visibility of the mole region when compared to the original samples. This upholds the major purpose of utilizing the pre-processor.

- *Parametric analysis on the influence of pre-processor*

We evaluated the effects of pre-processing using appropriate evaluation parameters. A comparison between performance scores for NIQE, PIQE, and BRISQUE with and without pre-processing is reported in Table 2.

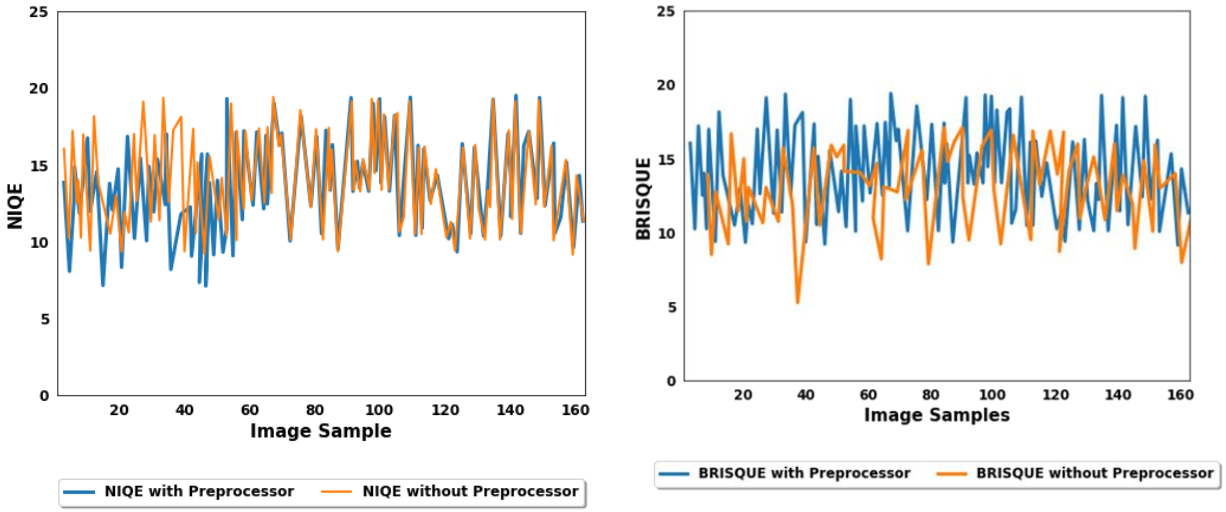
According to Table 2, lower values indicate better quality and less noise in the image. This indicates that our pre-processing has indeed had an impact on the image's quality. Figs. 3 and 4 depict the plot showing the comparison of evaluation scores obtained for each of the selected samples. Samples of each image are shown on the X-axis, and the evaluation score for a no-referential image quality parameter is shown on the Y-axis.

- *Evaluation of mole segmentation network*

To compare our mole segmentation, we adopted standard evaluation metrics. We have utilized the jaccard index and dice coefficient as evaluation metrics. The results of the jaccard Index and dice coefficient are shown in Table 3. Using this technique, heatmaps are drawn across all the vital spots to highlight areas or pixels where classifications were made. The heatmap is created by averaging the gradient vectors of all feature points for all selected pixels.

- *Analysis of the proposed texture encoder*

Our proposed network has also been compared to other state-of-the-art texture encoders to test its capability. From Table 3, it is evident that our mole segmentation network was able to isolate moles at an overall rate of 81%. Graphical plots depicted in Fig. 5 illustrates that after 217 epochs, the segmentation network reaches stability with more than 83%



(a) (b)

Figure. 3 Comparative analysis of (a) NIQE and (b) BRISQUE score for image samples with and without pre-processor

