

## Research Article

# Skin Lesion Segmentation Based on Edge Attention Vnet with Balanced Focal Tversky Loss

Majid Nour <sup>1</sup>, Hakan Öcal,<sup>2</sup> Adi Alhudhaif <sup>3</sup>, and Kemal Polat <sup>4</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>2</sup>Computer Engineering/School of Natural and Applied Sciences, Gazi University, Ankara, Turkey

<sup>3</sup>Department of Computer Science, College of Computer Engineering and Sciences in Al-Kharj, Prince Sattam Bin Abdulaziz University, P.O. Box 151, Al-Kharj 11942, Saudi Arabia

<sup>4</sup>Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant İzzet Baysal University, 14280 Bolu, Turkey

Correspondence should be addressed to Kemal Polat; [kpolat@ibu.edu.tr](mailto:kpolat@ibu.edu.tr)

Received 29 April 2022; Accepted 25 May 2022; Published 14 June 2022

Academic Editor: Ramin Ranjbarzadeh

Copyright © 2022 Majid Nour et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Segmentation of skin lesions from dermoscopic images plays an essential role in the early detection of skin cancer. However, skin lesion segmentation is still challenging due to artifacts such as indistinguishability between skin lesion and normal skin, hair on the skin, and reflections in the obtained dermoscopy images. In this study, an edge attention network (ET-Net) combining edge guidance module (EGM) and weighted aggregation module is added to the 2D volumetric convolutional neural network (Vnet 2D) to maximize the performance of skin lesion segmentation. In addition, the proposed fusion model presents a new fusion loss function by combining balanced binary cross-entropy (BBCE) and focal Tversky loss (FTL). The proposed model has been tested on the ISIC 2018 Task 1 Lesion Boundary Segmentation Challenge dataset. The proposed model outperformed the state-of-the-art studies as a result of the tests.

## 1. Introduction

Melanoma, a deadly skin cancer, is predicted to be the fifth most frequently diagnosed cancer for men (57,180 cases) and women (42,600 cases) by 2022 [1]. In the United States, treating skin cancers' annual cost is \$ 8.1 billion: about \$ 4.8 billion for nonmelanoma skin cancers and \$ 3.3 billion for melanoma [2]. The number of diagnoses and treatments for nonmelanoma skin cancers in the USA between 1994 and 2014 reached 77% [3]. Approximately 90% of nonmelanoma skin cancers of nonmelanoma are associated with ultraviolet (UV) radiation from the sun [4]. Board-certified dermatologists very often use dermoscopy, a noninvasive technique that can be helpful in diagnosing skin lesions. Dermoscopy is a magnifying device that allows the application of liquid between dermoscopy and the patient's skin or the use of cross-polarized light to make the epidermis translucent, allowing the structures in the epidermis and

superficial dermis to be visualized. Seeing very small skin lesions that are not visible to the naked eye is essential in deciding for a physician. Dermoscopy provides better performance than traditional methods thanks to ABCD criteria [5]. However, diagnosing the lesioned area with a dermoscopy device is time-consuming and challenging due to artifacts such as skin hairs, blood vessels, and similarities and contrast of light between ordinary and lesioned skin. Therefore, deep convolutional neural networks (DCNN), which will automatically identify images, have been used to overcome this difficult task. Such systems try to determine lesion boundaries and make high-accuracy decisions based on skin lesion segmentation [6–11].

Convolutional neural networks (CNNs) performed higher for many segmentation tasks [12]. DCNNs require large amounts of data for high performance. In addition, there is no clear boundary between the lesion and the surrounding skin. In addition, since the sizes and shapes of

the lesions vary, it is not easy to have information about their characteristics. In addition, ink marks and air bubbles are other difficulties [13]. An edge attention Vnet (ET-Vnet) model is proposed to overcome the artefacts in this study. ET-Vnet combines two attention mechanisms to extract information about the lesion boundary better. Specifically, ET-Vnet contains two paths in its decoder, each embedded with an attention module. In addition, BBCE and FTL functions are combined as a novel approach to increase the performance of the proposed model. Czajkowska et al. [14] proposed a model consisting of two DeepLab v3+ models with a ResNet-50 backbone and a fuzzy connectivity analysis module for fine segmentation. Hu et al. [6] proposed a new attention synergy network (AS-Net) to improve the discriminative ability for skin lesion segmentation by combining spatial and channel attention mechanisms. Arora et al. [15] proposed a modified U-Net-based segmentation model for automatic skin lesion segmentation. The proposed model used group normalization (GN) instead of batch normalization (BN) in both encoder and decoder layers [16, 17]. Besides, attention gates (AG), which focuses on minute details in the skip connection, and Tversky loss (TL) function, which provides higher success in class imbalances, were used. Gu et al. [18] proposed an attention-based modified U-Net (CA-Net) that comprehensively presents based on U-Net architecture by adding multiple spatial attention between layers and awareness of the most critical spatial locations and channel scales. A new channel attention module was used to further focus the proposed model on the lesion area's feature map. A scaling module has been presented to scale the images, highlighting the most prominent feature maps. Goyal et al. [19] proposed an ensemble of R-CNN and Deeplab V3C methods, to achieve high sensitivity and specificity in lesion boundary segmentation methods. The proposed ensemble method ensemble-S achieved better performance than FrCN, FCNs, U-Net, and SegNet. Jiang et al. [20] proposed the CSARM-CNN (canal and spatial attention residue module) model for automatic skin lesion segmentation based on deep learning. Each convolutional layer of CSARM has created a new attention module by combining channel attention and spatial attention to make segmentation training more effective. The spatial pyramid pool acquires multidimensional input images. Finally, two different cross-entropy methods were fused for higher segmentation training in the model, and the final loss function was obtained. Lei et al. [21] proposed a general adversarial network (GAN) to overcome automated lesion segmentation challenges. This network includes a fully dense U-Net-based skip connection and double discrimination (DD) layers. The proposed method (U-Net-SCDC) uses lower resolution up-sampling convolutional layers that preserve fine-grained information. In contrast, the DD module increases training performance by controlling each other in opposite directions. Thus, the two different methods work as if they try to find each other's fault, focusing on their wrong points. Therefore, using a conditional discriminatory loss, it has been said that the model that checks each other simultaneously provided superior performance compared to other models.

