Academic Journal of Electrical and Computer Engineering
Vol. 02, No. 01, August 2026
https://doi.org/10.31272/ajece.39
https://ajece.uomustansiriyah.edu.iq
ISSN 3079-2878
Research Article

Comparative Study of Deep Learning Architectures for Skin Lesion Segmentation

Abbas H. Miry[1]*, Ammar Al-Gizi[2], Gregor A. Aramice[3]

1,2,3 Department of Electrical Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq
*Email: abbasmiry83@uomustansiriyah.edu.iq
*ORCID: https://orcid.org/ 0000-0002-7456-287X

Received: 01/12/2025 Revised: 01/02/2026 Accepted: 07/02/2026

Abstract— Identifying cancer at an earlier stage gives possibility to obtain better treatment outcome. On the other hand, manually outlining the precise boundaries of skin lesions in dermoscopic images is time-consuming and varies between physicians. In this paper, we investigate the performance of two widely used deep-learning models for lesion segmentation U-Net and DeepLabV3+. For testing each of the model, we used two pre-trained feature extractors (encoders) which are ResNet-50 and EfficientNet-B0 (all pre-trained on ImageNet). The ISIC-2020 dataset was used with all images resized to 256×256 and the basic data augmentation was performed to enhance the training process. For performance measurement, we used Dice and IoU on a different test set. The highest performance was gotten from U-Net with EfficientNet-B0, which scored Dice = 0.920 and IoU = 0.859. In second best was DeepLabV3+ using EfficientNet-B0 (Dice = 0.914, IoU = 0.852). In summary, our results demonstrate that when keeping the training regime constant, the encoder choice makes a big difference. Skin lesion segmentation results Our experiments reveal that EfficientNet-B0 nearly always surpassed ResNet-50 across skin lesion segmentation.

Index Terms— Skin lesion segmentation, Deep learning, U-Net, DeepLabV3+, ISIC dataset.

I. INTRODUCTION

Skin cancer is one of the most common cancers worldwide, and melanoma is among the most dangerous types because it can progress quickly when diagnosis is delayed [1], [2]. Dermoscopy helps clinicians see pigmented lesions more clearly than the naked eye, but the interpretation of dermoscopic images can still differ from one clinician to another. This variation may cause inter-observer disagreement and can sometimes lead to missed cases or late diagnosis [3], [4].In computer-aided diagnosis systems, precise segmentation of the lesion area is a key step because it supports accurate measurement of clinically important features such as lesion size, shape, and border irregularity. Having good segmentation reduces the error propagation from earlier steps into later stages of the pipeline (in particular classification) [7], [8]. Nonetheless, dermoscopic images contain frequent artifacts, such as the presence of hair, rulers, and shadows or weak contrast between the lesion and background skin, that make them difficult to segment automatically. Moreover, among different patients and different imaging conditions, lesions can appear quite heterogeneous, which complicates the development of robust, reliable segmentation methods [9], [10].

Previous methods for segmentation were primarily based on classical image processing techniques like thresholding and active contour models, but they are less reliable on real clinical images due to artifacts and presence of indistinct boundaries [10]. Consequently, deep learning methods, particularly convolutional neural networks (CNNs), have become commonplace due to their ability to directly learn strong image representations from labelled datasets, while handling the variability in imagery more robustly [6], [7].

II. RELATED WORKS

The U-Net family and boundary refinement: U-Net [11] introduced an encoder-decoder design with skip connections to recover fine spatial details and has continued to be an important baseline for biomedical segmentation. Later on, residual/attention blocks and multi-scale features are introduced to improve boundary delineation, which are realized in MRP-UNet [16] and boundary-assisted designs, e.g. LB-UNet [19]. Light Transformed U-shaped models (i.e., LeaNet) are practical models that reduce computations and maintain accuracy for deployment [18].

Multi-scale context & atrous convolution: DeepLabV3+ uses atrous convolution in combination with Atrous Spatial Pyramid Pooling (ASPP) to integrate multi-scale context (in different receptive-field scales), which helps to learn better features when the lesions differ considerably in size/shape [13]. Different studies report improvements to boundary quality or feature selection in DeepLab-like pipelines over skin lesions dermoscopic images [14], [30]

Transformer and hybrid architectures: Following the success of transformer architectures in the field of natural

language processing, transformer modules have been incorporated into U-Net-like architectures to improve long-range dependency modeling and boundary consistency [17]. Hybrid CNN-transformer models (ESC-UNet) integrate convolutional induction bias with attention mechanism to enhance segmentation against difficult appearances [21].

