Melanoma Recognition and Lesion Segmentation using Multi-Instance Learning

Yangling Ma¹ and Zhouwang Yang^{2*†}

¹School of Mathematical Sciences, Suzhou University of Science and Technology, Suzhou Science and Technology City, Suzhou, 215163, Jiangsu, China.

^{2*}School of Mathematical Sciences, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, 230026, Anhui, China.

*Corresponding author(s). E-mail(s): yangzw@ustc.edu.cn;

Contributing authors: yangma@mail.ustc.edu.cn;

[†]These authors contributed equally to this work.

Abstract

Melanoma is one of the deadliest forms of skin cancer, but early and accurate identification can significantly improve the survival rate of patients. In this paper, an end-to-end framework based on multi-instance learning is proposed for melanoma recognition and lesion segmentation simultaneously. To make full use from the information of high-resolution images, we take each image block (super-pixel) as an instance in a bag and use multi-instance learning based on a graph convolutional network to recognize melanoma. Moreover, skin lesion segmentation is derived from attention weights and is calibrated by classification probability vectors. As a result, the AUC of our method for melanoma recognition reaches 0.93, which is much higher compared with other related methods. Also, the Jaccard index (JA) of our method for melanoma-related skin lesion segmentation reaches 0.699. In our end-to-end approach, segmentation and recognition are treated as intimately coupled processes, and hence, a high JA is also an indication of the reliability of melanoma recognition. Collectively, these findings confirmed that our method effectively assists melanoma diagnosis.

Keywords: Multi-instance learning, Melanoma recognition, Skin lesion segmentation, Structural relationship

1 Introduction

Melanoma is one of the deadliest forms of skin cancer, which accounts for about 75% of skin cancer deaths [1]. Early and accurate identification allows doctors to propose treatment plans for melanoma as soon as possible, effectively improving the survival rate of patients [2]. However, manual detection of melanoma has created a huge demand for well-trained experts, and there are differences between different observers. Existing dermoscopy techniques have improved diagnostic performance of melanoma [3]. Dermoscopy is a non-invasive skin-imaging technology that obtains enlarged high-definition images of skin regions to improve the clarity of spots [4] and enhances visualization of skin lesions by removing surface reflections. However, automatically identifying melanoma from dermoscopy images is still a task with many challenges. First, the low contrast between skin lesions and normal skin areas makes it difficult to segment lesion regions accurately. Second, melanoma and non-melanoma lesions may have a high degree of visual similarity, making it difficult to distinguish melanoma lesions from non-melanoma lesions. Third, changes in skin conditions, such as patient's skin color, hair, or veins, can produce different melanoma appearances in terms of color and texture. Methods based on deep learning that incorporate high-resolution images to recognize melanoma can alleviate these challenges. However, existing technologies for melanoma recognition in high-resolution images rely on image segmentation, which is regarded as merely a preprocessing, background-removing strategy that does not use disease-related information and lacks interpretability for the recognition results. Therefore, a reliable, automatic detection system for melanoma recognition would not only improve accuracy and efficiency for pathologists, but also be of great research significance.

There are two main methods for melanoma recognition and lesion segmentation. One method is melanoma recognition based on skin image segmentation results [5–7], which heavily depends on the performance of a segmentation algorithm. Another method for melanoma recognition and lesion segmentation is based on an end-to-end framework that uses resized or cropped images, which loses information from original high-resolution images [8–11].

One method based on skin images segmentation heavily depends on the performance of a segmentation algorithm, which is a key step [5, 12]. Accurate segmentation is conducive to the subsequent accurate recognition of skin lesions. Many researchers use fuzzy c-means clustering, mean shift, hybrid threshold technology, and the fully convolutional residual network (FCRN) for solving skin lesion segmentation in dermoscopic images [6, 7, 13, 14]. In addition, some researchers [15, 16] first use skin images segmentation algorithms, like FCRN, to obtain cropped images and then use deep networks and other methods to extract fused features for melanoma recognition.

Other methods for melanoma recognition and skin lesion segmentation are end-to-end frameworks based on deep networks and resized or cropped images; these methods are of important research significance, yet they do not retain all information from the original high-resolution images [8, 17]. In a recent study,

Codella et al. [8] establish a system that combines the latest advances in deep learning and machine learning methods for melanoma-related skin lesion image segmentation and recognition. Li et al. [9] propose a Lesion Indexing Network (LIN) composed of a multi-scale fully convolutional residual network and a lesion index calculation unit (LICU), which uses cropped images to conduct melanoma-related skin lesion segmentation and recognition simultaneously.

