| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 5, May 2020 ||

Automatic Skin Cancer Detection using Support Vector Machine

Nitin Kumar Nagle [1], Naveen Jain [2], Ankur Taneja [3]

PG Student, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[1]

Assistant Professor, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[2, 3]

ABSTRACT: In recent days, skin cancer is considered as one of the most harmful form of the Cancers generally found in Humans. Skin cancer can be classified in various types such as Basal, Melanoma, Actinic Keratosis, Basel Cell Carcinoma, Cherry Nevus, Dermatofibroma and Melanocytic Nevus. The skin cancer detection in early stage can be helpful to cure it properly. Computer vision is an area that can play important role in Medical Imaging Diagnosis. The input to the system is the skin lesion image that can be filtered through image processing techniques that analyses the presence of skin cancer from desired image. Here the system uses Support Vector Machine for detecting skin cancer with high level of accuracy. System will be capable to classify the cancerous area by determining its patterns or textures; not only on the basis of color and iterations will help to lead more accuracy as compare to the previous one. System can deal with unstructured data or non linear data points using kernel tricks.

KEYWORDS: Automatic Skin Cancer Detection, Support Vector Machine, Convolutional Neural Network, Machine Learning, Image Processing, Melanoma.

I. INTRODUCTION

Diagnosis of skin cancer usually begins with a visual examination. The Skin Cancer Foundation and the American Cancer Society recommend monthly self-screening and annual doctor visits for possible skin cancer. If a suspicious location is found, your doctor will first test the area, taking into account its size, shape, color, and texture, as well as any bleeding or scaling. Your doctor may also examine nearby lymph nodes to see if they are enlarged. If you are being seen by a primary care physician, you may be referred to a dermatologist who can perform more specific tests and make a diagnosis.

Fig. 1. Skin Cancer

Most skin cancers — particularly basal cell carcinoma, the most common form of skin cancer — remain local and do not spread to distant organs. Melanoma and Merkel cell carcinoma are at greater risk of spreading. In those cases, one of several medical imaging procedures can be used to determine whether cancer cells have metastasized to internal organs and bones. Melanoma is considered the deadliest form of skin cancer and is due to the development of a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 5, May 2020 ||

malignant tumor of melanocytes. The purpose of the skin cancer detection project is to develop a framework for analyzing and assessing the risk of melanoma using dermatographic photographs with a standard consumer-class camera [1].

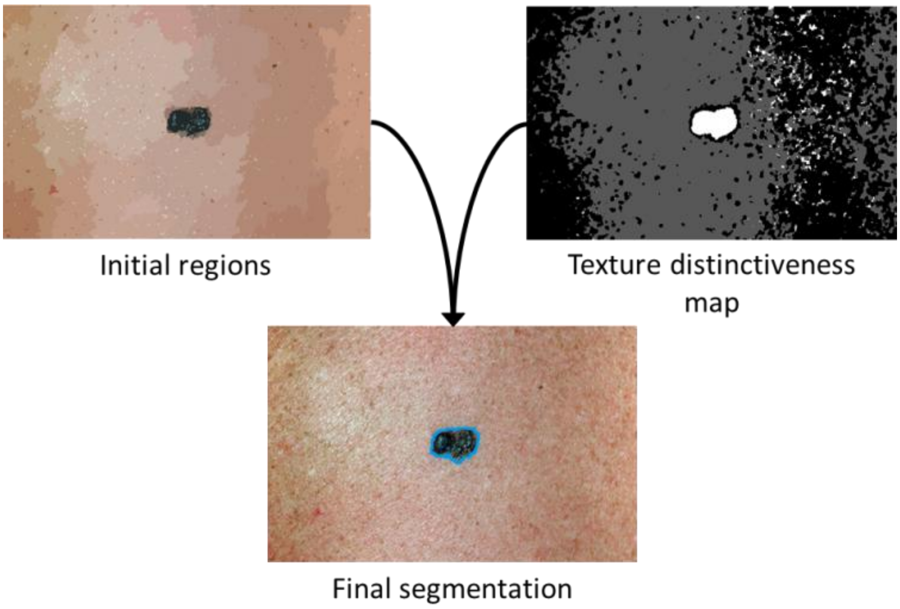


Fig. 2. Skin Cancer Segmentation [1]

The purpose of the skin wound segmentation step is to find out the extent of the skin wound. It is important that this step is performed correctly because many of the features used to assess the risk of melanoma are obtained based on lesion threshold. Our approach is a texture specificity based wound segmentation to find the wound boundary.

