Recognition of skin melanoma through dermoscopic image analysis

Catalina Gómez*[a] and Diana Sofia Herrera*[b]

[a,b]Departamento de Ingeniería Biomédica, Universidad de los Andes, Bogotá, Colombia

ABSTRACT

Melanoma skin cancer diagnosis can be challenging due to the similarities of the early stage symptoms with regular moles. Standardized visual parameters can be determined and characterized to suspect a melanoma cancer type. The automation of this diagnosis could have an impact in the medical field by providing a tool to support the specialists with high accuracy. The objective of this study is to develop an algorithm trained to distinguish a highly probable melanoma from a non-dangerous mole by the segmentation and classification of dermoscopic mole images. We evaluate our approach on the dataset provided by the International Skin Imaging Collaboration used in the International Challenge *Skin Lesion Analysis Towards Melanoma Detection*. For the segmentation task, we apply a preprocessing algorithm and use Otsu's thresholding in the best performing color space; the average Jaccard Index in the test dataset is 70.05%. For the subsequent classification stage, we use joint histograms in the YCbCr color space, a RBF Gaussian SVM trained with five features concerning circularity and irregularity of the segmented lesion, and the Gray Level Co-occurrence matrix features for texture analysis. These features are combined to obtain an Average Classification Accuracy of 63.3% in the test dataset.

Keywords: classification, color histograms, melanoma, segmentation, skin cancer, SVM.

1. INTRODUCTION

Skin cancer is the uncontrolled growth of abnormal skin cells. It is the result of unrepaired DNA damage that triggers mutations, genetic defects, and leads to the rapid multiplication of skin cells, ultimately forming malignant tumors.[1] The National Cancer Institute[2] states that skin cancer is the most common one since it usually appears in skin that has been exposed to sunlight, or can occur in any body part. Although there are different types of skin cancer, the two most frequent ones, Squamous cell cancer and Basal cell cancer, usually respond well to treatment and rarely extend to other body parts. In contrast Melanoma, the third most common type,[2] is the most deadly due to its likelihood to spread into other parts of the human body.[3] Melanoma is a type of cancer that occurs in melanocytes, which produce a dark pigment called melanin that gives extra protection to the skin from ultraviolet light. However, a cancer cell does not always produce a dark pigment, which makes the tumor take on a pale red or pink hue.

The American Cancer Society states that there are five characteristics known as **ABCDE** that can be visually examined to differentiate a possible melanoma from a regular mole, these are:[4]

1. **A**symmetry: half of the mole is not symmetric with the other.
2. **B**order: borders are irregular and not well defined.
3. **C**olor: is not uniform throughout the mole. It might have black, brown, pink, red, blue or white spots.
4. **D**iameter: the diameter of the mole is bigger than 5 mm, although melanomas can be smaller.
5. **E**volution: size, color or shape of the mole change through time.

The first four items can be studied through the analysis of a single picture of the skin lesion of the patient, but the fifth criterion, evolution, cannot be studied with a single image from the lesion, as it would be necessary

a: E-mail: c.gomez10@uniandes.edu.co
b: E-mail: ds.herrera10@uniandes.edu.co
* Equal contribution.

to have a series of pictures of the lesion through time. Likewise, the traditional medical approach to diagnose melanomas needs a sequence of observations over time and additional information provided from an appointment with a doctor to evaluate the symptoms. As the dataset of the International Skin Image Collaboration (ISIC) is composed of a single image for each patient, our algorithm only takes into account the first four items that can be evaluated with the available information.

Acquisition of the images to be analyzed is performed by a technique called dermoscopy, a noninvasive method that uses a dermatoscope and allows the examination of the skin by a magnification and lighting system, in order to distinguish malignant skin lesions, such as melanoma, from benign ones.[5] An accurate diagnosis can be achieved by experts, but the lack of training and interest can lead to an increase in the number of unnecessary excisions, which are disadvantages of using this tool.[6]

The ISIC started a challenge in 2016 with the main goal of developing image analysis to enable the automated diagnosis of melanoma from dermoscopic images.[7] In the context of the challenge, image analysis of skin lesions is composed of three parts: lesion segmentation, which is the task of identifying all the pixels that belong to the skin lesion; detection and localization of visual patterns/features, such as the ABCD, and disease classification in malignant or benignant. Hence, an automatic algorithm must be developed in order to fulfill the three tasks. Although we evaluate results only for the segmentation and classification stages, the features we design are used as an intermediate step to classify images.