In the field of melanoma-related skin lesion image analyses, although many strategies have been proposed, there is still room for improvement in melanoma recognition and lesion segmentation. Some methods use cropped or resized images for melanoma recognition and skin lesion segmentation, whereby information from the original high-resolution images is lost. We regard the original high-resolution image as a bag and each image block (super-pixel) as an instance in the bag, which solves the information utilization problem for high-resolution images. Other melanoma-recognition methods based on skin segmentation using high-resolution images highly depend on the performance of a segmentation algorithm, and these recognition results lack interpretability. In this study, we propose an end-to-end framework that uses a multi-instance learning method based on a graph convolutional network [18] (GCN-based MIL) to carry out melanoma recognition and skin lesion segmentation. We use skin lesion segmentation results to explain skin lesion categories obtained by a classification model. Moreover, our end-to-end framework does not require cropping or resizing. The main contributions are as follows:

- We propose an end-to-end framework based on multi-instance learning to conduct melanoma recognition and lesion segmentation simultaneously.
- Our segmentation is derived from attention weights and is calibrated by classification probability vectors. Moreover, the segmentation results also offer a reliable interpretation for melanoma recognition.
- For melanoma-related skin lesion images analyses, we take each image block (super-pixel) as an instance in a bag to overcome the difficulty of making full use of the information from high-resolution medical images.

The remaining parts of this paper are as follows: in Section 2, we propose an end-to-end method based on MIL for melanoma recognition and lesion segmentation; in Section 3, our results and analyses of melanoma recognition and lesion segmentation are presented; in Section 4, discussions about our method and other methods are offered; and our conclusion and future prospects are given in Section 5.

2 Methodology

In this study, we used the publicly available ISIC 2017 competition dataset [19] for experiments. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Fig. 1 The framework used for melanoma-related skin lesion recognition and segmentation.

In this section, we introduce the framework used for melanoma recognition and skin lesion segmentation. As shown in Fig. 1, we first use some strategies to preprocess skin images, then use GCN-based MIL [18] to classify skin lesion images, and finally use attention weights and category probability vectors to segment skin lesion images. Fig. 1 shows the whole framework of our algorithm.

2.1 Data Preprocessing

The International Skin Imaging Collaboration (ISIC) [19] focuses on automated skin lesion analysis, and since 2016, its dataset size has continued to expand. In the ISIC 2017 competition [19], in order to improve the accuracy of automatic detection methods for melanoma, they released an annotated dataset of three tasks related to skin lesion images. The ISIC 2017 dataset [19] contains three types of skin lesions, i.e. melanoma, seborrheic keratosis, and nevus. Since there are different numbers of images in different categories, we rotate images belonging to different categories to establish a class-balanced dataset. The dataset that is expanded through this data preprocessing method is called DR. Table 1 lists the numbers from the original training dataset and DR. The digits in brackets after the category names refer to the rotation angle. This data preprocessing method is described in [9]. In addition, when training a classification model, we randomly flip images in DR along the x-axis or y-axis, and then randomly rotate images by 90 degrees. During training, we use these data preprocessing strategies, as shown in the module on the far left of Fig. 1.

Table 1 Details of data enhancement.

	Melano- ma (18)	Seborrheic Keratosis (18)	Nevus (45)
Training Dataset	374	254	1372
DR	7480	5080	10976

2.2 Melanoma Recognition

To reduce the re-labeling of image blocks, we regard the original image as a bag and the image blocks taken from the image as the instances. Since the ISIC 2017 dataset contains image data, we use the adjacent relationships among the image blocks in the original image space domain to establish a graph structure in the bag, and then use GCN-based MIL [18] to classify skin lesion images in the three classification task.

As shown in Fig. 1, the yellow-dotted box and the red-realization box represent the melanoma recognition flowchart. The specific algorithmic steps for melanoma recognition are as follows. In first step, we extract $k \times k$ image blocks from the original high-resolution image; each image block represents the central pixel, and the moving stride is s . As shown in the left image in Fig. 2, we select an image block from the original image, which represents an instance. Then, as shown in the right image in Fig. 2, for the adjacency relationship between the red image block and the blue image blocks, we combine Eq. (1) to generate the adjacency matrix A in the graph structure.