Shan et al. [22] proposed a novel segmentation method named FC-DPN. The proposed method consists of the fully convolutional (FCN) and dual-path network (DPN). DPN is a model that enables the virtual feature maps of previous layers to be reused by using residual and densely connected ways. Sub-DPN projection blocks and sub-DPN processing blocks have been added instead of dense layers in the fully convolutional DenseNets (FC-DenseNets). It was stated that this method allows FC-DPN to acquire more representative and distinctive features to perform a more robust segmentation.

Dosovitskiy et al. removed the CNN-based encoder and replaced it with the vision transformer to improve the image recognition performance of the network [23]. Sarker et al. proposed a SLSNet network that reduces the computational cost by using 1-D core factor deep learning networks for sensitive skin lesion segmentation with minimal resources [24]. Wu et al. [25] proposed a new feature adaptive transformer network based on the classical encoder-decoder architecture called FAT-Net.

The organization of this manuscript is as follows. Section 2 provides information about image augmentation and preprocessing. It also includes information about the proposed method and evaluation metrics. In Section 3, details about the computational analysis of the application are given. In Section 4, we present the proposed model assessment. In this section, the results of the proposed method and other methods are presented comparatively. In Section 5, we analyzed and discussed our results and made various suggestions. Finally, the conclusion is given in Section 6.

## 2. Materials and Methods

*2.1. Preparing Dataset.* The proposed model was trained and tested in this study using the ISIC 2018 Lesion Boundary Segmentation dataset. The dataset used is provided by the International Skin Imaging Cooperation (ISIC) archive [12, 26]. The images in the dataset consist of 8-bit RGB dermoscopic images ranging in size from  $767 \times 576$  to  $6682 \times 4401$ . The dataset consists of 2594 training images obtained from different institutes, including various diagnostic challenges.

*2.2. Data Processing and Augmentation.* Deep learning models perform very poorly on datasets with low samples. A large amount of samples is needed for training these models. The ISIC 2018 dataset used for training the model consists of 2594 images. First, 2594 images in the ISIC 2018 dataset were divided into training (2075) and test (519) sets. Then, the amount of data was increased by applying boundary data augmentation with horizontal and vertical flips, random rotation, random distortion, elastic transformation, and scaling and clipping methods to the training samples obtained. Here, better detection of fine-grained features at the borders of the lesioned region was achieved with the border data augmentation method. In the random rotation method, augmented images are obtained from the original images by rotating the original image horizontally and vertically across

each row and column. With the applied data augmentation methods, 72000 training images were obtained. It is aimed to give more reliable and robust results by preprocessing the training images. Contrast stretching applied in data augmentation made the lesions more prominent. In addition, the sharpening algorithm was applied to the data during the contrast stretching. With sharpening and contrast stretching, the blurred edges of the lesion were made more prominent. All images are resized to  $512 \times 512$  as the images are of different sizes, and the suggested model is uniform. Figure 1 shows examples of preprocessed and enhanced images.

**2.3. Proposed Method.** Milletari et al. [27] proposed Vnet architecture for volumetric and fully convolutional 3D image segmentation. The ET-Vnet 2D model is shown in Table 1.

**2.3.1. The Decoder Stage.** Xavier weight initialization was used for the weight initialization of the model, and ReLU was used as the activation function for the layers [28]. In addition, ADAM was used as the optimizer [29]. In the ET-Vnet 2D model, each convolution layer consists of  $3 \times 3$  convolutions with GN to normalize the features in the channels [30]. The  $512 \times 512 \times 1$  input image is fed to the first block, as shown in Table 1. EGM highlighting low valence features in feature maps and ET-Net architecture with WAM emphasizing high valence features were proposed by Zhang et al. [31] for high-performance organ segmentation. The EGM and WAM modules of the proposed architecture have been applied to V-Net 2D. Here, the EGM module emphasizes low-value edge features in segmentation, while the WAM module emphasizes high-value features, allowing the proposed model to perform better segmentation. Also, the convolution layers doubled from 32 to 512 per block [32].

