



Classification of Various Skin Lesions using SVM and KNN Classifiers

Fathima Nizar, G.S Santhosh Kumar

M.Tech Student, Dept. of CSE, Amal Jyothi College of Engineering, Kottayam, Kerala, India

Asst. Professor, Dept. of CSE, Amal Jyothi College of Engineering, Kottayam, Kerala, India

ABSTRACT: This paper introduces a new computer-aided method for classifying both melanocytic skin lesions and non-melanocytic skin lesions. Several methods are developed to distinguish between Melanoma, the most fatal skin cancer and Nevus, which are categories of melanocytic skin lesions. However, most of the studies did not focus on the most common skin cancer, Basal Cell Carcinoma (BCC) and Seborrheic Keratosis (SK), which are categories of non-melanocytic skin lesions. This method distinguishes among melanomas, nevi, BCCs, and SKs and calculates 4 texture features: Contrast, Correlation, Homogeneity and Energy. Here introduced two types of classification methods: Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). This method is tested on 42 dermoscopic images: 16 melanomas, 9 nevi, 8 BCCs, and 9 SKs. The SVM classifier outperformed the KNN classifier, achieving an accuracy rate of 75.95%.

KEYWORDS: Melanocytic Skin Lesion (MSL); Non-Melanocytic Skin Lesion (NoMSL); Basal Cell Carcinoma (BCC); Seborrheic Keratosis (SK); Support Vector Machine (SVM); K-Nearest Neighbor (KNN); Gray Level Co-occurrence Matrix (GLCM).

I. INTRODUCTION

As skin cancer incidences are increasing over years, early treatment is becoming more important. According to the five year survival rate of melanoma, the most fatal skin cancer is only 9-15% at stage 4, while it increases to 85-99% if detected early at stage 2. The most common skin cancer called Basal cell carcinoma which is rarely fatal, but it can destroy the surrounding tissue if left untreated. Thus, early detection of such skin cancers and their appropriate treatments are essential.

Detection of skin cancers is difficult due to the confusing appearance of wide variety of skin lesions. Melanomas and nevi are especially difficult to differentiate. Even with dermoscopy, which uses a magnifying glass with a polarization filter and a uniform light source, the accuracy of melanoma diagnosis by expert dermatologists remains at 75-84%. Biopsy provides a definitive diagnosis, however, it can cause metastasis, and therefore, is only allowed based on the premise of following surgical operation within a month. In addition, these are invasive operations and make unpleasant experiences to the patient.

To avoid unnecessary biopsy, several researchers investigated noninvasive computer-aided methods to distinguish melanomas from nevi using dermoscopy images. These methods usually consist of three steps: 1) border detection of skin tumor; 2) feature extraction; and 3) classification. The border detection process finds the border of the tumor in the dermoscopy image, which is essential for an accurate skin lesion classification. Several methods have been proposed such as the dermatologist-like method, SRM, hybrid thresholding, threshold fusion, and so on. The feature extraction process obtains discriminating image features that facilitate classification such as general color statistics, contour shape, and texture information. Wavelet coefficients that capture color and shape information have also been investigated. The classification process determines the type of skin lesions from the extracted image features. General pattern classifiers such as linear discriminant classifier, k-NN, artificial neural networks, and support vector machines (SVMs) are often used. Based on the aforementioned three steps, researchers have improved the automated classification methods [1].

II. RELATED WORK

In [2] authors propose a novel hierarchical classification system based on the K-Nearest Neighbors (K-NN) model and its application to non-melanoma skin lesion classification. Color and texture features are extracted from skin lesion



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

images. The hierarchical structure decomposes the classification task into a set of simpler problems, one at each node of the classification. Feature selection is embedded in the hierarchical framework that chooses the most relevant feature subsets at each node of the hierarchy. The accuracy of the proposed hierarchical scheme is higher than 93% in discriminating cancer and potential at risk lesions from benign lesions, and it reaches an overall classification accuracy of 74% over five common classes of skin lesions, including two non-melanoma cancer types.