II. Related work

Farzam Kharaji Nezhadian et al. [2] proposed a system which is based on standard segmentation method where cancer affected skin area can be extracted and later classify the cancer using 2-D discrete wavelet transform. This paper uses SVM as well as color texture recognition for better classification. But transformation model is weak that resulted sharpen the image that looses the sensitivity of the disease. The purpose of this study was to diagnose benign and malignant melanomas.

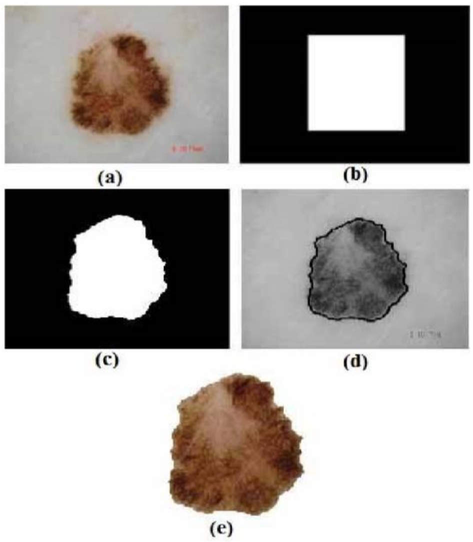


Fig. 3. Preprocessing steps for malignant lesion. original image (a), initial segmentation (b), active countor model segmentation (c), border of the sementation on gray scale image (d), colored ROI (e) [2]

Amulya P M et al. [3] proposed a system which is also based on SVM Classifier. The system depends upon the training samples and machine iterations. In this study, we have investigated various strategies for melanoma classification and detection. The process of melanoma detection is completed in various stages such as preprocessing, division, feature extraction, post handling, and arrangements that use advanced systems and software to obtain accurate results. As per the review conducted for ABCD rule based classification, we need proper and accurate feature extraction for accurate classification. Prachya Bumbrungkun et al. [4] proposed a system which is based on SVM classifier and snake model. This paper has been reported the causes of death in Thailand due to skin cancer. So, based on the SVM test; it helps to consider the template that compare with the input image. It has been considered that input skin cancer images may have curves, circles, rectangles and ellipse. It also uses snake model which is an edge detection technique that has been used for extracting the edges of the features that later compare with the templates. However the unwanted area can be eliminated and classified that does not affect the correct recognition rate. But snake model is not a good model for extracting edges from skin feature templates. Soniya Mane et al. [5] proposed a system which is based on preprocessing and classification techniques such as segmentation by thersholding and SVM classifier resp. Clustering is known as the process of grouping of similar data items together, while the items in the other clusters are as dissimilar as possible. This paper represents an efficient method to detect skin cancer. The system uses computer aided diagnosis to detect skin cancer. Detecting skin cancer manually is not only exhausting, but also a time-consuming task. The traditional method to detect skin cancer is biopsy. In this method, part of the suspected lesion is cut and sent for laboratory testing. So this method is aggressive, painful and time consuming. Shalu et al. [6] proposed a system which is based on Naïve Bayes, Decision Tree and KNN classifier. In this paper, a system has been developed to detect melanoma skin cancer. First, various preprocessing and segmentation techniques were used to enhance the image and extract the region of interest. Many features were removed from the HSV and YCbCr color space. Feature displays are tested on three types of classifiers: Na Bayve Bayes, Decision Tree and KNN Classifier. In comparison, it determines that the decision tree classifier performs better than other classifiers by achieving an accuracy of 82.35%. Noel B. Linsangan et al. [7] proposed a system which is based on k-Nearest Neighbors. The resulting system is a skin wound analyzer using image processing technology in a single board computer. Through the geometric features of the skin lesion, the researcher was able to classify suspected skin cancer samples from malignant, benign, and unknown. The system was able to perform tests to determine its accuracy. It was found that there are images that are not processed by the system and mostly give errors referring to the extraction of features in the image.