**2.3.2. The Encoder Stage.** The decoder stage takes only high-value features from the encoder and upsamples them to the input size in the decoder. As shown in Table 1, deconvolution, which upsamples feature maps from the down-sampling stage, is performed. A series of convolutional operations are applied at the encoder stage to enlarge the feature map from the encoder to the original low-resolution predicted input image. It is aimed to achieve higher performance segmentation of the network by synchronizing and comparing the decoder and encoder, thanks to the bypass connections from the encoder. In the decoder stages of the proposed model, the same convolution blocks used in the encoder are used in reverse. ReLU activation function is used in each layer, as in the encoder. In the last layer,  $1 \times 1$  convolution using the sigmoid activation function is used, since black and white image segmentation includes two classes. Figure 2.

**2.4. Evaluation Metrics.** The metrics used to score the performance of the models in the ISIC 2018 challenge competition were also used to test the performance of the

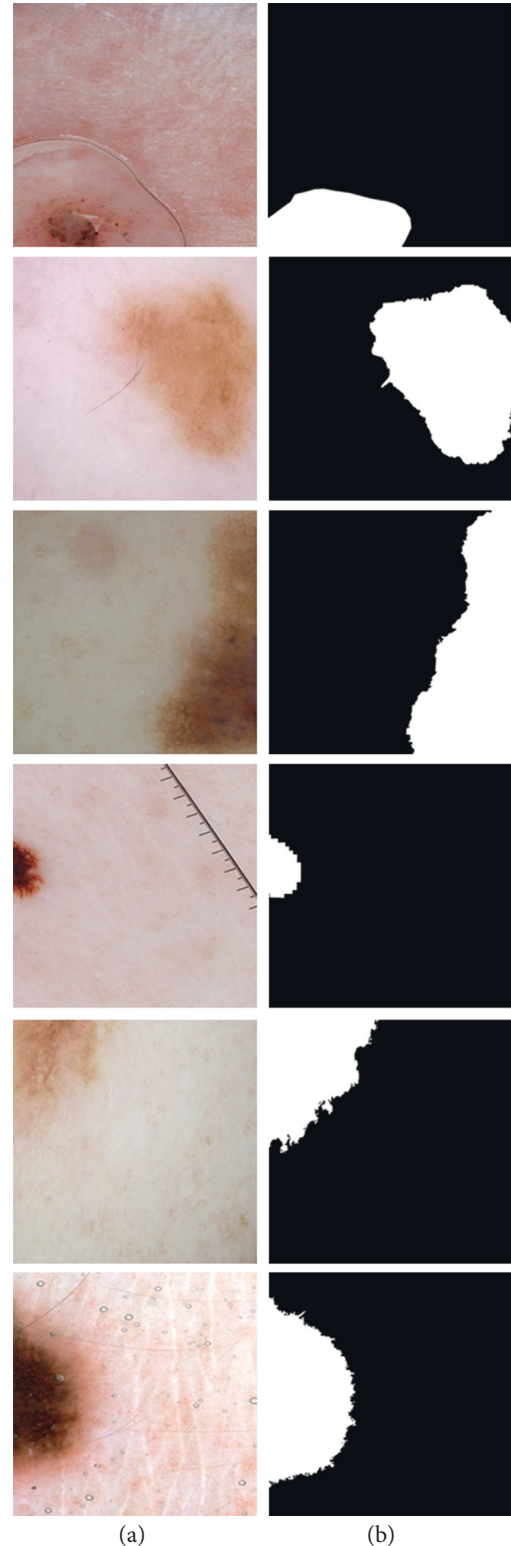


FIGURE 1: Image boundary augmentation. (a) Input image (b) Ground truth.

proposed model. The first of these metrics is the Sorensen-Dice coefficient (DSC), shown in equation (1). Dice is a simple measure that measures the similarity of two samples. The Jaccard index (Jaccard) is used to calculate the similarity

TABLE 1: ET-Vnet 2D models' details. Inputs (batch size, height, width, depth, Channels): (4,512,512,1,1).

Layer	Layers' details	Output size	Activate/Norm/Weight initializer
Layer 1	Conv $3 \times 3$ ResNet block	(4,512,512,1,32)	GN, ReLU, Xavier
Down 1	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,256,256,1,64)	GN, ReLU, Xavier
Layer 2	ResNet block		GN, ReLU, Xavier
Down 2	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,128,128,1,128)	GN, ReLU, Xavier
Layer 3	ResNet block		GN, ReLU, Xavier
Down 3	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,64,64,1,256)	GN, ReLU, Xavier
Layer 4	ResNet block		GN, ReLU, Xavier
Down 4	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,32,32,1,512)	GN, ReLU, Xavier
Layer 5	ResNet block		GN, ReLU, Xavier
Upsample 1	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,64,64,1,256)	GN, ReLU, Xavier
Layer 6	ResNet block		GN, ReLU, Xavier
Unsample 2	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,128,128,1,128)	GN, ReLU, Xavier
Layer 7	ResNet block		GN, ReLU, Xavier
Upsample 3	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,256,256,1,64)	GN, ReLU, Xavier
Layer 8	ResNet block		GN, ReLU, Xavier
Upsample 4	Filter $3 \times 3$ Filter $3 \times 3$ Filter $3 \times 3$	(4,512,512,1,32)	GN, ReLU, Xavier
Layer 9	ResNet block		GN, ReLU, Xavier
Layer 10 3D single output	ET-Net2D Filter $1 \times 1$	(4,512,512,1,1) + ET-Net2D	GN, ReLU, Xavier

