Texture and color feature extraction for classification of melanoma using SVM

Conference Paper · January 2016

DOI: 10.1109/ICCTIDE.2016.7725347

2 authors, including:

Kavitha Jc
Reva University

TEXTURE AND COLOR FEATURE EXTRACTION FOR CLASSIFICATION OF MELANOMA USING SVM

J.C.Kavitha
Research Scholar
Department of Computer Science & Engineering
Manonmaniam Sundaranar university
Tirunelveli - 627012
jck_kavitha@yahoo.co.in

Suruliandi.A
Professor
Department of Computer Science & Engineering
Manonmaniam Sundaranar university
Tirunelveli - 627012
suruliandi@yahoo.com

Abstract — **Feature plays a vital in the domain of image processing. The various features of an image are color, texture, shape or domain specific features. Texture is considered as one of the main feature of any image. The second order statistical features for an image is obtained by Gray level co-occurrence matrix (GLCM) and it operates on spatial domain. The haralick texture features are energy, entropy, homogeneity, correlation, contrast, dissimilarity and maximum probability. The aim of the paper is to classify the dermoscopy images into melanoma and non-melanoma by considering the texture and color features of an image. GLCM is used to extract the texture features of an image. Color histograms are used to extract the color features in three color spaces namely RGB, HSV and OPP. Support vector machine (SVM) is used for the process of classification. The performance of the proposed system is evaluated by the metrics sensitivity and specificity. The experimental result shows that the texture combined with RGB color space provides better classification accuracy.**

***Keywords*—Texture feature extraction, Gray level co-occurrence matrix (GLCM), Support vector machine (SVM), Color features.**

I. INTRODUCTION

Melanoma is one among the most malignant, metastatic and dangerous form of skin cancer. It is responsible for the majority of deaths related to skin cancer. The curability and survival rate of a patient can be increased if melanoma can be detected at an earlier stage. The skin lesions are diagnosed by dermatologists using a technique known as dermoscopy. This is a non-invasive procedure used for the in vivo observation of skin lesions. The skin cancers are classified into melanoma, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). Melanoma is defined to be a condition or a disorder that affects the melanocyte cells thereby increasing the synthesis of melanin [1]. The disorder is characterized by skin lesion development and it varies in shape, size, color and texture. The color and texture features are considered to be important in the detection of melanoma.

Texture is an important feature that identifies the object present in any image. The texture is defined by the spatial distribution of pixels in the neighbourhood of an image. The gray level spatial dependency is represented by a two dimensional matrix known as GLCM and it is used for texture analysis. The GLCM matrix specifies that how often the pairs of pixels with certain values occur in an image. The statistical measures are then derived using the GLCM matrix. The textural features represent the spatial distribution of gray tonal variations within a specified area. In images, the neighbouring pixel is correlated and spatial values are obtained by the redundancy between the neighbouring pixel values. The color features are represented by color histograms in six color spaces namely RGB, HSV, LAB, CIE, HUE and OPP.

Support Vector Machine (SVM) is defined as a supervised learning model. It is used for the purpose of classification and regression analysis. It analyse and recognise the image and data patterns used in image processing. The advantage of SVM is that in high dimensional spaces it works effectively and since it uses a subset of training points in the decision function it is considered as memory efficient. Different kernel functions can be used for decision function.

A. Related Works

There are several systems for the identification of melanoma in dermoscopy images. The classification of skin lesion is done in the global region extracted from the dermoscopy image. In the global method, the process of segmentation is done using simple adaptive thresholding algorihm. GLCM matrix is used for extracting the texture features in four different orientation angles. The color features are computed by color histograms in six color spaces. The classification is done using SVM-RBF for the detection of melanoma. In this system, the experimental results shows that the color feature outperforms the texture features [2].

In ELM or dermoscopy images, segmentation process is done using thresholding operation combined with color clustering. Dynamic thresholding provides good segmentation results. A set of features namely shape and color is used to describe the malignancy of a lesion. Feature selection is done by Sequential forward selection algorithm. The classification is done using KNN classifier and it classifies the dermoscopy image into beningn or malignant [3].

