ISIC 2018 - A Framework for Automatic Lesion Diagnosis based on Thresholding Segmentation and Hierarchical Classification

José G. M. Esgario and Renato A. Krohling

***Abstract*—International Skin Image Collaboration (ISIC) is an international initiative that encourages the development of image analysis tools for automatic skin cancer diagnosis. This work proposes a framework to address two out of three tasks proposed by ISIC Challenge 2018, i.e. segmentation (task 1) and classification (task 3). The segmentation is performed by an unsupervised method, a set of 82 handcrafted features were computed and an hierarchical classifier based on Extreme Learning Machine is proposed for the classification of skin lesions.**

I. Introduction

Skin cancer is considered one of the most common types of cancer worldwide. According to the World Health Organization [1], currently, between 2 and 3 million non-melanoma skin cancers and 132.000 melanoma skin cancer occur every year in the world. One in three diagnosed cancers is skin cancer. Early diagnosis of the disease plays an important role in reducing the mortality rate.

The diagnosis of skin cancer can be made by invasive and non-invasive methods. One of the most common noninvasive methods was presented by Soyer et al [2]. Which allows the visualization of morphological structures not visible to the naked eye with the use of an instrument called dermatoscope. Despite the help of the dermatoscope, the diagnosis is made by qualitative and semi-quantitative methods that present a lot of subjectivity and depend on the dermatologist's experience to obtain good results. Therefore, there is great research interest worldwide to develop automatic diagnostic systems that can help dermatologists correctly diagnose skin lesions.

This work proposes a framework for automatic diagnosis of skin lesions. The framework consists of a threshold-based segmentation algorithm with pre and post-processing steps. A set of 82 handcrafted features based on ABCD rule and finally an hierarchical classification method that classifies the lesion as one out of seven possible classes: Melanoma, Melanocityc Nevus, Basal cell carcinoma, Actinic keratosis, Benign keratosis, Dermatofibroma and Vascular lesion.

II. Proposed Approach

The proposed framework, presented as a block diagram in Fig. 1, consists of a set of four main steps: segmentation, feature extraction, data pre-processing and classification. The steps are detailed in the following sections.

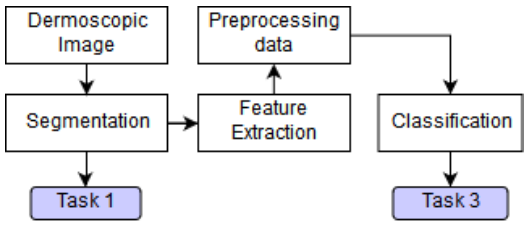

Fig. 1. Block diagram of the proposed framework.

A. Segmentation

The proposed method is based on the work of Agarwal [3] who proposed the combination of the channels of the Lab color space generating a final image in gray scale. This work improves the method proposed by Agarwal through the addition of steps of pre and post-processing of image and the individual use of the B channel that demonstrated better results. The fig. 2 presents a block diagram that details the proposed segmentation methodology.

The input images were standardized with size 300x225 (4:3 ratio) in order to reduce the segmentation computational time without significant loss in results quality. The first step of pre-processing consists of hair removal from dermoscopic images, most methods of hair removal are based on three main step, i.e. highlights hair, segmentation and inpainting [4]. The hair was highlighted with the Gaussian Laplacian filter, the hair mask was generated by fixed thresholding, finally, each masked pixel of original image was replaced by a RGB image modified with the morphological closing operator.

Researches point out that the B channel of the RGB color space produces better segmentation results and suggest that the lesions are more predominantly visible in this channel [5], [6]. Therefore, the pixels of the R and G channels were set to zero. The conversion to the Lab color space allows the separation of the illumination channel, in this way the illumination correction was performed in channel L and later the channels are combined forming a final image in gray scale. For this preprocessing step, the approach proposed by Glaister, Wong and Clausi [7] was taken as a basis. In the original approach the segmentation of the skin and lesion pixels were done with the Statistic Region Merging and the model parameters were

J. G. M. Esgario is with Graduate Program in Computer Science, Federal University of Espírito Santo, Brazil (e-mail: guilherme.esgario@gmail.com).

R. A. Krohling is with the Department of Production Engineering and Graduate Program in Computer Science, Federal University of Espírito Santo, Brazil (e-mail: krohling.renato@gmail.com).

