Computer Aided Detection And Recognition Of Malignant Melanoma In Dermoscopic Images

Megha Biswas, Manjunatha HIremath

Abstract: Malignant melanoma is a very dangerous form of skin cancer. It is caused due to pigment -producing cells called melanocytes which mutate and becomes cancerous in nature. It is a very dangerous form of skin cancer as it spreads very fast. The number of estimated cancer deaths in 2020 is 630,000. The objective of current research study is to develop a recognition system to identify malignant melanoma cells from the affected part of the human skin. For current study authors have considered ISIC skin dataset. The proposed method consists hybrid segmentation method which combines the outcomes of two individual methods, namely, watershed and active contour method outcomes. The geometrical feature values of the segmented outcomes are extracted and stored as knowledge base. Support vector machine (SVM) is used for classification of malignant melanoma skin region from the dermoscopic image. The proposed method provides 72.60% recognition accuracy. The model can be improved and enhanced in future by considering and combining various learning methods in future.

Index Terms: Melanoma, Marker-Controlled Watershed, Active Contour Segmentation, SVM.

1 INTRODUCTION

The Skin cancer is considered as one of the deadliest syndromes to ever affect humans, which is not an easily curable disease. Failure in recognizing the effected region in the early stages can result in death. To avoid such medical emergency conditions, computer aided algorithms are proposed / developed to help in easy and early recognition of the affected region. Dermoscopy is one method used to diagnose skin lesions and the various associated conditions related to human skin [1]. Dermoscopy is considered as the method to diagnose all kind of skin cancer and the various associated conditions related to human skin. It has been observed that inexperienced dermatologists have a lower diagnostic accuracy. These issues have been rectified by computer-aided analysis. Skin malignant growth is seen as of different kinds, for example, melanoma, basal and squamous cells. Among these, melanoma is considered the most unpredictable in its prognosis. If detected in its early stages, there are higher chances that melanoma cancers can be cured. Computer visualizations play a major role in identifying these hazardous cells by capturing images of these skin cells [2]. On an increased analysis of these images (such as texture analysis, detection, shape analysis, etc.), one can gain a better understanding of the cancerous cells. A dermatologist may not always be readily available for patients who have skin cancer. Hence, computer-aided systems can fill in for these periods of absences and help detect any issues that may arise [3]. Proper health care treatments can be given to affected individuals if it is diagnosed early as well. Hence, there should exist a system, which can help in diagnosing the disease early on. This is where computer-aided diagnostics come into play. Computer-aided diagnostics can help save several lives if the proceedings applied to the affected regions correctly [4]. With the help of health records from the medical professionals and using the above-proposed model - early identification of disease is possible to carry on further medications. Many algorithms and techniques that detect these cancerous cells are already available. However, these methods have a high rate of error as well. The main focus of this model is to identify the affected regions more efficiently and classify them accordingly than the other available models.

