International Journal of Artificial Intelligence
p-ISSN: 2407-7275, e-ISSN: 2686-3251

Original Research Paper

A Novel Sep-Unet Architecture of Convolutional Neural Networks to Improve Dermoscopic Image Segmentation by Training Parameters Reduction

Faezeh Sadeghi[1*], Mohammad Taheri[2], Maryam Rastgarpour[3], Arash Sharifi[1]

[1] *Department of Computer Engineering, Science and Research Branch, Islamic Azad University. Tehran, Iran.*

[2] *Department of Computer Engineering, Faculty of Engineering, Damghan University. Damghan, Iran.*

[3] *Department of Computer Engineering, Faculty of Engineering, Saveh Branch, Islamic Azad University. Saveh, Iran.*

Article History
Received:
03.07.2022

Revised:
21.07.2022

Accepted:
02.08.2022

***Corresponding Author:**
Faezeh Sadeghi
Email:
faezeh.sadeghi@srbiau.ac.ir

Abstract: Nowadays, we use dermoscopic images as one of the imaging methods in diagnosis of skin lesions such as skin cancer. But due to the noise and other problems, including hair artifacts around the lesion, this issue requires automatic and reliable segmentation methods. The diversity in the color and structure of the skin lesions is a challenging reason for automatic skin lesion segmentation. In this study, we used convolutional neural networks (CNN) as an efficient method for dermoscopic image segmentation. The main goal of this research is to recommend a novel architecture of deep neural networks for the injured lesion in dermoscopic images which has been improved by the convolutional layers based on the separable layers. By convolutional layers and the specific operations on the kernel of them, the velocity of the algorithm increases and the training parameters decrease. Additionally, we used a suitable preprocessing method to enter the images into the neural network. Suitable structure of the convolutional layers, separable convolutional layers and transposed convolution in the down sampling and up sampling parts, have made the structure of the mentioned neural network. This algorithm is named Sep-unet and could segment the images with 98% dice coefficient.

Keywords: Convolutional Layer, Convolutional Neural Network, Deep Neural Network, Segmentation, Separable Convolution.

1. Introduction

Automatic medical image segmentation is widely used in diagnosis and treatment of various disease [1]. One of these cases is skin images, which are widely used in the diagnosis and treatment of skin lesions, such as skin cancer.

Over the past years, various methods for medical image segmentation have been proposed by researches which according to the subject and the amount of processing required for dataset [2], we divide them in to four categories: low level, intermediate level, high level and advanced level. Advanced level algorithms use neural networks and deep neural networks. In this study, we have used a new architecture of convolutional neural networks for dermoscopic lesion segmentation. The main purpose of this research is accurate medical image segmentation using convolutional neural networks with fewer parameters. One of the important criteria in the neural network that affects its accuracy, is the hyperparameters. The hyperparameters include the number of convolution layers, the number of filters, the size of filters, stride, padding type, the number of epochs, dropout, the number of training parameters, the number of hidden layers and units, network weight initialization, activation function, learning rate, momentum, and batch size. Setting the parameters correctly can be a time- consuming process as it requires experience in its previous experiments and many comparisons are needed for the analysis. Numerous articles have used greedy search and random search algorithms to adjust the hyper parameters, which are among the meta heuristic algorithms in the category of strategies [3]. In this study, the main goal is to create an efficient method for medical image segmentation using CNN with initial desirable speed and accuracy. Reduce training parameters while maintaining the suitable accuracy. In the field of image processing, one of the practical and most used widely networks is deep convolutional neural network which inspired by the great success of this network, it has been used in image segmentation, including medical image segmentation [4]. This network is a successful network to solve many problems such as image classification, object recognition [5], image segmentation [6], feature extraction [7], face recognition [8], and human pose estimation [9]. Convolutional neural network (CNN) is one of the advanced image processing methods. The advantage of this method is its simplicity and high velocity. These networks utilize supervised classifiers to segment images. CNN is one of the key algorithms in machine vision science and has been achieved remarkable successes to solve real world problems. This paper is organized as following. Firstly, we discuss about the previous works in section II, then we introduce the details of the proposed algorithm in section III, then we report the experimental results and design in section IV, and eventually, we conclude the study in section V.

