

Comprehensive Review on U-Net Architectures for Skin Lesion Segmentation and Its Variants

Nimra Tanveer¹, Nimra Tariq^{1*}, Arooj Akram¹, and Faiza Shabbir¹

¹Department of Basic Sciences, The Superior University Lahore, Pakistan.

*Corresponding Author: Nimra Tariq. Email: nimra.tariq@superior.edu.pk

Received: June 19, 2024 Accepted: August 19, 2024 Published: September 01, 2024

Abstract: Skin lesion, ranging from benign growths to malignant tumors, like melanoma, pose significant diagnostic challenges because of their different morphological features. Early and accurate segmentation of these lesions from medical images is critical for effective diagnosis and treatment. Traditional image processing approaches, such as, thresholding and edge recognition, often fail to capture the complexity and variability of skin lesions. In contrast, deep learning methods, particularly Convolutional Neural Networks (CNNs), have modernized this field by providing robust solutions. The encoder-decoder architecture of the U-Net, for example design and skip connections, has showed significant consequence in describing lesion limits accurately. This review systematically examines the current state of skin lesions segmentation using u-net and its variants. It also highlights the performance of different models using metrics such as Dice coefficient and Jaccard index, addressing challenges such as the need for extensive annotated datasets. Our findings suggest that deep learning models like U-net significantly enhance the segmentation of skin lesions, improving diagnostic accuracy and clinical consequences.

Keywords: Medical Image Segmentation; Skin Lesions; Convolutional Neural Networks (CNN); U-Net.

1. Introduction

Skin lesions encompass a wide range of dermatological variations including harmless turn of events, pre-cancerous conditions, and harmful developments, for example, melanoma [1]. A significance of these lesions is remarkable medical test in light of their different morphological appearances, which require exact demonstrative and supportive involvements. Nevertheless, exact and irritate location representation of skin damages are vital in dermatology, as they can meaningfully affect results for patients Melanoma, for example, is mostly strong sort of skin sickness, at whatever point recognized early, can be successfully treated, which increments patience rates [2]. Then again, suspended finding usually brings about more terrible meditations, including the basic requirement for exact and reliable diagnostics devices. Next, exact skin lesions division from medical images is critical to the methodology for diagnosing all around, division frames the boundaries of the injury, segregating it from surrounding sound tissues [3].

Subsequently it allows inside and out assessment of the shape, size, variety, and surface of the sore, this step is essential. Which are key limits in the gap examination of skin conditions. Thusly, exact division makes it more straightforward to follow and assess the abilities of sores. Sore changes long term and guides supportive conclusions [4]. Moreover, its improvements the limit of robotized expressive systems to perform constantly, thusly supporting clinicians in settling on informed choices and improving the results of patient reflection. Machine learning, then again, has arisen as a remarkable system in medical picture division, which offer basic benefits over standard events [5]. Policies for predictable image handling like thresholding, edge recognition, and area developing oftentimes experience troubles with difficulty and unpredictability of skin flaws.

Deep learning models, in connection, and machine learning algorithms in general, excel at handling challenges because of their capability to gain and sum up from massive datasets. Utilizing annotated medical images, machine learning models can be prepared to observe complex samples and elements segmentation, but it also speeds up annual explanation by saving time and energy. Diagnostic technique [6] The primary goal of this review paper is to offer a detailed examination of the present status of clinical picture division for skin damages using AI methods for example U-Net. Moreover, the review will investigate late developments and rising trends in the field to bring consideration to innovative approaches and the potential significances they could have for clinical practice. In this paper, we explore advancements in U-Net architecture and their effectiveness in skin lesion segmentation. Our contribution includes a detailed comparison of traditional image processing techniques with deep learning-based methods, emphasizing the superiority of U-Net and its variants. Additionally, we investigate the challenges associated with segmentation accuracy, the need for large annotated datasets, and the role of performance metrics like Dice coefficient and Jaccard index in evaluating the efficacy of these models. Finally, we offer insights into future research directions aimed at enhancing segmentation performance, addressing existing limitations, and improving clinical applicability.

