

# Comparative Analysis and Automatic detection of various skin diseases and Implementation of leucoderma using support vector machine

Archana Singh  
Asstt. Professor  
Amity University  
Noida

Ajay Rana  
Director  
Amity University  
Noida

Shipra Srivastava  
M.Tech Research Scholar  
Amity University  
Noida

## ABSTRACT

This paper deals with various utilities of support vector machine in the field of dermatology, especially in skin disease detection. Skin diseases have common to everyone, different types of allergies are becoming more common. Many of these diseases are very dangerous, particularly if not treated at an early stage. The main problem with these diseases is the separation of affected skin from normal skin. Modern medicine is looking for solution which could help the doctors with any aspect of work using the new technology. One of the common approaches used in this area is Support Vector Machine. There are so many diseases that have got the solution using SVM such as automatic detection of ringworm, skin cancers.

## Keywords

Dermatology, Support vector machine, leucoderma.

## 1. INTRODUCTION

SVMs deliver state-of-the-art performance in real-world applications such as text Categorisation, hand-written character recognition, image classification, biosequences analysis, etc., and are now established as one of the standard tools for machine learning and data mining.

"The support-vector network is a new learning machine for two-group classification problems. The machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high-dimensional feature space. In this feature space a linear decision surface is constructed." Cortes and Vapnik (1995) with their introduction in 1995 stated, "Support Vector Machines (SVMs) marked the beginning of a new era in the learning from examples paradigm". Rooted in the Statistical learning theory developed by Vladimir Vapnik at AT&T, SVMs quickly gained attention from the pattern recognition community due to a number of theoretical and computational merits. These include: simple geometrical interpretation of the margin, uniqueness of the solution, statistical robustness of the loss function, modularity of the kernel function, and over fit control through the choice of a single regularization parameter. Support Vector machines (SVM) are a new statistical learning technique that can be seen as a new method for training classifiers based on polynomial functions, radial basis functions, neural networks, splines or other functions. Support Vector machines use a hyper-linear separating plane to create a classifier. For problems that can not be linearly separated in the input space, this machine offers a possibility to find a solution by making a non-linear transformation of the original input space into a high dimensional feature space, where an optimal separating

hyperplane can be found. Those separating planes are optimal, which means that a maximal margin classifier with respect to the training data set can be obtained.

## 2. RELATED WORK

In recent years a lot of work has been done in the area of skin disease detection using computer applications. Most of the implementation part has been done using Support Vector Machine. We have studied these researches some of which we want to discuss in this paper.

### 2.1 Automatic detection of Ringworm

Srimanta Kundu, Nibar Das and Mita Nasipuri used Svm for the automatic detection of ringworm in their paper [1]. For the Support Vector Machine classifier, an open source software LibSVM tool is used. In general, a classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features) [1].



Fig 1(a): sample of Ringworm images



Fig 1(b): sample of normal skins

According to their research they achieved 90% success rate in the automatic detection of ringworm using support vector machine.

### 2.2 Skin Cancer

Skin diseases are now very common all over the world. For example, the number of people with skin cancer has doubled in the past 15 years. "Currently, between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occurs globally each year [2]. Svm also applied to some medical tools. One of them, proposed in [3], identifies patients with breast cancer for whom chemotherapy could prolong survival

time. In the experiment, SVM was used for classifying patients into one of three possible prognostic groups: Good, Poor or Intermediate. After classification of 253 patients, 82.7% test set correctness was achieved. Another application of SVM in medicine is described in [4]. A system for cancer diagnosis used the DNA micro-array data as a classification data set. Diagnosis error obtained by above mentioned system was smaller than in systems which uses other known methods. Reduction of the error achieved 36%. SVM is relatively new method of classification and it expands very quickly. That will certainly cause wider use of SVM in different areas, also in medicine [5].

### 2.3 Skin Mole localization using DoG filters and a support vector machine

Mole pattern changes are important clues in detecting early signs of melanoma, a deadly skin cancer [6]. Early detection is especially important for melanoma because, while advanced cases are not curable, the disease can be cured if detected early [7]. However, a principled system to register mole pattern changes is currently lacking. In fact, a major burden on the dermatological workforce is in manual surveillance of pigmented lesions, which is both time consuming and prone to human error [8].