and diversity of sample sets, as defined in equation (2). We can also define Jaccard as the intersection of the union of two different sets. Accuracy (ACC), shown in equation (3), represents the percentage of correct predictions out of all predictions made by the model. Sensitivity (Sens), defined in equation (4), measures the proportion of samples predicted as true positive (TP) in the data set. Also, sensitivity can be defined as recall. As shown in equation (5), specificity (Spec) is a metric that calculates the proportion of correctly predicted nonlesional areas in an image. One drawback of the dice coefficient (Dice) is that false positive (FP) and false negative (FN) are equally weighted. While this increases the precision, it decreases the recall rate.

For this reason, Dice causes degradation in test performance in unbalanced datasets. The way to deal with this situation is to weigh FN more than FP. The Tversky similarity index is a generalization of the Dice coefficient that provides flexibility to its problems in balancing FPs and FNs as shown in equation (6).

$$\text{Dice} = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}}, \quad (1)$$

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

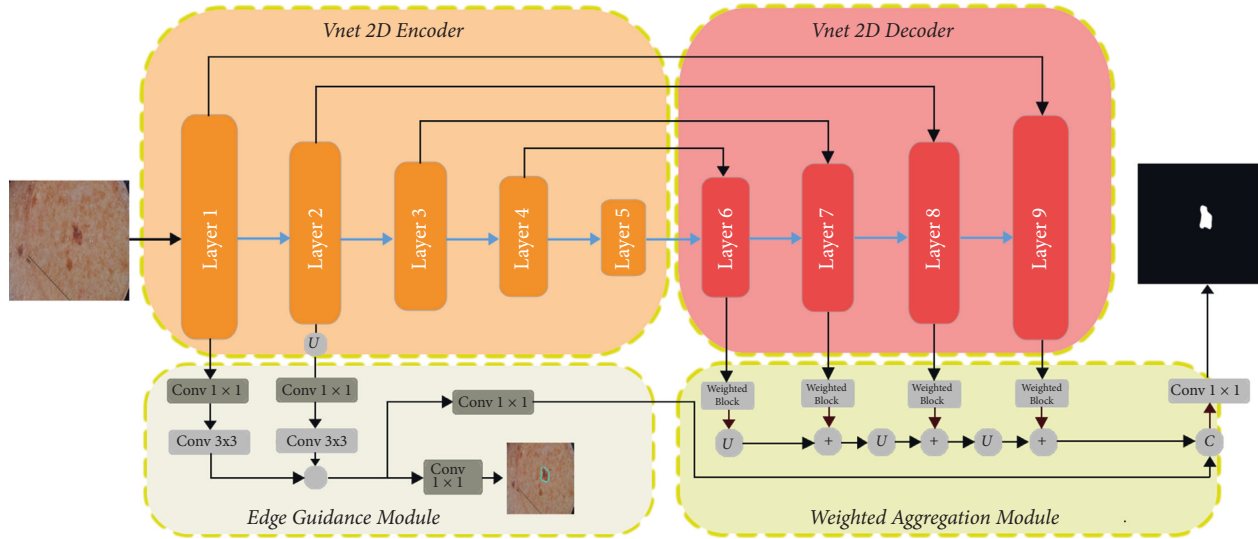


FIGURE 2: ET-Vnet 2D models' block diagram.