2. Overview of Skin Lesions

Skin lesions address a great range of dermatological situations and to figure out the differences it is essential for accurate diagnosis and treatment to distinguish between benign and malignant lesions. Healthy skin injuries like moles (nevi), seborrheic keratosis, and lipomas are normally non-carcinogenic developments that do not spread to other body parts. These lesions normally have uniform borders and regular colour and size that breaks the same over time. Uncooperatively, basal cell carcinoma, melanoma, and other malignant skin lesions carcinoma and squamous cell carcinoma are types of cancer that can spread to other parts of the body or metastasize too far off organs. Similarly, dangerous sores usually show random lines, frequent colours, as well as a tendency to alter in size, form, or colour. Early uncovering and therefore exact separation among meaningless and dangerous injuries are essential in light of the fact that hostile sores can put your life in danger [7].

However, there are significant problems in unravelling skin lesions from medical images because of the inherent inconsistency in how they appear. Like, skin sores can vary knowingly in shape, going from completely indirect to exceptionally random structures. Along these lines, size variety is another confusing factor, with sores going from a couple of millimetres to a few centimetres in width. Equally, variety fickleness inside and between sores further convolutes division events. Because Lesions can appear in a variety of colours, including white, brown, black, red, and blue [8]. It is challenging to apply uniform segmentation algorithms to a variety of lesions due to all of these factor's types, which requires advanced image handling techniques and AI models able to adjust to a variety of appearances.

Moreover, the presence of leftovers like hair, shadows, and reflections can inhibit exact segmentation. Due to the fact that shadows and reflections brought about by uneven lightning can rotation the injury's actual limits and tones [9]. The other thing is that the surrounding skin's changeability in skin-color and surface can incomprehensible the qualification between solid the lesion and the skin. This is particularly difficult when the lesion's edges are undefined. Or on the other hand mixtures into the surrounding tissue. To deal with this gathering of difficulties and issues, advanced methods, for example, convolutional brain organizations (U-Net variations), has shown guarantee in remaining these difficulties to recognize and separate complex examples in skin aching images [10]. This method force makes skin lesion segmentation more accurate and reliable, which helps in skin cancer classification and early detection. However, the training and development of such models call for a lot of data and a lot of computing power.

3. Image Segmentation Techniques for Skin Lesions

Medical image analysis depends on heavily on image segmentation, particularly when diagnosing and treating skin lesions. In this field, traditional image processing methods have been life-threatening. These approaches often include the use of thresholding, edge recognition, area-based division, and bunching calculations [11]. For example, thresholding is a direct methodology where pixel power is applied to make a paired picture, what separates the sore from the surrounding skin [12]. However, it be

likely to be exceptionally gentle to lighting conditions and changes in painful tone. Moreover, edge processing techniques, for example, the Vigilant or Sobel administrators decide limits in view of slopes in pixel power, which helps in portraying sore lines [13]. Despite this, blaring images and irregular lesion boundaries can make these methods difficult to use.

Image segmentation for skin lesions has been transformed in recent years by machine learning techniques, which offer significant advantages over traditional approaches. AI, mostly thoughtful learning measures like Convolutional Neural Networks (CNNs), can logically improvement and concentrate random highlights from pictures, prompting more exact and powerful division results [14]. CNNs are mainly useful for segmenting lesions of complex and irregular shapes because of their ability to capture spatial hierarchies in images thanks to their hierarchical structure [15]. Likewise, they can gain from huge explained datasets, which contributions in summing up with rising to new and covered information. This is very different from traditional methods, which trust a lot on features and heuristics that are made by hand and can be limited and less flexible to different lesion appearances.

Moreover, the capacity of machine learning models to include background information from the adjacent tissues improves their volume to separate between areas that appear to be similar. By proficiently combining high-resolution three-dimensional information with background features through skip connections, advanced architectures like U-Net and its variations have proved outstanding performance in medical image segmentation [16]. So, these models not only provide precise boundary delineation but also handle variations in size, shape, and texture of skin lesion more effectively than traditional methods. In addition to that, machine learning approaches can be continually improved with more data and enhanced algorithms, which makes them a dynamic and evolving solution for skin lesion segmentation [17].