In this section, DoG scale-space filters and the designed support vector machine (SVM) classifier will be introduced. Since the size of moles can vary, moles should be searched in a multi-scale fashion. The DoG filter is applied to RGB color channels separately, and the set union of the output maxima over scale in each channel are considered possible mole candidates. When combining the DoG maxima, any mole candidates occurring within a radius of another mole candidate is eliminated. Once mole candidates are localized, regions around mole candidates are cropped from the hair-removed image. The width of mole candidate patches is  $2 \times p \times 2$  times the radius of each DoG maximum.

Cropped mole candidates are classified as moles using a support vector machine (SVM) classifier. An SVM is a powerful tool to both generalize and classify objects. LIBSVM [1] was used to build the SVM classifier.

### 2.4 Acne Vulgaris

Acne Vulgaris (see Figure 2) is a chronic inflammatory disease characterized by cysts, open and closed comedones (plugged lesions containing a “cottage-cheese” like material), pus pockets, and raised red swellings. Often occurs in puberty, though may occur in the 20s or 30s. Several factors are involved in the pathogenesis: inheritance, increased sebum production, an abnormality of the microbial flora, cornification of the pilosebaceous duct and the production of inflammation [6, 7, 8].

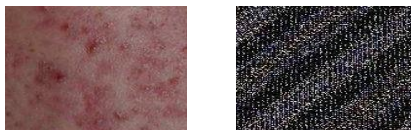


Fig 2: Acne Vulgaris

Before training programs can be used, all the pictures have to be compiled into proper dataset file. Michal has developed the Skinchecker-DataSet module which provides all necessary tools for image preprocessing. It also simplifies the organization of lists of all diseases and their images. User can simply add new diseases and assign pictures to them as well as remove any not used data. All lists can be saved and used in future, which is

very useful when there is a need to prepare a dataset from the same collection of pictures but with another parameters or when some images need to be added to the list (see figure 3). This version of Skinchecker uses LIBSVM library [9], which provides the support vector classification (SVC) in two versions: C-SVC and SVC. After using the C-SVC method it has appeared that results were not satisfactory. Therefore Michal decided to use SVC and they obtained satisfactory results for  $< 0, 3$ .

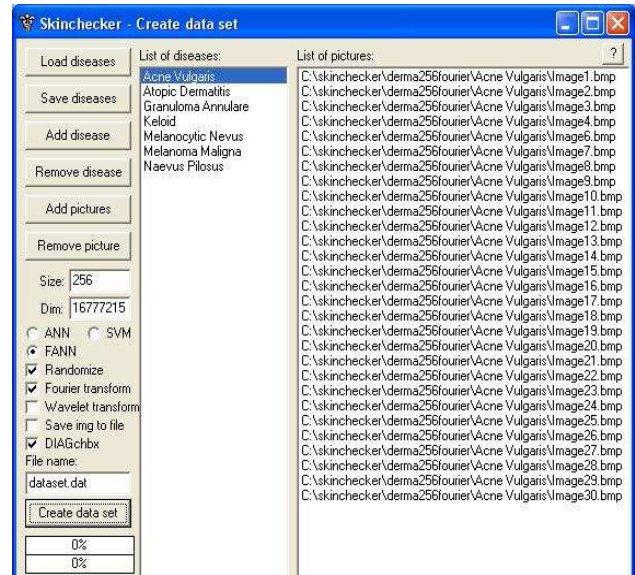


Fig 3: Data set

More information about above mentioned SVC methods can be found in [10]. He used SVC for classification of pictures. The results of testing the SVM for different value showed that results for acne vulgaris are nearly same comparing to other diseases.

