A Streamlined Lesion Segmentation Method Using Deep Learning and Image Processing for a Further Melanoma Diagnosis

Jinen Daghrir[1][a], Wafa Mbarki[1], Lotfi Tlig[1], Moez Bouchouicha[2], Noureddine Litaiem[3], Faten Zeglaoui[3] and Mounir Sayadi[1]

[1] *Université de Tunis, ENSIT, Laboratory SIME, Tunisia*

[2] *Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France*

[3] *University of Tunis El-Manar, Faculty of Medicine of Tunis, Department of Dermatology, Charles Nicolle Hospital, Tunis, Tunisia*

Keywords: Deep Learning, u-Net, Image Processing, Skin Lesion Segmentation, Majority Voting, Medical Imaging.

Abstract: Over the past two decades, the world has known a significant number of deaths from cancer. More specifically, melanoma which is considered as the deadliest form of skin cancer causes a remarkable percentage of all cancer deaths. Therefore, the health and disease management community has exceedingly invested in creating automated systems to help doctors better analyze such diseases. Correspondingly, we were interested in creating an automatic lesion detection task for further melanoma diagnosis. The lesion segmentation is considered to be a critical step in a pattern recognition system. Our proposed segmentation technique consists of finding lesions' masks using a baseline, edge-based, and more sophisticated and state-of-the-art method: thresholding using Otsu's technique, morphological snakes, and a fully CNN (Convolutional Neural Network) model based on the U-net architecture, respectively. These methods are commonly used when dealing with skin lesion segmentation, and each one of them has its advantages and drawbacks. The U-net architecture is improved by the use of the pre-trained encoder ResNet-50 on the ImageNet dataset. A majority voting is applied to generate the final segmentation map using these three methods. The experiments were conducted using a benchmark dataset and showed promising results compared to using these methods separately, the majority voting of the three methods can significantly improve the segmentation task by refining the borders of the masks issued by the Deep learning model.

1 INTRODUCTION

Skin cancer is considered one of the most frequent types of cancer. Melanoma which is a particular form of skin cancer, is the less common type but it is considered as the most malignant one. This fatal skin cancer can quickly spread to other parts of the body causing about 60 000 cancer deaths in 2018 as reported in (Khazaei et al., 2019), it is considered as 0.7% of all cancer deaths. Statistics examining the incidence rate from 1973 to 2009 show a rise in the number of cases (Heinzerling and Eigentler, 2021), which is particularly worrying in such a disease. Some types of skin lesions can be more dangerous than others, due to their ability to spread to other sites. About 5 to 15% of people having congenital nevi are more risked to develop melanoma. About one-third of cases will be certainly affected in their brains and can be associated with malformation of the central nervous system (Heinzerling and Eigentler, 2021). The incidence and mortality rates of melanoma differ from one country to another due to the variety of ethnic and racial groups(Schadendorf et al., 2018). The well-known definite risk factors for melanoma include ultraviolet (UV) radiation, life in low geographic latitudes, high alcohol consumption, consuming fatty foods, the presence of melanocytic or dysplastic naevi, a family or personal history of melanoma, phenotypic characteristics including fair hair, eye, and skin colors (Khazaei et al., 2019). Interestingly, it has been proved that the development index (HDI) is quite related to the number of mortality of melanoma. The more the HDI index increases, the more people have access to health services and early detection of diseases, and the less mortality will occur(Khazaei et al., 2019). UV protection is recom-

[a] https://orcid.org/0000-0002-1300-8939

Daghrir, J., Mbarki, W., Tlig, L., Bouchouicha, M., Litaiem, N., Zeglaoui, F. and Sayadi, M.
A Streamlined Lesion Segmentation Method Using Deep Learning and Image Processing for a Further Melanoma Diagnosis.
DOI: 10.5220/0013472400003938

