Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7765920

ANALYSIS AND PREDICTION FOR SKIN CANCER IMAGE DATASET USING IMAGE PROCESSING AND MACHINE LEARNING APPROACHES

[1]Vijay Arumugam R and [2]Saravanan S
[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India
Email: vijaynew@gmail.com

[2]Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India
Email: aucissaran@gmail.com

Abstract:

Now a days, skin cancer is the most common disease of death amongst human community. Skin cancer is exponential abnormal growth of skin cells that develops on the entire body exposed to the sunlight. Most of the skin cancers are remediable at beginning stages. In this case, the early analysis and prediction with fast detection of skin cancer to protect the human life. In this paper, we have proposed a skin cancer identification system using different soft computing techniques which are used to detect skin cancer disease in early stages. Diagnosing methodology uses image processing techniques, several machine learning and soft computing techniques for classification which includes SVM, ANN, KNN, and ensemble classifiers. The raw image of skin cancer is taken under various pre-processing techniques which are used to noise removal, enhancement and segmentation. Some features of the image have to be extracted using Gray Level Co-occurrence Matrix (GLCM) with features given as the input to classifier. SVM, ANN, KNN, and ensemble classifiers are used for classification and given image into skin cancer or normal. Numerical illustrations were also provided to prove the results and discussions.

Keywords: Grayscale image, image filter, image enhancement, image segmentation, prediction, and machine learning.

1.0 Introduction

The new technology, early detection of skin cancer is possible at the initial stage. Formal method for diagnosis skin cancer detection is Biopsy method [1]. It is done by removing skin cells and that sample goes to various laboratory testing. It is a painful and time-consuming process. Skin cancer is a deadly disease. Skin includes three types of layers with skin cancer beginning in outermost layer, which is made up of first layer squamous cells, second layer basal cells, and innermost or third layer melanocytes cell. Squamous cell and basal cell are sometimes called non-melanoma cancers. Non-melanoma skin cancer always responds to treatment and rarely spreads to other skin tissues. Melanoma is more dangerous than most other types of skin cancer [2]. If it is not detected at beginning stage, it is quickly invade nearby tissues and spread to other parts of the body. Formal diagnosis method to skin cancer detection is Biopsy method. A biopsy is a method to remove a piece of tissue or a sample of cells from a patient’s body so that it can be analyzed in a laboratory. It is an uncomfortable method. The Biopsy Method is time consuming for patient as well as doctor because it takes lot of time for testing. Biopsy is done by removing skin tissues (skin cells) and that sample undergoes series of laboratory testing [1]. There is a possibility of spreading disease into other parts of the body. It is riskier. Considering all the cases mentioned above, Skin cancer detection using SVM is proposed. This methodology uses digital image processing technique and SVM for classification. This technique has inspired the early detection of skin cancers and requires no oil to be applied to your skin to achieve clear sharp images of your moles. In this way, it's a quicker and cleaner approach. But, most importantly, due to its higher magnification, Skin Cancer Detection Using SVM can prevent the unnecessary excision of perfectly harmless moles and skin lesions.

2.0 Literature Review

Skin detection based on Maximum Entropy Threshold, feature extracted by using Gray Level Co-occurrence Matrix (GLCM), and c1assification using Artificial Neural Network (ANN). Back-Propagation Neural (BPN) Network is used for c1assification purpose [1]. The segmentation as various clustering technique, features can be extracted by using ABCD (Asymmetry Index Border Colour Index Diameter) method [3]. The system used rule based and forward chaining approach to detect skin disease. The authors explain the proposed system enables user to identify children skin diseases via online and provide useful medical suggestions. Used different data mining classification algorithms (AdaBoost, BayesNet, MLP and NaiveBayes) to predict and diagnose the skin disease. This only works for three skin diseases (Eczema, Impetigo and Melanoma) [4].

Various researchers have used various Data Pre-processing methods, Disease Diagnosis, Maximal Frequent Itemset Algorithm for training, K-means clustering for segmentation and significant frequent pattern for classification [5]. The research works the authors explain different method for melanoma diagnosis applied on a set of digital images. Features extracted by using Gray Level Co-occurrence Matrix (GLCM) and Using Multilayer Perceptron Classifier (MLP) to classify between cancerous and noncancerous images [6].

The arrangement was supervised love the predefined categories of the kind of carcinoma. Combining Self organizing map (SOM) and radial basis perform (RBF) for recognition and diagnosing of carcinoma is far and away higher than KNN, Naive Thomas Bayes and ANN classifier. The most effective classification accuracy (88%, 96.15% and 95.45% for Basal cell malignant neoplastic disease, malignant melanoma and epithelial cell malignant neoplastic disease respectively) were obtained victimization combining Kyrgyzstani monetary unit and RBF. the general classification accuracy was ninety-three 15% [7].