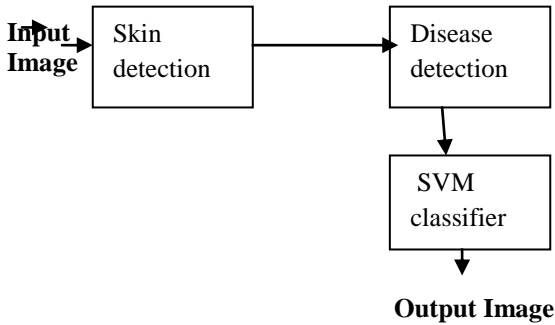
### 2.5 Skin Pigmentation

Several methods based on absorbance spectrum in visible and near infra-red have already been presented in the literature for quantifying the melanin in the skin [11, 12]. The main idea of these methods is to select specific spectral bands in the data in order to extract information on skin pigmentation. One of the most famous algorithms has been proposed by Stamatas in [13]. This algorithm is based on the analysis of the skin chromophores absorbance spectrum. An affine model of the melanin absorbance is proposed in the spectrum range of 630-700 nm: The spectrum can be corrected by subtracting the influence of melanin. This research concludes that whole spectrum instead of only few bands of interest can improve the skin hyper-pigmentation classification on multispectral images. They notice that the developed algorithms have an average computation time of 15 seconds on an image of 900 by 1200 by 18 pixels (with a Matlab code on a 2.2GHz core processor system) which is a reasonable time for a medical application.

## 3. PRESENT WORK

Leucoderma is a cutaneous condition, an acquired condition with localized loss of pigmentation of the skin that may occur after any number of inflammatory skin conditions, burns, intralesional steroid injections, postdermabrasion, etc. The disease is characterized by loss of normal color of the skin in

patches, resulting in various degrees of cosmetic disfigurement. It may cause a severe sense of humiliation, worry, anxiety & represent a social obstacle throughout life. The affected patient prefers seclusion & abstains from appearing in public, lest he should be looked upon as having a contagious disease [14]. The feature vector extraction procedure used in this work is delineated in Figure 4.



**Fig 4: Block diagram of svm**

The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Before considering the data directly from the linearly scaling each attribute to the range [-1, +1] or [0, 1].

Given a training set of instance-label pairs (xi, yi);

$$i = 1 \dots N;$$

Where xi ∈ Rn and yi ∈ {1, -1},

the support vector machines (SVM) require the solution of the following optimization problem:

$$\min (w, w_0, \epsilon, ) \frac{1}{2} w^T w + C \sum_{i=1}^N \epsilon_i$$

$$\text{Subject to } y_i (w^T x_i + b) \geq 1 - \epsilon_i, \quad \epsilon_i \geq 0$$

It's not difficult to generalize this linear program to the nonlinear case replacing xi with a nonlinear function M(xi):

$$\min (w, w_0, \epsilon, ) \frac{1}{2} w^T w + C \sum_{i=1}^N \epsilon_i$$

$$\text{Subject to } y_i (w^T M(x_i) + b) \geq 1 - \epsilon_i, \quad \epsilon_i \geq 0$$

### 3.1 Preparation of database

For the experiment we have taken different images of leucoderma skin and normal skin from internet. These collected images are not directly used in the experiment. Database of diseased skin and normal is undertaken in a sequence of steps.

A total of 129 images are used for the experiment. 9 images are of leucoderma affected skin and remaining 9 images are of normal skin. We have trained different classifiers using 15% of images from the leucoderma positive and negative image sets. Thus our training set is formed with 70 images with equal number of images from each class.

#### 3.1.1 Pre-processing steps

In the processing steps we have done the following things.

i. Prepare training data set of normal and diseased skin.

ii. Prepare test set of normal and diseased skin for feature extraction.

iii. Resizing all the images. All the images are of (144 x 144) pixels.

iii. Provide image as input through svm classifier and check the output of the image

iv. Cropping the specified zones of the diseased skin.

v. Converting those images into gray images.

After the pre-processing, histogram is calculated. For this, different region based approach is also considered. Each image is decomposed into different zones or regions and each region is assigned equal weight as all the regions are of equal importance.



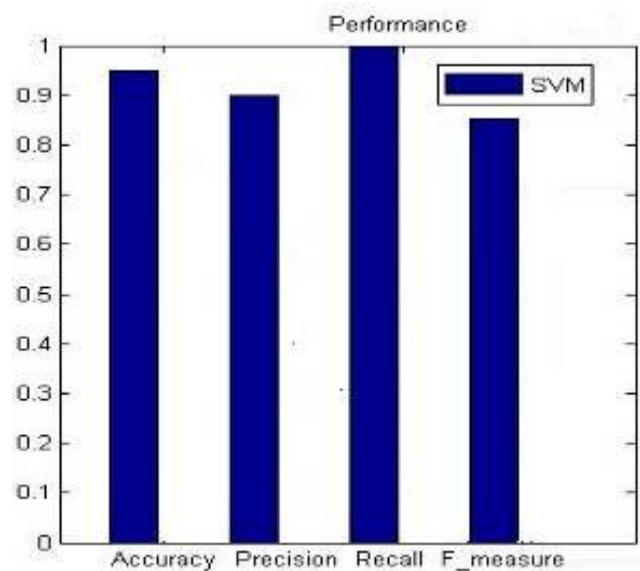
**Fig 5(a): sample of leucoderma images**