In [3] presents a classification method of dermoscopy images between melanocytic skin lesions (MSLs) and non melanocytic skin lesions (NoMSLs). The motivation of this research is to develop a pre-processor of an automated melanoma screening system. Since NoMSLs have a wide variety of shapes and their border is often ambiguous, developed a new tumor area extraction algorithm to account for these difficulties. This paper confirmed that this algorithm is capable of handling different dermoscopy images not only those of NoMSLs but also MSLs as well. The paper determined the tumor area from the image using this new algorithm, calculated a total 428 features from each image, and built a linear classifier. The paper found only two image features, the skewness of bright region in the tumor along its major axis and the difference between the average intensity in the peripheral part of the tumor and that in the normal skin area using the blue channel were very efficient at classifying NoMSLs and MSLs. The detection accuracy of MSLs by the classifier using only the above mentioned image feature has a sensitivity of 98.0% and a specificity of 86.6% in a set of 107 non-melanocytic and 548 melanocytic dermoscopy images using a cross-validation test.

In [4] a methodological approach to the classification of pigmented skin lesions in dermoscopy images is presented. First, automatic border detection is performed to separate the lesion from the background skin. Shape features are then extracted from this border. For the extraction of color and texture related features, the image is divided into various clinically significant regions using the Euclidean distance transform. This feature data is fed into an optimization framework, which ranks the features using various feature selection algorithms and determines the optimal feature subset size according to the area under the ROC curve measure obtained from support vector machine classification. The issue of class imbalance is addressed using various sampling strategies, and the classifier generalization error is estimated using Monte Carlo cross validation. Experiments on a set of 564 images yielded a specificity of 92.34% and a sensitivity of 93.33%.

In [5] Computer Aided detection system is a classification system which distinguishes malignant melanoma from benign melanoma using imaging techniques and softwares. This methodology uses Digital Image Processing technique and Artificial Intelligence for the classification. First stage is the input to the system. Only a dermoscopic image is required for the detection system. There is no physical contact with the skin. The dermoscopic image is a high resolution digital image of the lesion. This image undergoes some image processing for the noise removal and border refinement. The hairs in the images are removed using Dull razor software. Any noises present in the images are removed using filtering. The filtering method used here is mean filtering. The next step is to isolate the lesion from the skin background. The method used for the segmentation is Otsu Color Threshold segmentation. In order to simplify the classification, some unique features of the lesion is extracted. It is known as feature extraction. Here two feature extraction methods are used - GLCM (Gray Level Cooccurrence Matrix) and Normalized Red, Blue, Green color features. Seven features are selected. These features are given to the input of a classifier which classifies the given images into cancerous and non-cancerous. The classifier used here is a Hybrid Genetic Algorithm- Artificial Neural Network classifier.

III. PROPOSED METHOD

A. Description of the Proposed Method:

Aim of the proposed method is to classify various skin lesions based on the texture features using SVM and KNN classifiers. The proposed method is consists of three main steps.

Step 1: Segmentation

Segmentation subdivides an image into its constituent regions or objects. That is, segmentation should stop when the objects of interest in an application have been isolated. There are many methods for segmentation: watershed based

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

segmentation, split and merge, region growing, threshold based segmentation etc.,. The segmentation process is carried out to extract the lesion area from the selected image.



Figure 1. Segmented Images of Various Skin Lesions

Step 2: Feature Extraction

The features are extracted from the segmented image based on the Gray Level Co-occurrence Matrix (GLCM) method. GLCM is a method used to extract texture features from a gray level image. The various texture features are:

a) Contrast

It is the measure of local intensity variation of pixels.

b) Correlation

It is a measure of gray level linear dependency between the pixels at the specified positions relative to each other.

c) Energy

It shows how the gray levels are distributed, when the number of gray levels is low then energy is high.

d) Homogeneity

It is a measure of uniformity of an image.

The extracted texture features of different skin lesions are given as input to the classifiers. The different feature values obtained for various melanocytic and non- melanocytic skin lesions are given in table1 and table2.

Step 3: Classification

The various skin lesions are classified based on their texture features using Support Vector Machine (SVM) and K Nearest Neighbor (KNN) classifiers. The texture features extracted using GLCM method is given as input to the classifiers.

This system uses a hierarchical structure for classification. That is, it consists of 3 KNN classifiers and 3 SVM classifiers. The first KNN classifier is to classify melanocytic and non-melanocytic skin lesions. The second KNN classifier is to classify melanoma and nevus skin lesions. The third KNN classifier is to classify Basal Cell Carcinoma (BCC) and Seborrheic Keratosis (SK). Similarly, the three SVM classifiers are used. Only the correctly classified output of the first classifier is given as input to the second and third classifiers both in SVM and KNN.