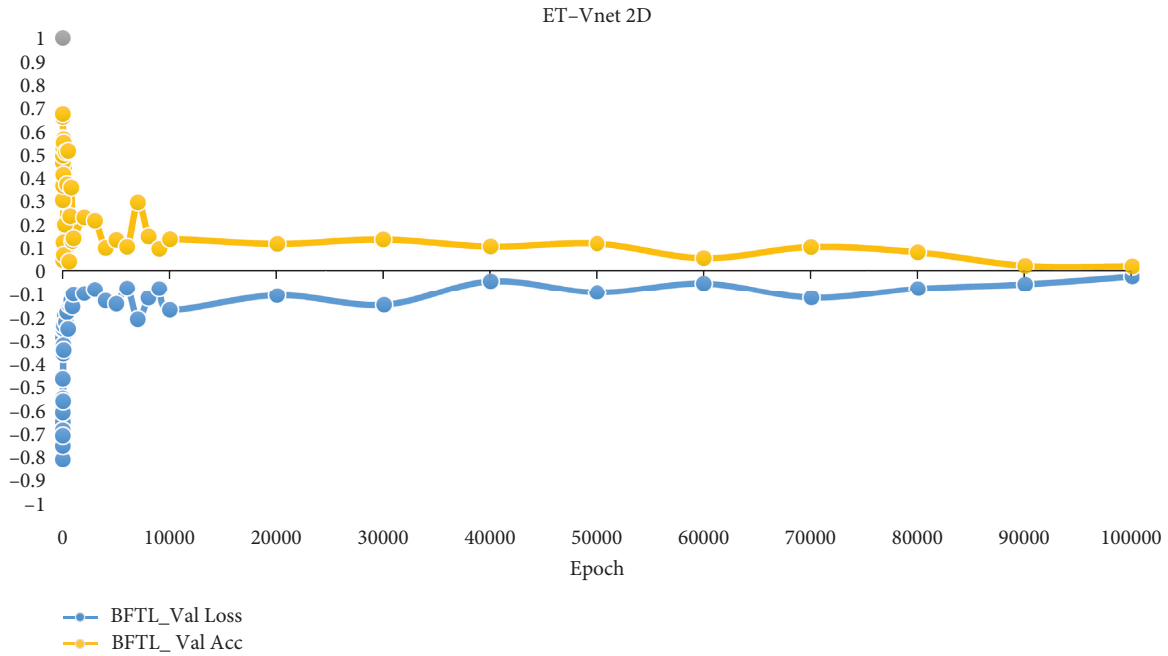


FIGURE 3: ET-Vnet 2D validation loss and validation accuracy results.

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

$$Sens = \frac{TP}{TP + FN}, \quad (4)$$

$$Spec = \frac{TN}{TN + FP}, \quad (5)$$

$$TverskyIndex = \frac{TP}{TP + \alpha FN + \beta FP}, \quad (6)$$

where True Positive (TP) represents the correctly labeled lesion pixels, False Positive(FP) represents the incorrectly

predicted lesion pixels, True Negative(TN) represents the correctly predicted nonlesion pixels, and False Negative(FN) represents the incorrectly predicted lesion pixels. In equation (6),  $\alpha$  and  $\beta$  are adjustable parameters to increase the weight of the recall rate in unbalanced datasets.

**2.5. Balanced Focal Tversky Loss Function (BFTL).** The BFTL loss function proposed in this study is developed from the loss function proposed by Zhou et al. [33]. The FTL loss function is explained in equation (7) which effectively solved the problems of Dice loss in class imbalances [34, 35]. The FTL function shown in equation (8) is presented as a solution to data imbalances. In the proposed model, a loss

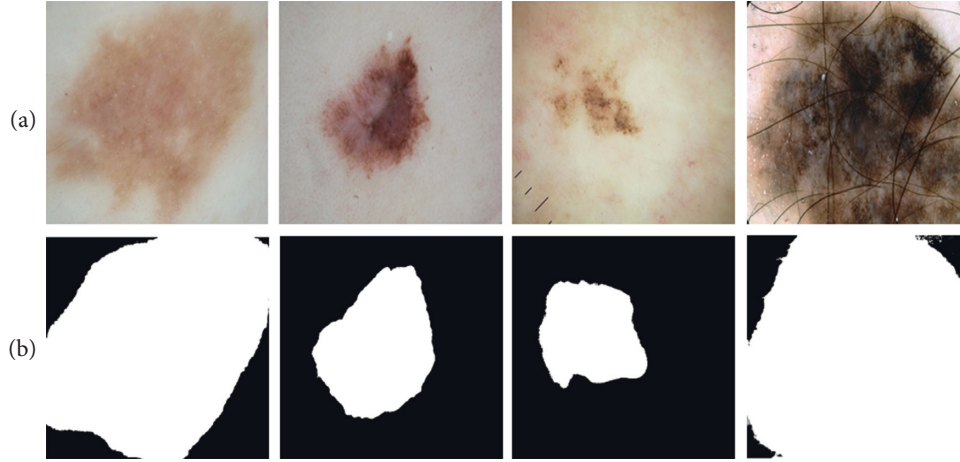


FIGURE 4: Test results for ISIC 2018 challenge 1000 test set.

TABLE 2: Performance scores of the proposed model with the latest models in the literature.

Methods	Performance metrics				
	Dice	Jaccard	Acc	Sens	Spec
DA-net	0.88	0.80	0.95	0.94	0.95
DRU-net [37]	0.86	0.76	—	0.88	0.92
Deeplabv3+	0.87	0.79	0.95	0.94	0.94
CKDNet [38]	0.88	—	0.95	0.91	<b>0.97</b>
DAGAN [24]	0.89	0.83	0.93	0.95	0.91
FAT-net [31]	0.89	—	<b>0.96</b>	0.91	<b>0.97</b>
Attn_U-net + GN [18]	0.91	0.83	0.95	0.94	0.95
SE_U-net [18]	0.91	0.83	0.95	0.89	0.96
ET-vnet (ours)	<b>0.91</b>	<b>0.83</b>	0.95	<b>0.95</b>	0.96

function is proposed that both take care of pixel type losses and solve data imbalances.

$$\text{FTL} = (1 - TI) \frac{1}{\gamma}, \quad (7)$$

$$\mathcal{L} \text{ Twersky} = \sum_{c=1}^{Nb} \left( 1 - \frac{\sum_{i=1}^N \text{pic}_{gic} + \epsilon}{\sum_{i=1}^N \text{pic}_{gic} + \alpha \sum_{i=1}^N \text{pic}_{gic} + \beta \sum_{i=1}^N \text{pic}_{gic} + \epsilon} \right)^{1/\gamma}. \quad (8)$$

The Tversky index is adapted to a loss function (TL) in [36] by minimizing  $\sum_c 1 - TI_c$ .