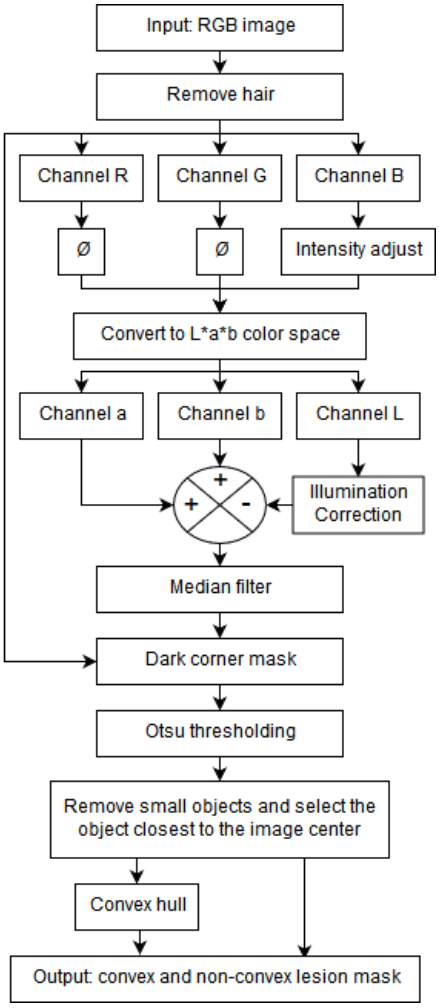

Fig. 2. Block diagram of the segmentation method.

estimated using the maximum likelihood method. In our work, the Otsu method and the least squares method were used.

A median filter was applied, this filter was chosen for its noise-removing ability without blurring the image. The filtered image passes through a dark corner removal mask (see appendix A) and finally the segmentation is performed based on thresholding by the Otsu method.

Post-processing consists of removing small elements, i.e. elements whose area is less than a certain threshold. The final lesion is selected as the resulting object that most closely approximates the image center. Two masks are generated as the final segmentation result and in one of them the convex hull is computed. The convex mask returned better results due to its greater similarity with the ground truth provided by the expert. The other mask was used in the process of feature extraction due to its greater fidelity to the lesion contour, highlighting the border irregularities.

B. Feature Extraction

After segmentation the feature extraction process is performed in order to quantify the features of each lesion. This work was based on the so-called ABCD-rule [8]. A set of 82 features were selected in the literature and are described in table I.

TABLE I
SET OF HANDCRAFTED FEATURES

Feature	*Description*	*Amount*
Shape Asymmetry [9]	the mask is rotated 180 degrees and the overlapping area ratio is calculated.	1
Color Asymmetry [10]	It divides the lesion into two parts on the horizontal and vertical axes and calculates the differences between the histograms.	3
Compactness [9], [11], [12]	Measures how compact the lesion is, taking into account its perimeter and area.	1
Fractal [9], [11], [12]	Divide image into grids and compute the contour length in the grid.	1
Radial Variance [9], [11], [12]	Estimates the border irregularity by the variance of the radial distance distribution.	1
Border Irregularity [13]	Detects irregularities by the derivative of the borderline function.	1
Pigmentation Transition [12]	Calculates the mean and variance of the gradient in the perimeter of the lesion.	2
Solidity [14], [15]	The ratio between the lesion area and its convex hull area.	1
Histogram Measures [16]	Calculates the mean, variance, skewness, uniformity and entropy of the histogram from each RGB and HSV channel.	30
Color Variegation [14]	Computes the log of the variance over mean of the RGB and HSV color channels for both lesion and skin.	12
Suspicious Colors [17], [18]	Count how many suspicious colors appear in the image. Max 6 colors.	1
Color Correlation [19]	Six correlation coefficients (red-green, green-blue, red-blue, red-gray, green-gray and blue-gray) were computed for both lesion and skin.	12
GLCM [9], [14], [20]	Compute Contrast, Correlation, Energy and Homogeneity for each orientations (0°, 45°, 90° and 135°).	16

C. Preprocessing Data

The quality of the classification results depends heavily on the quality of the training data. Regardless of the classifier used, if the training data are incorrect, poor models will result [21]. The automatic diagnosis of skin cancer directly faces the problem of imbalanced data due to the difficulty of collecting data and the rarity of certain diseases. In order to overcome these difficulties, four common approaches of literature were employed: Feature Scaling, Feature Selection, Remove Outliers and Undersampling the Majority Class.

1) Feature Scaling: The difference between the features range values may lead to poor classification results, therefore,

the min-max scaling method was used to normalize the data in the interval $[0, 1]$, it is calculated as

$$x'_{ij} = \frac{x_{ij} - \min x_j}{\max x_j - \min x_j}, i = 1, \ldots, m; j = 1, \ldots, n \quad (1)$$

2) Feature Selection: Features selection approaches are divided into: filter, wrapper and embedded. It consists of removing redundant and irrelevant variables, which helps to reduce computational time, reduce the curse of dimensionality and improve classifier performance [22]. Many papers have been proposed in the literature based on the use of evolutionary search techniques for this purpose [23]. In this work it was decided to use a wrapper approach with the Binary Particle Swarm Optimization (BPSO) search method [24].