In this automated approach of skin lesion identification, the segmentation process is done using laplacian filter and zero crossing algorithm. The texture features are extracted by computing the variables mean, contrast, entropy and fractality. The mean of the color space RGB inside and outside the lesion is calculated to detect whether the person is affected bt melanoma or not. ANN and Feed Forward ANN classifier is used to classify melanoma from the dermoscopy images [4].

An internet based melanoma screening system was proposed in which the server is opened for the public to upload the dermsocopy images. In this system, the digital dermoscopic image can be uploaded by the visitor and can register the clinical and pathological data. Once the image is accepted by the server, the tumor area is automatically extracted from the surrounding skin using automatic threshold decision algorithm. The parameters such as asymmetry, differential structures and color in RGB color space were calculated for the diagnosis and evaluation of lesion as melanoma or non-melanoma. The classification is done using linear discriminant analysis and ANN and the diagnostic result is given back to the to the client computer [5].

A novel method was proposed to classify the dermoscopy images. In this method, automatic border detection was done using JSEG algorithm to extract the lesion. The preprocessing of an image is done using color median filter and color reduction is done using variance based quantization. The texture features are extracted using Euclidean distance transform algorithm and GLCM. The color features are computed using color histograms in six color spaces. The classification of images is done using SVM classifier [6].

A real time automated approach using smart phone is used for the analysis of skin lesion. This system was proposed for the detection and prevention of melanoma. In this system, shape, texture and color features were taken into consideration. The smart phone captures the images of skin lesions and it automatically analyses whether the lesion is malignant or not. A wide research has shown that melanoma lesions are usually asymmetrical. It represents the lesions may have irregular borders with notched edges.

In this, segmentation is done using Gaussian filter and Otsu's method is used to compute global threshold. Feature extraction is done using 2D-Fast Fourier Transform, 2D-Discrete cosine Transform, Pigment network feature and color. The classification is done using SVM – RBF for the detection of melanoma [7].

In this paper, the textural features are considered for classifying the image. These textural features are calculated in the spatial domain and a set of gray tone spatial dependency matrix was computed. The textural features are computed using GLCM matrix in four different orientation angles. The textural features are based on the fact that describes how the gray tone appears in a spatial relationship to another [8].

B. Motivation and justification

There were many texture feature extraction methods explained above and of which GLCM texture feature extraction is widely used in the texture feature extraction from dermoscopy images [2][6][8]. The skin regions affected by melanoma are predominated by texture and color. In some cases texture may dominate the color and in some cases color may dominate the texture. Therefore it is essential to extract the texture and color separately. The color feature is computed using color histograms in six color spaces namely RGB, HSV, LAB, CIE, HUE and OPP [12]. SVM is used for the classification of the melanoma from dermoscopy images. Motivated by these facts, in the proposed system the texture and color features were extracted.

There are 14 textural features explained in GLCM and of which 8 features performs the best and it has been taken into consideration for the proposed system. Since SVM is a binary classifier and it is used for the identification of melanoma or non-melanoma from dermoscopic images. Justified by these facts, the classification of melanoma in dermoscopic images is done by combining the texture and color using SVM classifier.

C. Outline of the proposed approach

The proposed system deals with feature extraction using texture and color. The block diagram for texture feature extraction and color is shown in the below Fig.1.

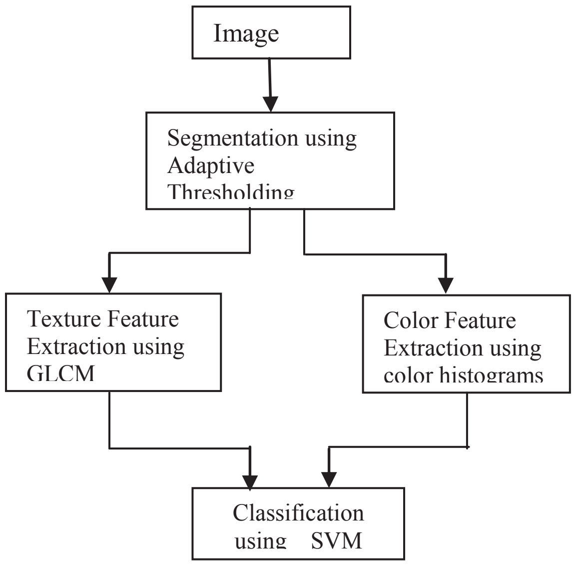

Fig.1 Feature Extraction and classification

In this, using segmentation process the skin lesion is extracted from the dermoscopy image by simple adaptive thresholding algorithm. Then the features such as texture and color of the segmented image are extracted. The texture features are extracted using GLCM in different orientation angles. The color features are computed using color histograms in three color spaces namely RGB, HSV and OPP. The SVM classifier is used for the classification of melanoma from dermoscopy images.