In *Proceedings of the 11th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2025)*, pages 368-376
ISBN: 978-989-758-743-6; ISSN: 2184-4984

mended since it is proved that using protection will lead to a decrease in the incidence rate. However, the use of sunscreen does not provide enough protection against the development of melanocytic nevi consequently the high risk of developing melanoma (Heinzerling and Eigentler, 2021). One typical solution to decrease the annual number of deaths caused by melanoma is diagnosed it in an early stage, where the relative survival rate will be critical (Schadendorf et al., 2018). Self-examination is so vital in this case. Thus, people should examine their skin head to toe regularly, looking for any lesions that might be turned into melanoma. Self-exams can help people identify skin cancers in an early stage when the odds of curing them are completely high. Yet, physicians encourage people to routinely do self-examination if they notice any suspicious-looking lesions. "when in doubt, cut it out", says Darrick Antell, MD, a board-certified plastic and reconstructive surgeon practicing in New York (Daghrir et al., 2020). Currently, the health and disease management community has entirely been interested in creating automated computer systems for skin lesion inspection and characterization. Thus, it becomes vital to use supportive imaging to improve the diagnosis of this life-threatening skin cancer. Conventionally, these systems use some visual cues to characterize the lesions as malignant or not. One typical used clue is the ABCD signs. This feature can differentiate malignant melanoma from a benign skin lesion based on five defined characteristics, namely asymmetry (A), border irregularity (B), color variability (C), and diameter greater than 6 mm (D)(Maglogiannis and Doukas, 2009). To extract information about the border, lesion segmentation should be performed. This task is considered to be a critical step in the whole process of melanoma detection, it concerns the isolation of pathological skin lesions from the surrounding healthy skin. The segmentation step is usually done by experts, who try to find the ROI (Region Of Interest). This method is accurate since the experts apply their knowledge and experiences to the segmentation process. Thought, this process will be time-consuming and exigent, due to the variability and reliability of the expert behavioral observation. One straightforward method is to use fully automatic systems where the segmentation process is done using computer vision techniques without the need for human intervention.

In medical image processing, ROI segmentation is defined as a semantic segmentation which is a pixel-level recognition problem. It refers to the process of assigning each pixel in the image to a particular label. Particularly, in the case of skin lesion segmentation, healthy skin is assigned a background label. However, the inspected lesion is assigned a foreground label. The result of this process is a segmentation map, where the background and foreground regions are assigned with two different class labels. Commonly, many classical image processing techniques are proposed for Medical image segmentation (Mbarki et al., 2020; Sharma and Ray, 2006). These techniques rely only on low-level pixel processing and mathematical modeling to construct rule-based systems. Thresholding has been widely used since it is simple and not time-consuming. It consists of dividing an image into two regions: one corresponds to the ROI and the other to the background. Thus, each pixel is assigned a class label if its intensity is greater than a threshold. The threshold can be either globally or locally fixed according to a small set of pixels' intensities. The threshold is defined using statistical analysis or by optimizing a certain criterion. Thus, using only thresholding can often fail to produce accurate results since it can generate disconnected objects. However, Active contours or snakes can guarantee a connected set of pixels for an ROI. These methods are considered edge-based methods. An active contour is formulated as an energy minimization problem, where the energy depends on the placement of the curve surrounding the object to be segmented. The initial contour is manually defined and then iteratively deformed to minimize the energy function, which is typically based on the intensity differences between the object and the background pixels. Obviously, these methods are very sensitive to noise and the initial curve placement. The large variability of the shape of substructures and organs in medical images is one of the most challenging problems for the segmentation task. Therefore, there is no universal segmentation technique. Some techniques do not use only color information of pixels, but also much more domain knowledge(Litjens et al., 2017). In order to define and implement this knowledge, a considerable amount of expertise and time is required. Thus, completely automated systems are proposed to model the domain knowledge by training some algorithms with labeled pixels. Recently, a significant trend in segmentation tasks is the use of deep learning(DL)due to its effectiveness and speediness. With the advance of DL models, scientific papers presenting organ and substructure segmentation have gradually increased in the last few years(about 90 papers were published in 2017)(Litjens et al., 2017). The motivation for the excessive use of DL is the ability to use one model in many different domain applications and with different data resources, such as in(Alom et al., 2018), authors have tested their proposed DL model on three different benchmark datasets such as blood vessel in the retina, skin lesion, and lung lesion datasets. One widely used architecture of DL is the CNN(Convolutional Neural Network) which ensures continued good performance in many tasks, such as classification, segmentation, detection, decision-making, etc. The number of papers published in 2017 applying CNNs in medical image analysis is over 200 papers, and most of them are dedicated to dealing with substructure segmentation(Litjens et al., 2017). Exceptionally, for semantic segmentation, a fully convolutional network is required, which contains two opposite units, encoder and decoder. The encoder is used to define the high-level information with a low-resolution spatial tensor, while the decoder contains also convolution layers but coupled with up-sampling instead of down-sampling layers that increase the size of the spatial tensor accordingly recapturing the segmented image. In(Ronneberger et al., 2015), the authors have proposed the U-net architecture that takes an input image and outputs a segmentation map that has the same dimension as the input. The U-net was first used to detect cell boundaries in biomedical images. The contribution of such architecture is to use skip connections, which connect opposite convolutional layers in order to preserve low-level information. Definitely, DL is efficiently and widely used for many tasks, but the fact that it is time-consuming and data-hungry makes it a little bit limited. The performance of systems using DL will degrade if the amount of input data is not big enough. In addition, the data can be only produced by an expert who will label every input. Time and resources are required to properly do data labeling to build highly accurate systems, but it is not a straightforward process when it comes to huge datasets. For melanoma diagnosis, the border-based features should be extracted after segmenting the skin lesion. Two main automatic segmentation methods are presented in the literature, region-based and contour-based methods. Other hybrid methods are proposed which use both color transformation and contour detection techniques (Maglogiannis and Doukas, 2009). These methods are also classified into two groups: Methods that use classical image processing with analysis techniques and those that involve Artificial Intelligence(AI) concepts. The most usual and basic techniques that are used for lesion segmentation, are thresholding and region-growing techniques. These methods are classified as region-based approaches since they involve the differentiation of pixels' intensities of both malignant and healthy skin regions. Another type of method is contour-based which detects the region of interest by finding its edge. However, with the advent of more powerful techniques, these techniques seem to be very naive. For that, hybrid approaches are proposed in the literature (Jamil et al., 2019) that use both color transformation and edge detection, snakes or active contours are considered to be the state-of-the-art of lesion segmentation before the extensive use of machine learning(Maglogiannis and Doukas, 2009). Other methods involving AI techniques like fuzzy borders, K-means segmentation, or Deep Learning (DL) techniques are considered nowadays as the state-of-the-art for most image segmentation problems (Jadhav et al., 2019). Melanoma detection at an early stage can highly extend the life expectancy, thus it becomes vital to use CAD to ensure a quick and accurate diagnosis of a huge number of people. Due to the sensitivity of the lesion segmentation process which plays an imperative role in melanoma diagnosis. Many research papers are proposed in the literature dealing with lesion segmentation, although research is still trending up sharply due to a lot of challenges.