**Fig 5(b): sample of normal images**

## 4. RESULT

For present work 9 images are chosen as training data. We have applied the above methodologies on the designed dataset. After training the classifiers we have used the test set to get actual recognition. Graph.4.1 shows the result of svm classifiers on test data. From the graph we can calculate accuracy, recall, precision etc.



**Fig 5: Result of disease detection using leucoderma**

We have also discussed different types of skin diseases implemented through svm in this paper including our own. The

test result for various diseases is summarised in table 4.1[15, 16].

**Table 1. Summary of result for various skin diseases**

Disease	Result(in terms of accuracy)	Remarks
Ringworm	90%	Better result than Bayesian classifier
Skin Cancer (Meanoma Maligna)	40-50%	Worst result (lack of database)
Skin mole	76-84%	Depends on test data
Acne Vulgaris	89.3%	Depends on test data
Skin pigmentation	High success rate in weak contrasted area.	Better result than Stamata algorithm
Leucoderma	97%-99%	For large set of data

## 5. CONCLUSION

In this paper, the applications of support vector machine used in medicine have been discussed. There are several studies which have worked on it using dermatology datasets. In this paper we have summarized some of them including our own work. In future, the success rate may be improved with addition of some other features extraction algorithm. Use of other classifiers like RBF instead of SVM may be ameliorated the success rate. Further a model can be extended to detect more than one disease with same or greater accuracy rate.

## 6. REFERENCES

- [1] Srimanta Kundu, Nibar Das, Mita Nasipuri. Automatic Detection of Ringworm using Local Binary Pattern (LBP).
- [2] Vapnik.V, 1999. An overview of statistical learning theory. IEEE Transaction on Neural Network. 10:989-999.
- [3] World Health Organization. Ultraviolet radiation and the INTERSUN Programme. <http://www.who.int>, accessed April 27, 2005,
- [4] Bernhard Sholkopf, 2001. Survival-time classification of breast cancer patients. Techni-cal Report 01-03, National Taiwan University of Science and Technology, Taipei, Taiwan.
- [5] Ming Huang and Vojislav Kecman.2005. Gene extraction for cancer diagnosis by support vector machines. *Artificial Intelligence in Medicine*, 35:185–194.
- [6] Michal Antkowiak. May 3, 2000 Master’s Thesis in Computing Science.
- [7] Hensin Tsao, Taeg Sang Cho, William T Freeman. A reliable skin mole localization scheme.
- [8] [8] Tsao.H, Atkins.M, and Sober.A .2004. Management of cutaneous melanoma. *The new England Journal of Medicine*, 351(10):998-1012.
- [9] ECureMe. Medical Dictionary. <http://www.ecureme.com/>, accessed March 30, 2006, 2006.
- [10] Hunter. J. A .A, Savin. J. A, and Dahl.M.V.1995 *Clinical Dermatology*. Blackwell Science,.
- [11] Dermis. March 2006. *Dermatology Information Service*. <http://www.dermis.net/>.
- [12] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a Library for Support Vector Machines,2001.<http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [13] Piotr Szczepaniak, Obliczenia inteligentne, szybkie przeksztalcenia, i klasyfikatory.2004 Akademyka Oficyna Wydawnicza EXIT, Warszawa, PL.
- [14] Stamatas. G. N, Zmudzka. B. Z, Kollias. N, and Beer.J.Z.2004 “Non-invasive measurements of skin pigmentation in situ,” *Pigment cell res*, vol. 17, pp. 618–626, 2004.
- [15] Stamatas .G.N, Zmudzka .B.Z, Kollias.N, and J. Z. Beer.2008.In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance spectroscopy with a commercial instrument,” *British Journal of Dermatology*, vol. 159, pp. 683–690.
- [16] Le Poole IC, Das P.K, Vandan Wijngaard RMJGJ, Bas. J. D, Westerhof. W.1993. Review of the epthiopathomechanism of vitiligo: a convergence theory. *Exp Dermatol*; 2:145-53.