Fig. 2 Image block selection method

$$A_{mn}^i = \begin{cases} 0, & y_m^i - y_n^i > \sqrt{2}k, \\ 1, & \text{otherwise,} \end{cases} \quad (1)$$

where y_m^i represents the center coordinates of m image block in i original image, and y_n^i represents the center coordinates of n image block in i original image. Assuming the resolution of the original image is 1022×767 , the coordinates of the upper left corner of the image matrix are $(0, 0)$, the coordinates of the lower right corner are $(1022, 767)$, and the coordinates of the other pixels are deduced by analogy.

In the second step, we use ResNet18 [20] to extract initial feature vectors from the image blocks. In the third step, we use a two-layer graph convolutional network [21] to extract the feature vectors from the image blocks to embed the relationships among the image blocks. In the fourth step, we use the graph-attention mechanism [18] to learn the attention weights of the instances, and then use the attention weights and the feature vectors of the instances to obtain a feature vector of the bag. In the fifth step, we use a fully-connected layer with a Softmax activation function (as shown in Eq. 2) to obtain a category probability vector for the image. The input of this fully-connected layer is a feature vector of the bag, and the output is a category probability vector p for the original skin lesion image. In addition, our classification model structure for melanoma recognition is shown in Table 2.

$$p_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2)$$

Table 2 The classification model structure. The parameters in $GCN(\cdot, \cdot)$ respectively indicate the number of input channels and the number of output channels. The parameters in $FC(\cdot, \cdot)$ respectively indicate the number of input channels and the number of output channels. *GraphAttention* represents the graph attention mechanism in the literature [18], and L represents a hyperparameter.

Layer	Type
1-18	ResNet18
19	$GCN(512, 512)$
20	$GCN(512, 512)$
21	$GraphAttention(L=128)$
22	$FC(512, 3) + \text{Softmax}$

2.3 Skin Lesion Segmentation

When using previous methods to segment skin lesions, the standard methods use mask-level label-based image segmentation algorithms to segment skin images. These methods are used as preprocessing steps for skin lesion image classification and are completely independent of lesion categories. To solve these problems, we use image-level labels to train a classification model and then use this classification model to segment skin lesion images.

As shown in Fig. 1, the solid blue outlined box illustrates skin lesion segmentation flowchart. We use skin a lesion classification model to obtain attention weights α of image blocks and a probability vector p of skin lesion categories, and then we calculate the scores of each pixel in the image block with respect to the skin lesion image. For pixels that are repeated between image blocks, we directly take the maximum product of the attention weight and probability. Moreover, a pixel in an image block corresponds to a pixel position in an original image. We use pixel value 1 to represent a skin lesion

region, and pixel value 0 to represent non-skin lesion region. We return pixel points in image blocks to subscripts in the original image, and we finally obtain lesion regions for skin lesion images. In the black and white image, black represents the background, and white represents skin lesion regions. The specific calculation formula is shown in Eq. (3).

$$Y_{y_{ij}} = \begin{cases} 0, & \text{softmax}(p_{y_{ij}} \times \alpha_{y_{ij}}) < \eta, \\ 1, & \text{otherwise,} \end{cases} \quad (3)$$

where η represents the threshold, y_{ij} represents pixel positions in original images, and softmax represents the Softmax activation function, as shown in Eq. (2).

To select the most suitable threshold value η for melanoma-related skin lesion segmentation, we use a threshold value from 0 to 1.0 every 0.05, and then test the threshold value multiple times on the ISIC 2017 training dataset. Finally, we evaluate the performance of our method and other methods on carrying out melanoma-related skin lesion segmentation on the ISIC 2017 test dataset.

3 Results and Analyses

In this section, we introduce the results and analyses of dataset, evaluation metrics, melanoma recognition, and melanoma-related skin lesion segmentation.