The diagnosing methodology uses Image processing methods and Support Vector Machine (SVM) algorithm. The dermo copy image of skin cancer is taken and it goes under various pre-processing technique for noise removal and image enhancement. Then the image is undergone to segmentation using Thresholding method. Some features of image have to be extracted using GLCM methodology. These features are given as the input to classifier. Support vector Machine (SVM) is used for classification purpose. It classifies the given image into cancerous or non-cancerous [8]. Explain different types of image processing techniques like filtering concepts namely gray level and median filter simply. The authors explain different pre-processing techniques in the image processing area [9].

Algorithms for common image processing applications have been developed that support sophisticated image processing without requiring an extensive background in mathematics. Authors fully updated with the newest of these, including 2D vision methods in content-based searches and the use of graphics cards as image processing computational aids. It's an ideal reference for software engineers and researchers [10]. The researchers clearly explain the utmost importance to early detect the skin cancers. Proper diagnosis is critical for the survival of the patient. Biopsy method of detection is much painful. We have proposed an automated system for detection and classification of one of the skin four types of skin cancers: Melanoma, Basal cell carcinoma, actinic Keratosis, Squamous cell carcinoma. There are certain features of these types of skin cancers, which can be extracted using proper feature extraction algorithm [11].

3.0 Proposed System

Skin cancer detection using SVM, ANN, KNN, and ensemble classifiers is basically defined as the process of detecting the presence of cancerous cells in image. Skin cancer detection is implemented by using GLCM and SVM, ANN, KNN, and ensemble classifiers. Gray Level Co-occurrence Matrix (GLCM) is used to extract features from an image that can be used for classification. SVM, ANN, KNN, and ensemble classifiers is a machine learning technique, mainly used for classification and regression analysis.

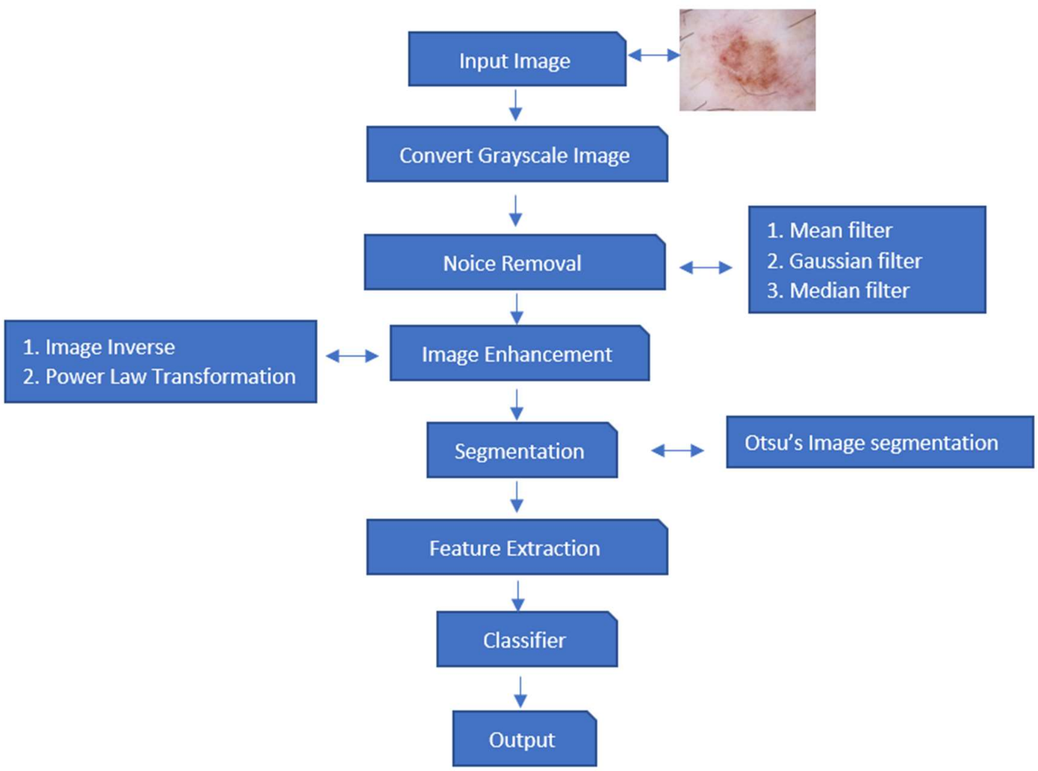

Fig 1. Block Diagram for Proposed Model

3.1 Implementation Details

As you probably know, RGB stands for Red, Green, and Blue. The RGB image is a composite of three different grayscale images that correspond to the intensity of Red, Green, and Blue color. These three grayscale images are recombined into a single image, which humans perceive as colored images. By combining different intensities of Red, Green, and Blue in grayscale, you can get a wide range of colored RGB images.

