Early Melanoma Detection Based on Chromatic Descriptors and Machine Learning Algorithms

Andreea Udrea

Department of Automatic Control and Systems Engineering,
University Politehnica of Bucharest, Bucharest, Romania,
e-mail: andreea.udrea@acse.pub.ro

Abstract: This article proposes a pigmented nevi classification methodology such to early detect melanoma. This is a deadly type of skin cancer with an alarming continuous growth in incidence. The proposed classification strategy can be applied on color conventional images; it is based on chromatic descriptors of the skin lesions and uses support vector machines for classification. For the identification of the most reliable chromatic descriptors (normalized histograms), the following color models are investigated: RGB, three normalized RGB versions, HSV, Ycbcr and Opponent color model. On a 604 conventional images database (138 images of melanomas), a sensitivity and specificity of 78% and, respectively, 82% in detecting melanoma is obtained. Out of 19 images of melanoma in situ (stage 0 tumors), only one is misclassified as benign. By using a larger database it is expected that the accuracy of the proposed classification procedure can be increased.

Keywords**:** chromatic descriptors, early melanoma diagnosis, support vector machines

1. INTRODUCTION

Skin cancer diagnosis based on automatic algorithms is a challenging task and a hot research subject because skin cancer is one of the most common malignancies in populations with I-III phototypes (fair skin). It is estimated that 50% of the cancers that are annually detected are of this type and the incidence increases from year to year with a substantial and alarming rate (Little et al., 2012). This increase is reflected in the degree of morbidity and mortality associated with these diseases and, respectively, with increased treatment costs. For example, in the US, 2 million BCC and SCC and 75 000 MM are detected annually.

Skin cancer can be largely divided (considering those with high incidence) in two main classes: melanoma (MM) and non-melanoma (NMSC). Melanoma is the most aggressive, with high morbidity and mortality associated rate if not detected in its early stages. It is characterized by abnormal color and shape and rapid development.

Basal cell carcinoma (BCC) is the most common form (75%) of NMSC and has an associated mortality rate much lower than MM. It appears in the form of small pearly elevations. Squamous cell carcinoma (SCC) represents about 20% of NMSC and its associated risks are lower than those of MM. It appears as small red elevations or ulcers that do not heal.

The diagnosis of these lesions is normally performed by dermatologists by clinical inspection (visual), dermoscopy or other non-invasive methods of diagnosis. Early diagnosis (incipient phase) of these lesions is of paramount importance, since it leads to improved treatment outcome and lower treatment costs. Diagnosis is a subjective process, and the doctors sensitivity and specificity in the diagnosis of MM, BCC and SCC varies from country to country and from urban to rural areas. Also, patients often address the general practitioner about these diseases, and these are not familiar with these heath conditions; this leads to poor diagnostic accuracy (sensibility around 65%). The gold standard in dermatology, for these cases remains the histopathology result, but this involves invasiveness (lesion excision).

In general, dermatologists make their diagnosis based on the color, contour, texture, elevation properties of the mole. This method is called ABCDE (Asymmetry, Border, Color, Diameter, Elevation). This is used for clinical inspection (and leads to sensitivity around 75%). The dermatologist sensitivity rises to around 85% when using a dermatoscope because this procedure reveals submacroscopic morphologic structures and vascular patterns that are located at different skin layers that are not visible with the naked eye.

Because of the difficulty and subjectivity of human interpretation and paramount importance of early detection (Weinstock et al., 2012), computerized analysis and classification of medical images obtained by standard (dermatoscopy, infrared and multispectral imaging, confocal microscopy) or conventional methods (high performance cameras) has become an important area of research.

In this context, diagnostic systems have been developed to offer dermatologists support in identifying skin cancers (e.g. MELAF - approved by the FDA in 2011 in the US). Typically, these systems are expensive, are not 100% accurate, are usually focused on the detection of MM and dedicated exclusively to healthcare professionals. Besides approved medical systems, there are a lot of proposals of automatic or semi-automatic diagnostic systems that analyze

standard medical or conventional images of skin lesions. Most aim to identify MM due to the high mortality associated with this disease (Alcón et al., 2010 ; Gilmore et al., 2009; Ramezani et al., 2014 ; Chang et al., 2005 ; Piantanelli et al., 2005). Fewer studies deal with the problem of identifying BCC and SCC (Stoecker et al., 2009; Cheng et al., 2011) because they are associated with lower morbidity, although, overall, they represent the skin cancers with higher incidence.