2. Literature Review

In recent years, many studies have been conducted in the field of medical image segmentation using artificial intelligence and computer vision algorithms that we have mentioned some of them in the following. According to Tan [10] a PSO algorithm is proposed and it is applied on two methods of segmentation, fuzzy clustering and deep convolutional neural networks and the field of research is medical imaging. These search algorithms work like a cascade, and this is not only allowed them to try out all the search algorithms, but also allows them to run multiple algorithms in parallel at the same time in each iteration. The other method was presented that implements evolutionary algorithm on the hyper parameters of convolutional neural network by using mnist dataset [11], to produce more desirable results. The researchers applied a multi objective optimization algorithm [12] to find the right kernel size. After applying the correction method, they compare the performance of the algorithm with the main method and believe that it is the most effective method and the most important parameters are hyper parameters. Another research pointed out the importance of laryngeal cancer and announced that up to year 2030, this cancer will have the most common statistics, aiming to isolate the saliva created from the background image [13]. To do this task, used the Bayesian classification algorithm. The importance of medical image segmentation has been explained in [14]. The challenge is that medical images can be very large (1024*1024) or very small (25*25). Also, sometimes the images are general and large, but we are looking for a small damaged lesion. To examine the images more closely, the researchers designed a kernel and compared their proposed kernel to a conventional kernel. Another author tried to increase the accuracy of the U-Net neural network and strengthen the data [15]. Also, in the Dropout by normalizing the categories, added a pixel-by-pixel suggestion to the image pixel to avoid overfitting. This network is used to process mammographic images. An author introduced a network which consists of subnets [16]. The output of the first network is the input of the next network. It has a kernel and two feature maps. This network is designed as a cascade. It examines the number and depth of neurons with themselves. NAS neural network was first introduced for image classification, and then the scientists and researchers also used this neural network for image segmentation [17]. This network implements UpSC and DownSC parts with u-shaped structure and has been introduced as NAS_UNET. This network has less training parameters than the primary Unet neural network.

In a study, the authors used a CNN for iso intense stage brain tissues segmentation by the multi-modality MRI brain images [18]. The other networks such as Livia-NET and Hyper-Dense-NET neural networks were used for MRI images of 6-month-old infants [19]. An article provided an overview of magnetic resonance imaging techniques focusing on the architecture, preprocessing, data preparation, and post processing strategies which are provided in these works [20]. The purpose of this research is in 3 parts. The main purpose of the report is how different CNN architectures evolve. It discusses state of the art strategies and examines their advantages and disadvantages. Other researchers have used a multi-scale residual and decoding method for skin lesion segmentation [21]. This method can segment a number of dermoscopic images safely and precisely. Additionally, there is a multi-resolution, multi-channel feature fusion module in this method which replaces the conventional encoding-decoding convolutional neural networks. The authors in [22] have made major improvements in Unet structure to improve its capacity for skin cancer segmentation. These improvements have been implemented in both encoding and decoding paths. Instead of using standard convolutional layers like the Unet neural network, the encoding path in the proposed network consists of 10 standard convolutional layers which are inspired from the Visual Geometry Group (VGG16) followed by a pyramid pooling module and a dilated convolutional block.

3. Methodology

Inspired by the deep encoder-decoder neural networks, we finally introduce the architecture of the proposed method to solve the problem ahead. In the encoding section, the input data is mapped to the feature space, and in the decoding section, they return from feature space to their original state. In fact, the main part of a deep encoder-decoder neural network is the hidden layer which is used as the extracted feature for segmentation. As preprocessing, we do normalization and divide each pixel by 256. The reason for doing this is that, the pixels have the same effect on the operation process and the proportions are balanced. In the next step, we divide the preprocessed data to train data, validation data, and test data. we have allocated 50% of dataset to training data, equally 25% to test data and 25% to validation data. In the last part, the data enter to the network, after down sampling and up sampling in the encoder and decoder sections respectively, the final segmented images will be as the output of network. The evaluation will be done based on the test data and the final segmented outputs. Although, the architecture of this algorithm looks similar to encoder-decoder algorithm, there are some differences that we will examine below. In the proposed Sep-unet neural network, we are going to reduce the number of training parameters and training time while improve the evaluation metrics. We have displayed a brief diagram of the proposed method in Figure 1.

