A Melanoma Skin Cancer Detection Using Machine Learning Technique: Support Vector Machine

Praveen Banasode*, Minal Patil, Nikhil Ammanagi

Department of Master of Computer Applications, Jain College of Engineering Belagavi-14
Affiliated to Visvesvaraya Technological University Belagavi, Karnataka state, India.

Email: praveenb.jce@gmail.com

Abstract. In this paper proposed is an easy way to detect the disease and help us to know before something turns out to be serious. The aim of this work is to detect skin cancer. People can get to know what skin disease they are having and what all precaution and measures to be taken at an early stage and it will help in treating the disease successfully. The major causes of skin cancer are air pollution, UV radiation, unhealthy life style etc. The concept of machine learning will be used to determine the disease and help us to detect the result. The most commonly used classification algorithms is support vector machine (SVM). First we are analyzing the skin image then converting the images into BGR-Gray and BGR-HSV for the computer to understand and enable it to read binary codes. The result of this study will help doctors to treat disease at the initial stage and further aggravation can be avoided.

Keywords: Melanoma, Lesion, Support Vector Machine

1. Introduction

Skin cancer is a rare disease which needs to be looked into it can cause many symptoms or will have different types which can be cured or may kill. It comes to any age group and can be deadly if not treated at an early stage. Approximately one in six people develop skin cancer it’s a rare disease which can affected in any way. In the medical field experts unanimously agrees that overexposure to sunlight may cause the increase of this disease.

This work is on skin cancer detection which helps us to know which type it is and what all necessary precautions should be taken. Skin cancer is the most vital thing in the medical field and we need to make the necessary arrangements to reduce this problem for the betterment of human life. Python and Machine Learning are used, which helps us to know the working and all other techniques in it. The work describes in three levels, which helps us to understand the procedure in a very precise and clear manner.

1.1 Deploying the image

The image should have a clear visibility and has to make the process in working manner. In this work some precautions are taken. The image should not be blur and low quality because if in case that is the situation then the outcome of the result will not be displayed and the accuracy won’t be showing us the correct result. So, load an image and then the further steps will be processed.

1.2 Pre-processing
The pre-processing method can be used to know which image has what type of disease. Firstly the image should be in high quality which gives the result instantly and greatly help us for better accuracy. Then we have to remove the background noise to know that only the selected part of image is being used. This has to be done for every image and the process keeps on repeating until all the images are fully covered and are ready to be processed.

1.3 Classification
In this, we will know the working of it so that it can be interconnected to the next level. First, we will be having a dataset (collections of data) where we have stored a large number of images in a folder. With this, we will be analysing the images to check for its clarity and high-quality resolution and not blur. Next, we have three categories which are the types of skin cancer. In the coding part we will be doing the classification folder, depending on the type the skin cancer belongs to, it goes to that folder. After that, we resize the images of all types for the machine to understand and it creates a folder of resized images in that dataset. Now we convert the images into BGR-Gray and BGR-HSV (hue, saturation, value) for the computer to understand and read it in binary codes.

1.4 Segmentation
Next comes the segmentation where we have to take the image and only cover the part where the skin cancer is there and remove the background, the covered part will be white coloured and the background will be black. This we call segmented image or masking where the image will be analysed and gives us the proper result.

2. Literature Survey

Alaa Haddad and Shihab A. Hameed [1] proposed to find skin disease by removing noise or unwanted things analysis result can support doctor to help initial diagnosis and to know the type of disease by using the low cost solution for the skin disease detection. Amir Mirbeket et. al [2] three semi-solid phantoms are used and accurately represent the dielectric properties of normal and malignant skin tissues the penetration depth of millimetre waves malignant skin disease is calculated penetrates the human skin tissue deep enough to reach the majority of skin structure. Azadeh Noori Hoshyar et. al [3], proposed automatic detection of skin cancer, where determining and describing the different steps of processing is achieved it is well designed as a substitute of clinician in melanoma diagnosis. Enakshi Jana et. al [4] worked on three types of cancers and tested using SVM and adaboost produced best results of image classification with good accuracy which is useful in classification of normal and abnormal skin cells different types of architecture of ANN and SVM for skin cancer is used to get result with accuracy various methods are used for classification of skin cells. Haseeb Younis et. al [5] proposed different types of skin cancer and skin classification method using deep learning where it classifies the skin within 2-3 seconds pre-trained MobileNet convolution neural network and trained over lesion dataset method used here is lightweight, fast and reliable shows a good accuracy of weightage 0.90 and 0.91 respectively. Kyamelia Roy et. al [6] proposed segmentation techniques on certain skin diseases using python language the adaptive Thresholding, edge detection, k-means clustering and morphology-based image segmentation are used to identify the skin diseases from the set. Pavels Osipovs et. al [9] proposed centralized algorithms for analysis and processing of data, thereby making sure the data is accessible and coordinates all other nodes in the system with certain limit.

3. Methodology
It helps us to know the result will have correct or somewhat similar answer to it. This algorithm is useful for the work that needs to be verified to find which is closure and which has a longer distance that can’t be used in this area.
In Fig 1, class 1 shows us the amount of space it is being maintained from the other class that is displayed in the diagram. It basically helps us to know that the closest it gets to the specified area that is allocated which we see in the image that is green in colour to the class 1 which is a neighbour because it is side by side that we call it as closer member is known as KNN. And the class 2 which are further away from the target can’t possibly reach the distance so this shows us the result of the diagram.

3.1 Support Vector Machine (SVM)

It is kind of an algorithm which helps us to analyse and gives the result. Let’s say it has data’s in that it has categorized in to two sets, one which helps us to mark the margin for the output purpose and other for the building a model to determine that which point it belongs to so this is how it works.

This can be a very useful algorithm, as you can see, we have co-ordinates which helps us to separate them on what basis it is being separated will be explained. The agenda of this algorithm is that it helps to know how it works as you can see, we have a dotted line in between the green side and the red side we call it a margin which helps us to divide them and see will it be able to work or no.

Fig 2: Co-ordinate points of supporting vector
(Source: https://www.analyticsvidhya.com)

Fig 3: System Design of Skin Cancer Prediction

Algorithm like Mining, KNN and Clustering can be used
4. Results and Discussions

The dataset collected from ISIC (International Skin Imaging Collaboration) containing more than 5341 images of melanoma and not melanoma, out of these images we choose 1751 images of melanoma. The training set of images is used for testing using the SVM, which classifies the affected skin area as melanoma. The feature like asymmetric behaviour, colour, and border irregularity are extracted. Finally this feature vector is given as input to SVM classifier which classifies the skin lesions [7]. Following tests are done.

![Diagram](image)

**Fig 4:** (a) Original Image (b) RGB to Gray Scale Image (c) Pre-processing Image (d) Labelling (e) Masking Image (f) Segmented Image (g) Image Transformation (h) Predicted Table (i) Predicted Table

a) The trained image is being deployed where it tells us the disease and the result.

b) This image is a conversion from one form to another it comes from light background.

c) It is an image which can be viewed in clear and helps us to find the spot of disease.
e) In this image, we have taken it from the original image and only taken the affected area which is white in colour so that it can be understood by anyone. This image we call it as masking.

f) This is the segmented part which helps us to know what difference does this make from masking.

h) This label tells us about the disease and how the number of images does the types have which are genuine through validation we get this model result.

g) Image transformation after SVM classification.

i) This label tells us about the accuracy of different types of diseases with the correct outcome.

By keeping other parameters in support of vector machine different tests are done to reduce the computational complexity. Following table shows the performance evaluation of SVM.

Table - 1 Performance Evaluation

|     | Sensitivity | Specificity | Accuracy  |
|-----|-------------|-------------|-----------|
|     | 95.7%       | 90.2%       | 96.9%     |

The performance are evaluated by calculating Specificity and Sensitivity

\[
\text{Specificity} = \frac{\text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive}} \quad \text{and} \quad \text{Sensitivity} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}
\]

The proposed method, using the SVM was successful. The result shows accuracy of 96.9% of detecting melanoma disease with other disease. The results explain the performance evaluation of SVM.

5. **Conclusion**

By using SVM algorithm we can be sure that the application which is going to be used in the coming future will be helpful for betterment of human life. Where the disease can be known at early stage or beforehand what type of disease the patient is affected with and what needs to be done before it’s too late for the person. Survival rate will be more if the Melanoma is detected in time. The accuracy of detecting disease is excellent. Since machine learning plays an important role of skin cancer detection it can be a helpful factor in the medical field. By using the latest technology we can make it easier. In future it is going to be a big boon in the medical field.

**References**

[1] Alaa Haddad, Shihab A. Hameed et. al "Image Analysis Model For Skin Disease. 2018 Advances in Wireless and Optical Communications.

[2] Amir Mirbeik-Sabzevari "Ultra-Wideband, Stable Normal and Cancer Skin Tissue Phantoms for Millimeter-Wave Skin Cancer Imaging". IEEE publication.

[3] Azadeh Noori Hoshyar, Adel Al- Jumaily, Riza Sulaiman "Review On Automatic Early Skin Cancer Detection". IEEE publication

[4] Enakshi Jana, Dr.Ravi Subban "Research on Skin Cancer Cell Detection using Image Processing". IEEE publication.

[5] Haseeb Younis et. al "Classification of Skin Cancer Dermoscopy Images using Transfer Learning". 2018 7th International Conference on Computer and Communication Engineering (ICCCE).

[6] Kyamelia Roy et. al "Skin Disease detection based on different Segmentation Techniques". IEEE Publication.

[7] Ms. Amulya P M, Prof Mr. Jayakumar T V "A Study on Melanoma Skin Cancer Detection Techniques". International Conference on Intelligent Sustainable Systems (ICISS 2017) IEEE Xplore Compliant - Part Number: CFP17M19-ART, ISBN:978-1-5386-1959-9.

[8] Naser Alfed, Fouad Khelifi and Ahmed Bouridane "Improving a Bag of Words Approach for Skin Cancer Detection in Dermoscopic Images". CoDIT'16 - April 6-8, 2016, Malta.

[9] Pavel Osipovs, Dmitrijs Bliznuks, "Cloud Infrastructure for Skin Cancer Scalable Detection System". IEEE publication.

[10] https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/.

[11] https://www.kaggle.com/drscarlat/melanoma.