The skin lesions diagnosis procedure performed on dermoscopic or conventional images generally follows these steps: preprocessing (applying filters for improved image visualization, removing noise and artifacts); segmentation and contour extraction; texture, color and contour analysis in the delimited lesion area; relevant descriptor-parameters identification, and classification using these descriptors.

The methods involved in these steps vary greatly between different studies (Premaladha et al., 2014). In general, the procedure closely follows the identification of parameters according to the ABCDE method (although this procedure is not necessarily the most effective).

A number of popular segmentation and boundary extraction methods are as described in (Celebi et al., 2009). In addition there have been proposed methods with superior results in (Li et al., 2011; Schaefer et al., 2011; Wang et al., 2011; Zhou et al., 2011). Each method behaves differently depending on the type of lesion of interest, respectively, image type.

The parameters of interest - descriptors refer to the color, texture and contour properties of the analyzed lesion (Lee et al., 2005; Mete et al., 2011; Stoecker et al., 2011; Dalal et al., 2011).

The classification methods range from discriminant analysis, to support vector machines type. A comprehensive review on this topic is presented in (Maglogiannis et al., 2011). The degree of sensibility and specificity varies from study to study, mainly depending on the parameters selected for classification; used classifier; images types; image database size.

In the last 5 years, with the advent of smartphones phones equipped with powerful cameras and processors, the idea of developing mHealth (mobile health) applications for the analysis of conventional images acquired by general users (non-medical staff) in order to provide diagnosis (or at least an associated degree of risk) for skin lesions become of high interest. In this context, the main goal was identifying MM.

These applications follow largely the same steps, in order to classify lesions, as those proposed for medical image analysis and classification. However, there are at least 2 angles from which the problem of creating a diagnosis mHealth app that differentiates lesions is difficult:

1. the analyzed signals are nonstandardized color images acquired with mobile phones and present noise (hair, other marks and/or lesions, shadows and/or overexposed regions due to inhomogeneous illumination) and cannot be compare in terms of details and quality with dermatoscopic images.
2. the analysis algorithms, however viable in the context of dermatoscopies will have lower performances in the context described above or, they could be even inapplicable (the case of those that analysis reveals submacroscopic morphologic structures and the vascular network).

The position of the medical world regarding these applications is controversial (Robson et al., 2012; Hamilton et al., 2012; Wolf et al., 2012); while the number of studies that present the accuracy of these applications is limited.

In this context, we are developing and constantly improving a medical app that offers a degree of risk associated to pigmented nevi. It was already subjected to a clinical trial and the conclusion was that it identifies MM with a sensibility of 73% and has a specificity of 83%. The risk assessment algorithm is based on fractal methods and classical image analysis methods and classification is made using a rule-based algorithm (Maier et al., 2015).

Following the study conclusions, we have developed an improved image acquisition methodology that acquires well focused, shadows free, centered and completely captured lesions (Udrea et al., 2014) and identify a series of additional shape and texture descriptors with high standalone sensitivity and specificity (Udrea, 2015; Udrea et al., 2015).

As discussed above, the skin lesions medical evaluation is based on the ABCDE method, and, to the present, we only evaluate texture and shape features.

In this study we present a series of color analysis methods and propose a classification procedure based on chromatic features analyzed over a series of color spaces and machine learning algorithms.

The chromatic descriptors analyzed in this study and the classification methods are presented in chapter 2 and the results, their interpretation and discussions follow in chapter 3.

2. MATERIALS AND METODS

2.1 Image preprocessing

For this study, 604 conventional images (acquired with a smartphone) are used. They contain normal nevi, dysplastic nevi and melanomas. The lesions' type is confirmed by the histopathological results.

The lesions are investigated from chromatic perspective, so the lesions' contours have to be identified in order to extract the lesion of interest from the image.

Segmentation and contour detection of skin lesions in non-standard color images is a difficult problem. Our approach is based on a series of steps:

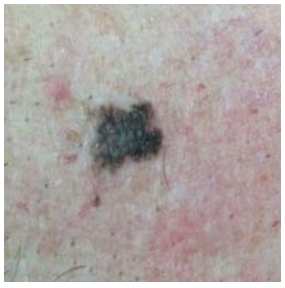
1. the color image (Fig.1.a)) is transformed in a grey scale image
2. a median filter (window of 5X5 pixels) is applied
3. the grey scale image histogram is computed

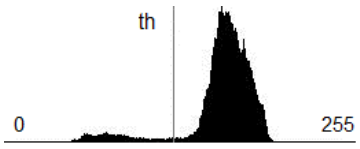
4. Otsu method (Otsu, 1979) is used to identify the threshold that best separates the lesion from the surrounding skin (Fig. 1.b)
5. the grey level image is transformed into a binary image (the previously identified threshold is used)
6. all the connected sets of black pixels are identified (Fig. 1.c))
7. the largest set is (the lesion) is localized and its contour is computed
8. the lesion is extracted from the color image based on the contour identified at the previous step (Fig. 1. d)).