Next, we introduce some experimental settings. When training the melanoma-related skin lesion classification model, we set the initial learning rate to 0.00001, the maximum epoch to 400, the learning rate drop to 0.0001 at 200 epochs, and the batch size to 1; we use the SGD optimizer with 0.9 momentum to update the parameters. For ResNet18, we use the model parameters trained on the ImageNet dataset to initialize the network parameters; for other layers, the weight parameters except for the biases are initialized to 1, the biases are initialized to 0. For different evaluation metrics, we take the average of five results as the final result, which reduces accidental model errors. We train the melanoma-related skin lesion classification model on a machine with two 1080Ti GPUs and 128G RAM.

3.1 Dataset

In this study, we use the publicly available ISIC 2017 competition dataset [19] for experiments. The ISIC 2017 competition consists of three tasks: skin lesion region segmentation, dermal feature extraction, and skin lesion classification. Our method mainly analyzes the first task and the third task from the ISIC 2017 competition. The ISIC 2017 competition provides 2,000 skin lesion images as a training set, including masks for segmentation, super-pixel masks for dermoscopic feature extraction, and labels for classification. Skin lesion images are divided into three categories: melanoma, seborrheic keratosis, and nevus.

Fig. 3 Skin lesion images

Melanoma is a malignant skin tumor that causes high mortality. The other two skin lesions, seborrheic keratosis and nevus are benign skin tumors derived from different cells. As shown in Fig. 3, the first column of images shows the original skin lesion images, and the second column of images shows the corresponding image masks used for lesion region segmentation. The ISIC 2017 competition also provides a publicly available validation dataset and a test dataset for evaluating method performance. Data distributions in the test dataset and the validation dataset are shown in Table 3.

This melanoma-related skin lesion image dataset includes images with different resolutions, such as 1022×767 , 600×450 , 3024×2016 , and 6688×4439 . To better train the melanoma-related skin lesion classification model, we obtain

image blocks with a size of 224×224 using a sliding window, where different strides s are taken according to different image heights h , and specific selections are as follows:

$$s = \begin{cases} 50, & h < 1000, \\ 120, & 1000 \leq h < 2000, \\ 200, & \text{otherwise.} \end{cases} \quad (4)$$

Table 3 ISIC 2017 dataset.

	Melanoma	Seborrheic Keratosis	Nevus
Validation	72	27	200
Test	117	90	393

3.2 Evaluation Metrics

In deep learning, we need to obtain some evaluation metrics for models on the test dataset to measure model generalization ability so as to obtain performance comparisons of different models. Model generalization ability refers to a model's performance ability on a test dataset. In our study, we mainly use Accuracy, Recall/Sensitivity, AUC, and the Jaccard index (JA). In this section, we express specific formulas of these evaluation metrics with binary classification tasks and multi-classification tasks. For the binary classification task, we assume that P represents the number of positive examples in the test dataset, N represents the number of negative examples in the test dataset, TP represents the number of positive examples predicted by the model in positive examples, and FP represents the number of negative examples predicted by the model in positive examples. TN represents the number of negative examples predicted by the model in negative examples, and FN represents the number of positive examples predicted by the model in negative examples. Therefore, formulas for the evaluation metrics are as follows:

$$Accuracy = \frac{TN + TP}{P + N} \quad (5)$$

$$Specificity = \frac{TN}{FP + TN} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$Precision = \frac{TP}{TP + FP}. \quad (8)$$

JA The Jaccard index (JA) is used to calculate the similarity or overlap of two samples. It is used to measure the performance of the segmentation model, the value range is $[0, 1]$. The formula is as follows:

$$JA = \frac{TP}{TP + FP + FN}. \quad (9)$$

AUC AUC refers to the area enclosed by the ROC curve and the abscissa. The closer the AUC value is to 1, the better the model classification performance. The abscissa of the ROC curve is a false positive rate ($FPR = FP/(FP + TN)$); the ordinate is a true positive rate ($TPR = TP/(TP + FN)$). The ROC curve reflects FPR and TPR trends when selecting different judgment category thresholds.

3.3 Melanoma Recognition and Skin Lesion Segmentation

For melanoma recognition and skin lesion segmentation, we use Multi-Channel-ResNet [22], LIN [9] and VGG16 [23], CNMP [16], AlexNet [24], FCN [25], FCRN [7], U-Net [26], and ResNet101 [20] to compare with our method. Among them, LIN [9] uses the full convolution residual network (FCRN) to obtain lesion segmentation results, and performs rough classification and segmentation refinement by calculating distance heat maps. We compute AUC, Accuracy, and Recall on recognition results for melanoma and non-melanoma and then compare them with other methods. In addition, we focus on the results of melanoma-related skin lesions segmentation, statistically count the Jaccard index (JA) on these results, and then compare the results with other methods.

