Semi-automatic Segmentation of Skin Lesions based on Superpixels and Hybrid Texture Information

Elineide Silva dos Santos* and Rodrigo de Melo Souza Veras*
*Department of Computing, Federal University of Piauí - Teresina, Brazil

***Abstract*—This article[1] exposes a semi-automatic method with the potential to aid the doctor while supervising the progression of skin lesions. The proposed methodology pre-segments skin lesions using the SLIC0 algorithm for the generation of superpixels. Following this, each superpixel is represented using a descriptor constructed of a mix from GLCM and Tamura texture features. The feature's gain ratios were utilized to choose the data applied in the semi-supervised clustering algorithm Seeded Fuzzy C-means. This algorithm uses certain specialist-marked regions to group the superpixels into lesion or background regions. Finally, the segmented image undergoes a post-processing step to eliminate sharp edges. The experiments were performed on a total of 3974 images. We used the 2995 images from PH^2, DermIS and ISIC 2018 datasets to establish our method's specifications and the 979 images from ISIC 2016 and ISIC 2017 datasets for performance analysis. Our experiments demonstrate that by manually identifying a few percentages of the generated superpixels, the proposed approach reaches an average accuracy of 95.97%, thus giving a superior performance to the techniques presented in the literature. Even though the proposed method requires physicians' intervention, they can obtain segmentation results similar to manual segmentation from a significantly less time-consuming task.**

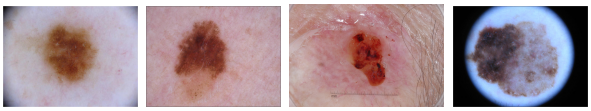
I. Introduction

The human body is similar to a machine made up of different components that work together to perform its functions. The main functions of the skin are excretory, protective, relationship, thermoregulatory, and metabolic. In skin areas, repeatedly over many years, exposed to the sun could occur skin cancer lesions.

These lesions could be benign or malignant. Benign lesions have more regular edges and uniform coloring (Figure 1a), whereas malignant lesions have more irregular borders and varied coloring (Figure 1b). Malignant lesions can lead to skin cancers, such as melanoma. Melanoma is aggressive cancer with easy proliferation. In a report, the Skin Cancer Foundation [1], estimated new cancer cases in 2021 increase by 5.8 percent. Also, it estimated more than 200,000 new cancer cases in the United States, with 101,280 involving melanoma at its most severe stage.

The detection of skin cancer is frequently made based on images that professionals visually analyze. However, manual image segmentation is a time-consuming task that demands skilled and trained specialists. Computer-assisted diagnosis systems (CAD) combines techniques related to artificial intelligence and digital image processing to help specialists in their clinical tasks [2].

[1]M.Sc. dissertation conducted between March 2019 and June 2021



(a) Benign lesions. (b) Malignant lesions.

Fig. 1. Skin lesion images samples.

Lesion segmentation an essential step in the CAD system to classify skin lesions. The image rising from this step provides the lesion classification characteristics as benign or malignant. According to [3], segmentation of a lesion's boundaries is vital to locate it accurately within the image under analysis and diagnosis. Also, the ABCD (asymmetry, border, color, and diameter) rule can help non-dermatologists differentiate between benign melanocytic naevi and melanoma lesions [3]. Recently, [4] demonstrated the importance of the ABCD rule in the automatic diagnosis of melanoma.

Deep learning has increased in demand in medical imaging analysis [5], and these systems also use a transparent segmentation step to the user. In some cases, the lesion under analysis is segmented before being classified by a convolutional neural network (CNN) [6].

There are various algorithms adopted in the processing and analysis of skin lesions images. However, researchers commonly applied strategies that depended on the image dataset characteristics used. Sometimes these datasets are private. Segmentation becomes more complicated when there is greater diversity in the skin lesion images under analysis, such as when the images are from people of different races or acquired under other conditions.

This article proposes a semi-automatic method for skin lesions segmentation from a wide diversity of skin lesion images collected by combining several publicly available image datasets. We executed various experiments with different texture descriptor algorithms to develop our method, including combining feature vectors. Our approach applies a pre-processing step for hair extraction and a pre-segmentation step based on superpixels. The specialist then selects a few of the generated superpixels, and the seeded fuzzy C-means (SFc-means), proposed by [7], cluster them into skin lesions or parts of the image background. Finally, we remove noise and specific fragments in the edges of the lesion.

The article is structured as follows. Section II shows a general survey of related works; Section III describes the proposed approach; Section IV gives more details of the im-

ages dataset, the evaluation metrics adopted, and the achieved results; Section V discusses the findings. Assumptions and recommendations for further studies are seen in the Section VI.

II. Related Work

Diverse methods were proposed regarding the skin lesions segmentation in images. According to the literature [8]–[10], the most common semi and fully automatic segmentation methods are ranked into three principal domains clustering, thresholding, and deformable models. More recently, the use of deep learning techniques has become popular [11]. However, despite the numerous segmentation methods that have been developed, there is still a demand for performance improvement [5].