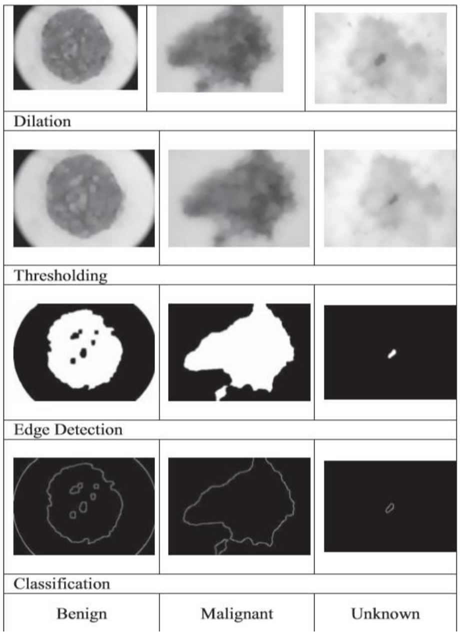


Fig. 4. Resulting Image [7]

Giulia Mansutti et al. [8] proposed a system that uses substrate integrated waveguide (SIW) technology. This work presents a novel millimeter-wave design near the skin large probe to detect skin cancer. The probe has been designed and tested through full-wave numerical simulations at CST: it can reach a lateral resolution of 0.2 mm and a depth of 0.4 mm, and is therefore capable of detecting early-stage skin tumors. Furthermore, since the probe relies on SIW technology and it uses microstrip feeds, it is inexpensive and easy.

III. PROBLEM IDENTIFICATION

Md. Zahid Hasan et al. [9] proposed a system that uses ANN (Artificial Neural Network) for classifying skin disease from infected area. System uses machine learning methodology for training datasets with various samples of skin diseases. System also erodes the non cancerous area on the basis of skin color and highlights the cancerous part. But train a system on the basis of supervised sample is not enough capable to dealing with irregular or undefined cancerous disease. Neural network is limited with samples and hard to predict correct decision on the basis of that. In machine learning a best classification method can classify the cancerous part effectively. System achieved 95 % of accuracy which could be bit higher. In this research, the effects of dimensionality are investigated using the Rough set on skin cancer decision support systems with multiple classifications such as Artificial Neural Network (ANN), Support Vector Machine (SVM) and Random Forest (RF). Results for the proposed model demonstrated that ANN is working with the aim of improving through a limited number of three types of at-risk skin cancer (common nevus, atypical nevus, and melanoma) datasets.

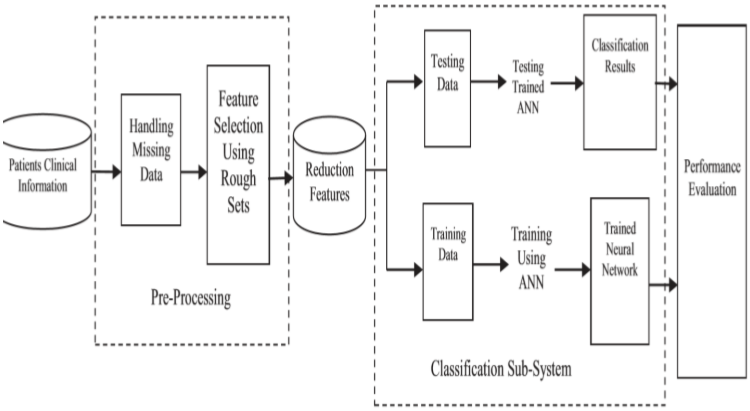


Fig. 5. Architecture [9]

IV. PROPOSED WORK

Here the proposed work is able to recognize various types of skin cancer with high level of accuracy with support vector machine. System will be capable to classify the cancerous area by determining its patterns or textures; not only on the basis of color and iterations will help to lead more accuracy as compare to the previous one. System can deal with unstructured data or non linear data points using kernel tricks. Skin cancer detection using SVM is basically defined as a process of detecting the presence of cancer cell in the image. SVM is machine learning technique, mainly used for classification and regression analysis.