Semi-supervised & Pipeline based methods: As pixel level labeling is expensive, weakly annotated or unlabeled data can be utilized for segmentation using semi-supervised or self-training strategies [23]. Another approach is to combine detection and segmentation (such as YOLO + SAM based methods) to locate lesions and then refine the mask [26], [27].

Summary and Gap: Two fundamental directions dictate the literature: (i) Boundary Accuracy, (ii) Model Complexity—the balance between a model's complexity and its practical operation. These are evident based on our exploration of reduced complexity architectures specifically designed for applications on devices with limited computational resources; Furthermore, we aggregate additional strategies directed towards improving generalization under data scarcity and across similar dataset instances (e.g., across ISIC subsets and preprocessing) [5], [22], [31]. Despite the big difference, many studies vary in data splits, preprocessing, and training protocols, making it hard to make direct comparisons.In this work, we therefore perform a controlled comparison of U-Net and DeepLabV3+ using two encoders under the same pipeline, and we report results on a held-out test set.

III. SKIN LESION SEGMENTATION

Skin lesion segmentation is the task of delineating the lesion region in a dermoscopic image by producing a binary mask that separates the lesion from surrounding healthy skin. This mask is then used to support downstream computer-aided diagnosis steps such as feature extraction, lesion size estimation, and longitudinal monitoring as shown in fig.1.

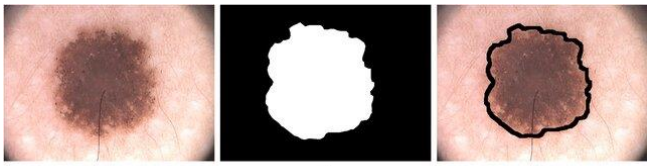

Fig. 1. An example of a skin lesion segmentation: (a) a dermoscopic image; (b) a binary mask; (c) a mask boundary overlaid on the image.

Previous works often utilized traditional techniques like threshold, edge detection, or active contours. They were presented in a relatively simple form and were less robust and often failed when the lesion was poorly defined, having ragged borders, or was affected by noise or artifacts such as hair and shadows [30]. Deep learning has changed this problem to a significant extent. These models are employed extensively according to their simple versions (U-Net) or complex ones (DeepLabV3+) that enable the deep networks to dislocate both global structures and local textures to perform better segmentation [31], [32], [34]. These models can handle a variety of lesions with different shapes and colours by learning from mass datasets composed of many labelled images. The goal of accurate segmentation is not only for detection but also, facilitates downstream tasks such as classification, lesion evolution monitoring and integration into screening pipelines. Importantly, automatic segmentation could help to relieve the burden for clinicians and potentially decrease error rates resulting from manual contouring of regions of interests (ROIs) [33].

IV. METHODOLOGY

In this section, we explain the experiment in a natural sequence, from dataset preparation, through model training, to result evaluation, as depicted in Fig. 2. It was also important to create a reproducible, transparent workflow, which other researchers could repeat under similar conditions.

1- Dataset and Data Organization:

Experimental Setup We used ISIC Challenge dataset which is widely used in skin-lesion research [15]. It comprises many paired dermoscopic color images and their corresponding binary mask, where the white is the lesion and the black is the background [15]. The files are divided into two folders named for model training and validation (Training_Data and Training_GroundTruth) and for final testing (Test_Data with Test_GroundTruth). These are stored using the same filename as their mask making it even simpler to link them. In order to have the same input shape, we preprocessed each image and mask to 256 x 256 pixels. Masks We applied nearest-neighbour resizing on the masks to avoid blurring and keeping them as binary masks. To avoid bias and conduct an appropriate evaluation, the dataset was split into 80% training, 10% validation and 10% test [5]. This was an entirely independent test set that was not at all used for training or model selection, allowing for an unbiased final assessment of the model performance.

