International Journal of Medical Imaging
2015; 3(2): 34-40
Published online March 16, 2015 (http://www.sciencepublishinggroup.com/j/ijmi)
doi: 10.11648/j.ijmi.20150302.15
ISSN: 2330-8303 (Print); ISSN: 2330-832X (Online)

Categorization of Non-Melanoma Skin Lesion Diseases Using Support Vector Machine and Its Variants

Immagulate I., Vijaya M. S.

PSGR Krishnammal College for Women, Coimbatore, India

Email address:
immagulate77@gmail.com (Immagulate I.) msvijaya@grgsact.com (Vijaya M. S.)

To cite this article:
Immagulate I., Vijaya M. S.. Categorization of Non-Melanoma Skin Lesion Diseases Using Support Vector Machine and Its Variants. *International Journal of Medical Imaging.* Vol. 3, No. 2, 2015, pp. 34-40. doi: 10.11648/j.ijmi.20150302.15

Abstract: Skin cancer is the growth of uncontrolled abnormal skin cells. There are two main types of skin cancers such as Melanoma and Non-Melanoma. The main objective of this research work is to focus on Non-Melanoma skin cancers and classify the types of it.The classification of non melanoma skin cancers is automated using machine learning approach and the model is built to predict the type of disease accurately using support vector machine and its variants. Various experiments have been carried out with skin lesion images and the results are analyzed.

Keywords: Classification, Machine Learning, Prediction, Support Vector Machine, Training

1. Introduction

Skin cancer starts in the cells of the skin. The skin is the body's largest organ. It protects the human being from injury, infection, heat and ultraviolet light from the sun. The skin helps control human being body temperature and gets rid of waste materials through the sweat glands. It also makes vitamin D and stores water and fat.

Skin cancer is often categorized into either melanoma or non melanoma. Non melanoma skin cancer occurs in either basal or squamous cells. These cells are located at the base of the outer layer of the skin or cover the internal and external surfaces of the body.

Non melanoma skin cancer is the most common type of cancer. Non melanoma skin cancer is mainly divided into 5 type's namely actinic keratosis, basal cell carcinoma, melanocytic nevus, squamous cell carcinoma, seborrhoeic keratosis.

An actinic keratosis, also known as a solar keratosis, is a scaly or crusty growth. It most often appears on the bald scalp, face, ears, lips, backs of the hands and forearms, shoulders, neck or any other areas of the body frequently exposed to the sun.Basal cell carcinoma begins in the basal cells - a type of cell within the skin that produces new skin cells as old ones die off. Basal cell carcinoma often appears as a waxy bump, though it can take other forms.

A melanocytic nevus is a type of lesion that contains nevus cells, a type of melanocyte. Some sources equate the term mole with melanocytic nevus. A melanocytic nevus blemishmay be akind of lesion that contains blemish cells. Squamous cell carcinoma is an uncontrolled growth of abnormal cells arising in the squamous cells, which compose most of the skin's upper layers called the epidermis. SCCs may occur on all areas of the body including the mucous membranes and genitals, but are most common in areas frequently exposed to the sun, such as the rim of the ear, lower lip, face, bald scalp, neck, hands, arms and legs.Seborrheickeratosis is a very common harmless, usually pigmented, noncancerous growth on the skin. It usually appears as a pale, black or brown growth on the back, shoulders chest or face, but can appear anywhere on the skin.

Skin cancer is the most common form of cancer in the United States.More than 3.5 million skin cancers in over two million people are diagnosed annually.Each year there are more new cases of skin cancer than the combined incidence of cancers of the breast, prostate, lung and colon.One in five Americans will develop skin cancer in the course of a lifetime 13 million white non-Hispanics living in the US at the beginning of 2007 had at least one non melanoma skin cancer, typically diagnosed as basal cell carcinoma or squamous cell carcinoma.

Biopsy is the standard diagnosis procedure carried out in diagnosing the skin cancer. Biopsy is a sample of tissue collected from an organ or other part of the body. A biopsy

can be done by cutting or scraping a small piece of the tissue or by using a needle and syringe to remove a sample, which is then examined for abnormalities, such as cancer, by a doctor trained to look at tissue samples. This procedure of diagnosing the diseases is time consuming and may lead to inaccurate prediction of the disease. Hence, in this research work it is proposed to automate the prediction of skin lesion using machine learning approach.

Based on the study of various literatures available on non melanoma skin lesion categorization, a brief report is presented in this section about the research directions in categorization of non melanoma in the last several years.

