

U-Net based Network Applied to Skin Lesion Segmentation: An Ablation Study

Graziela Silva Araujo, Guillermo Camara-Chavez

Federal University of Ouro Preto, Computer Science Department

Ouro Preto, Brazil, 35.400-000

graziela.araujo@aluno.ufop.edu.br, guillermo@ufop.edu.br

and

Roberta B. Oliveira

University of Brasilia, Computer Science Department

Brasilia, Brazil, 70.910-900

roberta.oliveira@unb.br

Abstract

Skin cancer is one of the types of cancer that requires an early diagnosis. The segmentation task plays a vital role in computer-aided diagnosis. Segmenting dermoscopic images is challenging for existing methods due to different image conditions. There is a significant variation in color, texture, shape, size, and location in dermoscopic images. Still, they may contain images with lighting variation and various artifacts, such as hair, ruler, air/oil bubbles, and color sample. The Convolutional Neural Network (CNN) model, U-Net, is widely used to segment dermoscopic images. This work proposes a model based on the U-Net architecture to segment dermoscopic images. Still, it presents an ablation study to justify the modifications made in the architecture, such as the number of training epochs, image size, optimization functions, dropout, and the number of convolutional blocks. Experiments were carried out on the ISIC 2017 and ISIC 2018 datasets and show that it is possible to arrive at a simple model capable of presenting competitive results compared to other state-of-the-art works with the appropriate adjustments to their parameters.

Keywords: Convolutional Neural Network, U-Net, Image Segmentation, Skin lesion, Melanoma.

1 Introduction

Skin cancer is among the 19th most common cancers worldwide, with melanoma and non-melanoma being the main types [1]. Nonmelanoma is the most common type of skin cancer, but it is not very aggressive. Basal cell carcinoma and squamous cell carcinoma are the most frequent cancers. Melanoma is a more aggressive type of skin cancer, and for that reason, it has a high mortality rate. The fact that melanoma can spread to other parts of the body (metastasis) further increases the need for an early diagnosis [2]. Its identification at an early stage allows for greater effectiveness in the treatment.

Some clinical methods are used internationally to assist in the diagnosis of melanomas, such as the *ABCDE* [3, 4] rule, and the 7-point list [5]. The *ABCDE* rule is based on according to the following criteria: (1) the “A” criterion is related to the **asymmetry** of the lesion. The lesion is divided into two parts based on axis of symmetry. If these parts are not symmetrical, the lesion may be considered suspicious; (2) the “B” criterion concerns the irregularity of the **border** of the lesion. A benign lesion has a normally regular border. Therefore, a lesion that contains irregular or ill border defined may be considered suspicious; (3) for the “C” criterion, the number of apparent **colors** in the lesion is significant. A wide range of colors in the lesion, such as red, white, black, and shades of blue-gray, can make the lesion suspicious; (4) in the “D” criterion, the **diameter** of the lesion is observed. Lesions with a diameter greater than or equal to 6 millimeters are considered suspicious, and (5) the “E” criterion handles **evolution** of the lesion, i.e., a lesion should be investigated if any modification to its structure is noted [6]. Regarding the 7-point list,

there are two categories: (1) major criteria that are the presence of atypical pigmentation, bluish-white veil, and non-typical vascularization pattern; and (2) minor criteria that are the irregular streaks, irregular pigmentation, irregular spots/globules, and regression structures. The major criteria may indicate a higher probability of cancer, therefore, the presence of these criteria receives more points. However, these clinical methods are applied manually, being a time-consuming process, it may lead to variable results, even being performed by an experienced dermatologist [7].

Due to the inefficiency of manual clinical inspections of dermoscopic images, the existence of an automated system for a faster and more accurate diagnosis of skin lesions is essential. An automated method can help the dermatologist in decision making and provide the patient with a more effective treatment [8]. Using computers to recognize melanoma is still a challenge. There is significant variation between melanoma classes in terms of color, texture, shape, size, and location in dermoscopic images and significant similarity between melanomas and non-melanomas. The acquired images sometimes have low contrast and illumination variation and the presence of some artifacts such as rulers, hairs, veins, and air/oil bubbles [9], which can make automated diagnosis difficult, as can be seen in Figure 1.

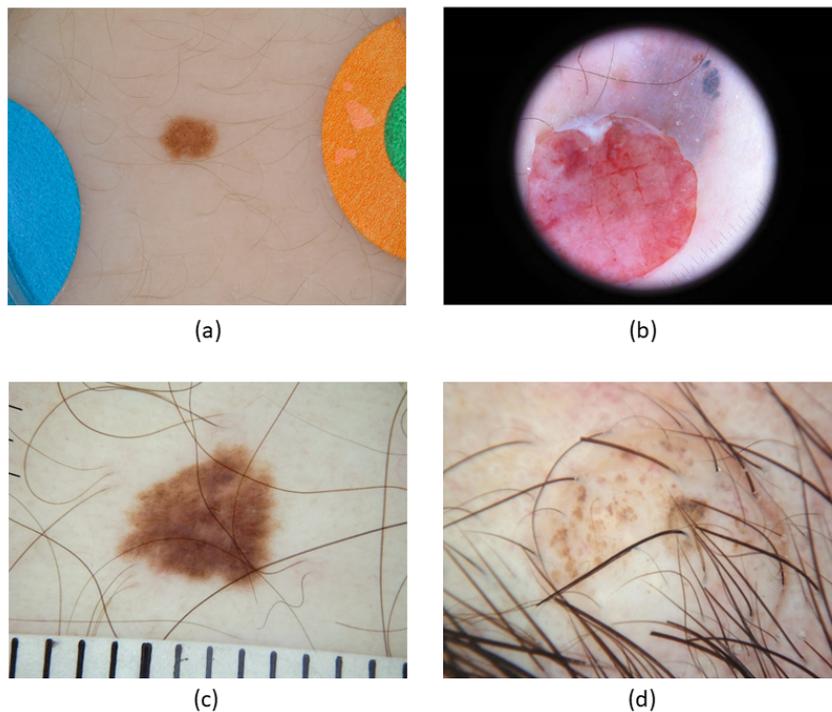


Figure 1: Examples of dermoscopic images of skin lesions with artifacts: (a) clinical color samples, (b) lighting variation, (c) clinical scale marker, and (d) skin with hair [10]

For a diagnosis using a Computer-Aided Diagnostic (CAD) system [11], the “B” criterion of the ABCDE rule, which checks for irregularities at the edge of the lesion, can be analyzed according to the result of the image segmentation process. The segmentation process separates the image into parts to identify regions of interest and, from these regions, obtain relevant information. The automatic diagnosis methods proposed in recent years have shown better results to solve skin lesion segmentation problems use deep learning techniques [8]. However, a disadvantage in using such techniques is related to training, as they require many images, and acquiring dermoscopic images and labeling them by a specialist is not a simple task. Furthermore, Graphics Processing Units (GPU) are needed to decrease the processing time. For this reason, the International Skin Imaging Collaboration (ISIC) [12] provides a database containing dermoscopic images for the skin lesion segmentation task.