The BFTL function  $L$  is formulated in equation (9).

$$\mathcal{L} = \mathcal{L}bbce + \log(\mathcal{L} \text{ Twersky}), \quad (9)$$

where  $\mathcal{L}bbce$  is the BBCE [37], unlike binary cross-entropy,  $\beta$  weight is added to BBCE. Where  $\beta$  is the number of negative samples/total number of samples. In other words,  $\beta$  is the proportion of the dominant sample in a data set.  $1 - \beta$  denotes the fraction of the other class. In addition, the

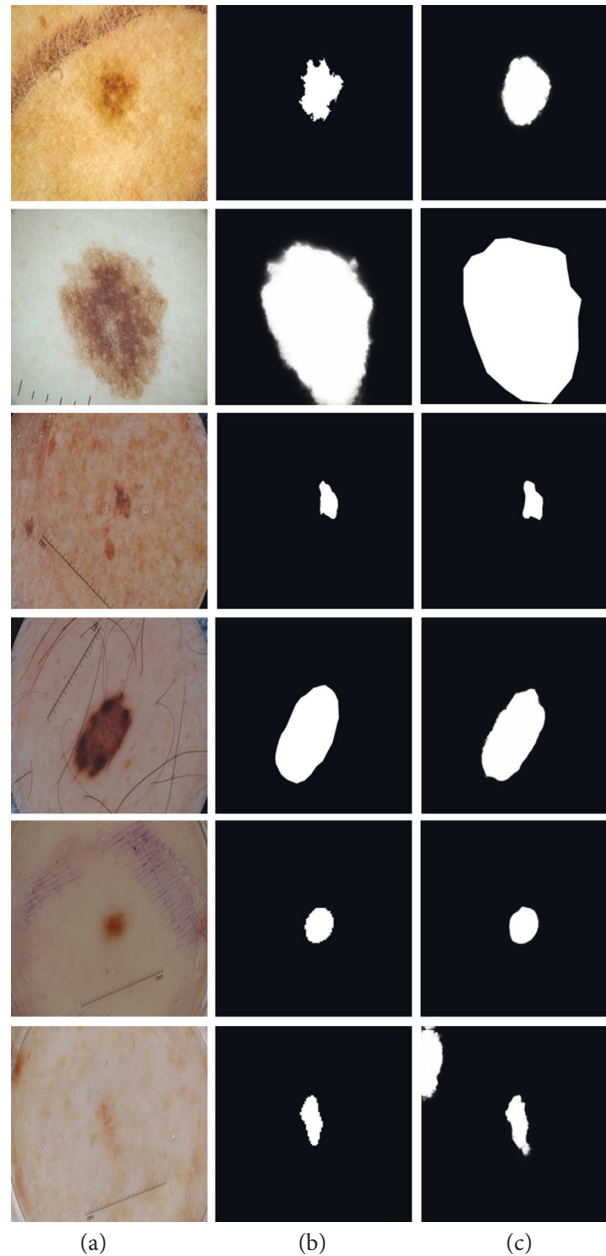


FIGURE 5: Test results are based on 519 test sets obtained from the ISIC 2018 training set. (a) Input image. (b) Image mask. (c) Predicted mask.

natural logarithm is applied to the FTL function to ensure consistency between the loss functions.

### 3. Calculation

**3.1. Application Details.** A computer with Intel I5–8300H processor and 32 GB of RAM and Nvidia GTX 1080 ti graphics card was used to train the proposed model. Besides, the computer's operating system is Windows 10–64 bit. The proposed deep learning model was created using the Python 3.6 programming language. In the proposed model, the learning rate was determined as  $1e-3$ , and the lot size = 4 due to the experiments. The model's training has been realized as 100,000 epochs, thanks to 72000 increased training images. Vnet 2D and ResUnet 2D models were separately trained in

the increased data set from the ISIC 2018 data set. The graphical analysis of the tests performed is shown in Figure 3.

The images of some samples from the proposed models' test results, their true lines, and the predicted segmentation result are shown in Figure 4. As can be seen from Picture 4, when the predicted results are compared with the actual lines, it is seen that the model exhibits very high performance.

**3.2. Qualitative Analysis.** The proposed model's segmentation and final results are shown in Figure 4. The figure demonstrates images obtained from the model with BBFTL loss function. Each column of the figure shows a unique introductory dermoscopic view. The proposed model efficiently analyses input images, mainly obtained through

boundary data augmentation, and predicts complex lesion boundaries with great precision.

**3.3. Hardware Analysis.** The proposed model has been compared to other cutting-edge models for many model parameters, storage requirements, and extraction rates using Nvidia Geforce GTX 1080 ti graphics. In addition, since the data set is stable, there is no need for pretraining scoring. The training of the model took approximately 8 hours. In addition, the model's estimation of the test data set took 52 seconds.