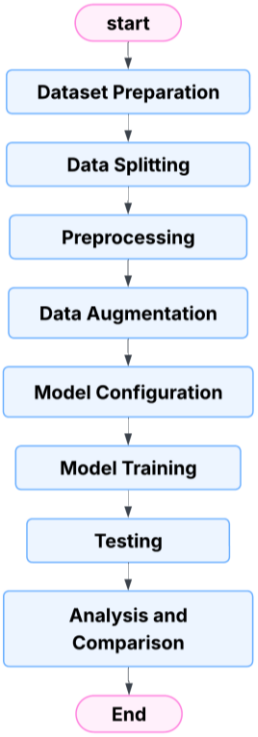

Fig. 2. Overall workflow of the experimental pipeline

2-Preprocessing and Augmentation : Dermoscopic images may have differences of illuminations, orientation and also contain artifical elements like a hair or a ruler mark. To improve robustness, we applied standard data augmentation techniques including horizontal/vertical flips, small rotations, random scaling, and brightness/contrast adjustments. Each image and its associated mask had all geometric transforms participated equally. Lastly, images were normalized using ImageNet statistics in order for pretrained encoders to be used properly.

3- Model Architectures - We show here two common segmentation deep learning models: U-Net - This is a type of encoder-decoder with skip connections (Fig. 3). The two skip connections are designed to retain the detail information on the input image, which is very important when the edges of lesion are thin or unclear. One reason that the U-Net has had so much success when it comes to medical imaging is its elegant simplicity [11].

DeepLabV3+ A more recent model employs atrous convolutions and ASPP (Atrous Spatial Pyramid Pooling) (Fig. 4). These help the model to observe features at various scales useful when detecting large, small, or irregular lesions. Nov steals some smarts from LAH with its lightweight decoder to sharpen the chirp borders [13].

Tables 1 and 2 summarize the layers of both networks. These tables summarize the architectures of U-Net and DeepLabV3+ and how they process the images step by step. U-Net takes an encoder–decoder shape with skip connections such that it preserves the fine details of the input image and recovers them in the decoder phase

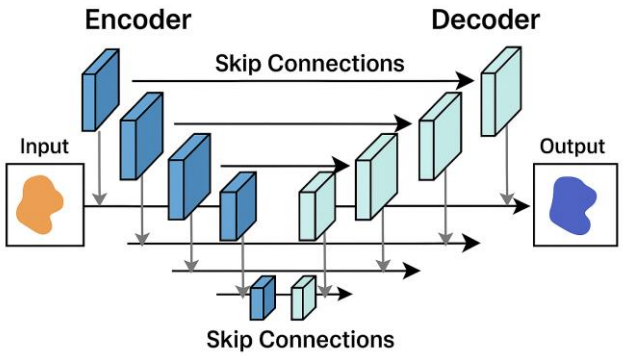

Figure 3.. U-Net architecture

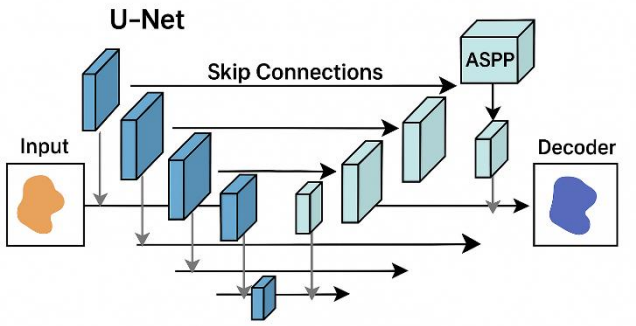

Fig. 4. DeepLabV3+ architecture

Table I. U-Net Layers

Stage	Main Operation	Output
Input	—	3 × 256 × 256
Encoder Stage 1	Conv/BN/ReLU + Downsample	64 × 128 × 128
Encoder Stage 2	Conv/BN/ReLU + Downsample	128 × 64 × 64
Encoder Stage 3	Conv/BN/ReLU + Downsample	256 × 32 × 32
Encoder Stage 4	Conv/BN/ReLU + Downsample	512 × 16 × 16
Bottleneck	Conv blocks	1024 × 8 × 8
Up + Skip (F4)	Upsample + Concat + Convs	256 × 16 × 16
Up + Skip (F3)	Upsample + Concat + Convs	128 × 32 × 32
Up + Skip (F2)	Upsample + Concat + Convs	64 × 64 × 64
Up + Skip (F1)	Upsample + Concat + Convs	32 × 128 × 128
Output Head	1×1 Conv → Logits → Sigmoid	1 × 256 × 256