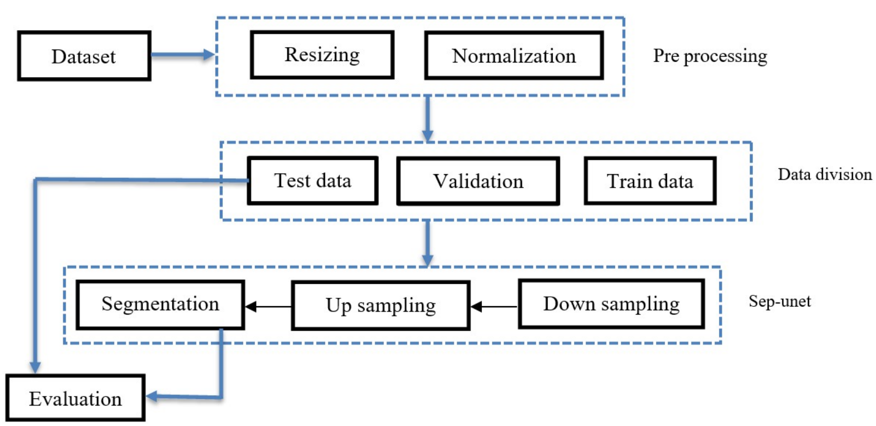

Figure 1. Diagram of the Proposed Method

In the training step, zero padding is used as a technique to not to reduce the image information. Activation function is relu which protects the algorithm from vanishing gradient. Adam optimization is used in place of the classical stochastic gradient descent to update the network's weights more efficiently. Separable convolution is a special type of convolution that by changing the dimensions of the kernels, will decrease the volume of calculation and increase the time of implementation. The batch normalization after each layer, accelerates convergence and it can reduce vanishing gradient too. Max pooling reduces the features of convolutional layers. In order to reach the best features, we deepen the network to get the most important features. This process has been repeated four times to deepen it and the purpose of this part is to extract the features which in our case means to separate the lesion from the background. The architecture of the purposed method is in Figure 2.