Our aim is to develop an algorithm that first segments the mole from the background, and then extracts visual features to classify the skin lesion. In general, our approach is based on the visual characteristics (ABCD) that dermatologists analyze, which means that color and shape information must be extracted from the dermoscopic images. The overall process that we define is presented in Fig. 1.

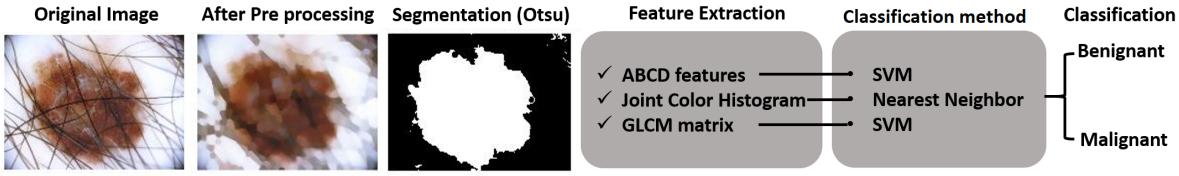

Figure 1. Overall view of our algorithm with an example, which includes preprocessing for hair removal, segmentation with Otsu's thresholding, feature extraction, and lastly, final categorization as benignant or malignant class through Machine Learning techniques.

A potential application of this algorithm would be in telemedicine to help patients identify potentially risky skin lesions when they cannot attend a doctor, or as a supporting tool for the diagnosis made by the specialist. This is particularly useful in a country like Colombia, in which 24.5% of the population has difficulty accessing the health system,[8] and where it can take months for a person to get an appointment with a doctor due to its inefficiency. Additionally, according to the Instituto Nacional de Cancerología-ESE,[9] from 2003 to 2011 there were 1,108 new cases of skin melanoma in Colombia, highlighting the relevance of the problem. These considerations compelled us to develop a simple, yet powerful algorithm that does not need high computational power to be executed, thus being affordable and widely applicable in low-income countries. It is paramount to note that our algorithm is created as a tool to help specialists identify a potential melanoma. Hence, an appointment with a dermatologist should be made to further study whether the mole is malignant or not with a biopsy.

The organization of the paper is as follows. Related work for this problem is described in Section 2. In Section 3 we describe our approach for the segmentation and classification tasks, Section 4 describes the experiments and evaluation results of our algorithm and, finally Section 5 presents some concluding remarks of our work.

2. RELATED WORK

The problem of skin melanoma detection has been addressed by different research groups in order to find an automated methodology for evaluating skin lesions based on digital images. The aim is to develop a robust

algorithm in order to provide a supporting tool to facilitate the diagnosis to decrease mortality rate, together with a reduction of the percentage of error.

In the experimental setup of the ISIC challenge,[7] for the segmentation task, results from automated segmentation were compared with human expert annotations and performance was measured with the dice coefficient, pixel level accuracy, specificity, sensitivity, and average precision at sensitivity of 100%. For the second task, automated predictions of dermoscopic features such as streaks (focusing on starburst pattern), globules, localization and classification were compared with annotations from expert dermatologists. Performance was measured with the dice coefficient, pixel level accuracy, specificity, sensitivity and final score was given by average precision. Finally, for the third task, each team could extract their own features to predict lesion disease state of benignant or malignant; centers with expert pathologists gave the annotations for comparison. Performance was measured according to accuracy, specificity, sensitivity, specificity at 99% sensitivity, and average precision evaluated at sensitivity of 100%. Again, average precision give the final result.

A preprocessing stage is important before addressing the segmentation problem.[10–12] This stage involves algorithms to enhance the quality of the images by highlighting relevant information and reducing noise or unwanted artifacts. For instance, a Median filter was reported by[10] and a Gaussian filter was applied to gray scale images,[11,12] and then converted to binary images by setting a threshold with Otsu's method. Additional methods for improving details (filling holes, preserve edges) were applied with morphological operations, such as closing, opening and reconstruction.[11] However, there is a drawback with thresholds if there is not a good contrast between the skin and the lesion, therefore a single threshold can not be applied to all the dataset. Thus, we put special care into evaluating color spaces other than RGB and optimizing the selection of a threshold in order to create a generalizable segmentation algorithm.

An alternative approach for the segmentation problem is to implement active contour methods, also known as *Snakes*, which deal with outlier edges.[13] This technique is based on the detection of contour segments (strokes) using edge linking and then approximates a subset using an expectation-maximization algorithm. On the other hand, H. Wang *et al.*[14]implemented a watershed technique for segmentation of dermoscopic images. Hair removal was achieved with the morphological closing operator to avoid over-segmentation because of similar intensities. This technique was improved by controlling the lesion size with an outer bounding box method to estimate the ratio between the area of the lesion and that of the whole image.