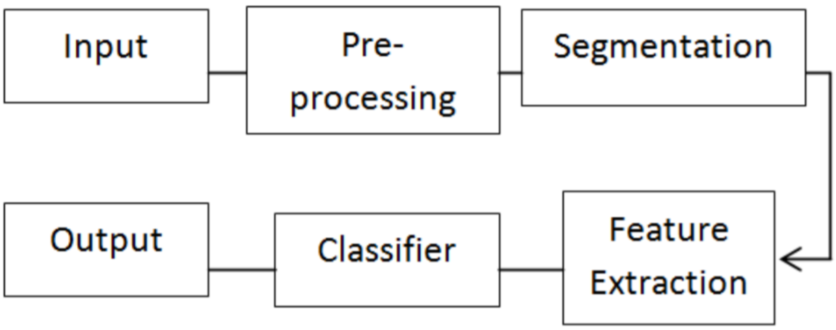


Fig. 6. Block Diagram

Pre-processing aims to improve image data that reduces and increases unwanted distortions some images are important for the front image processing. The objective of image enhancement is to process an image to increase visibility of feature of interest. Segmentation is process of removing region of interest from given image. Region of interest containing each pixel similar attributes. Here we are using maximum entropy thresholding for segmentation. The classifier is used to classify the image of cancer other skin diseases. Support vector for simplicity machine classifiers are used here. Takes set of SVM predicts images and for each input image which of the two categories of cancer and non-cancer. SVM aims to create hyper plane that separates the two squares maximum difference between them.

A. SUPPORT VECTOR MACHINE

In machine learning, support-vector machines (SVMs, support-vector networks) are monitored learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as one of one or two categories, an SVM training algorithm constructs a model that provides new examples for one category or another, Whereby it exists in a non-probable binary linear classifier (though methods) such as plot scaling to use SVM in a probabilistic classification classifier). An SVM model is a representation of instances as points in space, mapped so that instances of individual categories are divided by an explicit difference that is as detailed as possible. The new examples are then mapped to the same location and predicted to belong to a category based on the interval at which they fall.

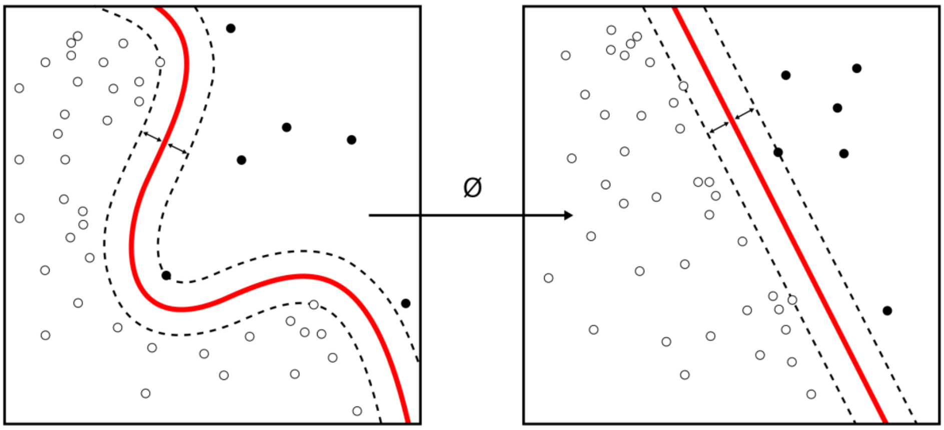


Fig. 7. Support Vector Machine [10]

When data are unlabled, supervised learning is not possible, and an untrained learning approach is required, which attempts to detect natural clustering of data in groups, and then map new data to these constituted clusters. Does. The support-vector clustering algorithm, created by Hwa Siegelman and Vladimir Vapnik, applies the statistics of support vectors, developed in the supplement vector machines algorithm, to categorize unlisted data, and is one of the most widely used clustering algorithms in industrial application.

