Proceedings of the International Conference on Edge Computing and Applications (ICECAA 2022)
DVD Part Number: CFP22BV8-DVD; ISBN: 978-1-6654-8231-8

Classification of skin cancer segmentation using Hybrid partial differential equation with fuzzy clustering based on machine learning techniques

P.Kavitha, Research Scholar,
Department of Computer Science,
School of Computing Sciences
Vels Institute of Science, Technology
& Advanced Studies (VISTAS)
Pallavaram, Chennai, India
pkavikamal@gmail.com

V.Jayalakshmi, Professor
Department of Computer Applications,
School of Computing Sciences
Vels Institute of Science, Technology
& Advanced Studies (VISTAS)
Pallavaram, Chennai, India
jayasekar1996@yahoo.co.in

S. Kamalakkannan, Associate Professor,
Department of Information Technology,
School of Computing Sciences
Vels Institute of Science, Technology
& Advanced Studies (VISTAS)
Pallavaram, Chennai, India
kannan.scs@velsuniv.ac.in

Abstract*— Skin cancer, particularly melanoma skin cancer is a serious public health concern today. The majority of skin cancersmay be treated if caught early enough. The fast proliferation of skin cancer necessitates the development of an automatic computerized detection system for skin cancer in its primary phases. The visual qualities of many skin cancer images are similar. The process of extracting characteristics from skin cancer image is a difficult one.The automated computerized diagnosis process aids dermatologists in improving the accuracy of skin disease analysis, allowing them tosave diagnostic time and provide better therapy for their patients. This research study presents an intelligent method for detecting malignant and non-cancerous cells using image processingtechniques. The Median Filter is first used to reduce artifacts, skincolor, hair, and other characteristics of the produced images by removing noise from the skin lesion. The lesion section is thensegmented individually using the suggested Hybrid Partial Differential Equation with Fuzzy Clustering (HPDE-FC) technique, which is also effective for feature extraction. Asymmetry, Border, Color and Diameter characteristics are extracted using ABCD scoring approach. Using several machine learning approaches like K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF) and Nave Bayes (NB) classifiers, the collected features are immediately fed to classifiers to categorize skin lesions amongst cancerous (malignant) and non-cancerous(melanoma). This research work, 325 images of normal skin lesions as well as 572 images of malignant skin lesions are downloaded from the International Skin Imaging Collaboration (ISIC) for this study. Using SVM classifiers, a classification result of 97.7% accuracy is obtained. Our goal is to evaluate the suggestedsegmentation technique's performance, excerpt the most applicablefeatures and associate the categorization results against those ofother algorithms in literature.

Keywords: Skin Cancer, Segmentation, Hybrid Partial Differential Equation with Fuzzy Clustering (HPDE-FC), feature extraction, Asymmetry, Border, Color and Diameter (ABCD), Support Vector Machine (SVM)

1. INTRODUCTION

Skin cancer contributes significantly to the causes of death all overthe world [1]. There are many different forms of malignancies that are identified and fought against. Conversely, skin cancer is one of the fastest-growing cancers today. According to recent study, every year the number of persons diagnosed with skin cancer grows faster than any other cancer kind [2]. The most common kind of skin cancer is melanoma, affecting melanocytes that line the skin's surface .It's made up of cells that darken your skin black [3]. Melanoma could be discovered in dark or darker colors, but itcan also appear in red, pink, blue, purple or white colors on the skin. [4].Identifying skin cancer at an early stage can help in reducing the risk factor in patients. Figure 1 depicts many kinds of skin cancer.

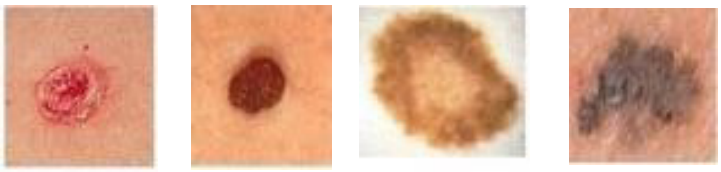


Figure.1 Types of skin cancer (a) Basal cell (b) Dysplastic Nevi (c) Melenoma (d) Nevus

Melanoma skin cancer can be treated if identified early on, and therapy for melanoma can begin sooner, potentially saving the patient's life. Melanoma, on the other hand, has a higher chance of spreading deep into the skin if it is identified in its later stages. If it has advanced to a deeper degree, it will be more difficult to cure. Melanocytes, which are present all over the body, are the main culprits in Melanoma [5]. Skin cancer is diagnosed using the biopsy procedure, which is the approved approach. This method entails taking a sample of a cell fromthe human body and sending it to a lab for analysis. It's the most difficult and terrible chore you'll ever face. It will take much longer for testing purposes. Testing requires extra time from both doctors and patients. It is riskier use the biopsy approach, as it may result in the illness spreading to other regions of the body [6]. The majority of the researchers contributed to this study and provided a number of detecting methods. The segmentation approach separates the lesion fromthe skin inorder to get the region of interest. Because of its efficiency and ease of implementation, the feature extraction technology is used in numerous detection systems. This computer-based analysis will cut down on diagnosis time and improve accuracy. There are development of malignant melanoma skin illnesses.This study intends to improve the accuracy of recognizing

and categorizing skin cancer using the unique HPDE-FC algorithm for segmenting the skin lesion, resulting in the development of a fully programmed computer-aided system to reliably detect melanoma cancer. Then, using a machine learning classifier, ABCD feature extraction algorithms are built, andaccuracy levels are calculated.