This procedure is fast and offers good results. There is however a small percentage of cases (aprox. 6%) where the lesion is not completely extracted or, besides the lesion, regions containing skin are included in the extracted region of interest.

This happens when the lesion's color is similar to the skin color and/or the shadows level is elevated (Fig. 2).

We have not excluded these cases from our analysis because this might be happening in real functioning conditions.


a)

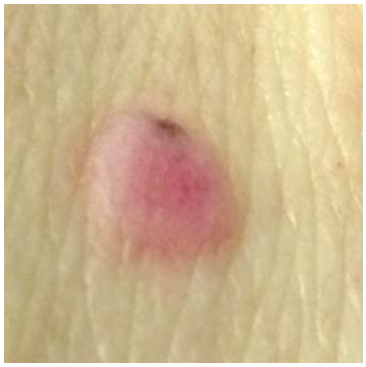


b)


c)


d)

Fig. 1. a) Color image of the lesion; b) the histogram and the selected threshold; c) the binary image; d) the extracted lesion.


a)


b)

Fig. 2. a) Lesion with color close to the surrounding skin color b) incorrect extracted region of interest.

2.2 Color descriptors based on normalized histograms

In this paper we analyze a series of color descriptors based on (normalized) histograms identified for a series of color spaces: RGB (red, green, blue), three normalized RGB versions, HSV (hue, saturation, value), Ycbcr and Opponent color model, where the transformations between RGB and the other color spaces are given below.

Each color space has its own properties and we are going to investigate each of them in order to detect the most suitable for lesion classification based on chromatic properties.

We investigate a large variety of color spaces because there is no general consent which performs best for image analysis of skin lesions.

The Opponent color space is obtained from RGB color space by using the next transformation (1):

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix} \quad (1)$$

The first normalized RGB color model, RGB_l, uses the mean value on each channel (μ) of the skin surrounding the lesion and this is subtracted from each image channel:

$$\begin{pmatrix} R_l \\ G_l \\ B_l \end{pmatrix} = \begin{pmatrix} R-\mu_{R_{skin}} \\ G-\mu_{G_{skin}} \\ B-\mu_{B_{skin}} \end{pmatrix} \quad (2)$$

For the second normalized RGB color model, RGB_N is obtained by dividing each pixel value on a specific channel to the sum of its red, green and blue values:

$$\begin{pmatrix} R_N \\ G_N \\ B_N \end{pmatrix} = \begin{pmatrix} \frac{R}{(\mathrm{R+G+B})} \\ \frac{G}{(\mathrm{R+G+B})} \\ \frac{B}{(\mathrm{R+G+B})} \end{pmatrix} \quad (3)$$

Note that the B_N channel is redundant (R_N+ G_N+ B_N=1).

For the third normalized RGB, RGB', the mean and standard deviation (σ) for each channel is used:

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} \frac{R-\mu_R}{\sigma_R} \\ \frac{G-\mu_G}{\sigma_G} \\ \frac{B-\mu_B}{\sigma_B} \end{pmatrix} \quad (4)$$

For each considered color space, the normalized histogram (probability density function) for each channel (except for B_N) is calculated for different histogram bins number.

Let I_c be a color channel. Let D be the color channel domain, q- the quantization level, then the number of bins for each channel histogram is D/q. To obtain the normalized histogram we divide all the bins' values $h(i)$ to A – the number of lesion's pixels (lesion's area):

$$h(\mathrm{i}) = \frac{1}{A}\sum_{j=1}^{A} I_c(\mathrm{j}),\ I_c(\mathrm{j}) \in (i, i+1],\ i = \overline{1, D/q} \quad (5)$$

The normalized histograms are the chromatic descriptors investigated in this study.

2.3 Features extraction and selection and machine learning algorithms

The problem to solve implies a classification of skin lesions, based on chromatic descriptors. The k-th input data for either the training/learning or the classification/test step can be denoted as a pair containing a vector $X(\mathrm{k}) \in R^d$ of chromatic features values and its correspondent class label $C(\mathrm{k}) \in \{0,1\}$, where 1 stands for melanoma and 0 for benign lesions.

