Integrated Intelligent Research (IIR)
International Journal of Computing Algorithm
Volume: 08, Issue: 01, June 2019, Page No.22-26
ISSN: 2278-2397

Skin Lesion Detection and Segmentation using SVM and Bayesian Classifier

Anisha Thangakani S[1], Sornam M[2]
[1]Research scholar, Department of computer science, University of Madras, Chennai
[2]Professor, Department of computer science, University of Madras, Chennai
Email : anishakani014@gmail.com, madasamy.sornam@gmail.com

Abstract - Skin cancer is a peculiar growth of skin cells; often develops on skin exposed to the sun. Skin cancer evolves predominantly on areas of sun-exhibited skin includes scalp, face, lips, ears, neck, chest, arm, hands and the legs in human. But it may appear on the areas that seldom see the sunshine of day like palms, at a lower place fingernails or toenails and reproductive organ space. However, it is curable if it gets detected at an early stage. To minimize the diagnostic error caused by the complexity of visual interpretation and subjectivity, it is important to develop a technology like computerized image analysis. This article presents a systematic approach for the classification of skin lesions in dermoscopic images. For detecting the melanoma skin cancer, the Otsu algorithm is applied to segment the lesion from the entire image. In the proposed method, feature extraction is performed by underlying shape, color, texture using GLCM (Gray-Level Co-Occurrence Matrix) features which are applied with support vector machine and Bayesian classifiers. While comparing, the Bayesian classifier achieved an overall classification accuracy of 90.44% on a dataset of 4151 images which is better than the existing method.

Keywords: Melanoma, benign, skin cancer, svm classifier, Bayesian classifier, GLCM features.

I. INTRODUCTION

Cancer refers to an abnormal growth of tissues, in which skin cancer is most common. It may of two types: Benign and malignant. Malignant tumours have highest mortality rate when compared with benign, due to its significant feature called metastasis. Metastasis is nothing but the ability to invade the faraway organs. Various diagnostic features are there to distinguish both. The most common method to distinguish both is the histological method, in which the affected tissues are taken and viewed under microscope. As it is an invasive procedure and chances of spreading to other normal underlying structures, other methods of diagnosis came into play. This work is one among them, which is a non invasive procedure and detected under computer vision by algorithms. To classify whether the given lesion is malignant or benign, the feature extraction is used texture features, color features and shape feature using GLCM features

- **Texture feature-** one half doesn't match the appearance of the other half.
- **Color features**- The color is not uniform. Shades of tan, brown, and black are presented dashes of red, white, and blue augment the patterned.
- **Shape features**- The edges are ragged, notched, blurred

The article is organized as follows: In section II, a general outline of the research related work is given. In section III, The proposed method IV followed by the corresponding result and evaluation criteria are described V. Finally, the conclusion of the research work is presented.

II. RELATED WORKS

Raman deep [12] proposed method uses the adaptive filter to find the threshold inspired by Swarm Intelligence (SI) optimization algorithms. The image processing techniques are implemented in first three streams: The grayscale image convert the texture feature. And then the RGB image [7] is used to compute the color moments like mean and standard deviation in the red, green and blue color domain. The k-means is used to extract the infected part and then local binary patterns are computed by threshold of segmented part using the [6] Otsu algorithm. The work results show the improvement in identifying the skin cancer melanoma at different stages based on textural feature analysis and SVM classifier and different stages of the Melanoma skin cancer. However, the SVM classifier takes the protect a good extent. Where the skin expert may get required information at fine accuracy. Krishnan [9] described the skin and faces for color pixel classification, some algorithms are proposed for skin color pixel classification. They included piecewise linear classifiers, in histogram technique used a Bayesian classifier, Gaussian classifiers and the Multi-layer perceptron. A dataset of

XM2VTS face is used to inspect the selection of color space can enhanced the density of the skin discriminability between [8] skin and non-skin class in 13 color spaces and 6 different skin classification for skin pixel algorithms An perusal of the pixel-wise skin segmentation approach that uses color pixel classification algorithms is presented. Multi-layer perceptron classifier produces positive results. The histogram technique with Gaussian classifier had a maximum CDR of 99.10 percent. But the Bayesian classifier [11] produces good results with 89%. The HSV and RGB color space with piecewise linear classifiers produce still higher classification rates than all other classifiers with good illumination conditions for an image.

In the case of melanoma, the characteristics are summarized by ABCD parameter [1] to detect the melanoma cancer. Most of the existing method extracted the features like Asymmetry, border, color and diameter. But the computations of these features were complex and were not much efficient.

III. PROPOSED METHOD

The proposed system is to classify benign and malignant skin lesion images by applying GLCM feature extraction method using [12] ISIC datasets and comparing with SVM and Bayesian classifiers.

A. Dataset Description

The ISIC standard for International Skin Imaging Collaboration is a global effort to strengthen the malignant melanoma identification. Recently [12], ISIC efforts to collect an accessible dataset for checking the image of the skin lesion by providing publicly available database ISIC skin image data archive. Images are collected from internationally influential clinical centers, obtained from different devices used in each center. In this work, we used 2,247 images for training and tested with 1,904 images.

