Baghdad Science Journal

Volume 23 | Issue 3 Article 30

3-25-2026

Dual Focus Attention Block UNet (DFAB-UNet) for Robust Skin Lesion Segmentation

Fallah H. Najjar
Department of Emergent Computing, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia AND Department of Computer Networks and Software Techniques, Technical Institute of Najaf, Al-Furat Al-Awsat Technical University, 54001 Najaf, Iraq,
fallahnajjar@atu.edu.iq

Farhan Mohamed
Department of Emergent Computing, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia AND Media and Game Innovation Center of Excellence, Institute of Human Centered Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia,
farhan@utm.my

Mohd Shafry Mohd Rahim
Department of Emergent Computing, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia, shafry@utm.my

Diana Sofia de Oliveira Bernardo
Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal,
diana.bernardo@research.fchampalimaud.org

Luís Filipe Almeida Bernardo
GeoBioTec, Department of Civil Engineering and Architecture, University of Beira Interior, 6201-001 Covilhã, Portugal, lfb@ubi.pt

See next page for additional authors

How to Cite this Article

Najjar, Fallah H.; Mohamed, Farhan; Rahim, Mohd Shafry Mohd; Bernardo, Diana Sofia de Oliveira; Bernardo, Luís Filipe Almeida; and Chan, Vei Siang (2026) "Dual Focus Attention Block UNet (DFAB-UNet) for Robust Skin Lesion Segmentation," *Baghdad Science Journal*: Vol. 23: Iss. 3, Article 30.
DOI: https://doi.org/10.21123/2411-7986.5233

Dual Focus Attention Block UNet (DFAB-UNet) for Robust Skin Lesion Segmentation

Authors

Fallah H. Najjar, Farhan Mohamed, Mohd Shafry Mohd Rahim, Diana Sofia de Oliveira Bernardo, Luís Filipe Almeida Bernardo, and Vei Siang Chan

BAGHDAD SCIENCE JOURNAL 2026;23(3):1100–1115

Scan the QR to view the full-text article on the journal website

RESEARCH ARTICLE

Dual Focus Attention Block UNet (DFAB-UNet) for Robust Skin Lesion Segmentation

Fallah H. Najjar [1,2,*], **Farhan Mohamed** [1,3,*], **Mohd Shafry Mohd Rahim** [1], **Diana Sofia de Oliveira Bernardo** [4], **Luís Filipe Almeida Bernardo** [5], **Vei Siang Chan** [1,3]

[1] Department of Emergent Computing, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[2] Department of Computer Networks and Software Techniques, Technical Institute of Najaf, Al-Furat Al-Awsat Technical University, 54001 Najaf, Iraq
[3] Media and Game Innovation Center of Excellence, Institute of Human Centered Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[4] Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal
[5] GeoBioTec, Department of Civil Engineering and Architecture, University of Beira Interior, 6201-001 Covilhã, Portugal

ABSTRACT

Skin cancer is the most fatal type of cancer worldwide. Prognosis is generally much better with early detection and effective treatment relies on an accurate diagnosis of skin lesions. Despite progress in deep learning, there is still a challenge to provide accurate segmentation of dermoscopic skin lesions due to image variations stemming from different illuminative input parameters, different resolutions/image sizes, image artifacts and varying skins. These differences hinder state-of-the-art models from learning pixel-precise lesion boundaries or more contextual features, consequently impairing their generalization and clinical applicability. We introduce a new architecture called Dual Focus Attention Block UNet (DFAB-UNet) with intrinsic capability of local and global feature extraction tasks. We trained and tested the proposed method on two commonly used sets of dermoscopy images, PH2 and HAM10000 with respect to six different parameters, namely, Accuracy (Acc), Precision (Pre), Sensitivity (Sen), Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and specificity (Spe) for comparative analysis Results. DFAB-UNET achieved an Acc of 97.24%, a DSC of 94.30%, an IoU of 89.21% and a Spe of 98.95% on the PH2 dataset. On the HAM10000, we achieved an Acc of 96.20%, DSC of 92.79%, IoU of 86.55% and Spe of 97.60%. These results show that the model generalizes across datasets by introducing boundary precision represented in the loss together with contextual lesion information. The DFAB-UNet model achieves high performance on segmentation, indicating the potential use and application of dermal disease research in clinic practice.

Keywords: DFAB-UNet, Medical image segmentation, Skin cancer, Skin lesion segmentation, UNet

Introduction

Skin cancer is one of the most common types of cancer worldwide, with melanoma being its most severe form.[1,2] Early detection and accurate diagnosis of skin lesions are critical steps in cancer treatment, so patients tend to have better outcomes when the disease is detected early.[3,4] Dermoscopy is a well-established tool in dermatology that enables non-invasive examination of cutaneous lesions. However, manual analysis of dermoscopic images is time-consuming and subject to inter-observer variability,[5] highlighting the need for automated and robust systems for skin lesion segmentation and classification.

Received 25 April 2025; revised 25 September 2025; accepted 27 September 2025.
Available online 25 March 2026

* Corresponding author.
E-mail addresses: fallahnajjar@atu.edu.iq (Fallah H. Najjar), farhan@utm.my (F. Mohamed), shafry@utm.my (M. S. M. Rahim), diana.bernardo@research.fchampalimaud.org (D. S. de Oliveira Bernardo), lfb@ubi.pt (L. F. A. Bernardo), vschan2@live.utm.my (V. S. Chan).

https://doi.org/10.21123/2411-7986.5233

Recent advances in deep learning techniques, particularly Convolutional Neural Networks (CNNs), have achieved unprecedented success in various medical image analysis tasks.[6] Among these, UNet architectures are widely used for image segmentation tasks[7] due to their encoder-decoder symmetrical scheme with skip connections that combine detailed and contextual information.[8] Despite their efficiency, UNet performance can be further enhanced with advanced attention mechanisms. Attention mechanisms are widely used in deep learning to enable models to focus on the most relevant areas of input data.[9] In image segmentation, attention mechanisms improve the model's ability to distinguish critical features from high-dimensional data.[10]

Furthermore, a large annotated dataset is essential for training deep learning models.[11] Public and private datasets, such as International Skin Imaging Collaboration (ISIC), ISIC2016,[12] ISIC2017,[13] ISIC2018,[14–16] Pedro Hispano Hospital (PH2),[17] and Human Against Machine with 10000 images (HAM10000),[18] have been instrumental in establishing benchmarking frameworks for segmentation and classification models in dermatology. Nevertheless, the data is still highly imbalanced, and the malignant lesions are outnumbered by the benign ones.[19] This disparity can result in a biased model that is good at segmenting common types of lesions, but its performance may suffer on less common but more dangerous variants.[20]

Another important problem in this field is image acquisition variability. Dermoscopic images could vary greatly in lighting and resolution, skin, and artifacts such as hair and bubbles.[21] These factors may compromise the generalization power of the model, and the need for a robust pre-processing pipeline and data augmentation mechanism during model training becomes inevitable. Future research is working on various data types, including patient metadata, clinical images, and histopathology combined with dermoscopic images to enhance diagnostic precision.[22] Furthermore, the explainability and interpretability of AI-based models are becoming increasingly popular, given that clinicians must comprehend and trust the rationale for automated decisions.[23] It is also possible to use visualizations like Grad-CAM or activation maps to see what the model is "attending" to understand its decisions.

Although U-Net-based architectures have achieved great success in skin lesion segmentation, there are still some limitations that hinder their performance in clinical applications. It may be problematic that traditional U-Net models are incapable of extracting multi-scale features from dermoscopic images and delivering them to a subsequent module, since these features are essential for representing the various sizes and shapes of skin lesions. In extreme cases where lesions are very small or very large, models lacking the ability to capture scale differences result in misclassification or the omission of regions. Moreover, U-Net employs multiple sequential convolutional settings that may work against an effective representation of complex texture patterns due to small receptive fields, resulting in poor contextual understanding of local and global regions.[24]

To alleviate this issue, approaches based on multi-scale fusion operations or dilated convolutions were proposed; nevertheless, they are more expensive in terms of computation and introduce redundant feature extraction. Moreover, even with the addition of attention mechanisms and squeeze-and-excitation (SE) modules to enhance channel-wise feature representation, both approaches still depend on aggregated base features obtained from the original encoder. However, this dependency becomes an issue in images when visual noise is present as in irregular illumination, noise or occlusion from hair and skin texture. Therefore, these models might be sensitive to low-contrast or indistinct lesion margins, resulting in degraded segmentation performance.[25]

The most prominent is the use of multi-resolution feature maps both without losing spatial resolution and semantic consistency. Traditional network implementations usually do not effectively combine low-level spatial information with high-level semantic meaning, which leads to inaccurate boundaries and fragmentary lesion masks. These challenges indicate that, despite significant progress, further research is needed to develop architectures that are accurate, computationally efficient, and robust to the variations commonly found in real-world dermoscopic images.[26]

In this work, we propose a solution based on the new Dual-Focus Attention Block (DFAB) to tackle these difficulties and improve the typical UNet model for skin lesion segmentation. DFAB combines both channel and spatial attention to highlight the discriminative channels and their informative spatial regions in every feature channel. The key contributions are summarized as follows: (i) A new attention mechanism is proposed to enable the UNet architecture to leverage channel-wise and spatial attention. This dual-focus input and output have compatibility, allowing the network to better extract useful features and spatial information, thereby improving segmentation accuracy. (ii) Thorough experimentation is conducted on two public dermoscopic image datasets to validate the proposed DFAB-UNet model. (iii) The model's effectiveness and reliability are assessed through a detailed analysis of performance

metrics, including precision, recall, F1 score, and other essential indicators.