After the identification of the skin lesion, the next step is to classify it as a melanoma or as a common mole. Several methods have been proposed for the classification into more categories, such as support vector machines (SVM), decision trees and neural networks. The input of the classification algorithms are the features extracted from the detected lesions, such as the ABCD characteristics mentioned before to identify a melanoma.[12] V. Jeya Ramya *et al.* reported four features, variance, energy, correlation, homogeneity and entropy, that have a fast extraction and resulted in higher discrimination accuracy.[15] An alternative method was proposed in,[10] in which the lesion was subdivided into regions (normal skin, peripheral, tumor) and features were extracted for each one.

A simple yet useful method is the one used by J. Shivangi,[16] which started by preprocessing the image with a contrast enhancement. Afterwards, automatic thresholding in each RGB plane was applied to facilitate the creation of binary masks to locate the skin lesion by finding the biggest blob, allowing thus a more precise segmentation of the lesion. Then, edge detection was used to segment only the skin lesion, to proceed to extract geometry-based features to evaluate ABCD criteria, the main information needed to classify the lesion. For this step, area, perimeter, greatest and shortest diameter, circularity index and irregularity indexes A, B, C and D were used.

With respect to available code, there is not an open source implementation for melanoma detection, but there are some applications that have been developed for this issue. By July 2014, 39 applications were available, most of them developed for iOS systems, nevertheless most of them were educative or based on taking and storing pictures for future appointments with doctor. Nine of them offered expert review of images and only four provided a risk assessment about the probability of a lesion to be malignant or benign. Additionally, one application calculates future risk of melanoma. However, none of them have been validated for diagnosis accuracy.[17] By the end of 2014, a German application, SkinVision, became available for iOS and Android systems, which claims to be the first CE certified melanoma application.[18] This one analyses the mole and allows the user to follow the

evolution, although scores and reviews on App Store and Google Play are very low. The latest innovation is the free application Mole Mapper, developed by Oregon Health & Science University one year ago to measure and monitor moles, which does not have enough reviews yet to be graded.[19]

Results comparisons of our method with the algorithms exposed in the related work section are not made because they were evaluated in different datasets, mainly composed of a smaller amount of images and with less diversity.

3. APPROACH

3.1 Dataset description

The International Skin Imaging Collaboration Archive provides a dataset of dermoscopic images that have undergone annotation by experts. The organization is an academia-industry partnership whose aim is to reduce melanoma mortality by providing high quality and standardized images and annotations to developers capable of creating an algorithm for skin cancer triage and diagnosis.[7] The annotations of each picture contain clinical information with a pathology diagnosis, lesion-level attributes such as symmetry and sub-lesion level features as pigment networks within a region of the lesion. The whole dataset is formed by five smaller data sets: UDA-1, UDA-2, MSK-1, MSK-2 and SONIC-1. The ones we used were:

1. UDA-1: includes 557 cutaneous melanocytic lesions with either a histopathological diagnosis or clinically benign history.

2. UDA-2: with 44 high quality lesion images of melanomas and benign lesions, containing metadata with patient age, diagnosis, gender and anatomic location.

3. MSK-1: includes lesions that were excised with diagnosis, anatomic location, gender, age and clinical impression included in metadata.

4. MSK-2: includes benign and malignant skin lesions.

Fig. 2 shows examples of moles classified as benign (top row) and malignant (bottom row) by the specialists. Visual patterns characterize each mole type, such as regular shape and color, but the difficulty of the task lies on moles that do not show typical features.

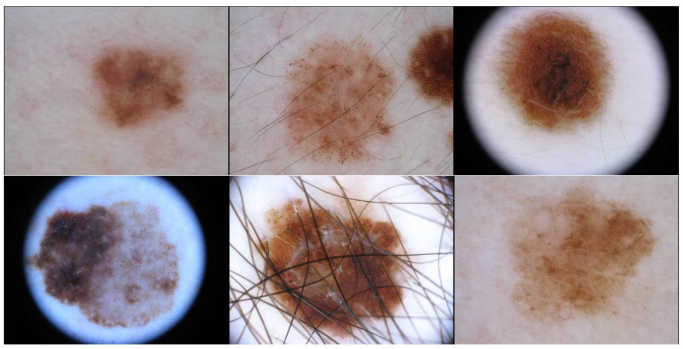

Figure 2. Examples of dermoscopic images. Top row: moles annotated as benign. Bottom row: moles annotated as malignant.