B. Preprocessing

Before the images can be segmented, they need to be preprocessed and cleaned to better the performance of the segmentation algorithm. First, the image is converted to grayscale. And then The homomorphic filter consists of performance an fft on the image and then it passes through a high pass filter [1]. Gaussian blur is applied to an image with a Gaussian function. This is referred to as a two- dimensional Weierstrass model. By contrast the bokeh effect, is used to reproduce a convolving by a circle would be more accurate. Since the Fourier model of a mathematician is a another mathematician applying a mathematician blur has the impact of reducing the image's high-frequency components; a mathematician blur is, therefore, allowing pass filter to remove any presence of salt and pepper noise by smoothing.

C. Segmentation

The cleaned image is then segmented using Otsu's thresholding algorithm, a global Thresholding algorithm [1, 2]. Otsu's algorithm differentiates between the image background and the image foreground by thresholding the image above a computed grayscale value. Otsu's algorithm requires us to a gray level histogram of an image before running the segmentation algorithm. Otsu's algorithm is aims to appear a threshold value [9] which realizes the intra-class variance and two classes is defined as weighted sum of variance as:

$$\sigma^2_{\omega} = \omega_0 (t) \sigma2_0(t) + \omega_1 (t) \sigma2_1(t) \qquad \dots (1)$$

Weights ω_0, ω_1 are the class probabilities separated by a t is defined as threshold value and σ^2_0, σ^2_1 are the variances [7] of these two classes. The image is then finalized above the threshold value and morphological operations are performed.

D. Feature Extraction

After the image segmentation process, the feature extraction process would be carried out to perform classification of the acquired segmented skin lesion. In this system, [3] we extract the GLCM features.

i. Shape Features: Melanoma is usually irregular, or not symmetrical in form. Benign moles are usually symmetrical. Based on whether the shape features [9] are extracted from the contour only or extracted from the complete shape region. The shape features are utilized as table.
ii. Color Features: The presence of further than one color like blue, black, white, red, light and brown, blue-gray. Or the rough distribution of color at time is a warning sign of melanoma. Benign moles are often a single shade of brown or tan. So, the lesion color variability can be quantified by identifying the occurrence of those color variability within the segmented lesion image. For the features to extract for the color variation in the lesion, we used RGB color channel. The first color moment can be interpreted as the

Volume: 08, Issue: 01, June 2019, Page No.22-26
ISSN: 2278-2397

average color in the image; the second color moment is the standard deviation, which is obtained by testing the square root of the variance of the color distribution.

iii. Texture Features: GLCM features are an efficient method to research the options of texture. Compared with a normal method. Textures of different diseases in the skin lesion image can be obtained, such as contrast, correlation, inverse_difference_moments, entropy etc,

Table 1: Shape Features

Features	Description
F1	Area: The ratio between the lesion (A)
F2	Perimeter: The ratio between the lesion (P)
F3	Physiological length: The length of the lesion(h)
F4	Physiological width: The width of the lesion(w)
F5	aspect ratio: float(w)/h
F6	Rectangularity: w*h/area
F7	Circularity: 4*3.14*area/((perimeter)**2)

Table 2. Color Features

Features	Description
F8	mean_r: The mean of red is μ_r
F9	mean_g: The mean of green is μ_g
F10	mean_b: The mean of blue is μ_b
F11	stddev_r: The standard deviation is σ_r
F12	stddev_g: The standard deviation is σ_g
F13	'stddev_b: The standard deviation is σ_b

Table 3.Texture Features

features	Description
F14	$contrast = \sum_{i,j=0}^{n=1} \mathrm{pij}(\mathrm{i}-\mathrm{j})$
F15	$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$
F16	$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{ij})P_{ij}$
F17	$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2}$

E. Feature Selection and Optimization

The method which trains the model is known as training algorithm. The gradient descent algorithm [7] is difficult if training data is too large. Thus, stochastic version algorithm used. To propel the use of stochastic optimization algorithms, often write deep learning models for training data, basically it achieve the target function as a sum of a finite number :

$$f(x) = a_0 + \sum \mathrm{i} = 1\mathrm{nfi(x)} \qquad \ldots (2)$$

Where fi(x) is a loss function. Hence, once n is too large, then per-iteration mathematical cost of gradient descent is very highly used. Considering, it offers a lighter-weight solution. At every iteration, compute the gradient ∇f(x), stochastic gradient descent with sample value (i) and measures ∇fi(x) instead.

F. Classification

In recent years, many classification algorithms exist to classify segmented images. In this work, two classifiers were used: SVM and Bayesian Classifier is used to classify the skin lesion as benign or malignant. Support vector machines used for regression and classification problems. In this it is used for classification. SVM is famous for kernel trick. The kernel calculates the two vectors as a dot product in certain feature space; kernel functions are sometimes called “generalized dot product”.

