



REGULAR ARTICLE

Segmentation of Skin Lesion Using Double U-Net Framework for Enhanced Feature Extraction

Pujari Madhuri* , Kunchala Supriya, Kodipyaka Sai Ganesh, B. Lakshmi Prasanna† 

Department of Computer Science and Information Technology, Institute of Aeronautical Engineering, Hyderabad, India

(Received 18 August 2025; revised manuscript received 25 October 2025; published online 30 October 2025)

Skin cancer is a frequent cancer globally, and the outlook for a patient and the effectiveness of treatment depend on its prompt identification. Dermoscopic images are very essential in the accurate and automatic segmentation of skin lesions for helping clinicians diagnose skin cancer. We, in this study, propose a new semantic segmentation model based on the DoubleU-Net architecture for improving the detail and accuracy of skin lesion detection. The proposed DoubleU-Net model works by integrating two U-Net networks in sequence, where the first U-Net extracts high-level features and provides an initial segmentation map. The second U-Net refines this output by learning from the residual errors of the first network and produces a more detailed and accurate segmentation. This dual network design helps in overcoming the challenges of blurred lesion boundaries and varying lesion sizes, which are common issues in skin lesion segmentation. We evaluated the performance of our model using the publicly available ISIC(2018) dataset which contains thousands of annotated dermoscopic images. Our model evolved with the Dice coefficient and losing cross-entropy in order to deal with class imbalance, which is frequently observed in medical datasets. Experimental results show that our proposed DoubleU-Net architecture performs more effectively than baseline U-Net model when using the metrics Intersection over Union (0.81589), Dice coefficient (0.88628), and overall segmentation accuracy (0.94437).

Keywords: Skin lesion segmentation, Double U-Net architecture, Semantic segmentation, Dermoscopic images.

DOI: [10.21272/jnep.17\(5\).05034](https://doi.org/10.21272/jnep.17(5).05034)

PACS numbers: 07.05.Tp, 87.19.xj

1. INTRODUCTION

Skin cancer is one of the main concerns in global health, as it leads to the majority of cases of all cancers around the world. Of all the types of skin cancer, melanoma is the worst, for its chances of metastasizing are very high at the time of diagnosis when it has been left undiagnosed. Chances of survival increase dramatically if there is an early diagnosis and treatment of skin cancer, especially melanoma. However, diagnosis of skin cancer is quite complicated as it involves highly specialized effort in distinguishing between benign and malignant lesions. Thermoscopic images that provide magnified views of the skin lesion are quite frequently used by dermatologists in making studies of the structure and patterns inside a lesion. Nazarov Jasurbek et al. [1] proposed a Genetic Algorithm-based approach for selecting the optimal image channels for enhancing skin lesion segmentation. Alafer et al. [2] introduced a model called L-UNet, which is one of the efficient deep learning-based approaches for high-quality image segmentation tailored for the various medical applications. L-UNet Architecture: The architecture is optimised to obtain the finest results in terms of efficiency and accuracy of segmentation within a computational model. A model is suitable for applications in different clinical

settings with limited resources. Tahir Hussain et al. [3] designed a MAGRes-UNet model, which is the refined version of U-Net for multi-class tumour and skin lesion segmentation using accurate MRI. The proposed model deals with various common problems that occur during the process of medical image segmentation like gradient vanishing issues. Yagmur Olmez et al. [4] proposed a developed algorithm based on the concept of PSO and focuses to upgrade the accuracy concerning skin cancer images for segmenting it. The enhanced PSO includes a visit table to minimize redundant search paths and utilizes multi-directional search strategy that enhances diversity, hence the segmentation process. Even with such advantages of the standard U-Net, it encounters problems not being able to detect well thin boundaries of lesions or even fail handling images possessing extreme variations of lesion features. To cope with all these, we develop a novel model with an architecture as Double U-Net that increases the quality of segmentation map further by incorporating the refinement with the second U-Net.