The organization of the paper is described as follows: Section 2 covers relevant studies, section 3 provides study methods based on preprocessing, segmentation, feature extraction, and classification, section 4 describes results and discussion based on performance assessment, and section 5 ends with conclusion.

2. LITERATURE REVIEW

Chen et al. [7] describe a technique for calculating parameter values based on their importance in the clustering process. The Weight Parameter Algorithm (WPFCM) is used to offer a novel possibilistic Fuzzy C-Means(FCM). On dataset X12, experimental findings suggest that WPFCM iteration times are roughly 25% faster than FCM and around 65 percent faster than PFCM. Sherin and Shayini [8] provide image processing algorithms for determining if a skin lesion is melanoma or nevus.

The input image is smoothed or blurred, the lesion region is segmented using the most rapid and exact multi thresholding technique, the characteristics of the lesion are extracted using an 18-feature vector, and a pre-trained Artificial Neural Network (ANN) classifier to differentiate between melanoma and nevus. According to Bhuiyan et al. [9], feature extraction is an important tool for properly analyzing a picture. Different digital images are examined using unsupervised segmentation algorithms. The segmented images are then subjected to feature extraction algorithms. According to Ansari and Anujasarode [10], dermoscopy images are processed based on the brightness of a grayscale picture is conveyed. Maximum entropy is used to the gray-scale photo histogram to segment the item. The Gray Level Co-Occurrence Matrix (GLCM) technique is used to extract texture characteristics. The output of GLCM is fed into SVM for categorization into malignant and non-cancerouscategories. The authors of Asht ami and Reshma [11] propose SVM for melanoma classification. To remove noise and tiny hairs, the median filter is utilized in the preprocessing stage. These preprocessed images are then segmented using K-means technique. The K-means clustering algorithm divides the images of a lesion into two groups. Even before the characteristics are retrieved, the territory is divided. The split of the region yielded two lesion regions: an inner lesion area and an outside lesion area.The color, texture, and boundary properties of the internal lesion Thao et al. [13] provide a suggested technique that focuses on texture and color feature extraction. The GCM approach is used to extract the textural properties. The attributes of mean and standard deviation are calculated using color characteristics. The features are chosen to categorize the most relevant characteristics for a better outcome, and the generated features are fed into the lesion using the k-NN classification algorithm. The categorization of an ANN is proposed by Wiem Abbes and DorraSellamin[14]. To eliminate noise, a median filter is utilized during the pre- processing stage. It has the advantage of being able to remember knowledge when on the go. High-level properties including color, border, asymmetry, and texture have been extracted and providedto ANN as training and testing input. Aleel et al. [15] provided a novel way to skin cancer diagnostics.

The input image collection,which included dermoscopic images, is pre-processed to eliminate air bubbles, thin hairs, and noise. For achieving the texture feature, GLCM is recommended. The GLCM may get data such as contrast, energy, mean, correlation, and so on. The lesion that is supplied as input is classified using an ANN-based classifier. Thecategorization is made based on the retrieved characteristics. Vijayakumar [20] applies the Combined Image Enhancement and the Restoration (CIEIR) on the input skin lesion images and makes it more presentable with the improved quality for the further imageprocessing steps in the identification of the normal skin and the skin affected by the Merkel cell tumour. Structure Similarity Index Measurement and the Mean Square Error are measured to evince the competence of the CIEIR against the convention methods of the image enhancement and the restoration. Krishna Babu and Joseph Peter [21] proposes an efficient skin cancer detection technique based on SVM with Histogram of Oriented Gradients (HOG) features. This proposed method is tested using and achieves 76% accuracy, 85% specificity, 84% precision, 76% recall and 75% F1-score. Kaur and Praveen [22] have used K- Means and PSO based cancer identification methods. RF is another popularly used cancer detection method that created decision trees for random data point selection. Murugan and Anu [23] have proposed KNN and RF method for skin cancer identification. Greedy fusion is another performing method in the field of cancerdetection. These methods are poor in performance in the real- world skin cancer identification problems and are time consuming and less accurate. Recently, SVM based methods are performing better in skin cancer identification

3. RESEARCH METHODOLOGY

The identification of skin cancer proceeds through several stages, as indicated in Figure 2. The procedure includes data collection, pre-processing, segmentation, feature extraction, and classification.