For the segmentation task, we randomly split the dataset into balanced training and validation subsets with 389 and 390 images, respectively. For this division we do not take into account the class of each image since the main goal was to distinguish the mole from the background. Then, for the classification stage, a balanced dataset composed of images from the segmentation dataset was built. In the training set, 177 images are used, where 87 are malignant melanoma and 90 are benign moles. For validation, 176 images are chosen with 86 malignant and

90 benign. Lastly, the test dataset is composed of 150 images, where 75 are malignant and 75 are benignant. Annotations for this stage are benign and malignant classes.

To evaluate the algorithm in the segmentation problem, the Jaccard Index is computed as the intersection over the union between the predicted mask and the annotations, to demonstrate the true positive over false positive plus false negative ratio. The classification problem will be evaluated by calculating the Average Classification Accuracy (ACA), computed as the average of the diagonal of the normalized confusion matrix, which permits the analysis of individual class performance in terms of true positives for both classes.

3.2 Segmentation of skin lesions

Based on our attempt to reproduce the feature analysis that dermatologists perform, we process the images starting from color spaces to distinguish the mole from the background for the segmentation task. First, we apply a preprocessing step to remove hair from the images, as in the instance shown in Fig.2 bottom row at the center, to eliminate non-relevant information that could affect the segmentation. Then, with the images from the training set we consider the channels of different color spaces (RGB, Lab, HSV, YCbCr) and determine the channel in which the mole was most easily distinguished from the background (based on pixel intensity). In order to extract the mole, we apply Otsu's thresholding in the most relevant channel to have pixels labeled as mole and background. To improve the quality of the segmentation mask for the posterior classification, we make sure that only the largest connected component was part of the segmentation mask, removing black borders present in some images (see Fig. 2). To evaluate the performance of our segmentation method, we compute the Jaccard Index of all the candidates by comparing it to the respective annotation, and report the average over all the validation and test sets.

3.3 Feature extraction and Classification

Before classifying lesions into benign and malignant, it is necessary to balance the classes in the dataset to avoid overfitting when training the model. This process is accomplished by randomly selecting the same number of images of benign and malignant lesions. Each class is represented with features that distinguish it from the other, such as color, texture and shape of the moles. This information can be obtained from every pixel, or from a region of interest.

For the color features approach, we represent the original images with joint color histograms (all the channels of the color space) in different color spaces. The common metrics used to compare two normalized histograms are the Intersection Kernel and the Chi-Squared Distance in order to implement a nearest neighbor classifier based on color. We compare the histograms of the images from the validation set with each histogram from the images on the training set. Then, for the classification task we choose the minimum Chi-Squared Distance and the maximum Intersection Kernel that reflects a highest affinity measurement, and finally, assign the label of the image with the highest match.

For shape and texture analysis, the extracted features are the inputs for a classifier trained to predict a label. These features are the shape attributes of the moles according to the ABCD criteria, and texture features based on local information. The classifier is a Support Vector Machine (SVM), for which different kernels are tested such as linear, cubic, RBF Gaussian, among others. For shape analysis, feature representation is extracted from the images with the segmentation mask. Features such as the border of the object, area, orientation, major (MajorAL) and minor axis length (MinorAL) and the centroid are computed. Additionally, the circularity (equation 1) and irregularity indexes (equations 2, 3, 4, 5) are also calculated.[20]

$$Circularity = 4 \times \pi \times \frac{Area}{Perimeter^2} \tag{1}$$

$$Irregularity(A) = \frac{Perimeter}{Area} \tag{2}$$

$$Irregularity(B) = \frac{Perimeter}{MajorAL} \tag{3}$$

$$Irregularity(C) = Perimeter \times \left(\frac{1}{MinorAL} - \frac{1}{MajorAL} \right) \tag{4}$$

$$Irregularity(D) = MajorAL - MinorAL \tag{5}$$

Once the five features are extracted from the training set, they are concatenated within a matrix and normalized. The matrix of features is the input to train the model, and the predictors, response and model type are specified within the Classification Learner application in Matlab. Since the application allows training many models in short time, we keep as the final model the one with the best accuracy in the training dataset, and export it to predict the response for the validation set.

Since the five features of shape are extracted from the mole delimited by the segmentation mask, the quality of our segmentation procedure could affect the results of the posterior classification. If there are mistakes in the mole segmentation (background inclusion or missing the mole), then the classification would not be accurate and we could determine whether the limitation lies in the segmentation or classification algorithm. Hence, to verify the effect of our segmentation algorithm in the classification task, we perform an Oracle experiment in which we extract the features from the moles marked by the annotation masks and compare the results of the classification accuracy.