We searched to identify the literature's available works based on the applied segmentation technique(s), accuracy, and publication year. Our search was performed on three public databases, Scopus, IEEE Xplore, and Web of Science. We applied three search strings: "skin lesion segmentation", "melanoma segmentation" and "dermoscopy images segmentation".

We established the following selection criteria: (1) articles must have been published after 2016 in computer science, engineering, or medicine fields; (2) articles must use dermoscopy images; (3) articles must use public image datasets in the evaluation of segmentation accuracy; (4) articles must use evaluation metrics suggested by the International Skin Imaging Collaboration (ISIC).

Regardless of the segmentation technique(s), most skin lesion segmentation approaches from images included both pre and post-processing steps.

Table I summarises the works found in the literature based on the year of publication, the segmentation technique(s), the number of validation images, the number of image datasets, and the segmentation accuracy achieved. It is important to note that in all of these works, the images used for validation were extracted from the same image dataset(s) used to adjust their parameters.

After analyzing the literature's works, we observe that each method's performance decreased when the images under analysis were more difficult to segment, mainly due to uneven illumination, artifacts, noise, and low contrast. To subdue these obstacles, we introduce a semi-automatic method that requires only a quick and easy intervention by the user. Although the segmentation results depend on this user interaction, our approach relieves the user of the challenging act of delineating the lesion border. Also, the time required for diagnosis is reduced. The proposed method offers the user high confidence in segmentation since they take an active and essential role in the segmentation process.

TABLE I
Sumário dos trabalhos encontrados na literatura.

Work	Year	Technique(s)	Images	Datasets	Accuracy(ies)
Li et al. [12]	2018	*2D DenseUNet-167 architecture*	600	1	ISIC 2017: 94.30%
Li and Shen [13]	2018	*Lesion Indexing Network*	150	1	ISIC 2017: 95.00%
Unver and Ayan [14]	2019	*GrabCut Algorithm*	800	2	PH^2: 92.99% ISIC 2017: 93.39%
Filali and Belkadi [15]	2019	*Multi-scale contrast segmentation*	201	1	DermIS: 97.86%
Nida et al. [16]	2019	*Deep regional CNN + Fuzzy C-means*	380	1	ISIC 2016: 94.20%
Navarro, Escudero-Viñolo and Bescós [17]	2019	SLIC algorithm + *Gaussian distribution*	600	1	ISIC 2017: 95.50%
Garcia-Arroyo and Garcia-Zapirain [18]	2019	*Cuts and their image masks obtained for different levels of probability from fuzzy sets and probability images + Histogram thresholding + Mask post-processed*	979	2	ISIC 2016: 93.40% ISIC 2017: 88.40%
Goyal et al. [3]	2020	*R-CNN with DeeplabV3C mask*	800	2	PH^2: 91.90% ISIC 2017: 94.08%
Xie et al. [19]	2020	*Mutual bootstrapping deep CNNs*	800	2	PH^2: 96.50% ISIC 2017: 94.70%
Lei et al. [20]	2020	*U-Net network + Dual discrimination*	1279	3	ISIC 2016: 96.00% ISIC 2017: 93.50% ISIC 2018: 92.90%
Arora et al. [21]	2021	*Modified U-Net network + Gaussian blurring + Median filtering*	550	1	ISIC 2018: 95.00%
Qamar, Ahmad and Shen [22]	2021	*Combination of DenseNet network and ResNet+ Atrous spatial pyramid pooling + dense skip connection*	518	1	ISIC 2018: 96.25%
Tong et al. [23]	2021	*Extended version of U-Net with triple attention mechanism*	1179	3	PH^2: 92.60% ISIC 2016: 95.40% ISIC 2017: 94.30%

III. Proposed Segmentation Method

We propose a semi-automatic segmentation approach using seeds, selected by the specialist, related to the lesion and non-lesion regions. The selected seeds are then used to outfit the semi-supervised clustering algorithm. Following this, post-processing is performed to exclude misclassified areas and reduce noise from the segmented lesion border. Figure 2 represents the proposed segmentation method.

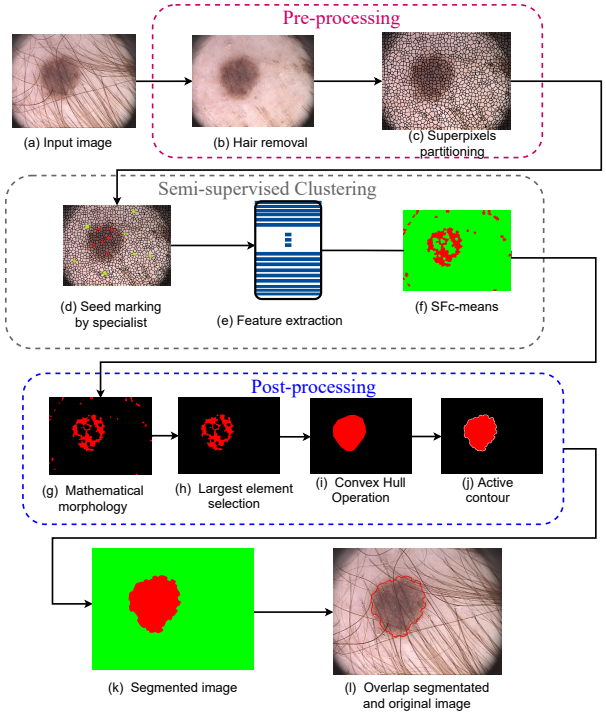


Fig. 2. Flowchart of the proposed segmentation approach.