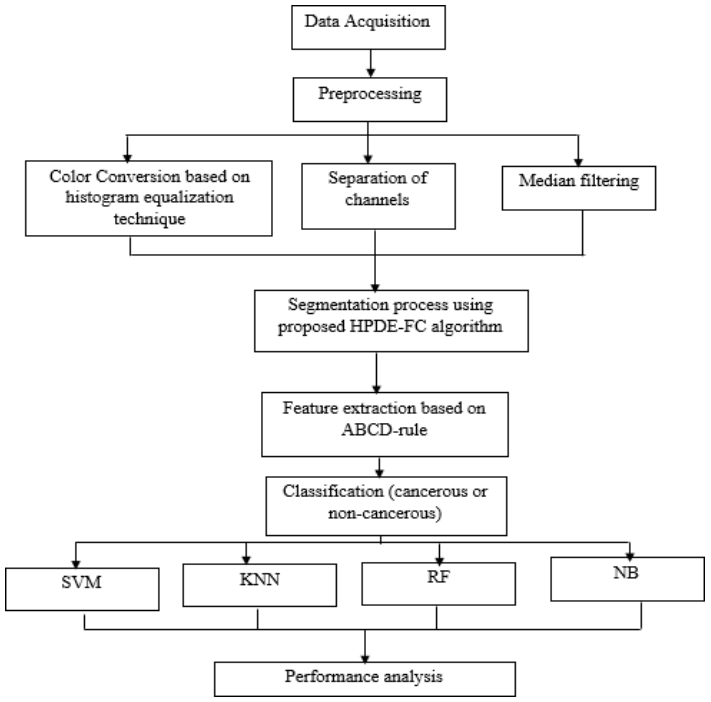


Figure.2 workflow diagram

3.1. Data Acquisition

The International Skin Imaging Collaboration (ISIC) databases of skin lesion images were used in the first portion of this study. The data sets utilized in this study include both normal and malignant melanoma skin lesions. There are 325 photos of normal lesions and 572 photographs of malignant melanoma lesions. ISIC 2017 databases were used to collect images of skin lesions. The images are in JPEG format. Figure. 3 Skin cancerous Melanoma images. For training and testing, the skin lesion photos are separated into 80:20 ratios.

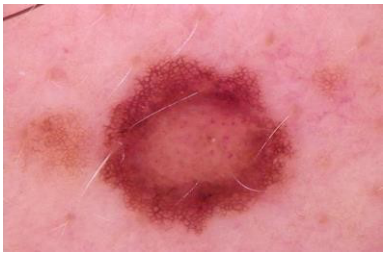
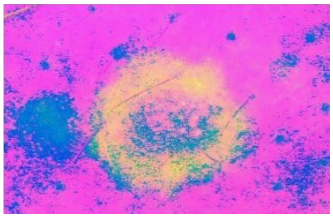


Figure.3 Input skin lesion images of Melanoma (cancerous region)

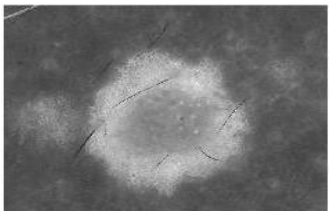
3.2 Pre-processing

In the second phase, the skin lesion datasets are pre-processed. Pre-processing eliminates undesirable items other than the lesion, allowing the lesion to be identified in subsequent steps. Artifacts, low contrast, hairs, veins, skin tones, moles, and other undesirable features are examples. The following preprocessing processes are then depicted in the diagram. Figure 4 (a) to (f), which are as follows: (i) color conversion using the histogram equalization approach, which involves converting a color image to the Hue Saturation Value (HSV) color space and then adjusting the saturation. Every nook and cranny is illuminated with HSV. (ii) Secondly, channel separation which display a grayscale representation of each color channel, (iii) The grayscale image is processed for noise removal using median filtering, which develops the skin lesion image, and the median filtered image is used for hair detection and removal. It is employed median filtering, which is far more successful than preserving edges at removing noise. The median filter's working structure is to go over the image element by element. Then, for each pixel, replace it with the median of nearby pixels. The window is the pattern of neighbors that moves across the whole image element, element by element. The element values from the window are numerically ordered after the first sorting of all the medians. An area filling morphology with inward interpolation on the pixel is used to remove the hair.

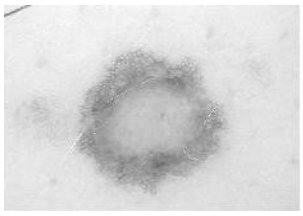


(a)

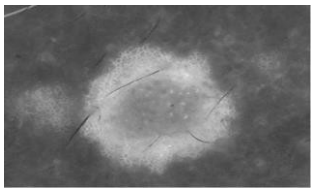
(b)



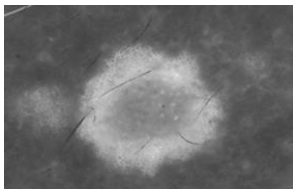
(c)



(d)



(e)



(f)

Figure.4 Pre-processed image of skin lesion (a) Color conversion using RGB to HSV (b) Channel separation of Red color (c) Channel separation of Green color (d)Channel separation of blue color (e) Median filter image (f) binary image(only grey level)