In this matter, we were interested in creating a new method for lesion segmentation that combines three different methods: thresholding using Otsu's technique, morphological Snakes, and a deep convolutional neural network based on U-net. The aforementioned techniques are used independently to generate three lesion masks. Then, using majority voting a new lesion segmentation map is generated. Majority voting defines the class of each pixel whether it belongs to the lesion or to the background based on the other decisions. This paper is organized as follows. In the following section, we extendedly introduce our proposed method. All the conducted experiments and the results evaluating the performance of the proposed method are given in the third section. At the end, we highlight and discuss the basic concept of our proposed method.

2 PROPOSED METHOD

One critical process in melanoma diagnosis is lesion segmentation. It consists of identifying the lesion's pixels among the others belonging to the healthy skin. In this paper, a new method has been proposed that combines three different methods using majority voting. The fusion of the three results issued by thresholding using Otsu's technique, morphological snakes, and CNN-based U-Net architecture, significantly ameliorates the segmentation mask by refining the lesion's border.

2.1 Thresholding Method

Thresholding has been widely used in skin cancer due to its effectiveness and integrity. Researchers have proposed many local adaptative and global thresholding techniques, such as Kapur et al.'s Entropy technique(Armato et al., 2002), Abutaleb's entropy technique(Ashwin et al., 2012), and Kittler and Illingworth's minimum error technique(Bae et al., 2005). However, Otsu's algorithm has been extensively applied in many systems. Thresholding is done by finding one single threshold that classifies pixels into two classes without image preprocessing. The threshold is determined by minimizing intra-class intensity variance σ_w, which is defined as the weighted sum of class variances:

$$\sigma_w(t) = w_{C_0}(t)\sigma^2_{C_0}(t) + w_{C_1}\sigma^2_{C_1}(t) \quad (1)$$