Table 4 Comparison of our method with other methods. The first column of the table represents the different methods; the second column represents whether original high-resolution images are used as naked image inputs (no additional information required); the third column represents whether the method is end-to-end; the fourth column represents whether melanoma recognition is conducted; and the fifth column represents whether skin lesion segmentation is performed.

Methods	Naked Image Input	End-to-end Method	Recognition	Segmentation
Multi-Channel-ResNet [22]	✗	✗	✓	✗
VGG16 [23]	✗	✗	✓	✗
AlexNet [24]	✗	✗	✓	✗
ResNet101 [20]	✗	✗	✓	✗
CNMP [16]	✗	✗	✓	✗
Gessert [17]	✗	✗	✓	✗
FCN [25]	✗	✗	✗	✓
FCRN [7]	✗	✗	✗	✓
U-Net [26]	✗	✗	✗	✓
LIN [9]	✗	✓	✓	✓
GCN-based MIL(ours)	✓	✓	✓	✓

As shown in Table 4, the most important thing is that our method is not only an end-to-end approach to melanoma recognition and skin lesion segmentation, but also the input of this method is original high-resolution images.

Table 5 Melanoma recognition and skin lesion segmentation in different models.

Models	Seborrheic Keratosis-AUC	AUC	Accuracy	Recall	JA
Multi-Channel-ResNet [22]	0.959	0.8743	—	—	—
VGG16 [23]	—	0.892	0.847	0.586	—
AlexNet [24]	—	0.859	0.823	0.343	—
ResNet101 [20]	—	0.869	0.840	0.336	—
CNMP [16]	—	0.920	0.901	—	—
Gessert [17]	—	0.923	0.882	—	—
FCN [25]	—	—	—	—	0.696
FCRN [7]	—	—	—	—	0.607
U-Net [26]	—	—	—	—	0.651
LIN [9]	—	0.912	0.857	0.490	0.753
GCN-based MIL(ours)	0.988	0.930	0.892	0.761	0.699

As shown in Fig. 1, we input an original high-resolution image, divide the image into multiple image blocks, and then input the image blocks into our method to obtain the recognition and segmentation results. Therefore, we overcome the difficulty of making full use of high-resolution image information. In addition, AlexNet [24], VGG16 [23], and ResNet101 [20] use the segmentation method to crop high-resolution images, resize these cropped images, and then feed these processed images into networks to obtain skin lesion classifications. FCN [25], FCRN [7], and U-Net [26] only use resized images to segment images. Multi-Channel-ResNet [22] uses cropped images based on images segmentation and multiple ResNet to obtain skin lesion classifications. CNMP [16] uses fused features based on a variety of traditional image feature extraction methods and deep networks to obtain skin lesion classifications; the input of this method is cropped images based on image segmentation. LIN [9] is an end-to-end approach to melanoma-related skin lesion image classification and segmentation; however, the input of this method is multiple scale images that consist of center cropped images based on background removal algorithms.

As shown in Table 5, our method yields better recognition and segmentation results compared other methods. Our method boats a high Recall for melanoma recognition, which helps prevent failed diagnoses or misdiagnoses. This is useful for assisting doctors in the treatment of melanoma. Our method makes full use of information in high-resolution images and structural relationships among image blocks, therefore, our method has a higher Accuracy and AUC in melanoma recognition tasks, which shows that our method's overall performance exceeds others. In addition, it can be seen from AUC that our method is also better than the multi-channel ResNet model fusion method [22] in seborrheic keratosis recognition tasks. These results indicate our method is most successful for melanoma recognition in melanoma-related skin lesion recognition tasks. In addition, for skin lesion segmentation, our method's JA is superior to FCN [25] and U-Net [26]. Although our method is not as effective as LIN [9], it does not need multiple scale images with center cropped images based on background removal algorithms. Moreover, as shown in Fig. 4, skin lesion segmentation results provide an intuitive explanatory display for melanoma recognition. As shown in Table 4 and Table 5, in terms of method

input, processing tasks, and numerical results, our method is more suitable for practical scenarios and yields better output.