Related works

Advancements in deep learning have significantly improved skin lesion segmentation by addressing specific challenges and enhancing accuracy and robustness. For example, Tomar et al.[27] proposed a new deep-learning framework, FANet, to enhance the segmentation of biomedical images. By reusing features from each training epoch, FANet progressively refined the prediction masks, yielding more accurate segmentation outputs. It included a feedback mechanism that reused the previous segmentation mask for the same image and combined it with the feature maps of the current epoch, allowing the network to pay intricate attention to selected features. The FANet was trained on the HAM10000 dataset, achieving an F1 score of 0.8731 and a sensitivity of 0.8650, outperforming BCDU-Net by 2.21% in F1 and 8.00% in sensitivity. Gupta et al.[28] published a combined method for skin lesion segmentation integrating multi-level closing-based hair removal with an ensemble of deep networks, including U-Net and Res-Net. This model obtained a Dice Similarity Coefficient (DSC) of 0.9555, an Intersection over Union (IoU) of 0.8545, and an accuracy of 95.87 % on the HAM10000 and ISIC datasets. Zhang et al.[29] developed SPDD-PatchGAN for accurate dermoscopic skin lesion segmentation, which implemented superpixel-guided discriminators and dual-branch patch-based discriminators to enhance local segmentation details. The SPDD-PatchGAN achieved an IoU of 89.92% and an accuracy of 95.94% on the HAM10000 dataset, outperforming other state-of-the-art methods using CNN or GAN techniques. Li et al.[30] proposed SG-MIAN, a skin lesion segmentation algorithm that employed image-level labels reproducibly instead of pixel-level labels and reduced the significant spending on labor-intensive annotation. It employed multiple spatial perceptrons (MSP) to localize and activate lesion areas robustly. This model improved the previous weakly supervised segmentation techniques by attaining an IoU of 74.33% and a DSC of 83.10% on the HAM10000 dataset. Zhi et al.[31] introduced TEDMAE, a Teacher-Student architecture based on Masked Autoencoders (MAE) for more accurate skin lesion segmentation. By combining self-distillation and optimization techniques, TEDMAE can learn local and global features from dermoscopy images. On the PH2 dataset, this model achieved a DSC of 91.2% and an IoU of 84.5%, surpassing several existing methods. Kaur et al.[32] tackled the problem of skin lesion segmentation in dermoscopic images with an advanced U-Net-based architecture incorporating residual learning, attention gates, and a modified decoder - ARU-Net-MD. Experimental studies on various datasets showed that ARU-Net-MD achieved accuracy rates up to 0.97, indicating its effectiveness in automatic skin lesion segmentation. Lie et al.[33] introduced USL-Net, a new method of determining dermoscopic image segmentation without manual labeling. USL-Net utilizes contrastive learning for image feature extraction and Class Activation Maps (CAMs) to generate pseudo-labels for lesion regions. On the ISIC-2017 dataset, USL-Net outperformed state-of-the-art unsupervised methods, improving mean accuracy by 1.7%, DSC by 6.6%, IoU by 4.0%, and sensitivity by 10.6%.

Zhao et al.[34] proposed LUT-SLS, a lightweight Transformer-based U-Net variant, designed to improve skin lesion segmentation by integrating the Convolution Module, the Pooling-based Lightweight Transformer Structure (PLTS), and the Multi-Scale Bottleneck Network (MSBN). The Convolution Module uses depth-separable convolution and SE attention to better extract features and minimize computation. The PLTS Module uses average pooling instead of self-attention, greatly reducing computational costs while retaining segmentation capability. On the other hand, the MSBN Module utilizes multi-level spatial pyramid pooling to capture lesion features at multiple scales. Optimally performing using a Binary Cross-Entropy and Dice loss function, trained the model with Adam optimizer and CosineAnnealingLR scheduler.

Garbaz et al.[35] proposed DMFC-UFormer, a depth-wise multi-scale factorized convolution (DMFC) transformer-based U-Net, specifically adapted for medical images, such as skin lesions, by combining the strengths of CNNs and Transformers for local and global features in fewer parameters. The architecture consisted of four main parts: DMFC-Transformer, which explored the multi-scale factorized feature extraction (MFFE) and DMFC to capture feature diversity; Enhanced Contextual Feature Integration (ECFI) to model contextual correlation; Spatial-Channel Partitioned Feature Attention (SCPFA) to increase the receptive field range and strengthen the feature information; Attention-Based Feature Stabilization (AFS) to further optimize the skip connection and keep the necessary semantics. DMFC-UFormer was validated on two skin lesion segmentation datasets, ISIC 2017 and PH2. The model reached state-of-the-art results with a Dice Coefficient of 82.39% on ISIC 2017 and 89.99% on PH2. This promising model that balances speed and accuracy is suitable for real-time skin lesion analysis, particularly in the context of

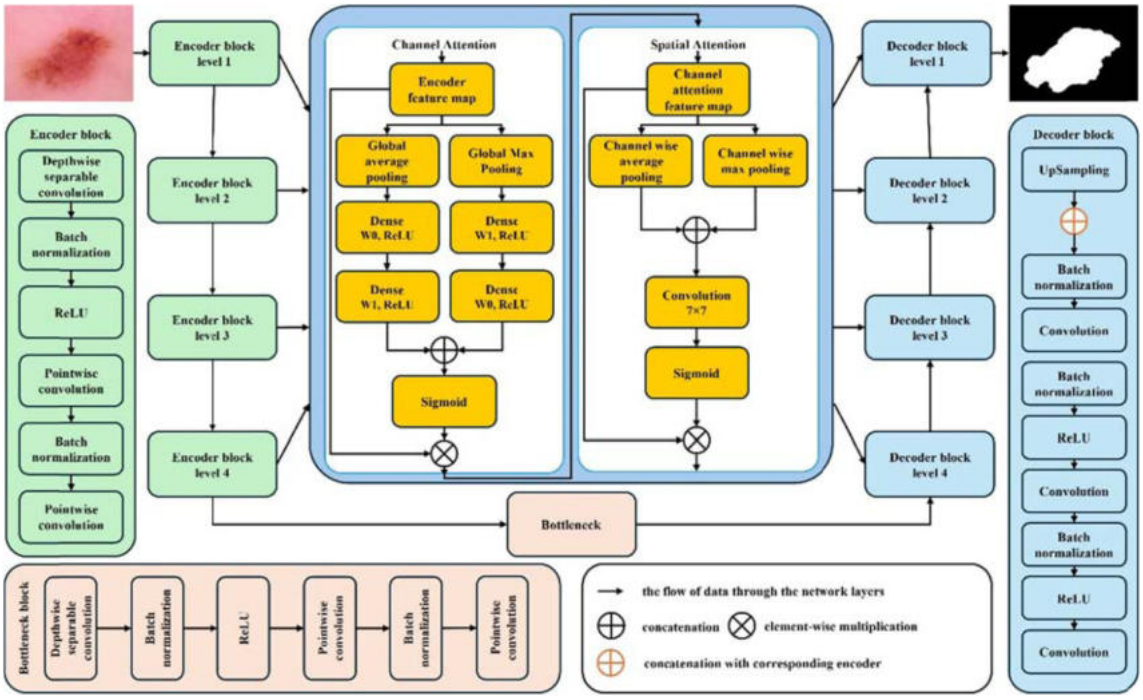

Fig. 1. Overview of the proposed DFAB-UNet architecture.

mobile and IoT-based healthcare. Finally, our previous study[36] presented DeepSkinSeg, a lightweight U-Net-inspired model specifically designed for precise skin lesion segmentation in dermoscopic images. The proposed model solves two common problems, hard-to-identify boundaries and low efficiency, of medical image segmentation. The authors suggested a simplified version of the conventional U-Net, using fewer layers and retaining important skip connections to compromise the performance of the segmentations and the computational needs. State-of-the-art results have been achieved by DeepSkinSeg on the PH2 and HAM10000 datasets, using Dice coefficient 96.36% and Intersection over Union (IoU) 92.98%, respectively, which several recent models do not better. The authors' work highlighted the feasibility of practical, efficient, real-time applications for AI in clinical dermatology while preserving the accuracy of segmentation. Therefore, it is a strong contender for integration into automated diagnostic platforms.

Methodology

In this paper, a DFAB is proposed and incorporated into a U-Net, resulting in the DFAB-UNet architecture to improve skin lesion segmentation. It combines channel and spatial attention mechanisms, which significantly benefits the model by focusing on more informative characteristics. This section elaborates on the encoder and decoder components that the DFAB has improved throughout the network. See Fig. 1.