w_{C_0} and w_{C_1} are the probabilities of the two classes separated by a threshold t, and $\sigma^2_{C_0}$ and $\sigma^2_{C_1}$ are the variances of the two classes. After computing the probabilities of each class using the histogram, $w_i(0)$ and $\sigma^2_i(0)$ are initially set. Using different values of threshold t from 1 to the maximum pixel's intensity, $w_i(t)$ and $\sigma^2_i(t)$ will be computed to get the intra-class intensity variance. The global threshold will be the one that minimizes the $\sigma_w(t)$. This threshold is used then to differentiate the foreground which is the pathological skin lesion and the background. The segmentation map is found by assigning the value 1 to all the pixels having an intensity greater than the threshold t and the value 0 to all the other pixels. Two steps are performed to filter the unwanted objects. First, an opening is used twice, the morphological opening is an erosion followed by dilation using a disk as a structuring element. The erosion of the image helps to remove the small objects. Consecutively, dilation helps in ameliorating the background region by closing all the gaps. For further filtering, big regions having a centroid that does not belong to 75% of the image area will be eliminated. This process ensures that only the skin lesion mask will stand. Fig.1 illustrates the result of using thresholding and filtering the segmentation map.

2.2 Morphological Snake Method

Active contours, also known as Snakes are one of the most used techniques in computer vision, more specifically for image segmentation. Frequently, these methods are edge-based which requires clean boundaries indicting and the absence of artifacts and noises. These problems were then handled by the invention of more practical approaches such as the Active Contours Without Edges (ACWE)(Marquez-Neila et al., 2013). These approaches are known as morphological snakes since they use morphological operators over a binary array. The morphological ACWE can accurately segment the object when the pixels of the region of interest and those of the background have different averages. In this work, the Chan-Vese algorithm(Chan and Vese, 1999) which is a form of the snake curve is adopted. It is an energy minimization method that is widely used in medical applications. This method does not require defining the contours to separate the heterogeneous objects and it is sensitive to noise. This algorithm can also accurately detect the objects and increase the convergence speed when using a good initial curve placement. Since this method can detect more than one region of interest and obviously the skin lesion will be placed in the center of the image, we have used the initial curve presented in Fig.2. The result of using the snakes is shown in Fig.2. The final segmentation mask is generated after postprocessing the output of the Snake process. We have applied the morphological opening to refine the segmentation map.

2.3 U-Net Based Resnet50

The U-Net architecture is a fully convolutional neural network that was mainly developed for biomedical image segmentation. It consists of two blocks: encoder and decoder. CNNs are commonly used for classification tasks, the input image is fed to the network and downsampled into a low-resolution tensor containing high-level information which can then be classified using a fully connected layer. However, to generate a segmentation map that has the same size as the input, the output of the encoder should pass by an upsampling path(decoder). This makes the network layout has a U shape.

The left side of the U follows the typical architecture of a convolutional network. It consists of the repeated application of two 3×3 convolutions, which is followed by a normalization layer and a rectified linear unit (ReLU) activation function. Dropout and 2×2 max-pooling layers are used to downsample the feature map. At each downsampling step, the number of feature channels is increased. The purpose of the encoder is to capture the high-level information of the input image to do the segmentation process. To prevent losing the low-level information, the decoder should have connections with the encoder layers, that is known as skip connections.

Many CNNs are previously trained for image classification tasks, that contain meaningful information about how to perform predictions effectively. These

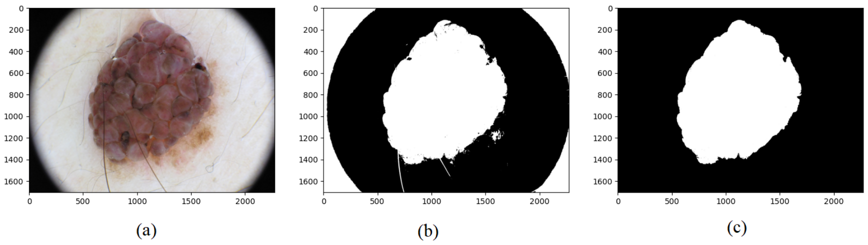

Figure 1: Segmentation result of the thresholding process: (a) Original image, (b) segmentation map using Otsu's algorithm, and (c) segmentation map after filtering.