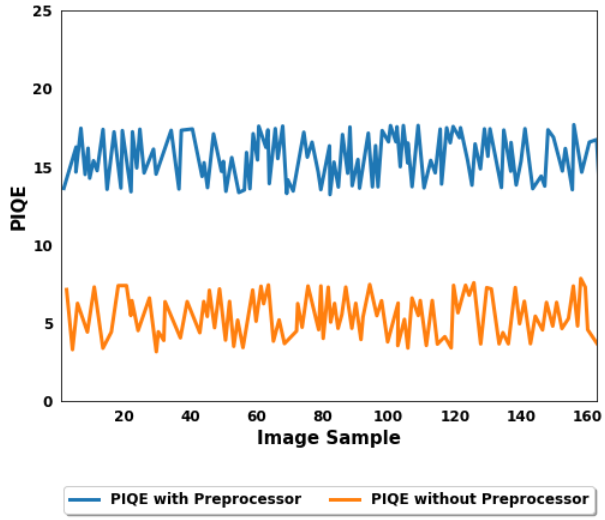


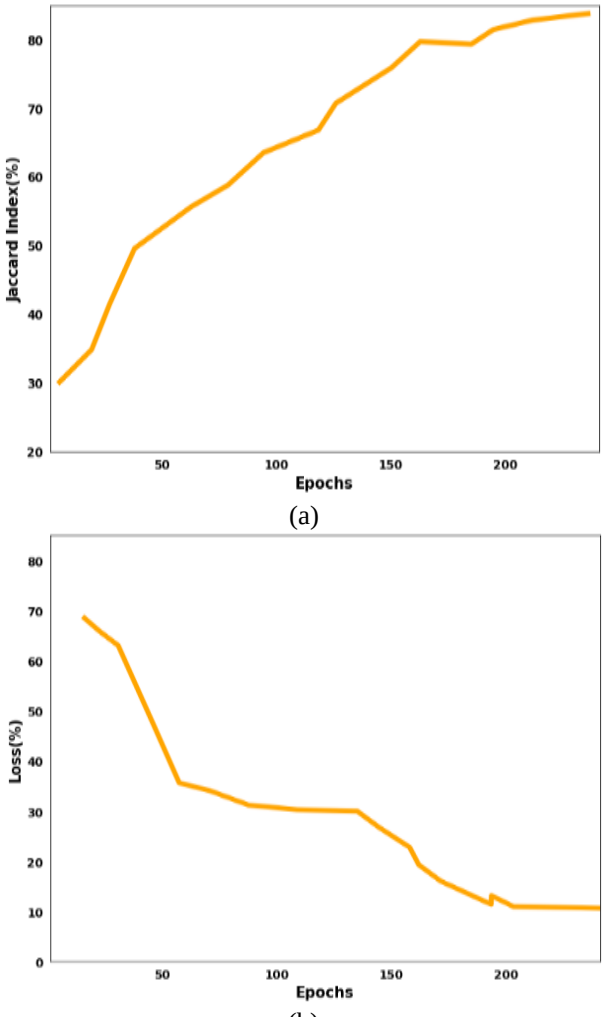
Figure. 4 Comparative analysis of PIQE score for image samples with and without pre-processor

Table 3. Semantic quality assessment of mole extraction technique

Assessment Parameter	Image Testing and Validation Set				Average
	Set 1	Set 2	Set 3	Set 4	
No of Images	2704	4875	715	1706	2500
Jaccard Index↑	81%	82%	80%	81%	81%
Dice Coefficient↑	73%	74%	72%	68%	71.75%

accuracy and less than 10% loss. This proves that 250 training epochs are sufficient to train the network for performing segmentation.

- *Visual representation of contextually aware semantic segmentation using gradient activation maps*



(a)

(b)

Figure. 5 Graphical plots depicting the accuracy and loss values of the mole segmentation network

Semantic segmentation via gradient-weighted class activation mapping was also used to interpret

Table 4. Efficiency comparative analysis of various texture encoders score with the proposed system

Techniques	Texture Set			
	Set1	Set 2	Set 3	Set 4
FV-SIFT	49.3	48.2	47.1	49.2
FV-CNN (VGG.)	63.2	64.3	62.3	64.3
FV-CNN (ResNet)	69.1	70.3	68.4	71.2
Deep TEN	80.6	80.2	84.3	82.3
Proposed DRPN	83.5	85.6	85.4	85.4

Table 5. Comparative analysis of various skin filtering

Evaluation Metric	Technique Name			
	SFAT	Weighted Image Filtering	Spatial Filtering	Proposed System
PSNR↑	37.1	36.2	39.2	36.1
MSE↓	63.2	62.2	62.3	58.2
Pixel Error Rate↓	53.2	52.3	68.4	50.6

Table 6. Accuracy comparison with existing techniques

Technique	Accuracy (%)
CIS [17]	86.71
16-layered CNN model and HDCT [18]	88.39
Proposed	89.21

the contextually aware mole separator's capabilities. From Table 4, we can infer that our proposed encoder does indeed outperform other state of art encoders.

- *Experimental investigation of skin tone analyzer*

In this section, we discuss in detail our system's performance in analyzing skin tones. Our proposed system has been compared in terms of skin filtering, equalization, and overall comparison. We have evaluated the capability using three metrics: PSNR, MSE, and pixel error rate.