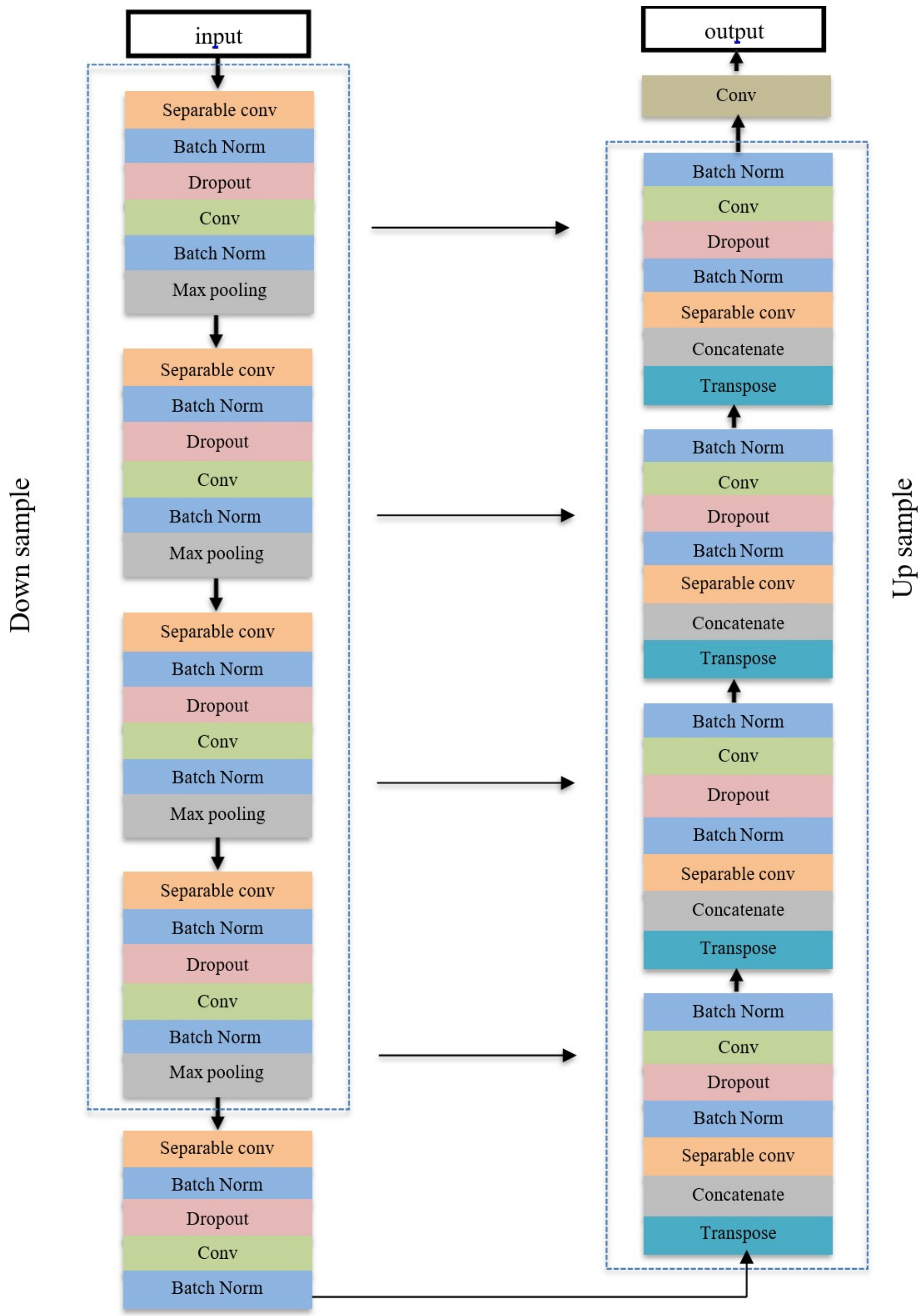

Figure 2. The Architecture of Sep-Unet Deep Convolutional Neural Network

In the proposed network, there are Down sampling and Up sampling parts which are clarified clearly in the following section.

3.1. Down Sampling

The background of medical images can create lots of redundancy in the image which is not required. Down sampling is one of the most efficient feature extraction procedures in supervised algorithms. In all of the convolutional layers, a value of 0 is placed around the images as zero padding so as not to reduce the information of images. The activation function is relu that bypasses the gradient fade. Separable convolution by changing the dimensions of the kernels, will reduce the volume of calculations and accelerate the process of the algorithm. There is batch normalization after each convolution layer that accelerates convergence and also the problems such as gradient fading less occurs. The main goal of down sampling section is that the algorithm learns how to identify the desired features. The feature vector is stored as a compressed vector in a block after down sampling and is used as input for up sampling section. As mentioned before, the main goal of this part is that, the algorithm learns how to compress the desired information.

3.2. Up Sampling

The main goal of this part is to restore the image to its original size and resolution. The deleted information is returned to the images and the feature map is applied to the images step by step. The transpose convolution layers are the reverse of convolution layers in the down sampling part. To restore the image to its original size, we have repeated the operation four times.

3.4. Experimental Design

3.4.1. Dataset

The PH2 dataset [23]is compiled from a joint study between the two research centers of the Porto, Tecnico University, as well as the Pedro and Hispano dermatological hospitals in Portugal. The Tuebinger Mole Analyzer device is used for dermoscopic imaging. These images are 8bit RGB images with 560*768 dimensions. The database contains 200 dermoscopic images, including 80 images of normal wounds, 80 images of deep wounds, and 40 images of melanoma skin cancer. The ISIC dataset [24] contains high quality dermoscopic images of skin lesions. Currently, ISIC contains more than 13000 dermoscopic images which is prepared from prominent international clinical centers and from various devices with different qualities, but we have used its 2017 version with 2000 dermoscopic images.