Lucia Ballerina and Robert B. Fisher et al., [2]have used *k*-nearest-neighbor classification. Watershed segmentation was the main pre-processing steps. Color and texture features were extracted from the skin lesion images. The image database comprised of 960 lesions that belonged to 5 classes. K-nearest-neighbor classification technique gave 74% accuracy. But difficulty of this method was segmentation of skin lesion becomes difficult in the presence of shadows and bright areas caused by illumination variation.

JoséFernándezAlcón, CalinaCiuhu, Warner ten Kate, Adrienne Heinrich et al., [4] proposed automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. The system included a dedicated image processing system for feature extraction and classification, and patient-related data decision support machinery for calculating a personal risk factor. A robust segmentation algorithm had been developed.Texture feature was extracted from the image. And this method produced 81% of accuracy.

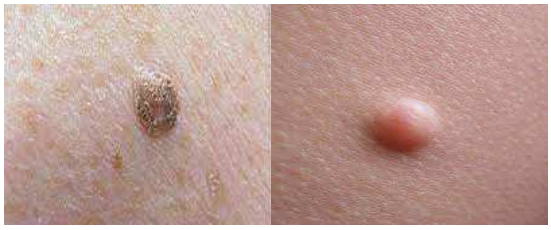
In existing research work, various researchers proposed classification technique like decision tree,K-nearest-neighbor classification. From the literature it was well appreciated that decision tree is a powerful classification technique with high generalization than other classification algorithms. In this research support vector machine and its variants such as active support vector machineand proximal support vector machine have been adopted for constructing the non melanoma skin lesion prediction model.

2.Proposed Work

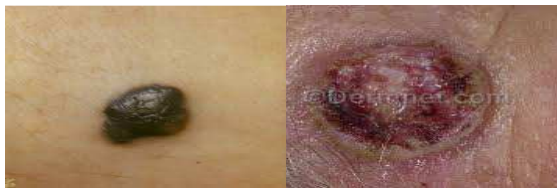
The proposed skin lesion categorization problem is modeled as multiclass classification task and offers solution using supervised learning approach.The skin lesion images related to five types of diseases such as actinic kertosis, basal cell carcinoma, melanocytic nevus, squmous cell carcinoma, sebhorreickertosis have been considered for this research and collected from two different websites namely Dermnet.com and DERMOFIT.com. The images are preprocessed using image adjustment and image resize. The color and texture features are extracted from the preprocessed skin lesion images and training dataset is developed. The training dataset is trained using multiclass support vector machine and its variants such as proximal support vector machine and active support vector machine. Various tasks carried out in modeling skin lesion prediction task are detailed below.

2.1.Data Acquisition

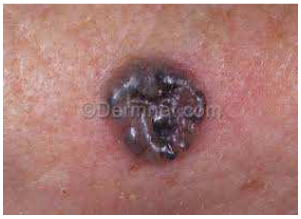
The data preparation is an important task in any machine learning activity. The images related to five types of non melanoma skin lesions have been collected from Dermnet.com and DERMOFIT.com. Five sets of images namely actinic kurtosis, basal cell carcinoma, sqmous cell carcinoma, melanocytic nevus, sebhorric kurtosis, each consisting of 100 images are taken into consideration. The sample images are shown in Fig.1.



AcitinicKertosisBasal Cell Carcinoma



Melanocytic Nevus SebhorreicKertosis



Squmous Cell Carcinoma

Fig. 1.*Sample images of non melanoma skin lesion*

2.2.Preprocessing

The aim of preprocessing is to improve the image data that suppresses unwanted distortions for further processing. Preprocessing is the process of removal of noise and transformation of the image to produce a well defined pattern for texture analysis. In this work, two steps of pre processing are performed. The first step is image resize and the second is image adjustment.

2.2.1.Image Resize

As the size of the images vary for different images, it is essential to make them uniform. The images are resized to 256*256 dimensions using interpolation method. Interpolation is the process used to estimate an image value at a location in between image pixels. When imresize function of matlab is used, the image enlarges and the output image contains more pixels than the original image.