The fitness function used is the result of the classification performed by the Nearest Centroid Classifier (NCC) evaluated with the F1-score metric. The NCC is a classification model that assigns the new samples (observations) the label of the training class whose centroid is closest to the observation.

3) Remove Outliers: The segmentation algorithm is not always able to correctly detect the lesions, so it can result in spurious data. To make the training data more concise and with less noise, a statistical approach was applied in order to remove outliers. Given a data matrix $M_{m\times n}$ described by

$$M = \begin{bmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \ldots & x_{mn} \end{bmatrix} \quad (2)$$

where m is the amount of data and n is the number of features.

Calculating the Euclidean distance between each sample of matrix $M_{m\times n}$ results in a matrix $D_{m\times m}$ denoted by

$$D = \begin{bmatrix} d_{11} & \ldots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{m1} & \ldots & d_{mm} \end{bmatrix} \quad (3)$$

The mean distance of each sample to all others can be obtained by averaging the values of each line of the matrix

The mean distance from each sample to all others can be obtained by calculating the mean values of each row of matrix $D_{m\times m}$ resulting in a vector $\tilde{D}$ described by

$$\tilde{D} = \begin{bmatrix} \tilde{d}_1 \\ \vdots \\ \tilde{d}_m \end{bmatrix} \quad (4)$$

Considering $\tilde{D} \sim N(\mu, \sigma^2)$, therefore, the *i*-th sample M_i can be considered an outlier respecting the following constraint

$$\begin{cases} true, & if\ |\tilde{D} - \mu| > k\sigma \\ false, & \text{otherwise} \end{cases} \quad (5)$$

where k is a constant that defines how many standard deviations the sample can deviate from the mean before being considered an outlier.

4) Undersampling Majority Class: One of the best known methods for treating imbalanced data is called sampling. When a class is underrepresented a methodology is to apply the oversampling on it, generating new samples from the existing ones. If a class is overrepresented the undersampling can be accomplished by removing samples from that class. In this work we applied a simple method called Random Undersampling that selects and removes from the dataset random samples from a certain class in order to balance data.

D. Classification

In order to improve the classification results we used in this work a methodology based on a binary tree structure that transforms the multiclass problem into sub binary classification problems. Vural and Dy [25] proposed an approach called Divide-by-2 Method which hierarchically divides the dataset into two subsets until each subset consists of only a single class. Based on this hierarchical approach, several papers have proposed the use of clustering algorithms to group classes into two [25]–[27]. Using Divide-by-2 Method we construct a decision tree (Fig. 3) using the K-Means clustering method.

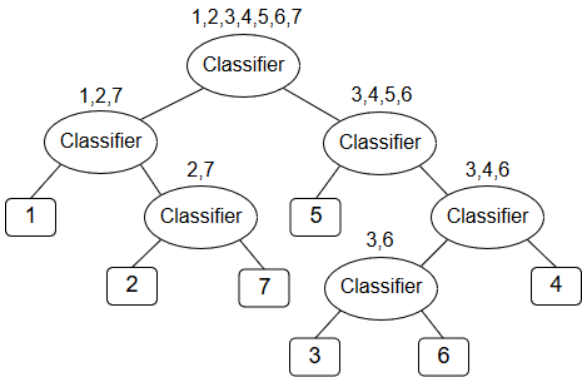

Fig. 3. Binary tree structure for classify skin cancer.

In this work we used the Extreme Learning Machine (ELM) proposed by Huang [28] as the classifier in each node of the binary tree. The total of required classifiers is equal to $K - 1$ classes.

III. Experimental Results

The experiments were conducted using training and validation data extracted from the "ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection" grand challenge datasets [29], [30]. The size of the datasets are shown in the table II.

TABLE II
Dataset size

	Training	*Validation*
Task 1	2594	100
Task 3	10015	193

To perform the preprocessing and classification some parameters were defined empirically. The PSO for feature selection was set up with 20 particles and executed 200 iterations, the removal of outliers was performed with $k = 1$ and the

Random Undersampling was applied only in class 2 because it had a much greater amount of data than the other classes (66.9%). The number of neurons for the ELM was set to 200. Table III shows the details of the Task 3 training dataset before and after preprocessing, the result shows a significant data reduction of class 2 and the number of features.