In a previous version of this paper [13], we proposed a method based on a U-Net with modifications in its original architecture, such as increasing the number of network layers and adjustments in the optimization and activation function. This version extends the discussion by making an ablation study of the number of training epochs, image size, optimization functions, dropout, and the number of convolutional blocks. In the case of the image size, we explore the advantages or disadvantages of working with high-resolution images. We also evaluate the values and places of the dropout in the different convolutional blocks.

Section 2 presents a literature review of methods used to perform image segmentation. Traditional

methods and machine learning methods are discussed. Section 3 details the U-Net model covering the proposed adjustments. It also discusses the steps used to perform the dermoscopic image segmentation task. Section 4 discusses the results of the proposed model. An ablation study is carried out to verify the model's effectiveness concerning its parameters. Section 5 presents a brief conclusion of the studies carried out.

2 Related Works

Image segmentation is an essential step in image skin lesion analysis, as it separates the diseased area from the healthy region. For a diagnosis using a CAD system, the segmentation task is fundamental. There are several computational methods in the literature to perform the segmentation task. These methods can be classified as conventional methods and methods based on deep learning [14].

A conventional method like K-means was applied by Alvarez *et al.* [15], Agarwal *et al.* [16] and Garg *et al.* [17]. In Agarwal *et al.* [16], the authors proposed an approach divided into three steps: (1) pre-processing; (2) segmentation, in which the K-means method is applied using two clusters, one to represent the lesion and another to represent the background skin (normal skin). The generated binary image results from the grouping submitted to an intensity-based threshold; and (3) post-processing, using the K-means algorithm output. Bibi *et al.* [18] also used conventional methods to perform the segmentation. A contrast enhancement in the image is applied, followed by a transformation to Hue, Saturation, Value (HSV) on the color channels. First, the method selects the best channel to generate a lesion map. Finally, the map is converted to binary.

Methods based on deep learning have demonstrated greater precision in the segmentation task in the literature. Chen *et al.* [19] proposed a multitasking Convolutional Neural Network (CNN) architecture, including FeatureNet, SegNet, and ClsNet architectures, to perform feature extraction, segmentation, and classification, respectively. A feature passing module was proposed to transmit information between the segmentation and classification modules, improving the performance of both tasks. An approach using encoder-decoder technique for semantic segmentation of pixel-wise based on SegNet is proposed in Youssef *et al.* [20]. Similar to the SegNet network [21], each decoder layer is connected to a corresponding encoder layer. The network is trained from scratch, with no fine-tuning or transfer of learning. The dropout technique is applied to each convolutional layer of the encoder for overfitting issues. Finally, a softmax layer is used for the output containing three classes: lesion, skin, and artifacts.

Some authors [22, 23, 24] performed the lesion segmentation task in dermoscopic images and evaluated the U-Net network in the ISIC 2018 database. Abder-Rahman *et al.* [22] compared a supervised approach with an unsupervised one and analyzed the U-Net network with a "SELU" activation function, which, according to them, enabled an improvement in the results. Other authors [24] proposed a two-stage approach: in the first stage, the segmentation of the skin lesion is performed using the U-Net; and, in the second, the edge is extracted from the segmented image using *FuzzEdge* method. Codella *et al.* [25] proposed a U-Net-like architecture to segment the skin lesion. In their proposal, the images are inserted into the network with six color channels, Red, Green, Blue (RGB), and Hue, Saturation, Value (HSV). More recently, Arora *et al.* [26] proposed the 11-layer CNN with two segmentation models. In Arora *et al.* [27], and Tong *et al.* [28], the authors presented a proposal inserting an Attention Gate (AG) module, which is a mechanism used in human perception, *i.e.*, more attention is paid to a certain region and demands less attention to less important details. Arora *et al.* [27] modified the U-Net network by inserting Group Normalization (GN) layers in the encoder and decoder path to normalize the resources. They were later activated by an activation unit ReLU. The bottleneck of the proposed architecture is composed of dilated convolutions, and AG focuses on skip connections details.

Hasan *et al.* [29] proposed a hybrid-CNN that has three feature extractor modules. The segmentation, rebalancing, and augmentation tasks were performed as pre-processing. Finally, DermoExpert's weights were passed to a web application. The authors made available the code and segmented masks on GitHub. The authors evaluated the model's performance on three different bases: ISIC 2016, ISIC 2017, and ISIC 2018. Wu *et al.* [30] proposed a feature adaptive transformer network named FAT-Net based on the classical encoder-decoder architecture; they also do an ablation study. The proposal was evaluated in the ISIC 2016, ISIC 2017, ISIC 2018, and PH2 datasets. Khoulood *et al.* [31] proposed an architecture based on double encoder-decoder named W-net for segmentation task. The W-net is the junction between a ResNet Encoder-Decoder and a ConvNet Encoder-Decoder. This method also evaluated its performance in ISIC 2016, ISIC 2017, ISIC 2018, and PH2 datasets. Mohakud *et al.* [32] proposed a Fully Convolution Encoder-Decoder Network (FCEDN). Instead of manually defining the parameters, its hyperparameters were optimized by an Exponential Neighborhood Gray Wolf Optimization (EN-GWO) algorithm. Finally, Kaur *et al.* [33] proposed a two-step system. The first step, pre-processing, consists of a hair removal module; according to the authors, it improved the general performance. The second step segments the skin lesion using a CNN

based on encoder-decoder architecture.

Based on the excellent performance of the U-Net network in medical image segmentation tasks, this work proposed a convolutional neural network based on U-Net [34]. We apply some modifications to its original architecture, such as increasing the number of layers in the network and adjusting the optimization and activation functions. Furthermore, we insert dropout layers to achieve better generalization. An ablation study is still carried out, demonstrating some of the experiments carried out to justify using specific parameters. Finally, the ISIC 2017 [35], and ISIC 2018 [36] databases were used to evaluate the network considering the accuracy and Jaccard metrics. In addition, the results achieved are compared with other research works in the literature.