## 4. Prediction Results

The proposed segmentation model was tested on ISIC 2018 Task 1 dermoscopic lesion images. The BBFTL loss function combining BBCE and FTL function is used as a new approach. One of the main factors in applying the border data augmentation method is to enable the model to recognize better the lesion borders that separate the lesioned area from the skin.

**4.1. Comparative Analysis with the Latest Models.** Table 2 shows the comparative results of the proposed model with other models in the ISIC 2018 dataset. The proposed model surpassed the latest literature methods in the tests in the 2018 dataset, scoring 0.91 points in the Dice score, 0.83 points in the Jaccard score, and 0.95 points in the sensitivity, and the other models achieved superior performance.

Figure 4 shows the comparative results of the proposed model with other models in the ISIC 2018 dataset.

Figure 5 shows the visually estimated output of some of the complex samples in the 519 test images separated from training with the proposed approach. The figure compares the basic reality and predicted image based on the original input image. Each row in the figure shows the test results of a test input image in the data set. The third column shows the segmentation result predicted by the proposed fusion model of the test input image presented in the first column.

## 5. Discussion

In the proposed model, we get the ET-Vnet 2D network by fusing the EGM, WAM, and Vnet 2D networks from the ET-Net 2D network. In addition, the ET-Vnet 2D network training was completed in 8 hours. EGM and WAM from the ET-Net 2D network significantly improved the performance of the Vnet 2D network. Also, in the proposed model, balanced binary cross-entropy and focal Tversky loss functions are hybrids combined. The proposed hybrid loss function caused a slight prolongation of the training time. But it enabled the model to give more robust results.

## 6. Conclusion

In this study, EGM and WAM modules, which are two modules used in the ET-Net network, are combined with Vnet 2D to create the proposed model—tested on ET-Vnet

2D ISIC 2018 Lesion Boundary segmentation dataset. The proposed model consists of data set preprocessing and data augmentation, lesion identification, and prediction operations. In the dataset preprocessing stage, artifacts such as color inconsistency, exposure problem, and visibility of lesion borders are corrected. The visualization of a limited number of labeled skin lesions has been achieved sufficiently to train the proposed model, mainly thanks to the border and other data augmentation methods. In addition, a new loss function is proposed by combining BBCE, which calculates pixel type losses, and FTL, which is presented as a solution to class imbalances. The loss function proposed as a novel approach significantly improves the performance of the proposed model. The proposed hybrid loss function is thought to play a crucial role in challenging segmentation tasks. The data set's Dice and Jaccard similarity metrics were recorded as 0.95 and 0.83, surpassing the latest segmentation techniques. The proposed model will be tested in other organ segmentations to prove its robustness in future studies.

## Data Availability

We can send the datasets at the request of the authors.

## Ethical Approval

This article does not contain any studies with human participants. No animal studies were involved in this review.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

All authors contributed equally to this work. In addition, all authors have read and approved the final manuscript and given their consent to publish the article.

## References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics," *CA: A Cancer Journal of Clinicians*, vol. 71, 2021.
- [2] G. P. Guy, S. R. Machlin, D. U. Ekwueme, and K. R. Yabroff, "Prevalence and costs of skin cancer treatment in the US, 2002-2006 and 2007-2011," *American Journal of Preventive Medicine*, vol. 48, no. 2, pp. 183-187, 2015.
- [3] S. V. Mohan and A. L. S. Chang, "Advanced basal cell carcinoma: epidemiology and therapeutic innovations," *Current Dermatology Reports*, vol. 3, no. 1, pp. 40-45, 2014.
- [4] H. K. Koh, A. C. Geller, D. R. Miller, T. A. Grossbart, and R. A. Lew, "Prevention and early detection strategies for melanoma and skin cancer. Current status," *Archives of Dermatology*, vol. 132, no. 4, pp. 436-443, Article ID 8629848, 1996.
- [5] M. E. Vestergaard, P. Macaskill, P. E. Holt, and S. W. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669-676, 2008.
- [6] K. Hu, J. Lu, D. Lee, D. Xiong, and Z. A. S.-N. Chen, "AS-Net: attention Synergy Network for skin lesion segmentation,"



