Journal of Engineering Science and Technology
Vol. 16, No. 6 (2021) 4896 - 4909

EVALUATING THE EFFECT OF LESION SEGMENTATION ON THE DETECTION OF SKIN CANCER BY PRE-TRAINED CNN MODELS

JALE BEKTAŞ[1,*], YASIN BEKTAŞ[2], EVRIM E. KANGAL[1]

[1]Mersin University, Erdemli School of Applied Technology and Management, Department of Computer Technology and Information Systems, Mersin, Turkey
[2]Mersin University, Vocational School of Erdemli, Department of Computer Technologies, Mersin, Turkey
*Corresponding Author: jale@mersin.edu.tr

Abstract

Early diagnosis of melanoma, which is considered to be one of the deadliest skin cancers, via medical imaging can significantly improve the course of the disease. However, expert assessments are subjective and open to errors due to large variations in dermoscopy images. To cope with this problem, a two-stage framework is proposed for the detection of melanoma in dermoscopic images. Firstly, by eliminating the presence of natural structures such as veins or hair and the variations in the pattern region, segmented images are obtained from raw image data with the help of pixel-wise image processing techniques. The second part of this framework is the recognition stage of the skin lesions by Pre-Trained Deep Networks (PTN). By using segmented input images, the PTN classifiers are optimized with hyperparameters and classification is performed. The performance of the framework was evaluated separately before and after segmentation for six PTN used in this study. An improvement of 5.1% in the Accuracy metric and 7.7% in the Jaccard metric have been observed in the lesion recognition process with the segmentation framework according to the averages of network performances.

Keywords: Deep learning, Dermoscopy, Lesion segmentation, Melanoma recognition, Pre-trained CNNs

1. Introduction

Melanoma is a cancer type, which usually begins and develops in the skin. It is considered being one of the deadliest skin cancers, and approximately 9.6 million cancer deaths are detected in a year [1]. Early detection and proper treatment improve the survival rate significantly.

Dermoscopy Technique (DT henceforth) was developed to enhance the detection ability of melanoma. It is a non-invasive skin imaging technique that is employed to obtain an enlarged and illuminated image of a specific area on the skin. Elimination of the surface reflection and its effects, better reveals the details in the skin lesions. The use of dermoscopy evaluation is very common and gives higher accuracy rates when compared to the naked eye in melanoma diagnosis [2]. However, the examinations based on dermoscopy images carried out by dermatologists may generally be time-consuming and subjective; and even experienced dermatologists may have different diagnostic results on the same images [3]. Because of these considerations, automatic pre-diagnosis systems receive high demand.

In recent years, automatic melanoma diagnosis using dermoscopic images is a topic that has attracted a lot of attention from researchers. However, the low-contrast lesion at the early stages of the disease and the presence of natural structures such as hair and veins has made automatic melanoma diagnosis quite difficult [4]. For this reason, considerable efforts have been dedicated to developing computer-aided analysis systems using dermoscopic image analysis. By and large, these automatic solutions contain the following steps: (i) Lesion segmentation [5-10]; (ii) Determining the images and the dermoscopic features (feature extraction) [11-15]; (iii) Detecting the disease with classifiers [16-22].

The importance to review the state-of-the-art deep learning architectures that optimizes the hyperparameters and developing an automatic lesion segmentation system with image processing techniques for a cancer diagnosis on dermoscopic images is proposed in this study. The importance of applying image processing techniques within a framework in the correct flow is emphasized. The classification success of state-of-the-art networks on segmented dermoscopic lesion images is compared with the classification success on non-segmented images.

2. Previous Work

Image segmentation is an important step in extracting high-level features in an image and is included as a pre-process of the pattern recognition process [5]. In addition, the recognition process has been proposed in previous works along with the accurate acquisition of lesions by segmentation approaches in different techniques. For this reason, these types of approaches can be broadly divided into threshold-based, region-based [6], clustering approaches, and hybrid approaches. Replacing colour images with density images in a threshold-based approach and using Otsu's threshold method can produce good segmentation results [7].

Performing a median filter transform in an image in a region-based approach and then applying morphological operations at the pixel level can offer distinctive results for a region of Interest (ROI) field inference [8]. Converting the original image to RGB channels and analysing boundary irregularities by performing some

morphological features on skin lesions increased the sensitivity performance at the pixel level [9].

Hybrid approaches combine multiple approaches and are implemented in a schema structure in an ordered manner. Integration of threshold, morphological operations and k-means usage in segmentation applications with clustering approach in computer vision area were proposed and its effectiveness on firefly algorithm was measured [10]. Again in another study, mathematical morphology is combined with the k-means method with an ensemble technique [11]. Preprocess stages like feature extraction and feature selection, which are the most influential subjects that affect the segmentation performance, were mentioned in detail in a hybrid approach [12].