All the images are represented in RGB because, in colored images, each pixel can be represented as the vector of three numbers for the three primary colors: red, green, and blue. Each of these numbers is ranging from 0 to 255. [12]

3.2 Input image

Input to proposed system is dermoscopic images, dermoscopic images are images taken by dermatoscope. It is kind of magnifier used to take pictures of skin lesions (body part). It is hand held instrument make it very easier to diagnose skin disease. The dataset collected through the Kaggle website [13].

3.3 Pre-processing

Goal of pre-processing is an improvement of image data that reduces unwanted distortions and enhances some image features important for further image processing. Image pre-processing involves three main things gray scale conversion, noise removal and image enhancement.

3.3.1 Grayscale Conversion

It is simply one in which the colors are the shades of gray. You can get different shades of gray ranging from 0 to 255 (black to white). It varies according to the given intensities. **Gray scaling** is the process of converting an image from other color spaces e.g. RGB, CMYK, HSV, etc. to shades of gray. It varies between complete black and complete white.

- Dimension reduction: For example, In RGB images there are three color channels and three dimensions while grayscale images are single-dimensional.
- Reduces model complexity: Consider training neural articles on RGB images of 10x10x3 pixels. The input layer will have 300 input nodes. On the other hand, the same neural network will need only 100 input nodes for grayscale images.
- For other algorithms to work: Many algorithms are customized to work only on grayscale images e.g. Canny edge detection function pre-implemented in the OpenCV library works on Grayscale images only.

In grayscale conversion colour image figure (2) is converted into grayscale image shows in figure (3). Grayscale images are easier and faster to process than colored images. All image processing techniques are applied on grayscale image. In our proposed system colored or RBG image is converted into grayscale image by using weighted sum method by using following equations

$$\text{Grayscale intensity} = 0.299\,\mathrm{R} + 0.587\,\mathrm{G} + 0.114\,\mathrm{B} \ldots (1)$$

3.3.2 Noise Removal

The objective of noise removal is to detect and removed unwanted noise from digital image. The difficulty is in deciding which features of an image are real and which are caused by noise. Noise is random variations in pixel values.

In our proposed system we are using median filter to remove unwanted noise shows in figure (4) to figure (6). Median filter is nonlinear filter, it leaves edges invariant. Median filter is implemented by sliding window of odd length. Each sample value is sorted by magnitude, the centermost value is median of sample within the window, is a filter output.

I. MEAN FILTER

The mean filter is used to blur an image in order to remove noise. It involves determining the meaning of the pixel values within a *n x n* kernel. The pixel intensity of the center element is then replaced by the mean. This eliminates some of the noise in the image and smooths the edges of the image shown in figure 4. The blur function from the Open-CV library can be used to apply a mean filter to an image. When dealing with color images it is first necessary to convert from RGB to HSV since the dimensions of RGB are dependent on one another whereas the three dimensions in HSV are independent of one another.

II. GAUSSIAN FILTER

The Gaussian Filter is similar to the mean filter however it involves a weighted average of the surrounding pixels and has a parameter sigma. The kernel represents a discrete approximation of a Gaussian distribution. While the Gaussian filter blurs the edges of an image shown in figure 4. It does a better job of preserving edges than a similarly sized mean filter. The 'Gaussian Blur' function from the Open-CV package can be used to implement a Gaussian filter. The function allows you to specify the shape of the kernel. You can also specify the standard deviation for the x and y directions separately. If only one sigma value is specified, then it is considered the sigma value for both the x and y directions.

III. MEDIAN FILTER

The median filter calculates the median of the pixel intensities that surround the center pixel in a *n x n* kernel. The median then replaces the pixel intensity of the center pixel. The median filter does a better job of removing salt and pepper noise than the mean and Gaussian filters. The median filter preserves the edges of an image, but it does not deal with speckle noise. The 'median Blur' function from the Open-CV library can be used to implement a median filter result shown in figure 5.