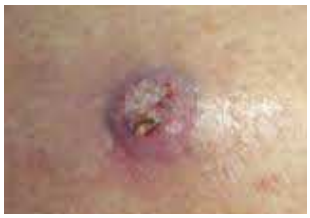


Fig. 2.*Before Preprocessed Image*

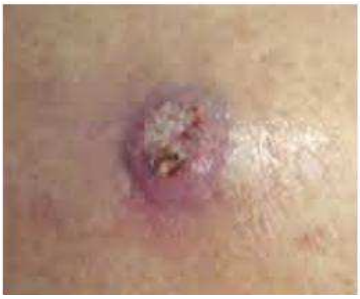


Fig.3. *Resized image*

2.2.2.Image Adjustment

This preprocessing step is mainly used to adjust the image intensity values or color map values. As color features are based on the color pixel rate, the image adjustment is required to change the color pixel values into correct and accurate format. Another function carried out in this step is to adjust the intensity of the color images. It brings down the high intensity and increases the low intensity of the images to convert them into a medium intensity images.

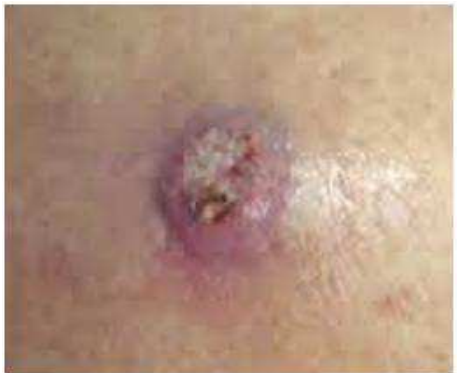


Fig.4. *Resized image*

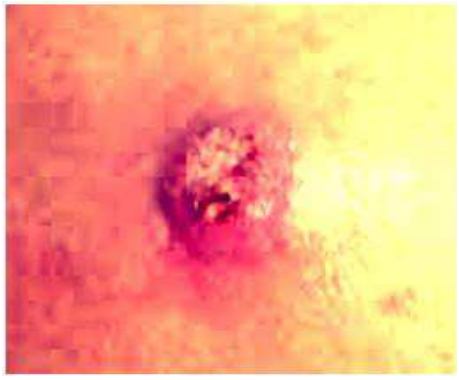


Fig.5. *After Preprocessed Image*

2.3.Feature Extraction

Feature extraction plays a vital role in data mining. It can be used to improve the classification effectiveness and computational efficiency. Drawing out specific features from the preprocessed images is called feature extraction. Feature extraction is carried out with all the preprocessed skin lesion images. The features describe the distinguishing pattern of the skin lesion images. Two types of features such as color and texture features are extracted.

2.3.1.Color Features

Color features play a decisive role in the classification of skin lesions. A color image is a combination of some basic colors. Each individual pixel of a color image is broken down into red, green and blue values.

The RGB color features such as meanR, meanG, meanB, normalizedR, normalizedG, normalized are extracted from skin lesion images. First a color image is split into three separate RGB color planes using the following functions (2).

Image_red = Image_rgb(:,1);

Image_green=Image_rgb(:,2);

Image_blue = Image_rgb(:,3);

Then the red, green and blue pixel values are calculated. Finally the mean values of the color planes R, G and B are calculated using the following functions.

mean_r=mean2(Red);

mean_g=mean2(Green);

mean_b=mean2(Blue);

Using minmax normalization, the meanR, meanG, meanB features are normalized and normalizedR, normalizedG, normalizedB features are computed. The following min max normalization is used for normalization.

$$v' = \frac{v - \min A}{\max A - \min A}(new\_\max A - new\_\min A) + new\_\min A$$

For example, consider the sample preprocessed image is shown in the Fig.3.The values of the features calculated are

Mean R = 155.42

Mean G = 119.72

Mean B = 99.95

Normalized R = 0.87

Normalized G = 0.81

Normalized B = 0.67

2.3.2.Texture Features

The texture features are extracted using Grey Level Co-occurrence Matrix (GLCM). A GLCM is a matrix where number of rows and columns is equal to number grey levels

G in an image. It is defined over an image to be the distribution of co-occurring values in the given offset. It is a way of extracting second order statistical features. It is used to measure the spatial relationships between pixels. This method is based on the belief that texture information is contained in such relationships. Few of the common statistics applied to co-occurrence probabilities are given the following table with related formulas.