A. Pre-processing

1) Hair removal: The first method step is hair removal. For this, we applied the DullRazor [24] algorithm. The algorithm mentioned uses a morphological closing operation on greyscale images to recognize regions with hair. Since hairs are long, thin structures, a bilinear interpolation is carried out on the identified hair regions' pixels. Finally, a median adaptive filter is applied to soften the modified hair pixels.

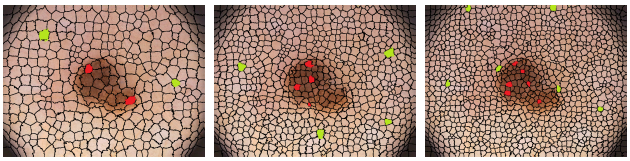
2) Superpixel partitioning: Traditionally, in digital image processing, the input image pixels are individually processed; however, this requires a high computational effort in images with large dimensions. We use superpixels to reduce computational cost. Superpixels are usually generated by pixel clustering algorithms. They collect local redundancies in the image and merge the corresponding pixels into superpixels. The proposed method uses Simple Linear Iterative Clustering algorithm with zero parameters (SLIC0) [25] for the generation of superpixels. This algorithm's efficiency has been proven relative to other algorithms in terms of its speed, memory consumption, and segmentation performance.

In search of optimal superpixels number, we evaluated different values for c, starting from 400 and adding increments of 200 until the segmentation accuracy had stabilized. In total, seven values of c were assessed: 400, 600, 800, 1000, 1200, 1400 and 1600.

B. Semi-supervised clustering

After superpixel representation, we cluster them using Seeded Fuzzy C-means (SFc-means). The SFc-means groups elements based on knowledge obtained from the labeled data, i.e., the seeds are selected by the user. The algorithm is based on approximate reasoning using fuzzy logic, which permits this method to treat the imprecision inherent to the scenario.

1) Selection of seeds: Our proposed semi-automatic segmentation method is based on a specialist's selections from the lesion and non-lesion regions in the dermoscopic image under analysis to feed a semi-supervised clustering algorithm. Figure 3 shows examples of superpixels that can be chosen by the specialist. The black lines delineate the region of each superpixel (sp) The superpixels in red represent lesion areas, while those in green represent the background of the images.



(a) Image with 400 sp. (b) Image with 800 sp. (c) Image with 1200 sp.

Fig. 3. Images with selected superpixel (sp) seeds (where red indicates the lesion region and green the image background).

There is also a ground truth image, i.e., a binary mask with the ideal segmentation for each image utilized in this work. To simulate the specialist's seeds selection, we chose 1%, randomly, of the superpixels for use as seeds for the SFc-means. Those seeds contained superpixels from both regions, i.e., skin lesion and image background.

2) Extraction of texture features: Features can be obtained from an input image, such as texture, color, and shape. Attributes can be grouped into feature vectors known as image descriptors. In this case, we extracted texture features from six color channels: RGB (red, green, blue) and L*a*b* (lightness, green-to-red variation, blue-to-yellow variation). The texture information extracted were from: Grey-level co-occurrence matrix - GLCM [26] (102 features), Grey-level run-length matrix - GLRLM [27]–[29] (264 features), Geostatistics [30], [31] (576 features), Local binary pattern - LBP [32] (1536 features) and Tamura [33] (18 features).

3) Feature selection: After building the features vector and evaluating the best results, we combined different texture descriptors to create a hybrid texture descriptor. Following this, we performed feature selection to simplify the prediction model and thus eliminate unnecessary features. Therefore, the required computational cost is decreased, and the clustering algorithm's performance can be improved.

For feature selection, we used the gain ratio information [34] algorithm. It is a filter that overvalues features with multiple values. Hence, this algorithm selects features that maximize the information gain while minimizing an attribute's number of values.

The gains ratio information approach is resilient, combines various methods of research and evaluation. We applied the Ranker technique. It ranks the attributes based on their evaluations. In this technique, a score is used for each attribute. So, a ranking is constructed according to attribute relevance degree. In this context, we performed 24 tests using 1%, 2%, 3%, 4%, 5%, and from 10% to 100% with a step of 5% of the feature vector length.

4) Seeded Fuzzy C-means: Based on the fuzzy C-means, the SFc-means [7] is a semi-supervised clustering algorithm. The authors introduced the use of seeds and a clustering threshold.