According to Table 5, our technique outperforms other techniques in terms of standard evaluation metrics. Table 6 shows the accuracy comparison of proposed method with the exiting technique. The proposed method is compared with the recent approaches such as: CIS and 16-layered CNN model and high dimension contrast transform (HDCT). These existing approaches uses HAM10000 lesion dataset. The HAM10000 lesion dataset helps to predict seven skin lesions, the skin lesions have both even and uneven lesions.

The existing approaches attains skin lesion classification accuracy as 86.71% and 88.39%. The accuracy of proposed method is 89.21% which is higher than the existing approaches.

5. Conclusion

A flexible criminal identification system utilizing multiple soft biometrics with mole and skin pattern identification has been developed and successfully implemented in this research work. Our proposed system has achieved all four objectives defined in our research effort. A correct matching score was determined by analyzing skin tone. This was accomplished via a basic histogram equalization technique coupled with adaptive pixel noise cancellation. Based on the results of this research, the following perspectives have been identified as limitations:

- If there is no exposure to a body part containing a mole, the proposed system cannot be used to identify criminals.
- A body part covered in hair fails to identify mole patterns and skin tone. In exceptional cases, such as scars and marks caused by temporary skin diseases, the system does not apply.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Mr. Rohan Don Salins and Dr. Ananth Prabhu G proposed the idea. Dr. Ananth Prabhu G took the lead in writing the manuscript. Practical investigations and comparative analysis were done by Dr. Json Elroy Martis and Dr. Sannidhan M. S. Mr. Rohan Don Salins supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

References

[1] R. D. Salins, T. S. Ashwin, G. A. Prabhu, M. Basthikodi, and C. K. Mallikarjun, “Person identification from ARM’s hair patterns using CT-twofold siamese network in forensic psychiatric hospitals”, *Complex & Intelligent Systems*, Vol. 8, No. 4, pp. 3185–3197, 2022.

[2] Y. B. Ayzeren, M. Erbilek, and E. Celebi, “Emotional state prediction from online handwriting and signature biometrics”, *IEEE Access*, Vol. 7, pp. 164759–164774, 2019.

[3] J. J. Engelsma, K. Cao, and A. K. Jain, “Learning a fixed-length fingerprint representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, No. 6, pp. 1981–1997, 2021.

[4] I. Chtourou, E. Fendri, and M. Hammami,

"Person re-identification based on gait via part view transformation model under variable covariate conditions", *Journal of Visual Communication and Image Representation*, Vol. 77, p. 103093, 2021.

[5] B. Hartung, D. Rauschning, H. Schwender, and S. R. Timme, "A simple approach to use hand vein patterns as a tool for identification", *Forensic Science International*, Vol. 307, p. 110115, 2020.

[6] Q. Leng, M. Ye, and Q. Tian, "A survey of open-world person re-identification", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 30, No. 4, pp. 1092–1108, 2020.

[7] S. Soleymani, A. Dabouei, H. Kazemi, J. Dawson, and N. M. Nasrabadi, "Multi-level feature abstraction from Convolutional Neural Networks for multimodal biometric identification", In: *Proc. of 2018 24th International Conference on Pattern Recognition (ICPR)*, 2018.

[8] M. Regouid, M. Touahria, M. Benouis, and N. Costen, "Multimodal biometric system for ECG, ear and Iris recognition based on local descriptors", *Multimedia Tools and Applications*, Vol. 78, No. 16, pp. 22509–22535, 2019.

[9] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer", *IEEE Access*, Vol. 7, pp. 38306–38313, 2019.

[10] E. Rezk, M. Eltorki, and W. E. Dakhakhni, "Improving skin color diversity in cancer detection: Deep Learning Approach", *JMIR Dermatology*, Vol. 5, No. 3, 2022.

[11] S. Jain, U. Singhania, B. Tripathy, E. A. Nasr, M. K. Aboudaif, and A. K. Kamrani, "Deep learning-based transfer learning for classification of Skin cancer", *Sensors*, Vol. 21, No. 23, p. 8142, 2021.

[12] B. Hewitt, M. H. Yap, J. Ng, and M. Goyal, "The effect of color constancy algorithms on semantic segmentation of skin lesions", In: *Proc. of Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2019.

[13] M. Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images", *Multimedia Tools and Applications*, Vol. 78, No. 15, pp. 20835-20854.