3 Proposed model

The architecture proposed in this study is presented in Table 1. Our proposal is based in a U-net architecture formed by next blocks: an encoder, a bottleneck and a decoder. In the encoder path, that is, on the left side of the network, regular convolutions and max-polling are applied. In this step, the dimensions of the images are reduced (x and y axes) while their depth (z axis) gradually increases. For example, considering the network entry of an image $256 \times 256 \times 3$, at the end of the encoding, there is an image $8 \times 8 \times 128$. This process at the encoder enables the network to extract relevant information about the image. The layers that make up the encoder block are described below:

- 2D convolution, with a 3×3 kernel and activation function “ReLU”;
- 2D convolution, with a 3×3 kernel and activation function “ReLU”;
- Max-pooling 2D with pool (2, 2) and strides (2, 2)).

We add a dropout layer of 0.2 in “block 4”.

In the decoder path, the image size (x and y axes) increases, and the depth (z axis) decreases, both gradually. From the application of up-sampling and jump connections concatenating the output of the convolution performed in the encoder with the same level in the decoder, the image information is retrieved, and thus it is reconstructed. The decoder block consists of the following layers:

- Up-sampling 2D com size 2×2 ;
- Concatenation between the ‘*us*’ and ‘*skip*’ layers, where ‘*us*’ is the output of the up-sampling layer, and ‘*skip*’ is the output of the convolution at the same level as the encoder;
- 2D convolution, with a 3×3 dimension kernel and “ReLU” activation function;
- 2D convolution, with a 3×3 dimension kernel and “ReLU” activation function.

Between the encoder and decoder, we have the bottleneck block, formed by two convolutional layers:

- 2D convolution, with a 3×3 dimension kernel and “ReLU” activation function;
- 2D convolution, with a 3×3 dimension kernel and “ReLU” activation function.

A final convolutional layer was applied over the final feature map, containing the following parameters:

- 2D convolution layer using filter 1, with kernel size 1×1 and “Sigmoid” activation function.

In the present work, some modifications in the network U-Net were proposed. One of the modifications is related to its structure. The original U-Net [34] architecture has four blocks in the encoder path and four blocks in the decoder path. From experiments carried out with the shallower and deeper net, it can be seen that the net with four blocks is sensitive to noise found in the images. The network became less sensitive by including one more block in the encoder and decoder. Therefore, as proposed, the architecture now has five blocks on both paths. We use the Sigmoid function as the active function of the last convolutional layer of the network. The other parameters related to its original structure were kept.

Other important changes made to the network were in the optimization and loss functions. For the optimization function, we use Ada [37] with $\epsilon = 2e - 4$ and $\beta_1 = 0.5$, and for the loss function the binary crossentropy. After some adjustments in the network parameters, the proposal significantly improved the segmentation result. In Table 2, the original U-Net is compared to the U-Net with the proposed adjustments. The network with its adjustments was trained for 100 epochs and with batch size of 1.

Table 1: Proposed U-Net based convolutional neural network for skin lesion segmentation

ENCODER			DECODER		
Input 1: InputLayer	input:	(256, 256, 3)	Block 1		
	output:	(256, 256, 3)	up sampling2d: UpSampling2D	input:	(8, 8, 256)
Block 1				output:	(16, 16, 256)
conv2d: Conv2D	input:	(256, 256, 3)	concatenate: Concatenate	input:	[(16, 16, 256), (16, 16, 128)(conv2d 9)]
	output:	(256, 256, 8)		output:	(16, 16, 384)
conv2d 1: Conv2D	input:	(256, 256, 8)	conv2d 12: Conv2D	input:	(16, 16, 384)
	output:	(256, 256, 8)		output:	(16, 16, 128)
maxpooling2d: MaxPooling2D	input:	(256, 256, 8)	conv2d 13: Conv2D	input:	(16, 16, 128)
	output:	(128, 128, 8)		output:	(16, 16, 128)
Block 2			Block 2		
conv2d 2: Conv2D	input:	(128, 128, 8)	up sampling2d 1: UpSampling2D	input:	(16, 16, 128)
	output:	(128, 128, 16)		output:	(32, 32, 128)
conv2d 3: Conv2D	input:	(128, 128, 16)	concatenate 1: Concatenate	input:	[(32, 32, 128), (32, 32, 64)(conv2d 7)]
	output:	(128, 128, 16)		output:	(32, 32, 192)
max pooling2d 1: MaxPooling2D	input:	(128, 128, 16)	conv2d 14: Conv2D	input:	(32, 32, 192)
	output:	(64, 64, 16)		output:	(32, 32, 64)
Block 3			conv2d 15: Conv2D	input:	(32, 32, 64)
conv2d 4: Conv2D	input:	(64, 64, 16)		output:	(32, 32, 64)
	output:	(64, 64, 32)	Block 3		
conv2d 5: Conv2D	input:	(64, 64, 32)	up sampling2d 2: UpSampling2D	input:	(32, 32, 64)
	output:	(64, 64, 32)		output:	(64, 64, 64)
max pooling2d 2: MaxPooling2D	input:	(64, 64, 32)	concatenate 2: Concatenate	input:	[(64, 64, 64), (64, 64, 32)(conv2d 5)]
	output:	(32, 32, 32)		output:	(64, 64, 96)
Block 4			conv2d 16: Conv2D	input:	(64, 64, 96)
conv2d 6: Conv2D	input:	(32, 32, 32)		output:	(64, 64, 32)
	output:	(32, 32, 64)	conv2d 17: Conv2D	input:	(64, 64, 32)
conv2d 7: Conv2D	input:	(32, 32, 64)		output:	(64, 64, 32)
	output:	(32, 32, 64)	Block 4		
max pooling2d 3: MaxPooling2D	input:	(32, 32, 64)	up sampling2d 3: UpSampling2D	input:	(64, 64, 32)
	output:	(16, 16, 64)		output:	(128, 128, 32)
dropout: Dropout	input:	(16, 16, 64)	concatenate 3: Concatenate	input:	[(128, 128, 32), (128, 128, 16) (conv2d 3)]
	output:	(16, 16, 64)		output:	(128, 128, 48)
Block 5			conv2d 18: Conv2D	input:	(128, 128, 48)
conv2d 8: Conv2D	input:	(16, 16, 64)		output:	(128, 128, 16)
	output:	(16, 16, 128)	conv2d 19: Conv2D	input:	(128, 128, 16)
conv2d 9: Conv2D	input:	(16, 16, 128)		output:	(128, 128, 16)
	output:	(16, 16, 128)	Block 5		
max pooling2d 4: MaxPooling2D	input:	(16, 16, 128)	up sampling2d 4: UpSampling2D	input:	(128, 128, 16)
	output:	(8, 8, 128)		output:	(256, 256, 16)
BOTTLENECK			concatenate 4: Concatenate	input:	[(256, 256, 16), (256, 256, 8) (conv2d 1)]
conv2d 10: Conv2D	input:	(8, 8, 128)		output:	(256, 256, 24)
	output:	(8, 8, 256)	conv2d 20: Conv2D	input:	(256, 256, 24)
conv2d 11: Conv2D	input:	(8, 8, 256)		output:	(256, 256, 8)
	output:	(8, 8, 256)	conv2d 21: Conv2D	input:	(256, 256, 8)
				output:	(256, 256, 8)
			LAST CONVOLUTIONAL LAYER		
			conv2d 22: Conv2D	input:	(256, 256, 8)
				output:	(256, 256, 1)