In addition to these approaches, Deep Convolutional Neural Network (DCNN henceforth) based studies have also been conducted using different deep learning networks for skin lesion segmentation. The entropy-based CNN technique is used for colour, brightness and texture features extraction to obtain more sensitive segmentation results [13]. In some studies, CNN techniques were evaluated in the same framework in both segmentation and classification [14,15].

Many studies have been conducted on the automatic detection of melanoma over dermoscopic images, including benign/malignant classification. Therefore, the CNN Architecture, which has come to the forefront in recent years as a very important for benign/malignant classification [16], has been employed. The lesion segmentation and the classification processes were carried out with DRN (Deep Residual Networks henceforth), which is another deep learning method [17]. In another study, the classification of all non-melanoma, melanoma, or atypical lesions was classified by combining deep learning, sparse coding, and support vector machine (SVM) algorithms in one framework [18].

Pre-trained CNN models, such as Alexnet and VGG16 has utilized for recognition of the benign/malign differentiation and PCA is used for feature selection [19]. To classify seven-class based skin lesions, GoogLeNet and Inception-v3 are used on ISICC 2018 dataset [20]. There are also studies developed using CNN to distinguish skin cancer from rashes and to help detect skin cancer in the skin [21]. A CNN model community is defined to improve the accuracy of lesions to related categories such as melanoma, benign, and seborrheic keratosis. High-level deep attributes are defined and classified based on the sum of the final maximum probabilities, resulting in significant performance [22].

3. Proposed Method

The rest of the paper is organized as follows: Section 3.1 presents lesion segmentation by determining dermoscopic image feature extraction. Section 3.2 illustrates a detailed description of the CNN methods and hyperparameters involved in the diagnosis of skin lesion images. Meanwhile, Section 4 lays performance evaluation measures by giving the design evaluations of the classifiers and transfer learning optimisations. Finally, Section 5 concludes this review. The flowchart from segmentation steps to application on pre-trained DCNN networks on ImageNet for classification of ISBI dataset is given in Fig. 1.

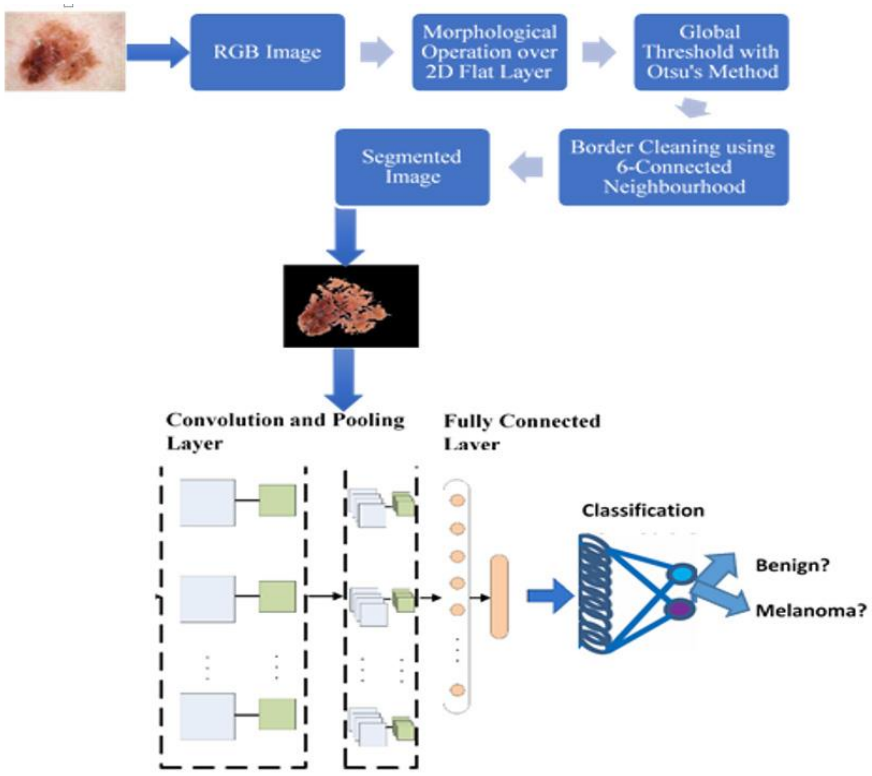

Fig. 1. Flowchart of the two stages DCNN model for skin lesion analysis and classification.