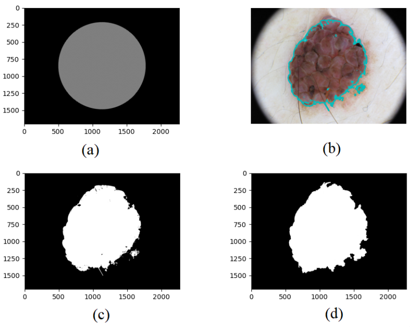

Figure 2: Segmentation result of using morphological snakes: (a) initial curve, (b) evolution of the curve, (c) final segmentation map, and (d) final segmentation map after applying morphological operations.

models can be re-used to vastly speed up the training process, this concept is called transfer learning. Further, the convolution layers of the pre-trained models can be re-used in the encoder layers of the segmentation model. In this work, a Resnet50 is adopted which was pre-trained on the ImageNet dataset(Russakovsky et al., 2015). This dataset is widely used when dealing with medical imaging problems since it contains a repository of annotated clinical (radiology and pathology) images. Thus, the U-net model will use the weights of a pre-trained ResNet50 model systematically having knowledge about the disposed features.

CNN has known many improvements over time, one common trend in research was to add more regularization layers to overcome overfitting. AlexNet has five convolutional layers, however, other architectures have gone deeper to ameliorate their performance, such as VGG and GoogleNet networks. Therefore, increasing the convolutional layers does not regularly improve the performance, contrarily, its performance can get saturated due to vanishing gradient issues. Thus, ResNet was introduced to overcome this issue by adding identity shortcut connections and skipping one or more convolutional layers as shown in Fig.3 (He et al., 2016a). Then the residual block has been ameliorated generating a pre-activation variant of the residual block where the gradient can flow through

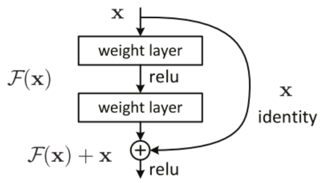

Figure 3: The residual block.

the shortcut connections to any other earlier convolutional layer(He et al., 2016b).

Fig.4 illustrates the segmentation process using the U-Net-based ResNet50 encoder.

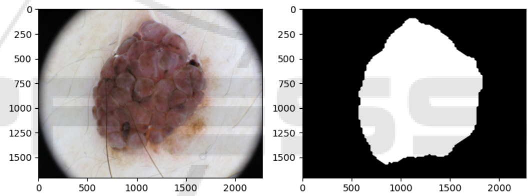

Figure 4: Semantic segmentation result when using the U-net architecture based ResNet50 encoder: Original image, and segmentation map.

2.4 Majority Voting

After applying the three techniques separately, a fusion method on the three segmentation maps is used. The fusion of multiple result predictions has been used by many researchers as an accurate strategy to ameliorate the performance of many pattern recognition systems, rather than relying on one single model prediction Many techniques have been proposed such as majority voting, bayesian decision fusion(Yue et al., 2012), or Dempster-Shafter theory(Kang et al., 2020) for fuzzy information fusion. In this paper, we used the majority voting method which is one of the basic and intuitive approaches. It assigns a sample based on the most frequent class assignment(Ballabio et al., 2019). For semantic segmentation, each pixel in the output of the majority voting process is assigned a background or foreground class based on the three other corresponding pixels of the three segmentation maps issued by thresholding, snakes, and U-Net-based ResNet.

3 EXPERIMENTS AND RESULTS

Experiments used to evaluate the performance of DL models are not only an academic practice but even when creating commercial systems, experiments are necessary to create the most performant models. Thus, to demonstrate the performance of our proposed method, we have tested the methods separately using different strategies and then we have reported the result of fusing them. The experiments were conducted using the ISIC2017 (International Skin Imaging Collaboration) dataset(Codella et al., 2018).

3.1 Technical Materials

Implementing these methods was done using Keras, Sklearn, OpenCV, and TensorFlow frameworks. The Keras and TensorFlow are used on a single machine with an NVIDIA GEFORCE GTX-1050 Ti.

3.2 Dataset Representation

The experiments were conducted using a public dataset which is collected from the ISIC archive, the dataset was taken from a competition on skin lesion segmentation that held in 2017(Codella et al., 2018). The images were collected from a dermatoscopy tool, and it contains about 2000 images of malignant and benign lesions with their corresponding segmentation maps. The dermoscopic lesion images are in JPEG format, the size of almost all images is 1022×767 pixels. Their ground truth images are mask images in PNG format. They are encoded as grayscale(8-bit), where each pixel is 0 for pixels belonging to the healthy skin(Background) and 255 for pixels belonging to the suspicious lesion.