Table 2: Comparison between the parameters of the Original U-Net and the U-Net with the proposed adjustments

Parameters	Original U-Net	Proposal U-Net
optimizer	stochastic gradient descent	Adam
Loss	cross entropy	binary_crossentropy
Dropout	Não	0,2
Number of blocks	8	10
Number of convolutional layers	23	28
Number of filters used	[64, 128, 256 , 512, 1024]	[8, 16, 32, 64, 128, 256]

Finally, Table 1 presents, in more detail, the general structure of the proposed network. In each block, the applied operation and the input and output size of the layer in question are demonstrated. In the decoder, the concatenation is represented as a layer, and information regarding the layers to be concatenated is presented in all blocks. At the end of the net, as output of the last convolutional layer, there is a generated mask of size 256×256 pixels with only one color channel.

4 Experiments Performed

4.1 Dermatoscopic Image Database

In the present work, the images were extracted from the public sets of ISIC 2017 challenges [35] and ISIC 2018 [38, 39]. The ISIC 2017 image set includes 2,750 dermoscopic images, with 2,000 images for the training stage, 150 images for validation, and 600 images for testing. For all these images, the corresponding ground truths are available.

The ISIC 2018 image set contains 3,694 dermoscopic images. For training it is available 2,594 skin lesion images, 100 images for validation, and 1,000 images for testing. The ground truths of the images for validation and testing are not available. Therefore, for the present work, only 2,594 training images were used. The others were discarded.

In both sets, the images were collected in leading international clinical centers with a dermoscopic device and submitted to annotations and markings by dermatologists. The markings, which originated the ground truth, include global dermoscopic characteristics, which discriminate types of lesions. Some examples of the images and their ground-truth, available in both the ISIC 2017 and the 2018 base, can be seen in Figure 2.

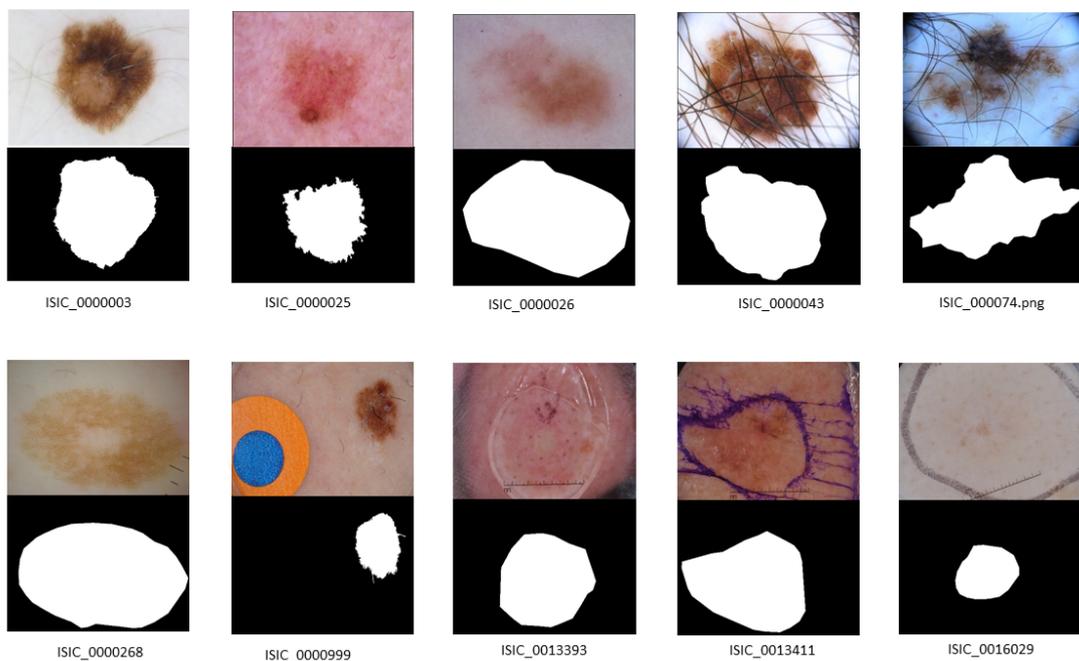


Figure 2: Examples of images of skin lesions and ground-truths from the ISIC 2017 and ISIC 2018 databases. For each image, it is shown below, its corresponding ground truth.

4.2 Image Pre-processing

Images in their original form are often not ready to be analyzed. The presence of artifacts in the images can harm the performance of the segmentation task, and, therefore, its treatment is necessary in order to improve its quality [40]. The dermoscopic images of bases ISIC 2017 and 2018 are of different sizes with high resolution, which can considerably increase the processing time. In order to decrease processing time, images have been resized to $256 \times 256 \times 3$ pixels and ground-truth to $256 \times 256 \times 1$. Normalization was performed, changing the pixel interval from $[0,225]$ to $[0,1]$. The application of the data augmentation technique, which produces alternative data from the existing dataset [41], is essential to train a neural network for the segmentation task when working with a database containing a few images [42]. 90° , flip horizontal and vertical rotation were applied to the resized images.

4.3 Implementation

The proposed algorithm for the segmentation task was implemented with Keras based on Tensorflow. The experiments were carried out on a Dell computer, with an Intel(R) Core™ i7-3770 CPU, 3.40GHz CPU with

16GB of memory, and an Nvidia GeForce GT 640 GPU with 1GB of memory. The CentOS Linux 7 operating system was used.

4.4 Evaluation Metrics

In the present work, the Accuracy, Jaccard, sensitivity, specificity and Dice coefficient metrics were adopted, calculated from the following information obtained from the confusion matrix:

- True Positive (**VP**): when the class to be predicted (lesion pixels) is correctly predicted;
- False Positive (**FP**): when the class to be predicted (lesion pixels) is incorrectly predicted;
- True Negative (**VN**): when the class you don't want to predict (skin pixels) is correctly predicted;
- False Negative (**FN**): when the class you don't want to predict (skin pixels) is incorrectly predicted.