3.1. Lesion segmentation

The segmentation step has unquestionable importance in the classification and similarity-matching steps as it is where feature extraction depending on the boundaries of the lesion is done. However, segmentation might be hard because of variations in shape, size, colour, tissue, and lesion site together with different skin types. In addition, some lesions may have irregular boundaries, or there may be a swift transition between the lesion and the skin. Other situations that need to be coped with are the hairy structures, veins and specular entity reflections that surround the lesions. In literature, many image processing methods were employed after combining one or more methods to determine lesion boundaries. Among these, adaptive thresholding, morphological operations, edge-based model, colour clustering come to the forefront. Using more than one method increases the success of the classifiers [23-25].

As seen in Fig. 1., the flowchart of the skin lesion segmentation stage consists of a combination of (i) Morphological Operations, (ii) Global thresholding with the Otsu's, and (iii) Border cleaning methods, respectively. The images were firstly subjected to a morphologic dilation process using a $6x6$ spatial filter. By doing so, the aim was to purify the images from artefacts like veins, hair details, and marks in their structure. Firstly, the Grayscale IG Image has been filtered:

$$D_1 = \frac{1}{90}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{1}$$

$$IG_{dil} = IG * D_1 \tag{2}$$

where IG is a grayscale image and $*$ is a convolution operation. Then the grayscale image IG_{dil}, which has passed through a dilation process belonging to the three RGB channels was combined. The Otsu Global Thresholding is applied to the image that belongs to the blue plane. By optimizing the Otsu Threshold for the dermoscopy images, the Otsu Threshold is expanded after reducing it by 10 [7].

$$T = \begin{cases} 0 & OtsuT - 10 > 10 \\ eps & OtsuT - 10 < 10 \end{cases} \tag{3}$$

After applying the Global Threshold Procedure with the Otsu Method, all possible pixels belonging to the lesion were eliminated. However, after the binarization process, outside the lesion, there may appear areas that need to be cleared, particularly in the boundaries of the image. A method was chosen which clears the brighter pixels that were associated with the boundaries of the image. By employing a 3D connection, a 6×6 size filter was preferred as a basis of the neighbourhood status of the pixel. The images were segmented by employing the Morphologic Operation, Otsu Global Thresholding and Border Cleaning Methods, respectively. The segmentation phases are given in Fig. 2, respectively. Then, sample images with segmentation boundaries are given in Fig. 3.

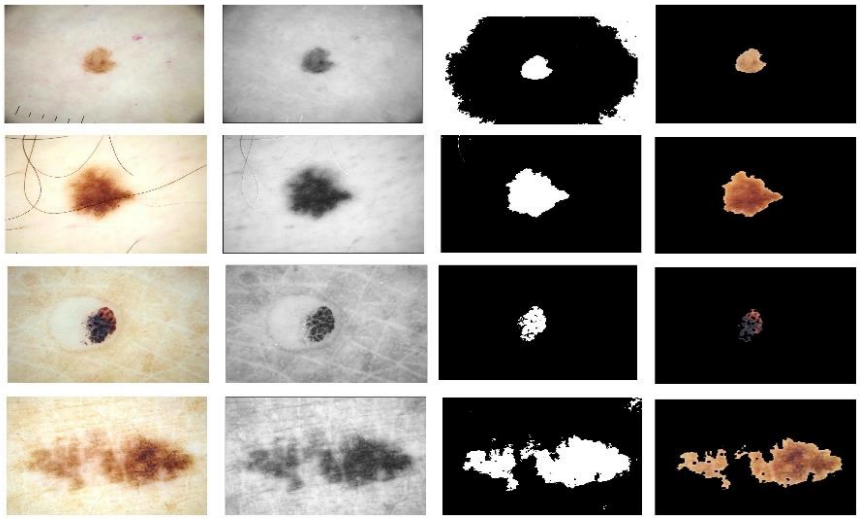

Fig. 2. The first two rows belong to benign; the last two rows belong to malignant images (a) original RGB image (b) Morphological dilation (c) after the Otsu's method (d) after border cleaning operation and the result of segmentation of the lesion region.

