PROCEEDINGS OF
INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS AND NEW APPLICATIONS
https://proceedings.icisna.org/
2nd International Conference on Intelligent Systems and New Applications (ICISNA'24), Liverpool, April 26-28, 2024.

Performance Evaluation of U-Net Based Methods for Lesion Segmentation from Dermoscopy Images

Musa Doğan[1], Ilker Ali Ozkan[1]

[1]*Computer Engineering Department, Selcuk University, Konya, Türkiye*
musa.dogan@selcuk.edu.tr, ORCID: 0000-0001-8757-2869

[1]*Computer Engineering Department, Selcuk University, Konya, Türkiye*
ilkerozkan@selcuk.edu.tr, ORCID: 0000-0002-5715-1040

***Abstract*— Skin cancer accounts for approximately half of all cancer cases worldwide, making it one of the most prevalent types of cancer. Melanoma, which develops from melanocytes that give skin its color, is the most lethal among skin cancers. Early diagnosis of melanoma, particularly through dermoscopy images, is of vital importance. To this end, automated diagnostic systems significantly aid dermatologists in their decision-making processes. In recent years, advancements in deep learning and machine learning have improved diagnostic accuracy. Specifically, CNN-based deep learning algorithms are utilized for medical image analysis and skin lesion segmentation. While traditional methods struggle to capture fine details and broader context, the U-Net architecture overcomes these challenges, providing more accurate segmentation. This study evaluates U-Net, Residual U-Net, and Attention U-Net models for skin lesion segmentation. The performance of the models is measured using Dice Score, Jaccard Index, and train loss metrics. The results reveal that Attention U-Net demonstrates the highest performance, with a Dice Score of 0.8063 and a Jaccard Index of 0.7203.**

***Keywords*— residual, attention gate, semantic segmentation, deep learning**

I. Introduction

Skin cancer is one of the most frequent types of cancer and accounts for about half of all cancer diagnoses worldwide [1, 2]. Melanoma, a type of skin cancer that can appear anywhere in the skin, is the most lethal skin cancer. Melanoma develops from melanocytes, which are pigment-containing cells that give skin color [3].

The early detection of melanoma from dermoscopy images is critical, beacuse it can respond to treatment in its early stages [4]. Therefore, automated diagnosis of melanoma aids dermatologists in the decision-making process [5]. Visual examination of skin lesions during medical examinations can be challenging because of the similarities between lesions and normal tissues, leading to misdiagnoses. Dermoscopy has emerged as a significant technique for dermatologists in monitoring melanoma lesions in the last decade. [6]. Recent advancements in deep learning and machine learning have helped reduce the margin of error in diagnosis. Deep learning algorithms, particularly CNNs, are used to analyze medical images and for the classification and segmentation of the relevant area. Traditional image segmentation methods often struggle to capture both fine details (localization) and broader context in the segmentation of skin diseases. U-Net addresses this challenge by employing an intelligent encoder-decoder architecture. With U-Net, images are processed without being converted into vectors, which could potentially lead to information loss, thereby preserving both critical details and the overall context, contributing to more accurate skin disease segmentation. Numerous studies have been conducted on the segmentation of lesions. Examples of these studies include the following. Goyal et al. [7] presented a fully automated deep learning ensemble method that combines Mask R-CNN and DeeplabV3+ for accurate skin lesion boundary segmentation. Trained on the ISIC-2017 dataset, this method achieved a sensitivity of 89.93% and a specificity of 97.94%. Ramadan and Aly [8] introduced three novel U-Net variations tailored for skin lesion segmentation, encompassing single, dual, and triple input models. Each model utilizes different color spaces to enhance segmentation performance. A channel-wise attention module integrates features from multiple encoder sub-networks, and a composite loss function further improves the results. Evaluations of the ISIC 2017, ISIC 2018, and PH2 datasets demonstrated that these models provide robust segmentation despite color variations in dermoscopy images. Yuan, et al. [9] introduced a fully automated skin lesion segmentation method using an end-to-end trained 19-layer deep convolutional neural network. To address issues of limited training data and imbalanced pixel distribution, the researchers implemented effective training strategies and designed an innovative Jaccard distance-based loss function. The method, evaluated on the ISBI 2016 and PH2 datasets, highlighted its potential for broader medical image segmentation applications. This study evaluates the performance of a CNN-based approach, U-Net, and its modified architectures in the task of skin lesion segmentation.