This is especially useful if the lesion margins are small or unclear. DeepLabv3+ adopts another strategy: Atrous Spatial Pyramid Pooling (ASPP) to extract the multi-scale features to enhance robustness for small as well as large, irregular lesions. The encoder (ResNet50, EfficientNet-b0, etc.) that are used are also quite important for capacity and generalization. A simpler way to compare the two architectures can be found in

Tables I–II: U-Net focuses on skip connections to capture fine detail, while DeepLabV3+ uses ASPP to capture contextual scale. ImageNet-pretrained encoders can accelerate convergence and achieve higher final accuracy with both approaches.

Table II. DeepLabV3+ Layers

Stage	Main Operation	Output
Input	—	3 × 256 × 256
Encoder (low-level)	Early features (then reduced by 1×1 Conv)	48 × 64 × 64
Encoder (high-level)	Deep features from backbone	2048 × 16 × 16
ASPP	1×1 Conv + 3×3 Atrous (r=6,12,18) + Pooling	256 × 16 × 16
Upsample	Bilinear upsampling	256 × 64 × 64
Decoder fusion	Concat(ASPP↑ , low-level) + 3×3 Convs	256 × 64 × 64
Output Head	Upsample → 1×1 Conv → Logits → Sigmoid	1 × 256 × 256

All experiments were conducted in a cloud GPU environment (Google Colab). We trained for up to 20 epochs with a batch size of 8 using the Adam optimizer (learning rate = 1e-3). Early stopping was enabled to mitigate overfitting by selecting the checkpoint with the best validation Dice. Unless otherwise stated, the dataset split used throughout this study is 80% for training, 10% for validation, and 10% for final testing. All reported results in Section VI correspond to evaluation on the held-out test set.

4- Loss/Objective Function:

The Dice function is a measure of overlap widely used to assess segmentation performance when a gold standard or ground truth is available. Used as a loss function, the 2-class variant of the Dice loss, can be expressed as [34]

$$\text{Dice} = 1 - \frac{\sum_{n=1}^{N} p_n r_n + \epsilon}{\sum_{n=1}^{N} p_n + r_n + \epsilon} - \frac{\sum_{n=1}^{N} (1-p_n)(1-r_n) + \epsilon}{\sum_{n=1}^{N} 2 - p_n - r_n + \epsilon} \quad (1)$$

Where p_n is predicted probability for pixel n , r_n is ground-truth mask , where ε is a small constant to avoid division by zero. At inference time, the probability map is converted to a binary mask using a fixed threshold (0.5). During training, we minimized the Dice loss, which directly targets overlap between the predicted and true masks:

V. EVALUATION METRICS

To check how well the models worked, we used two metrics: Dice and IoU (Intersection over Union). Both are very common in medical image segmentation and appear in many recent studies [5].Dice score tells us how much the predicted mask overlaps with the true mask. It is calculated like this:

$$\text{Dice} = 2TP / (2TP + FP + FN) = 2|P \cap G| / (|P| + |G|) \quad (2)$$

$$\text{IoU} = TP / (TP + FP + FN) = |P \cap G| / |P \cup G| \quad (3)$$

Where P denotes the predicted binary mask and G denotes the ground-truth mask. TP, FP, and FN are computed at the pixel level. Dice emphasizes overlap and is commonly used in medical segmentation, while IoU is a stricter measure because it divides the intersection by the union. Reporting both metrics provides a more complete view of segmentation quality, especially when lesions occupy a small fraction of the image.

VI. RESULTS AND DISCUSSION

Table III summarizes the quantitative performance of the evaluated configurations on the held-out ISIC-2020 test set. Overall, configurations using EfficientNet-b0 consistently achieved higher Dice and IoU than those using ResNet-50 under the same training protocol.

Table III. Test-set results on ISIC-2020 (our experiments)

Model	Encoder	Dice	IoU
U-Net	EfficientNet-b0	0.920	0.859
DeepLabV3+	EfficientNet-b0	0.914	0.852
U-Net	ResNet-50	0.901	0.831
DeepLabV3+	ResNet-50	0.900	0.833

Table 4. Comparison with representative recent studies (reported metrics), Metrics in Table 4 are taken from other works and not be directly comparable due to differences in datasets (ISIC subsets) protocols.

Table V. Comparison

Study / Method	Category	Dice	IoU
YOLOSAMIC [26]	Detection + segmentation	0.92	0.85
LeaNet [18]	Lightweight U-Net	0.91	0.84
ELANet [22]	Lightweight attention net	0.90	0.83
Noisy-Student [23]	Semi-supervised / self-training	0.91	0.84

U-Net with EfficientNet-b0 obtained the best scores (Dice = 0.920, IoU = 0.859). DeepLabV3+ with EfficientNet-b0 followed closely (Dice = 0.914, IoU = 0.852), suggesting that both multi-scale context (ASPP) and skip connections can be effective when paired with a parameter-efficient encoder.

When using ResNet-50, performance decreased for both architectures (U-Net: Dice = 0.901, IoU = 0.831; DeepLabV3+: Dice = 0.900, IoU = 0.833). A plausible explanation is that, given the fixed input resolution and limited training budget, the higher-capacity encoder may not provide additional benefits compared with a lighter encoder that generalizes well in this setting; similar trends have been reported for lightweight segmentation models in the literature [18], [22].

Qualitative analysis: Fig. 5 shows representative examples of predicted masks and overlays. In images with clear contrast, both models capture the lesion region with accurate overall shape. Failure cases are more likely when boundaries are faint or when hair artifacts partially occlude the lesion, which can lead to small boundary deviations or under-segmentation near thin protrusions. These observations highlight the importance

of boundary-aware training and artifact handling for robust clinical use.

Statistical reliability and limitations: The present study reports point estimates from the adopted training protocol. In future revisions, we will extend the evaluation to multiple random seeds and report mean ± standard deviation and significance testing across runs to quantify variability more rigorously.

Where we use ResNet50 encoder, the numbers go a bit down. U-Net with ResNet50 got Dice = 0.901 and IoU = 0.831, while DeepLabV3+ with ResNet50 was almost the same, Dice = 0.900 and IoU = 0.833. So here the two models are very close, but still less than the EfficientNet-b0 cases. Maybe because ResNet50 is heavy network and need more data, or it not capture the small features so good.

Overall, from the table we can say that EfficientNet-b0 encoder is better choice for skin lesion segmentation in our dataset. U-Net gave the best results, but DeepLabV3+ was also has good result . We can see that U-Net with EfficientNet-b0 gave the best Dice (0.920) and IoU (0.859) due to EfficientNet-b0 is light but strong encoder, and U-Net structure with skip connections makes it good at keeping the details of the lesion borders. Both models were able to capture good features without losing too much information.

For DeepLabV3+ with EfficientNet-b0, the results are good (Dice 0.914, IoU 0.852). This show that both models have good performance, but U-Net got just a little better. The difference due to the method of U-Net passes low-level features with skip connections, which helps in segmentation tasks like skin lesion where edges are very important.

At ResNet50 encoder, both U-Net and DeepLabV3+ scored lower results (Dice =0.90, IoU =0.83). Not as good as EfficientNet-b0, but this is still fine. Perhaps because ResNet50 was heavier and with our small data set, may have overfitted or not generalized well. Moreover, each ResNet block might not be as optimal as EfficientNet for small fine lesion boundary representations.

Overall, EfficientNet-b0 serves as a better encoder for this task than ResNet-50 when trained with the same protocol. U-Net outperformed DeepLabV3+ slightly across our experiments on ISIC-2020, potentially due to the skip connections that retain finer boundary detail important for segmentation of lesions. But the gaps between the best setups are small, and both architectures deliver solid performance. The qualitative examples in Fig. Figure 5 reinforce that models exhibit capability of modelling lesion shape and borders on challenging lesions, with irregular edges and colour, and present artifacts such as hair. Clinically, lesion boundaries legibly impact the diagnostic value of shape and border characteristics of lesions, as size and shape are important diagnostic cues.