Thus, accuracy is the proportion of correctly predicted cases, both true positives, and true negatives. The Equation (1) represents the accuracy metric:

$$Accuracy = \frac{VP + VN}{VP + FN + VN + FP}. \quad (1)$$

The Jaccard Index is a metric commonly used to assess the detection of an object. It can be defined as the intersection ratio over the union of the segmented image predicted by the network and the ground truth. The larger the overlap, the better the Jaccard value. The Jaccard index is defined by the Equation (2):

$$Jaccard = \frac{VP}{VP + FN + FP}. \quad (2)$$

The Dice coefficient is the metric used to measure the similarity of two samples. The Equation (3) defines the Dice Coefficient.

$$Dice\text{-Coefficient} = \frac{2.VP}{2.(VP + FN + FP)}. \quad (3)$$

The sensitivity, also known as true positive rate, calculates the proportion of samples that are genuinely positive.

$$Sensitivity = \frac{TP}{TP + FN}. \quad (4)$$

The specificity, also known as true negative rate, is the proportion of samples that are genuinely negative.

$$Specificity = \frac{TN}{TN + FP}. \quad (5)$$

4.5 Experimental Results

A neural network model needs to generalize training learning in the test set. One of the biggest obstacles faced in the generalization was the difficulty in obtaining a sufficiently large base of images. The approach usually applied for databases with many images divides the dataset into training, validation, and testing subsets. For databases that contain few images, such as ISIC 2017 and ISIC 2018, the model may not generalize as expected. Thus, to minimize the impact due to the small number of images, the data augmentation [41] technique was applied. From ISIC 2017 and ISIC 2018, we used 2,750, and 2,594 images, respectively. The datasets were divided: 80% of the images for training and 20% of the images for testing. Thus, the ISIC 2017 base was left with 2,200 images for training and 550 images for testing. The ISIC 2018 base was left with 2,075 images for training and 519 images for testing. We ran the experiments five times and reported the mean value.

In the proposed model, the number of blocks of the convolution path and the expansion path, the optimizer, the Loss function, the application of techniques such as data augmentation, dropout, and normalization, the function used in the activation layer (both in internal layers of the network and in the last layer) were tested. Input image size, batch size, and base splitting were also evaluated. Finally, we arrive at the following configuration shown in Table 3.

Table 3: Parameters Defined for the Proposed Model

Architecture	U-Net (5 blocks)
computer used for processing	GPU GT640
Optimizer	Adam ($2e - 4$, $\beta_1 = 0.5$)
Loss	binary_crossentropy
Data augmentation	yes
Dropout	yes
Activation (inner layers)	ReLU
Activation (last layer)	Sigmoid
Size of images	256×256
Normalization	[0, 1]
Epoch	100
Batch size	1
training for images	80%
Images for validation	20%

4.5.1 Experimental Results for ISIC 2017 Base

In Table 4, we can see the results achieved by the proposed model evaluated in the ISIC 2017 image dataset. Furthermore, the results of some methods in the literature that performed the segmentation task are presented, including the U-Net, considering the Accuracy evaluation metrics and the Jaccard Index. Analyzing the Accuracy metric, the model obtained 0.949, which represents a better performance compared to Liu *et al.* [43], which obtained 0.930. Furthermore, the proposed model surpassed the results presented in Goyal *et al.* [44], Vesal *et al.* [45] and the original U-Net, that was implemented in this work according to Ronneberger *et al.* [34], leaving only 0.001 behind the result of Li and Shen *et al.* [46]. Observing the results for the Jaccard Index, our model obtained 0.833, that is, a higher result than the proposed in Liu *et al.* [43], Goyal *et al.* [44], Vesal *et al.* [45], and Li and Shen *et al.* [46].

Table 4: Performance comparison of the proposed model from the Accuracy and Jaccard Index metrics for the 2017 ISIC image dataset

Model	Accuracy	Jaccard index
Liu <i>et al.</i> (2019) [43]	0.930	0.752
Goyal <i>et al.</i> (2019) [44]	0.941	0.793
Vesal <i>et al.</i> (2018) [45]	0.851	0.767
Li <i>et al.</i> (2018) [46]	0.950	0.753
Original U-Net [34]	0.877	0.325
Proposed model	0.949	0.833

4.5.2 Experimental Results for 2018 ISIC Base

In Table 5, we can see the results achieved from the experiments performed, referring to the accuracy and Jaccard obtained. The values resulting from the tests were compared to the methods in the literature, which propose a solution to the problem of skin lesion segmentation using machine learning techniques and tests in the ISIC 2018 database. The proposed model achieved an accuracy of 0.954, and its Jaccard was 0.850. The model proposed in this article, even being simpler, surpassed in terms of accuracy the method proposed by [47], which reached 0.935. The authors in [47] applied the Shades of Gray method in the pre-processing and used an ensemble for segmentation with two neural networks, the VGG-UNet, and the DeeplabV3. It presents a better Jaccard Index compared to the proposed model by [48]. The authors evaluated their proposal also based on U-Net, called C-UNet, by applying a Dice Finetune technique.

In Figures 3 and 4 are presented some segmentation results made by the proposed model, containing images with different levels of difficulty for the segmentation task, resulting from the evaluation of the bases ISIC 2018 and ISIC 2017, respectively. In the first column are the original images acquired by a specialist through a dermatoscopic. In the second, we can find the ground truth representing the segmentation defined

Table 5: Performance comparison of the proposed model from the Accuracy, Jaccard Index, Sensitivity, Specificity and Dice Coefficient metrics for the 2018 ISIC image dataset

Model	Accuracy	Jaccard	Sensitivity	Specificity	Dice Coefficient
Ali <i>et al.</i> (2019) [47]	0.935	0.816	-	-	0.886
Koohbanani <i>et al.</i> (2018) [49]	0.968	0.847	0.921	0.973	0.910
Wu <i>et al.</i> (2019) [48] -	-	0.775	-	-	0.869
Arora, Ginni <i>et al.</i> (2021) [26]	0.903	0.820	0.918	0.888	0.901
Arora, Ridhi <i>et al.</i> (2021) [27]	0.950	0.830	0.940	0.950	0.910
Original U-Net [34]	0.954	0.688	-	-	-
Proposed model	0.953	0.847	0.913	0.975	0.913

by specialists in the area. In the third column, the mask is predicted by the original U-Net network, and in the fourth column, the mask is predicted by the proposed modified U-Net. In some images, the net segments very well, but in others, the net fails, not correctly segmenting the lesion. This is due to different formats of lesions, which may indicate a need for a more significant number of images for training.