The set of features to be examined are denoted by:

$$F = \{\mathrm{f}_1, .., \mathrm{f}_\mathrm{d}\} \quad (6)$$

The training set contains N_{tr} pairs and is given by:

$$S_{tr} = \{(\mathrm{X(i)}, \mathrm{C(i)}) \mid \mathrm{i} = \overline{1, \mathrm{N_{tr}}}\} \quad (7)$$

while the test set is given by :

$$S_{te} = \{(\mathrm{X(i)}, \mathrm{C(i)}) \mid \mathrm{i} = \overline{1, \mathrm{N}_{te}}\} \quad (8)$$

where N_{te} is the number of pairs in S_{te}.

In this case, we are dealing with a large number -*d*- of chromatic features *F*. If we consider all channels of all proposed color spaces for q=1, d is larger than 4000.

We can diminish the number of bins for each histogram by 2, 4, 8, etc. in order to lower the number of features in *F*, but the number is still too large so we need a feature extraction or feature selection method to perform dimensionality reduction (Khalid et al., 2014). This improves overall learning performances and lowers computational complexity.

Feature selection methods preserve the features meaning and select a sub-set of lower dimension d_S of the most reliable and relevant features -F_s- to be used for classification. The features in *F* can be ranked by using different evaluation criteria/ ranking evaluation functions.

Feature extraction methods produce a new feature space, of lower dimension - d_e:

$$F_e = \{\mathrm{f}'_1, .., \mathrm{f}'_{d_e}\}, \mathrm{d}_e < \mathrm{d} \quad (9)$$

The new features are combinations of the original features and their initial meaning is lost. Principal component analysis (PCA) and linear discriminant analysis (LDA) are two widespread methods for features extraction. They both extract features by projecting the original features vectors into a new feature space. The difference between them is in the way the linear transformation matrix is computed.

PCA (Lindsay et al., 2002) generates a transformation matrix based on the largest variations in the initial feature space, while LDA takes into account the largest ratio of between-class variation and within-class variation when projecting the initial feature to a reduced space.

On the set of extracted or selected chromatic features a series of classification algorithms can be applied. In this paper we consider the following two: support vector machines (SVM) and K nearest neighbor (KNN). In our case, the classification algorithm output is categorical: 0 (benign) and 1 (melanoma).

KNN (Bhatiaet al., 2010) is a simple algorithm based on the idea of storing all available training points and classifying new entries using a similarity measure (e.g. distance functions). A new point is classified using the majority vote of its neighbors, and it is assigned to the class most common amongst its K nearest neighbors measured by a distance function.

SVM (Byun et al., 2002) is a large margin classifier. This method identifies the hyperplane(s) such that the nearest training point position is the furthest from the considered hyperplane(s). The support vectors are the training points that are the closest to the hyperplane(s).

We measure the classifiers' performances in terms of sensibility (*Ss*) and specificity (*Sp*) in detecting melanoma:

$$Ss = \frac{TP}{TP+FN}\cdot 100$$
$$Sp = \frac{TN}{TN+FP}\cdot 100 \quad (10)$$

where: TP (true positive) stands for the number of melanomas correctly identified, FN (false negative) is the number of melanomas that are not is correctly identified, TN (true negative) is the number of benign lesion that are correctly identified and FP (false positive) the number of benign lesions that are not correctly identified as normal.

We are interested in both high sensibility and specificity.

It is known that the classification features number (d_e or d_s) influence the classification results so in the next section we are going to test the performances varying the number of features used for the learning and classification phases. We are interested to determine the classifier with the best *Ss* and *Sp* for a range of features number.

3. RESULTS, DISCUSSION AND PROPOSED CLASSIFICATION METHODOLOGY

3.1 Statistical results for different color channels and models, classification features number and classifiers

For this study, 604 images were acquired: 342 contain normal nevi (benign), 124 dyspastic nevi (benign) and 138 melanomas. The lesions' type is confirmed by the histopathological results. Considering the system proposed by the American Joint Committee on Cancer (AJCC) for The Melanoma of the Skin Staging, 37 images present clinical Stage 0, Tis primary tumours (melanomas in situ). This being the earliest stage in the development of the tumour, detection is of paramount importance.

The images are divided into a training set and a test set each containing 233 images of benign lesions and 69 images of melanomas. In the test set contains 19 images of melanomas in situ.

The normalised histograms for each color channel of each color model, for different quantisation levels (q=1,2,4,8,16) are computed. For q=4 the best results were obtained. They are presented and discussed below.

Two data dimensionality reduction methods: feature selection and features extraction methods are used and compared. For feature selection we use t-test and for feature extraction we apply PCA.

For our study we consider a linear SVM algorithm and a KNN algorithm with the Euclidian distance as similarity measure.