3.4 Image enhancement

The objective of image enhancement is to process an image to increase visibility of feature of interest. Here contrast enhancement is used to get better quality results shown in fig (5).

3.4.1 Image Inverse

As you might have guessed from the title of this section, image inverse aims to transform the dark intensities in the input image to bright intensities in the output image, and bright intensities in the input image to dark intensities in the output image shown in figure 6. In other words, the dark areas become lighter, and the light areas become darker [14].

Say that I(i, j) refers to the intensity value of the pixel located at (i, j). To clarify a bit here, the intensity values in the grayscale image fall in the range [0,255], and (i, j) refers to the row and column values, respectively. When we apply the image inverse operator on a grayscale image, the output pixel O(i, j) value will be:

$$O(i, j) = 255 - I(i, j)$$

3.4.2 Power Law Transformation

This operator, also called *gamma correction*, is another operator we can use to enhance an image. Let's see the operator's equation. At the pixel (i, j), the operator looks as follows:

$$p(i, j) = kI(i, j)^{gamma}$$

I(i, j) is the intensity value at the image location (i, j); and k and gamma are positive constants. I will not go into mathematical details here, but I believe that you can find thorough explanations of this topic in image processing books. However, it is important to note that in most cases, k=1, so we will mainly be changing the value of gamma. The above equation can thus be reduced to:

p(i, j) = I(i, j)^gamma

3.5 Segmentation

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. First of all we have to take gray level of original image then calculate histogram of gray scale image then by using maximum entropy separate foreground from background. After maximum entropy we obtained binary image that is black and white image.

3.5.1 Otsu’s Image segmentation (Threshold-based segmentation)

It comes under threshold-based segmentation. In Otsu’s Segmentation, the input image is first processed, and then we try to obtain the histogram of the image, which will show the distribution of pixels in the image. Here we focus on peak value. The next step is to compute the threshold value and compare it with the image pixels. Set the pixel to white; if they are greater than the threshold else, set it to black [15]. Thus, it performs automatic thresholding. This method is not suitable for noisy images shown in figure 8. Applications include scanning documents and recognizing patterns.

3.6 Feature extraction

Feature extraction is very different from Feature selection: the former consists in transforming arbitrary data, such as text or images, into numerical features usable for machine learning. The latter is a machine learning technique applied on these features. Feature extraction plays an important role in extracting information present in the given image. Here we are using GLCM for texture image analysis. GLCM is used to capture spatial dependency between image pixels. GLCM works on gray level image matrix to capture the most common features such as contrast (equation 2), energy (equation 3), homogeneity (equation 4) and mean (equation 5).

Contrast

$$\sum_i \sum_j (i-j)^2\, C(i,j) \qquad \ldots (2)$$

Energy

$$\sum_i \sum_j C^2\,(i,j) \qquad \ldots (3)$$

Homogeneity

$$\sum_i \sum_j \frac{C(i,j)}{1+|i-j|} \qquad \ldots (4)$$

Mean (μ)

$$\frac{\sum_i^m \sum_j^n C(i,j)}{M*N} \qquad \ldots (5)$$

The purpose of feature extraction (GLCM) is to suppress the original image data set by measuring certain values or features that helps to classify different images from one another.

3.7 Classifier

Classifier is used to classify cancerous images from other skin diseases. For simplicity Support Vector machine classifier is used here. SVM, ANN, KNN, and ensemble classifiers take set of images and predicts for each input image belongs to which of the two categories of cancerous and non-cancerous classes. The purpose of SVM is create hyper plane that separates two classes with maximum gap between them. In our proposed system output of GLCM is given as input to SVM classifier which takes training data, testing data and grouping information which classifies whether given input image is cancerous or non-cancerous shows in figure (8).

4.0 Performance of Test Statistics

The term of test statistics tells us out of the total number of people who have skin cancer disease, the number that is correctly classified as having the disease by the model. The specificity of the model is its ability to determine healthy cases correctly. The positive predictive value (or precision) of the model means out of the total who are detected as disease positive by the model, how many of them are in fact disease positive. The accuracy of the model is the total number of persons correctly identified (true diseases positive plus true disease negative) out of the total number of people tested. F1 score is a harmonic mean of two factors i.e., precision and recall (or sensitivity). The performance of the model was measured by several factors including sensitivity, specificity, accuracy, and F1 score. False positive fractions and true positive fractions were used to demonstrate the ROC curve using table 1.