2. LITERATURE SURVEY

Zhou et al. (2019) suggest a modifying form the called UNet++. With regard to good structures for

* Correspondence e-mail: madhuripujari12@gmail.com

† lakshmiprasanna447@gmail.com



planting plans, the U-Net approach is interesting because it helps to enhance good model recognizes images containing complicated medical data through segmentation and localization of data over multiple scales [5]. GU et al. (2020) present CA-Net, a convolutional neural network with thick attention mechanisms. The model target clearness in medical image segmentation by showing important regions in images working attention-based layers, rendering it of particular use in sensitive statuses [6]. Jha et al. (2021) brings DoubleU-Net, extend depth and feature extraction capabilities by stacking two U-Nets on over one another. forward segmentation accuracy is the aim of this design, particularly on rough medical images where little details are required [7]. Valanarasu et al. (2021) develop KiUNet, an overcomplete convolutional architecture. To perform high-resolution biomedical image segmentation, KiUNet strongly captures volumetric and fine features, reaching the complex details of biomedical images with overcomplete layers that refine feature retention [8]. Taghanaki et al. (2021) present a review of deep semantic segmentation techniques for both natural and medical images. They check at different medical image segmentation architectures and challenges, exhibit the demand of value and workable models for a limit of imaging contexts [9].

3. EXISTING METHOD

The existing systems of skin lesion segmentation are dominated by the use of architectures of U-Net. U-Net relies on the encoder-decoder structure with skip connections. However, there are some drawbacks using this architecture. For example, U-Net actually loses fine details in some cases such as images that contain subtle lesions with diverse sizes or unclear boundaries. Other traditional methods for skin lesion segmentation include thresholding, and clustering techniques, but these approaches tend to perform poorly on complex images and are highly sensitive to variations in illumination and noise.

4. PROPOSED METHOD

To overcome these problems with the current systems, we introduce a novel semantic segmentation model using a Double U-Net for the detection of skin lesions. This model has two U-Net networks stacked sequentially with the first U-Net generating an initial segmentation map and the second U-Net refining the output with learning from the errors in the first network. The system can capture more complex features and produce more accurate segmentation with the dual network design. Combining the two models are compared with strengths of two U-Net models, this proposed system can easily handle variations in lesion size, shape and texture along with more accurate boundary detection of the lesions. Furthermore, the proposed system also uses the data augmentation techniques to resolve the class imbalance problem ensuring that the model learnt is robust and generalizable to new data. Such an innovative architecture brings forth a more precise segmentation for a more decisive clinical diagnosis and proper treatment planning.

The section shows the general structure of the Double U-Net model, as depicted in Fig.1, and goes into details about each of the blocks.

4.1 Input Image

The process starts with the raw input image, which goes through the first U-Net to extract preliminary feature maps. These feature maps are later passed to the second U-Net to refine and improve the segmentation outcome.