3.4.2. Performance Evaluation

The output of the proposed algorithm are binary images. The performance of the proposed algorithm is determined by comparing the computer-generated masks as segmented outputs with the binary images as labeled masks in the dermoscopic dataset. We have assessed our algorithm by the following metrics:

3.4.3. Precision

Precision is the ratio of the true positives to all the pixels that were part of the lesion pixels in the image. This criterion is obtained from the following Equation 1.

$$Precision = \frac{TP}{TP+FN} \tag{1}$$

Which TP (True Positive) means that the lesion in the output is detected as lesion correctly. FN (False Negative) means that the pixel in the image is located in the lesion and the algorithm does not mistakenly recognize this pixel as a lesion.

3.4.4. Accuracy

Accuracy means a ratio of negatives that the test correctly marks as negative. Here is the ratio of pixels that the algorithm correctly describes as pixels that do not belong to the lesion. Mathematically, dividing the correct negatives by the sum of correct negatives and false positive.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + PP} \tag{2}$$

TN (True Negative) are the pixels in the image that are not in the lesion and the algorithm does not correctly identify these pixels as a lesion. FP (False Positive) means that the pixels are detected mistakenly as a lesion.

3.4.5. Recall

The focus of this criterion is on pixels that are a lesion and the lesion is correctly diagnosed and seeks to cover the lesion correctly in the whole data. This metric is calculates as following:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

This metric seeks to find a balance between the accuracy and precision. We can use it in cases where the FN and FP are different from each other.

$$F1 = \frac{2 * precision * recall}{precision + recall} \tag{4}$$

3.4.6. Dice Coefficient

It is one of the most common metrics to assess the performance of the image segmentation algorithms. It is calculated from the following formula:

$$DSC(A, B) = 2\frac{|A \cap B|}{|A| \cup |B|} \tag{5}$$

Which A, B are the reference mask and resulted mask of segmentation algorithm, respectively. The larger value of the dice coefficient shows the better performance of the algorithm. The upper limit of this amount is 1, because when A and B exactly match each other, the value of the dice is equal to 1.

4. Finding and Discussion

We have displayed the final results of our algorithm in Figure 3 and Figure 4. The proposed method in this study is a robust method toward dermoscopic image segmentation and can segment the lesions with no blurry borders which dermatologists can detect the lesions precisely.

We have implemented the proposed algorithm on two famous dermoscopic datasets, including PH2 and ISIC2017 to segment the skin lesions that can help the physicians to detect the skin better and precisely. We used Google colab [25] to implement our algorithm and the final results have been implemented in matlab R2018b. The numerical results of the proposed algorithm on PH2 dataset have been displayed in Table 1.

Table 1. Numerical Results Based on Evaluation Metrics for PH2 Dataset

Method	**Accuracy**	**Precision**	**Recall**	**F Measure**	**Dice Coefficient**
Sep-Unet	0.96	0.95	0.97	0.95	0.96

We have implemented the proposed algorithm on two famous dermoscopic datasets, including PH2 and ISIC2017 to segment the skin lesions that can help the physicians to detect the skin better and precisely. We used Google colab to implement our algorithm and the final results have been implemented in matlab R2018b. The numerical results of the proposed algorithm on PH2 dataset have been displayed in Table 1.