Table 1.*Texture Features.*

Feature	Formula
Energy	$energy(ene) = \sum_i \sum_j g_{ij}^{\ 2}$
Entropy	$entropy(ent) = -\sum_i \sum_j g_{ij} \log_2 g_{ij}$
Contrast	$contrast(con) = \sum_i \sum_j (i-j)^2 g_{ij}$
Variance	$variance(var) = \sum_i \sum_j (i-\mu)^2 g_{ij} where \mu is the mean of g_{ij}$
Homogeneity	$homogeneity(hom) = \sum_i \sum_j \frac{1}{1+(i-j)^2} g_{ij}$
Correlation	$correlation(cor) = \frac{\sum_j \sum_j (ij) g_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} g_{ij}$
Autocorrelation	$\frac{1}{(L_x-\lvert x\rvert)(L_y-\lvert y\rvert)}$ $p(x,y) = \frac{\iint_{-\infty}^{\infty} I(u,v) I(u+x,u+v) du dv}{\frac{1}{L_x L_y} \iint_{-\infty}^{\infty} I^2(u,v) du dv}$
Sum Average	$sumaverage(sa) = \sum_{i=2}^{2N_g} i g_{x+y}(i)$
Sum Entropy	$sumentropy(se) = -\sum_{i=2}^{2N_g} i g_{x+y}(i) \log\{g_{x+y}(i)\}$
Sum Variance	$sumvariance(sv) = \sum_{i=2}^{2N_g} (i-sa)^2 g_{x+y}(i)$
Difference variance	$differencevariance = variance of g_{x-y}$
Difference Entropy	$differenceentropy(se) = -\sum_{i=0}^{N_g-1} g_{x-y}(i) \log\{g_{x-y}(i)\}$

From each preprocessed image the above features are computed using matlab code and feature vectors of size 29 is formed. The training dataset is then developed with 500 feature vectors.

3. SupervisedPattern Clasification

3.1.Support Vector Machine

Support vector machines [7] (SVMs) are a set of related supervised learning methods used for classification and regression. A support vector machine constructs a hyper plane or set of hyper planes in a high-dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the nearest training data points of any class called functional margin, since in general the larger the margin the lower the generalization error of the classifier. Let's introduce the notation used to define formally a hyperplane:

$$f(x) = \beta_0 + \beta^T x$$

where β is known as the weight vector and β_o as the bias.The optimal hyperplane [8] can be represented in an infinite number of different ways by scaling of β and β_o. As a matter of convention, among all the possible representations of the hyperplane, the one chosen is

$$|\beta_0 + \beta^T x| = 1$$

In general, the training examples that are closest to the hyperplane are calledsupport vectors. This representation is known as the canonical hyperplane. Now, the result of geometry that gives the distance between a point x and a hyperplane (β,β_o):

$$Distance = \frac{|\beta_0 + \beta^T x|}{\|\beta\|}$$

In particular, for the canonical hyperplane, the numerator is equal to one and the distance to the support vectors is

$$Distance_{supportvectors} = \frac{|\beta_0 + \beta^T x|}{\|\beta\|} = \frac{1}{\|\beta\|}$$

Recall that the margin introduced in the previous section, here denoted as, is twice the distance to the closest examples:

$$M = \frac{1}{\|\beta\|}$$

Finally, the problem of maximizing M is equivalent to the problem of minimizing a function L (β) subject to some constraints. The constraints model the requirement for the hyper plane to classify correctly all the training examples x_i. Formally,

$$min_{\beta,\beta o} L(\beta) = \frac{1}{2}\|\beta\|^2 subjecttoy_i(\beta^T x_i + \beta_0) \geq 1\forall_i$$

3.2.Proximal Support Vector Machine (PSVM)

The proximal SVM [10] also uses a hyper plane w.x+b=0 as the separating surface between positive and negative training examples. But the parameter w and b are determined by solving the following problem.

$$min\frac{1}{2}(\|w\|^2 + b^2) + C\sum_i \xi_i^2$$

The main difference between SVM and proximal SVM is the constraints. SVM employs an inequality constraint whereas proximal SVM employs an equality constraint. SVM only considers points on the wrong side of $w.x_i+b=1$ and $w.x_i+b=-1$ training errors. However, in proximal SVM, all the points not located on the two planes are treated as training errors. In this case the value of training error ξ_i in PSVM may be positive or negative. The second part of the objective function in PSVM uses a squared loss function $\sum_i \xi_i^2$ instead of $\sum_i \xi_i$ to capture this new notion of error.

3.3.Active Support Vector Machine (ASVM)

Thealgorithm [9]consistsof determiningapartition of thedual variableuintonon-basic andbasicvariables. The non-basic variables are those which are set to zero.The values of the basic variables are determined by finding thegradientof theobjective function of SVM with respectto thesevariables, setting thisgradientequalto zero, and solvingtheresulting linearequationsforthebasicvariables.If anybasic variable takeson anegativevalueafter solvingthelinear equations,itissetto zero and becomesnon-basic.This is the essence of the algorithm.In order to make the algorithm converge and terminate, afewadditionalsafeguardsneed to be putin placein orderto allowus to invoke thefiniteterminationresult.