Volume: 08, Issue: 01, June 2019, Page No.22-26
ISSN: 2278-2397

$$\sum_{i=1}^{n} \propto i - 1/2 \sum_{j=1}^{n} \propto i \propto j \, K(X_i^T . X_j) \qquad \ldots (3)$$

Gaussian RBF is another Kernel method used in SVM models. RBF kernel is a function whose value confide on the distance from the specific point and parameter is called gamma (if increase over fit, if decrease under fits) and c (transposed of the strength of regularization) .Gaussian Kernel is of the consequent format:

$$K(X_{1,} X2) = \text{exponent } (-\gamma \|X_1 - X_2\|^2) \qquad \ldots (4)$$

Bayesian classifier is applied to achieve an efficient classification for large dataset. It was very useful for statistical classification. The basic idea of a Bayesian classifier [4, 5] is that it assumes any one feature of a classifier is that it is not related to any of another feature; here the system classifies an image into extracted features. Working of Bayesian classifier is based on Bayes theorem. According to Bayes theorem, posterior probability can found which is given by the equation.

$$P (C|X) = P (X|C) P(C) / P(X) \qquad \ldots (5)$$

Where,
P(X) is a prior probability of predictor,
P(C) is predictor probability of class and
P (X|C) is the like- hood which is the probability of predictor given class.

IV. RESULT

The image dataset was split into training and testing dataset. For SVM Classifier, a different type of the kernel was used for accuracy. Performance metrics are shown in Table 4.

Table 4: Accuracy and performance metrics of SVM

Different kernel	accuracy	precision	recall	F1- score	Support
RBF	68.14	0.71	0.25	0.37	121
linear	67.95	0.68	1.00	0.81	564
Linear, RBF	70.98	0.25	1.00	0.40	1

In Bayesian classifier, comparison was carried out by dividing the dataset into training and testing dataset. 2,247 training images were used and tested with 1,904 test images for various categories of benign and melanoma characteristics. With an overall accuracy are 90.44% and performance metrics for the dataset from [12] are shown in Table 5.

Table 5: Accuracy and Performance metrics of Bayesian classifiers

dataset	accuracy	precision	recall	F1-score	support
Benign [12]	90.44	0.91	0.99	0.95	1892
Malignant [12]		0.04	0.01	0.01	183

V. CONCLUSION

In this research work, segmentation and classification of skin lesion images were performed to help the patient to identify the skin cancer without going to hospitals. The diagnosis research work included segmentation with Otsu algorithm features extracted using GLCM features. Classification using SVM classifiers with different kernels along with Bayesian classifier were compared. Bayesian classifiers give much better result with the highest accuracy of 90.44%.

Reference

[1] Dubal, P., Bhatt, S., Joglekar, C., & Patil, S. (2017, November). Skin cancer detection and classification. In 2017 6th international conference on electrical engineering and informatics (ICEEI) (pp. 1-6). IEEE..

International Journal of Computing Algorithm
Volume: 08, Issue: 01, June 2019, Page No.22-26
ISSN: 2278-2397

[2] Soumya, R. S., "Advanced earlier melanoma detection algorithm using color correlogram." 2016 International Conference on Communication Systems and Networks (ComNet). IEEE, 2016.

[3] Lynn, Nay Chi, and Zin Mar Kyu. "Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images." 2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT). IEEE, 2017.

[4] Esteva, Andre, "Dermatologist-level classification of skin cancer with deep neural networks." Nature 542.7639 (2017): 115.

[5] Nguyen-Trang, T. (2018). “A New Efficient Approach to Detect Skin in Color Image Using Bayesian Classifier and Connected Component Algorithm”. Mathematical Problems in Engineering, 2018.

[6] Noboyuki Otsu, “A Threshold Selection Method from Gray-Level Histogram”,, IEEE Transactions On Systems, Man, And Cybernetics, Vol. SMC-9, No. 1, January 1979.

[7] Sagar, Chiranjeev, and Lalit Mohan Saini. "Color channel based segmentation of skin lesion from clinical images for the detection of melanoma." In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), pp. 1-5. IEEE, 2016.

[8] R. Moussa, F. Gerges, C. Salem, R. Akiki, O. Falou, and D. Azar,” Computer-aided detection of Melanoma using geometric features”, 2016 3rd Middle East Conference on Biomedical Engineering (MECBME), Beirut, 2016, pp. 125-128.

[9] Miss Hetal J. Vala and Prof. Astha Baxi,” A Review on Otsu Segmentation Algorithm”, International Journal of Advanced Research in Computer Engineering and Technology, Vol. 2, Issue 2, February 2013

[10] Dan Foresee, F., & Hagan, M. T. (1997, June). “Gauss-Newton approximation to Bayesian learning”. In International Conference on Neural Networks (Vol. 3, pp. 1930-1935).

[11] Anisha Thangakani S , M Sornam, “A Survey On Skin Classification Methods Using Deep Learning Techniques”, International Journal of Research in Advent Technology (IJRAT), Special Issue, March 2019E-ISSN: 2321-9637

[12] International Skin Imaging Collaboration Website http://www.isdis.net/index.php/isic-project