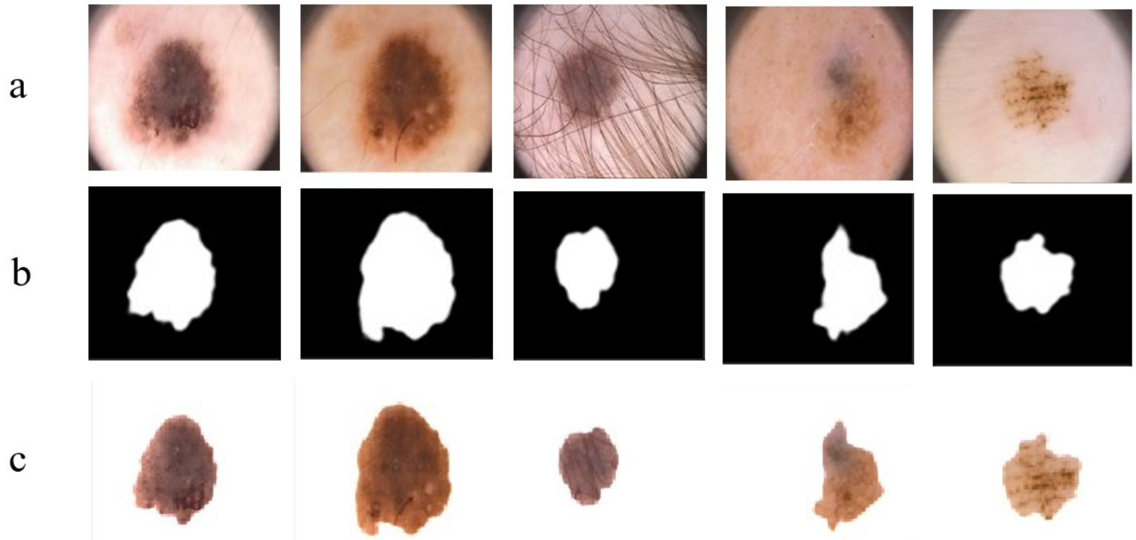

Figure 3. The Results of Implementation on PH2 Dataset
(A) The Samples of PH2 Skin Lesions
(B) The Segmentation Results for Each Sample
(C) The Final Skin Lesion Segmentation.

Additionally, we have displayed the numerical results of the proposed algorithm on ISIC2017 in Table 2.

Table 2. Numerical Results Based on Evaluation Metrics for ISIC2017 Dataset

Method	**Accuracy**	**Precision**	**Recall**	**F Measure**	**Dice Coefficient**
Sep-Unet	0.96	0.96	0.98	0.97	0.96

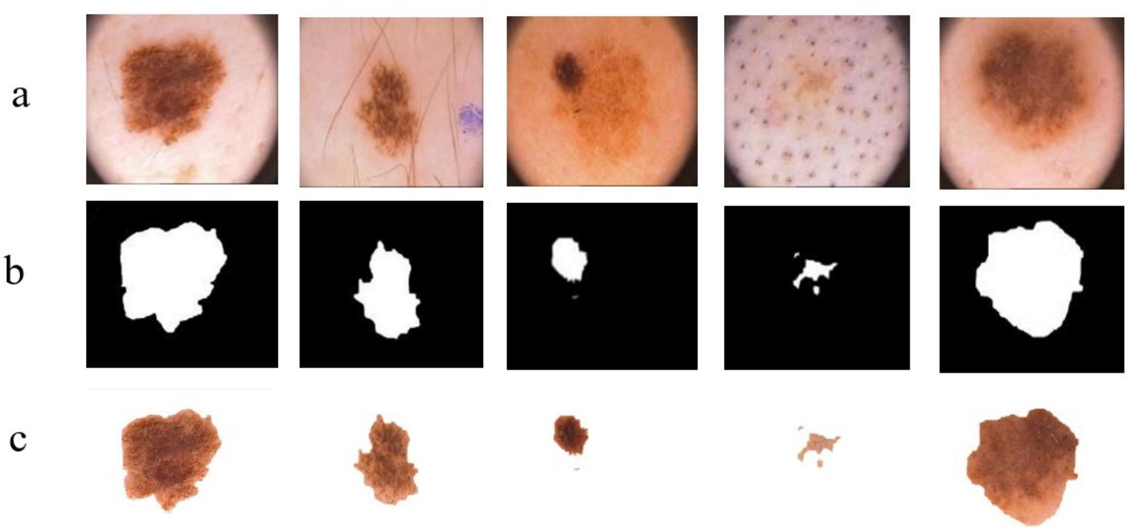

Figure 3. The Results of Implementation on ISIC Dataset
(A) The Samples of PH2 Skin Lesions
(B) The Segmentation Results for Each Sample
(C) The Final Skin Lesion Segmentation.