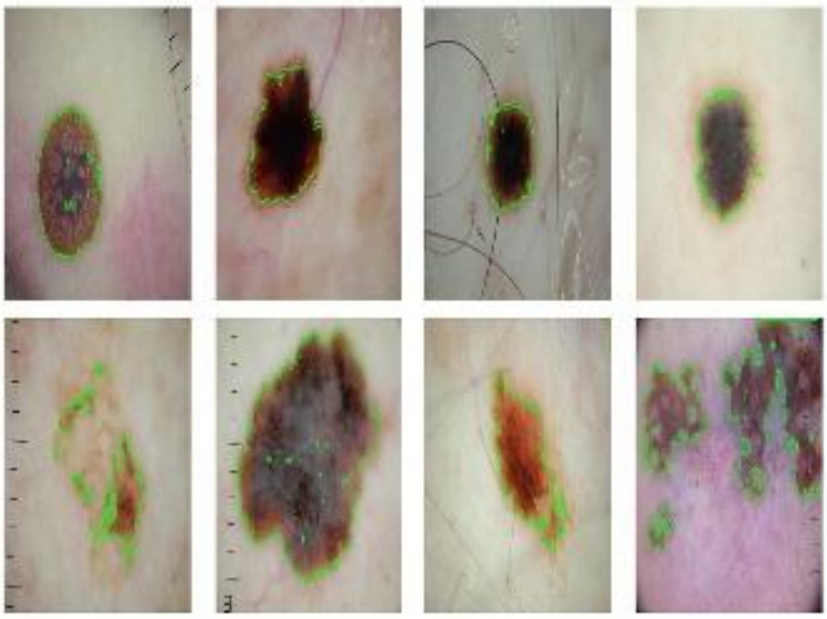

Fig. 3. Examples of some skin lesion segmentation results from dermoscopic images. Non-melanoma and melanoma lesion images were randomly mixed. The green contours indicate the segmentation results of the proposed method on the original images.

3.2. CNN methods

In the image, it is possible to make a distinction between the categories with the Feature Extraction Method; and furthermore, an upper-level information level may be deduced based on the raw pixel values. The Feature extraction is carried out in an unaudited way. Here, the image classes have nothing in common with the information extracted from the pixels. For this reason, the Feature extraction cannot be corrected according to the classes and images. If the extracted features do not have the qualifications needed for distinguishing the classes, problems occur in the accuracy rates of the model regardless of the classification strategy. Another common strategy that is preferred for this may be selecting a great number of feature extractors and bringing them together in a creative way to achieve a better feature; however, this is also very difficult in practice. For this reason, achieving a good accuracy level involves a great deal of intuitivism to change the parameters manually depending on the domain.

With the help of Convnet, which was developed to reduce this complexity, the transitions of the parameters may be more stable, and the application development process may become more practical. A network training starting from zero was performed by employing a great dataset like ImageNet including Convnet. To perform this, the parameters between the neurons and the sparse connections in the convolution layer must be shared. Since 2010, ImageNet has been supporting the participants for studies in the field of visual recognition with 1.2 million images of 1000 different classes from the ImageNet data set.

3.2.1. Complexity in calculation

Most ConvNets have large memory and calculation requirements, particularly in the training step. For this reason, this becomes an important issue. As it may be predicted, to achieve higher rates of accuracy, a more calculation-intensive network is required, which always necessitates a balance between accuracy and calculation. Aside from these, there are a great number of factors like the facility of training and the good generalization of the network.

3.2.2. AlexNet

This architecture is one of the initial Deep Networks that employed the ImageNet Classification with accuracy compared to traditional methodologies. It consists of 5 convolution layers, 3 of which consist of layers that are connected completely to each other. Instead of the Tanh or Sigmoid function, which is the previous standard for conventional neural networks, it employs the ReLu (Rectified Linear Unit) for the non-linear part. The advantage of ReLu over Sigmoid is that the sigmoid derivative becomes very small in the saturation area; and together with this, the learning furthers in a faster manner because of the weakening in the update in the weights. ReLu is given by $f(x) = max\ (0, x)$.

In the network, after each convolution and after each fully bound layer (FC), the ReLu layer is inserted. Another problem that is resolved by this architecture is reducing the over-fitting by employing a Dropout Layer after each FC layer. The Dropout Layer is related to a probability (p) and is applied separately for each neuron of the counterpart map.

3.2.3. VGG16

Large Kernel-Sized Filters (11 and 5 in the first and second layers, respectively) are replaced with more than $3x3$ Core-Size filters one after the other to improve AlexNet. In the domain size of the input image, which holds a certain output, the multiple-stacked small-size kernel is better than the larger-size kernel, since multiple non-linear layers increase the depth of the network. In this way, it makes it possible to learn more complex features at a lower cost.

Blocks that are applied more than once to the same filter size are employed in VGG-D to make the feature extraction process more complex and descriptive. The VGG Convolution Layers are followed by 3 fully bound layers. The width of the network starts with at least the "64" value after each sub-sampling/pooling layer and increases twice.

3.2.4. GoogLeNet

GoogLeNet is based on the notion that neurons are unnecessary in a deep network due to their correlations. The case in which a deep network is most efficient is the case in which the connection between the output neurons is of low size. There are some techniques employed to reveal these connections, which will result in weight-connection in a sparse structure.