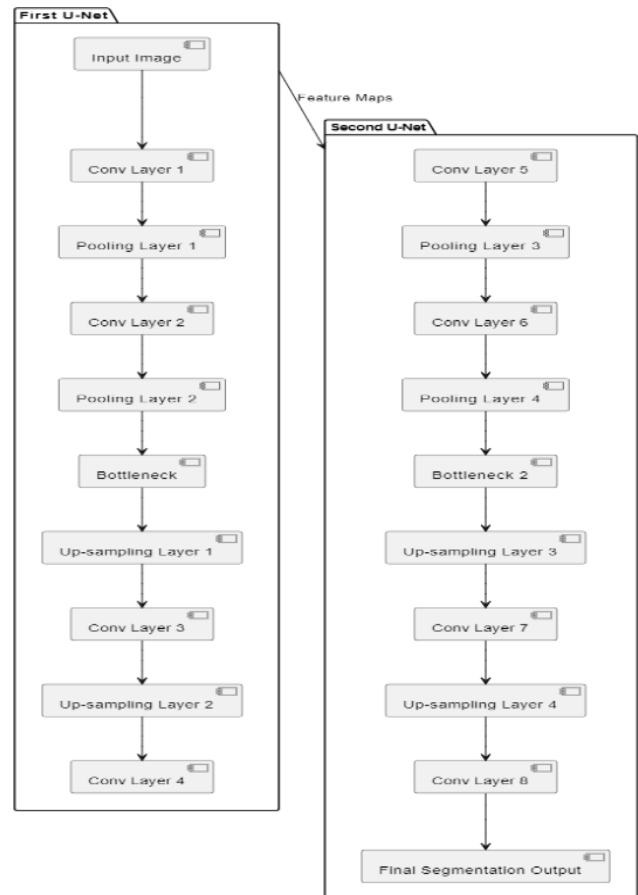


Fig. 1 – Proposed Double U-Net Architecture

4.2 Conv Layer

Convolutional layers in both U-Nets use filters to identify image features. Convolution Layers 1 and 2 in the First U-Net capture fundamental patterns and edges, while Convolution Layers 5 and 7 in the Second U-Net enhance these patterns.

4.3 Pooling Layer

The network can learn a hierarchical structure of the input image thanks to pooling, which gradually compresses the feature maps in both U-Nets.

4.4 Bottleneck

As a bridge between the encoding and decoding track, the bottleneck layer minimize extracted lesion point into a minim format while reconstructing important segmentation information. The bottleneck in

the first U-Net settle features for a segmentation, while Bottleneck 2 in the second U-Net further compresses and refines the data for a more precise result.

4.5 Up-sampling Layer

Up-sampling layers essentially reverse the pooling operations by restoring the spatial dimensions of feature maps. Within each U-Net, up-sampling layers take the feature mappings through the decoder route to regain the initial input size (that is, to 'upsample').

4.6 Final Conv Layer

Each U-Net merites its final convolution layers clear the up-sampled feature maps, producing in an accurate segmentation mask. Conv Layers 3 and 4 in the First U-Net gives an initial segmentation output, while Conv Layer 8 in the Second U-Net creates the final adjustments.

4.7 Final Segmentation Output

The final segmentation map is output of the second U-Net's linking of all the learnt features from the two U-Nets. By combining initial and refined feature extraction, this two-stage design successfully the target regions and includes an accurate segmentation of the input image.

5. RESULTS AND EVALUATION METRICS

IoU, DSC, Pre, Accuracy, Sensitivity, and Specificity were the six parameters that were used in the trials to calculate the segmentation performance of different models. Intersection over Union (IoU), the Jaccard index is one of the most essential metrics for semantic segmentation. The intersection area and union of the ground truth and expected masks are closely computed. In mathematics, IoU is defined as follows in equation (1):

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (1)$$

True positives (TP) are the number of pixels correctly categorizing the object. False negatives (FN) are the number of pixels incorrectly classifying the object. False positives (FP) are the number of pixels wrongly classifying the object.

Common metric used to root out image segmentation is the Dice Similarity Coefficient (DSC). Below is the equation (2) defines DSC in mathematics:

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (2)$$

Acc is expressed as the number of pixels correctly classified divided by the total number of pixels. It indicates measure of performance in all parts of the images. Regarding Acc in Math it is defined by the equation (3) as follows.

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

where True Negatives (TN) stands for pixels that have

been precisely found to be outside of the object. Sensitivity (Sen), regularly referred to as recall, evaluates how well the model can segment lesion pixels. The equation (4) is definition of Sen in mathematics is as follows:

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Specificity (Spe) quantifies the accuracy of the model in identifying non-lesion pixels correctly. The equation (5) is definition of Spe in mathematics is as follows:

$$\text{Sen} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

The percentage of true positive predictions among all predicted positives is called as Precision (Pre), and the equation (6) is definition of Pre in mathematics is as follows:

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

Insights into the segmentation model's performance are provided by each of these metrics, that support identifying how well it separates lesion zones from the background and reduces false positives and false negatives.