As shown in Table 1 and Table 2, recall is more than precision and recall means that the lesion in the output is detected as lesion correctly. The precision indicates that, it can correctly identify the cases that were a lesion and the cases that were not a lesion are not identified as a lesion. For medical cases the amount of recall is more important because the opinion is about the lesion, so the proposed algorithm works well. In the next part, we will compare the performance of the proposed method with the other algorithms.

Table 3. The Results of Comparison Based on PH2 Dataset

Algorithm	Dice Coefficient
Unet [9]	0.8370
Evolving FCM based on the proposed PSO [23]	0.8748
Sep-Unet	0.9603

Table 4. The Results of Comparison Based on ISIC2017 Dataset

Algorithm	Dice Coefficient
Unet [9]	0.71
Evolving FCM based on the proposed PSO [23]	0.76
Sep-Unet	0.96

In total, the number of training parameters in the Sep-unet algorithm reaches to 19.926.666, while the number of training parameters in the Unet is equal to 31.042.369. In other words, we have about 36 percent reduction in training parameters in the proposed algorithm. A comparison of training parameters between the proposed method and Unet has been displayed in Figure 5. In addition to the training parameter reduction, we could increase the evaluation metrics by the proposed algorithm.

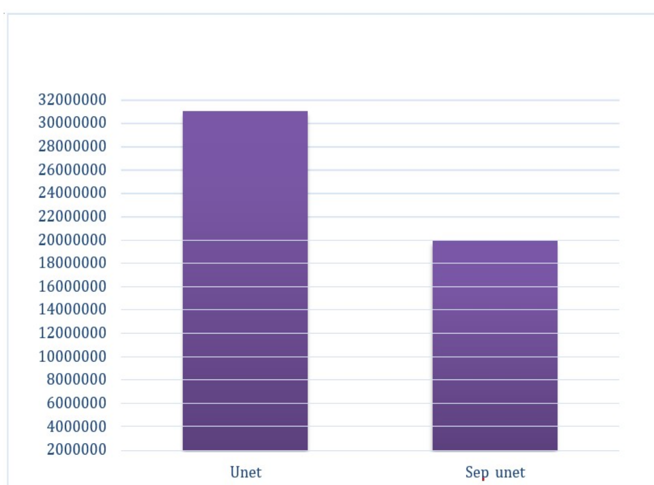

Figure 5. An Analysis Based on Training Parameters

5. Conclusion

In this study, the authors used a novel architecture for dermoscopic image segmentation by convolutional neural networks based on the separable layers. The main goal of this study is training parameters reduction. By examining the results, the proposed algorithm, despite the fact that it has less complexity based on the number of training parameters than the other methods, it has a good precision compared to the other similar algorithms. Using this segmentation algorithm has increased the precision of training pixels classification and by evaluating the center as well as the location of the lesion in the image accurately, the computational criteria has been increased. As shown in Figure 4 and Figure 5, the borders of the segmented lesions are clear and there is not any ambiguity in the lines and borders around the lesions. Additionally, based on the numerical results in the Table 1 and Table 2, the evaluation metrics are more than 95% which shows the robustness of this method toward medical image segmentation. By comparing the amount of dice coefficient that is an effective metric in image segmentation, we see that the proposed algorithm is better than the other similar algorithms for skin lesion segmentation. This method is implemented on the PH2 and ISIC datasets and in the future works, we can develop it on the other skin lesion images and utilize lighter architectures to decrease the implementation time. In addition, there are many other medical images such as chest Xray and brain MRI images that can be segmented by the proposed algorithm for tumor and lesion segmentation.

References

[1] C. Chen, J. Leng and G. Xu, “A general framework of piecewise-polynomial Mumford–Shah model for image segmentation,” *International Journal of Computer Mathematics*, vol.94, no.10, pp. 1981-1997, 2017.

[2] C. Kaushal, K. Kaushal and A. Singla, “Firefly optimization-based segmentation technique to analyse medical images of breast cancer,” *International Journal of Computer Mathematics*, vol. 98, no. 7, pp. 1293-1308, 2021.