3.3 Hybrid Partial Differential Equation with Fuzzy Clustering (HPDE-FC) algorithm

The pre-processed images are segmented in the third phase. A skin lesion's specific location is determined via the segmentation process. Existing segmentation approaches are plagued by poor accuracy and great complexity. Because of the existence of artifact and large light changes, the cost has increased. The because of the limited visibility of melanoma, correct diagnosis is extremely difficult. Techniques in dermoscopy, such as Nevoscopy, are becoming more popular, High-frequency ultrasound, acoustic microscopy, and other techniques as well as high-frequency ultrasound in the third dimension (3D). Image is utilised to build the 3D volumes and to estimate the volume of the underlying structures. The depth of lesions is important in making an accurate melanoma diagnosis. In this research work, proposed method is performed using HPDE-FC algorithm for segmentation process. To extract the lesion part's border during the segmentation step, contours are employed. Segmentation is a technique for extracting a ROI from an image. Each pixel with common properties is contained within the ROI. To initialize the active contour, use the initialize level set function, or load an existing initial contour, follow these steps.
Equation (1) may be used to define the histogram equalized image

$$Z_{x,y} = base((Z-1)\sum_{y=0}^{c_4} x\ q^z)) \quad (1)$$

Where the base is the integer that comes closest to the value supplied. This is the same as converting pixel intensity, but in the other direction.

$$\frac{\partial O}{\partial y}\left(\int_0^O x\ qO(y)da\right) = \partial O\,(O)(y^{-1})(O))d/dO \quad (2)$$

This is where, at long last, the probability distributed uniformity function may be expressed as $\frac{\partial O}{\partial y}$,

The pre-processing of the input image is represented by Eqn.1 and Eqn.2. While the findings reveal that the equalisation technique

DVD Part Number: CFP22BV8-DVD; ISBN: 978-1-6654-8231-8

creates flat histograms, it also softens and enhances histograms' look.

In response to these issues, a semi-automatic segmentation method is developed that generates the seed locations for establishing the cancer's perimeter interactively. As a result of this procedure, the Laplace transformation is first applied to Eqn.3 in relation to the variable t, yielding

$$v(y) = L[O[v(y)] + S[v(y)]] + L[\phi(y,x)] \quad (3)$$

By using I.C. (3), eqn (4) is obtained

$$t^{o}\tilde{v}(y,t) = i^{-}(y,t) + L[O \mid v(y,x)] + G[v(y,x)]] + \tilde{\phi}(y,t) \quad (4)$$

Where x, y indicate the rows and columns of pixels, and t denotes the intensity level and smoothening range.

$$f^{\sim}(y,t) = \frac{i^{-}(y,t)}{t^{o}} + \frac{o^{-}(x,s)}{s^{n}} \quad (5)$$

Now, inverse Laplace transformation is applied on equation (5) concerning $^{2}8^{7}$, and then eqn (6) is obtained

$$Vx = C^{-1}[f^{\sim}(x,y)] + C^{-1}[\frac{1}{y^{o}}L[O[v(y,u)] + S[v(y,t)]] \quad (6)$$

In the second step, a differential transformation on the equations (5) and (6) concerning $^{4}x'$ is applied and eqn (7) is obtained.

$$V_0(u) = h_0(u), U_1(u) = i_1(u) \quad (7)$$

where $V_u(u)$ and $G_l(v)$ are the differential transform of $v(x,y)$ and $z(x,y)$ respectively.
The closed-form of the solution may be expressed as using the preceding recurrence equation (7) and the initial conditions (8). The segmentation method used in this study is grey scales, which are used in combination with a computer. The size difference between big and small pixels surrounding the object's edges is measured.

$$X^{segment} = \sum_{[x,y]\in S_2}^{o} x\, V_2\left(V_x, V_y\right) \cdot Vo \cdot log_{c_i} + \gamma \int c_i dy \quad (8)$$

In Eqn. 9, the SPF function is created in the following manner

$$fn(j(y)) = \frac{v(y) - \frac{d_1+d_2}{2}}{n\left((y) - \frac{d_1+d_2}{2}\right)}, y \in \Omega \quad (9)$$

open function of $H^{2}, j(x)$ is the given image in Ω, d_1 and d_2 are defined in Equations 10 and 11 respectively.

$$d_1(\phi) = \frac{\int_{\Omega}(y)I(\phi)dy}{\int_{i} I(\phi)dy} \quad (10)$$

Here the Heaviside function, $I(\phi)$ is approximated by a smoothed functional I_F which is defined by Equation 11.

$$I_d(A) = \frac{1}{2}\left(1 + \frac{2}{\pi} arctan\left(\frac{\pi}{j}\right)\right) \quad (11)$$

Equation 11's relevance may be explained in the following way. Even if the intensities within and outside the object are homogenous and c 1c 2, it is clear that Min (Jy)d1, d2Max (J y) and the equal signs cannot be obtained at the same time anywhere the dominant contour is present. As a consequence, Equation 12 is obtained.

$$Min(1(x) < \frac{d_1+d_2}{2} Max(1(y)), y \in \Omega \quad (12)$$

When the item is in the opposite position, it is derived as in Eqn 13 by substituting the fn function in Eqn 12 of the Level set formulation.

$$\frac{\partial\phi}{\partial t} = fn(j(y)) \cdot (div\left(\frac{\nabla\phi}{|\nabla\phi|} + \alpha\right) \mid \nabla\phi 1 + \nabla fn(m(y)) \cdot \nabla\phi, y \in \Omega \quad (13)$$

As a result, the level set formulation of HPDE-FC may be simplified in Eqn 14

$$\frac{\partial\phi}{\partial t} = spf\,(I(x)) \cdot a|\nabla\phi|, x \in \Omega \quad (14)$$

The slices, or finely divided images, are the outcome of a leveling technique that provides proper proximity on the skin cancer image.
The segmented portion of a skin cancer image is depicted in Figure 5.