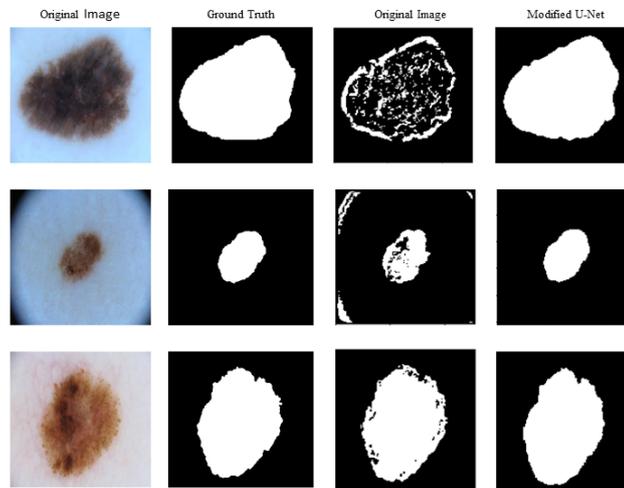


Figure 3: Images resulting from the original U-Net model and modified U-Net model evaluated in the ISIC 2017 image base

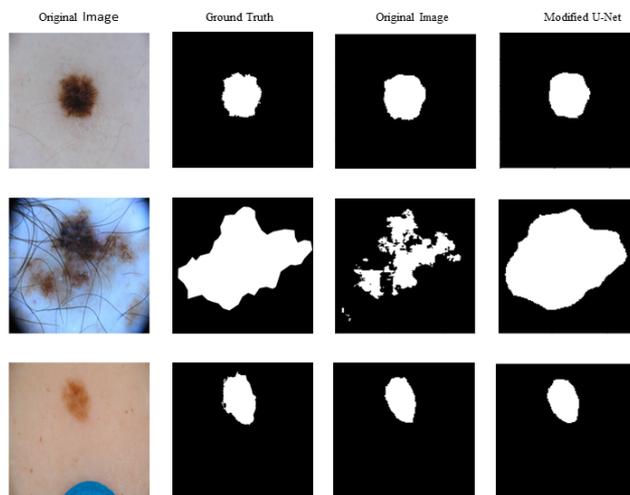


Figure 4: Images resulting from the original U-Net model and U-Net modified for ISIC 2018 image base

4.6 Ablation Study for U-Net Parameters

In order to verify the effectiveness of the model regarding its parameters, an ablation study was carried out. Initially, the model’s performance is compared considering different numbers of epochs. The experimental results are reported in Table 6, where it can be seen that up to epoch 100, there is an improvement in the Jaccard value and the Dice Coefficient. After that, the network does not improve significantly for both metrics.

Table 6: The performance of U-Net with different numbers of epochs

Number of epochs	Loss	Accuracy	Jaccard	Sensitivity	Specificity	Dice Coefficient
10	0.1259	0.9518	0.7583	0.9031	0.9771	0.8549
30	0.1337	0.9552	0.8235	0.8993	0.9822	0.8995
50	0.1881	0.9547	0.8422	0.9155	0.9743	0.9107
100	0.2638	0.9546	0.8511	0.9249	0.9710	0.9151
150	0.3053	0.9551	0.8537	0.9214	0.9727	0.9170
200	0.3462	0.9550	0.8552	0.9201	0.9738	0.9181
300	0.3915	0.9545	0.8548	0.9130	0.9763	0.9178

The databases used in the experiments are composed of high-resolution images, so resizing to reduce the size of the images to be processed is necessary. Tests were performed with images of 128×128 , 160×160 , 256×256 and 512×512 , as can be seen in Table 7. So, in the final experiments, the image size was set to 256. The experiments show that it is unnecessary to work with larger images, thus reducing the computational cost and the necessity of powerful computers. Based on the experiments, we conclude that high-resolution images introduce more noise in the training process.

Table 7: The Performance U-Net for Different Image Sizes

Experiments	Image size	Loss	Accuracy	Jaccard
Train 1	128×128	0.2712	0.9505	0.8417
Train 2	160×160	0.2601	0.9508	0.8408
Train 3	256×256	0.2831	0.9518	0.8443
Train 4	512×512	0.2812	0.9494	0.8385

Table 8 shows the performance of the experiments using various optimization functions. When observing the results, it is clear that the Adam function with the adjustments in its parameters favors the network to reach the best result found during the trainings.

Table 8: U-Net performance for different optimization functions

Experiments	Optimizer	Loss	Accuracy	Jaccard
Run 1	Adam	0.3329	0.8643	0.4517
Run 2	Adam(beta_1=0.9)	0.1840	0.9399	0.7284
Run 3	Adam(2e-4, beta_1=0.9)	0.4009	0.9438	0.7920
Run 4	Adam(2e-4, beta_1=0.5)	0.2672	0.9544	0.8527
Run 5	SGD	0.2072	0.9509	0.8348
Run 6	Adamax	0.3261	0.9521	0.8465
Run 7	Nadam	0.2567	0.8937	0.5257
Run 8	Adagrad	0.2201	0.9463	0.8163
Run 9	Adamax(epsilon= 2e-4)	0.3149	0.9535	0.8482
Run 10	Adamax(epsilon= 2e-4, beta_1=0.5)	0.2845	0.9547	0.8508
Run 11	Adadelta	0.1921	0.9242	0.6706

In Table 9, there are some tests performed considering a dropout layer. By comparison, the test presented in the first row of the table shows the performance without dropout. As can be seen, the Jaccard variation for the experiments is minimal, even when testing different values and in different blocks.

Table 9: U-Net Performance with Different Dropout Executions

Dropout	Encoder Blocks					Decoder Blocks					Accuracy	Jaccard
	1	2	3	4	5	1	2	3	4	5		
Test 1	-	-	-	-	-	-	-	-	-	-	0.9539	0.8491
Test 2	-	-	-	-	-	0.2	-	-	-	-	0.9523	0.8459
Test 3	-	-	-	0.4	0.4	0.4	-	-	-	-	0.9528	0.8452
Test 4	-	-	-	-	0.5	-	-	-	-	-	0.9524	0.8434
Test 5	-	-	-	0.4	-	-	-	-	-	-	0.9544	0.8506
Test 6	-	-	-	-	0.4	-	-	-	-	-	0.9546	0.8461

Experiments carried out with different numbers of blocks show that the U-Net with 5 blocks in encoding and 5 blocks in decoding presents a better result, these data can be visualized in Table 10.