Table 1. Performance of the test statistics

Predicted Model	Disease Detection	Normal	Total
Predicted Model Positive	True Positive (TP)	False Positive (FP)	TP + FP
Predicted Model Negative	False Negative (FN)	True Negative (TN)	FN + TN
Total	TP + FN	FP + TN	TP + FP + FN + TN

- $\text{Sensitivity of Recall} = \text{TP}/(\text{TP} + \text{FN})$
- $\text{Specificity} = \text{TN}/(\text{FP} + \text{TN})$
- $\text{Positive Predictive Value or Precision} = \text{TP}/(\text{TP} + \text{FP})$
- $\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{FN} + \text{TN})$
- $\text{Average of precision and recall F1 Score} = 2(\text{Recall} * \text{Precision})/(\text{Recall} + \text{Precision})$
- $\text{False Positive Fraction} = \text{FP}/(\text{FP} + \text{TN})$
- $\text{True Positive Fraction} = \text{TP}/(\text{TP} + \text{FN})$

5.0 Results and Discussions

Numerical illustrations are used to explain the proposed models with tables and graphs for easy-to-understand concepts. In this section, which is used to presents the experimental results and their model performance and evaluation of the proposed skin cancer disease detection outcomes are categorized whether raw infected skin images for skin cancer or normal. The proposed model was implemented using Python for publicly available skin cancer disease datasets on the Kaggle repository. Gray-Level Co-occurrence matrix (GLCM) is a texture analysis method in digital image processing. This method represents the relationship between two neighboring pixels that have gray intensity, distance, and angle. In general, we use GLCM to get texture features in images such as dissimilarity, correlation, homogeneity, contrast, and others. The proposed model GLCM is compared to different similar studies with accuracy and training and detection time. The dataset contains 580 samples in two classes and the model predicts the outcome called skin cancer or normal. In this case the predicted number of samples was normal (90) and skin cancer (490). In this model taking consideration into the raw datasets are divided into 70% as training data and 30% as a testing data for the analysis of proposed system. Various numerical illustrations for use in these sections which are used to prove the results and discussions using tables 1 to table 6 subsequently explain the same through figure 2 to figure 12. The proposed system has been analyzed and predicted using the familiar detection accuracy namely F1-score, precision, recall, and specificity. The comparative analysis with the existing methods through accuracy and computational time.

Table 2: Dataset and their splitting

Total skin cancer images	580
Normal skin cancer images	90
Predicted skin cancer images	490
Training samples (70%)	410
Testing samples (30%)	170

The experiments and their accuracy performance of the proposed system using different soft computing techniques namely K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN) and Ensemble which is used to demonstrate the detection accuracy, precision, recall, specificity, and F1score respectively.

Fig. 2. Input image

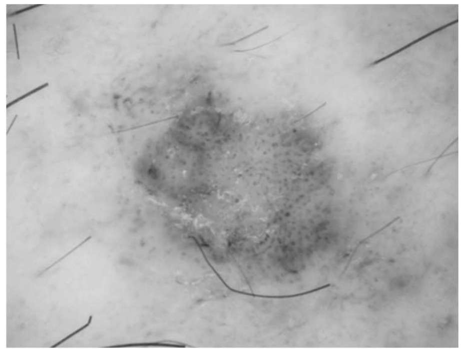

Fig. 3. Gray scale image

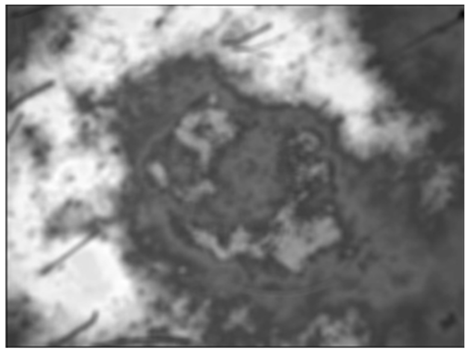

Fig. 4. Image without noise (Mean filter)

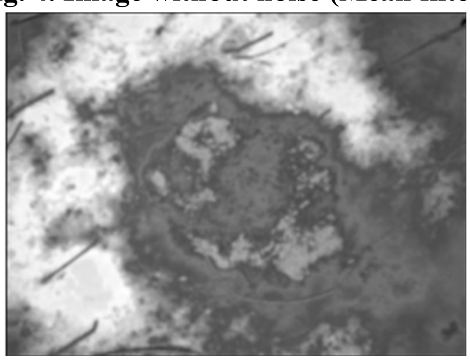

Fig. 5. Image without noise (Gaussian filter)