TABLE III
TASK 3 TRAINING DATASET BEFORE AND AFTER PREPROCESSING

Class	*Before preprocessing*		*After preprocessing*	
	Amount	%	Amount	%
Melanoma	1113	11.1	887	25.0
Melanocytic nevus	6705	66.9	855	24.1
Basal cell carcinoma	514	5.1	451	12.7
Actinic keratosis	327	3.3	262	7.4
Benign keratosis	1099	11.0	870	24.5
Dermatofibroma	115	1.2	103	2.9
Vascular lesion	142	1.4	122	3.4
All	10015	100.0	3550	100.0
Features	82	—	39	—

Task 1 consists of finding the binary mask that best segments the skin lesion. The obtained results were scored using the Jaccard index metric. Images with a score below 0.65 received score zero, the final result is the average of all scores obtained. Task 3 is defined as a multiclass classification problem, where the proposed algorithm should classify each lesion into one out of seven classes present in the training dataset. The results were evaluated using the normalized multiclass accuracy metric. The experimental results obtained in both Tasks are described in the table IV.

TABLE IV
EXPERIMENTAL RESULTS FOR VALIDATION DATASETS

Task	*Metric*	*Score*
1	Jaccard index	0.723
3	Normalized Multiclass Accuracy	0.714

IV. DISCUSSION AND CONCLUSION

Few papers presented in the previous challenges used unsupervised approaches, so in this work a simple and effective approach was proposed for segmentation that compared to previous results suggests that the approach is promising. Unsupervised approaches are a good alternative for problems where data is scarce. The classification problem presented itself as a more complex problem than the previous ones, the classes proved to be quite overlapping and difficult to classify, so comparing current results with past binary problems does not seem to make sense. Future research will investigate new features and compare the results with different classifiers in order to find a more appropriate model that returns better results.

REFERENCES

[1] "World health organization: How common is skin cancer?." http://www.who.int/uv/faq/skincancer/en/index1.html. Accessed: 2018-07-26.

[2] H.-P. Soyer, J. Smolle, H. Kerl, and H. Stettner, "Early diagnosis of malignant melanoma by surface microscopy," *The Lancet*, vol. 330, no. 8562, p. 803, 1987.

[3] A. Agarwal, A. Issac, M. K. Dutta, K. Riha, and V. Uher, "Automated skin lesion segmentation using k-means clustering from digital dermoscopic images," in *40th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 743–748, July 2017.

[4] Q. Abbas, M. Celebi, and I. F. García, "Hair removal methods: A comparative study for dermoscopy images," *Biomedical Signal Processing and Control*, vol. 6, no. 4, pp. 395–404, 2011.

[5] S. Khalid, U. Jamil, K. Saleem, M. U. Akram, W. Manzoor, W. Ahmed, and A. Sohail, "Segmentation of skin lesion using Cohen–Daubechies–Feauveau biorthogonal wavelet," *SpringerPlus*, vol. 5, no. 1, p. 1603, 2016.

[6] S. Pathan, K. G. Prabhu, and P. Siddalingaswamy, "Techniques and algorithms for computer aided diagnosis of pigmented skin lesions - A review," *Biomedical Signal Processing and Control*, vol. 39, pp. 237–262, 2018.

[7] J. Glaister, A. Wong, and D. A. Clausi, "Illumination correction in dermatological photographs using multi-stage illumination modeling for skin lesion analysis," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 102–105, Aug 2012.

[8] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig, "The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions," *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551–559, 1994.

[9] M. Messadi, M. Ammar, H. Cherifi, M. Chikh, and A. Bessaid, "Interpretable aide diagnosis system for melanoma recognition," *Journal of Bioengineering & Biomedical Sciences*, vol. 4, no. 1, p. 1, 2014.

[10] N. Smaoui and S. Bessassi, "A developed system for melanoma diagnosis," *International Journal of Computer Vision and Signal Processing*, vol. 3, no. 1, pp. 10–17, 2013.

[11] M. A. H. Bhuiyan, I. Azad, and K. Uddin, "Image processing for skin cancer features extraction," *International Journal of Scientific & Engineering Research*, vol. 4, no. 2, pp. 1–6, 2013.

[12] T. Yamunarani, "Analysis of skin cancer using ABCD technique," *International Research Journal of Engineering and Technology*, 2018.

[13] J. Jaworek-Korjakowska, "Novel method for border irregularity assessment in dermoscopic color images," *Computational and mathematical methods in medicine*, 2015.

[14] J. Jaworek-Korjakowska, "Computer-aided diagnosis of micro-malignant melanoma lesions applying support vector machines," *BioMed research international*, 2016.