Theotherkeyfeatureof thealgorithm isacomputation a loneandmakesuseof the SMWformula.This feature allows us to invert an (n + 1) x(n + 1) matrix at each step insteadof amuch biggermatrix of orderm xm.

H = D [A - e], Q= I /v + HH'.

With these definitions the dual problem becomes

$$min_{0\leq u\in R^m} f(u) := \frac{1}{2}u'Qu - eu$$

It will be understood that within the ASVM Algorithm, Q - 1 will always be evaluated using the SMW formula and hence only an (n+l) x (n+l) matrix is invertedthis new notion of error.

4.Experimrnts and Results

Three experiments have been carried out by implementing support vector machine and its variants such as active support vector machine and proximal support vector machine. The training dataset is developed using 500 images related to five types of skin lesion diseases. Image resize and image adjustment are the two main preprocessing tasks carried out. The color and texture features are extracted from the preprocessed skin lesion images to form feature vectors. For each feature vector, class labels 1 to 5 are assigned. Label 1 indicates acitinickertosis disease, class label 2 indicates basal cell carcinoma disease, class label 3 indicates melanocytic lesion disease, class label 4 indicates squmous cell carcinoma disease and class label 5 indicates sebhorreic kertosis disease.

In the first experiment, the non melanoma skin lesion prediction model is built by implementing multi class SVM. The mex code of SVM from LIBSVM is interfaced with MATLAB code during implementation. It consists of a training module (svm_train) and a prediction module (svm_predict).The RBF kernel is used and the regularization parameter c is assigned values ranging from 0 to 1. It is observed that the training was stabilized for c=0.9.The predictive accuracy of about 86% is reported and the average number of support vectors recognized by learning is 228.

In the second experiment, the same training dataset has been used to implement PSVM. The linear kernel is used and the regularization parameter c is assigned values ranging from 0 to 1. It is observed that the training was stabilized for c=0.9. The performance of the PSVM classifier is evaluated using 10 fold cross validation and the predictive accuracy of 93% is reported. The average number of support vectors identified during training is 463. The result of PSVM in terms of accuracy and learning time is shown in Table II.

Table 2. *Results of PSVM.*

Parameter c	Prediction Accuracy	Learning Time (in secs)
0.1	91.24	0.62
0.2	86.24	0.61
0.3	87.65	0.56
0.4	84.32	0.53
0.5	85.43	0.59
0.6	90.1	0.56
0.7	86	0.59
0.8	91.4	0.54
0.9	**93**	**0.50**
1	87.3	0.64

The same training dataset has been used to implement ASVM in the third experiment. The linear kernel is used and the regularization parameter C is assigned values ranging from 0 to 1. It is observed that the training was stabilized for C=0.9. The performance of the ASVM classifier is evaluated using 10 fold cross validation and the predictive accuracy of 92% is reported. The average number of support vectors documented here is 362. The result of ASVM in terms of accuracy and learning time is shown in Table III.

Table 3. *Results of ASVM*

Parameter c	Prediction Accuracy	Learning Time (in secs)
0.1	81.02	0.69
0.2	84.32	0.67
0.3	89.26	0.65
0.4	85.32	0.61
0.5	91	0.62
0.6	85.33	0.68
0.7	83.71	0.66
0.8	91	0.62
0.9	92	0.60
1	90.23	0.64

The comparative results of SVM, PSVM and ASVM in terms of accuracy learning time and number of support vectors shown in Table IV and illustrated in Fig.6.

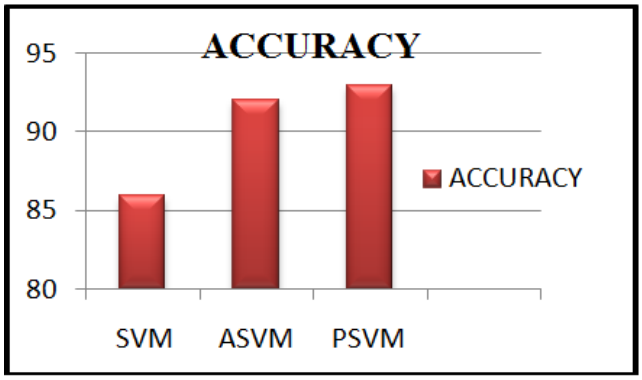


Fig.6.*Comparison of Classifiers accuracy*

Table 4. *Comparison of Classifiers.*

Classifier	SVM	ASVM	PSVM
Accuracy	86%	92%	93%
Learning time	0.85	0.60	0.50
Number of support vectors	228	362	463