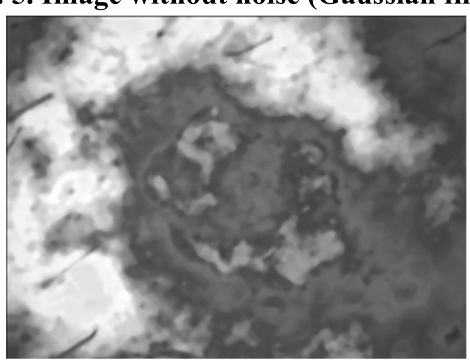

Fig. 6. Image without noise (Median filter)

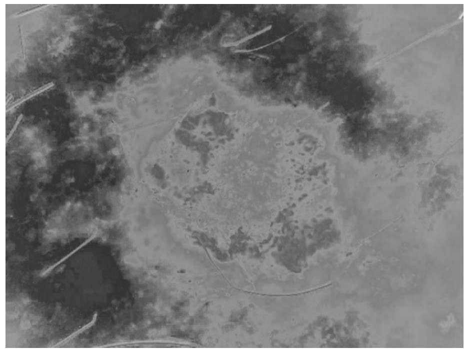

Fig. 7. Image enhancement

Fig. 8. Segmented image (Otsus binary threshold)

Different image processing techniques implemented using Google CoLab cloud environment with python. The configuration of the system Python 3 Google compute engine background with RAM: 1.10 GB/12 GB and Disk: 23.56 GB/107.72 GB. The following table indicate the processing time for different image processing techniques.

Table 3: Image processing techniques and its computational time

Image Processing Techniques	**Computational Time (In seconds)**
Gray scale image	0.1669
Image without noise (Mean filter)	0.0417
Image without noise (Gaussian filter)	0.0442
Image without noise (Median filter)	0.0605
Image enhancement (Image Inverse)	0.0106
Segmented image (Otsus binary threshold)	0.0935

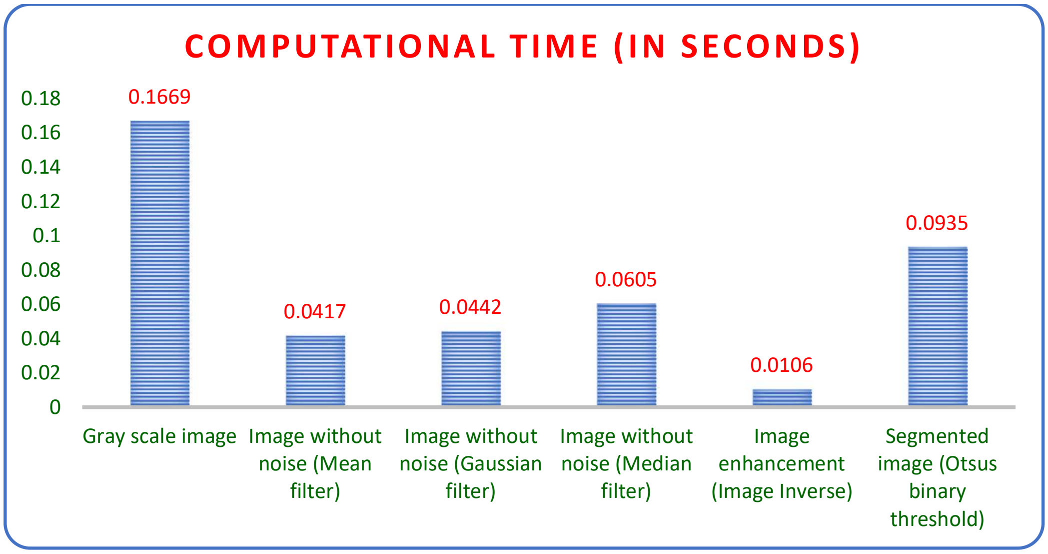

Fig. 9. Image processing techniques and its computational time

Table 4 Accuracy Performance

Classifier	Raw Features	Proposed Model
KNN	86.47	89.24
SVM	87.29	91.47
ENSEMBLE	90.28	92.57
ANN	93.89	94.47
Proposed Model	95.32	97.89