It has developed a module that is named "an inception module", which is similar to a sparse CNN by GoogLeNet with a normal density. As mentioned above, the width/number of the convolution filters of a certain core size is kept small as a small

number of neurons are effective. Furthermore, it also uses 5×5, 3×3 and different sizes of folds similar to these to capture the details at different sizes. Another point that attracts attention to the module is that it includes a Bottleneck Layer. With the help of this layer, the calculation requirement is reduced a great deal. GoogLeNet replaces the fully-connected layers with a Simple Global Average Pooling, which evaluates and averages the channel values after the last convolutional layer, reducing the total number of parameters at a significant level. Using a wide and deep network enables GoogLeNet to extract without influencing the accuracy of the FC layers.

3.2.5. Resnet

The global average pooling is employed right after the classification layer. ResNets perform the learning process with a network depth, which is as big as 152. The architecture is similar to VGGNet, which consists mostly of $3x3$ filters. Residual Networks (Resnet) carry out deep network training through the modules that are called Residual Models by configuring the network. With this structure, the degradation problem might be eliminated. A network that has a flat structure with 30 layers may have higher error rates than a flat network that has fewer layers. In such a situation, since the network is converted into a residual network structure, it becomes possible that fewer training mistakes are caught.

4. Experiments and Results

4.1. Dataset

The International Skin Imaging Cooperation (ISIC) dataset has one of the biggest non-polarized dermoscopy image collections, which offers the characteristics and metadata of the lesions for analyses. The dataset used in the present study is a part of the ISBI 2017 dataset and presents the most difficult cases for the experts in the way that it includes 334 image melanoma and 144 nevus images and 2146 explicit benign lesions (in total 2624).

The images are 96 dpi and might be taken as standard format in $1024x768$ resolution. The images in this dataset are presented without lesion segmentation. For this reason, the limiting areas surrounding the skin lesions have been removed sensitively and then analysed with the classifiers to eliminate the erroneous areas of the image at the diagnosis stage of the disease. In this respect, three tasks must be carried out. For the rest of the paper, the experimental results that were obtained from the classifiers created with the defined hyperparameters of the proposed method will be given.

4.2. The design evaluations of the classifiers and transfer learning parameters

The RGB images that were segmented have been re-organized according to the image size required by the classifiers as an initial step. While expecting the arrangement of the entry data at VGG-16 and VGG-19, 224×224 pixel sizes, GoogLeNet, InceptionV3 and AlexNet have been arranged according to 227×227 pixels. 350 sample images have been used, 200 for benign and 150 for malign, for the test phase. After the input data has been presented to the networks,

they have been progressed to the classifier layer in line with the architecture. Some arrangements have been made in the architecture. In this respect:

- All the data were adapted to train/validation datasets with a 70-30% separation.
- The Fully Connected Layer was re-arranged according to the Class Number, and the Connected Class Layer was re-arranged according to the Binary Output (not according to the 1000-class).
- According to the transfer learning strategy, the pre-trained weight values are loaded; however, the weight values of the first 4 layers are frozen to avoid that they are updated again during the training step. In this way, the dominant effects of the general features of the natural images on the network are avoided. The classification of skin lesions might be more specific by training only the final layers.
- The connections between the layers are created again.
- The Data Augmentation Process, which is applied to extend the dataset in limited numbers, is very effective on the performance of the classifiers [26]. For this reason, it is carried out separately by applying them to the training and validation groups. By carrying out a scale between 0.9-1.4, the rotation at right angles enables us to obtain sensitive samples without losing the image size. The proper rotation might be obtained as there is a black background behind the boundaries of the image. As most lesions might be inserted almost anywhere in the image, this zooming method is versatile. Facing every angle with CNN was made difficult for the network.
- The Stochastic Gradient Descent with the Momentum Learning Method was preferred with a 10-6 learning rate and 0.9 momentum value.
- 3-fold cross-validation was executed.

4.3. Results

4.3.1. Evaluation metrics

The test dataset has been adjusted in terms of balance with lesion mask images that have been created with the lesion segmentation process, and the main metrics have been applied to evaluate the performance of the method. The main metrics that were employed in the study include accuracy (AC), sensitivity (Se), specificity (Spec), dice coefficient (DI) and Jaccard index (JA). The metric formulizations have been presented in Fig. 4.

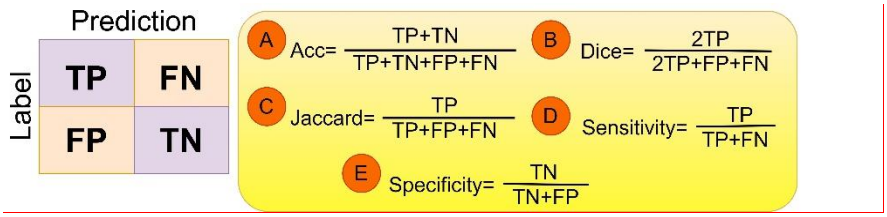

Fig. 4. Confusion matrix distribution and main metric formulas.