And finally the accuracy of both SVM and KNN classifiers are calculated and found that SVM gives a better accuracy than KNN. That is, SVM gives better classification results than KNN. The output of both types of classifiers is given in figure 2 and figure 3.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Image Types	Contrast	Correlation	Energy	Homogeneity
melanoma1	0.1991	0.9365	0.4761	0.9736
melanoma2	0.5997	0.8873	0.2961	0.8634
melanoma3	0.4159	0.8564	0.4216	0.9155
melanoma4	0.2467	0.9179	0.5114	0.958
melanoma5	0.4525	0.9012	0.3907	0.9509
melanoma6	0.3914	0.8492	0.349	0.9164
melanoma7	0.3672	0.928	0.4029	0.9676
melanoma8	0.3679	0.8979	0.5995	0.9402
melanoma9	0.245	0.921	0.6481	0.9614
melanoma10	0.2049	0.8755	0.7241	0.9625
melanoma11	1.2392	0.8203	0.2805	0.8155
melanoma12	0.5908	0.8885	0.483	0.8991
melanoma13	0.3026	0.9238	0.5662	0.9455
melanoma14	0.2854	0.9303	0.4008	0.915
melanoma15	0.306	0.9401	0.4701	0.9325
melanoma16	0.2971	0.9345	0.4483	0.9431
nevus1	0.2917	0.9207	0.4986	0.9523
nevus2	0.1778	0.9491	0.822	0.9872
nevus3	0.1609	0.9232	0.5859	0.9647
nevus4	0.1782	0.8871	0.8115	0.9778
nevus5	0.1	0.9683	0.7766	0.9864
nevus6	0.1068	0.8924	0.8093	0.9846
nevus7	0.2344	0.9616	0.5332	0.9541
nevus8	0.2007	0.9065	0.483	0.9665
nevus9	0.255	0.9109	0.5076	0.9636

Image Types	Contrast	Correlation	Energy	Homogeneity
BCC1	0.18	0.8971	0.5254	0.9674
BCC2	0.1963	0.9158	0.4295	0.9488
BCC3	0.0515	0.8766	0.843	0.9891
BCC4	0.087	0.9654	0.7403	0.9864
BCC5	0.0489	0.8791	0.8839	0.988
BCC6	0.2124	0.9657	0.5873	0.9654
BCC7	0.113	0.9229	0.8363	0.981
BCC8	0.181	0.8969	0.6395	0.9643
SK1	0.0237	0.8529	0.9795	0.9971
SK2	0.0646	0.736	0.9807	0.9965
SK3	0.024	0.8258	0.9896	0.998
SK4	0.0595	0.8338	0.9549	0.994
SK5	0.0155	0.8024	0.9792	0.9968
SK6	0.0041	0.6694	0.9957	0.9994
SK7	0.0326	0.9319	0.9403	0.9954
SK8	0.0961	0.9105	0.8924	0.9884
SK9	0.0434	0.9338	0.9237	0.9942