3.3 Evaluation Strategy

Our proposed method consists of fusing different types of methods, those that are trainable and those that use classical image processing methods. To demonstrate the performance of our proposed method the data set is split into two groups: a testing set containing 600 images and 1400 images used for building the DL model. The selected test images are used to test the DL model and to evaluate the performance of the two other methods, as well. The normal procedure when evaluating a hypothesis is to validate it. The validation helps to formulate the hypothesis to see other proprieties and relations that can make it better understood. Thus, this process will be iteratively repeated to find a better version of the hypothesis, also the same process is needed when creating ML and DL models. While training the DL model, a validation step should take place at each epoch. It is necessary to validate its performance to report how well the model is learning. Yet, a part of the training set will be held out for validation, and when the experiments seem to be successful, the final evaluation will be done on the testing set. Therefore, splitting the training set into two sets one for training and the other for validation may fail in accurately building the best machine learning model due to the reduction of the number of samples used for learning. Many validation strategies are introduced for evaluating the model(James et al., 2013), such as k-fold,leave-one-out Cross-Validation, etc. The latter strategy consists of validating one image using the model that has been trained on the other remaining images. However, the k-fold Cross-Validation is more interesting and relevant since it minimizes the number of validation processes by dividing the N data points into k subsets. The training and the validation are then repeated k iterations, in each iteration one of the subsets is used for validation and the remaining for training. The error is also averaged over all the k trials. Basically, k can be either 5 or 10. Thus, we have used this strategy for the validation of the DL model. We have also evaluated its performance by splitting the dataset into two sets: 1250 sample images for training and 150 for validation. This is the same conventional splitting strategy used in(Codella et al., 2018).

3.4 Experimental Results

Fine-tuning the parameters depends on the nature of the ML model. For deep learning, generally, the number of epochs will vary, simultaneously the training and validation accuracy will be computed. This experiment shows how the DL model is proceeding at different epoch numbers. Regarding the experiments done by Md Zahangir Alom *et al.*(Alom et al., 2018), the training and validation accuracy becomes slightly stable when using more than 50 epochs. Thus, this number is adopted in this work.

We have trained the U-Net based Resnet encoder with 1250 images, the 150 images will be used for validation during training with a batch size of 12, we have used the ADAM optimizer technique. Also, the images were resized to 256×256 pixels and augmented by applying many transformations such as rotation, scale, color, and flipping. Using these transformations on the images will increase the size of the dataset, automatically ameliorating the training

Figure 5: Validation and training accuracy using U-net based ResNet50.

Figure 6: Validation and training accuracy using DeepLabv3+ based ResNet50.

performance. Fig5 shows both the training and validation accuracy when using 50 epochs. The accuracy is stable for both validation and training. We have also tested the DeepLabv3+ based Resnet50 encoder(Chen et al., 2018). The training and validation accuracy are shown in Fig.6 which shows many oscillations compared to Fig.5. It is caused by the use of a small batch. It is observed that when using a small batch for some models, it will appear a huge degradation in the training process. Thus, the batch size is an important hyperparameter that leads to either sharp or flat minimizers. We have used many strategies for evaluating the performance of the proposed method. For that, using the U-Net based ResNet50 encoder with a batch size of 12 is good enough for accurately training the model and for minimizing the amount of memory consumption. For a deeper study of the U-net architecture, we have used more than one validation strategy with 20 epochs. First, training and validation were done using the conventional splitting strategy. Validation was performed using randomly selected 150 images. Another strategy is the k-fold cross-validation was performed. Using the k-fold strategy can notably increase the diversity of data used for both validation and training. The model will

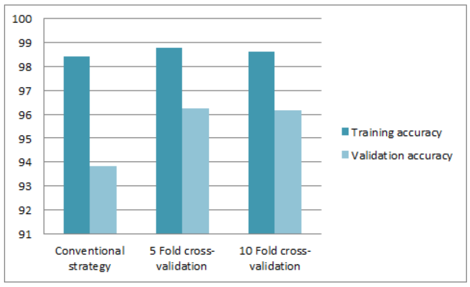

Figure 7: Validation and training accuracy using U-net based ResNet50 with different validation strategies.

be more generalized and unbiased. The training accuracy for both k-fold strategies is slightly higher than using the conventional strategy (see Fig.7). However, the validation accuracy is considerably greater when using k-fold cross-validation for k=5 and k=10. Evidently, models were stable in their predictions due to the effectiveness of the training phase and the diversity of validation samples.