Table 1 – ISIC2018 segmentation metrics evaluation comparison between DoubleU-Net,U-Net model

Model	IoU	DSC	Acc	Sen	Spec	Pre
U-Net++	0.8158	0.8862	0.9443	0.9311	0.9554	0.8751
U-Net	0.8067	0.8801	0.9412	0.9421	0.9473	0.8566

The performance metrics shown in Table 1 is that DoubleU-Net outperforms U-Net in IoU (0.8158 vs. 0.8067), DSC (0.8862 vs. 0.8801), and accuracy (0.9443 vs. 0.9412), indicating stronger segmentation accuracy and overlap with the ground truth masks. While U-Net has a slightly higher sensitivity (0.9421 vs. 0.9311), Double U-Net's advantages in specificity (0.9554 vs. 0.9473) and precision (0.8751 vs. 0.8566) make it the preferred model, as it balances accurate segmentation with fewer false positives, leading to more effective and consistent segmentation performance.

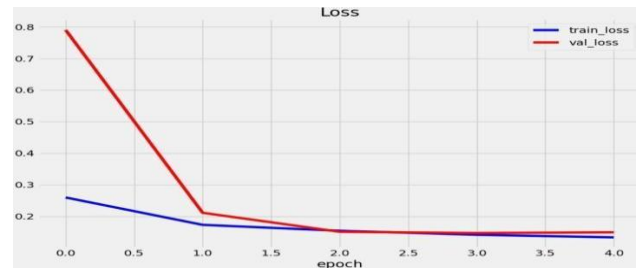


Fig. 2 – Training and validation loss curves

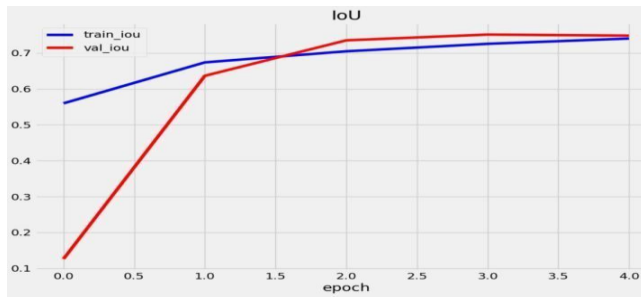


Fig. 3 – Training and validation IoU curves



Fig. 4 – Training and validation DSC curves

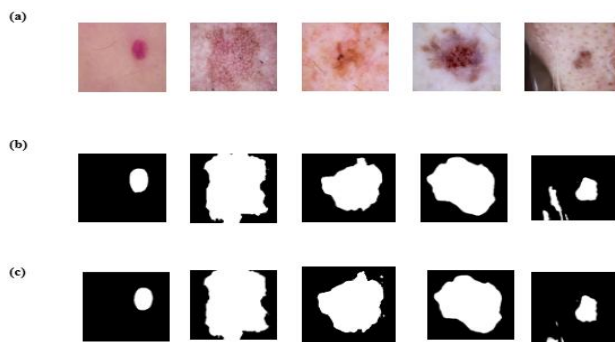


Fig. 5 – The ISIC dataset was segmented as follows: (a) original image; (b) U-Net results; and (c) Double U-Net result

The graphs shown in Fig. 2, Fig. 3, Fig. 4 is the training and validation performance of a segmentation model over four epochs, focusing on loss, IoU, and Dice Coefficient metrics.

The Fig. 5 presents segmentation results on the ISIC dataset for skin lesions, where performance of both U-Net and Double U-Net models were compared. The original images of skin lesions can be seen in row (a), and row (b) showed segmentation masks predicted by the U-Net model, whereas row (c) illustrates masks obtained by the Double U-Net model. Visually, Double U-Net (c) provides more accurate segmentation contours and captures finer details of the lesions, suggesting a better performance as compared with the U-Net model (b), showing less regular boundaries and missing region. This comparison illustrates how the Double U-Net is better suited for lesion area delineation in a skin image.