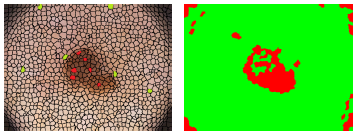
The SFc-means do not use the original concept of centroids. The algorithm uses all available labeled examples (i.e., all the seeds) as group representations. During clustering, SFc-means measures the membership degree of each input sample for each seed. The algorithm places the results in a descending sequence in a membership array.

Another SFc-means feature is the establishment of a clustering threshold value (t). This threshold is responsible for the association (or not) of an input sample with a cluster.

Figure 4 depicts the results of the SFc-means algorithm in grouping the superpixels presented in the input image. Red color represents superpixels in lesion area, green color represents background superpixels.

C. Post-processing

Although the SFc-means clustering has a high degree of certainty, it often produces errors, as we can see from Figure 5a. For this reason, we applied a post-processing step using mathematical morphology techniques (MM) and geodesic active contour (GAC).



(a) (b)

Fig. 4. Results of a semi-supervised segmentation obtained using the SFc-means (a - Input image with 1200 sp; b - Sp. of (a) after clustering).

To eliminate the small amounts of noise that persist after clustering, we applied an opening morphological operation succeeded by an erosion. The structuring element is a disc with a radius equal to 0.5% of the number of lines of the input image. We then separated the largest connected component, as shown in Figure 5c.

After selecting the region of the lesion, we removed sharp-edged fragments from the margins. To do this, we calculated the convex hull of the lesion region (Figure 5c and Figure 5d). We approximated the two contours by executing 100 iterations of the GAC algorithm [35] (Figure 5e). Hemalatha et al. [36] proposed the GAC algorithm to medical image analysis. The GAC modifies, proportional to the curvature, the convex hull in the Euclidean plane by moving its points perpendicularly to the lesion region.

(a) (b) (c) (d) (e)

Fig. 5. Example of the post-processing step of the proposed method (a - Result clustering; b - Mathematical morphology; c - Selection the largest element component; d - Convex Hull operation; e - Geodesic active contour.)

IV. Experiments

A. Image Datasets

One of the difficulties of dermoscopic image segmentation is segment lesions' capacity in datasets with distinct traits. For example, several experts collected images using different cameras, from patients of different races, varying resolutions, including ground truth images given by distinguished specialists. Numerous authors have, therefore, evaluated their systems using just individual datasets. In this research, we arranged the datasets into development and performance datasets.

The development dataset is created using images from the PH^2 [37], DermIS [38] and ISIC 2018 [39] datasets. The PH^2 base was brought together through a partnership between the University of Porto and the Dermatology Service of Hospital Pedro Hispano. Comprises 200 images with a resolution of 768×560 *pixels*. While the DermIS base is the result of the partnership between the Universities of Heidelberg and Erlangen. DermIS base presents more challenging skin lesions when compared to PH^2 base, as they have a greater similarity between benign and malignant lesions and make segmentation difficult.

The ISIC 2018 database provided by the *International Skin Imaging Collaboration Challenge* (ISIC) has RGB images with sizes ranging from 767×576 and 6682×4401. The set of images is composed of 2594 training images acquired from different patients from different institutions and different dermoscopy. We employed the development dataset for the algorithm parameter tuning.

To verify that the proposed method can be successfully used in broad heterogeneous images, we applied it to the images in the performance set, composed of images from the ISIC 2016 [40] and ISIC 2017 [41] datasets. We chose to use only the images from the testing phase of these challenges. The images in the ISIC 2016 and 2017 datasets are very challenging since they include quite heterogeneous images.

B. Evaluation Metrics

In this work, we use the term 'positive' to designate lesion areas, and 'negative' to designate non-lesion areas. Based on confusion matrix we computed the segmentation accuracy (Acc), specificity (Spe), sensitivity (Sen), and Jaccard index (Jac).

Codella et al. [39] demonstrated that the direct use of the Jaccard index as a measure of performance does not accurately reflect the number of images in which the computational segmentation fails or falls outside the expert interobserver variability. It is, therefore, more suitable to use the threshold Jaccard index metric (TJI).

In the ISIC challenge of 2018, the organizers chose 0.65 as a basis value for the threshold Jaccard index metric to indicate segmentation failure for a skin lesion image. To compute the TJI metric, a score for each image (i) is calculated based on the Jaccard index:

$$\begin{cases} score_i = 0, \text{if } Jac < 0.65, \\ score_i = Jac, \text{otherwise.} \end{cases} \quad (1)$$

Given a dataset with n images, the final TJI value is defined as the mean of all per-image scores:

$$TJI = \frac{\sum_{i=1}^{n} score_i}{n}. \quad (2)$$

C. Results

Our method requires four parameters: the number of superpixels generated by the SLIC0 algorithm (400, 600, 800, 1000, 1200, 1400 or 1600); the color channels for feature extraction (RGB, L* a* b* or RGB + L* a* b*); the texture descriptor (GLCM, GLRLM, Geostatistics, LBP, Tamura or combinations of these descriptors); and the percentage of the most significant features.