We are interested to identify the color channels/ spaces that can be used for melanoma identification with best results. To that end, we compute the *Ss* and *Sp* for each channel and color space when using feature selection and feature extraction methods combined with a KNN and respectively SVM classifier.

We vary the number of selected/extracted features. We observe the variations in *Ss* and *Sp* with the features number (Fig. 3). We consider that the appropriate features number is d_e or d_s such that *Ss* and *Sp* maintain approximately the same values in the interval $[d_e\text{-}\Delta,\ d_e\text{+}\Delta]$ or, respectively, $[d_s\text{-}\Delta,\ d_s\text{+}\Delta]$, where Δ large enough.

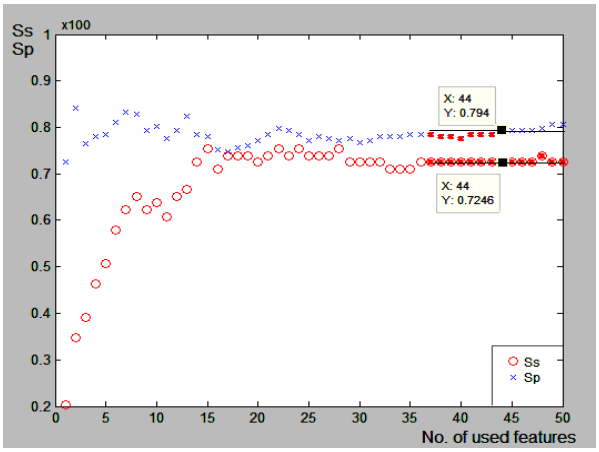


Fig. 3. Sensibility and specificity function of the features number used for classification.

The *Ss* and *Sp* for all combinations of dimensionality reduction methods and classifications methods, for all color spaces and, respectively, each of their channels are presented in tables 1-4.

Table 1. Ss and Sp for each channel when using feature selection for dimensionality reduction and KNN and, respectively, SVM as classification methods

Color Channel	KNN		SVM	
	Ss	Sp(d_s)	Ss	Sp(d_s)
H	41	80(30)	42	78 (34)
S	38	82(42)	45	77(35)
V	40	80(42)	44	84(29)
R	40	81(37)	50	72(23)
G	40	79(35)	55	75(26)
B	47	83(36)	46	78(31)
O1	45	83(33)	60	74(35)
O2	46	80(35)	55	69(30)
O3	42	79(36)	47	70(40)
Y	41	77(27)	65	72(20)
Cb	39	84(35)	49	76(25)
Cr	43	85(37)	47	75(23)
R_N	36	80(25)	56	72 (35)
G_N	36	82(25)	59	81(27)
R'	36	82(27)	49	78(30)
G'	38	78(27)	61	72(32)
B'	39	82(27)	57	77(32)
R_I	42	80(37)	53	77(24)
G_I	38	77(35)	59	78(25)
B_I	47	84(34)	67	73(21)

Table 2. Ss and Sp for each color space when using feature selection for dimensionality reduction and KNN and, respectively, SVM as classification methods

Color space	KNN		SVM	
	Ss	**Sp(d_s)**	**Ss**	**Sp(d_s)**
HSV	65	75(35)	68	76(38)
RGB	43	88(37)	68	75(40)
Opponent	61	77(25)	73	72(35)
YcbCr	66	83(23)	82	76(45)
RGB_N	36	80(40)	68	80(37)
RGB'	31	76(40)	73	78(45)
RGB_l	43	88(37)	67	75(42)

Table 3. Ss and Sp for each channel when using feature extraction for dimensionality reduction and KNN and, respectively, SVM as classification methods

Color Channel	KNN		SVM	
	Ss	**Sp(d_e)**	**Ss**	**Sp(d_e)**
H	38	81(20)	70	78(33)
S	40	71(30)	55	75(30)
V	61	91(37)	57	87(32)
R	62	91(27)	68	87(23)
G	70	74(25)	80	52(30)
B	65	72(20)	85	67(35)
O1	60	80(23)	60	78(18)
O2	46	73(20)	70	48(37)
O3	30	76(15)	53	63(25)
Y	65	80(20)	56	75(45)
Cb	58	79(25)	47	71(26)
Cr	56	83(30)	60	80(35)
R_N	30	82(10)	20	98(13)
G_N	43	80(10)	10	92(15)
R'	28	69(10)	30	90(35)
G'	26	70(10)	28	78(18)
B'	28	64(10)	35	72(14)
R_l	62	91(25)	68	87(22)
G_l	74	70(28)	75	55(37)
B_l	65	72(20)	80	68(35)