Table 1. Texture Feature Values for Melanocytic Skin Lesions

Table 2. Texture Feature Values for Non-Melanocytic Skin Lesions

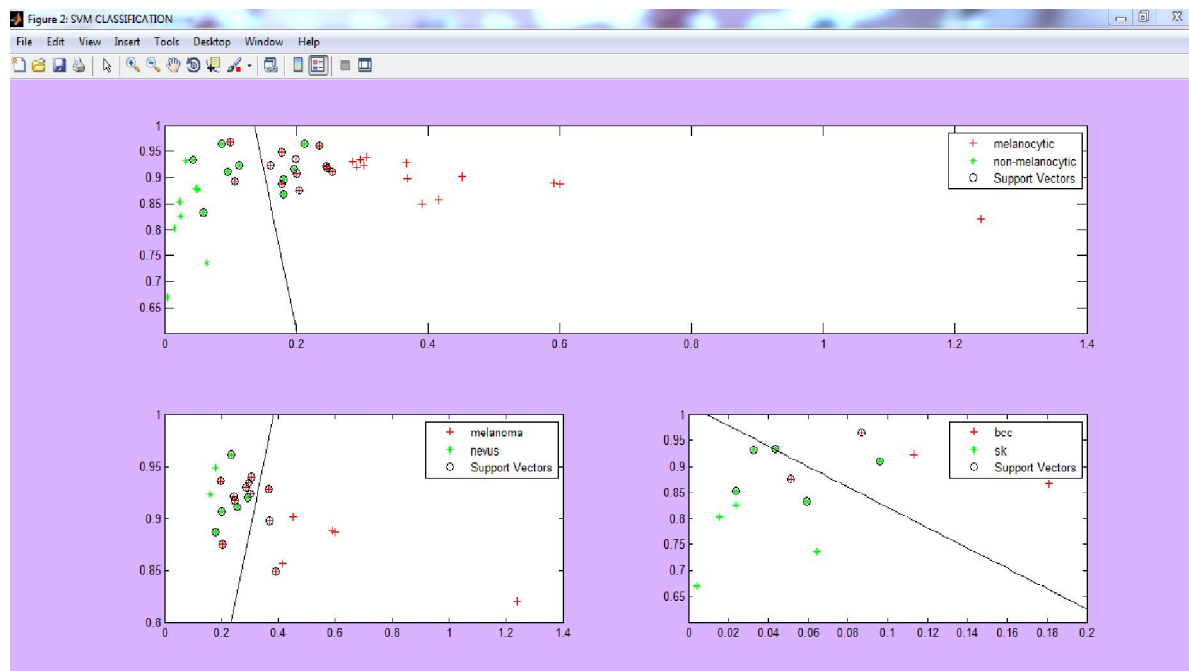


Figure 2. SVM Classification

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

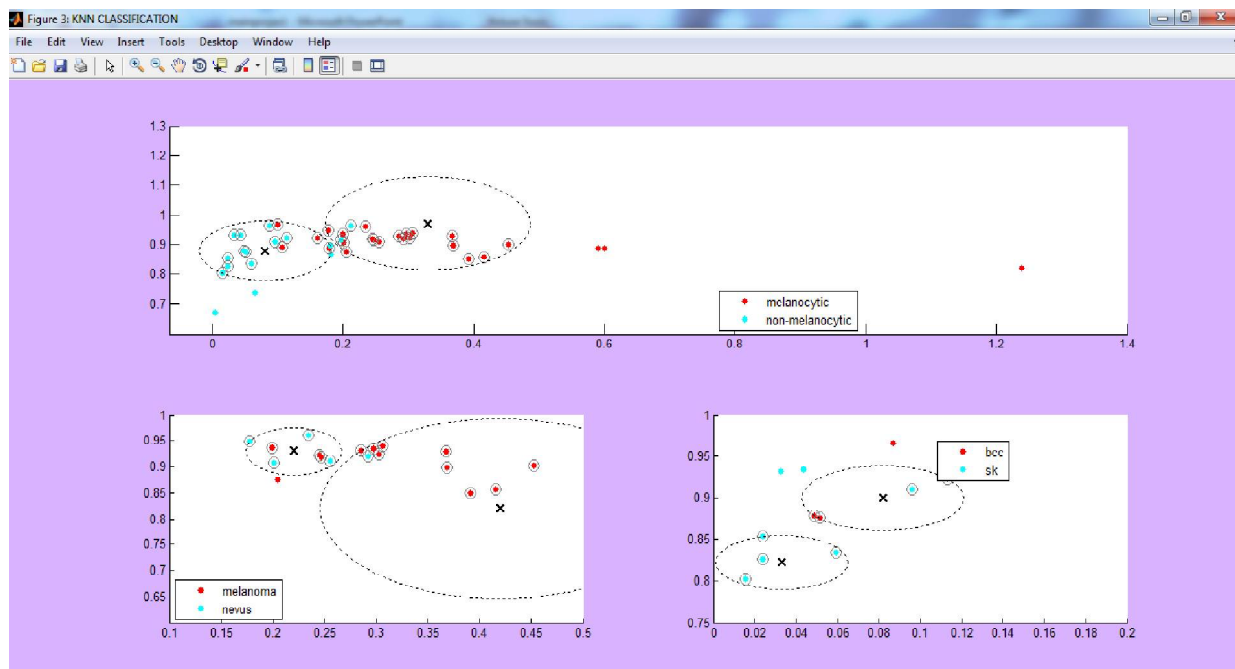


Figure 3. KNN Classification

IV. SIMULATION RESULTS

The proposed method is implemented with MATLAB. In this experiment, the performance of the system is expressed in terms of the accuracy of the two classifiers. The accuracy obtained for both SVM and KNN classifiers is given in table 3 and table 4.