Fig. 4 Segmentation diagram of melanoma lesion region images.

To visually analyze the segmentation and recognition performance of our method, Fig. 4 and Fig. 5 respectively show segmentation results for melanoma images and non-melanoma images. As shown in Fig. 4, the appearance of these lesion regions also fully explains why the recognition results have identified melanoma. These skin lesion images are recognized by our method as both melanoma and non-melanoma. Green and red lines respectively represent segmentation contours predicted by our method and contours marked by expert doctors. These examples illustrate some of the primary challenges in the skin lesion image processing field. In Case 2 and Case 3 in Fig. 4, the contrast discrimination between the lesion and the skin region is low. The hair near the lesion region in Case 4 in Fig. 4 affects the segmentation result. For Case 1 and Case 2 in Fig. 4, and Case 3 in Fig. 5, manual labeling results are another type of noise information that is used for lesion region segmentation. However, it can be seen from Fig. 4 and Fig. 5 that our method has achieved satisfactory segmentation results for all challenging cases.

4 Discussion

In this section, we introduce discussions about our method and others.

Fig. 5 Segmentation diagram of non-melanoma lesion region images

4.1 Comparison

In the field of melanoma-related skin lesion image analyses, although many strategies have been proposed, there is still room for improvement in melanoma recognition and lesion segmentation. Melanoma-related skin lesion recognition methods based on skin images segmentation heavily depend on the performance of a segmentation algorithm, which is a key step [5]. Accurate segmentation is conducive to the subsequent accurate recognition of skin lesions [16]. An end-to-end framework based on deep networks and resized images for melanoma recognition and skin lesion segmentation has important research significance, but it loses information from original high-resolution images [3, 8]. In our study, we regard an original high-resolution image as a bag, and each image block (super-pixel) as an instance in the bag, which solves the information utilization problem for high-resolution images. In addition, we propose an end-to-end framework by using GCN-based MIL [18] to carry out melanoma recognition and skin lesion segmentation, and the skin lesion segmentation results are used to explain skin lesion categories obtained by a classification model (as shown in Fig. 4). As shown in Table 4 and Table 5, in terms of method input, processing tasks, and numerical results, our method is more suitable for practical scenarios and yields better output. Collectively, these findings confirm that our method effectively facilitates melanoma diagnosis.

4.2 Limitation and Future Work

In the proposed method, the segmentation task is only calibrated on melanoma data, thus it often produces unexpected results on non-melanoma data. As shown in Fig. 5, with an image recognized as non-melanoma, the hairs are segmented as lesion regions. In the future, we will consider constructing another threshold for the segmentation model on non-melanoma data and apply segmentation according to this threshold on images classified as non-melanoma. Moreover, the proposed recognition and segmentation models for skin lesion images only leverage vision information. Text information, such as age, gender, and recorded medical history, would be an excellent supplement. In the future, we will consider merging such text information into the graph structure to improve the overall performance. On the other hand, we are also able to use the proposed method for other problems, such as fault diagnosis of electric impact drills [27].

5 Conclusions

In melanoma-related skin lesion image analysis, we propose an end-to-end framework based on an original high-resolution image in order to recognize melanoma images and segment lesion regions simultaneously. We regard the original image as a bag, the image blocks as instances, and then use the GCN-based MIL [18] method to obtain a probability vector of the image and attention weights of the image blocks. Finally, we use a probability vector of the image and attention weights of the image blocks to obtain recognition and segmentation results from the original image. Our method surpasses other methods in the melanoma recognition task in the ISIC 2017 competition; without human mask-level labeling, our method is more supervised than some methods for the skin lesion segmentation task in the ISIC 2017 competition. In addition, lesion region segmentation results provide an explanation for skin lesion categories. In the skin lesion image segmentation and recognition, there is not only image information, but also information such as age, gender, and medical history from a doctor's inquiry. In the future, we can merge this text information into the graph structure, and then use the multi-instance learning method based on a graph convolutional network to classify and segment skin lesion images.

Declarations

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Compliance with Ethical Standards

Compliance with Ethical Standards:

- **Funding:** This study had no funding.
- **Conflict of Interest:** The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.
- **Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.
- **Informed consent:** Informed consent was obtained from all individual participants included in the study.