Table 4. Ss and Sp for each color space when using feature extraction for dimensionality reduction and KNN and, respectively, SVM as classification methods

Color space	KNN		SVM	
	Ss	**Sp(d_e)**	**Ss**	**Sp(d_e)**
HSV	52	84(33)	70	85(33)
RGB	72	82(25)	82	78(45)
Opponent	54	72(20)	61	80(25)
YcbCr	63	78(40)	70	83(33)
RGB_N	50	75(10)	61	74(15)
RGB'	29	72(10)	60	70(15)
RGB_l	76	83(37)	82	75(45)

3.2 Discussion

On the quantisation level: For low quantisation levels (eg, q=1 or 2) the noise influences on the histogram has greater impact compared to the case of higher quantisation levels. Also, neighbour bins refer to colors that are not substantially different.

For high quantisation levels (eg, q=8 or 16) the information contained on each color channel can be dramatically lost, thus leading to poor results.

As it can be observed from Fig. 4, one cannot find a stable enough domain in terms of features number such that *Ss* and *Sp* have approximately the same value.

On the color spaces and color channels: The poorest results were obtained for the RGB' and RGB_N color spaces, followed closely by Opponent color space. However, O1 color channel leads to good results.

On the dimensionality reduction method: Fracture extraction (PCA) leads to better results than feature selection (t-test), both in terms of classification results and classification stability function of features number for both classifiers.

On the classification method: Overall, in terms of specificity and sensitivity, SVM performs better than KNN, independent to the dimensionality reduction method applied to the features set.

Moreover, in the case of SVM, the domain for which *Ss* and *Sp* maintain approximately constant values tends to be more pronounced than in the case of KNN.

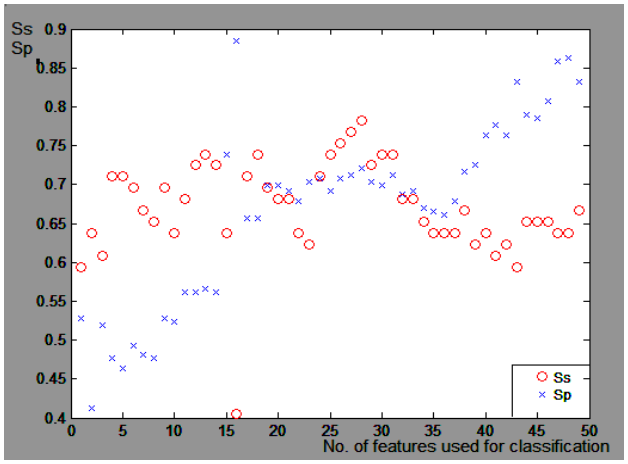


Fig. 4. Sensibility and specificity function of the features number used for classification for a to small quantisation value (q=2).

3.3 Proposed classification methodology and obtained results

Taking into consideration all these facts, we consider for analysis, as features set, the normalised histograms for the RGB, HSV, RGB_l and Ycbcr color models, and, respectively, for the Opp1 channel. The normalised histograms are computed for q=4. Data dimensionality reduction is

performed using PCA, SVM is used as classification method and we investigate the classification results (*Ss* and *Sp*) function of the number of extracted features. The results are presented in Fig.4.

For the following extracted features domain: [30,47], we expect that the sensitivity is at least 78% and the specificity 82% (Fig 5). The best obtained results are: 83% sensitivity and 86% specificity, when considering 36-39 extracted features for classification.

Considering the extracted features domain: [30, 47], out of 19 images of melanomas in situ, only one was misclassified as benign. This shows that the proposed strategy is indeed viable for early detection of this deadly type of cancer.

The obtained sensibility and specificity surpass by far the general practitioner sensitivity and specificity and approach the accuracy of the dermatologist clinical diagnosis.

This is a valuable result considering that this level of accuracy is obtained using only the color information that we have on the imaged lesion.

By combining this information on the lesion's color with a series of texture and shape descriptors, we are confident to attain superior classification results.