4.3.2. Comparison of pre-trained networks within two-stage framework

Lesion segmentation and melanoma/benign classification were carried out from dermoscopic images with a two-step framework in this study. After the data were

given a balanced structure, model evaluations were carried out individually according to before and after segmentation. To confirm the necessity and importance of the segmentation, the classification performance was compared with the pre-trained Deep Learning Models by keeping the image samples, with and without segmentation. The same image samples were executed with the same augmentation and the same training parameters in all of the networks. The final outcomes of TPR/FPR are presented in Fig. 5.

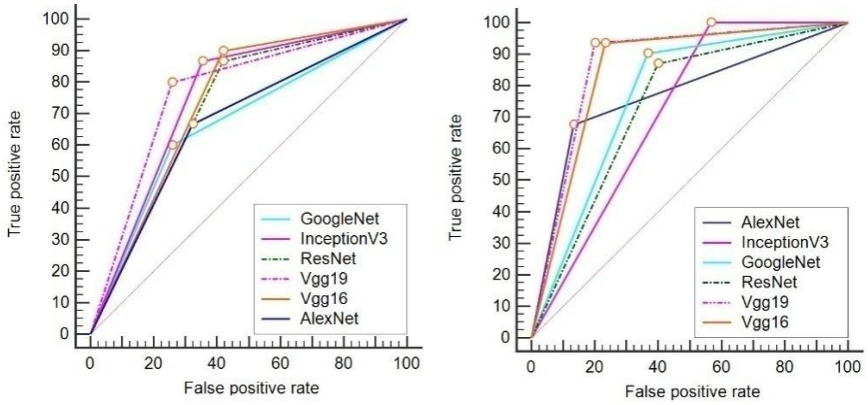

Fig.5. TPR/FPR results for pre-trained networks.
(a) Classification without segmentation. (b) Classification with segmentation.

Figure 5 further compares the TPR/FPR results of the pre-trained networks for the classification without segmentation stage and classification with segmentation stage, respectively. It can be seen from Fig. 4 that the result of Resnet, VGG16 networks for the classification without segmentation, and VGG19, VGG16 networks for the classification with segmentation are better than other comparative pre-trained networks. This shows that the proposed framework can improve the detection performance of melanoma and non-melanoma images.

In Table 1, the results obtained from 6 pre-trained networks that are most commonly used are listed. Accuracy, specificity, sensitivity, dice, jaccard evaluation metrics were used in this regard.

A possible reason for this is that features are extracted from the layer right before the Softmax, which learns the most effective specific features from the ImageNet data in the deepest layer. However, it is also possible that it extracts the most suitable features in the feature extraction process with the same context as the ISBI dataset.

Training the network in line with the results of the segmentation makes it possible to avoid the training deviations stemming from other artefacts and structures in the image during the training, and to obtain features for classifying the classes in a better manner, and as a result, to observe the increase in the performance in the evaluation metrics. With this two-stage framework, it was observed in previous studies that better results were obtained in original dermoscopy images without segmentation when compared to the direct use of pre-trained networks. According to the averages given in Table 2 for the evaluations of the results, there

was a co-recovery at a rate of 5.1% for the Accuracy metric and 7.7% for the Jaccard metric. One probable cause for this is the variations of the lesion sizes.

Table 1. Comparisons of different networks with and without segmentation.

Method	AC	Se	Spec	DI	JA
Classification without Segmentation					
InceptionV3	0.750	0.742	0.900	0.807	0.676
Alexnet	0.672	0.738	0.733	0.742	0.590
GoogleNet	0.672	0.742	0.733	0.742	0.590
Resnet50	0.738	**0.903**	0.700	0.824	0.700
VGG16	0.738	0.677	0.933	0.778	0.636
VGG19	**0.771**	0.806	0.867	**0.833**	**0.714**
Classification with Segmentation					
InceptionV3	0.717	0.600	1	0.750	0.600
Alexnet	0.766	0.867	0.833	0.852	0.743
GoogleNet	0.767	0.800	0.933	0.857	0.750
Resnet50	0.767	0.767	0.867	0.807	0.676
VGG16	0.800	0.833	0.933	0.877	0.781
VGG19	**0.833**	0.900	0.900	**0.900**	**0.818**

Table 2. Relative comparisons of results over original data and segmented data.

Method	AC	Se	Spec	DI	JA
Original	0.724	0.769	0.811	0.788	0.651
Segmented	0.775	0.794	0.911	0.841	0.728

Performances of comparison with other studies made on the ISBI challenge dataset for skin lesion segmentation are given in Table 3.