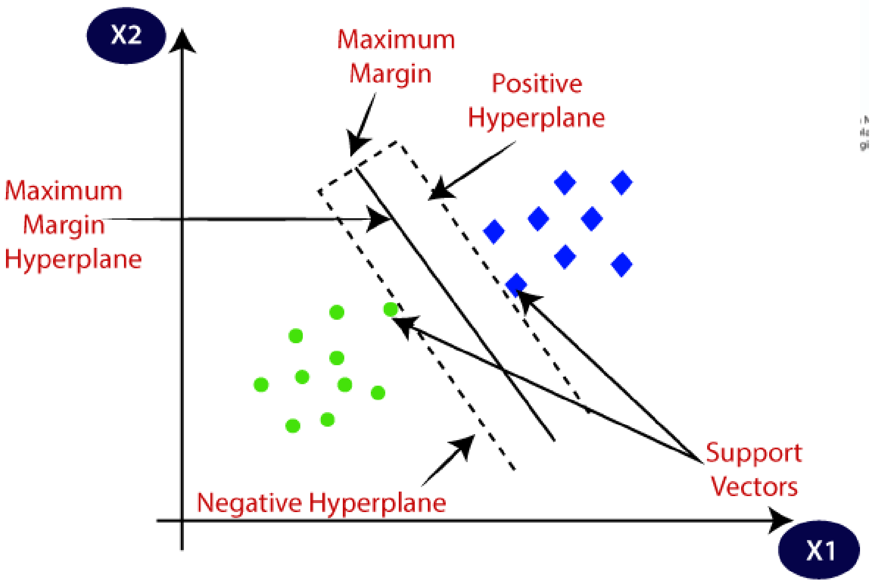


Fig. 8. SVM Classification [11]

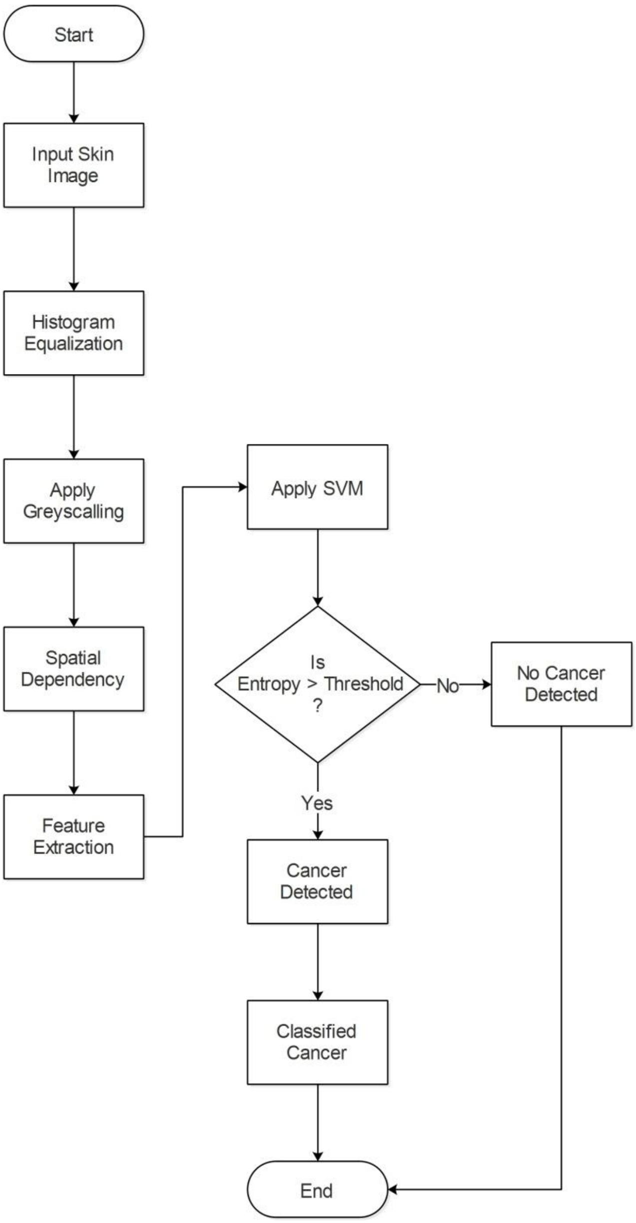


Fig. 9. Flow Chart

A. SVM Algorithm:

INPUT: A ← Input skin image

OUTPUT: H ← Absolute gradient magnitude

Step 1: Input 2-dimentional image as array

Step 2: Convert RGB image to grayscale

Step 3: Adjust contrast using histogram equalization

$$cdf_x(i) = \sum_j^0 P_x(j)$$

where cdf is cumulative distribution function, x is grayscale image, i is gray levels and P is probability