To set the value of each parameter, we executed the method ten times. Furthermore, we multiplied the TJI by 100 to homogenize the visualization and to allow the tables in this section to be understood more easily.

When the TJI was used as the primary performance metric, a hybrid of the RGB and L* a* b* color channels gave the six best combinations, as shown in Table II. Three values for the number of superpixels (1000, 1200, and 1400) and two descriptors (GLCM and Tamura) were optimal. To set up our hybrid texture feature vector, we decided to combine the

settings that gave the best results. Therefore, we concatenated the GLCM and Tamura descriptors to create three additional feature vectors, using 1000 superpixels in the RGB and L* a* b* color channels. We then applied a similar setup using 1200 and 1400 superpixels.

TABLE II
BEST SEGMENTATION PERFORMANCE FOR THE DEVELOPMENT DATASET USING THE PROPOSED METHOD.

GLCM						
N° of Sp.	Color Channel	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
1400	RGB+L*a*b*	96.88	96.68	91.57	87.41	**96.85**
1200	RGB+L*a*b*	96.73	96.61	91.61	87.45	**96.27**
1000	RGB+L*a*b*	96.41	96.10	90.87	86.34	**95.71**
Tamura						
N° of Sp.	Color Channel	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
1400	RGB+L*a*b*	97.17	97.19	92.04	88.61	**96.85**
1000	RGB+L*a*b*	96.78	96.35	91.59	87.71	**96.13**
1200	RGB+L*a*b*	97.04	97.07	91.64	88.18	**95.98**

We executed empirical tests to define the definitive feature vector's dimensionality, always searching for bargaining between TJI and the vector length. In these tests, we applied the gain ratio algorithm and assembled the built vectors with the features ranked in descending order of gain ratio. When performing the experiments, we used these ranked features, starting from the highest gain ratio and adding the other ones incrementally to include all features. Thus, we used feature vectors with the percentage of the most significant features in a sum of 216 tests, i.e. 24 feature vectors with different lengths multiplied by nine descriptors. Figure 6 depicts the results obtained from these experiments. Using TJI as the most relevant metric, we obtained the best results for 1400 superpixels using 80% of the GLCM and Tamura features.

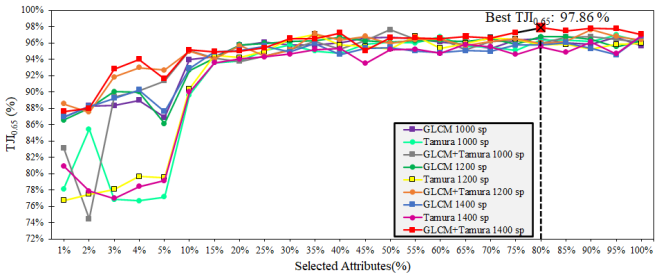


Fig. 6. Values of TJI achieved using different numbers of features, sorted by gain ratio.

Figure 7 represents the features chosen from descriptors and the vector positions sorted by the gain ratio algorithm in the GLCM+Tamura feature vector. It can be seen from this figure that the GLCM descriptor the part larger on the crucial features subset.

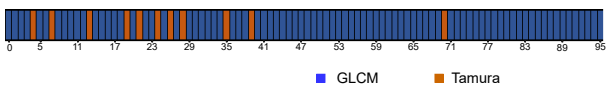


Fig. 7. Origin of the 96 selected features (the bar at the bottom indicates the source of the characteristics at each position).

To demonstrate that the proposed method works well on images from different datasets, we applied it to the performance image dataset (ISIC 2016 + ISIC 2017 datasets) using the best parameters found in the previous stage. We obtained the results shown in Table III. The obtained results and the average processing time, in seconds. The tests were carried out on a computer with an Intel Core i7-3630QM CPU and 8GB of RAM memory using MATLAB [42].

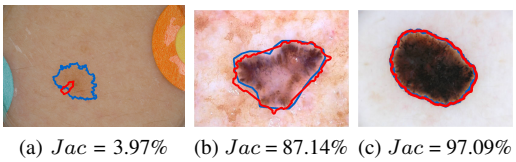
TABLE III
AVERAGE AND STANDARD DEVIATION OF THE SEGMENTATION QUALITY METRICS OBTAINED BY THE PROPOSED METHOD ON THE PERFORMANCE DATASET AND THE RESPECTIVE AVERAGE PROCESSING IN SECONDS.