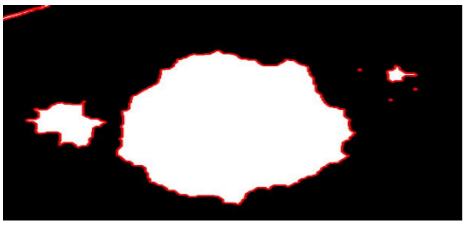


Figure.5 segmented region using HPDE-FC

3.4 Feature Extraction using ABCD rule

The lesion is classed as malignant or benign once it has been segmented off of the background skin. The optimal feature descriptors for machine learning modeling are necessary for improved categorization outcomes. Asymmetry, border structure, color variation and dermatoscopical structure or diameter of the lesion are the ABCDs that form the basis for a dermatologist's diagnosis. The following steps are used to extract features.

A. Asymmetry

Symmetry is a significant component of shape comprehension that comes in handy when analyzing patterns. A symmetric pattern requires only one half of a design with the axis of symmetry. If a component of pattern is lost or noisy, symmetry can be used to fill in the gaps or eliminate the noise. Two asymmetry characteristic values that may be used to determine the degree of symmetry are the Asymmetry Index (AI) and the Lengthening Index. In this paper, the asymmetry index is calculated.
The following equation (15) is used to calculate the AI.

$$AI = \frac{\Delta A}{A} x100 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

Where, A= total Image's Area, ΔA = Difference in area between the entire picture and the lesion region.

B. Border Irregularity

It is measured using different metrics are described below:

Compact Index (CI):

The CI is the most used type of barrier for estimating dimensionless 2D objects. However, noise around the border

DVD Part Number: CFP22B V8-DVD; ISBN: 978-1-6654-8232-8

makes this measurement extremely sensitive. The following equation (16) can be used to find this

$$CI=\frac{P_L^2}{4\pi A_L} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(16)$$

P_L = Perimeter of the lesion, A_L = Area of the lesion.

- **Fractal Dimension (FD):**

Self-similarity and scale/size qualities are features of fractals. Each part has a fractal that is scaled differently than the entire fractal. These qualities make fractal compression algorithms a good fit. Another aspect is the fractal dimension. A dimension's size is often an integer; for example, a line has one dimension, a field has two, a cube has three, and so on. The fractal dimension is remarkable in that it may be valued only a fraction of what it appears to be. This FD can be used as a feature in an image. The method of box calculation may be used to compute the FD (box-counting). The Haussdorf dimension computation technique is a simple and effective way to calculate the FD of an image. Examine a squiggly line drawn on a sheet of paper. Allow N (e) to be the smallest e-sided cube that can cover this line with a 2-dimensionallaa cube with side e. The dimension of this line is expressed in Equation 17.

$$D=\lim\ e\to 0\ \frac{\ln(N\ (e))}{\ln(\frac{1}{e})} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(17)$$

- The fractal dimension of a picture may be simply computed using equation (17).

Edge Abruptness:

The radial distance of a lesion with irregular borders (abruptness edge) is significantly different. The distribution of radial distance difference in equation 18 is used to assess barrier irregularity.

$$C_r=\frac{\frac{1}{}\sum_{p\in c\ d_2}(P.GL-m_d)^2}{\frac{P_L}{m_d^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(18)$$

Where, m_d is the mean distance of d_2 between the centered point and the barrier P.GL.

C. Color Variation

One of the earliest indicators of melanoma is the occurrence of color changes. Melanoma cells are typically brown or black in color, depending on the creation of melanin pigment at various depths in the skin. Melanoma cells increase in increasing pigment. Color descriptors are statistical variables derived from several color channels namely, the average value and standard deviation of the RGB or HSV color channels. In this paper, the color variance of the RGB image is computed using the HSV channel.

D. Diameter

Melanomas develop larger than typical moles, particularly those with a diameter of 6mm. Because wounds are typically uneven in shape, the diameter is measured by drawing a line from all the edge pixels to the pixel edges via the midway and averaging it.

3.5Classification

After segmentation and feature extraction, the feature vector is submitted to classifiers to detect the malignant and non-cancerous. Machine learning algorithms are the most extensively utilized for categorization of lesion kinds. The most widely used algorithms include SVM, KNN, NB, RF, and other approaches in which retrieved features are immediately passed to the classifier. These classification algorithms are trained and tested with their default parameter values to attain high accuracy which are described below.

SVM: The most frequent type of classifier is the SVM. The main advantage is the unified framework; that is, by selecting a kernel, many types of machine learning architectures may be built. An SVM classifier is given the information acquired from malignant and non-cancerous skin lesions.

KNN: It's the simplest, most straightforward, quickest, and most effective technique. To categorize an image, the majority votes of its neighbors are employed. The KNN classifier is given the training and testing samples, and each class is determined using the closest distance. The KNN classifier is used to classify malignant and non-cancerous features.