Pre-processing

This study only performed resizing on the initial dermoscopic images. All images were resized to 224 × 224 pixels for standardization of input dimensions in the network. This resizing was exclusively performed to optimize batch processing while being adapted to the model architecture, without post-processing operations like the removal of artifacts, contrast adjustment, or normalization. This design choice demonstrates the robustness of the model to raw dermoscopic images and the reduction of pre-processing steps.

Encoder blocks

The input image (I) is forwarded through several convolutional blocks to generate multi-scale feature maps. The convolutional blocks consist of a convolutional layer followed by Batch Normalization (BN) and Rectified Linear Unit (ReLU) activation functions. More concretely, this input image flows through four convolutional blocks that result in intermediate feature maps E_1, E_2, E_3, and E_4. Then, E_4 is further processed by the Bottleneck block E_{NB}. Fig. 2 displays the architecture of the encoder blocks.

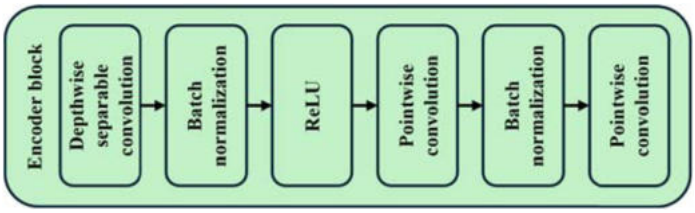

Fig. 2. Architecture of the encoder blocks.

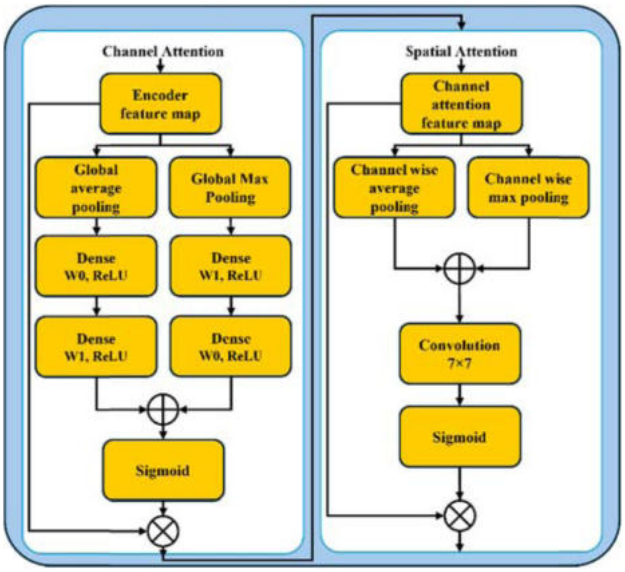

Fig. 3. Architecture of the proposed DFAB.

The proposed DFAB

Channel Attention (CA) is applied to each feature map produced by the encoder. The CA mechanism enables the feature maps to focus more attention on the informative channels. This process is done by channel-pooling the statistics using average and max-pooling operations. The query, key, and value come from the same data, and SoftMax is applied to create attention weights between the input and the output. These statistics are then passed through two fully connected layers that use ReLU and sigmoid activations to compute the attention weights. The resulting attention map is applied to the feature map through element-wise multiplication, producing the channel-attended feature maps $E_{CA,\,1}$, $E_{CA,\,2}$, $E_{CA,\,3}$, $E_{CA,\,4}$, and $E_{CA,\,BN}$. Fig. 3 illustrates the architecture of the proposed DFAB.

Eqs. (1) and (2) are used to compute the global average pooling and global max pooling for each channel c, respectively.

$$avg_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{ijc} \tag{1}$$

$$max_c = \max_{1 \le i \le H,\ 1 \le j \le W} x_{ijc} \tag{2}$$

Where x_{ijc} represents the input feature map, H and W denote the height and width, respectively, and c represents the channel index.

Pooled features are passed through the shared fully connected layers described in Eqs. (3) and (4).

$$F_{avg} = W_1 \gamma \left(W_0\ avg = b_0\right) + b_1 \tag{3}$$

$$F_{max} = \ W_1 \gamma \left(W_0\ max = b_0\right) + b_1 \tag{4}$$

Where W_1 and W_0 represent the weights, b_0 and b_1 indicate the biases, respectively, and γ denotes the ReLU activation function.

The outputs from both global average pooling and global max pooling are combined, and the sigmoid function is applied using Eq. (5).

$$M_c = \sigma(F_{avg} + F_{\max}) \tag{5}$$

Where σ represents the sigmoid activation function.

Finally, the feature map was refined by CA using Eq. (6).

$$X_{CA} = M_C \odot X \tag{6}$$

Where $\odot$ represents element-wise multiplication, X signifies any of the intermediate feature maps of the E_1, E_2, E_3, E_4, and E_{NB} in the encoder or their respective channel-attended versions $E_{CA,\,1}$, $E_{CA,\,2}$, $E_{CA,\,3}$, $E_{CA,\,4}$, and $E_{CA,\,BN}$.

After applying CA, Spatial Attention (SA) is employed on the channel-attended feature maps, as presented in Fig. 3. The SA mechanism highlights the most critical spatial regions in the feature maps. This process is done by computing the channel-wise average and max-pooling the spatial statistics. Then, the concatenated pooled features are fed into a convolutional layer with a sigmoid activation to produce one attention map. The full channel-attended feature map allows spatially attended feature maps

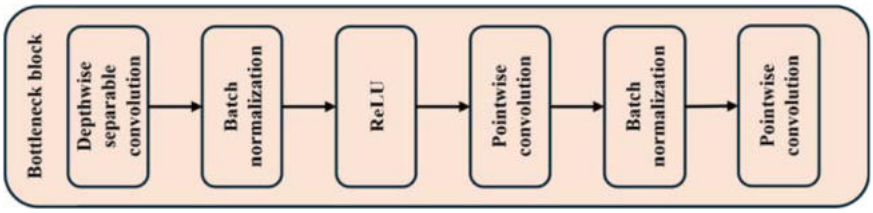

Fig. 4. Architecture of the bottleneck block.

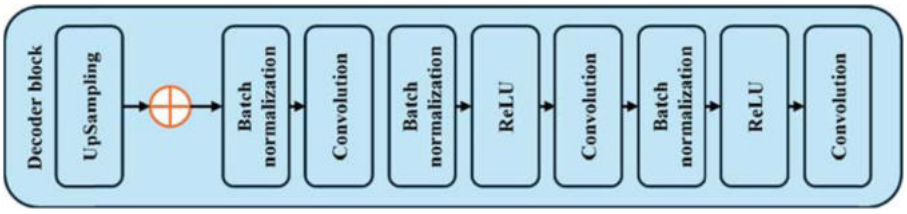

Fig. 5. Architecture of the decoder blocks ⊕ representing connection with the corresponding encoder.

through element-wise multiplication $E_{SA,\ 1}$, $E_{SA,\ 2}$, $E_{SA,\ 3}$, $E_{SA,\ 4}$, and $E_{SA,\ BN}$.

Eqs. (7) and (8) are used to compute the channel-wise average pooling and channel-wise max pooling across each channel C, respectively.

$$avg_{SA} = \frac{1}{C}\sum_{c=1}^{C} x_{ijc} \tag{7}$$

$$max_{SA} = \max_{1\le c\le C} x_{ijc} \tag{8}$$

Where x_{ijc} represents the input feature map after the CA.

Eq. (9) is used to compute the spatial attention mask.

$$M_s = \sigma\left(Conv\left(Concat\left(avg_{SA}, max_{SA}\right), kernal\,(7\times 7)\right)\right) \tag{9}$$

Where σ represents the sigmoid activation function, and *Conv* is the convolutional layer with a 7 × 7 kernel. Finally, the feature map was refined by SA using Eq. (10).

$$X_{SA} = M_S \odot \sigma \tag{10}$$

Where X_{SA} represents the spatial-attended feature map after applying SA to the X_{CA}.

Bottleneck block

The bottleneck block serves as a bridge connecting the encoder with the decoder to capture high-level semantics. It uses multiple layers that capture complex patterns that are crucial for precise lesion delineation, as shown in Fig. 4.

Decoder blocks

The decoder processes the upsampled feature maps, which are concatenated with the corresponding feature maps from the encoder that the DFAB has refined. The process begins with the upsampled feature map, $E_{SA,\ BN}$, concatenated with $E_{SA,\ 4}$ to form D_4. This process is iterated over every higher level, concatenating and upsampling the feature maps D_4, D_3, and D_2 with $E_{SA,\ 3}$, $E_{SA,\ 2}$, $E_{SA,\ 1}$, respectively, producing intermediate feature maps D_3, D_2, and D_1. Fig. 5 shows the architecture of the decoder blocks.

Furthermore, each of the concatenated feature maps is processed by a decoder convolutional block to polish the upsampled features. These blocks consist of convolutional layers followed by BN and ReLU activations, yielding refined feature maps $D_4^{'}$, $D_3^{'}$, $D_2^{'}$, and $D_1^{'}$. The output is the final segmentation map, generated through the 1×1 convolution with sigmoid activation after the last decoder block $D_1^{'}$, resulting in the segmentation map BW_{out}.