Table 3. Relevant studies on results of ISBI dataset on skin lesion classification

Authors	Method	Size	Sen	Spe	Acc
Codella[18]	Caffe CNN	2624	0.81	0.86	0.85
Lequan [17]	GoogleNet	900	0.51	0.93	0.85
This Study	**VGG19**	**350**	**0.90**	**0.90**	**0.83**
Lopez [27]	VGGNet	900	0.78	0.79	0.81
Mahbod [28]	VGGNet	2000	-	-	0.81
Majtner [16]	AlexNet	900	0.53	0.87	0.81
Subha [21]	CNN	-	-	-	0.80
Burdick [14]	InceptionV3	900	0.76	0.72	0.69
Harangi [20]	InceptionV3, GoogleNet	6705	0.66	-	0.64

Table 3 lists the studies which employ pre-trained networks and ranks according to accuracy values. These studies also include image segmentation methods for their investigations. Codella [18] made the distinction between Melanoma and Atypical and Benign images with a classification performance of accuracy of 85%, specificity of 86%, and sensitivity of 80%. In this study, better results were obtained with ensemble methods, but according to pre-trained network assessment, the Caffe CNN result was included in the first row of Table 3. Lequan [17] used GoogleNet and achieved a classification performance with an accuracy of 85%, a specificity

of 93%, and a sensitivity of 50%. Lopez [27] used VGG16 ConvNet in three different ways, consisting of transfer learning paradigm and fine-tuning combinations, reaching a classification performance with an accuracy of 81.33%, specificity of 78.66%, and sensitivity of 79.74%. Mahbod et al. [28] utilized AlexNet and VGG16 pre-trained networks for feature extraction. In this study, the features were mainly extracted from the last fully connected (FC) layers. VGGNet (FC8) with the SVM Classifier achieved a classification performance with an accuracy of 81%. Majtner [16] employed RSurf Features and Local Binary Patterns for preprocessing and used the SVM classifier based on AlexNet, achieving a classification performance with an accuracy of 80.5%, specificity of 87.2%, and sensitivity of 53.3%. Burdick [14] investigated the effect of expanding the segmentation border for skin lesion segmentation and achieved a classification performance with InceptionV3 with an accuracy of 69.3%, specificity of 72.3%, and sensitivity of 76%.

5. Conclusion

In this study, the effects of lesion segmentation on dermoscopic images in image classification on pre-trained CNNs have been evaluated. Segmentation and skin lesion recognition processes have been combined without any dramatical missegmentation and overfitting problems, and an automatic framework has been created without the requirement of manual interaction. With this approach, the ISBI 2017 data were evaluated for skin lesion segmentation. The best-obtained result for the "classification with segmentation" framework was 83.3%, and an improvement has been reached at a rate of 7.7% for the Jaccard compared to the participation of the original dataset with a "classification without segmentation". Furthermore, CNNs with very deep, effective training parameters, pre-trained have been used to solve complex medical image analysis problems. It has been observed that the CNNs used can produce satisfactory results even with very few training samples when used in conjunction with appropriate hyperparameters.

Consequently, the DCNN model has been used with fewer samples than other studies done and dealt with overfitting by organizing hyperparameters. Although it is not possible to investigate all possible options in studies due to long training periods, the DCNN enables researchers to compare pre-trained networks by using the same benchmark data with a reduced number of samples with effective performance values.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A cancer journal for Clinicians*, 68(6), 394-424.
2. Bayraktar, M.; Kockara, S.; Halic, T.; Mete, M.; Wong, H.K.; and Iqbal, K. (2019). Local edge-enhanced active contour for accurate skin lesion border detection. *BMC Bioinformatics*, 20(2), 87-97.
3. Chakraborty, S.; Chatterjee, S.; Ashour, A.S.; Mali, K.; and Dey, N. (2018). *Intelligent computing in medical imaging: A study*. In Advancements in Applied Metaheuristic Computing. IGI global,143-163.