Based on the approach for classification developed by Kavitha, J. C. and Suruliandi, A.,[21] additional features based on texture are extracted from the training images to train the SVM model. The Gray level co-occurrence matrix (GLCM), a statistical method for texture analysis, is used to extract additional features such as contrast, correlation, energy and homogeneity. Contrast is related with a measure of the intensity contrast between a pixel and its neighbor over the whole image, the energy is the sum of the squared elements in the GLCM, correlation is a measurement of how correlated is a pixel to its neighbor over the whole image, and homogeneity measures the closeness of the distribution of elements in the GLCM to its diagonal.[22] These features are used to train a new SVM, with a RBF kernel, in which C and γ are optimized to obtain the best accuracy for predictions.

We compute the Average Classification Accuracy (ACA) from the confusion matrix in order to evaluate all the approaches described above.

4. EXPERIMENTAL VALIDATION

4.1 Experimental setup

Evaluation of the algorithms is conducted in both validation and test datasets of the ISIC challenge.

4.1.1 Segmentation

The preprocessing step to remove hair includes image opening preceded by image closing, both with the same structuring element. To differentiate the mole from the background, we use Otsu's thresholding in channels of different color spaces and clustering with k-means (k=2) on RGB images. Given that we expect clearly separated peaks in the histogram for the lesion and the skin, the clustering technique is appropriate for images where there is color contrast between the mole and the skin, but it is not the general case, so this method was discarded. The training images are used to analyze in which channel and color space, among RGB, Lab, HSV and YCbCr, the mole could be better separated from the skin or background. Afterwards, we apply Otsu's thresholding in the chosen channel. Then, we calculate connected components within the binary image and keep the largest one. After removing black borders, we check if the main object was removed too, which implies that either the image did not contain black borders or the mole component was connected with the border. Finally, we compute Jaccard Index to evaluate the performance of our algorithm.

4.1.2 Classification

We use three different representation spaces to classify the segmented skin lesions: color histograms, texture and irregularity features (ABCD irregularity indexes and circularity). We list the different experiments below:

- **Joint Color Histograms:** compute joint histograms in RGB, HSV, Lab and YCbCr color spaces and compare validation with training histograms with the Chi-Squared Distance and the Intersection Kernel metrics. We keep the color space and affinity metric that yields the best results with ACA.

- **ABCD irregularity indexes and circularity features:** we compute these features with the resulting segmentation mask, after black corners removal and taking into account only the largest connected component. We follow equations 1, 2 and 3 shown on Section 3.3. We use the Classification Learner application in Matlab to train several different classifiers and choose the one with best ACA score. Afterwards, we modify the hyperparameters from the chosen classifier to find optimal values. To evidence the effect of segmentation results, we also compute the features with the annotation masks to compare the resulting ACA.

- **GLCM features:** we extract texture features (contrast, correlation, energy and homogeneity) at different pixel orientations (0°, 45°, 90° and 135°). These features are the input to train the SVM with a Radial Basis Function (RBF) kernel, optimizing the parameters of penalty C and gamma with grid search on the train set. The accuracy of classification for the training model was 68.92%.

- **Combination of texture and shape features:** we concatenate within a matrix the five features of irregularity, circularity and the texture features at four orientations, and use it as the input to train an SVM model. Using the Classification Learner application we train different models and finally keep the one with the highest accuracy and highest true positive rate for malignant class. The ACA for the training model with a RBF Gaussian SVM was 63.8%.

- **Combination of color and five features:** we train a second layer SVM using a matrix concatenating joint color histogram and irregularity features.

- **Relationship between outputs from color and five features:** Given the probability output from the SVM trained with the five irregularity features together with the classification made according to color histograms, we calculate an optimized linear relationship between them.

An alternative classification system is created by combining joint color histograms in the YCbCr color space and the features of irregularity and circularity. In this case, the feature classification returns the probability of the image to belong to a certain class, benign (-1) or malignant (1). Also, given the information in the confusion matrices for color and feature classification on their own, we observe that color gives important information, leading to a high classification score for malignant moles while feature representation is more accurate with the benign class. Taking this behavior into account, four cases are considered and determined empirically:

1. If the color and feature models predict the benign class, the image is classified as benign.

2. If the color and feature models predict the malignant class, the image is classified as malign.

3. If the color model predicts malignant and the feature model benignant, a threshold (TA) (equation 6) is considered. In the case that the score given by features is higher than TA, the image is classified as benign, if not it is classified as malignant.

$$TA = BMS + BMSSD \times \alpha \tag{6}$$

   BMS = Benign mean score.
   BMSSD = Benign mean score standard deviation.
   Where α is optimized in the training set.