After training the U-Net based Resnet50 model using 1250 images during the 50 epochs, the network weights are saved when the validation accuracy is at its highest. These weights will be used then to make predictions on the 600 testing images. Each one of the three techniques will result in an output segmentation mask. Table1 reports the quantitative results of the experiments. The accuracy and IoU scores are shown for each technique as well as for fusing them. Fig.8 highlights some of the output examples of our proposed method, it is obvious that the fifth column contains the most accurately segmented skin lesions compared to the output of the DL model predictions.

Fig.8 shows that the DL model segments the lesion accurately when thresholding fails in segmenting it. However, using majority voting to fuse the three techniques can refine the segmentation result. The irregularity is an important border-based feature to recognize Melanoma, the proposed method can perfectly draw attention to sharp borders contrarily to the DL model. The ISIC2017 dataset has been proposed in the 2017 skin lesion segmentation challenge. Yading Yuan has proposed a framework based on a deep fully convolutional-deconvolutional neural network(CDNN) to automatically detect skin lesions(Yuan, 2017). This method yielded a Mean IoU of 0.784 on the testing set, our proposed method provides a testing Mean IoU of 0.912 with a high accuracy of 0.983.

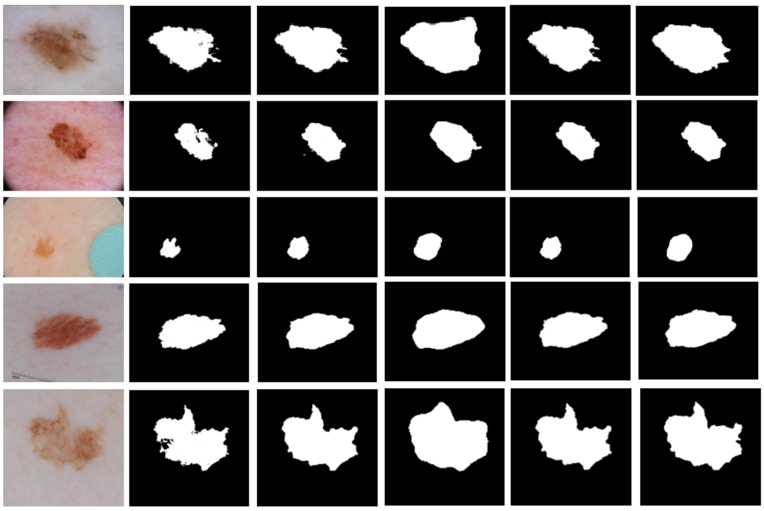

Figure 8: Experimental results for skin lesion segmentation: (a) original lesion image, the segmentation map from: (b) thresholding, (c) snakes, (d) U-net based Resnet50, (e) the proposed method, and (f) ground truth segmentation mask.

Table 1: Testing accuracy of the different methods and fusing them using majority voting.

Method	Thresholding	Snakes	ResNet	Majority Voting
Accuracy	87.61%	87.45%	96.64%	98.31%
Mean IoU	66.18%	66.66%	89.01%	91.29%

4 CONCLUSIONS

Statistics on melanoma incidence rates show that it can be a non-life-threatening cancer if detected early. Early diagnosis and treatment significantly improve the chances of successful recovery. The healthcare community has been highly invested in creating automatic systems that can help doctors and patients quickly and accurately inspect melanoma. These systems are built based on image processing and machine learning techniques. One typical process is segmenting the skin lesion that will be inspected and then recognizing it as malignant or benign. In this work, we have proposed a new method for lesion segmentation that fuses three different methods: thresholding using Otsu's algorithm, morphological snake, and U-Net-based Resnet50. Data fusion, specifically decision fusion combines the decisions of multiple decisions into a common one. The majority voting was used to generate the final segmentation map. Using other decision fusion techniques can be adapted further to ameliorate the performance of the proposed system. The U-Net-based Resnet50 can be also explored using data fusion in the encoder or decoder units.