D .Organisation of the paper

The second section presents the texture feature extraction using GLCM matrix. Color feature extraction using color histograms is also explained in this chapter. The third section explains the classification algorithm using SVM. The fourth section presents a detailed description of the experiments conducted on dermoscopy images. The proposed system performance is also analysed in this section. The final section presents the results of the experiments and thus gives the conclusion.

II. TEXTURE FEATURE EXTRACTION

A. GRAY LEVEL CO-OCCURENCE MATRIX (GLCM)

In statistical texture analysis, from the distribution of intensities the texture features are obtained at specified position relative to one another in an image. The statistics of texture are classified into first order, second order and higher order statistics. The method of extracting second order statistical texture features is done using Gray Level Co-occurence Matrix (GLCM). First order texture measure is not related to pixel neighbour relationships and it is calculated from the original image. GLCM considers the relation between two pixels at a time, called reference pixel and a neighbour pixel. A GLCM is defined by a matrix in which the number of rows and columns are equal to the number of gray levels G in an image. The matrix element P $(i, j \mid \Delta x, \Delta y)$ is the relative frequency where i and j represents the intensity and both are separated by a pixel distance Δx, Δy. The different textural features such as energy, entropy, contrast, homogenity, correlation, dissimilarity, inverse difference moment and maximum probability can be computed using GLCM matrix.

B. COLOR FEATURE EXTRACTION

Color is considered as another important feature descriptor for the classification of melanoma. If the particular region is affected the skin lesions region changes the color effectively. Relative color histograms in different color spaces are constructed to identify the melanoma. The 3-D histogram is constructed for the color spaces such as RGB, LAB, HSV, HUE and OPP [12]. RGB color space represents a mixture of Red, Green and Blue. The color component is represented by the mixture coefficients of these three colors. The drawbacks of the color spaces are not perceptually uniform, and it depends on the acquisition setup. It provides a high correlation among these three color channels. Different color representations have been proposed to overcome these drawbacks. e.g: the edges are strengthened in biologically inspired color spaces such as the opponent color space (OPP), hue saturation and brightness (HSV and HIS) are the color spaces related to human description of color, CIE La*b* and L*uv are perceptually uniform color spaces. These two color spaces are device independent. For these histograms 8x8x8 = 512 color bins are generated and it is considered as a single feature vector. This feature vector is given as an input to the SVM classifier to calculate the sensitivity and specificity.

III. Support Vector Machine (SVM) Algorithm

Support vector machines [16] are considered as the supervised learning models. It is a kind of learning algorithm which is used to analyze the data and recognize the data patterns. This algorithm is used for the purpose of classification. It is mainly applicable for solving the binary problems.

IV. EXPERIMENTS AND RESULTS

A. Experimental Data

The experiments are carried out using 150 dermoscopy images which comprises of training set and testing set. The sample contains 100 training set images and 50 testing set images. The training set images are trained using SVM classifier for the identification of melanoma. The experiments were carried out on the testing and training image dat set. The experiments are performed to extract the texture features by using GLCM and color features by using color histograms.

The input data images are given in the below Fig.2.

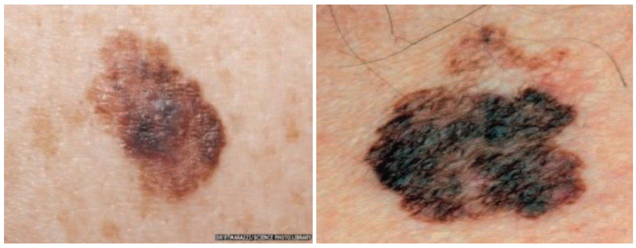

Fig.2.Sample Dermoscopy Images

B. Experimental Results

The dermoscopy images are taken as an input and the skin lesion is extracted from the image using Simple Adaptive Thresholding Algorithm. Then the different textural features are calculated in four different orientation angles. This is considered as the statistical nature of texture. The textural information of an image is defined by the spatial relationship. A gray tone spatial dependency matrix is computed for any given image and 8 textural features are calculated from this matrix. The textural features extracted by GLCM is given in the below Table.1.