II. Material and Methods

This study mainly focuses on lesion segmentation from dermoscopy images. In this section, a brief overview of the dataset used, the models employed for segmentation, and the performance metrics is provided.

A. Dataset

In this study, the ISIC 2018 [10] dataset was used for lesion segmentation. This dataset comprises an extensive compilation of dermoscopy images published by the International Skin Imaging Collaboration (ISIC). It contains 2594 RGB dermoscopic images in jpg format, each with a resolution ranging from 556 x 679 pixels to 4499 x 6748 pixels. The ISIC-2018 skin lesion competition dataset was classified according to parameters defined by the ISIC. This dataset includes images of both benign and malignant oncological conditions, with images evenly distributed among subsets, slightly dominated by moles and melanomas. Figure 1 shows an original image and its corresponding lesion segmentation mask.

Fig. 1 Sample original image and segmentation mask in ISIC-2018 [10]

B. U-Net-based Methods for Lesion Segmentation

In 2015, Ronneberger, et al. [11] introduced one of the most popular FCN-based approaches for semantic medical image segmentation, known as U-Net. U-Net employs a fully convolutional network to perform the task of semantic segmentation. It can generally be considered as an encoder-decoder architecture. The first layer, referred to as the encoder or downsampling layer, functions as a "contracting path" that helps retain contextual information within feature maps. The second layer, referred to as the decoder or upsampling layer, functions as an "expanding path." This layer gradually restores the details of the object and increases the spatial dimensions. Additionally, there are skip connections between the encoder and decoder, which are used to transfer contextual information from the encoder to the decoder. Figure 2 provides a diagram illustrating the U-Net architecture.

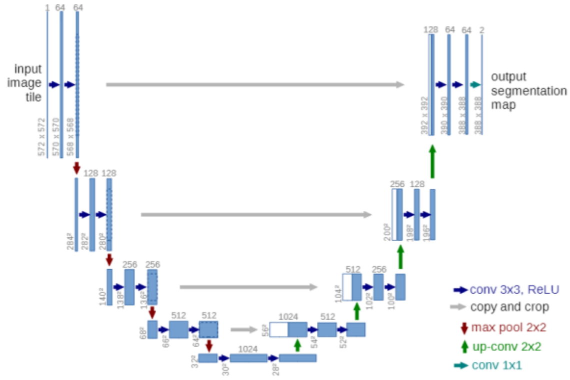

Fig. 2 Sample original image and segmentation mask in ISIC-2018

The U-Net architecture has been continuously improved through various studies. Oktay, et al. [12] introduced the use of an attention unit to focus on specific important objects while disregarding irrelevant areas. This approach allows different parts of the network to focus on segmenting different objects in the U-Net architecture. The attention gate implements a function that assign weights to the feature map based on each class. He, et al. [13] introduced Residual Neural Networks (ResNets) to address the vanishing gradients problem by incorporating skip connections. This helps the model learn faster and more accurately. In their study, Zhang, et al. [14] presented the Residual U-Net, incorporating residual blocks in the encoder and decoder parts. This improved the model while simultaneously addressing the gradient vanishing problem. The model was tested on a public road dataset and compared with both the U-Net and other deep learning-based road extraction techniques in the literature.

C. Performance Metrics

The performance of the models for lesion segmentation from dermoscopy images was evaluated using the Dice Score (DS) and Jaccard Index (JI). The DS represents the overlap between the predicted image (SM) and the ground truth (GT), as shown in Equation 1. The JI measures the ratio of the intersection over the union of the objects by dividing their overlap by their combined area, as detailed in Equation 2.

$$\mathrm{DS} = \frac{2 \times |SM \cap GT|}{|SM| + |GT|} \tag{1}$$

$$\mathrm{JI} = \frac{|SM \cap GT|}{|SM \cup GT|} \tag{2}$$

The values provided in the experimental results of the study represent the mean values of Jaccard Index and Dice Coefficient metrics.