4. Lan, J.; Wen, J.; Cao, S.; Yin, T.; Jiang, B.; Lou, Y.; .and Tao, J. (2020). The diagnostic accuracy of dermoscopy and reflectance confocal microscopy for amelanotic/hypomelanotic melanoma: a systematic review and meta-analysis. *British Journal of Dermatology*, 183(2), 210-219.
5. Pedro, M.M.; Pereira, R.; Fonseca-Pinto, R.; Paiva, R.P.; Assuncao, P.A.; Tavora, L.M.; Thomaz, L.A.; and Faria, S.M. (2020). Dermoscopic skin lesion image segmentation based on local binary pattern clustering: Comparative study. *Biomedical Signal Processing and Control*, 59, 101924.
6. Rout, R.; and Parida, P. (2020). Transition region based approach for skin lesion segmentation. *Procedia Computer Science*, 171, 379-388.
7. Zortea, M.; Flores, E.; and Scharcanski, J. (2017). A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. *Pattern Recognition*, 64, 92-104.
8. Chandra, K.; and Murari, B. (2020). Confocal corneal endothelium dystrophy's analysis using a hybrid algorithm. *Journal of Engineering Science and Technology* (*JESTEC*), 15(5), 3419-3432.
9. Chatterjee, S.; Dey, D.; and Munshi, S. (2019). Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification. *Computer Methods and Programs in Biomedicine*, 178, 201-218.
10. Garg, S.; and Jindal, B. (2021). Skin lesion segmentation using k-mean and optimized fire fly algorithm. *Multimedia Tools and Applications*, 80(5), 7397-7410.
11. Braiki, M.; Benzinou, A.; Nasreddine, K.; and Hymery, N. (2020). Automatic human dendritic cells segmentation using K-means clustering and Chan-Vese active contour model. *Computer Methods and Programs in Biomedicine*, 195, 105520.
12. Oliveira, R.B.; Pereira, A.S.; Tavares, J.; and Manuel, R.S. (2017). Skin lesion computational diagnosis of dermoscopic images: ensemble models based on input feature manipulation. *Computer Methods and Programs in Biomedicine*, 149, 43-53.
13. Pathak, M.K.; Srinivasu, N.; and Bairagi, V. (2020). Support value based fusion matching using iris and sclera features for person authentication in unconstrained environment. *Journal of Engineering Science and Technology* (*JESTEC*), 15(4), 2595-2609.
14. Burdick, J.; Marques, O.; Weinthal, J.; and Furht, B. (2018) Rethinking Skin Lesion Segmentation in a Convolutional Classifier. *Journal of Digital Imaging*, 31(4), 435-440.
15. Yang, X.; Zeng, Z.; Yeo, S. Y.; Tan, C.; Tey, H. L.; and Su, Y. (2017). A novel multi-task deep learning model for skin lesion segmentation and classification. arXiv 2017, arXiv:1703.01025.
16. Majtner, T.; Yildirim, Y.S.; and Hardeberg, J.Y. (2016). Image processing theory tools and applications (IPTA). C*A: IEEE Computer Society and* 6*th International Conference*. Oulu, Finland, ,1-6.
17. Lequan, Y.; Chen, H.; Dou, Q.; Qin, J.; and Heng, P.A. (2017). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 36(4), 994-1004.

18. Codella, N.; Cai, J.; Abedini, M.; Garnavi, R.; Halpern, A.; and Smith, J.R. (2015). Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images. *International Workshop on Machine Learning in Medical Imaging*, Munich, Germany, 118-126.
19. Amin, J.; Sharif, A.; Gul, N.; Anjum, M.A.; Nisar, M.W.; Azam, F.; and Bukhari, S.A.C. (2020). Integrated design of deep features fusion for localization and classification of skin cancer. *Pattern Recognition Letters*, 131, 63-70.
20. Harangi, B.; Baran, A.; and Hajdu, A. (2020). Assisted deep learning framework for multi-class skin lesion classification considering a binary classification support. *Biomedical Signal Processing and Control*, 62, 102041.
21. Subha, S.; Wise, D. J. W.; Srinivasan, S.; Preetham, M.; and Soundarlingam, B. (2020). Detection and differentiation of skin cancer from rashes. 2020 *IEEE International Conference on Electronics and Sustainable Communication Systems* (*ICESC*), 389-393.
22. Harangi, B. (2018). Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of Biomedical Informatics*, 86, 25-32.
23. Hassan, M.R.; Ema, R.R.; and Islam, T. (2017). Color image segmentation using automated K-means clustering with RGB and HSV color spaces. *Global Journal of Computer Science and Technology*, 17(2), 1-10.
24. Maglogiannis, I.; Zafiropoulos, M.; and Kyranoudis, C. (2006). *Intelligent segmentation and classification of pigmented skin lesions in dermatological images*. Hellenic Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, 214-223.
25. Silveira, M.; Nascimento, J.C.; Marques, J.S.; and Rozeira, J. (2009). Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE Journal on Selected Topics in Signal Processing*, 3(1), 35-45.
26. Wang, J.; and Perez, L. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv [Preprint]. arXiv:1712.04621.
27. Lopez, A.R.; Giro-i-Nieto, X.; Burdick, J.; and Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. *In* 2017 13*th IASTED International Conference on Biomedical Engineering* (*BioMed*). Innsbruck, Austria, 49-54.
28. Mahbod, A.; Schaefer, G.; Wang, C.; Ecker, R.; and Ellinge, I. (2019). Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 *IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*), 1229-1233.