Table 1.Results of Texture and Color

		Melanoma					Non-Melanoma				
					min	max				min	max
Orientation = 0	Contrast	0.49	0.49	0.44	0.44	0.49	0.47	0.39	0.27	0.27	0.47
	Correlation	0.97	0.97	0.97	0.97	0.97	0.98	0.98	0.99	0.98	0.99
	Dissimilarity	0.19	0.16	0.21	0.16	0.21	0.10	0.06	0.09	0.06	0.10
	Energy	0.32	0.40	0.40	0.32	0.40	0.44	0.51	0.49	0.44	0.51
	Entropy	1.76	1.48	1.62	1.48	1.76	1.09	0.75	0.94	0.75	1.09
	Homogeneity	0.93	0.94	0.92	0.92	0.94	0.97	0.99	0.97	0.97	0.99
	Maximum probability	0.53	0.60	0.62	0.53	0.62	0.59	0.59	0.63	0.59	0.63
	Inverse difference moment	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00
Orientation = 45	Contrast	8.03	9.14	7.13	7.13	9.14	7.06	7.97	4.25	4.25	7.97
	Correlation	0.58	0.54	0.50	0.50	0.58	0.69	0.67	0.78	0.67	0.78
	Dissimilarity	1.49	1.63	1.49	1.49	1.63	1.13	1.15	0.67	0.67	1.15
	Energy	0.19	0.23	0.23	0.19	0.23	0.30	0.36	0.44	0.30	0.44
	Entropy	2.49	2.27	2.36	2.27	2.49	1.56	1.18	1.19	1.18	1.56
	Homogeneity	0.71	0.70	0.70	0.70	0.71	0.82	0.85	0.89	0.82	0.89
	Maximum probability	0.40	0.46	0.46	0.40	0.46	0.47	0.47	0.61	0.47	0.61
	Inverse difference moment	0.92	0.91	0.92	0.91	0.92	0.93	0.93	0.96	0.93	0.96
Orientation = 90	Contrast	15.06	16.86	13.84	13.84	16.86	15.02	17.01	8.35	8.35	17.01
	Correlation	0.20	0.16	0.06	0.06	0.20	0.35	0.30	0.53	0.30	0.53
	Dissimilarity	2.66	2.87	2.71	2.66	2.87	2.32	2.44	1.26	1.26	2.44
	Energy	0.10	0.13	0.12	0.10	0.13	0.21	0.27	0.40	0.21	0.40
	Entropy	2.88	2.66	2.73	2.66	2.88	1.79	1.39	1.31	1.31	1.79
	Homogeneity	0.54	0.53	0.51	0.51	0.54	0.66	0.69	0.82	0.66	0.82
	Maximum probability	0.24	0.29	0.27	0.24	0.29	0.33	0.33	0.60	0.33	0.60
	Inverse difference moment	0.85	0.84	0.85	0.84	0.85	0.86	0.85	0.93	0.85	0.93
Orientation = 135	Contrast	14.64	17.11	14.93	14.64	17.11	17.62	19.57	8.78	8.78	19.57
	Correlation	0.21	0.11	-0.05	-0.05	0.21	0.24	0.20	0.51	0.20	0.51
	Dissimilarity	2.70	3.00	2.95	2.70	3.00	2.68	2.80	1.33	1.33	2.80
	Energy	0.10	0.12	0.12	0.10	0.12	0.21	0.26	0.38	0.21	0.38
	Entropy	2.92	2.63	2.61	2.61	2.92	1.75	1.41	1.34	1.34	1.75
	Homogeneity	0.52	0.51	0.48	0.48	0.52	0.61	0.65	0.80	0.61	0.80
	Maximum probability	0.24	0.29	0.27	0.24	0.29	0.33	0.32	0.58	0.32	0.58
	Inverse difference moment	0.85	0.83	0.84	0.83	0.85	0.84	0.83	0.92	0.83	0.92

The features described contain information about image textural characteristics such as homogenity, contrast, correlation, energy, entropy, dissimilarity and maximum probability. These measures indicate the complexity and gray tone transition present in an image. These eight texture features are given as an input to the SVM-RBF and SVM-Linear kernel classifier. It has been shown that SVM-RBF performs better when compared to SVM-Linear kernel classifier.

The given image is converted into features descriptors using three different color spaces. The color features are extracted using color histograms in different color spaces namely RGB, HSV, OPP. For each color space 512 bins were generated and it is considered as a single feature vector. This feature vector is given as an input to the SVM- Linear kernel classifier. The different parameters such as sensitivity and specificity are calculated. The color histograms are given in the below Fig.4.(a) RGB color space descriptors (b) HSV color space descriptors (c) OPP color space descriptors.

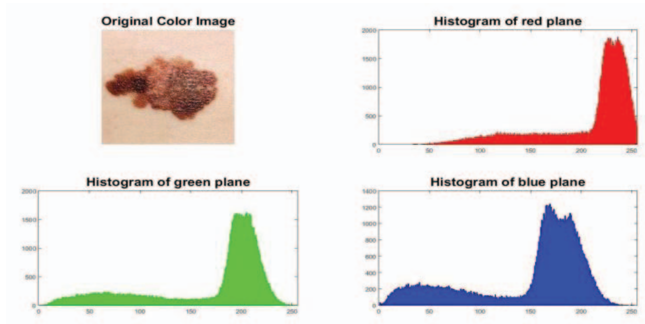

Fig.4(a) RGB color space Descriptors

Histogram Of "H" Histogram Of "S" Histogram Of V"

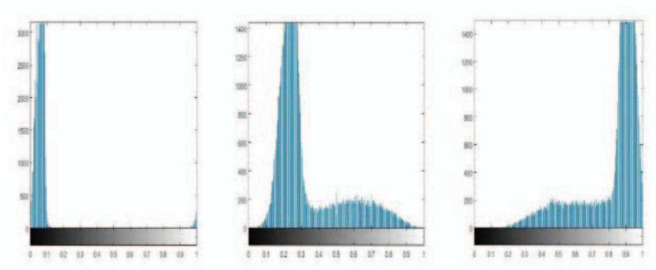

Fig.4(b) HSV color space Descriptors

Histogram of O1 Histogram of O2 Histogram of O3

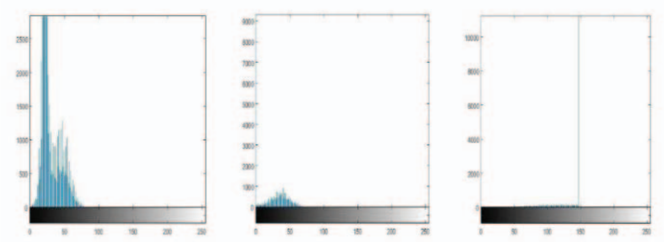

Fig.4(c) OPP color space Descriptors

C. PERFORMANCE ANALYSIS

(a).Performance Metrics

The proposed system performance is calculated on the basis of different metrics namely sensitivity and specificity. The accuracy of the proposed system is classified using SVM classifier.

Sensitivity is defined as the measure of true positive rate and it is given by

$$\text{Sensitivity} = \frac{TP}{TP+FN} \qquad (1)$$

Specificity is defined as the measure of true negative rate and it is given by

$$\text{Specificity} = \frac{TN}{TN+FP} \qquad (2)$$

where TP – True Positive, TN – True Negative, FN – False Negative, FP – False Positive.

(b).Result Analysis

The classification result shows that the accuracy of texture feature extraction is 76%. The classification accuracy of RGB and HSV color space is 92% and OPP is 89%. It has been shown that the texture feature when combined with RGB color space provides a better classification accuracy of 93.1%. The classification results are shown in the below Table 2.

TABLE.2 CLASSIFICATION RESULTS FOR THE DETECTION OF MELANOMA

FEATURE (texture+color)	SENSITIVITY	SPECIFICITY	ACCURACY
GLCM	72.9	71.1	76
RGB	86.2	100	92
HSV	86.2	100	92
OPP	84.2	95.3	89
RGB + TEXTURE	88.2	85.5	93.1

For the input image data set taken for experimental analysis, color feature outperforms the texture in our system. In our system, color dominates the texture feature and it provides better accuracy. But it may not be the same for all cases.

V. CONCLUSION

A feature extraction model using texture and color was proposed for the identification of melanoma. The image is preprocessed to increase the resolution and it is segmented using Simple Adaptive Thresholding Algorithm. Then the filtered image is subjected to feature extraction. These texture features are used to evaluate skin lesion discrimination using GLCM matrix. Histogram analysis technique is used for color feature extraction. Classifying the texture and color features is done using SVM and it identifies the melanoma from dermoscopy images. The experimental result shows that the when the texture feature is combined with RGB color space it provides a better classification results of 93%. This new texture feature provides sensitivity 88.2 % and specificity 85.5%.

REFERENCES

[1] S.Suer, S.Kockara, and M.Mete, "An improved border detection in dermoscopy images for density based clustering", BMC bioinformatics.,vol.12, No.10,p.S12,2011.

[2] Catarina Barata, Margarida Ruela, Mariana Francisco, Teresa Mendonca and Jorge S.Marques, "Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features", IEEE systems Journal, Vol.8,No.3, Sep.2014.

[3] H.Ganster, A .Pinz, E. Wilding, M. Binder and H.Kittler, "Automated melanoma recognition," IEEE Trans.Med. Imaging, Vol.20,no. 3, pp. 233-239, Mar, 2001.

[4] P. Rubegni, G. Cevenini, M. Burroni, R. Perotti, G. Dell'Eva, P. Sbano, and C. Miracco, "Automated diagnosis of pigment skin lesions," Int. J. Cancer, vol. 101, no. 6, pp. 576–580, Oct. 2002.

[5] H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K. Ogawa, "An Internet-based Melanoma Diagnostic System - Toward the Practical Application" IEEE, 2005.

[6] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi,Y. Aslandogan, W. Stoecker, and R. Moss, "A methodological approach to the classification of dermoscopy images," Science Direct, Computerized Medical Imaging and Graphics, vol. 31, no. 6, pp. 362–373, Sep. 2007.

[7] Omar A., Buket d.Barkana, Non invasive Real Time Automated skin Analysis system for Melanoma Early Detection and prevention", IEEE Journal of Translation Engineering in Health and Medicine, April 2015.

[8] Robert .M.Haralick, Shanmugam.K, Itshak Einstein, "Textural Features for Image Classification", IEEE Transactions on systems, man and cybernetics, Nov.1973.

[9] M. Sheha, M. S. Mabrouk, and A. Sharawy, "Automatic detection of melanoma skin cancer using texture analysis," Int. J. Comput. Appl., vol. 42, no. 20, pp. 22–26, Mar. 2012.

[10]Md.Amran Hossen Bhuiyan, Ibrahim Azad, " Image Processing for Skin cancer Features Extraction " , IJSER, Vol.4, Issue 2, Feb.2013.

[11] G.Di Leo, A.Paollillo, P.Sommella and C.Liguori, " An improved procedure for the automatic detection of dermoscopic structures in digital elm images of skin lesions, in Proc.2008 IEEE computer soc. Pp.190-194.

[12] W.Stoecker, M.Wronkiwwiez, R.Chowdhury, R. J. Stanley, J. Xu, A. Bangert, B. Shrestha, D.A. Calcara, H.S. Rabinovitz, M. Oliviero, F. Ahmed, L.A. Perry and R. Drugge,.

"Detection of Granularity in dermoscopy images of malignant melanoma using color and texture features", Compu.Med Imaging Graph, Vol.32.N0.8, pp 670-677, Dec.2008.

[13] M.Sneha, M. S.Mabrouck and A.Sharawy, "Automatic detection of melanoma skin cancer using texture analysis, "Int.Journal of Computer Applications, vol.42, No.20,PP.22-26,March.2012.

[14] M.E.Celebi, H.A. Kingravi, B. Uddin, H. Iyatomi, Y. Aslandogan, W. Stoecker and R. Moss., " Amethodological approach to the classification fo dermoscopy images", Compu.Med.Imaging and Graphics,Vol.31,No.6,pp.362-373, Feb.2007.

[15]X.Yuan, Z.Yang and G.Zouridakis, "SVM based texture classification and application to early melanoma detection", Proc. IEEE EMBS Int.Conf.2006, PP.4775-4778.

[16] www.supportvectormachines.org