[15] N. C. Lynn and Z. M. Kyu, "Segmentation and classification of skin cancer melanoma from skin lesion images," in *International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, pp. 117–122, IEEE, 2017.

[16] W. V. Stoecker, K. Gupta, B. Shrestha, M. Wronkiewiecz, R. Chowdhury, R. J. Stanley, J. Xu, R. H. Moss, M. E. Celebi, H. S. Rabinovitz, *et al.*, "Detection of basal cell carcinoma using color and histogram measures of semitranslucent areas," *Skin Research and Technology*, vol. 15, no. 3, pp. 283–287, 2009.

[17] R. Kasmi and K. Mokrani, "Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule," *IET Image Processing*, vol. 10, no. 6, pp. 448–455, 2016.

[18] N. K. E. Abbadi and Z. Faisal, "Detection and analysis of skin cancer from skin lesions," 2017.

[19] W.-Y. Chang, A. Huang, C.-Y. Yang, C.-H. Lee, Y.-C. Chen, T.-Y. Wu, and G.-S. Chen, "Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study," *PloS one*, vol. 8, no. 11, p. e76212, 2013.

[20] S. Chatterjee, D. Dey, and S. Munshi, "Mathematical morphology aided shape, texture and colour feature extraction from skin lesion for identification of malignant melanoma," in *2015 International Conference on Condition Assessment Techniques in Electrical Systems (CATCON)*, pp. 200–203, IEEE, 2015.

[21] F. Kamiran and T. Calders, "Data preprocessing techniques for classification without discrimination," *Knowledge and Information Systems*, vol. 33, pp. 1–33, Oct 2012.

[22] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014. 40th-year commemorative issue.

[23] Neha and J. Vashishtha, "Particle swarm optimization based feature selection," *International Journal of Computer Applications*, vol. 146, pp. 11–17, Jul 2016.

[24] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 5, pp. 4104–4108 vol.5, Oct 1997.

[25] V. Vural and J. G. Dy, "A hierarchical method for multi-class support vector machines," in *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, (New York, NY, USA), p. 105, ACM, 2004.
[26] G. Madzarov, D. Gjorgjevikj, and I. Chorbev, "A multi-class SVM classifier utilizing binary decision tree," *Informatica*, vol. 33, no. 2, 2009.
[27] M. Athimethphat and B. Lerteerawong, "Binary classication tree for multiclass classication with observation-based clustering," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 6, no. 2, pp. 136–143, 2012.
[28] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, vol. 2, pp. 985–990 vol.2, July 2004.
[29] 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)," 2017.
[30] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 Dataset: A Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions," *Sci. Data 5, 180161 doi:10.1038/sdata.2018.161*, 2018.

Appendix A
Dark Corner Remove

In this work a new method for removing dark corners is proposed. The method consists of generating a circle mask centered on the image, the circle radius is optimized for the size that best removes the image dark corners. Fig. 4 shows an example of method operation.

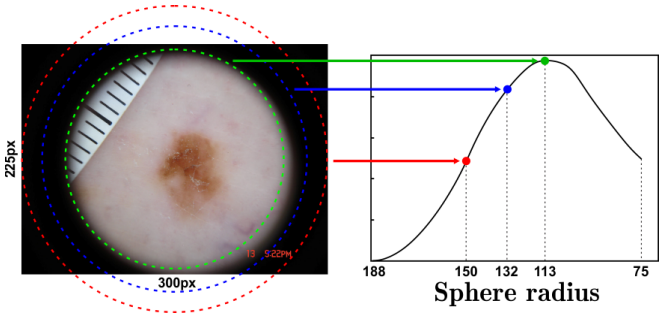

Fig. 4. Example of Dark Corner Remove method operation.

The curve as a function of the circle radius is described by the following equation

$$F(r) = \pi r^2 (J - \tilde{J}) \tag{6}$$

where r is the circle radius, J is the amount of black pixels in the entire image and $\tilde{J}$ is the amount of black pixels in the inner area of the circle. J is computed by the equations 7 and 8.

$$J = \sum_{i=1}^{n} \sum_{j=1}^{m} \phi(I_{ij}) \tag{7}$$

$$\phi = \begin{cases} 1, & I_{ij} < \delta \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$\tilde{J}$ is calculated exactly equal to J with the difference that only the pixels inside the circle area are counted.

Due to some noises present in the images the signal is filtered by performing a convolution operation between the signal and a Gaussian filter G described by

$$H(k) = (F * G)(k) \tag{9}$$

Maximizing H we have the ideal value of circle radius, with optimal circle radius in hand it is possible to generate a final binary mask to remove the dark corner. Images whose H curve has near zero or zero variance at the beginning of the curve are treated as having no dark corner.