2. RELATED WORK

Prashant Bhati and Manish Singhal [5] have proposed a system that helps to classify skin lesions as malignant or benign. The system is based on the Total Dermoscopic Score. The preparing and OTSU's division technique is utilized for the grouping of sores as non-malignant or dangerous.This along with the ABCD algorithm can lead to the result being accurate enough in classifying the skin lesion. Additionally, Farzam Kharaji, Nezhadian, and Saeid Rashidi proposed a new algorithm to categorize the dermoscopic images into malignant and benign. The snake's Algorithm method was used for segmentation purposes. From the segmented image, the color features and textural features were extracted. Texture-based features were analyzed first and it was seen that this method was highly efficient. By the use of the support vector classifier, they managed to get an accuracy of 97% (as was mentioned in the dataset of international skin imaging collaboration) [6]. M.Chaitanya Krishna, S.Ranganayakulu and Dr. P. Venkatesan proposed a methodology through the clustering technique and feature extraction of skin cancer. This Lesion Image analysis tools checked for the various Melanoma factors Like Asymmetry, Border, Color, Diameter (ABCD), etc. This was done by a surface, estimate and shape examination for image segmentation and highlight. The mined feature parameters are used to classify the image as Normal skin or Melanoma cancer lesion [7]. Pratik Dubal, Sankirtan Bhatt, Chaitanya Joglekar and Dr. Sonali Patil proposed a new procedure for Skin cancer detection and classification. The images are segmented, and features are extracted using ABCD rule. A neural network is used for the classification of the lesion to attain a higher degree of accuracy. This trained Neural Network achieved an overall classification accuracy of 76.9% on a dataset of 463 images (divided into six distinct classes) [8]. Abdol-al-Jalil Et Al identified skin cancer using the neural network. He used artificial intelligence and image processing techniques for the identification of skin cancer. The images were taken through the preprocessing methods by various image enhancement techniques in order to remove any undesirable objects and noises. This led to improved image quality. 2-D wavelet transformation was used to extract the features from the segmented image. Researchers used the extracted features as the input to the neural network and backpropagation neural network was used to classify the images into those that have the presence of melanoma and those which don't [9]. Hamd and Asa predicted skin cancer melanoma by treating the pigment. A digital method based on symmetric color pigments of the skin lesion was used. For

additional processing, the edge of the lesion was identified and segmented. The symmetry level of all the images was then calculated to identify the category of the tumor. The images of the skin lesions were classified into three groupings that were Melanoma, Basal cell carcinoma, and Squamous cell carcinoma. The accuracy of the results after classification using the first method is 80% and using the second method is 92.5%. The results can be increased if the segmentation of the skin lesion is increased [10].

3. PROPOSED METHODOLOGY:

The primary objective of the current study is to develop a computational system, which recognizes the affected skin melanoma regions. The entire process is divided into two major parts. The first part is training phase and second phase is Testing Phase. The block diagram of proposed work is presented in figure 1. In this, each phase consists of five stages in it, namely, data acquisition, preprocessing, segmentation, feature extraction and knowledge base. In the following sections these stages are explained in details.

Block Diagram:

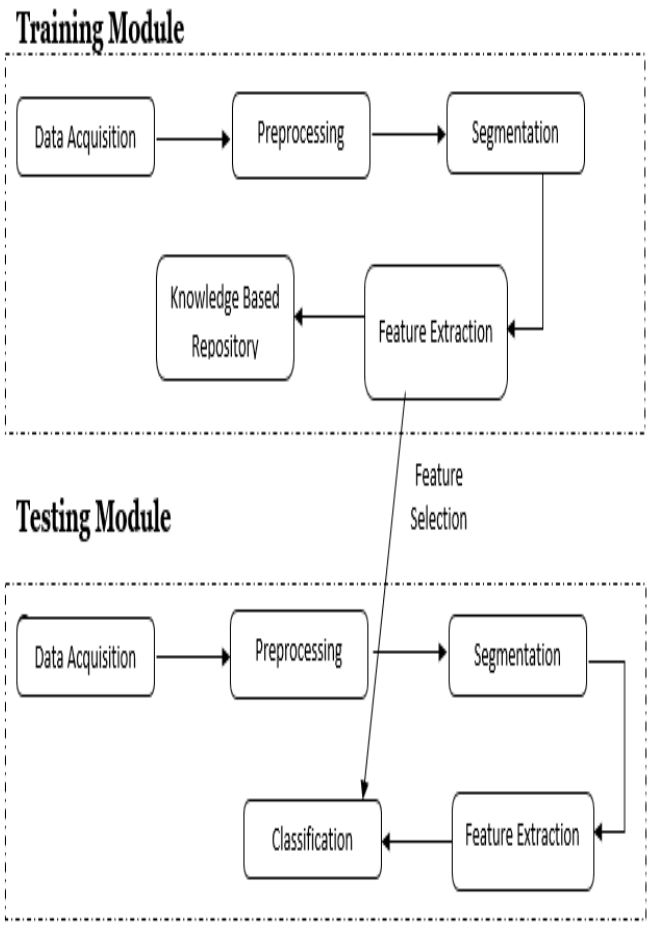

***Fig. 1**. Proposed research model – block diagram.*

Stage 1 - Data Acquisition:

This stage includes collecting the image data from ISIC archives and other archives. There is a need to check whether the collected images are a combination of benign and melanoma cells. Digital images obtained using photo dermatoscope have adequate resolution that allows for a precise analysis in terms of differential structures appearance [11]. A dermatologist can make exact documentation of the accumulated images, opening a way for PC investigation, where images are regulated so as to separate data that can later use to order those images. The dataset flooded in the ISIC archive has been used by many researchers for developing and proposing various models to identify the cancerous cells.

Stage 2 - Pre-Processing

Pre-processing is to be performed on every one of the images procured. This movement is applied so as to ensure that every one of the images are dependable in the ideal trademark. The collected dermoscopic images have to be pruned for noise removal. The generic procedures that are highly applied through the image pre-processing technique, that are of conversion of the image from the color format to a grayscale image, which makes the extraction of needful features from the converted images, a helping hand for further segmentation proceedings. Pre-processing techniques such as resizing, equalization and morphological processing can be applied [12]. So, implying more than one pre-processing technique over the collected image generally helps to classify the lesion either as melanoma or benign.

Stage 3 - Segmentation

Segmentation helps to segment the region of interest from the pre-processed image. There is continuous regression of segmentation algorithms that have been imposed on the images, to acquire a proposed model. The algorithms such as the Active Contour, Marker controlled watershed algorithm and a combination of both algorithms. The marker-controlled watershed division has been demonstrated to be a hearty and adaptable technique for the division of articles with shut forms, where the limits are communicated as edges. The marker pictured utilized for the watershed division is a paired picture comprising of either single marker focuses or bigger marker areas, where each associated marker is set inside an object of intrigue. Each underlying marker has a coordinated relationship to a particular watershed area, in this way, the number of markers will be equivalent to the last number of watershed locales. After division, the limits of the watershed districts are masterminded on the ideal edges, in this way isolating each article from its neighbors. The markers can be physically or naturally chosen, yet high throughput analyzes frequently utilize consequently produced markers to spare human time and assets. Representing the individuality among the algorithms gives better refined segmented regions, but the proposed model implies to attain a more refined region from the hybridization of the algorithms. The generic ideology behind using the segmentation technique is to achieve the border of the skin lesions.

Stage 4 - Feature Extraction

Feature extraction is considered as the crucial step that helps towards the classification. Based on the extracted features the classification of the images is made out to rectify whether the lesions are cancerous or non-cancerous. The features can be extracted using geometrical parameters. The geometrical

parametric features include:

- Area
- Major Axis
- Minor Axis
- Roundness.
- Perimeter.
- Totality
- Eccentricity
- Length.
- Compactness.

Stage 5 - Classification

Classification plays a major role in identifying the lesions by categorizing the extracted regions. The classification algorithm being used is SVM. The more accurate value is obtained from the classifier which helps for easy identification of the lesions from the affected regions of the skin.

Stage 6 - Knowledge Repository

A knowledge-based repository is constructed from the features being extracted. These features are being stored as knowledge for further data interpretations and enhancements. These values are being cross verified with testing parameters chosen from the single images. The Knowledge base helps to improve the efficiency of the features being extracted. Repository helps for easy identification of the image.

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the proposed technique is explained in detail. The model briefly works on the two segmentation techniques and the hybrid method of those algorithms, which marked as a third proposed algorithm. The segmentation techniques are followed by feature extraction and classification algorithms. The classification algorithm consideres of SVM algorithm. There are similar techniques being used to find out the cancerous cells but hybridized model are a handful.