III. Results

In this study, three different models (U-Net, Residual-U-Net, and Attention-U-Net) were evaluated for the task of skin lesion segmentation. The performance of the models was measured using the Dice Score (DS), Jaccard Index (JI), and train loss (TL) metrics. The experimental study was conducted using PyTorch (version 1.12.1) and trained on an 48 GB GDDR6 NVIDIA Quadro RTX 8000 GPU, an Intel Xeon Gold 6226R CPU @ 2.90GHz with 12 processors, and 64 GB of memory. During the training phase, the batch size was set to 4. The Dice Coefficient was used as the loss function, and the models were trained for 50 epochs. The dataset was split into 90% for training and 10% for validation. The models were trained with the same parameters for the lesion segmentation task, and the results are presented in Table 1.

TABLE I
BEST RESULTS OF MODELS ON LESION SEGMENTATION TASK

Models	**DS**	**JI**	**TL**
U-Net	0.7070	0.6088	0.1190
Residual-U-Net	0.8004	0.7079	0.6770
Attention-U-Net	0.8063	0.7203	0.2439

IV. CONCLUSIONS

In this work, we evaluated the performance of U-Net and U-Net based methods for lesion segmentation from dermoscopy images using ISIC-2018 dataset. The results indicate that the developed Residual U-Net and Attention U-Net models achieved comparable results similar to the basic U-Net model. The Attention U-Net achieved the highest Dice Score (0.8063) and the highest Jaccard Index (0.7203). Additionally, the Attention U-Net performed a faster learning process with a lower training loss (0.2439). As a result, we argue that more sophisticated models including attention mechanisms outperform simpler models in lesion segmentation tasks. These models successfully capture both local and global information, which improves segmentation performance. This research highlights that utilizing the Attention-U-Net model for lesion segmentation from dermoscopy images can yield more successful results. Future studies could further improve model architectures and validate findings across additional datasets.

ACKNOWLEDGMENT

We would like to thank the Scientific Research Coordinatorship of Selcuk University for their support with the project titled “Data-Intensive and Computer Vision Research Laboratory Infrastructure Project” numbered 20301027.

REFERENCES

[1] T.-C. Pham, C.-M. Luong, M. Visani, and V.-D. Hoang, "Deep CNN and data augmentation for skin lesion classification," in *Intelligent Information and Database Systems: 10th Asian Conference, ACIIDS 2018, Dong Hoi City, Vietnam, March 19-21, 2018, Proceedings, Part II 10*, 2018: Springer, pp. 573-582.

[2] M. Azeem, K. Kiani, T. Mansouri, and N. Topping, "SkinLesNet: Classification of Skin Lesions and Detection of Melanoma Cancer Using a Novel Multi-Layer Deep Convolutional Neural Network," *Cancers,* vol. 16, no. 1, p. 108, 2023.

[3] Z. A. Nazi and T. A. Abir, "Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with u-net and dcnn-svm," in *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018*, 2020: Springer, pp. 371-381.

[4] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors,* vol. 18, no. 2, p. 556, 2018.

[5] N. H. Quang, "Automatic skin lesion analysis towards melanoma detection," in *2017 21st Asia Pacific symposium on intelligent and evolutionary systems (IES)*, 2017: IEEE, pp. 106-111.

[6] H. Sharen, M. Jawahar, L. J. Anbarasi, V. Ravi, N. S. Alghamdi, and W. Suliman, "FDUM-Net: An enhanced FPN and U-Net architecture for skin lesion segmentation," *Biomedical Signal Processing and Control,* vol. 91, p. 106037, 2024.

[7] M. Goyal, A. Oakley, P. Bansal, D. Dancey, and M. H. Yap, "Skin lesion segmentation in dermoscopic images with ensemble deep learning methods," *Ieee Access,* vol. 8, pp. 4171-4181, 2019.

[8] R. Ramadan and S. Aly, "CU-net: a new improved multi-input color U-net model for skin lesion semantic segmentation," *IEEE Access,* vol. 10, pp. 15539-15564, 2022.

[9] Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE transactions on medical imaging,* vol. 36, no. 9, pp. 1876-1886, 2017.

[10] N. Codella *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.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, 2015: Springer, pp. 234-241.

[12] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999,* 2018.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[14] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters,* vol. 15, no. 5, pp. 749-753, 2018.