[14] J. H. Ng, M. Goyal, B. Hewitt, and M. H. Yap, "The effect of color constancy algorithms on semantic segmentation of skin lesions", In: *Proc. of SPIE 10953, Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Vol. 10953, pp. 138-145.

[15] M. A. Anjum, J. Amin, M. Sharif, H. U. Khan, M. S. Malik, and S. Kadry, "Deep semantic segmentation and multi-class skin lesion classification based on Convolutional Neural Network", *IEEE Access*, Vol. 8, pp. 129668–129678, 2020.

[16] M. Nawaz, Z. Mehmood, T. Nazir, R. A. Naqvi, A. Rehman, M. Iqbal, and T. Saba, "Skin cancer detection from dermoscopic images using deep learning and fuzzy *k* -means clustering", *Microscopy Research and Technique*, Vol. 85, No. 1, pp. 339–351, 2021.

[17] D. Popescu, M. E. Khatib, and L. Ichim, "Skin Lesion Classification Using Collective Intelligence of Multiple Neural Networks", *Sensors*, Vol. 22, No. 12, p. 4399, 2022.

[18] M. A. Khan, K. Muhammad, M. Sharif, T. Akram, and V. H. D. Albuquerque, "Multi-class skin lesion detection and classification via teledermatology", *IEEE Journal of Biomedical and Health Informatics*, Vol. 25, No. 12, pp. 4267-4275, 2021.

[19] M. Nawaz, M. Masood, A. Javed, J. Iqbal, T. Nazir, A. Mehmood, and R. Ashraf, "Melanoma localization and classification through faster region-based convolutional neural network and SVMm", *Multimedia Tools and Applications*, Vol. 80, No. 19, pp. 28953-28974, 2021.

[20] V. Pugazhenthi, S. K. Naik, A. D. Joshi, S. S. Manerkar, V. U. Nagvekar, K. P. Naik, C. G. Palekar, and K. Sagar, "Skin disease detection and classification", *International Journal of Advanced Engineering Research and Science*, Vol. 6, No. 5, pp. 396-400, 2019.

[21] M. Kim and M. H. Song, "High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN", *Applied Sciences*, Vol. 13, No. 2, p. 989, 2023.

[22] P. Bose and S. K. Bandyopadhyay, "Facial spots detection using convolution neural network", *Asian Journal of Research in Computer Science*, pp. 71-83, 2020.

[23] K. K. V. Toh, H. Ibrahim, and M. N. Mahyuddin, "Salt-and-pepper noise detection and reduction using fuzzy switching median filter", *IEEE Transactions on Consumer Electronics*, Vol. 54, No. 4, pp. 1956-1961.

[24] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer", *IEEE Signal Processing Letters*, Vol. 20, No. 3, pp. 209-212, 2013.

[25] L. S. Chow and H. Rajagopal, "Modified-BRISQUE as no reference image quality

assessment for structural MR images", *Magnetic Resonance Imaging*, Vol. 43, pp. 74-87, 2017.
[26] A. Pandey, D. Yadav, A. Sharma, D. Sonker, C. Patel, C. Bal, and R. Kumar. "Evaluation of Perception based Image Quality Evaluator (PIQE) no-reference image quality score for 99mTc-MDP bone scan images", *Journal of Nuclear Medicine*, Vol. 61, 2020.
[27] N. Yamanakkanavar and B. Lee, "A novel M-SegNet with global attention CNN architecture for automatic segmentation of brain MRI", *Computers in Biology and Medicine,* Vol. 136, p. 104761, 2021.
[28] S. Mao, D. Rajan, and L. T. Chia, "Deep residual pooling network for texture recognition", *Pattern Recognition,* Vol. 112, p. 107817, 2021.
[29] Y. George, M. Aldeen, and R. Garnavi, "Skin hair removal for 2D psoriasis images", In: *Proc. of 2015 International Conference on Digital Image Compu-ting: Techniques and Applications (DICTA)*, IEEE, pp. 1-8, 2015.
[30] S. Patel and M. Goswami, "Comparative analysis of Histo-gram Equalization techniques", In: *Proc. of 2014 International Conference on Contemporary Computing and Informatics (IC3I), IEEE*, pp. 167-168, 2014.
[31] M. S. Sannidhan, G. A. Prabhu, K. M. Chaitra, and J. R. Mohanty, "Performance enhancement of generative adversarial network for photograph–sketch identification", *Soft Computing*, Vol. 27, No. 1, pp. 435-452, 2023.
[32] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions", *Scientific Data*, Vol. 5, No. 1, pp. 1-9, 2018.