Algorithm 1: Training Method

Input: Dermoscopic Image
Output: Malignant Melanoma cell features

Step 1: Input the training image dataset T_n to proposed model.
Step 2: Apply preprocessing method to remove noise exist in the dermoscopic image.
Step 3: Implement Active contour segmentation method followed by Marker controlled watershed method.
Step 4: Superimpose the outcomes of both segmentation methods.
Step 5: The geometrical feature values such as area, eccentricity, perimeter, major axis length, minor axis length, length- width ratio, roundness are extracted and stored as knowledgebase.
Step 6: Store the feature set as knowledgebase.

Algorithm 2: Testing Method

Input: Dermoscopic Image
Output: Malignant Melanoma cell recognition

Step 1: Input the Testing image dataset T_m to proposed model.
Step 2: Apply preprocessing method to remove noise exist in the dermoscopic image.
Step 3: Implement Active contour segmentation method followed by Marker controlled watershed method.
Step 4: Superimpose the outcomes of both segmentation methods.
Step 5: The geometrical feature values such as area, eccentricity, perimeter, , major axis length, minor axis length, length- width ratio, roundness are extracted and stored as knowledgebase.
Step 6: The malignant cell recognition is achieved by comparing the test feature with the knowledge base.

The below figure shows the sample training images:

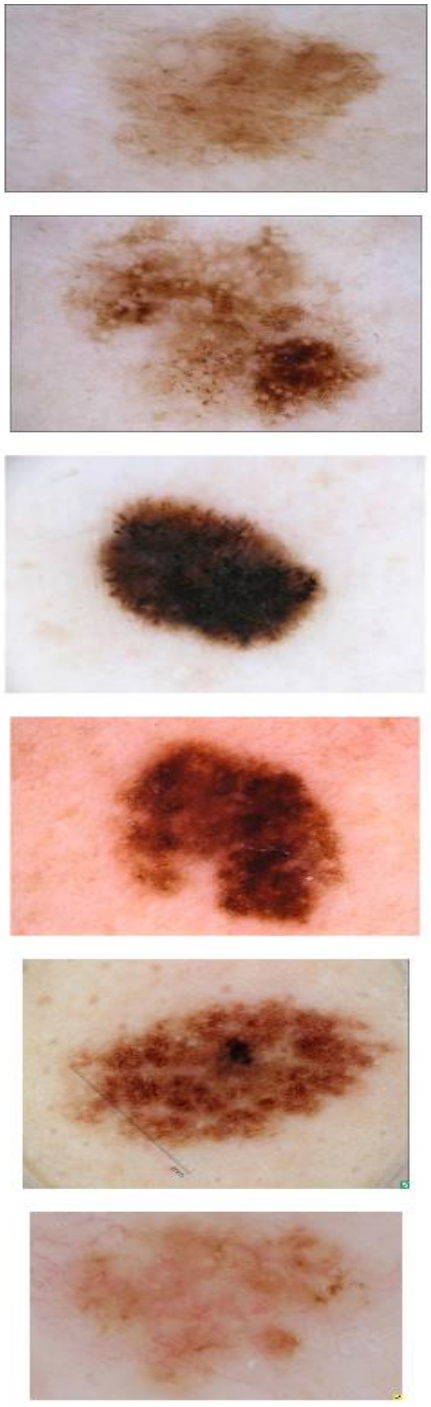

Figure 2: *Sample training images of proposed model*

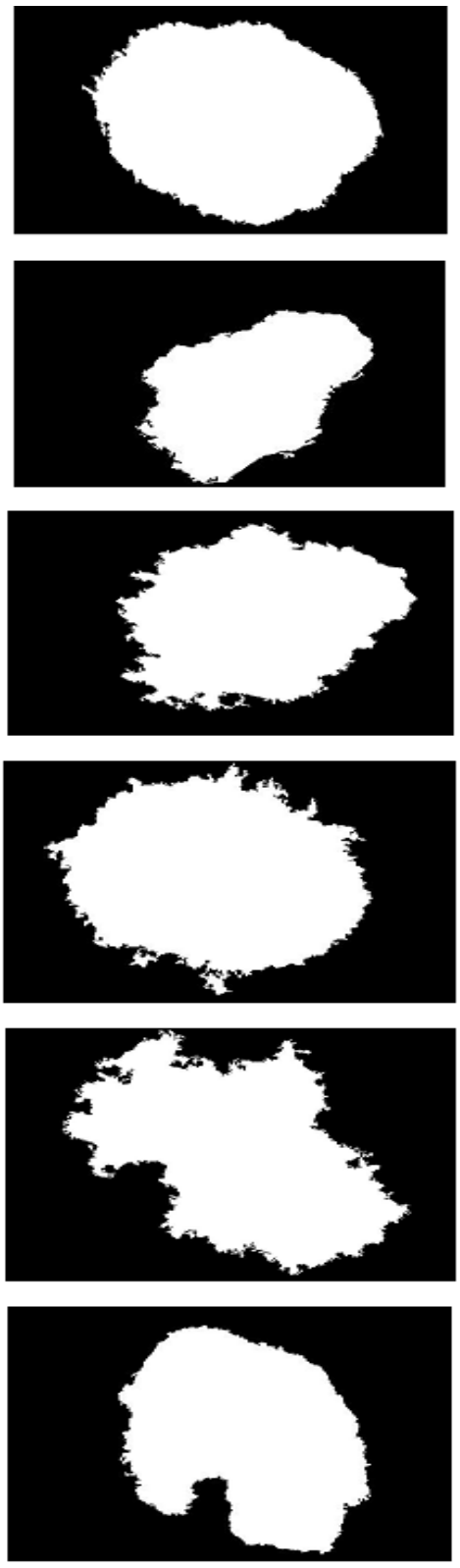

Figure 3: *Segmented images of proposed model*

The 9 features extracted from the above images are as follows:

Table 1*: Extracted Feature set values of proposed method*

Area	Mj. Axis	Mn. Axis	Eccentricity	Perimeter	Roundness	Totality	Length	Compactness
440.31557	302.54664	103208	0.726551	1738.5177	0.429107	0.253271	1.455366	2.320422
282.41244	248.13155	53985	0.477533	1390.6488	0.350791	0.203088	1.138156	2.85007
206.71911	141.70166	22846	0.728093	793.667	0.455768	0.260461	1.458834	2.194101
317.50611	250.73522	59582	0.613499	2814.4.392	0.945227	0.112815	1.26663	10.57899
318.07222	257.84977	63882	0.585512	1411.6077	0.402866	0.225326	1.233557	2.482214
246.39445	214.042	387291	0.495345	2156.298	0.104672	0.114267	1.151115	9.553696
580.43488	366.10299	16198 2	0.775995	4653.2477	0.094008	0.124738	1.585442	10.63744

```
In [15]:
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.svm import SVC
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix
from sklearn.metrics import accuracy_score
```

```
In [37]:
data = pd.read_csv(r"C:\Users\Megha\Desktop\final m images\mimages.csv",header = None)
```

```
In [38]:
data.head()
```

Out[38]:

	0	1	2	3	4	5	6	7	8	9
0	338.664498	297.406303	75237	0.478294	2863.110	0.133310	0.127165	1.138893	7.501306	1
1	416.450376	394.426642	127214	0.320893	2286.833	0.305686	0.182108	1.055837	3.271328	1
2	36.506311	20.666376	560	0.824334	112.111	0.559888	0.326626	1.766459	1.788070	0
3	240.268879	147.682244	27160	0.788797	1385.959	0.177880	0.173359	1.828931	5.828091	1
4	612.297360	481.825310	220869	0.866587	3554.340	0.219898	0.172288	1.325820	4.551697	1

```
In [39]:
X = data.values[:, 0:9]
Y = data.values[:,9]
```

```
In [40]:
X_train, X_test, y_train, y_test = train_test_split( X, Y, test_size = 0.3, random_state = 42)
```

```
In [41]:
svc_model = SVC()
```

```
In [42]:
svc_model.fit(X_train, y_train)
```

```
Out[42]: SVC(C=1.0, cache_size=200, class_weight=None, coef0=0.0,
  decision_function_shape='ovr', degree=3, gamma='auto_deprecated',
  kernel='rbf', max_iter=-1, probability=False, random_state=None,
  shrinking=True, tol=0.001, verbose=False)
```

```
In [43]:
y_predict = svc_model.predict(X_test)
```

```
In [44]:
cm = np.array(confusion_matrix(y_test, y_predict, labels=[1,0]))
confusion = pd.DataFrame(cm, index=['is_melanoma','is_unknown'],
                         columns=['predicted_malignantcells','predicted_unknown'])
confusion
```

Out[44]:

	predicted_malignantcells	predicted_unknown
is_melanoma	31	0
is_unknown	17	10

```
In [35]:
print(accuracy_score(y_test,y_predict)*100)
```

```
72.41379310344827
```

Figure 4 : *Sample screenshot of experimental result code, which depicts SVM execution*

The above SVM algorithm gives an accuracy of 72.6%.

4 CONCLUSION

Malignant melanoma is considered dangerous in nature as maximum amount of death is caused by it. Early stage recognition can save lives. Image Processing and computer aided diagnostics has earned more modeling and greater importance in the modern world of medical diagnosis. Image processing techniques are used for recognition of malignant melanoma skin cells presence. A recognition system to differentiate and identify malignant cells from the affected part of the skin, so that it can classify the presence of the cancerous cells.

REFERENCES

[1]. Farzam Kharaji Nezhadian and Saeid Rashidi, "Melanoma skin cancer detection using color and new texture features", artificial intelligence and signal processing.

[2]. M.H.Hamd,K.A.EssaandA.Mustansirya,"SkinCancerPrognosisBasedPigment Processing", International

Journal of Image Processing, vol. 3, no. 7, pp.227,2013.

[3]. M.Chaithanya Krishna 1 , S.Ranganayakulu 2 , DR.P.Venkatesan3, "Skin Cancer Detection and Feature Extraction through Clustering Technique Vol.

4, Issue 3, March 2016”
[4]. Prashant Bhati, Manish Singhal, “Early Stage Detection and Classification of Melanoma”
[5]. Pratik Dubal, Sankirtan Bhatt, Chaitanya Joglekar, Dr. Sonali Patil, “Skin Cancer Detection and Classification”

[6]. J.A. Jaleel, S. Salim and R.B. Aswin, “Artificial neural-network-based detection of skin cancer”, International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, vol. 1, no. 1, pp. 200-205, 2012.
[7]. An Introduction to Digital Image Processing with Matlab Notes for SCM2511 Image Processing 1 Alasdair McAndrew School of Computer Science and Mathematics Victoria University of Technology.
[8]. Sagar Adatrao, Mayank Mittal, "An Analysis of Different Image Preprocessing Techniques for Determining the Centroids of Circular Marks Using Hough Transform".
[9]. R.M. Nishikawa, M.L. Giger, K. Doi, C.J. Vyborny and R.A. Schmidt. 1995. "Computer aided detection of clustered microcalcifications in digital mammograms," Med. Biol. Eng. Comp. Vol. 33, pp. 174-178.
[10]. Chandrahasa M, Varun Vadigeri and Dixit Salecha,"Detection Of Skin Cancer Using Image Processing Techniques " International Journal of Modern trends in engineering and research
[11]. DermLite, 2017a. https://dermlite.com/. (Accessed 20 July 2017).
[12]. Erdem Okur a, Mehmet Turkan, “”A survey on automated melanoma detection" Engineering Applications of Artificial Intelligence 73 (2018) 50–67.