NB: The cornerstone of NB analysis is the prior probability belief, which is based on Bayes theorem. The key advantages are that no vast data is necessary, that it is quick, and that the qualities are conditionally independent (i.e. there is no reliance between them). The supervised machine learning approach of NB is used.

RF: RF is made up of a number of decision trees, each of which has a position arrangement impact due to the use of dissimilar categorization. This method allows for the random sample methodology to be used to evaluate the sampling allocation,which is very useful for some minute models.

Algorithm: Skin Cancer Classification

Step: 1 procedure Color Conversion as Img = rgb2hsv (img1)
Step: 2 Separation of Channels
Step:3 Median Filter (img2=medfilt(1))
Step:4 Set parameters
Maximum no of iteration = 10

Lamba =1 coefficient of cv term

nu = 0.5 coefficient of smooth term, increase it can obtain more smooth segmentation result

mu = 1 coefficient of penalty term with respect to level set function \phi

epsilon = 1 parameter of regularized Heaviside function and its derivative

timestep = 0.1 time step

Step: 5 initialize level set function or load an existing initial contour (Fuzzy Clustering)

Step: 6 Evolution process represents current evolving contour every 5 iterations

function phi = Evolution(phi, u0, nu, mu, timeStep, lambda, epsilon)

Step: 7 compute dirac function using epsilon

DE= (epsilon/pi)./(epsilon^2.+phi.^2);

DVD Part Number: CFP22B5-DVD; ISBN: 978-1-6654-8231-8

Step: 8 compute average intensity from inside and outside

in = find(phi <= 0); the current pixel is assigned from inside

out = find(phi > 0); the current pixel is assigned an edge label from outside

Step:9 Feature extraction (HPDE-FC)

(f1,f2,f3,f4..fn) = ABCD features

Step:10 Classification based on segmented features (f1,f2,f3,f4…...fn)

Model = training and testing (f1,f2,f3,f4…...fn)

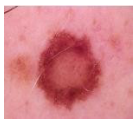
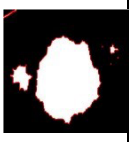
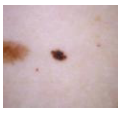
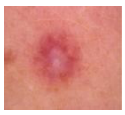
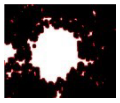
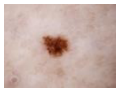
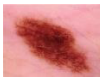
Accuracy = Confusionmat (predicted labels, groud labels)

End procedure

4. Results and Discussion

The suggested approach is used to analyze skin lesion pictures from ISIC. There are 325 normal images and 572 melanoma images in the databases. The next section discusses the experimental findings of the extracted feature. The segmented output is shown in Table 1 together with the relevant feature values.

TABLE.1 PROPOSED SEGMENTED SKIN CANCER IMAGES WITH OBSERVED FEATURE VALUES

Input image	Segmente doutput image	Asym metri c	Compa ctness index	Fractal Dimen sion	Edge Abrup tness	Color variance
		15%	1.9256	0.9012	0.025	0.110
		16%	1.5245	0.8512	0.0361	0.1031
		20%	1.8245	0.8563	0.0548	0.1458
		21%	1.6523	0.8956	0.0641	0.1254
		25.6%	1.9563	0.9142	0.0745	0.1546

If the asymmetric index value of a correctly segmented skin lesion falls between 10% and 25%, it might be deemed asymmetric. To obtain the real value of the compactness index, adequate segmentation is also necessary. Based on the information gathered, the compactness index value for melanoma skin cancer is estimated to be between 1.5 and 2.0. In a multi-resolution representation, the fractal dimension is a measure of similarity. All of the skin lesions for melanoma skin cancer have fractal dimensions ranging from 0.83-0.91. If the fractal dimension value of a skin lesion falls within this range and other feature fulfillments, it might be considered a melanoma. The results of this investigation indicated that the value of skin cancer edge abruptness ranges from 0.02-0.08. It is also been discovered that some non-melanoma cancers fall into this category. As a result, other factors should be considered when determining whether or not a person has skin cancer. The color variance of all skin lesions falls between the ranges of 0.10-0.15, according to the color variance statistics.

Different train and test ratios are used to train the classifiers. The Image of skin lesions are separated into two categories: training and testing. The acquired findings are briefly discussed as classification performance. True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are the four potential results for binary class categorization. A TP indicates that the system has identified melanoma (skin cancer) as malignant (melanoma), whereas a TN indicates that the lesion is non-cancerous and has been classed as such. As shown in the table, a projected FP value arises when a melanoma lesion is diagnosed as non-cancerous, and a FN occurs when a non-cancerous lesion is labeled as malignant in table 2.

TABLE.2 CONFUSION MATRIX FOR VARIOUS CLASSIFIER BASED ON PROPOSED SEGMENTED FEATURE VALUES

Model Description	TP	TN	FP	FN
SVM	140	79	3	2
KNN	135	80	3	6
RF	137	75	8	4
NB	135	72	11	6

When applied to FP and FN examples, Table 3 demonstrates how the performance of binary class classification may be examined in terms of accuracy, specificity, recall, sensitivity and precision.