Table 10: U-Net Performance with Different block number Executions

Experiments	Number of Blocks	Filters	Loss	Accuracy	Jaccard
Run 1	Cod: 3; Decod: 3	32, 64, 128, 256	0.2853	0.9366	0.7478
Run 2	Cod: 4; Decod: 4	16, 32, 64, 128, 256	0.3641	0.9447	0.7978
Run 3	Cod: 5; Decod: 5	8, 16, 32, 64, 128, 256	0.2672	0.9544	0.8527
Run 4	Cod: 6; Decod: 6	8, 16, 32, 64, 128, 256, 512	0.3563	0.9457	0.8056
Run 5	Cod: 8; Decod: 8	2, 4, 8, 16, 32, 64, 128, 256, 512	0.3171	0.9393	0.7787

Finally, we have the experiments carried out to find the best Activation and Loss Functions. The results regarding the activation functions are shown in Table 11, and it is concluded that the Relu function favors the network in the lesion segmentation task. As for the Loss function, shown in Table 12, the binary cross-entropy function has a better result compared to the Poisson Loss function. Other Loss functions were tested, however, they were discarded for presenting unsatisfactory results.

Table 11: U-Net Performance with Different Activation Function Executions

Activation Function	Loss	Accuracy	Jaccard
Elu	0.3818	0.9458	0.7976
LeakyRelu	0.3153	0.9471	0.8010
Relu	0.2672	0.9544	0.8527
Sigmoid	0.2102	0.9407	0.7610
Softmax	0.2922	0.8919	0.4864
Softplus	0.4222	0.9403	0.7806
Tanh	0.1939	0.9470	0.7870

Table 12: U-Net Performance with Different Loss Function Executions

Loss Function	Loss	Accuracy	Jaccard
binary_crossentropy	0.2672	0.9544	0.8527
poisson	0.4661	0.9471	0.7970

5 Conclusion

We presented a U-Net-based approach for skin lesion segmentation in this work. The ISIC 2017 and ISIC 2018 dermoscopic image databases evaluated the model. Modifications were proposed to improve generalizability. Some of the tests performed are demonstrated from an ablation study performed. Some optimization functions, like the Adam function, were tested with some parameter adjustments, which enabled the network to obtain a better result. The Loss function achieved the best result is the binary cross-entropy function. Other functions were also tested but with a lower performance. Furthermore, tests were performed on the network with variations in its number of layers, making it deeper and deeper, concluding that the five blocks in the encoding path and five blocks in the decoding path made the network less sensitive to noise. The

impact of the dropout layer in different situations was also evaluated. Comparing the proposed model with models found in the literature, it is recognized that there are still adjustments to be made, but the results prove the feasibility and potential of the proposed model. The data set with a small number of images was limited, as they are medical images, and the difficulty in having a sufficiently large database is understood. Another parameter that has an impact on this problem is the image size. Higher-resolution images introduce noise in the training process since minor variations appear more evident.

References

- [1] W. C. R. F. International., “Wcrf international, american institute for cancer research,” <https://www.wcrf.org/dietandcancer/skin-cancer-statistics/>, 2020, accessed: Apr, 2022.
- [2] A. A. R. O. Halpern, Allan C. Marghoob, “Melanoma overview: A dangerous skin cancer.” <https://www.skincancer.org/skin-cancer-information/melanoma/>, 2020, accessed: May, 2021.
- [3] INCA, “Instituto nacional de câncer,” <https://www.inca.gov.br/tipos-de-cancer/cancer-de-pele-melanoma>, 2020, accessed: May, 2021.
- [4] H. Tsao, J. M. Olazagasti, K. M. Cordoro, J. D. Brewer, S. C. Taylor, J. S. Bordeaux, M.-M. Chren, A. J. Sober, C. Tegeler, R. Bhushan, and W. S. Begolka, “Early detection of melanoma: Reviewing the abcdes,” *Journal of the American Academy of Dermatology*, vol. 72, no. 4, pp. 717–723, 2015. doi: <https://doi.org/10.1016/j.jaad.2015.01.025>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0190962215000900>
- [5] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, “Epiluminescence Microscopy for the Diagnosis of Doubtful Melanocytic Skin Lesions: Comparison of the ABCD Rule of Dermatoscopy and a New 7-Point Checklist Based on Pattern Analysis,” *Archives of Dermatology*, vol. 134, no. 12, pp. 1563–1570, 12 1998. doi: 10.1001/archderm.134.12.1563. [Online]. Available: <https://doi.org/10.1001/archderm.134.12.1563>
- [6] S.-R. (2021), “Sociedade brasileira de dermatologia.” <https://sbdrs.org.br/autoexame-do-abcde-ajuda-no-diagnostico-de-cancer-de-pele-melanoma/>, 2020, accessed: Jan, 2021.
- [7] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, “Automated melanoma recognition in dermoscopy images via very deep residual networks,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2017. doi: 10.1109/TMI.2016.2642839
- [8] Z. Ma and S. Yin, “Deep attention network for melanoma detection improved by color constancy,” in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, 2018. doi: 10.1109/ITME.2018.00037 pp. 123–127.
- [9] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, “Automated melanoma recognition in dermoscopy images via very deep residual networks,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2017. doi: 10.1109/TMI.2016.2642839
- [10] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, “Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering,” *International Journal of Medical Informatics*, vol. 124, pp. 37–48, 2019. doi: <https://doi.org/10.1016/j.ijmedinf.2019.01.005>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1386505618307470>
- [11] Y. Yuan and Y.-C. Lo, “Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 519–526, 2019. doi: 10.1109/JBHI.2017.2787487
- [12] ISIC, “The international skin imaging collaboration.” <https://challenge.isic-archive.com/>, accessed: May, 2021.
- [13] G. S. Araujo, G. Cámara-Chávez, and R. B. Oliveira, “Convolutional neural networks applied for skin lesion segmentation,” in *2021 XLVII Latin American Computing Conference (CLEI)*, 2021. doi: 10.1109/CLEI53233.2021.9640189 pp. 1–10.