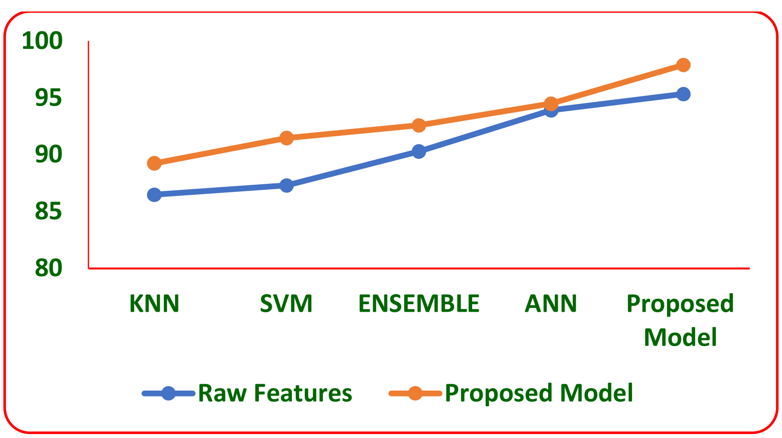

Fig. 10 Analysis of Accuracy Performance

Table 5: Precision Performance

Classifier	Raw Features	Proposed Model
KNN	80.58	83.74
SVM	82.45	85.16
ANN	84.24	86.14
ENSEMBLE	85.74	87.25
Proposed Model	91.85	93.45

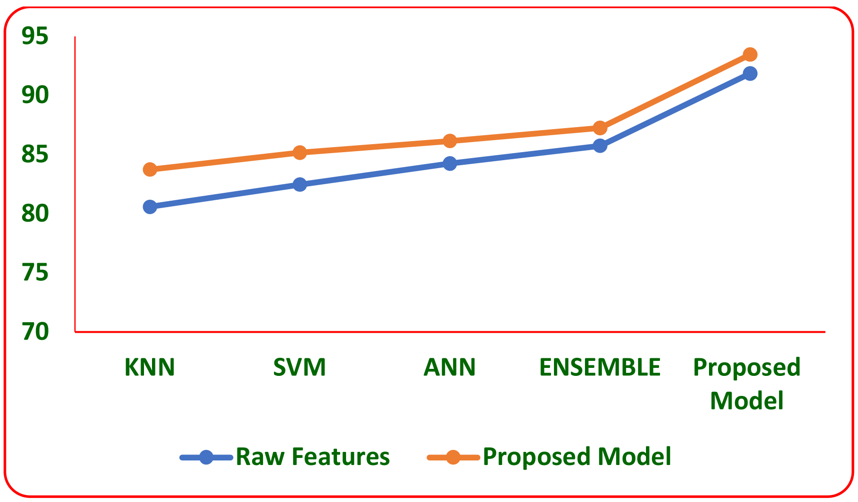

Fig. 11 Analysis of Precision Performance

Table 6: Recall Performance

Classifier	Raw Features	Proposed Model
KNN	88.47	91.85
SVM	87.24	90.74
ANN	89.74	92.74
ENSEMBLE	82.77	89.24
Proposed Model	91.54	94.78

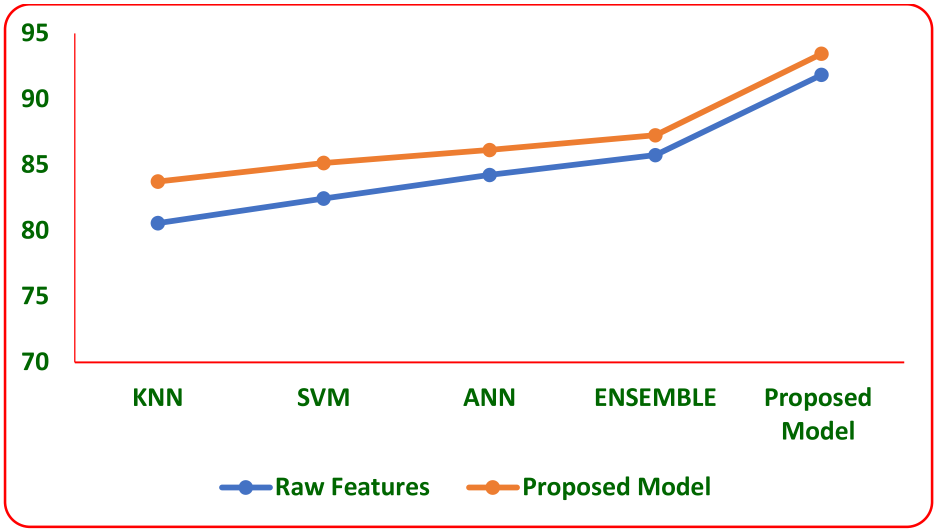

Fig. 12 Analysis of Recall Performance

Based on table 3 and figure 9, take 0.1669 seconds to convert raw input image to grayscale image. In image processing techniques noise reduction is one of the main processes for archive the better outcomes. In this case the proposed system taking consideration into three different methos like mean filtering, gaussian filtering and median filtering. These types of filtering techniques mean filtering have less time (0.0417) for conversion, subsequently gaussian (0.0442) and median filtering (0.0605). After grayscale and noise reduction, the image enhancement technique using image inverse is one the essential pre-processing techniques for archiving better outcomes with 0.0106 seconds. Finally, the image segmentation using GLCM takes 0.0935 seconds to complete the segmentation process.