Dataset	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$	Time(s)
ISIC 2016	96.24±0.08	95.62±0.22	86.12±0.09	82.43±0.16	94.28±0.01	4.12±0.35
ISIC 2017	95.80±0.05	94.62±0.06	78.71±0.23	74.13±0.29	78.97±0.02	4.75±0.40

The values in Table III indicate that the proposed method achieved an accuracy of bigger than 95.80% for both datasets. Moreover, the Jaccard index values reveal that, on average, the similarity between the segmented regions and their ground truths was higher than 82% for the ISIC 2016 dataset and higher than 74% for the ISIC 2017 dataset. These results reflect the value obtained for TJI. They indicate that in 94.28% of the ISIC 2016 dataset images, the agreement was higher than 65% between the segmented regions and their ground truth. The ISIC 2017 dataset was a particularly challenging case, and a TJI of 78.97% was achieved. The method took, on average, between 4 and 5 seconds to execute an image with a tiny standard deviation. Thus, demonstrating that the execution time is no susceptible to the dimensions in the image.

Skin lesion segmentation algorithms can be used to assist the specialist in extracting from the input image information related to the ABCD rule, for example, or in producing inputs for lesion classification systems. It is, therefore, essential to determine the effectiveness of segmentation for benign and malignant lesions.

Figure 8 depicts the results of the proposed segmentation method (in red) overlapped by the correspondent ground truths (in blue). For each individual dataset in the performance dataset, the figure shows the worst (Figures 8a and 8d), the median (Figures 8b and 8e), and the best (Figures 8c and 8f) values of the Jaccard index.



(a) Jac = 3.97% (b) Jac = 87.14% (c) Jac = 97.09%

(d) Jac = 0% (e) Jac = 82.25% (f) Jac = 97.84%

Fig. 8. Examples of skin lesions from the ISIC 2016 (a, b, c) and ISIC 2017 (d, e, f) datasets (for each example, the ground truth is shown in blue and the segmentation obtained by the proposed method in red).

V. Discussion

The results obtained in this work are equivalent to those of state-of-the-art alternatives. In some situations, researchers have used private-labeled datasets, and it is difficult to reproduce the results reported in their work. Nevertheless, since the PH^2, DermIS, ISIC 2018, ISIC 2016, and ISIC 2017 datasets are publicly available, a comparison against other state-of-the-art methods that used these datasets was viable. Table IV shows the results of this comparison for the development dataset. The values shown in this table for existing methods were acquired from the original articles.

For the PH^2 dataset, the accuracy and Jaccard values obtained with the proposed method were lower than those reported by [19], with differences of 0.4% and 1.76%, respectively. For the DermIS dataset, the proposed method's accuracy was 1.62% lower than that reported by [15]. In the examples where our method gave a worse average performance, the differences were not statistically significant. The ISIC 2018 dataset the Acc, Spe, Jac and TJI values obtained by the proposed method was higher than for the other methods. One aspect that should be noted is that although the proposed method attained a high accuracy for the DermIS dataset, the value of 73.63% for the TJI indicates that the method demands to be refined.

Table IV exhibits the results of the proposed method against state-of-the-art methods for the images in the performance dataset. Table V presents the results of this comparison.

TABLE IV
COMPARISON OF THE RESULTS OF THE PROPOSED METHOD AGAINST THOSE OF STATE-OF-THE-ART ALTERNATIVES FOR THE DEVELOPMENT DATASET.

PH^2					
Method	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
[14]	92.99	94.02	83.63	79.54	–
[3]	93.80	92.90	93.20	83.90	–
[19]	**96.50**	**94.60**	**96.70**	**89.40**	–
[23]	94.30	93.70	96.00	84.20	–
Proposed	96.10±0.06	93.72±0.17	91.93±0.38	87.64±0.26	**98.83**±0.00
DermIS					
Method	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
[15]	**97.86**	99.00	**91.86**	–	–
Proposed	96.24±0.08	**99.58**±0.05	72.33±1.49	**70.28**±1.28	**73.63**±0.03
ISIC 2018					
Method	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
[21]	95.00	95.00	94.00	83.00	–
[20]	92.90	91.10	95.30	–	82.40
[22]	96.25	97.00	**96.50**	83.30	–
Proposed	**97.08** ±0.13	**97.42**±0.15	93.48±0.89	**89.89**±0.58	**99.67**±0.01

For the ISIC 2016 dataset, the proposed method achieved higher accuracy than the alternative methods. The Jac value obtained by the proposed method was higher than for the other methods except for that of [16]. For the ISIC 2017 dataset, the accuracy obtained by the proposed method surpassed all of the other methods under comparison, although the Jac value was only higher than that obtained by the method of [18].

An important aspect that should be noted about the state-of-the-art methods in Table V is that they used images from the same dataset for both parameter tuning (for example, the training of classifiers) and performance evaluation. This means that these methods will be more efficient for these pre-known images, unlike our approach, which not used the ISIC datasets in the development; this is further confirmation of the superior robustness portability of our method.

TABLE V
COMPARISON OF THE RESULTS OF PROPOSED METHOD AGAINST THOSE OF THE STATE-OF-THE-ART ALTERNATIVES FOR THE PERFORMANCE DATASET.