[3] J. Bergstra, and Y. Bengio, “Random search for hyper-parameter optimization,” *Journal of machine learning research*, vol. 13, no. 2, 2012.

[4] M. Taheri, M. Rastgarpour and A. Koochari, “A novel method for medical image segmentation based on convolutional neural networks with SGD optimization,” *Journal of Electrical and Computer Engineering Innovations (JECEI)*, vol. 9, no. 1, pp. 37-46, 2021.

[5] S. Ren, “Faster r-cnn, towards real-time object detection with region proposal networks,” *Advances in neural information processing systems,* vol.28, pp. 91-99, 2015.

[6] C. Farabet, “Learning hierarchical features for scene labeling,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915-1929, 2015.

[7] A. Karpathy, and L. Fei-Fei, “Deep visual-semantic alignments for generating image descriptions,” *IEEE conference on computer vision and pattern recognition*, 2015.

[8] Y. Taigman, “Closing the gap to human-level performance in face verification deepface,” In Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR).

[9] A. Toshev, and C.D. Szegedy, “Human pose estimation via deep neural networks,” *CVPR*. Columbus, Ohio, pp. 1653-1660, 2014.

[10] T. Y. Tan, “Evolving ensemble models for image segmentation using enhanced particle swarm optimization,” *IEEE access,* vol. 7, pp. 34004-34019, 2019.

[11] O. Ronneberger, P. Fischer and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *International Conference on Medical image computing and computer-assisted intervention,* 2015.

[12] Q. Zhang, B. Li, and F. Zhang, “A MOEA/D approach to exploit the crucial structure of convolution kernels,” *Tenth International Conference on Advanced Computational Intelligence (ICACI),* 2018.

[13] T. Y. Tan, L. Zhang, C.P. Lim, B. Fielding, Y. Yu, and E. Anderson, “Evolving ensemble models for image segmentation using enhanced particle swarm optimization,” *IEEE access*, vol. 7, pp. 34004-34019, 2019

[14] C. C. Chen, “Medical image segmentation with adjustable computational complexity using data density functionals,” *Applied Sciences*, vo. 9, no. 8, pp. 1718, 2019.

[15] D. Abdelhafiz, "Convolutional neural network for automated mass segmentation in mammography," *BMC bioinformatics*, vol.21, no.1, pp. pp. 1-19, 2020.

[16] N. Feng, X. Geng, and L. Qin, "Study on MRI medical image segmentation technology based on CNN-CRF model," *IEEE Access*, vol. 8, pp. 60505-60514, 2020.

[17] N. Dong, "Neural architecture search for adversarial medical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2019.

[18] W. Zhang, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *Neuro Image*, vol. 108. pp. 214-224, 2015.

[19] Y. Ding, "Using deep convolutional neural networks for neonatal brain image segmentation, Frontiers in neuroscience," vol. 14, pp. 207, 2020.

[20] J. Bernal, K. Kushibar, D. S. Asfaw, S. Valverde, A. Oliver, R. Martí and X. Lladó, "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging,"vol. 95, pp. 64-81, 2019.

[21] D. Dai, C. Dong, S. Xue, Q. Yan, Z. Li, C. Zhang and N. Luo, "A novel multi-scale residual encoding and decoding network for skin lesion segmentation," *Medical Image Analysis*, vol. 75, pp. 102293, 2022.

[22] B. Hafhouf, A. Zitouni, A. C. Megherbi, and S. Sbaa, "An Improved and Robust Encoder–Decoder for Skin Lesion Segmentation," *Arabian Journal for Science and Engineering*, pp. 1-15, 2022.

[23] T. Mendonca, Data from: A pusblic database for the analysis of dermoscopic images, Dermoscopy image analysis, [Online] available at https://www.fc.up.pt, 2015. [Accessed: Jan. 5, 2022].

[24] N. C. Codella, "Data from: Skin lesion analysis toward melanoma detection," International symposium on biomedical imaging (isbi), 2017, [Online] available at https://challenge.isic-archive.com, [Accessed: Jan. 8, 2022].