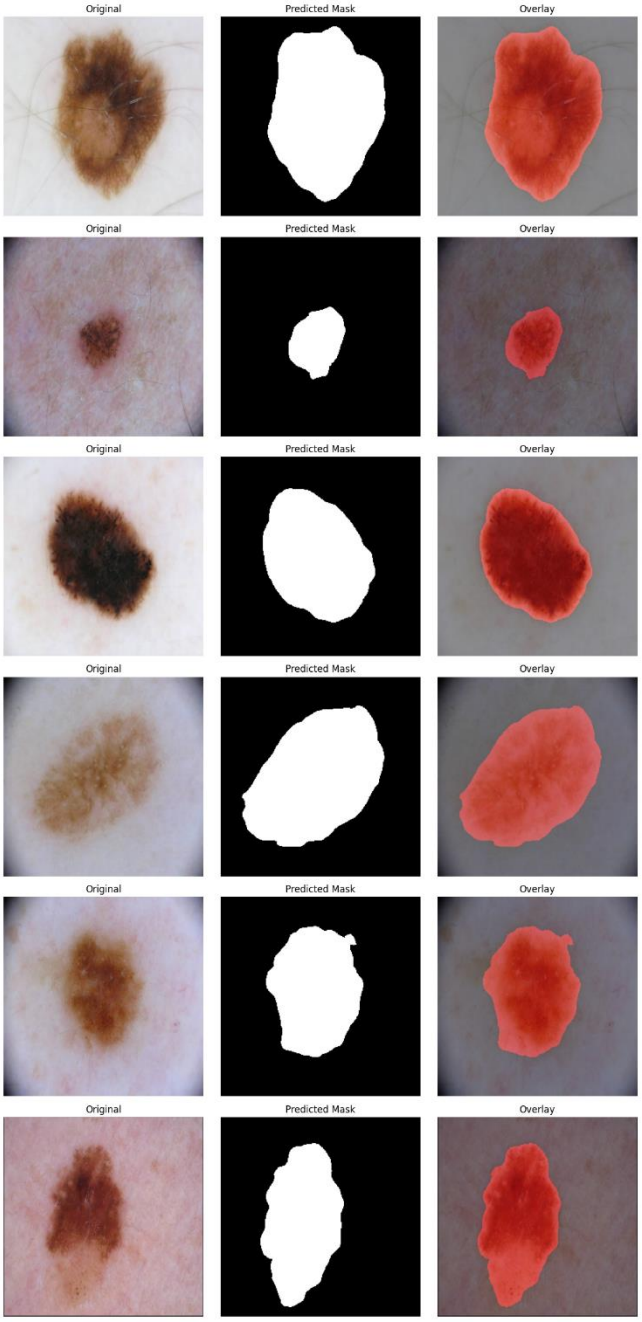

Fig. 5. Experimental outcome

While some small differences may appear in how well the borders are captured, the general performance suggests that deep learning is a powerful tool for supporting clinicians in identifying and monitoring skin lesions.

VII. CONCLUSION

This paper compared two state of the art deep - learning segmentation architectures (u-net and deeplabv3+) as a controlled comparison of skin lesion segmentation on isic-2020. The results presented in the paper were obtained under a single preprocessing and training protocol, and indicated that across both architectures, efficientnet-b0 provided consistently higher dice and iou vs resnet-50, with the best performing configuration being u-net with: efficientnet-b0 (dice = 0.920,

IoU = 0.859) Qualitative examples additionally indicated the models can robustly recapture the shape of the lesion, while boundary ambiguity and hair artifacts were the primary sources of error.

Future work, it can be evaluate stability across multiple random seeds and report statistical significance, investigate boundary-aware losses and artifact removal to enhance fine margins.

ACKNOWLEDGEMENTS

The authors would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq), Baghdad - Iraq, for supporting this work.

REFERENCES

[1] World Health Organization,(2024). "Cancer fact sheet," WHO. https://www.who.int/news-room/fact-sheets/detail/cancer

[2] National Cancer Institute – SEER,(2024). "Melanoma of the Skin — 5-Year Relative Survival by Stage at Diagnosis," SEER. https://seer.cancer.gov

[3] G. Argenziano and H.P. Soyer,(2001). "Dermoscopy of pigmented skin lesions—a valuable tool for early diagnosis of melanoma," Lancet Oncology, 2(7), 443–449, doi: https://doi.org/10.1016/S1470-2045(00)00422-8

[4] P. Carli, V. De Giorgi, L. Naldi, G. Dosi, and F. Giannotti,(1998). "Reliability and inter-observer agreement of dermoscopy in the diagnosis of melanocytic skin lesions," European Journal of Cancer Prevention, 7(4), 397–402.