Classifiers	Accuracy
Classifier1	85.71%
Classifier2	65.21%
Classifier3	76.92%
Average	75.95%

Table 3. Accuracy of SVM Classifiers

Classifiers	Accuracy
Classifier1	69.04%
Classifier2	72.22%
Classifier3	63.63%
Average	68.30%

Table 4. Accuracy of KNN Classifiers

From the above results, it is clear that SVM classifiers have better accuracy than KNN classifiers. So, the analysis proves that SVM gives a better classification results than KNN classifier.

V. CONCLUSION AND FUTURE WORK

This system performs preprocessing, segmentation, feature extraction and classification of various skin lesions using SVM and KNN classifiers and also calculates the accuracy of both classifiers. This system obtained an accuracy of 75.95% for SVM classifiers and 68.30% for KNN classifiers. From that we can that conclude SVM classifier performs better than KNN.

This system extracts only texture features of the selected image to perform classification. Other features such as color, shape etc can be extracted to get better classification output which can be done as a future work.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

REFERENCES

1. Kouhei Shimizu, Hitoshi Iyatomi, M. Emre Celebi, Kerri-Ann Norton, and Masaru Tanaka, "Four Class classification of skin lesions with task decomposition strategy", IEEE transactions on biomedical engineering, vol. 62, no. 1, January 2015.
2. Lucia Ballerini, Robert B. Fisher, Ben Aldridge, Jonathan Rees, "A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions", Department of Dermatology, University of Edinburgh, Edinburgh, UK.
3. Hitoshi Iyatomi, Kerri-Ann Norton, M.Emre Celebi, Gerald Schaefer, Masaru Tanaka, and Koichi Ogawa, "Classification of melanocytic skin lesions from non-melanocytic lesions", 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina, August 31 - September 4, 2010.
4. M. Emre Celebi, Hassan A. Kingravi, Bakhtiyar Uddin, Hitoshi Iyatomi, Y. Alp Aslandogan, William V. Stoecker, and Randy H. Moss, "A methodological approach to the classification of dermoscopy images", Comput Med Imaging Graph. Author manuscript; available in PMC 2011 October 13.
5. Abdul Jaleel, Sibi Salim and Aswin.R.B, "Hybrid Genetic Algorithm – Artificial Neural Network Classifier for Skin Cancer Detection", International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014.
6. Kiran Ramlakhan and Yi Shang, "A Mobile Automated Skin Lesion Classification System", 23rd IEEE International Conference on Tools with Artificial Intelligence, 2011.
7. Mariam A.Sheha, Mai S.Mabrouk and Amr Sharawy, "Automatic Detection of Melanoma Skin Cancer using Texture Analysis", International Journal of Computer Applications (0975 - 8887) Volume 42- No.20, March 2012.
8. Jayant Ghode and Ashutosh Datar, "Classification of Skin Melanoma using ANN", International Journal of Computer Applications (0975 - 8887) Volume 128 - No.10, October 2015.
9. J Abdul Jaleel, Sibi Salim and Aswin.R.B, "Computer Aided Detection of Skin Cancer", 2013 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2013].
10. Lucia Ballerini, Robert B. Fisher, Ben Aldridge and Jonathan Rees, "Non-Melanoma Skin Lesion Classification using Color Image Data in a Hierarchical K-NN Classifier", School of Informatics, University of Edinburgh, Edinburgh, UK.
11. Sigurdur Sigurdsson, Peter Alshede Philipsen, Lars Kai Hansen, Jan Larsen, Monika Gniadecka, and Hans Christian Wulf, "Detection of Skin Cancer by Classification of Raman Spectra", IEEE transactions on biomedical engineering, no 10, vol. 51, October 2004.
12. M. Emre Celebi, Y. Alp Aslandogan, William V. Stoecker, Hitoshi Iyatomi, Hiroshi Oka, and Xiaohu Chen, "Unsupervised border detection in dermoscopy images", Skin Res Technol. 2007 November ; 13(4): 454-462.