Table 1. Comparison of Traditional Segmenting Techniques

Techniques	Description	Advantages	Disadvantages
Thresholding	Converts an image to a binary image based on pixel intensity values.	Simple and computationally efficient.	Sensitive to lighting conditions and variations in lesion color
Edge Detection	Identifies boundaries in an image by detecting changes in intensity.	Effective at highlighting edges and boundaries.	Can struggle with noisy images and irregular lesion boundaries.
Region Growing	Starts from seed points and expands based on predefined criteria such as intense similarity.	Provide accurate segmentation for well-defined regions.	Requires careful parameter selection and can be computationally expensive.
Clustering (e.g., K-means)	Group pixels with similar characteristics into clusters.	handle multi-model intensity distributions	May not effectively capture complex lesion structures.
Gaussian Mixture Model	Use probabilistic models to represent the distribution of pixel intensities.	Model complex distribution better than simple clustering.	Computationally intensive and sensitive to initial parameters.

4. Machine Learning for Skin Lesion Segmentation

Machine learning algorithms have profoundly impacted the field of medical imaging, particularly in the segmentation of skin lesions [18]. As mentioned earlier, traditional methods often relied on manual delineation by dermatologists, which is time-consuming, subjective, and prone to inter-observer variability. So, machine learning offers automated, objective, and reproducible solutions that significantly enhance the efficiency and accuracy of skin lesion segmentation. Among the machine learning techniques, convolutional neural networks (CNNs) have emerged as the most prominent for skin lesion segmentation due to their exceptional ability to capture spatial hierarchies in images [19]. CNNs, with architectures like U-Net and its variants, leverage encoder-decoder structures that efficiently process and reconstruct high-resolution segmentation maps. The encoder part extracts features at various levels of abstraction, while the decoder part reconstructs the segmentation map to ensure precise localization of lesion boundaries. These networks are typically trained on large datasets such as ISIC (International Skin Imaging Collaboration) archive, which provides diverse and annotated images of skin lesions [20]. As a result, this enables the models to generalize well across different types of lesions and imaging conditions. Other than CNNs, there are thoughtful learning models, for example, fully convolutional networks (FCNs) and generative adversarial networks (GANs). Similar to CNNs, FCNs make use of fully convolutional layers that are capable of producing segmentation maps that are the same size as the images that they take in to guarantee pixel-wise classification precision [21]. On the other hand, GANs are made up of a generator that creates segmentation maps as well as a discriminator to determine how accurate they are. In addition to outdated machine learning techniques like support vector machines (SVMs), such as deep learning, skin lesion segmentation has also utilized random forests and k-means clustering, with opposing levels of achievement [22]. Handmade features like consistency are characteristically required in these methods. Descriptors of color and shape to differentiate lesions from healthy skin. Although practical in some due to their dependence on features, these methods are classically less healthy and flexible in situations. Designing and limited ability to demonstrate complex examples in high-layered information. Nevertheless, critical developments, challenges stay in the space of skin sore division. For example, variation in sore appearance remembering for variability, surface, and shape as well as artifacts like hair, shadows, and reflection present significant flat areas [23]. In addition, the requirement for huge, explained datasets for preparing thoughtful learning models features the significance of cooperative happenings in information sharing also, comment. In addition, it is essential to guarantee that these models can be applied to a wide range of populations. What's more, imaging conditions for their certain clinical reaction.

5. Application of U-Net for Skin Lesions Segmentation

The use of U-Net for skin sore division addresses a significant jump advancing in the field of medical image investigation, especially in dermatology [24]. Convolutional neural network U-Net network (CNN) engineering scheduled clearly for biomedical picture partition, which succeeds in tasks demanding exact boundary description and localization. This capability is essential in dermoscopic images to differentiate lesions from healthy skin, a vital step in the early finding and finding of skin diseases including melanoma. This is present in the U-Net architecture. Unique structure with a symmetric expanding path and a tightening path to capture context for detailed positioning. As an encoder, the contracting path makes use of repeated utilizations of convolutions and max pooling to energetically lessen spatial features while holding significant level highlights. Crucially, this path merges the up-sampled features with corresponding high-resolution features from the contracting path through skip connections. The main purpose is to ensure that fine details are preserved and the segmentation map is accurate. In the context of skin lesion segmentation, U-Net's architecture proves particularly effective due to its ability to manage the diverse shapes, sizes, and colours of skin lesions. Like, U-Net's capability to integrate contextual information from different scales enables it to differentiate between lesions and normal skin more effectively. In addition to that, the skip connections play a vital role in retaining spatial information lost during down sampling, which is crucial to accurately delineate the often-irregular shapes of skin lesions [25].