[5] Z. Mirikharaji, K. Abhishek, A. Bissoto, C. Barata, S. Avila, E. Valle, M.E. Celebi, and G. Hamarneh,(2023). "A survey on deep learning for skin lesion segmentation," Medical Image Analysis, 88, 102863, doi: https://doi.org/10.1016/j.media.2023.102863

[6] H.K. Jeong, C. Park, R. Henao, and M. Kheterpal,(2022). "Deep Learning in Dermatology: A Systematic Review of Current Approaches, Outcomes, and Limitations," JID Innovations, 3(1), 100150, doi: https://doi.org/10.1016/j.xjidi.2022.100150

[7] A. Mahbod, P. Tschandl, G. Langs, R. Ecker, and I. Ellinger,(2020). "The effects of skin lesion segmentation on the performance of dermatoscopic image classification," Computer Methods and Programs in Biomedicine, 197, 105725, doi: https://doi.org/10.1016/j.cmpb.2020.105725

[8] Z. Chen, J. Chen, J. Zhao, et al.,(2023). "What Matters in Radiological Image Segmentation? Effect of Segmentation Errors on the Diagnostic Related Features," Journal of Digital Imaging, 36(5), 2088–2099, doi: https://doi.org/10.1007/s10278-023-00865-2

[9] A.L. Sharma, K. Sharma, and P. Ghosal,(2025). "Skin lesion segmentation: A systematic review of computational techniques, tools, and future directions," Computers in Biology and Medicine, 196(Pt C), 110842, doi: https://doi.org/10.1016/j.compbiomed.2025.110842

[10] M.E. Celebi, H. Iyatomi, G. Schaefer, and W.V. Stoecker,(2009). "Lesion border detection in dermoscopy images," Computerized Medical Imaging and Graphics, 33(2), 148–153, doi: https://doi.org/10.1016/j.compmedimag.2008.11.002

[11] O. Ronneberger, P. Fischer, and T. Brox,(2015). "U-Net: Convolutional networks for biomedical image segmentation," MICCAI 2015, LNCS 9351, 234–241, doi: https://doi.org/10.1007/978-3-319-24574-4_28

[12] R. Ramadan and S. Aly,(2022). "CU-Net: A New Improved Multi-Input Color U-Net Model for Skin Lesion Semantic Segmentation," IEEE Access, 10, 15539–15564, doi: https://doi.org/10.1109/ACCESS.2022.3148402

[13] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam,(2018). "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation (DeepLabv3+)," ECCV 2018, 833–851, doi: https://doi.org/10.1007/978-3-030-01234-2_49

[14] M. Zafar, M.U. Akram, and M.R. Khan,(2023). "DeepLabv3+-Based Segmentation and Best Features Selection for Skin Cancer Diagnosis," Mathematics, 11(2), 364, doi: https://doi.org/10.3390/math11020364

[15] V. Rotemberg, N. Kurtansky, B. Betz-Stablein, et al.,(2021). "A patient-centric dataset of images and metadata for identifying melanomas using clinical context," Scientific Data, 8, 34, doi: https://doi.org/10.1038/s41597-021-00815-z

[16] Z. Liu, J. Hu, X. Gong, and F. Li,(2025). "Skin lesion segmentation with a multiscale input fusion U-Net incorporating Res2-SE and pyramid dilated convolution," Scientific Reports, 15, 7975, doi: https://doi.org/10.1038/s41598-025-92447-1

[17] L. Cai, K. Hou, and S. Zhou,(2024). "Intelligent skin lesion segmentation using deformable attention Transformer U-Net with bidirectional attention mechanism in skin cancer images," Skin Research and Technology, 30(8), e13783, doi: https://doi.org/10.1111/srt.13783

[18] .B. Hu, X. Li, Y. Wang, and Q. Zhang,(2024). "LeaNet: Lightweight U-shaped architecture for high-performance yet efficient skin lesion segmentation," Computers in Biology and Medicine.