References

- [1] Jerant, A.F., Johnson, J.T., Sheridan, C.D., Caffrey, T.J.: Early detection and treatment of skin cancer. *American family physician* **62**(2), 357–368 (2000)
- [2] Namozov, A., Im Cho, Y.: Convolutional neural network algorithm with parameterized activation function for melanoma classification. In: 2018 International Conference on Information and Communication Technology Convergence (ICTC), pp. 417–419 (2018). IEEE
- [3] Yu, Z., Jiang, X., Zhou, F., Qin, J., Ni, D., Chen, S., Lei, B., Wang, T.: Melanoma recognition in dermoscopy images via aggregated deep convolutional features. *IEEE Transactions on Biomedical Engineering* **66**(4), 1006–1016 (2019). <https://doi.org/10.1109/TBME.2018.2866166>
- [4] Celebi, M.E., Wen, Q., Iyatomi, H., Shimizu, K., Zhou, H., Schaefer, G.: A state-of-the-art survey on lesion border detection in dermoscopy images. *Dermoscopy image analysis* **10**, 97–129 (2015)
- [5] Erkol, B., Moss, R.H., Joe Stanley, R., Stoecker, W.V., Hvatum, E.: Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. *Skin Research and Technology* **11**(1), 17–26 (2005)
- [6] Zhou, H., Schaefer, G., Sadka, A., Celebi, M.E.: Anisotropic mean shift based fuzzy c-means segmentation of skin lesions. In: Proceedings of the 5th International Conference on Soft Computing as Transdisciplinary Science and Technology, pp. 438–443 (2008)
- [7] Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P.A.: Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging* **PP**(99), 994–1004 (2016)

- [8] Codella, N.C.F., Nguyen, Q.B., Pankanti, S., Gutman, D.A., Helba, B., Halpern, A.C., Smith, J.R.: Deep learning ensembles for melanoma recognition in dermoscopy images. *Ibm Journal of Research & Development* **61**(4), 5–1515 (2017)
- [9] Li, Y., Shen, L.: Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* **18**(2), 556 (2018)
- [10] Harangi, B.: Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of Biomedical Informatics* **86**, 25–32 (2018)
- [11] Zhang, R.: Melanoma detection using convolutional neural network. In: 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 75–78 (2021). <https://doi.org/10.1109/ICCECE51280.2021.9342142>
- [12] Canalini, L., Pollastri, F., Bolelli, F., Cancilla, M., Allegretti, S., Grana, C.: Skin lesion segmentation ensemble with diverse training strategies. In: International Conference on Computer Analysis of Images and Patterns, pp. 89–101 (2019). Springer
- [13] Zhou, H., Li, X., Schaefer, G., Celebi, M.E., Miller, P.: Mean shift based gradient vector flow for image segmentation. *Computer Vision Image Understanding & Cviu* **117**(9), 1004–1016 (2013)
- [14] Pennisi, A., Bloisi, D.D., Nardi, D., Giampetruzzi, A.R., Mondino, C., Facchiano, A.: Skin lesion image segmentation using delaunay triangulation for melanoma detection. *Computerized Medical Imaging & Graphics*, 89–103 (2016)
- [15] Kawahara, J., BenTaieb, A., Hamarneh, G.: Deep features to classify skin lesions. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1397–1400 (2016). IEEE
- [16] Lai, Z.F., Deng, H.F.: Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. *Computational Intelligence and Neuroscience* **2018**, 1–13 (2018)
- [17] Gessert, N., Nielsen, M., Shaikh, M., Werner, R., Schlaefer, A.: Skin lesion classification using ensembles of multi-resolution efficientnets with meta data. *MethodsX* **7**, 100864 (2020)
- [18] Ma, Y., Yang, Z.: Multi-instance learning by utilizing structural relationship among instances. *arXiv preprint arXiv: 2102.01889* (2021)
- [19] Codella, N., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A.,

- Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H.: 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) (2018)
- [20] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
- [21] Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
- [22] Guo, S., Yang, Z.: Multi-channel-resnet: An integration framework towards skin lesion analysis. *Informatics in Medicine Unlocked* **12**, 67–74 (2018)
- [23] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
- [24] Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
- [25] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
- [26] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 234–241 (2015). Springer
- [27] Glowacz, A.: Fault diagnosis of electric impact drills using thermal imaging. *Measurement* **171**, 108815 (2021)