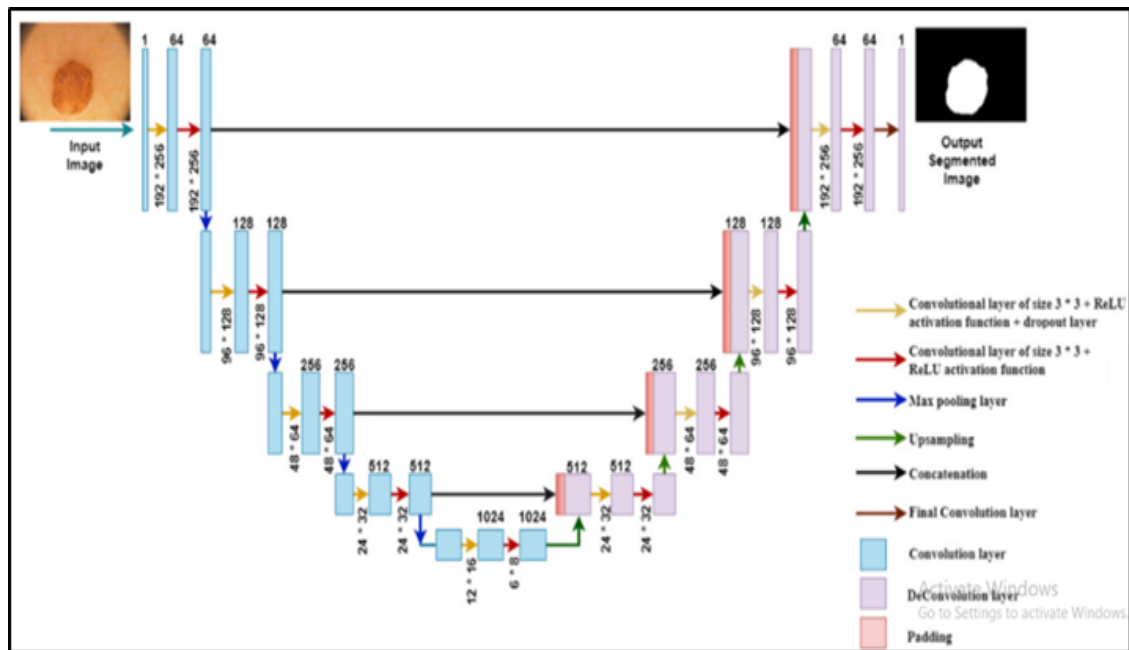


Figure 1. U-Net Model in Skin lesion Segmentation [25]

However, training a U-Net model for skin lesion segmentation utilizes a large dataset of annotated dermoscopic images, where each pixel is labelled as either lesion or non-lesion. Basically, the model is trained to minimize a loss function, adjusted to handle the class imbalance which is typically present in medical image datasets. Also, researchers have consistently demonstrated that U-Net based models achieve high performance in segmenting dermoscopic images. For instance, the research showed that U-Net models, when trained on ISIC challenge datasets can achieve state-of-art performance to highlight their effectiveness and robustness in clinical scenarios. In this way, the clinical implications of U-Net is skin lesion segmentation extend beyond accurate segmentation. Because automated segmentation tools like U-Net can reduce diagnostic workload and inter-observer variability, leading to a more consistent and reliable assessments.