6. CONCLUSION

This project demonstrates deep learning's possibility applied in medical image segmentation from a new method, being developed to be utilized in skin lesion detection using an architecture of Double U-Net. The model achieves better segmentation accuracy for skin lesions, including difficult situations with uneven boundaries and small lesions, by stacking two U-Net networks to capture both fine and coarse data. By offering more accurate lesion segmentation, the model continuously beats conventional techniques, proving the value of using deep learning architectures for sophisticated medical picture analysis.

REFERENCES

1. N. Jasurbek, D.F. Ivan, O.S. Ajani, R. Mallipeddi, *IEEE Access* **12**, 135692 (2024).
2. F. Alafer et al., *IEEE Access* **12**, 140769 (2024).
3. T. Hussain, H. Shouno, *IEEE Access* **12**, 40290 (2024).
4. Y. Olmez, G.O. Koca, A. Sengür, U.R. Acharya, H. Mir, *IEEE Access* **12**, 840 (2024).
5. Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, *IEEE Trans. Med. Imag.* **39** No 6, 1856 (2019).
6. R. Gu, G. Wang, T. Song, R. Huang, M. Aertsen, J. Deprest, *IEEE Trans. Med. Imag.* **40** No 2, 699 (2020).
7. Debesh Jha, Michael A. Riegler, Dag Johansen, *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* (2021).
8. J.M.J. Valanarasu, V.A. Sindagi, I. Hacıhaliloglu, V.M. Patel, *IEEE Trans. Med. Imag.* **41** No 4, 965 (2021).
9. S. Asgari Taghanaki, K. Abhishek, J. Paul Cohen, J. Cohen-Adad, G. Hamarneh, *Artif. Intell. Rev.* **54**, 137 (2021).

Сегментація уражень шкіри за допомогою Double U-Net Framework для покращеного вилучення ознак

Pujari Madhuri, Kunchala Supriya, Kodipyaka Sai Ganesh, B. Lakshmi Prasanna

Department of Computer Science and Information Technology, Institute of Aeronautical Engineering, Hyderabad, India

Рак шкіри є поширеним видом раку в усьому світі, і прогноз для пацієнта та ефективність лікування залежать від його швидкої ідентифікації. Дерматоскопічні зображення дуже важливі для точної та автоматичної сегментації уражень шкіри, що допомагає клініцистам діагностувати рак шкіри. У цьому дослідженні ми пропонуємо нову модель семантичної сегментації, засновану на архітектурі DoubleU-Net, для покращення деталізації та точності виявлення уражень шкіри. Запропонована модель DoubleU-Net працює шляхом послідовної інтеграції двох мереж U-Net, де перша U-Net витягує високорівневі ознаки та надає початкову карту сегментації. Друга U-Net уточнює цей результат, намагаючись на залишкових помилках першої мережі, та створює більш детальну та точну сегментацію. Така конструкція подвійної мережі допомагає подолати проблеми розмитих меж уражень та різних розмірів уражень, що є поширеними проблемами при сегментації уражень шкіри. Ми оцінили продуктивність нашої моделі, використовуючи загальнодоступний набір даних ISIC(2018), який містить тисячі анотованих дерматоскопічних зображень. Наша модель розвивалася з урахуванням коефіцієнта Dice та втрати перехресної ентропії, щоб впоратися з дисбалансом класів, який часто спостерігається в медичних наборах даних. Експериментальні результати показують, що запропонована нами архітектура DoubleU-Net працює ефективніше, ніж базова модель U-Net, при використанні метрик Intersection over Union (0,81589), коефіцієнта Dice (0,88628) та загальної точності сегментації (0,94437).

Ключові слова: Сегментація уражень шкіри, Архітектура Double U-Net, Семантична сегментація, Дерматоскопічні зображення.