ACKNOWLEDGEMENTS

The authors would like to appreciably thank all the members of LIS laboratory, for their direct help and collaboration. This research was done during an internship at LIS laboratory, SeaTech, University of Toulon.

REFERENCES

Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., and Asari, V. K. (2018). Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*.

Armato, S. G., Li, F., Giger, M. L., MacMahon, H., Sone, S., and Doi, K. (2002). Lung cancer: performance of automated lung nodule detection applied to cancers missed in a ct screening program. *Radiology*, 225(3):685–692.

Ashwin, S., Ramesh, J., Kumar, S. A., and Gunavathi, K. (2012). Efficient and reliable lung nodule detection using a neural network based computer aided diagnosis system. In *2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM)*, pages 135–142. IEEE.

Bae, K. T., Kim, J.-S., Na, Y.-H., Kim, K. G., and Kim, J.-H. (2005). Pulmonary nodules: automated detection on ct images with morphologic matching algorithm—preliminary results. *Radiology*, 236(1):286–293.

Ballabio, D., Todeschini, R., and Consonni, V. (2019). Recent advances in high-level fusion methods to classify multiple analytical chemical data. *Data Handling in Science and Technology*, 31:129–155.

Chan, T. and Vese, L. (1999). An active contour model without edges. In *International Conference on Scale-Space Theories in Computer Vision*, pages 141–151. Springer.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818.

Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al. (2018). 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 *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 168–172. IEEE.

Daghrir, J., Tlig, L., Bouchouicha, M., and Sayadi, M. (2020). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–5. IEEE.

He, K., Zhang, X., Ren, S., and Sun, J. (2016a). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

He, K., Zhang, X., Ren, S., and Sun, J. (2016b). Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer.

Heinzerling, L. and Eigentler, T. K. (2021). Skin cancer in childhood and adolescents: Treatment and implications for the long-term follow-up. In *Late Treatment Effects and Cancer Survivor Care in the Young*, pages 349–355. Springer.

Jadhav, A. R., Ghontale, A. G., and Shrivastava, V. K. (2019). Segmentation and border detection of melanoma lesions using convolutional neural network and svm. In *Computational Intelligence: Theories, Applications and Future Directions-Volume I*, pages 97–108. Springer.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Jamil, U., Sajid, A., Hussain, M., Aldabbas, O., Alam, A., and Shafiq, M. U. (2019). Melanoma segmentation using bio-medical image analysis for smarter mobile healthcare. *Journal of Ambient Intelligence and Humanized Computing*, 10(10):4099–4120.

Kang, B., Zhang, P., Gao, Z., Chhipi-Shrestha, G., Hewage, K., and Sadiq, R. (2020). Environmental assessment under uncertainty using dempster–shafer theory and z-numbers. *Journal of Ambient Intelligence and Humanized Computing*, 11(5):2041–2060.

Khazaei, Z., Ghorat, F., Jarrahi, A., Adineh, H., Sohrabivafa, M., and Goodarzi, E. (2019). Global incidence and mortality of skin cancer by histological subtype and its relationship with the human development index (hdi); an ecology study in 2018. *World Cancer Res J*, 6(2):e13.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., and Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88.

Maglogiannis, I. and Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE transactions on information technology in biomedicine*, 13(5):721–733.

Marquez-Neila, P., Baumela, L., and Alvarez, L. (2013). A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1):2–17.

Mbarki, W., Bouchouicha, M., Frizzi, S., Tshibasu, F., Farhat, L. B., and Sayadi, M. (2020). Semi-automatic segmentation of intervertebral disc for diagnosing herniation using axial view mri. In *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–6. IEEE.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252.

Schadendorf, D., van Akkooi, A. C., Berking, C., Griewank, K. G., Gutzmer, R., Hauschild, A., Stang, A., Roesch, A., and Ugurel, S. (2018). Melanoma. *The Lancet*, 392(10151):971–984.

Sharma, N. and Ray, A. K. (2006). Computer aided segmentation of medical images based on hybridized approach of edge and region based techniques. In *Proc. Int. Conf. Math. Biol*, pages 150–155.

Yuan, Y. (2017). Automatic skin lesion segmentation with fully convolutional-deconvolutional networks. *arXiv preprint arXiv:1703.05165*.

Yue, D., Guo, M., Chen, Y., and Huang, Y. (2012). A bayesian decision fusion approach for microrna target prediction. *BMC genomics*, 13(8):1–11.