The comparative results indicate that PSVM based classification model yields a better performance when compared to other models. As far as the skin lesion prediction model is concerned, the PSVM model is more efficient in terms of predictive accuracy plays a major role than learning time. Once the model is developed, it can be integrated with medical diagnosing system. Therefore it is not necessary to pay much attention to learning time.

5. Conclusion

In this research, an automated medical decision support model for non-melanoma skin cancer has been developed. Support vector machine and its variants such as active support vector machine and proximal support vector machine has been used to train the model. The outcome of the experiments proves that proximal support vector machine is effectual. As future enhancement of this research work, more features that help to increase the classification accuracy can be identified, extracted and used for learning.

References

[1] G. R. Day & R. H. Barbour. "Automated melanoma diagnosis: where are we at?" Skin Research and Technology 6, pp. 1–5,2000.

[2] Lucia Ballerina, Robert B. Fisher Ben Aldridgey, Jonathan Reesy "Non-melanoma skin lesion classification using colour image data in a Hierarchical k-nn classifier"

[3] M. E. Celebi, W. V. Stoecker, and R. H. Moss, "Advances in skin cancer image analysis," Computerized Medical Imaging and Graphics, vol. 35, No. 2, Pp. 83 – 84, 2011.

[4] M. EmreCelebi and Hassan and A. Kingravi"A methodological approach to the classification of dermoscopyimages"Comput Med Imaging Graph. 2007 September; 31(6): 362–373.

[5] José FernándezAlcón, CalinaCiuhu, Warner ten Kate, Adrienne Heinrich, NatalliaUzunbajakava, GertruudKrekels, Denny Siem, and Gerard de Haan-2009"Automatic Imaging System With Decision Support for Inspection of Pigmented Skin Lesions and Melanoma Diagnosis"

[6] .I. Maglogiannis And C. N. Doukas, "Overview Of Advanced Computer Vision Systems For Skin Lesions Characterization," Ieee Transactions On Information Technology In Biomedicine, Vol. 13, No.,Pp.721–733,2009.

[7] Ruben Nicolas,1 Albert Fornells," DERMA: A Melanoma Diagnosis Platform Based onv Collaborative MultilabelAnalog Reasoning".

[8] Cortes, Corinna, Vapnik and Vladimir N, "Support Vector Networks, Machine Learning".

[9] Xin Li and YuhongGuo "Active Learning with Multi-Label SVM Classification" Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence.

[10] Andreas Vlachos," Active Learning with Support Vector Machines

[11] GLENN M. FUNG," Multicategory Proximal Support Vector Machine Classifiers" 2005 Springer Science + Business Media, Inc. Manufactured in The Netherlands.

[12] C.-W. Hsu and C.-J. Lin. A comparison of methods for multi-classsupport vector machines. IEEE Transactions on Neural Networks,2002.]

[13] Kapoor, K. Grauman, R. Urtasun, and T. Darrell. Active learningwith Gaussian Processes for object categorization. In ICCV, 2007.

[14] M. Li and I. Sethi. Confidence-based active learning. IEEE Trans.PAMI, 2006.

[15] H.-T. Lin, C.-J. Lin, and R. C. Weng. A note on Platt's probabilisticoutputs for support vector machines. Machine Learning, 2007.

[16] T. Mitchell. Machine Learning. Boston: McGraw-Hill, 1997.

[17] A. Dhawan. "An expert system for the early detection of melanoma using knowledge-based image analysis." *Anal Quant CytolHistol*10, pp. 405–416, 1988.

[18] D. Gutkowicz-Krusin, M. Elbaum, P. Szwaykowski et al. "Can early malignant melanoma be differentiated from atypicalmelanocytic nevus by in vivo techniques?" *Skin Res Technol*pp. 3:15–22, 199.

[19] S. McDonagh. "Skin Cancer Surface Based Classification." *Undergraduate Thesis, School of Informatics, University of Edinburgh*2008.

[20] H. Ganster, A. Pinz, R. Rohrer et al. "Automated melanoma recognition." *IEEE Transactions on Medical Imaging* 20, pp. 234–239, 2001.