Testing was performed on 580 sample images with 90 samples are normal skin cancer images. Predicted skin cancer images 490, training samples of 70% is 410, similar way testing also occur of 30% is 170. The related results and discussions are shown in table 2. Accuracy is calculated by using various machine learning approaches and its performance mentioned through table 4 to table 6 and similarly through figure 10 to figure 12. Based on different machine learning approaches and their accuracy performance, the proposed system has 97.89 %, the results shown in table 4 and figure 10. The proposed system and their precision performance compare different approaches namely KNN, SVM, ANN and Ensemble comparison, the proposed system having 93.45% which is shown in table 5 and figure 11. The percentage of recall performance compared to other classifications of the proposed system having 94.78% compared to other approaches. The results and discussion are shown in table 6 and figure 12. The proposed system delivered an high accuracy performance parameter namely accuracy, precision, recall, specificity, and F1-score metrics.

6.0 Conclusion and further studies

The proposed system of skin cancer detection can be implemented using gray level cooccurrence matrix and support vector machine, k-nearest neighbors, artificial neural network and ensemble comparison the proposed system having maximum accuracy. In future, improve the segmentation process try to get maximum accuracy to classify easily whether image is cancerous or non-cancerous.

References

1. Jaleel, J.A., Salim, S. and Aswin, R.B., 2013, March. Computer aided detection of skin cancer. In *2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT)* (pp. 1137-1142). IEEE.
2. Achakanalli, S. and Sadashivappa, G., 2014. Statistical Analysis Of Skin Cancer Image–A Case Study. *International Journal of Electronics and Communication Engineering (IJECE)*, *3*(3), pp.1-10.
3. Krishna, M.C., Ranganayakulu, S. and Venkatesan, D.P., 2016. Skin cancer detection and feature extraction through clustering technique. *International Journal of Innovative Research in Computer and Communication Engineering*, *4*(3), p.8.
4. Amarathunga, A.A.L.C., Ellawala, E.P.W.C., Abeysekara, G.N. and Amalraj, C.R.J., 2015. Expert system for diagnosis of skin diseases. *International Journal of Scientific & Technology Research*, *4*(01), pp.174-178.
5. Ahmed, K., Jesmin, T. and Rahman, M.Z., 2013. Early prevention and detection of skin cancer risk using data mining. *International Journal of Computer Applications*, *62*(4).
6. Sheha, M.A., Mabrouk, M.S. and Sharawy, A., 2012. Automatic detection of melanoma skin cancer using texture analysis. *International Journal of Computer Applications*, *42*(20), pp.22-26.
7. Rao, C.N., Sastry, S.S., Mallika, K., Tiong, H.S. and Mahalakshmi, K.B., 2013. Co-occurrence matrix and its statistical features as an approach for identification of phase transitions of mesogens. *Int. J. Innov. Res. Sci. Eng. Technol*, *2*(9), pp.4531-4538.
8. Ansari, U.B. and Sarode, T., 2017. Skin cancer detection using image processing. *Int. Res. J. Eng. Technol*, *4*(4), pp.2875-2881.
9. Jayaraman, S., Esakkirajan, S. and Veerakumar, T., 2009. *Digital image processing* (Vol. 7014). New Delhi: Tata McGraw Hill Education.
10. Parker, J.R., 2010. *Algorithms for image processing and computer vision*. John Wiley & Sons.
11. Maurya, R., Singh, S.K., Maurya, A.K. and Kumar, A., 2014, March. GLCM and Multi Class Support vector machine based automated skin cancer classification. In *2014 International conference on computing for sustainable global development (INDIACom)* (pp. 444-447). IEEE.
12. https://ninza7.medium.com/image-processing-convert-rgb-image-into-grayscale-with-few-lines-of-python-code-b2d2745ccaea
13. https://www.kaggle.com/code/manudeep123/23-skin-diseases-images-classification-40-acc/data
14. https://towardsdatascience.com/image-filters-in-python-and%20Gaussian%20filters
15. https://code.tutsplus.com/articles/image-enhancement-in-python--cms-29289
16. https://www.analyticsvidhya.com/blog/2021/09/image-segmentation-algorithms-with-implementation-in-python/