Sensitivity: The percentage of positive values that the classifier correctly recognizes.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

Specificity: The proportion of incorrectly recognized negatives by the classifier.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN}+\text{FP}}$$

Accuracy: The following formula is used to calculate the proportion of correctly recognized instances to the total number of test cases:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: The percentage of instances that are projected to be cancerous.

$$\text{Precision} = \frac{TP}{TP+FP}$$

DVD Part Number: CFP22B V8-DVD; ISBN: 978-1-6654-8231-8

TABLE 3 PERFORMANCE EVALUATION RESULTS FOR PREDICTING SKIN CANCER BASED ON PROPOSED SEGMENTATION FEATURE VALUES

Model Description	Accuracy	Precision	Recall	Sensitivity	Specificity
SVM	0.9777	0.9790	0.9859	0.9859	0.9634
KNN	0.9598	0.9783	0.9574	0.9574	0.9639
RF	0.9464	0.9448	0.9716	0.9716	0.9036
NB	0.9241	0.9247	0.9574	0.9574	0.8675

Accuracy is the metric which can be calculated by number of correct prediction by total number of patients considered for observation. The accuracy score for SVM is 0.9777 which illustrates the prediction of skin cancer with most accurate while compare with other existing model KNN, RF and NB is shown in figure 6. Similarly, the precision can be calculated as number of positive predictions by total amount of positive prediction whereas the proposed HPDE-FC segmented feature values with SVM is 0.9790 which is comparatively greater than KNN, RF and NB. The major metrics considered for evaluating the performance of the skin cancer is sensitivity and specificity. The sensitivity is measured as total correct positive prediction by total number of positive prediction in dataset. The sensitivity of proposed HPDE-FC segmented feature values with SVM is 0.9859 which is higher while compared with other classifier such as KNN, RF and NB. In other hand, the specificity can be measured as number of correct negative prediction by total number of negative prediction in dataset. The HPDE-FC segmented feature values with SVM is 0.9634 comparatively lesser than KNN whereas SVM is higher while compared with NB and RF is shown in figure 6.

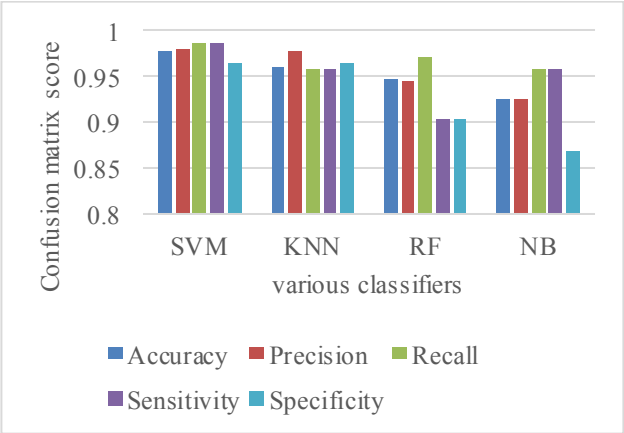


Figure 6 Performance Evaluation of proposed HPDE-FC segmented feature values with various classifiers

Hence, the proposed HPDE-FC segmented feature values with SVM plays a main role for better way of feature extraction which makes the dataset with significant variables related to the target of the dataset. The proposed model has accumulated the significance features from the four individual classifiers namely SVM, KNN, RF and NB. Thus, the accumulated SVM classifier prediction has enhanced the prediction level with more accuracy score of 0.9777. Table 4 shows the suggested segmentation approach, how to extract the most appropriate characteristics, and how to compare the classification results with those of other techniques in the literature.

TABLE.4 COMPARISON OF PROPOSED TECHNIQUE WITH OTHER LITERATURE

Authors	Techniques used	Accuracy
Vidya and Karki [16]	Segmented using Geodesic Active Contours (GAC), feature extraction based on ABCD, classification using SVM	96%
Qasim Khan et.al[17]	K means clustering, GLCM features and SVM classifier	96%
Christy Bobby [18]	Otsu global , SVM based on Relieff algorithm	90%
Thaajwer and Ishanka [19]	Adaptive image thresholding using Otsu's thresholding method, GLCM based SVM classifier	83%
Proposed method	Segmentation of HPDE-FC, ABCD rule with SVM classifier	**97.7%**

5. Conclusion

The suggested model is utilized in this paper to categorize skin lesions as malignant or non-cancerous. Machine learning techniques are used to identify skin lesions automatically utilizing the ABCD rule for feature extraction and classification. The HPDE-FC approach is recommended for skin lesion segmentation. The suggested segmentation algorithm is more resistant to air bubbles, thick hair, and low contrast than existing segmentation algorithms, which are inaccurate and complicated. The ABCD rule is suggested for feature extraction based on skin lesion's asymmetry, color and diameter. The classification is then addressed using several machine learning approaches such as SVM, RF, KNN and NB. The suggested approach is tested using ISIC dataset's skin lesion image. When compared to other classification algorithms, SVM outperforms them all with a 97.7% accuracy rate. SVM has a sensitivity of 98.59 percent and a specificity of 96.34 percent, respectively. The data shows that following augmentation, the accuracy achieved improves. Since it is a timeless and painless approach, it is more efficient and pleasant for both patients and physicians than biopsy. This strategy may also be applied to Deep Learning to improve accuracy.