4. If the color model predicts benignant and the feature model malignant, a threshold (TB)(Equation 7) is taken into account. The feature score has to be higher than TB in order to classify the image as malignant, if not it is benignant.

$$TB = MMS + MMSSD \tag{7}$$

   MMS = Malignant mean score.
   MMSSD = Malignant mean score standard deviation.

4.2 Results

The results over the validation and test sets of the experiments described in the previous section are shown in Table 2, for both segmentation and classification tasks. All the values are within a range from 0 to 1. The average Jaccard Index for segmentation in both validation and test datasets is acceptable when compared with the winner of the challenge, whose result was 84.3%.[23] Hence, the agreement with the annotations increases with the distinguishable features of YCbCr color space. We only present results for the best performing color space due to space limits. An example of a segmentation with our method is shown in Fig. 3, in which the annotation covers almost all the image, while our segmentation mask comprises the mole with some background included. We consider this as one of the limitations of the dataset, many annotations are not precise.

Similarly, the majority of results are favorable and some examples can be seen in Fig. 4. From left to right, columns one and two evidence that light-colored skin lesions are properly segmented, even with better edge preserving results than the annotations, as can be seen in the second column. The third column shows that the hair removal process is successful, as the resulting mask is similar to the ground truth even though the original image had a lot of hair. Lastly, the fourth column shows that the YCbCr color space yields good results even with images comprising black borders or the corners, thus reducing the negative effect of these artifacts.

Additionally, Fig. 5 presents examples of the main limitations of our segmentation algorithm. From left to right, the first column shows that images in which there are objects other than mole or skin cannot be segmented properly with the YCbCr color space, although they could be easily segmented using other color spaces. In the middle column, the resulting mask has only one connected component comprising mole and black corners together, hence, removing connected components in contact with the edge would remove the mole itself. Lastly, the right column evidences that the biggest connected component is not always the skin lesion, especially in images with black corners. This is why, although we tested our algorithm removing connected components that touch the image boundaries and keeping only largest connected component resulting after this process, we decided not to execute these steps for final segmentation because the average Jaccard Index decreased about 15%.

Table 1. Results for Validation and Test datasets in the segmentation task using Otsu's thresholding method in YCbCr color space. Values displayed correspond to Jaccard Index.

Segmentation (Jaccard Index)		
Method	Validation	Test
Otsu in YCbCr (Cr channel)	0.691	0.701

Table 2. Overall Results for Validation and Test datasets in the classification task in terms of ACA. Results for all the experiments described in Section 4.1.2 are displayed.

Classification (ACA)		
Method	Validation	Test
Only GLCM features	0.665	0.527
Shape features	0.654	0.613
Shape features WITH ANNOTATIONS	0.559	0.500
Shape features and GLCM features	0.462	**0.633**
YCbCr joint color histograms	0.646	0.493
Color Histograms + shape features	**0.734**	0.580

Figure 3. Comparison between the segmentation we obtain with our method and the annotation in an image of the validation set, where borders are better preserved in our result.

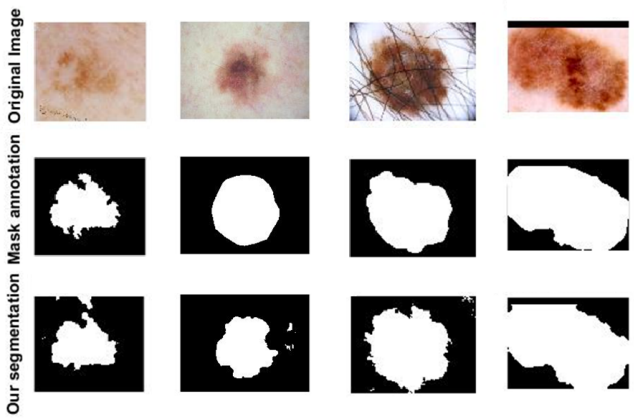

Figure 4. Examples of appropriate segmentations on the segmentation task obtained with our method. Each column corresponds to a single case. The first row contains the original image, second row the annotation and third one our segmentation.