6. Discussion and Results

To evaluate the image segmentation for skin lesion, different metrics assess the accuracy and effectiveness of the segmentation algorithms to ensure that they meet the clinical standards necessary for practical application. There are many commonly used evaluation metrics such as the Dice Coefficient, Jaccard Index, Sensitivity, and Specificity, each of which provides unique insights into performance of the models. For instance, the study of R. Kaur et al. (2024) explored various U-Net architectures for skin lesion segmentation. The study compares the original U-Net with two modified versions: FCN-1, which uses PreLU (Parametric Rectified Linear Unit) activation and transpose convolution for up-sampling and FCN-2, which employs PreLU and bilinear interpolation for up-sampling. Based on the outputs, it was observed that models using transpose convolution (as in FCN-1) exhibited more false negative and slight under-fitting challenges in learning the features accurately. Conversely, the interpolation method with PreLU resulted in the detection of more outliers, which was mitigated through incorporating a dropout technique. When the validation accuracies are compared over training epochs, it was evident that transpose convolution-based models (original U-Net and FCN-1) achieved peak accuracy faster. In contrast, interpolation-based methods (FCN-2 and the proposed method) showed a slower learning curve due to dropout effect and the absence of learnable weights in the interpolation process [26].

Table 2. Test accuracy of various U-Net models [26]

Model	Pixel Accuracy	Dice	IoU
U-Net	91.12%	78.23%	39.26%
FCN-1	90.47%	81.85%	40.51%

FCN-2	91.34%	83.26%	41.84%
Proposed Method	94.36%	88.33%	44.05%

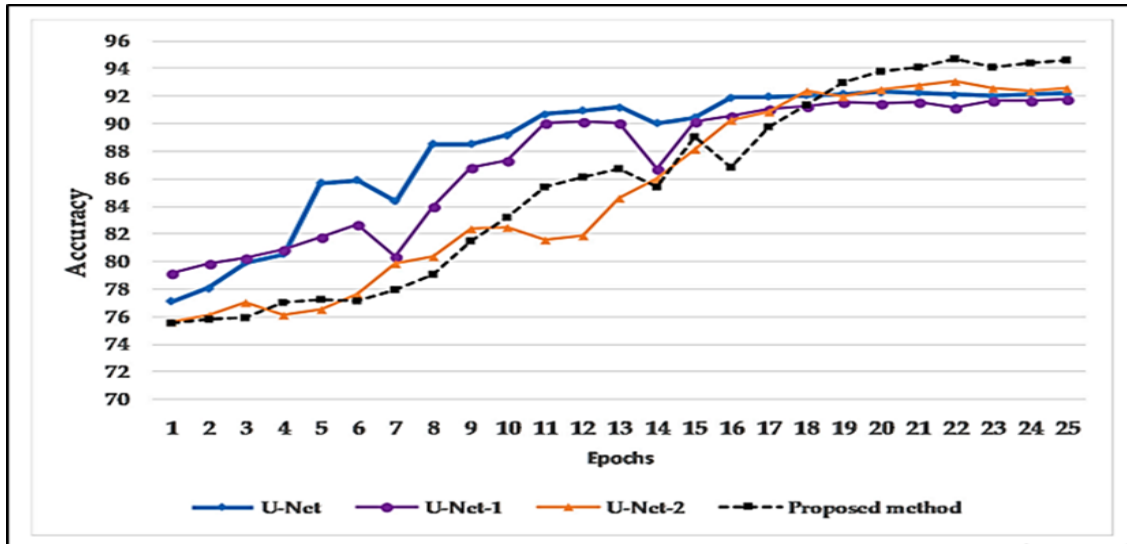


Figure 2. Comparing Learning Curves per epoch [27]

The paper of Iranpoor et al. (2020) investigates the performance of convolutional neural networks (CNNs) on a specific dataset consisting of 200 images e.g. 160 are used for training and 40 for testing. The study uses the Dice Coefficient to measure classification accuracy. The Dice Coefficient assesses the overlap between predicted and target masks through comparing the number of common pixels to the total pixels across both masks. As far as parallel examination, the examination presents different organization models including the first U-Net, RelayNet, and proposed system involving a pre-prepared 34-layer ResNet in the encoding segment. The outcomes show that the first U-Net accomplishes 84% exactness, RelayNet accomplishes 81%, and the proposed ResNet-based strategy arrives at 89% precision. Overall, the results show that including a ResNet encoder significantly improves classification accuracy, proving this method's suitability for image segmentation tasks [28].