- Expert Systems with Applications*, vol. 201, Article ID 117112, 2022.
- [7] N. Daldal, M. Nour, and K. Polat, "A novel demodulation structure for quadrature modulation signals using the segmentary neural network modelling," *Applied Acoustics*, vol. 164, Article ID 107251, 2020.
- [8] N. Daldal, A. Sengur, K. Polat, and Z. Cömert, "A novel demodulation system for base band digital modulation signals based on the deep long short-term memory model," *Applied Acoustics*, vol. 166, Article ID 107346, 2020.
- [9] N. Daldal, Z. Cömert, and K. Polat, "Automatic determination of digital modulation types with different noises using Convolutional Neural Network based on time - frequency information," *Applied Soft Computing*, vol. 86, Article ID 105834, 2020.
- [10] M. Nour, N. Daldal, M. F. Kahraman, H. Sindi, A. Alhudhaif, and K. Polat, "A novel tilt and acceleration measurement system based on Hall-effect sensors using neural networks," *Mathematical Problems in Engineering*, vol. 2022, Article ID 7000486, 13 pages, 2022.
- [11] M. F. Kahraman and S. Öztürk, "Experimental study of newly structural design grinding wheel considering response surface optimization and Monte Carlo simulation," *Measurement*, vol. 147, Article ID 106825, 2019.
- [12] N. C. F. Codella, D. Gutman, M. E. Celebi 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 *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, April 2018.
- [13] S. A. Mahdiraji, Y. Baleghi, and S. M. Sakhaei, "Skin lesion images classification using new color pigmented boundary descriptors," in *Proceedings of the 2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA)*, pp. 102–107, Shahrekord, Iran, April 2017.
- [14] J. Czajkowska, P. Badura, S. Korzekwa, and A. Płatkowska-Szczerek, "Automated segmentation of epidermis in high-frequency ultrasound of pathological skin using a cascade of DeepLab v3+ networks and fuzzy connectedness," *Computerized Medical Imaging and Graphics*, vol. 95, Article ID 102023, 2023.
- [15] R. Arora, B. Raman, K. Nayyar, and R. Awasthi, "Automated Skin Lesion Segmentation Using an Attention-Based Deep Convolutional Neural Network," *Biomedical Signal Processing and Control*, vol. 65, 2020.
- [16] Y. Wu and K. He, "Group normalization," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, Munich, Germany, March 2018.
- [17] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, <http://www.arxiv.org/abs/1502.03167>.
- [18] R. Gu, G. Wang, T. Song et al., "CA-net: comprehensive attention convolutional neural networks for explainable medical image segmentation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 699–711, 2021.
- [19] 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, 2020.
- [20] Y. Jiang, S. Cao, S. Tao, and H. Zhang, "Skin lesion segmentation based on multi-scale Attention convolutional neural network," *Skin lesion segmentation based on multi-scale Attention convolutional neural network*, vol. 8, 2020.
- [21] B. Lei, Z. Xia, F. Jiang, X. Jiang, Z. Ge, and Y. Xu, "Skin Lesion Segmentation via Generative Adversarial Networks with Dual Discriminators," *Medical Image Analysis*, vol. 64, 2020.
- [22] P. Shan, Y. Wang, C. Fu, W. Song, and J. Chen, "Automatic skin lesion segmentation based on FC-DPN," *Computers in Biology and Medicine*, vol. 123, Article ID 103762, 2020.
- [23] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, and T. Unterthiner, "An image is worth 16x16 words: transformers for image recognition at scale," in *Proceedings of the International Conference on Learning Representations*, OpenReview.net, Canada, April 2021.
- [24] M. M. K. Sarker, H. A. Rashwan, F. Akram et al., "SLSNet: skin lesion segmentation using a lightweight generative adversarial network," *Expert Systems with Applications*, vol. 183, Article ID 115433, 2021.
- [25] 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, Article ID 102327, 2022.
- [26] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, and D. Gutman, "Skin Lesion Analysis toward Melanoma Detection 2018: A challenge Hosted by International Skin Imaging Collaboration (ISIC)," 2019, <https://arxiv.org/pdf/1902.03368>.
- [27] F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image," in *Proceedings of the 2016 Fourth International Conference on 3D Vision*, California, CA, USA, October 2016.
- [28] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *Aistats*, vol. 9, pp. 249–256, 2010.
- [29] J. Ba and D. P. Kingma, "Adam: a method for stochastic optimization," in *Proceedings of the International Conference on Learning Representations*, pp. 1–11, ICLR, San Diego, CA, USA, May 2015.
- [30] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," 2016, <https://arxiv.org/pdf/1607.08022>.
- [31] Y. Zhang, J. Wu, W. Chen, Y. Chen, and X. Tang, "Prostate segmentation using Z-net," in *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 11–14, Venice, Italy, April 2019.
- [32] Q. Jin, H. Cui, C. Sun, Z. Meng, and R. Su, "Cascade knowledge diffusion network for skin lesion diagnosis and segmentation," *Applied Soft Computing*, vol. 99, Article ID 106881, 2021.
- [33] W. Zhou, X. Tao, Z. Wei, and L. Lin, "Automatic segmentation of 3D prostate MR images with iterative localization refinement," *Digital Signal Processing*, vol. 98, Article ID 102649, 2020.
- [34] N. Abraham and N. M. Khan, "A novel focal Tversky loss function with improved attention U-net for lesion segmentation," in *Proceedings of the 2019 IEEE 16th International*

- Symposium on Biomedical Imaging (ISBI 2019)*, pp. 683–687, IEEE, Venice, Italy, April 2019.
- [35] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, 2020.
- [36] P. Wang and A. C. S. Chung, “Focal dice loss and image dilation for brain tumor segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, Cham, Switzerland, pp. 119–127, 2018.
- [37] H. Jiang, T. Shi, Z. Bai, and L. Huang, “AHCNet: an application of attention mechanism and hybrid connection for liver tumor segmentation in CT volumes,” *IEEE Access*, vol. 7, Article ID 24898, 2019.
- [38] M. Jafari, D. Auer, S. Francis, J. Garibaldi, and X. Chen, “DruNet: An Efficient Deep Convolutional Neural Network for Medical Image Segmentation,” in *Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA, April 2020.