ISIC 2016					
Method	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
[16]	94.20	94.00	**95.00**	**93.00**	–
[18]	93.40	**97.80**	87.00	79.10	–
[20]	96.00	96.80	93.70	–	87.10
[23]	95.40	96.10	92.70	84.50	–
Proposed	**96.24**±0.08	95.62±0.22	86.12±0.09	82.43±0.16	**94.28**±0.01
ISIC 2017					
Method	$Acc(\%)$	$Spe(\%)$	$Sen(\%)$	$Jac(\%)$	$TJI(\%)$
[12]	94.30	95.30	87.90	79.80	–
[13]	95.00	97.40	85.50	75.30	–
[14]	93.39	92.68	**90.82**	74.81	–
[17]	95.50	–	–	76.90	–
[18]	88.40	92.30	86.90	66.50	–
[3]	94.08	95.00	89.93	79.34	–
[19]	94.70	96.80	87.40	**80.40**	–
[20]	93.50	**97.60**	83.50	–	77.10
[23]	92.60	96.50	82.50	74.20	–
Proposed	**95.80**±0.05	94.62±0.06	78.71±0.23	74.13±0.29	**78.97**±0.02

VI. Conclusion and Future Work

In this article, we presented a semi-automatic method for the segmentation of skin lesions in dermoscopic images. Our approach is based on superpixels, the union of texture information, and semi-supervised clustering.

The proposed method requires the user to identify areas in each region of the image (i.e., the lesion and the background). However, this corresponds only to 1% of the existing superpixels in the image. In practical terms, this means that a minimal amount of data is required from the specialist. The physician can obtain segmentation results similar to manual segmentation from a significantly less time-consuming task in our method.

To simulate an environment closer to physicians' real-world problems, we used 3974 images. The results were promising but could be improved, and we, therefore, think to evaluate various other post-processing image techniques and other types of texture information.

Acknowledgment

This study was financed in part by the "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES), in Brazil, Finance Code 001.

Publications

- A Skin Lesion Semi-Supervised Segmentation Method. Online version: https://ieeexplore.ieee.org/document/9145240.
- Semi-automatic Segmentation of Skin Lesions based on Superpixels and Hybrid Texture Information (In review on Medical Image Analysis).

References

[1] T. S. C. Foundation, "Skin cancer facts & statistics," https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/, 2021.

[2] J. Scharcanski and M. E. Celebi, *Computer vision techniques for the diagnosis of skin cancer*. Springer, 2013.

[3] M. Goyal, A. Oakley, P. Bansal, D. Dancey, and M. H. Yap, "Skin lesion segmentation in dermoscopic images with ensemble deep learning methods," *IEEE Access*, vol. 8, pp. 4171–4181, 2020.

[4] N. Moura, R. Veras, K. Aires, V. Machado, R. Silva, F. Araújo, and M. Claro, "Abcd rule and pre-trained cnns for melanoma diagnosis," *Multimedia Tools Appl.*, vol. 78, no. 6, p. 6869–6888, Mar. 2019. [Online]. Available: https://doi.org/10.1007/s11042-018-6404-8

[5] S. Chan, V. Reddy, B. Myers, Q. Thibodeaux, N. Brownstone, and W. Liao, "Machine learning in dermatology: Current applications, opportunities, and limitations," *Dermatology and Therapy*, p. 6869–6888, April 2020.

[6] V. B.N., P. J. Shah, V. Shekar, H. R. Vanamala, and V. Krishna A., "Detection of melanoma using deep learning techniques," in *2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, 2020, pp. 391–394.

[7] L. Santos, R. Veras, K. Aires, L. Britto, and V. Machado, "Medical image segmentation using seeded fuzzy c-means: A semi-supervised clustering algorithm." in *International Joint Conference on Neural Networks*. IEEE, 2018, pp. 1–8.

[8] M. Celebi, Q. Wen, H. Iyatomi, K. Shimizu, H. Zhou, and G. Schaefer, "A state-of-the-art survey on lesion border detection in dermoscopy images," in *Dermoscopy Image Analysis*. CRC Press, sep 2015, pp. 97–129. [Online]. Available: https://doi.org/10.1201%2Fb19107-5

[9] R. B. Oliveira, M. E. Filho, Z. Ma, J. P. Papa, A. S. Pereira, and J. M. R. Tavares, "Computational methods for the image segmentation of pigmented skin lesions: A review," *Computer Methods and Programs in Biomedicine*, vol. 131, pp. 127–141, 2016.

[10] Z. Ma and J. M. R. S. Tavares, "A novel approach to segment skin lesions in dermoscopic images based on a deformable model," *Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 615–623, 2016.

[11] N. N. Sultana and N. B. Puhan, "Recent deep learning methods for melanoma detection: A review," in *Mathematics and Computing*, D. Ghosh, D. Giri, R. N. Mohapatra, E. Savas, K. Sakurai, and L. P. Singh, Eds. Singapore: Springer Singapore, 2018, pp. 118–132.

[12] X. Li, L. Yu, H. Chen, C. Fu, and P. Heng, "Semi-supervised skin lesion segmentation via transformation consistent self-ensembling model," *arXiv:1808.03887*, 2018.