Table 3 . AVG. Accuracy of Network Output in 50 Runs [28]

Architecture		Error rate (%)	Accuracy of test data (%)	Error rate (%)	Accuracy of train data (%)
Other Methods	U-Net	0.26	0.84	0.26	0.97
	U-Net without pooling	0.28	0.80	0.35	0.96
	RelayNet	0.30	0.81	0.31	0.96
Proposed Method	U-Net Modified by ResNet in decoder and convolutional layer	0.25	0.89	0.34	0.92

Another paper of Vesal (2018) presents an innovative approach to address a key limitation of the conventional U-Net architecture used for skin lesion segmentation. The traditional U-Net, despite its efficacy, has a drawback at its lowest level due to a small receptive field. This restricted receptive field limits the network's ability to capture non-local image information, which is crucial for accurate segmentation of skin lesions. To overcome this limitation, Dilated Convolutions incorporate an additional parameter called the dilation rate to determine the spacing between values in the convolutional kernels. This adjustment enables the network to maintain computational efficiency while significantly enhance the ability to process and integrate non-local features. For instance, the novel architecture of SkinNet showed superior performance in skin lesion segmentation. By effectively incorporating local and global information through dilated convolutions and dense connections, SkinNet achieved higher Jaccard Index (JI) and Dice Coefficient (DC) scores of 76.67% and 85.10% respectively, along with the high sensitivity of 93% [29].

Table 4. Lesion segmentation performances of different frameworks [29]

Method	AC	DC	JI	SE	SP
Yading Yuan	0.934	0.849	0.765	0.825	0.975
Math Berseth	0.932	0.847	0.762	0.820	0.978
Auto-ED	0.936	0.842	0.738	0.836	0.966
LIN	0.952	0.753	0.753	0.855	0.974
SkinNet	0.932	0.767	0.767	0.930	0.905

7. Conclusion

To conclude, the integration of deep learning architectures, particularly U-Net, into the medical image segmentation holds significant promise for early detection and treatment of skin lesions. Traditional image segmenting methods, while foundational, often falter in the face of the inherent variability and complexity of skin lesions. So, deep learning models, with their capacity to learn and generalize from extensive annotated datasets, offer a transformative solution. Through leveraging these models, clinicians can achieve more precise and reliable segmentation, which is crucial for the early detection and treatment of malignant lesions such as melanoma. However, the advancements in U-Net architectures have further refined segmentation performance, which address challenges such as varying lesion shapes, sizes, and colours, as well as the present of artifacts. As the field continues to evolve, further improvements in U-net based architectures and the availability of large annotated datasets will be crucial in ensuring that these technologies can be widely adopted in clinical settings ultimately improving patient outcomes.