[19] J. Xu and L. Tong,(2024). "LB-UNet: A Lightweight Boundary-Assisted UNet for Skin Lesion Segmentation," MICCAI 2024, LNCS 15009.

[20] M. Arshad, M.A. Khan, J.M. Górriz, J. Baili, D.A. AlHammadi, C. Kim, and Y. Nam,(2025). "Skin cancer segmentation and recognition from dermoscopy images: a novel framework based on improved DeepLabV3+ and network-level fused deep architectures," Journal of Advanced Research, Online ahead of print, doi: https://doi.org/10.1016/j.jare.2025.08.039

[21] A. Jimi, N. Zrira, O. Guendoul, and H. Belkacemi,(2025). "ESC-UNET: A hybrid CNN and Swin Transformers for skin lesion segmentation," Intelligence-Based Medicine, 12, 100257, doi: https://doi.org/10.1016/j.ibmed.2025.100257

[22] T. Nie, Y. Zhao, and S. Yao,(2024). "ELA-Net: An Efficient Lightweight Attention Network for Skin Lesion Segmentation," Sensors (Basel), 24(13), 4302, doi: https://doi.org/10.3390/s24134302

[23] A. Dzieniszewska, P. Garbat, and R. Piramidowicz,(2024). "Improving Skin Lesion Segmentation with Self-Training," Cancers (Basel), 16(6), 1120, doi: https://doi.org/10.3390/cancers16061120

[24] D. Bardou, A. Chaki, and K. Dutta,(2022). "Hair artifact removal in dermoscopy using variational autoencoders," Journal of Imaging.

[25] K. Nawaz, A. Zanib, I. Shabir, J. Li, Y. Wang, T. Mahmood, and A. Rehman,(2025). "Skin cancer detection using dermoscopic images with convolutional neural network," Scientific Reports, 15, 7252, doi: https://doi.org/10.1038/s41598-025-91446-6

[26] S. Gül, G. Cetinel, B.M. Aydin, D. Akgün, and R. Öztaş Kara,(2025). "YOLOSAMIC: A Hybrid Approach to Skin Cancer Segmentation with the Segment Anything Model and YOLOv8," Diagnostics (Basel), 15(4), 479, doi: https://doi.org/10.3390/diagnostics15040479

[27] .K. Manzoor, N.U. Gilal, M. Agus, and J. Schneider,(2025). "Dual-stage segmentation and classification framework for skin lesion analysis using deep neural network," Digital Health, 11, eCollection 2025, doi: https://doi.org/10.1177/20552076251351858

[28] N.R. Kurtansky, V. Rotemberg, and H.P. Soyer,(2024). "SLICE-3D: Skin lesion crops from 3D total-body photography," Scientific Data.

[29] L. Wang, Q. Zhang, and H. Liu,(2024). "Performance of U-Net variants in dermoscopy segmentation," BMC Medical Imaging, 24, 112, doi: https://doi.org/10.1186/s12880-024-01012-1

[30] R. Zhao, Y. Wang, and Q. Li,(2024). "Refining lesion boundaries with DeepLabV3+," Computers in Biology and Medicine, 170, 105728, doi: https://doi.org/10.1016/j.compbiomed.2024.105728

[31] N. Codella, V. Rotemberg, and D. Gutman,(2024). "Diversity of ISIC dermoscopy datasets," Scientific Data, 11, 101, doi: https://doi.org/10.1038/s41597-024-01412-9

[32] The International Skin Imaging Collaboration (ISIC),(2018). "ISIC 2018 Challenge – Task 1: Lesion Boundary Segmentation," ISIC Challenge. https://challenge.isic-archive.com

[33] A. Yang, Z. Zeng, S.Y. Yeo, C. Tan, H.L. Tey, and Y. Su,(2017). "A Novel Multi-task Deep Learning Model for Skin Lesion Segmentation and Classification," arXiv preprint, arXiv:1703.01025. https://arxiv.org/abs/1703.01025.

[34] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, (2017). "Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, *Lecture Notes in Computer Science*, 10553, pp. 240–248, doi: https://doi.org/10.1007/978-3-319-67558-9_28.