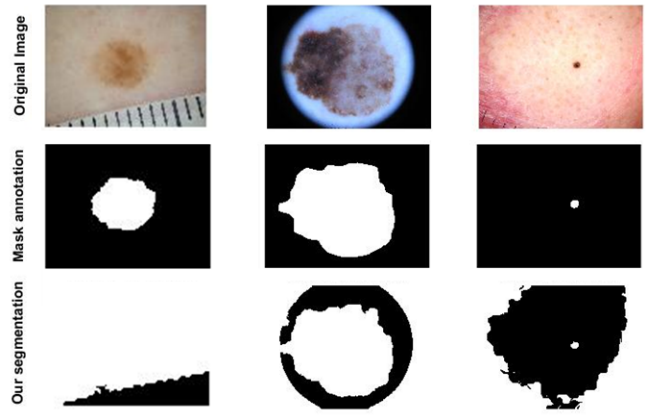

Figure 5. Examples of limitations on the segmentation algorithm. Each column corresponds to a single case. The first row contains the original image, the second row the annotation and the third one our segmentation.

When comparing the results of the classifiers, the three approaches (color histograms, shape features and texture features) exhibit a similar behavior, with an accuracy around 65% on the validation set. On the one hand, the accuracy in the validation set was improved by combining joint color histograms with the five features, 73.4 % which is greater than the accuracy of a random classification (50%). The combination results in a good classification methodology since, on the one hand, the classification based on the color histograms recognizes

almost all of the malignant melanomas (with an accuracy of 83.7%), but misclassifies benign moles (45.6%). On the other hand, the model trained with the 4 features learns to classify benign moles (87.8%), but does not perform well for malignant ones (43%). The performance of the color classifier is better in terms of prevention to alert the patient to contact a specialist and check the condition of the lesion.

The best accuracy for the test set is obtained with the model that combines texture and irregularity features with an ACA of 63.3%, despite the fact that the ACA of the validation set is below the value of random classification. The divergence of the results in the validation and the test sets could be due to different statistics between these sets, and also to the non-homogeneous nature of each class.

5. CONCLUSIONS

Dermoscopic skin lesion images are not homogeneous and contain non-relevant information that might affect results, as can be seen in Fig. 5; hence preprocessing must be done to remove this information. The YCbCr color space demonstrated to be useful for segmentation by Otsu's thresholding when compared to other color spaces and clustering methods such as k-means. Additionally, the segmentation results have a great impact on the classification output as only color, texture and feature information from the mole should be taken into account, ignoring the background. The resulting mask is especially important for irregularity features and circularity value, as they are computed based on the mask rather than the RGB image. Thus, comparing the ACA obtained with these features based on our own segmentation against the annotation masks gives information about the effect of our segmentation algorithm in the classification results. When these values are compared, they demonstrate that our segmentations did not affect classification, and actually helped to improve results in both Validation and Test datasets, as the ACA with our segmentations was 65.4% and with the annotations of the segmentation masks it was 55.9% for the first one and 61.3% against 50% in the second one. We hypothesize that this could be because our algorithm preserves edges and small details, whereas some annotations do not track these details, as shown in Fig. 3 and Fig. 4 column 2.

Furthermore, a wide range of features can be taken into account to classify melanomas. Nevertheless, many of them increase significantly the computational cost and do not increase the ACA. The features chosen in this paper were carefully evaluated individually and mixed according to their confusion matrices. Using different features is important because classification made only through joint color histogram leads to poor results just as judgment made from a non-expert point of view. We are not aware of any other approach that analyzes and combines features through a deep analysis of their behavior between classes.This could be a possible reason to explain why, when compared to the winner of the 2016 challenge (when our algorithm was created), whose ACA was 63.7%,[24] our result of 73.4% is significantly higher in Validation dataset, and comparable in Test dataset, with an ACA of 63.3%.

Lastly, it is paramount to note that deep learning algorithms can be applied to solve skin lesion classification, as did the winners of ISIC Challenge on 2017, who obtained a 87.4% score on the classification task.[25] Nevertheless, our simple yet effective algorithm, developed following the specialists' insights, obtains a competitive score, does not require the usage of advanced computer power and takes short time to be executed, providing an affordable method to classify melanomas when compared to neural networks. Our approach can be specially useful for low-income countries like Colombia in which an important part of the population does not have access to specialized medical centers and is in need of opportune medical attention or early diagnosis.

ACKNOWLEDGMENTS

We gratefully acknowledge professor Pablo Arbeláez for sharing his pearls of wisdom with us during the course of this research and to Laura Bravo, Gustavo Pérez, Andrés Romero and Alejandro Pardo for comments that greatly improved the manuscript.