References

1. C. F. Heal, B. A. Raasch, P. Buettner, and D. Weedon, "Accuracy of clinical diagnosis of skin lesions," *British Journal of Dermatology*, vol. 159, no. 3, pp. 661-668, 2008.
2. B. Ahmed, M. I. Qadir, and S. Ghafoor, "Malignant melanoma: skin cancer– diagnosis, prevention, and treatment," *Critical Reviews™ in Eukaryotic Gene Expression*, vol. 30, no. 4, 2020.
3. A. A. Adegun, S. Viriri, and M. H. Yousaf, "A probabilistic-based deep learning model for skin lesion segmentation," *Applied Sciences*, vol. 11, no. 7, p. 3025, 2021.
4. D. A. Okuboyejo and O. O. Olugbara, "A review of prevalent methods for automatic skin lesion diagnosis," *The Open Dermatology Journal*, vol. 12, no. 1, 2018.
5. X. Liu, L. Song, S. Liu, and Y. Zhang, "A review of deep-learning-based medical image segmentation methods," *Sustainability*, vol. 13, no. 3, p. 1224, 2021.
6. M. M. Taye, "Understanding of machine learning with deep learning: architectures, workflow, applications and future directions," *Computers*, vol. 12, no. 5, p. 91, 2023.
7. F. Navarro, M. Escudero-Vinolo, and J. Bescós, "Accurate segmentation and registration of skin lesion images to evaluate lesion change," *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 501-508, 2018.
8. K. Korotkov and R. Garcia, "Computerized analysis of pigmented skin lesions: a review," *Artificial intelligence in medicine*, vol. 56, no. 2, pp. 69-90, 2012.
9. R. B. Oliveira, E. Mercedes Filho, Z. Ma, J. P. Papa, A. S. Pereira, and J. M. R. Tavares, "Computational methods for the image segmentation of pigmented skin lesions: a review," *Computer methods and programs in biomedicine*, vol. 131, pp. 127-141, 2016.
10. J. Zhang, F. Zhong, K. He, M. Ji, S. Li, and C. Li, "Recent Advancements and Perspectives in the Diagnosis of Skin Diseases Using Machine Learning and Deep Learning: A Review," *Diagnostics*, vol. 13, no. 23, p. 3506, 2023.
11. A. Bali and S. N. Singh, "A review on the strategies and techniques of image segmentation," in *2015 Fifth International Conference on Advanced Computing & Communication Technologies, 2015: IEEE*, pp. 113-120.
12. M. Zortea, E. Flores, and J. Scharcanski, "A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images," *Pattern Recognition*, vol. 64, pp. 92-104, 2017.
13. M. E. Celebi, Q. Wen, H. Iyatomi, K. Shimizu, H. Zhou, and G. Schaefer, "A state-of-the-art survey on lesion border detection in dermoscopy images," *Dermoscopy image analysis*, vol. 10, no. 1, pp. 97-129, 2015.
14. M. Jogin, M. Madhulika, G. Divya, R. Meghana, and S. Apoorva, "Feature extraction using convolution neural networks (CNN) and deep learning," in *2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT), 2018: IEEE*, pp. 2319-2323.
15. S. Maqsood and R. Damaševičius, "Multiclass skin lesion localization and classification using deep learning based features fusion and selection framework for smart healthcare," *Neural networks*, vol. 160, pp. 238-258, 2023.
16. K. Sanjar, O. Bekhzod, J. Kim, J. Kim, A. Paul, and J. Kim, "Improved U-Net: Fully convolutional network model for skin-lesion segmentation," *Applied Sciences*, vol. 10, no. 10, p. 3658, 2020.
17. K. M. Hosny, D. Elshora, E. R. Mohamed, E. Vrochidou, and G. A. Papakostas, "Deep learning and optimization-based methods for skin lesions segmentation: a review," *IEEE Access*, 2023.
18. R. Baig, M. Bibi, A. Hamid, S. Kausar, and S. Khalid, "Deep learning approaches towards skin lesion segmentation and classification from dermoscopic images-a review," *Current medical imaging*, vol. 16, no. 5, pp. 513-533, 2020.
19. R. Arora, B. Raman, K. Nayyar, and R. Awasthi, "Automated skin lesion segmentation using attention-based deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 65, p. 102358, 2021.
20. N. C. Codella 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: IEEE*, pp. 168-172.

21. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431-3440.
22. T. G. Debelee, "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review," *Diagnostics*, vol. 13, no. 19, p. 3147, 2023.
23. K. A. A. Eltayef, "Segmentation and lesion detection in dermoscopic images," Brunel University London, 2017.
24. M. D. Alahmadi, "Multiscale attention U-Net for skin lesion segmentation," *IEEE Access*, vol. 10, pp. 59145-59154, 2022.
25. Z. Mirikharaji et al., "A survey on deep learning for skin lesion segmentation," *Medical Image Analysis*, p. 102863, 2023.
26. B. S. Lin, K. Michael, S. Kalra, and H. R. Tizhoosh, "Skin lesion segmentation: U-nets versus clustering," in 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017: IEEE, pp. 1-7.
27. R. Kaur and S. Kaur, "Automatic skin lesion segmentation using attention residual U-Net with improved encoder-decoder architecture," *Multimedia Tools and Applications*, pp. 1-27, 2024.
28. R. Iranpoor, A. S. Mahboob, S. Shahbandegan, and N. Baniyadi, "Skin lesion segmentation using convolutional neural networks with improved U-Net architecture," in 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), 2020: IEEE, pp. 1-5.
29. S. Vesal, N. Ravikumar, and A. Maier, "SkinNet: A deep learning framework for skin lesion segmentation," in 2018 IEEE nuclear science symposium and medical imaging conference proceedings (NSS/MIC), 2018: IEEE, pp. 1-3.