[13] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, p. 556, Feb 2018. [Online]. Available: http://dx.doi.org/10.3390/s18020556

[14] H. M. Ünver and E. Ayan, "Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm," *Diagnostics*, vol. 9, no. 3, p. 72, Jul 2019. [Online]. Available: http://dx.doi.org/10.3390/diagnostics9030072

[15] I. Filali and M. Belkadi, "Multi-scale contrast based skin lesion segmentation in digital images," *Optik*, vol. 185, pp. 794 – 811, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030402619304917

[16] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering," *International Journal of Medical Informatics*, vol. 124, 04 2019.

[17] M. E.-V. Fulgencio Navarro and J. Bescós, "Accurate segmentation and registration of skin lesion images to evaluate lesion change," *IEEE Journal of Biomedical and Health Inf*, vol. 23, no. 2, pp. 501 – 508, 2019.

[18] J. L. Garcia-Arroyo and B. Garcia-Zapirain, "Segmentation of skin lesions in dermoscopy images using fuzzy classification of pixels and histogram thresholding," *Computer Methods and Programs in Biomedicine*, vol. 168, pp. 11 – 19, 2019.

[19] Y. Xie, J. Zhang, Y. Xia, and C. Shen, "A mutual bootstrapping model for automated skin lesion segmentation and classification," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2020.

[20] B. Lei, Z. Xia, F. Jiang, X. Jiang, Z. Ge, Y. Xu, J. Qin, S. Chen, T. Wang, and S. Wang, "Skin lesion segmentation via generative adversarial networks with dual discriminators," *Medical Image Analysis*, vol. 64, p. 101716, 2020.

[21] R. Arora, B. Raman, K. Nayyar, and R. Awasthi, "Automated skin lesion segmentation using attention-based deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 65, p. 102358, 2021.

[22] S. Qamar, P. Ahmad, and L. Shen, "Dense encoder-decoder–based architecture for skin lesion segmentation," *Cognitive Computation*, vol. 13, no. 2, pp. 583–594, 2021.

[23] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, "Ascu-net: Attention gate, spatial and channel attention u-net for skin lesion segmentation," *Diagnostics*, vol. 11, no. 3, p. 501, 2021.

[24] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, "Dullrazor®: A software approach to hair removal from images," *Computers in biology and medicine*, vol. 27, no. 6, pp. 533–543, 1997.

[25] B. F. Nan and Z. C. Mu, "Slico-based superpixel segmentation method with texture fusion," *Chinese Journal of Scientific Instrument*, vol. 35, no. 3, pp. 527–534, 2014.

[26] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[27] M. M. Galloway, "Texture analysis using gray level run lengths," *Computer graphics and image processing*, vol. 4, no. 2, pp. 172–179, 1975.

[28] A. Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution of run lengths for texture analysis," *Pattern Recognition Letters*, vol. 11, no. 6, pp. 415–419, 1990.

[29] B. R. Dasarathyand and E. B. Holder, "Image characterizations based on joint gray-level run-length distributions," *Pattern Recognition Letters*, vol. 12, no. 8, pp. 497–502, 1991.

[30] A. C. Silva, P. C. P. Carvalho, and M. Gattass, "Analysis of spatial variability using geostatistical functions for diagnosis of lung nodule in computerized tomography images," *Pattern Analysis and Applications*, vol. 7, no. 3, pp. 227–234, 2004.

[31] J. Sousa, A. Paiva, J. Almeida, A. Silva, G. Junior, and M. Gattass, "Texture based on geostatistic for glaucoma diagnosis from fundus eye image," *Multimedia Tools and Applications*, vol. 76, 09 2017.

[32] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[33] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 6, pp. 460–473, 1978.

[34] L. E. Raileanu and K. Stoffel, "Theoretical comparison between the gini index and information gain criteria," *Annals of Mathematics and Artificial Intelligence*, vol. 41, pp. 77 – 93, 2004.

[35] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[36] R. Hemalatha, T. Thamizhvani, A. J. A. Dhivya, J. E. Joseph, B. Babu, and R. Chandrasekaran, "Active contour based segmentation techniques for medical image analysis," in *Medical and Biological Image Analysis*, R. Koprowski, Ed. Rijeka: IntechOpen, 2018, ch. 2, pp. 17–34.

[37] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, "Ph2-a dermoscopic image database for research and benchmarking," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 5437–5440.

[38] T. L. Diepgen and G. Yihune, "Dermatology information system – dermis," http://dermis.net/, 2012, retrieved September 12, 2016.

[39] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "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, pp. 168–172.

[40] D. Gutman, N. C. F. Codella, M. E. Celebi, B. Helba, M. A. Marchetti, N. K. Mishra, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)," *arXiv*, 2016. [Online]. Available: arXiv:1605.01397

[41] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. K. Mishra, H. Kittler, and A. Halpern, "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)," *IEEE*, 2017.

[42] MATLAB, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.