REFERENCES

[1] "Skin Cancer Information - SkinCancer.org", Skincancer.org, 2017. [Online]. Available: http://www.skincancer.org/skin-cancer-information. [Accessed: 12- Jun- 2017].

[2] National Cancer Institute. National Institutes of Health, 2016. [online] Available at: http://www.cancer.gov/types/skin
[3] Copyright American Cancer Society, 2016, Melanoma Skin Cancer, melanoma skin cancers pp 2.
[4] Copyright American Cancer Society, 2016, Melanoma Skin Cancer, melanoma skin cancers pp 13.
[5] DermNet New Zeland. Dermatoscopy. Introduction to dermatoscopy. [online] Available at: http://www.dermnetnz.org/cme/dermoscopy-course/introduction-to-dermoscopy/
[6] V. Jeya et al. Detection of melanoma skin cancer using digital camera images. ARPN Journal of Engineering and Applied Sciences. 10 (7):3082-3085, 2015.
[7] International Skin Imaging Collaboration: Melanoma Project, 2014. [online] Available at https://isic-archive.com/#
[8] J. Ayala. La salud en Colombia: más cobertura pero menos acceso. Banco de la República. Centro de Estudios Económicos Regionales. [online] Available at: http://www.banrep.gov.co/sites/default/files/publicaciones/archivos/dtser_204.pdf
[9] Instituto Nacional de Cancerología- Ministerio de Salud y Portección Social- Colombia. [online] Available at: http://www.cancer.gov.co/cancer_en_cifras
[10] R. Suganya. An Automated Computer Aided Diagnosis of Skin Lesions Detection and Classification for Dermoscopy Images. 2016, Fifth International Conference On Recent Trends In Information Technology.
[11] O. Abuzaghleh, B. Barkana and M. Faezipour. Automated Skin Lesion Analysis Based on Color and Shape Geometry Feature Set for Melanoma Early Detection and Prevention. Systems, Applications and Technology Conference (LISAT), 2014 IEEE Long Island.
[12] T. Satheesha, D. Satyanarayana, M. Giriprasad and K. Nagesh. Detection of Melanoma Using Distinct Features. 2016, 3rd MEC International Conference on Big Data and Smart City
[13] M. Silveira *et al.*.Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 3, NO. 1, FEBRUARY 2009.
[14] H. Wang *et al.*. Watershed segmentation of dermoscopy images using a watershed technique. Skin Res Technol. 2010 Aug; 16(3): 378-384.
[15] V. Jeya Ramya, J. Navarajan, R. Prathipa and L. Ashok Kumar. Detection of Melanoma Skin Cancer using Digital Camera Images. ARPN Journal of Engineering and Applied Sciences. VOL. 10, NO. 7, APRIL 2015.
[16] J. Shivangi, V. Jagtap, N. Pise, 2016, Computer Aided Melanoma Skin Cancer Detection Using Image Processing. Procedia Computer Science 48 (2015): 735-740. Web. 16 Oct.
[17] A. Kassianos, J. Emery, P. Murchie, F. Walter, 2015, Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. British Journal of Dermatology, Volume 172, Issue 6, pp. 1507-1519.
[18] Skin Vision, 2014, Check any mole for skin cancer risk. [online] Available at https://skinvision.com/
[19] OHSU Dermatology, 2015, Mole Mapper. [online] Available at https://goo.gl/sazE6A
[20] S. Jain, V. Jagtap, N. Pise. Computer Aided Melanoma Skin Cancer Detection Using Image Processing, 2015. Procedia Computer Science Volume 48, pp735-740.
[21] Kavitha, J. C., & Suruliandi, A. Texture and color feature extraction for classification of melanoma using SVM. 2016. In Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on (pp. 1-6). IEEE.
[22] Documentation. (n.d.). Retrieved December 4, 2016 from: https://goo.gl/wWu4TH
[23] Covalic. ISBI 2016: Skin Lesion Analysis Towards Melanoma Detection Part 1: Lesion Segmentation. 2016. [online] Available at https://challenge.kitware.com/#phase/566744dccad3a56fac786787
[24] Covalic. ISBI 2016: Skin Lesion Analysis Towards Melanoma Detection Part 3: Lesion Classification. 2016. [online] Available at https://challenge.kitware.com/#phase/5667455bcad3a56fac786791
[25] ISBI 2017: Skin Lesion Analysis Towards Melanoma Detection Part 3: Lesion Classification. 2017. [online] Available at https://challenge.kitware.com/#phase/584b0afccad3a51cc66c8e38