The integration of CA enables the model to learn to weigh different channels for each input image, thereby improving feature representation and segmentation performance. This using SA allows the model to enrich the representation of spatial locations, referred as spatial importance, in order to achieve accurate localization of lesions, therefore, it can get improved segmentation performance.

Dataset collection

This study utilized dermoscopic images of skin lesions and their corresponding masks from the PH2[17] and HAM10000[18] datasets for the proposed DFAB-UNet experiment. The PH2 database contains 200 dermoscopic images, each with a resolution of

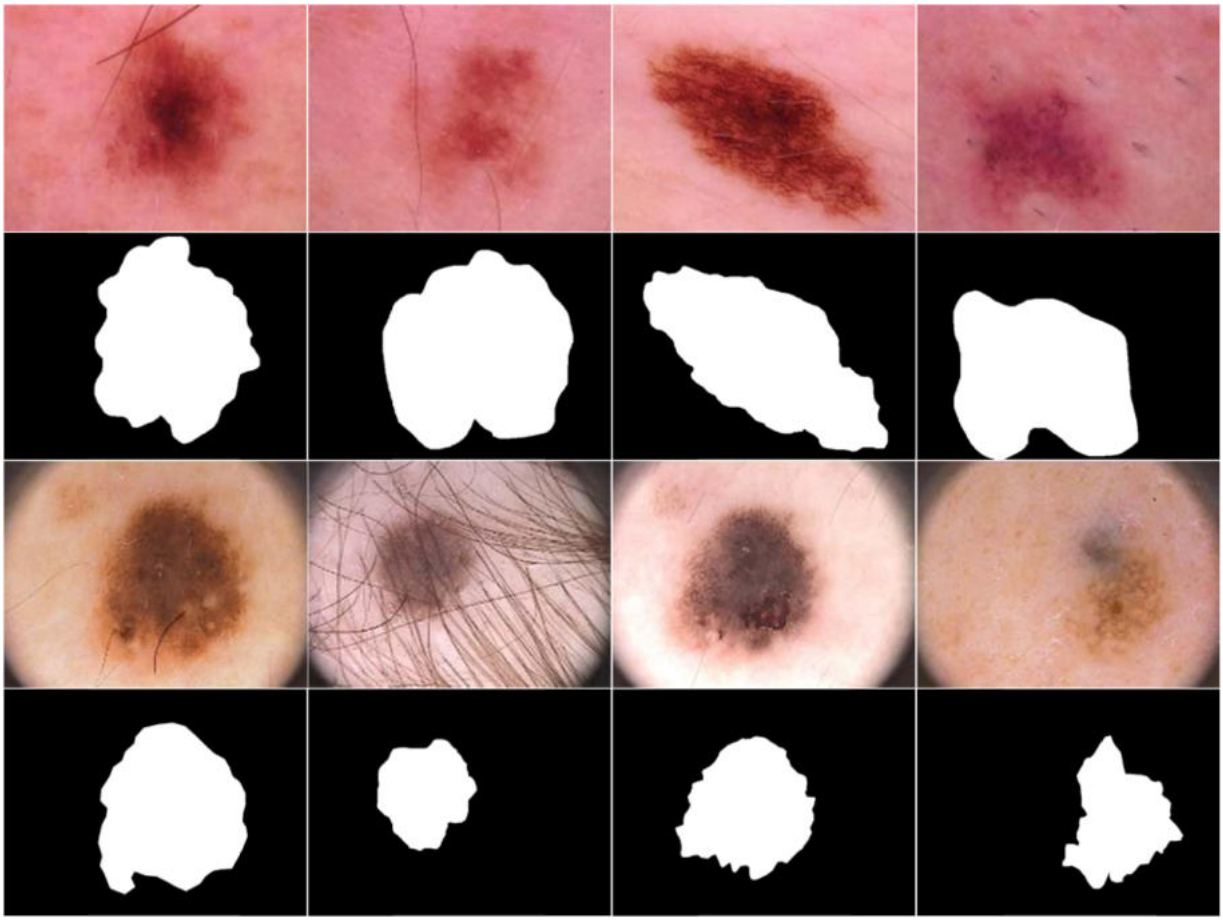

Fig. 6. Sample images: (1st and 2nd rows) HAM10000 dataset[18] and (3rd and 4th rows) PH2 dataset.[17]

768 × 560 pixels, although the HAM10000 dataset comprises 10015 dermoscopic images of flattened skin lesions with 600 × 450 pixels. Fig. 6 illustrates representative image samples and their corresponding ground truth from both datasets. Each dataset was randomly split into 80% for training, 10% for validation, and 10% for testing. This approach ensures sufficient data for training and validation, while leaving a subset unseen for final evaluation. All images and masks were resized to a uniform resolution of 224 × 224 pixels.

Results and discussion

This section presents the results which evaluate how well the proposed DFAB-UNet model performs in segmenting skin lesions. We have tested Adam and SGD optimizers with different parameter settings. The Adam optimizer with 0.0001 initial learning rate outperformed all others. The model was trained for 50 epochs, using a batch size of 16 and binary cross-entropy as the loss function.

Fig. 7 illustrates the training and validation performance of the proposed DFAB-UNet approach on the PH2 dataset across key metrics, including Loss, Acc, DSC, and IoU. The loss curves for both the training and validation show a gradual decrease, reaching optimal values around epoch 20. The training accuracy exceeds 95%, with validation accuracy following closely. Moreover, the DSC and IoU metrics further highlight the efficiency of the model, as both metrics exhibit steady improvement during the initial epochs, before stabilizing around 0.95 and 0.90, respectively, for the training and validation datasets. The strong alignment in performance between training and validation across several metrics suggests that DFAB-UNet can provide segmentations that are both precise and accurate with respect to the ground truth annotations.

Fig. 8, displays the segmentation confusion matrix for both datasets, highlighting the model's ability to distinguish between lesion and background with minimal misclassification.

Based on the PH2 confusion matrix as shown in Fig. 8, we can observe that the segmentation model

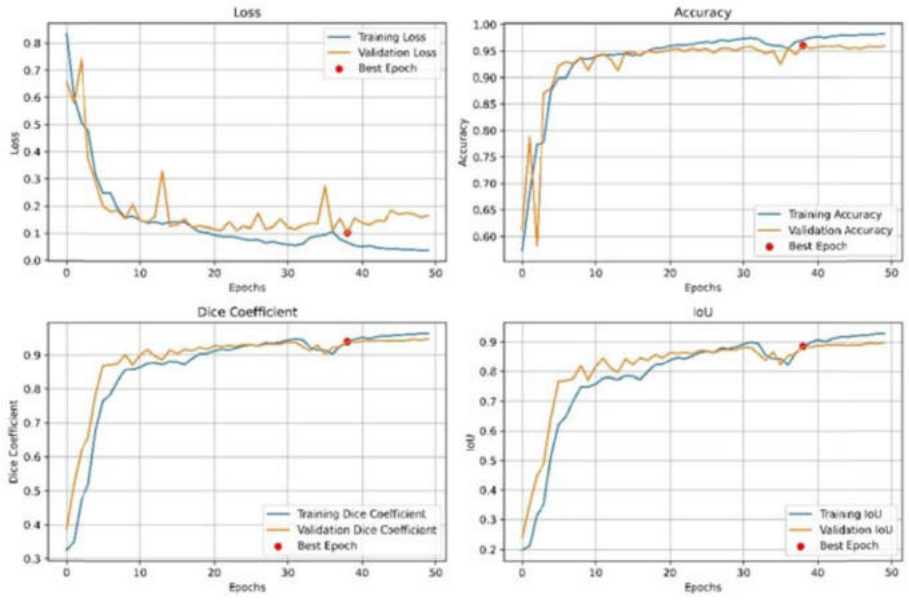

Fig. 7. PH2 training and validation performance: loss, Acc, DSC, and IoU.

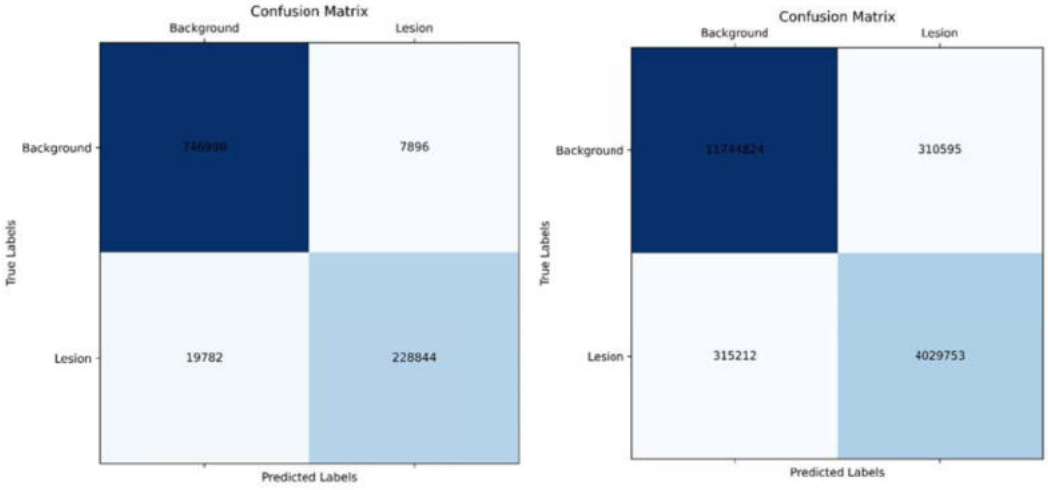

Fig. 8. Pixel-level confusion matrix: left: PH2 (20 images) and right: HAM10000 (1004 images).

achieves a high overall performance: it classifies background pixels and lesion pixels accurately. The model achieves very high precision, referring to how rarely it assigns background pixels as lesions, and good recall, indicating that it generally identifies most lesion pixels.

In addition, the Dice coefficient of overlap between the predicted lesion regions and the ground truth is high, indicating good segmentation performance. In contrast, the result demonstrates the existence of false negatives, where pixels of the lesion in question were classified as background, which indicates that some parts of the lesion are missed altogether, which is an undesirable behavior in a medical use case.

Overall, the confusion matrix shows that the segmentation model performs well, with a high number of lesion pixels and background pixels, which means that the model could discriminate well between the lesion and the non-lesion regions. Nevertheless, there are false positives (background pixels misclassified as lesions) and false negatives (lesion pixels misclassified as background), areas where the model could still improve. To improve this, more work could be devoted to boosting the model's ability to identify smaller or less prominent regions of the lesion. These could involve more complex data augmentation approaches, possible modifications to the loss function to emphasize the identification of lesions or

Table 1. Segmentation results comparison of different methods on the PH2 dataset.

Author(s) and Year	Acc	Pre	Sen	DSC	IoU	Spe
Li et al., 2024[30]	-	-	-	85.36	76.30	-
Zhi et al., 2024[31]	-	-	-	91.20	84.50	-
Kaur and Kaur 2024[32]	96.56	-	91.61	92.65	88.57	97.86
Li et al., 2024[33]	92.40	-	93.60	88.90	80.10	93.10
Zhao et al., 2024[34]	96.24	91.85	95.57	93.54	87.92	-
Garbaz et al., 2025[35]	-	-	-	89.99	82.06	-
DFAB-UNet	97.24	96.66	92.04	94.30	89.21	98.95

Table 2. Segmentation results comparison of different methods on the HAM10000 dataset.

Author(s) and Year	Acc	Pre	Sen	DSC	IoU	Spe
Tomar et al., 2022[27]	92.35	96.11	86.50	-	80.23	-
Zhang et al., 2024[29]	95.94	90.41	92.31		89.92	97.51
Li et al., 2024[30]	-	-	-	83.10	74.33	-
Zhao et al., 2024[34]	95.63	90.79	87.99	88.94	80.59	-
DFAB-UNet	96.20	93.28	92.31	92.79	86.55	97.60

architectural changes, such as extracting features at multiple scales or making use of attention decision-making mechanisms. Using ground truth data, the model can learn to make continual fine-tuning like these to invest more attention to data points, leading to a more effective segmentation, especially on a dynamic application form. In our proposed system, six different parameters, Accuracy (Acc), Precision (Pre), Sensitivity (Sen), DSC, IoU, and Specificity (Spe), are considered for comparative analysis and are calculated using Eqs. (11) to (16)[37,38].

$$\text{Accuracy} = \text{Acc} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{11}$$

$$\text{Precision} = \text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

$$\text{Sensitivity} = \text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

$$\text{DSC} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \tag{14}$$

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{15}$$

$$\text{Specificity} = \text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{16}$$

Where TP is true positive, FP represents false positive, TN denotes true negative, and FN indicates false negative. Nevertheless, Table 1 displays the segmentation performance of the proposed DFAB-UNet regarding the PH2 dataset.

The stable performance of DFAB-UNet on vital statistics confirms its effectiveness and reliability, making it an ideal candidate for skin lesion segmentation. DFAB-UNet achieved the highest accuracy on the PH2 dataset, 97.24% on a pixel-level task, thus improving upon the state-of-the-art. It recorded only 3.34 % false positive errors, demonstrating the high precision of its lesion detection. Moreover, DFAB-UNet achieved a sensitivity of 92.04%.

Another example of superior performance across all basic binary segmentation evaluation metrics, DFAB-UNet achieved a DSC of 94.30% and IoU of 89.21%. Higher values in these metrics indicate a well-predicted segmentation that aligns closely with the ground truth.

DFAB-UNet also achieved the highest specificity of 98.95%, i.e., the highest proportion of true non-lesional voxels while minimizing false negatives.

Moreover, DFAB-UNet significantly outperforms alternative techniques in these metrics, showcasing it as a robust, stable, and promising solution for automated skin lesion segmentation on the PH2 dataset. Furthermore, Table 2 displays the segmentation performance of the proposed DFAB-UNet regarding the HAM10000 dataset.

Finally, Figs. 9 and 10 present the segmentation results of DFAB-UNet in comparison to the baseline UNet on the PH2 and HAM10000 datasets, respectively.

The DFAB-UNet was evaluated on the HAM10000 dataset and compared with other state-of-the-art techniques. The results indicate that DFAB-UNet achieved the highest pixel-level classification accuracy of 96.20%, highlighting its strong potential

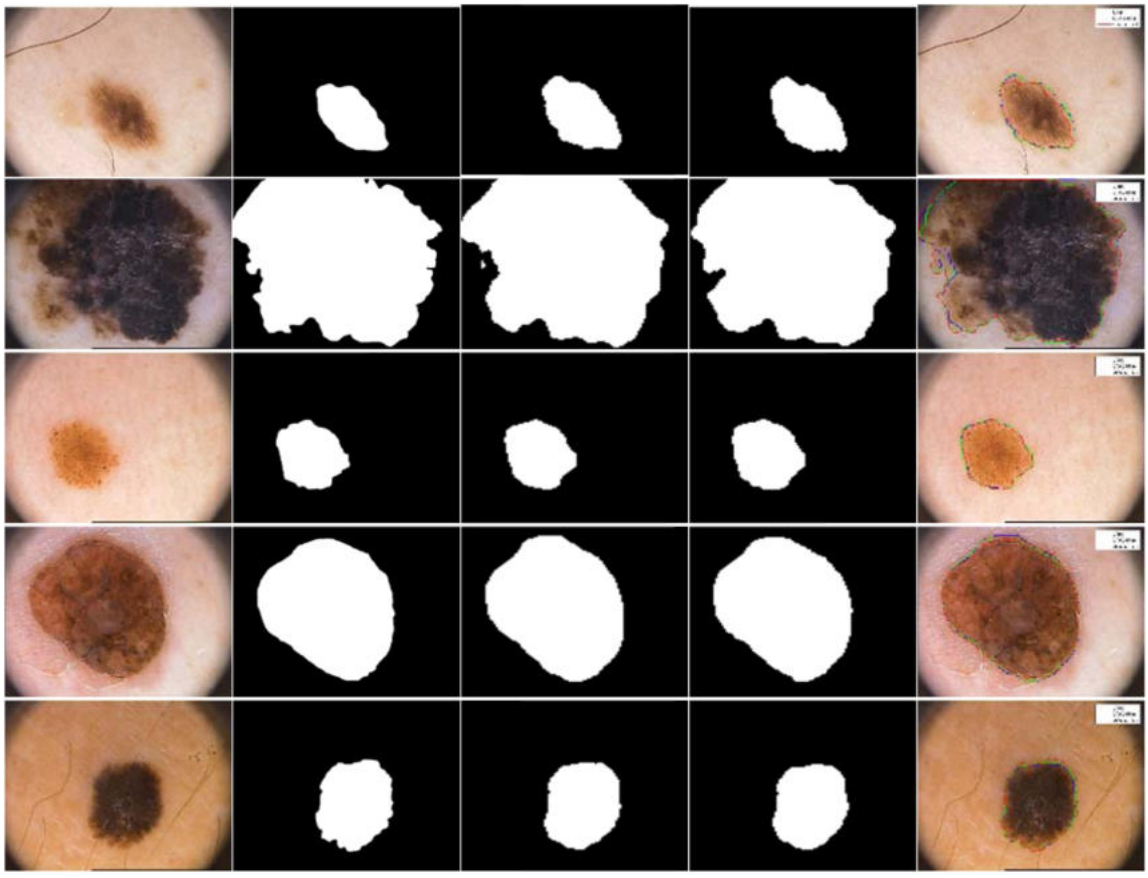

Fig. 9. PH2 dataset segmentation results: (left to right) Original image, ground truth, UNet, DFAB-UNet, and boundaries.

for pixel-wise classification segmentation. Additionally, DFAB-UNet achieved the best Sen together with Zhang et al.,[29] with a value of 92.31%. This underscores its capability in true positive detection and lesion localization.

DFAB-UNet also achieved the highest DSC of 92.79%, indicating superior performance in handling deformations and ensuring high-quality segmentations. Finally, the Spatial Distance Maps (SDMs) were used to visualize the segmentation performance of DFAB-UNet across different cases, as shown in Fig. 11.

The SDM visualizations show three regions: the lesion core (red), the lesion rim (green) and background (blue). Red: the lesion core, where pixel intensity is inversely proportional to the distance to the lesion center. The boundary is delineated with green edges, indicating the model's ability to get edge details perfectly. Blue represents the background outside of lesion. The SDM examples show distinct transition at their boundaries and little background penetration, which demonstrates the applicability of our model in discerning lesions from non-lesion regions, as well as proving its stability to maintain structural integrity and contour precision across various lesion morphologies.

Ablation study

The hyperparameters of DFAB-UNet were selected based on extensive experiments and established best practices in deep learning for medical image segmentation. To validate these choices, an ablation study was performed to systematically search for the optimal hyperparameter settings for training DFAB-UNet. The key hyperparameters of the DFAB-UNet are displayed in Table 3.

Table 3. Hyperparameters and their selection.

Hyperparameter	Value
Batch Size	10
Number of Epochs	50
Learning Rate	1×10^{-5}
Optimizer	Adam Optimizer
Loss Function	Binary Cross-Entropy

In addition, an ablation study was performed, where these hyperparameters were systematically modified to evaluate their contribution to

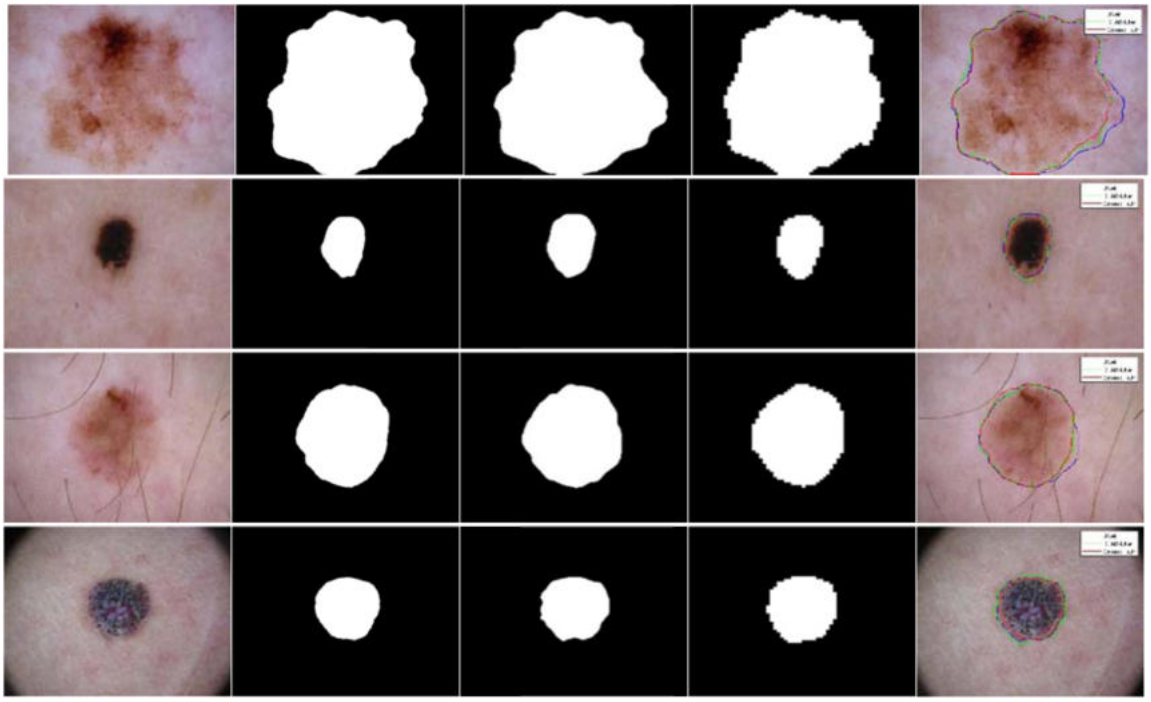

Fig. 10. HAM10000 dataset segmentation results: (left to right) Original image, ground truth, UNet; DFAB-UNet, and boundaries.

segmentation performance. Their impact was determined by using different metrics, including Acc, Pre, Sen, Dice, and IoU. The impact of batch size is demonstrated in Table 4, where different batch sizes are evaluated.

Table 4. Impact of batch size.

Batch Size	Acc	Pre	Sen	DSC	IoU
5	93.78	96.01	94.10	93.72	89.25
10	96.71	97.97	96.84	95.56	91.49
20	92.80	94.65	93.52	92.94	88.71

A batch size of 10 demonstrated the best balance between computational efficiency and model performance. On the other hand, smaller batch sizes resulted in noisier training conditions or noisier gradients. In comparison, larger batch sizes caused a decline in the network's performance due to a loss of generalization. Moreover, the impact of the learning rate is displayed in Table 5, where different learning rate values are evaluated.

Table 5. Impact of learning rate.

Learning Rate	Acc	Pre	Sen	DSC	IoU
1×10^{-4}	91.82	93.56	92.10	92.01	87.80
1×10^{-5}	96.71	97.97	96.84	95.56	91.49
1×10^{-6}	93.20	94.67	93.56	93.14	88.62

With a learning rate of 1×10^{-5} we were able to achieve a smooth convergence without excessive jumps or decreased learning speed.

Larger learning rates caused random oscillation during training, while smaller reduced convergence speed. Table 6 shows the performance evaluation of SGD, Adam, and RMSprop optimizers.

Table 6. Impact of optimizer.

Optimizer	Acc	Pre	Sen	DSC	IoU
SGD	89.78	91.68	90.12	90.53	85.72
Adam	96.71	97.97	96.84	95.56	91.49
RMSprop	93.88	95.14	94.12	93.46	89.32

The Adam optimizer yielded better results than other optimizers due to its adaptive adjustment of the learning rate and efficient handling of sparse gradients. Furthermore, Table 7 presents a detailed performance evaluation of the impact of the loss function.

Table 7. Impact of loss function.

Loss Function	Acc	Pre	Sen	DSC	IoU
Mean Squared Error	88.65	91.02	90.10	89.22	84.10
Focal Loss	94.89	96.02	95.28	94.18	90.42
Binary Cross-Entropy	96.71	97.97	96.84	95.56	91.49

Binary cross-entropy proved to be the most effective loss function for binary segmentation problems, as it strikes a good balance between precision and sensitivity. Finally, Table 8 provides the impact of the number of epochs.

Training for 50 epochs provided optimal balance between training duration and performance, since additional epochs lead to diminishing improvements.

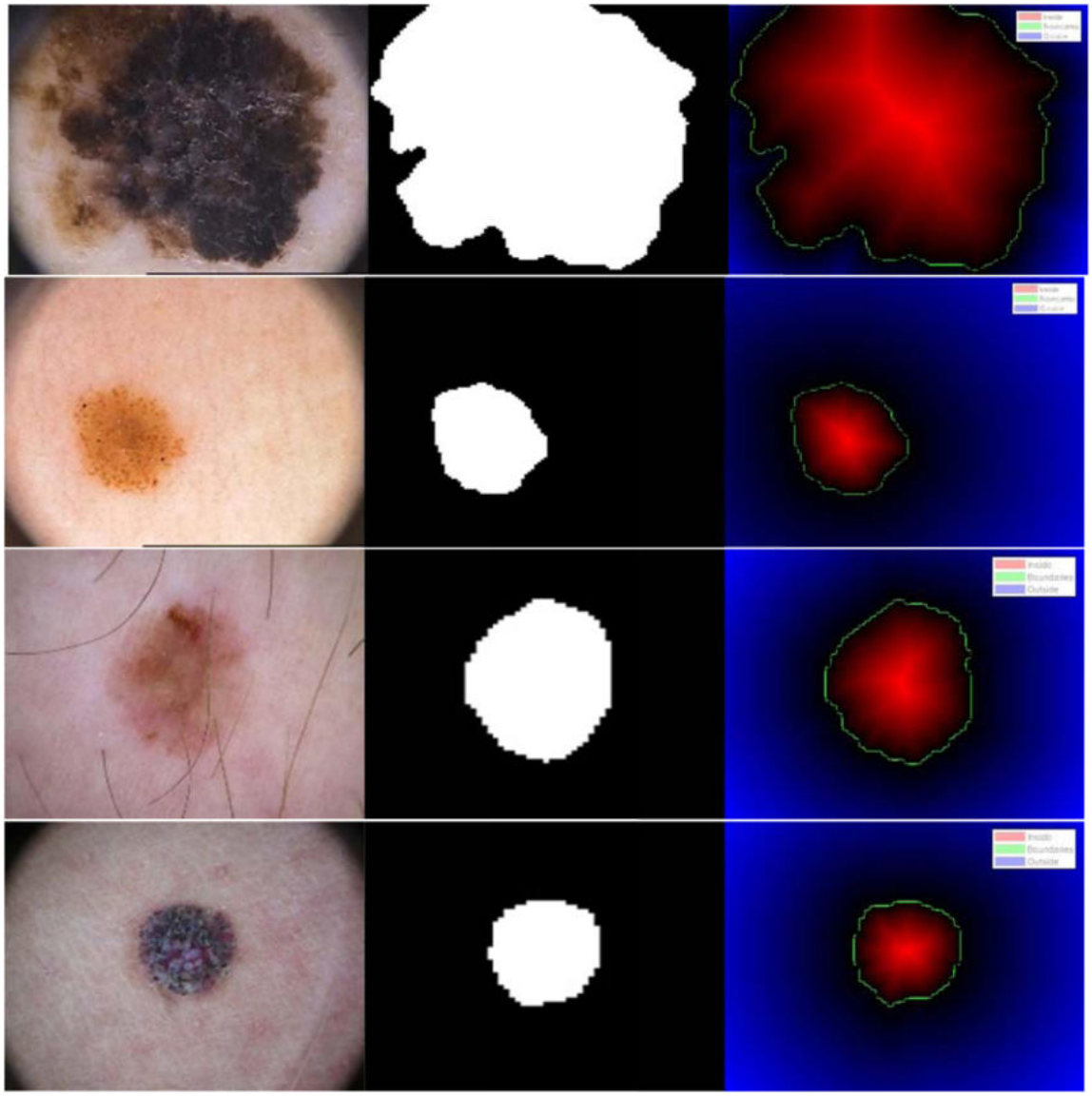

Fig. 11. DFAB-UNet SDM results on PH2 and HAM1000 datasets.

Table 8. Impact of number of epochs.

Number of Epochs	Acc	Pre	Sen	DSC	IoU
20	92.89	94.78	93.58	93.01	88.45
50	96.71	97.97	96.84	95.56	91.49
100	95.89	97.02	96.30	95.12	90.98

This subsection presents a comprehensive analysis and validation of the DFAB-UNet hyperparameters, demonstrating that the selected parameters maximize performance in distinguishing skin lesions from normal tissue while maintaining computational efficiency. Our systematic evaluation shows that the best performance was obtained using a batch size of 10, a learning rate of 1×10^{-5}, and Adam optimization. For this binary segmentation task, the appropriate loss function is the binary cross-entropy. Training for 50 epochs provides a good balance between performance and computational time. Thus, these hyperparameter choices guarantee that DFAB-UNet achieves state-of-the-art segmentation performance with minimal computational cost.

Conclusion

The goal of this study was to design an accurate and robust model for skin lesion segmentation, targeting effective operation on dermoscopic images with minimal pre-processing. The presented DFAB-UNet framework achieves the goal by introducing a dual focus attention block to the conventional

U-Net, which has the capability of capturing local detailed structure and global contextual information. In contrast to many current approaches, the deep network works on downscaled input images with minimal pre-processing and post-processing and thus can be applied in real clinical work. The efficiency of the proposed model was successfully validated by performing extensive experiments on the PH2 and HAM10000 benchmark datasets. The DFAB-UNet performed remarkably well in the chosen metrics, accuracy, sensitivity, precision and and the DSC, while outperforming or being on par with the results of more sophisticated and more pre-processed models. These results indicate that the major aim of this work, to establish a consistent, effective and weakly dependent segmentation model for limb lesion analysis, is well achieved. However, though the proposed DFAB-UNet performs favorably on benchmark data sets, a few challenges remain before clinical translation. First, the dermoscopic datasets, such as PH2 and HAM10000 used here, do not fully model variability in natural clinical conditions where image quality, acquisition modality and patient demographics can vary significantly. This may affect model generalization and usefulness in different populations. Secondly, our previous framework is limited to binary lesion segmentation and is unable to address multi-class discrimination of lesions, which is also essential in clinical diagnosis. Thirdly, although the model reduces pre-processing of input volume to a minimum, artifacts like hair, bubbles and uneven light are still probable pitfalls that could influence the accuracy of segmentation. Finally, integration into clinical practice requires attention to computational efficiency, interpretability, and alignment with regulatory standards. These limitations will be an important focus of future work, which includes additional validation on larger-scale multi-center data sets, introduction of artifact suppression modules and construction of explainable AI methods to boost clinician confidence in automated output. Future work could consider integrating lightweight artifact suppression modules or extending the model to multi-class lesion classification to improve its diagnostic capability.

Authors' declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images that are not ours have been included with the necessary permission for republication, which is attached to the manuscript.
- No animal studies are present in the manuscript.
- Author(s) signed on ethical consideration's approval.
- Ethical Clearance: The project was approved by the local ethical committee at Universiti Teknologi Malaysia.

Data availability

The data supporting the findings of this study are completely available online. PH2 dataset: https://www.fc.up.pt/addi/ph2%20database.html and HAM10000 dataset: https://challenge.isic-archive.com/data/#2018.

Authors' contribution statement

F. H. N. participated in the concept and design, model structure design, data preparation, and writing of the original draft. F. M. project oversight, overall study design, and review and editing of the manuscript, model generation, data curation, and formal analysis. M. S. M. R. contributed to optimizing deep learning, evaluating the models, data analysis, visualization and resources. D. S. B. helped with the literature review, manuscript structuring, and validation results. L. F. B. gave technical support and helped implement the evaluation metrics. V. S. C. final manuscript review, proofreading, and critical revision.

References

1. Leiter U, Keim U, Garbe C. Epidemiology of skin cancer: update 2019. Adv Exp Med Biol. 2020:123–39. https://doi.org/10.1007/978-3-030-46227-7_6.
2. Albakaa ZH, Alhasani AT, Alkattan H, Alfilh RH, Abotaleb M, Dhoska K. Machine learning-based time series prediction of student academic performance: A comparative analysis of RF, SVM, and k-NN. In: International Conference on Intelligence-Based Transformations of Technology and Business Trends 2025 May 29:(pp. 370–379). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-07373-0_27.
3. Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, *et al.* Skin cancer detection: a review using deep learning techniques. Int J Environ Res Public Health. 2021 May 20;18(10):5479. https://doi.org/10.3390/ijerph18105479.
4. Najjar FH, Waheed SR, Mahdi DA. Coronavirus classification based on enhanced X-ray images and deep learning. Mal J Fund Appl Sci. 2023 May 26;19(3):369–78. https://doi.org/10.11113/mjfas.v19n3.2909.
5. Du G, Cao X, Liang J, Chen X, Zhan Y. Medical image segmentation based on U-net: A review. J Imaging Sci Technol. 2020 Mar 1;64(2). https://doi.org/10.2352/J.ImagingSci.Technol.2020.64.2.020508.
6. Kumar RR, Shankar SV, Jaiswal R, Ray M, Budhlakoti N, Singh KN. Advances in deep learning for medical image analysis:

A comprehensive investigation. J Stat Theory Pract. 2025 Mar;19(1):9. https://doi.org/10.1007/s42519-024-00422-2.
7. Lim SS, Shin K, Mun JH. Dermoscopy for cutaneous fungal infections: A brief review. Health Sci Rep. 2022 Jan;5(1):e464. https://doi.org/10.1002/hsr2.464.
8. Kazerouni IA, Dooly G, Toal D. Ghost-UNet: An asymmetric encoder-decoder architecture for semantic segmentation from scratch. IEEE access. 2021 Jul 5;9:97457–65. https://doi.org/10.1109/ACCESS.2021.3094925.
9. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing. 2021 Sep 10;452:48–62. https://doi.org/10.1016/j.neucom.2021.03.091.
10. Soydaner D. Attention mechanism in neural networks: where it comes and where it goes. Neural Comput. Appl. 2022 Aug;34(16):13371–85. https://doi.org/10.1007/s00521-022-07366-3.
11. Daneshjou R, Yuksekgonul M, Cai ZR, Novoa R, Zou J. SkinCon: A skin disease dataset densely annotated by domain experts for fine-grained model debugging and analysis. arXiv preprint arXiv: 2302.00785. 2023 Feb 1. https://doi.org/10.48550/arXiv.2302.00785.
12. Codella NC, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, *et al.* Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) 2018 Apr 4:(pp. 168–172). IEEE. https://doi.org/10.1109/ISBI.2018.8363547.
13. Gutman D, Codella NC, Celebi E, Helba B, Marchetti M, Mishra N, *et al.* Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). arXiv preprint arXiv: 1605.01397. 2016 May 4. https://doi.org/10.48550/arXiv.1605.01397.
14. Codella N, Rotemberg V, Tschandl P, Celebi ME, Dusza S, Gutman D, *et al.* Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC). arXiv preprint arXiv: 1902.03368. 2019 Feb 9. https://doi.org/10.48550/arXiv.1902.03368.
15. Combalia M, Codella NC, Rotemberg V, Helba B, Vilaplana V, Reiter O, *et al.* Bcn20000: Dermoscopic lesions in the wild. arXiv preprint arXiv:1908.02288. 2019 Aug 6. https://doi.org/10.48550/arXiv.1908.02288.
16. Rotemberg V, Kurtansky N, Betz-Stablein B, Caffery L, Chousakos E, Codella N, *et al.* A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci. Data. 2021 Jan 28;8(1):34. https://doi.org/10.1038/s41597-021-00815-z.
17. Mendonça T, Ferreira PM, Marques JS, Marcal AR, Rozeira J. PH 2-A dermoscopic image database for research and benchmarking. In 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2013 Jul 3:(pp. 5437–5440). IEEE. https://doi.org/10.1109/EMBC.2013.6610779.
18. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data. 2018 Aug 14;5(1):1–9. https://doi.org/10.1038/sdata.2018.161.
19. Gurcan F, Soylu A. Synthetic boosted resampling using deep generative adversarial networks: A novel approach to improve cancer prediction from imbalanced datasets. Cancers. 2024 Dec 2;16(23):4046. https://doi.org/10.3390/cancers16234046.
20. Vaidya A, Chen RJ, Williamson DF, Song AH, Jaume G, Yang Y, *et al.* Demographic bias in misdiagnosis by computational pathology models. Nat. Med. 2024 Apr;30(4):1174–90. https://doi.org/10.1038/s41591-024-02885-z.
21. Khalaf M, Dhannoon BN. MSRD-Unet: multiscale residual dilated U-Net for medical image segmentation. Baghdad Sci J. 2022 Dec 5;19(6 (Suppl.)):1603. https://doi.org/10.21123/bsj.2022.7559.
22. Dong C, Dai D, Zhang Y, Zhang C, Li Z, Xu S. Learning from dermoscopic images in association with clinical metadata for skin lesion segmentation and classification. Comput Biol Med. 2023 Jan 1;152:106321. https://doi.org/10.1016/j.compbiomed.2022.106321.
23. Nasarian E, Alizadehsani R, Acharya UR, Tsui KL. Designing interpretable ML system to enhance trust in healthcare: A systematic review to proposed responsible clinician-AI-collaboration framework. Inf Fusion. 2024 Apr 6:102412. https://doi.org/10.1016/j.inffus.2024.102412.
24. Karimi A, Faez K, Nazari S. DEU-NET: Dual-encoder u-net for automated skin lesion segmentation. IEEE Access 2023;11:134804–134821. https://doi.org/10.1109/ACCESS.2023.3337528.
25. Yang L, Fan C, Lin H, Qiu Y. Rema-Net: An efficient multi-attention convolutional neural network for rapid skin lesion segmentation. Computers in Biology and Medicine 2023;159:106952. https://doi.org/10.1016/j.compbiomed.2023.106952.
26. Liu Z, Hu J, Gong X, Li F. Skin lesion segmentation with a multiscale input fusion U-Net incorporating Res2-SE and pyramid dilated convolution. Sci Rep 2025;15(1):7975. https://doi.org/10.1038/s41598-025-92447-1.
27. Tomar NK, Jha D, Riegler MA, Johansen HD, Johansen D, Rittscher J, *et al.* A feedback attention network for improved biomedical image segmentation. IEEE Trans Neural Netw Learn Syst. 2022 Mar 25;34(11):9375–88. https://doi.org/10.1109/TNNLS.2022.3159394.
28. Srivastava V, Gupta S, Singh R, Gautam V. A multi-level closing based segmentation framework for dermatoscopic images using ensemble deep network. Int J Syst Assur Eng Manag. 2024 Aug;15(8):3926–39. https://doi.org/10.1007/s13198-024-02393-w.
29. Zhang J, Che M, Wu Z, Liu Y, Liu X, Zhang H, *et al.* Segmentation of skin lesion using superpixel guided generative adversarial network with dual-stream patch-based discriminators. Biomed Signal Process Control. 2024 Aug 1;94:106304. https://doi.org/10.1016/j.bspc.2024.106304.
30. Li Z, Zhang N, Gong H, Qiu R, Zhang W. SG-MIAN: Self-guided multiple information aggregation network for image-level weakly supervised skin lesion segmentation. Comput Biol Med. 2024 Mar 1;170:107988. https://doi.org/10.1016/j.compbiomed.2024.107988.
31. Zhi Y, Bie H, Wang J, Ren L. Masked autoencoders with generalizable self-distillation for skin lesion segmentation. Med Biol Eng Comput. 2024 Apr 24:1–5. https://doi.org/10.1007/s11517-024-03086-z.
32. Kaur R, Kaur S. Automatic skin lesion segmentation using attention residual U-Net with improved encoder-decoder architecture. Multimed. Tools Appl. 2025 Mar;84(8):4315–41. https://doi.org/10.1007/s11042-024-18895-5.
33. Li X, Peng B, Hu J, Ma C, Yang D, Xie Z. USL-Net: Uncertainty self-learning network for unsupervised skin lesion segmentation. Biomed. Signal Process. Control. 2024 Mar 1;89:105769. https://doi.org/10.1016/j.bspc.2023.105769.
34. Zhao M, Zhou B, Hung LJ. LUT-SLS: A lightweight transformer network based on U-net for skin lesion segmentation. J Internet Technol. 2024 Sep 1;25(5):795–805. https://doi.org/10.70003/160792642024092505015.

35. Garbaz A, Oukdach Y, Charfi S, El Ansari M, Koutti L, Salihoun M. DMFC-UFormer: Depthwise multi-scale factorized convolution transformer-based UNet for medical image segmentation. Biomed Signal Process Control. 2025 Mar 1;101:107200. https://doi.org/10.1016/j.bspc.2024.107200.
36. Najjar FH, Kadhim KA, Mohamed F, Mohd Rahim MS, Haidar Abdullah AN. A lightweight U-net model for accurate skin lesion segmentation. Iraqi J Comput Sci Math. 2025;6(2):1. https://doi.org/10.52866/2788-7421.1230.
37. Najjar FH, Kadhim KA, Kareem MH, Salman HA, Mahdi DA, Al-Hindawi HM. Classification of COVID-19 from X-ray images using GLCM features and machine learning. Mal J Fund Appl Sci. 2023 May 26;19(3):3 89–98. https://doi.org/10.11113/mjfas.v19n3.2911.
38. Nisa SQ, Ismail AR. Dual U-net with ResNet encoder for segmentation of medical images. Int J Adv Comput Sci Appl. 2022;13(12). https://doi.org/10.14569/IJACSA.2022.0131265.

تجزئة الآفات الجلدية بدقة عالية وموثوقية فائقة باستخدام DFAB-UNT

فلاح حسن نجار[1، 2]، فرحان محمد[1، 3]، محمد شفري محمد رحيم[1]، ديانا برناردو[4]، لويس برناردو[5]، تشان سيانغ[1، 3]

[1] قسم الحوسبة الناشئة، كلية الحوسبة، الجامعة التكنولوجية الماليزية، جوهور باهرو، ماليزيا.
[2] قسم تقنيات شبكات وبرامجيات الحاسوب، المعهد التقني النجف، جامعة الفرات الأوسط التقنية، النجف، العراق.
[3] مركز التميز للابتكار في الوسائط والألعاب، معهد الهندسة المرتكزة على الإنسان، الجامعة التكنولوجية الماليزية، جوهور باهرو، ماليزيا.
[4] مركز شامباليمود للمجهول، لشبونة، البرتغال.
[5] مركز التقنيات البايوتكنولوجية، قسم الهندسة المدنية والمعمارية، جامعة بيرا الداخلية، كوفيليا، البرتغال.

الملخص

سرطان الجلد هو أكثر أنواع السرطان شيوعًا في العالم، ويُعدّ الميلانوما (سرطان الجلد الخبيث) الأكثر فتكًا. يُعد الاكتشاف المبكر والتشخيص الدقيق للآفات الجلدية أمرين حاسمين لعلاج السرطان، وعادةً ما يحقق المرضى نتائج أفضل عندما يتم اكتشاف المرض مبكرًا. تختلف الصور الجلدية المأخوذة بالدرموسكوب من حيث الإضاءة والدقة ونوع البشرة والشعر، بالإضافة إلى وجود بعض الفقاعات. يمكن أن تؤثر هذه العوامل سلبًا على قدرة النموذج على التعميم، ولذلك هناك حاجة إلى اعتماد عملية معالجة مسبقة صارمة واستراتيجية فعّالة لتعزيز البيانات أثناء تدريب النموذج. تقدم هذه الورقة البحثية هيكلية جديدة لتجزئة آفات الجلد تُعرف باسم (DFAB-UNet). يهدف النموذج المقترح إلى الدمج بين مزايا التقاط السمات المحلية الدقيقة والمعلومات العامة للمظهر، مما يضمن استخراج سمات عالية الجودة ويُعزز من أداء النموذج في عملية التجزئة. تم تقييم فعالية النموذج باستخدام مجموعتين من بيانات الصور الجلدية المأخوذة بواسطة التنظير الجلدي، وهما PH2 وHAM10000. أظهر النموذج أداءً متميزًا على مجموعة بيانات PH2، حيث بلغت الدقة 97.24%، الدقة الإيجابية 96.66%، الحساسية 92.04%، Dice 94.30%، IoU 89.21%، النوعية 98.95%. أما على مجموعة بيانات HAM10000، فقد حقق النموذج دقة بلغت 96.20%، دقة إيجابية 93.28%، حساسية 92.31%، Dice 92.79%، IoU 86.55%، نوعية 97.60%. تعكس هذه النتائج متانة وموثوقية نموذج DFAB-UNet، مما يجعله أداة تشخيصية ورقابية واعدة في مجال أبحاث الأمراض الجلدية.

الكلمات المفتاحية: سرطان الجلد، شبكة U-Net، شبكة DFAB-UNet، تقسيم الصور الطبية، تقسيم الآفات الجلدية.