978-1-6654-8232-5/22/$31.00 ©2022 IEEE

References

[1] M. Aleem, N. Hameed, and A. Anjum, "m-Skin Doctor : A Mobile Enabled System for Early Melanoma Skin Cancer Detection Using Support Vector Machine," vol. 2, pp. 468–475, 2017.
[2] B. Kong, S. Sun, X. Wang, Q. Song, and S. Zhang, "Invasive cancer detection utilizing compressed convolutional neural network and transfer learning," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2018.
[3] Q. Li, L. Chang, H. Liu, M. Zhou, Y. Wang, and F. Guo, "Skin cells segmentation algorithm based on spectral angle and distance score," Opt. Laser Technol., vol. 74, pp. 79–86, 2015.
[4] R. Kasmi and K. Mokrani, "Classification of malignant melanoma and benign skin lesions : implementation of automatic ABCD rule," vol. 10, no. 1, pp. 448–455, 2016.
[5] V. M. M, "Melanoma Skin Cancer Detection using Image Processing and Machine Learning," International Journal of Trend in Scientific Research and Development (IJTSRD), vol. 3, no. 4, pp. 780-784, 2019.
[6] B. A. Uzma and T. Sarode, "Skin Cancer Detection Using Image Processing," International Research Journal of Engineering and Technology (IRJET), vol. 04, no. 04, pp. 2875-2881, 2017.
[7] Jiashun Chen , Hao Zhang, Dechang Pi, Mehmed Kantardzic, Qi Yin, and Xin Liu, "A Weight Possibilistic Fuzzy C-Means Clustering Algorithm", Hindawi Scientific Programming, Volume 2021, Article ID 9965813, 10 pages, https://doi.org/10.1155/2021/9965813.
[8] K C Shahana sherin and R Shayini , "Classification of Skin Lesions in Digital Images for the Diagnosis of Skin Cancer", IEEE Xplore, 2020, ISBN: 978-1-7281-5461-9.
[9] Md.Amran Hossen Bhuiyan, Ibrahim Azad and Md.Kamal Uddin, "Image Processing for Skin Cancer Features Extraction", International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013, ISSN 2229-5518.
[10] Uzmabano Ansari, T anujasarode, "Skin Cancer Det ect ion Using SVM", Proceedings of WRFER International Conference, 16thApril2017, Pune, India.
[11] Asht ami K M, Ms. Reshma V K, "Classificat ion of Melanoma on Dermoscopy Images using SVM Classifier", Int ernational Journalfor Research in Applied Science and Engineering Technology (IJRASET), ISSN: 2321-9653, vol:06, issue:04, April 2018.
[12] Akila Vict or, Muhammad Rukunuddin Ghalib, Ali Douik, "Automatic Detect ion and Classification of Skin Cancer", International Journal of intelligent Engineering and Systems, vol:10, issue no: 03, March 2017.
[13] Le T hu T hao, Nguyen Hong Quang, "Aut omat ic Skin Lesion Analysis towards Melanoma Detection", International Conference on Communication, Control and Intelligent Systems (CCIS),2015.
[14] Wiem Abbes and DorraSellami, "High-Level Features for Automatic Skin Lesions Neural Network based Classificat ion", IEEE IPAS'16:International Image Processing Applications And Systems Conference 2016.
[15] Dr. J. Abdul Jaleel, Sibi Salim, Aswin. R . B, "Diagnosis and Detection of Skin Cancer Using Artificial Intelligence", International Journal of Engineering and Innovative Technology(IJEIT), volume 3, issue 2, August 2013.
[16] M Vidya and Dr. Maya V Karki, "Skin Cancer Detection using Machine Learning Techniques", IEEE Xplore, 2020.
[17] Muhammad Qasim Khan, Ayyaz Hussain, Saeed ur Rehman, Umair Khan, Muazzam Maqsood, Kashif Mehmood and Muazzam A. Khan, " Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer", DOI 10.1109/ACCESS.2019.2926837, IEEE Access,2017.
[18] T. Christy Bobby, " Benign and Malignant Dermatoscopy Image Classification", IEEE, 2021.
[19] MA. Ahmed Thaajwer and UA. Piumi Ishanka, "Melanoma Skin Cancer Detection Using Image Processing and Machine Learning Techniques", IEEE Xplore, 2020.
[20] Vijayakumar, Dr T. "Selective Image Enhancement And Restoration For Skin Cancer Identification." Journal of Innovative Image Processing 1, no. 1 (2019): 1-10.
[21] G. Neela Krishna Babu and V. Joseph Peter, "Skin Cancer detection using Support Vector Machine with histogram of oriented gradients features", ICT ACT Journal on soft computing, January 2021, Volume: 11, Issue: 02.
[22] R. Kaur and Praveen Kumar, "An Enhanced and Automatic Skin Cancer Detection using K-Mean and PSO Technique", International Journal of Innovative Technology and Exploring Engineering, Vol. 8, No. 9, pp. 634-639, 2019.
[23] A. Murugan and S. Anu, "Detection of Skin Cancer using SVM, Random Forest and kNN Classifiers", Journal of Medical Systems, Vol. 1, No. 9, pp. 269-275, 2019.