- [14] B. Olimov, K. Sanjar, S. Din, A. Ahmad, A. Paul, and J. Kim, "Fu-net: fast biomedical image segmentation model based on bottleneck convolution layers," *Multimedia Systems*, vol. 27, no. 4, pp. 637–650, 2021. doi: 10.1007/s00530-020-00726-w
- [15] D. Alvarez and M. Iglesias, "k-means clustering and ensemble of regressions: an algorithm for the isic 2017 skin lesion segmentation challenge," *arXiv preprint arXiv:1702.07333*, 2017.
- [16] A. Agarwal, A. Issac, M. K. Dutta, K. Riha, and V. Uher, "Automated skin lesion segmentation using k-means clustering from digital dermoscopic images," in *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2017. doi: 10.1109/TSP.2017.8076087 pp. 743–748.
- [17] S. Garg and B. Jindal, "Skin lesion segmentation using k-mean and optimized fire fly algorithm," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7397–7410, 2021. doi: 10.1007/s11042-020-10064-8
- [18] A. Bibi, M. A. Khan, M. Y. Javed, U. Tariq, B.-G. Kang, Y. Nam, R. R. Mostafa, and R. H. Sakr, "Skin lesion segmentation and classification using conventional and deep learning based framework," 2022. doi: 10.32604/cmc.2022.018917
- [19] S. Chen, Z. Wang, J. Shi, B. Liu, and N. Yu, "A multi-task framework with feature passing module for skin lesion classification and segmentation," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018. doi: 10.1109/ISBI.2018.8363769 pp. 1126–1129.
- [20] A. Youssef, D. D. Bloisi, M. Muscio, A. Pennisi, D. Nardi, and A. Facchiano, "Deep convolutional pixel-wise labeling for skin lesion image segmentation," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2018. doi: 10.1109/MeMeA.2018.8438669 pp. 1–6.
- [21] A. Kendall, V. Badrinarayanan, and R. Cipolla, "Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding," 2015. [Online]. Available: <https://arxiv.org/abs/1511.02680>
- [22] A.-R. Ali, J. Li, and T. Trappenberg, "Supervised versus unsupervised deep learning based methods for skin lesion segmentation in dermoscopy images," in *Canadian Conference on Artificial Intelligence*. Springer, 2019, pp. 373–379.
- [23] S. N. Hasan, M. Gezer, R. A. Azeez, and S. Gülseçen, "Skin lesion segmentation by using deep learning techniques," in *2019 Medical Technologies Congress (TIPTEKNO)*, 2019. doi: 10.1109/TIPTEKNO.2019.8895078 pp. 1–4.
- [24] A.-R. Ali, J. Li, S. J. O'Shea, G. Yang, T. Trappenberg, and X. Ye, "A deep learning based approach to skin lesion border extraction with a novel edge detector in dermoscopy images," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019. doi: 10.1109/IJCNN.2019.8852134 pp. 1–7.
- [25] N. C. F. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5:1–5:15, 2017. doi: 10.1147/JRD.2017.2708299
- [26] G. Arora, A. K. Dubey, Z. A. Jaffery, and A. Rocha, "Architecture of an effective convolutional deep neural network for segmentation of skin lesion in dermoscopic images," *Expert Systems*, p. e12689, 2021. doi: 10.1111/exsy.12689
- [27] 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. doi: <https://doi.org/10.1016/j.bspc.2020.102358>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1746809420304663>
- [28] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, "Ascu-net: attention gate, spatial and channel attention u-net for skin lesion segmentation," *Diagnostics*, vol. 11, no. 3, p. 501, 2021. doi: 10.3390/diagnostics11030501
- [29] M. K. Hasan, M. T. E. Elahi, M. A. Alam, M. T. Jawad, and R. Martí, "Dermoexpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation," *Informatics in Medicine Unlocked*, vol. 28, p. 100819, 2022. doi: <https://doi.org/10.1016/j.imu.2021.100819>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352914821002835>

- [30] H. Wu, S. Chen, G. Chen, W. Wang, B. Lei, and Z. Wen, "Fat-net: Feature adaptive transformers for automated skin lesion segmentation," *Medical Image Analysis*, vol. 76, p. 102327, 2022. doi: <https://doi.org/10.1016/j.media.2021.102327>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841521003728>
- [31] S. Khoulood, M. Ahlem, T. Fadel, and S. Amel, "W-net and inception residual network for skin lesion segmentation and classification," *Applied Intelligence*, pp. 1–19, 2021.
- [32] R. Mohakud and R. Dash, "Skin cancer image segmentation utilizing a novel en-gwo based hyper-parameter optimized fcedn," *Journal of King Saud University - Computer and Information Sciences*, 2022. doi: <https://doi.org/10.1016/j.jksuci.2021.12.018>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1319157821003645>
- [33] R. Kaur, H. GholamHosseini, and R. Sinha, "Skin lesion segmentation using an improved framework of encoder-decoder based convolutional neural network," *International Journal of Imaging Systems and Technology*, 2022.
- [34] 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*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015. ISBN 978-3-319-24574-4 pp. 234–241.
- [35] I. challenge, <https://challenge.isic-archive.com/landing/2017>, accessed: May, 2021.
- [36] I. 2018, "The international skin imaging collaboration." <https://challenge2018.isic-archive.com/>, 2018, "accessed: Jun, 2021".
- [37] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond," 2019. [Online]. Available: <https://arxiv.org/abs/1904.09237>
- [38] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic)," 2019.
- [39] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, no. 1, pp. 1–9, 2018. doi: <https://doi.org/10.1038/sdata.2018.161>
- [40] M. Jafari, N. Karimi, E. Nasr-Esfahani, S. Samavi, S. Soroushmehr, K. Ward, and K. Najarian, "Skin lesion segmentation in clinical images using deep learning," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016. doi: 10.1109/ICPR.2016.7899656 pp. 337–342.
- [41] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," 2017. [Online]. Available: <https://arxiv.org/abs/1711.04340>
- [42] Y. Xie and D. Richmond, "Pre-training on grayscale imagenet improves medical image classification," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.
- [43] L. Liu, L. Mou, X. X. Zhu, and M. Mandal, "Skin lesion segmentation based on improved u-net," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, 2019. doi: 10.1109/CCECE.2019.8861848 pp. 1–4.
- [44] M. Goyal, A. Oakley, P. Bansal, D. Dancy, and M. H. Yap, "Skin lesion segmentation in dermoscopic images with ensemble deep learning methods," *IEEE Access*, vol. 8, pp. 4171–4181, 2020. doi: 10.1109/ACCESS.2019.2960504
- [45] 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. doi: 10.1109/NSSMIC.2018.8824732 pp. 1–3.
- [46] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, 2018. doi: 10.3390/s18020556. [Online]. Available: <https://www.mdpi.com/1424-8220/18/2/556>

- [47] R. Ali, R. C. Hardie, B. Narayanan Narayanan, and S. De Silva, "Deep learning ensemble methods for skin lesion analysis towards melanoma detection," in *2019 IEEE National Aerospace and Electronics Conference (NAECON)*, 2019. doi: 10.1109/NAECON46414.2019.9058245 pp. 311–316.
- [48] J. Wu, E. Z. Chen, R. Rong, X. Li, D. Xu, and H. Jiang, "Skin lesion segmentation with c-unet," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019. doi: 10.1109/EMBC.2019.8857773 pp. 2785–2788.
- [49] N. A. Koohbanani, M. Jahanifar, N. Z. Tajeddin, A. Gooya, and N. M. Rajpoot, "Leveraging transfer learning for segmenting lesions and their attributes in dermoscopy images," *arXiv preprint arXiv:1809.10243*, 2018.