Step 4: Apply spatial dependencies i.e. convert it into square matrix

Step 5: Apply SVM by defining data points

$(x_n, y_n) \rightarrow (x_1, y_1).....(x_n, y_n)$

Step 6: Separate the data points by hyperplane

$\vec{w}.\vec{x}$ - b = 1

$\vec{w}.\vec{x}$ - b = 0

$\vec{w}.\vec{x}$ - b = -1

where $\vec{w}$ is the normal vector of the hyperplane

Step 7: Plotting Data

Step 8: Compute Entropy

Step 9: **if** Entropy > T_r **then**

Cancer Detected;

else

No Cancer Detected;

end else

end if

Step 10: Classify Cancer

Step 11: End

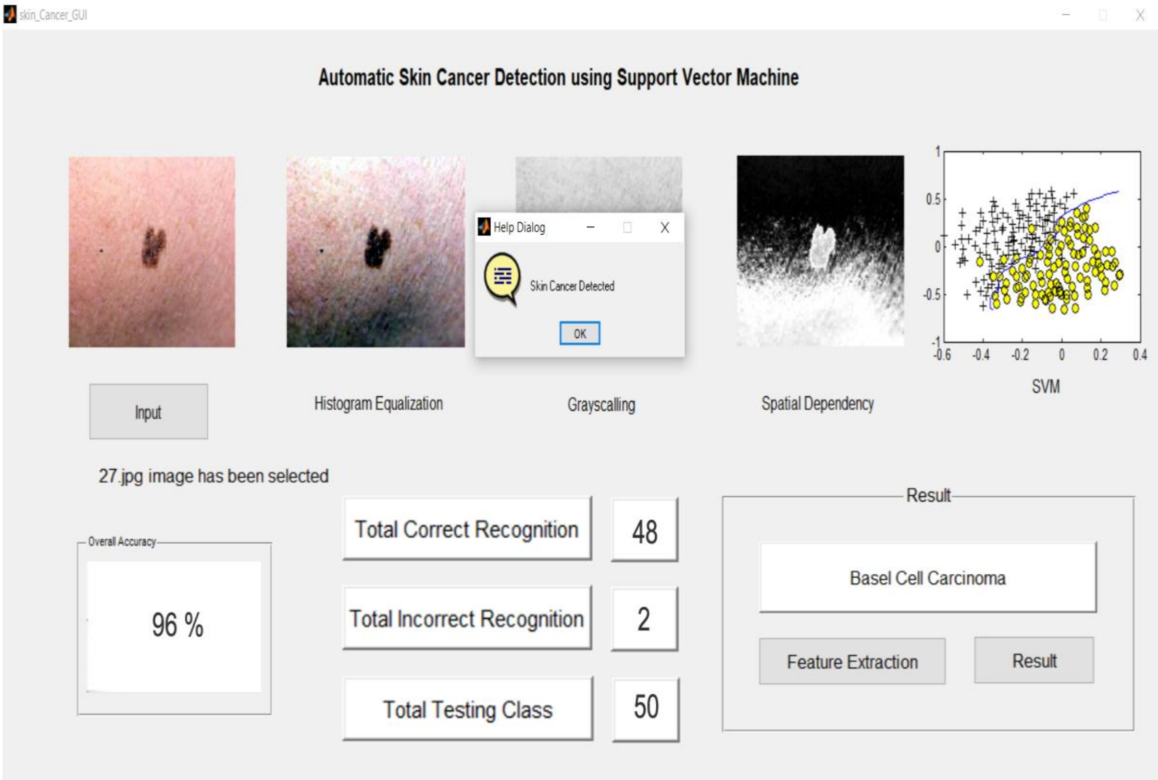


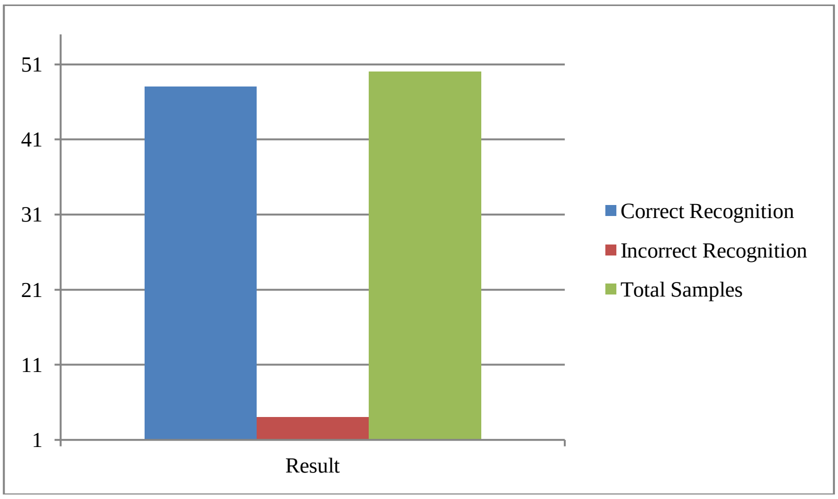
Fig. 10. Graphical User Interface

Classification of machines is a common function in machine learning. Suppose some given data points belong to one of every two classes, and the goal is to decide which class a new data point will be in. In the case of support-vector machines, a data point is viewed and can be separated with a 2 -dimensional hyperplane. There are several hyperplanes that can classify data. An appropriate choice as the best hyperplane is the one that represents the largest separation or margin between the two classes. Therefore we choose the hyperplane so that its distance from the nearest data point on each side is maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier defining it is known as the maximum-margin classifier; or equivalently, indicating optimal stability.

V. SIMULATION RESULTS

The simulation studies involve the various trails with distinct skin cancer images. There are total number of 50 trails where 48 trails recorded as correct recognition and 2 as incorrect that may includes true positive, true negative, false positive and false negative. True positive means that there are certain trails that positively detected which returns correct recognition and few images that may contain cancer but system is not able to detect; that entertained in the category of true negative. Similarly as false negative where image is detected as cancerous but actual there is no cancer in the image, whereas false positive means having no cancer image rejected positively. So, by observing all these datasets, the perceived accuracy is 96.00 %.

Graph 1 Result Analysis



Graph 1 shows the correct recognition, incorrect recognition and total number of samples tested by the system.

Table 1 Result Comparison

	Accuracy %
Md. Zahid Hasan [9]	95.00
Proposed	**96.00**

VI. CONCLUSION AND FUTURE WORK