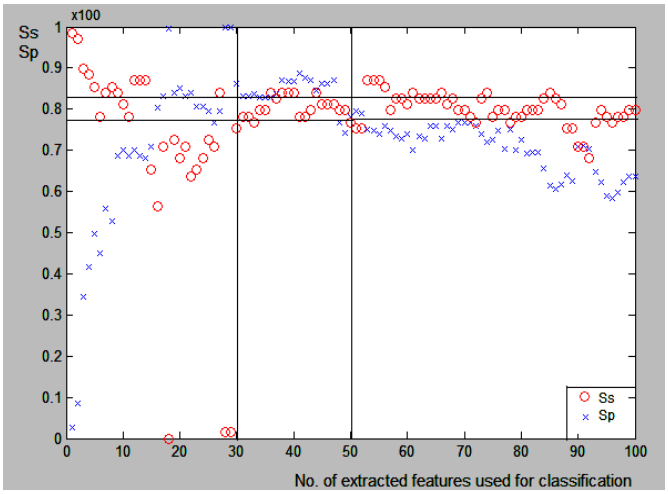


Fig. 5. Sensibility and specificity function of the features number when considering the normalised histograms for the RGB, HSV, RGB_l and Ycbcr, and O1 channel, a quantisation level q=4 and using PCA for dimensionality reduction and SVM for classification.

4. CONCLUSIONS

This article discusses the performances of different color spaces and classification algorithms in terms of offering an accurate classification of benign lesions and melanoma based on chromatic descriptors.

The proposed methodology starts with a fast algorithm for lesion extraction, then the normalised histograms for the RGB, HSV, RGB_l and Ycbcr color spaces, and O1 color channel of the Opponent color space are computed for the extracted lesion. These give the features set associated to a case. Dimensionality reduction is performed by a feature extraction method (PCA). The classification algorithm is SVM.

The proposed strategy has an overall sensitivity and specificity in detecting melanoma of at least 78% and, respectively, of 82% (considering 30 to 47 extracted features).

Moreover, in this study only, out of 19 images of melanomas in situ (incipient stage) just one was misclassified. This is a very important result because the most stringent problem is recognising this type of cancer early in its development, when mortality rate is low and curative measures can be taken.

By using a larger training set and by including the chromatic descriptors along with shape and texture descriptors for classification, we expect to obtain higher sensitivity and specificity values.

At this point our goal is to detect melanoma, especially melanoma in situ (earliest stage). Future work will establish if we can extend the classification method based on chromatic descriptors to classify melanoma accordingly to the clinical staging system proposed by AJCC.

ACKNOWLEDGEMENT

The work of Andreea Udrea was supported by Project SOP HRD - PERFORM /159/1.5/S/138963.

REFERENCES

Alcón J.F., Ciuhu C., ten Kate W., Heinrich A., Uzunbajakava N., Krekels G., Siem D., and de Haan G. (2009). Automatic Imaging System With Decision Support for Inspection of Pigmented Skin Lesions and Melanoma Diagnosis. *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1

Bhatia N. et al, (2010) Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security*, Vol. 8, No. 2

Byun H. and Lee S.W. (2002). Applications of Support Vector Machines for Pattern Recognition: A Survey SVM 2002, Springer-Verlag Berlin Heidelberg, pp. 213-236

Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V (2009). Lesion border detection in dermatoscopy images. *Computerized Medical Imaging and Graphics* 33(2), pp. 148-153

Chang, Y., Stanley, R. J., Moss, R. H., & Van Stoecker, W. (2005). A systematic heuristic approach for feature selection for melanoma discrimination using clinical images. *Skin Research and Technology : Official Journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging* (ISSI), 11(3), 165–178. doi:10.1111/j.1600-0846.2005.00116.x

Cheng B., Erdos D., Stanley R., Stoecker W., Calcara D. and Gómez D. (2011). Automatic detection of basal cell carcinoma using telangiectasia analysis in dermoscopy

skin lesion images, *Skin Research and Technology*; 17(3), pp. 278–287

Dalal, A., Moss, R.H., Stanley, R.J., Stoecker, W.V., Gupta, K., Calcara, D.A., Xu, J., Shrestha, B., Drugge, R., Malters, J.M., Perry, L.A. (2011). Concentric decile segmentation of white and hypopigmented areas in dermatoscopy images of skin lesions allows discrimination of malignant melanoma. *Computerized Medical Imaging and Graphics* 35(2), pp. 148 – 154

Gilmore S., Hofmann-Wellenhof R., Soyer P. (2010). A support vector machine for decision support in melanoma recognition. *Experimental Dermatology*, vol. 19, Issue 9, pp. 830–835

Hamilton A.D., Brady R.R., (2012). Medical professional involvement in smartphone 'apps' in dermatology, *British Journal of Dermatology*, vol.167, pp 220-221

Khalid, S. , Khalil, T., Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. *Science and Information Conference (SAI)*, pp. 372 - 378

Li, X., Aldridge, B., Ballerini, L., Fisher, R., Rees, J.(2009). Depth data improves skin lesion segmentation. *Proc. 12th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, London, pp. 1100-1107

Lindsay I Smith. (2002). A tutorial on principal components analysis. Cornell University, USA

Little E. G., Eide M. J. (2012). Update on the current state of melanoma incidence. *Dermatol Clin.*, vol.30, pp 355-361

Lee, T.K., Claridge, E. (2005). Predictive power of irregular border shapes for malignant melanomas. *Skin Research and Technology* 11(1), pp. 1-8

Maglogiannis, I., Doukas, C.N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine* 13(5), pp. 721-733

Maier T., Kulichova D., Schotten K., Astrid R., Ruzicka T., Berking C., Udrea A., (2015). Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result, *Journal of the European Academy of Dermatology and Venereology,* Volume 29, Issue 4, pp. 663–667

Mete, M., Kockara, S., Aydin, K.(2011). Fast density-based lesion detection in dermatoscopy images. *Computerized Medical Imaging and Graphics* 35(2), pp. 128- 136

Otsu N. (1979). A threshold selection method from gray-level histograms, *IEEE Trans. Sys., Man., Cyber.*, vol.9, pp. 62–66

Piantanelli A., Maponi P., Scalise L., Serresi S., Cialabrini A., Basso A., (2005).Fractal characterisation of boundary irregularity in skin pigmented lesions. *Med Biol Eng Comput.*, Volume 43, Issue 4, pp.436-42.

Premaladha J., Sujitha S., Lakshmi Priya M. and Ravichandran K.S. (2014). A Survey on Melanoma Diagnosis using Image Processing and Soft Computing Techniques. *Research Journal of Information Technology*, 6, pp. 65-80

Ramezani, M., Karimian, A., & Moallem, P. (2014). Automatic Detection of Malignant Melanoma using Macroscopic Images. *Journal of Medical Signals and Sensors*, *4*(4), 281–290.

Robson Y., Blackford S.,Roberts D., (2012). Caution in melanoma risk analysis with smartphone application technology, *Br J Dermatol*.,vol. 167, pp. 703-704

Schaefer, G., Rajab, M.I., Celebi, M.E., Iyatomi, H. (2011). Colour and contrast enhancement for improved skin lesion segmentation. *Computerized Medical Imaging and Graphics* 35(2), pp. 99 – 104

Stoecker W., Gupta K., Shrestha B., Wronkiewiecz M., Chowdhury R., Stanley J., Xu J., Moss R., Celebi M.E., Rabinovitz H., Oliviero M., Malters J., and Kolm I. (2009). Detection of Basal Cell Carcinoma Using Color and Histogram Measures of Semitranslucent Areas. *Skin Res Technol.*; 15(3), pp. 283–287

Stoecker, W., Wronkiewiecz, M., Chowdhury, R., Stanley, R.J., Xu, J., Bangert, A., Shrestha, B., Calcara, D.A., Rabinovitz, H.S., Oliviero, M., Ahmed, F., Perry, L.A., Drugge, R. (2011). Detection of granularity in dermatoscopy images of malignant melanoma using color and texture features. *Computerized Medical Imaging and Graphics* 35(2), pp. 144 -147

Udrea A., Lupu C. (2014). Real-time acquisition of quality verified nonstandardized color images for skin lesions risk assessment – a preliminary study. *18th International Conference on System Theory, Control and Computing*, pp. 199-204

Udrea A. (2015). Pigmented nevi risk assessment based on the correlation dimension of the associated lesion's attractor, *Proceedings of the 18th Conference on Control Systems and Computers Science*

Udrea A., Popescu D., Miron C. (2015). An analysis on the reliability of a series of texture and shape descriptors for melanoma diagnosis, *Bull. UPB, C series*

Wang, H., Moss, R.H., Chen, X., Stanley, R.J., Stoecker, W.V., Celebi, M.E., Malters, J.M., Grichnik, J.M., Marghoob, A.A., Rabinovitz, H.S., Menzies, S.W., Szalapski, T.M. (2011). Modified watershed technique and post-processing for segmentation of skin lesions in dermatoscopy images. *Computerized Medical Imaging and Graphics* 35(2), pp. 116- 120

Weinstock M. A. (2012). Reducing Death from Melanoma and Standards of Evidence. *Journal of Investigative Dermatology*, vol. 132**,** pp. 1311–1312

Wolf JA, Moreau JF, Akilov O, et al. (2012). Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA Dermatol*.;149, pp. 422-426

Zhou, H., Schaefer, G., Celebi, M.E., Lin, F., Liu, T. (2011). Gradient vector flow with mean shift for skin lesion segmentation. *Computerized Medical Imaging and Graphics* 35(2), pp. 121-127