Automatic skin cancer detection using support vector machine is a trending approach for diagnosing skin cancer effectively. It has been implemented using spatial dependencies and support vector machine. Most of the system uses machine learning methods to train the system with various samples. But a large dataset can consume the large amount of memory that increases the execution time where it is very important to communicate as earlier as possible with high level of accuracy. Here the system trained with various samples and able to recognize and classifies various skin cancers with relative features. System acquired 96.00 % of accuracy with less error rate or incorrectly recognition. In future; accuracy can be enhanced with various image enhancement technique and feature extraction method that later can be implemented for real time skin cancer detection.

REFERENCES

1. https://uwaterloo.ca/vision-image-processing-lab/research-demos/skin-cancer-detection
2. F. K. Nezhadian and S. Rashidi, "Melanoma skin cancer detection using color and new texture features," 2017 Artificial Intelligence and Signal Processing Conference (AISP), Shiraz, 2017, pp. 1-5, doi: 10.1109/AISP.2017.8324108.

3. P. M. Amulya and T. V. Jayakumar, "A study on melanoma skin cancer detection techniques," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 764-766, doi: 10.1109/ISS1.2017.8389278.
4. P. Bumrungkun, K. Chamnongthai and W. Patchoo, "Detection skin cancer using SVM and snake model," 2018 International Workshop on Advanced Image Technology (IWAIT), Chiang Mai, 2018, pp. 1-4, doi: 10.1109/IWAIT.2018.8369708.
5. S. Mane and S. Shinde, "A Method for Melanoma Skin Cancer Detection Using Dermoscopy Images," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697804.
6. Shalu and A. Kamboj, "A Color-Based Approach for Melanoma Skin Cancer Detection," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 508-513, doi: 10.1109/ICSCCC.2018.8703309.
7. N. B. Linsangan, J. J. Adtoon and J. L. Torres, "Geometric Analysis of Skin Lesion for Skin Cancer Using Image Processing," 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), Baguio City, Philippines, 2018, pp. 1-5, doi: 10.1109/HNICEM.2018.8666296.
8. G. Mansutti, A. T. Mobashsher and A. Abbosh, "Design of a Millimeter-Wave Near-Field Probe for Early-Stage Skin Cancer Detection," 2019 13th European Conference on Antennas and Propagation (EuCAP), Krakow, Poland, 2019, pp. 1-4.
9. M. Z. Hasan, S. Shoumik and N. Zahan, "Integrated Use of Rough Sets and Artificial Neural Network for Skin Cancer Disease Classification," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2019, pp. 1-4, doi: 10.1109/IC4ME247184.2019.9036653.
10. https://en.wikipedia.org/wiki/Support_vector_machine
11. https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm