Multimedia Tools and Applications (2021) 80:8909–8929
https://doi.org/10.1007/s11042-020-10056-8

Integration of modified ABCD features and support vector machine for skin lesion types classification

K. Melbin [1] · Y. Jacob Vetha Raj [1]

Received: 28 December 2019 / Revised: 2 September 2020 / Accepted: 7 October 2020 /
Published online: 5 November 2020

Abstract
The abnormal growth of skin cells often leads to skin cancer due to the exposure of skin cells to the sun. The skin disease is primarily caused by bacteria, fungus, viruses, UV radiation, and chemical substances. Generally, clinicians have been a trouble to categorize melanoma, seborrheic keratosis and lupus Erythematosus diseases due to the resemblance in the features of pigmented diseases. The paper presents an integrated approach for detecting the skin lesion from the dermoscopic images. The proposed integrated cumulative level difference mean (CLDM) based modified ABCD features and Support vector machine (SVM) have used for the detection and classification of skin lesion images. The proposed modified ABCD features employed for extracting the skin features like shape, size, color, and texture from the skin lesion images. Prior to the classification method, the Eigenvector Centrality feature ranking and selection (ECFS) method has utilized for better classification. After the feature selection method, a skin lesion image is classified by the Support vector machine (SVM). The performance of segmentation has been assessed by evaluating the Jaccard Similarity Index (JSI), Dice similarity coefficient (DSC), sensitivity, specificity, and accuracy. The proposed SVM method classifies the three skin lesion classes and produces excellent classification results with the classification accuracy is 97%, specificity is 98%, sensitivity is 97%, JSI is 97% and DSC is 98% for melanoma, seborrheic keratosis, and lupus Erythematosus respectively. The proposed approach classifies the three skin lesion classes (melanoma, seborrheic keratosis and lupus Erythematosus) with high accuracy. The integrated method not only enhances the accuracy level but also delivers significant information for better classification.

Keywords Skin lesion · Dermoscopic images · Cumulative level difference mean · Eigen-vector centrality · Support vector machine

✉ K. Melbin
melbink432@gmail.com; melbinmean@gmail.com

[1] Department of Computer Science, Nesamony Memorial Christian College, Marthandam, India

1 Introduction

The human skin engages different responsibilities to protect the body from external attack. Skin delivers strong protection and performs as an obstacle against ultraviolet radiation, viruses, bacteria, and dangerous chemicals. Generally, individual skin has epidermal and dermal layers. Also, it shields the body from mechanical stress as well as strain. Skin diseases have commonly induced by microscopic organisms like bacteria, viruses, fungus, and dangerous chemicals. According to the survey, different types of skin diseases acquire worldwide every year [26]. An unrestrained increase of skin diseases can be injurious to health; therefore, suitable medicine is significant, and also transmittable diseases can confirm to be most hazardous hence required protections to be taken for avoiding the diseases.

Skin lesions like Actinic Keratosis (AK), Rosacea (ROS), Seborrheic Keratosis (SK), Lupus Erythematosus (LE), Squamous Cell Carcinoma (SCC) and Basal Cell Carcinoma (BCC) are some examples of benign or malignant lesions. Skin cancers are the malignant lesions which divide the cells rapidly and spread to other regions of the body parts [28]. BCC and SCC are usually familiar types of skin cancer, and these diseases are non-melanoma skin cancer [41]. The most fatal sort of skin tumor is melanoma, which spread the cells rapidly to other regions hence very risky to cure this disease. Based on the research, the death rate can be decreased by 90% if the skin tumor is detected by an earlier period. Early-phase detection is helpful to survive the patient from malignant lesions [11, 19].

In worldwide, there is a lack of dermatologists; thus, the circumstances have further weakened. In this situation, the imaging process acts as a vital function in the rapid resolution creating during the detection of skin lesions [43]. The medical imaging has utilized for providing internal organ information concealed by the skin and also used to detect and treat disease [27]. Dermatologist supports the identification of skin injury on the visual evaluation of the skin. Moreover, computer-aided treatment has helpful for collecting information from skin imagery and also human visualization reduces the accuracy level [17]. Dermatoscopy has a non-invasive method to provide a complete observation of the skin lesions. The most extensively agreed diagnosticated method utilized for earlier identification of melanoma which is mainly a fatal type of skin cancer. Dermoscopy images encompass immense possible for the earlier identification of skin diseases. However, once the ailment has detected at an earlier period, effectual healing treatment can be done to cure the diseases [25].

Various machine learning algorithms has extensively used in the medical image processing area for the intention of skin lesion classification and feature extraction method. These methods have used for assessing a different variety of skin lesions at an earlier period, followed by analyzing the diverse nature of the lesions. Several classification algorithms are employed in this study to classify skin lesions based on skin color, texture, and skin types. Consequently, the identification and classification of skin lesions have significant for the diagnosis of skin diseases. In recent years, evolutionary methods have been achieved better performance in computer vision applications [1, 30, 34, 35, 36, 38, 39, 40].

In this research, the main intention is to the classification of three skin lesion classes that are melanoma, seborrheic keratosis and lupus Erythematosus precisely. For feature extraction, the CLDM based modified ABCD features has established to extract the skin lesion features like shape, color, size and texture from the dermoscopy images. Then, to select the skin lesion features, the ECFS method have employed. After the feature

ranking and selection method, the skin lesions have classified by SVM. The proposed SVM method provides improved classification results with better accuracy.

The major contributions of the research paper have illustrated as follows.

- A novel CLDM based modified ABCD feature extraction technique has proposed for extracting the skin lesion features.
- We proposed the SVM classification method to categorize the three skin lesion classes, such as seborrheic keratosis, melanoma, and lupus Erythematosus, with better performance.

This paper is structured as follows: The existing techniques for detection and classification of skin lesion have given in Section 2. The proposed methodology of skin lesion detection and classification `has illustrated in Section 3. The results and discussion for the classification of skin lesions have explained in Section 4. At last, the research paper has completed in Section 5.

2 Related work

In the past few decades, many researchers have proposed various methods to solve different problems of skin lesions feature extraction and classifications in Table 1.

In [10], Chatterjee S et al. have presented the multi-label ensemble multiclass skin lesion classifier method for the detection of skin lesions such as BCC, melanoma, nevus, and SK. However, the overall performances of skin lesion classification become low. Munshi S et al. [9] discussed the Gray level co-occurrence matrix (GLCM) based fractal-based regional texture analysis (FRTA) method for the feature extraction techniques of skin lesions. A multiclass classification technique has used in this research for skin lesion classifying strategy, but this method impedes the speed of the computation. Lu, Juan, et al. [22] described a Gabor method and SVM based MRF classification process for the removal of skin features and the categorization of skin diseases. In this study overall performances of sensitivity, specificity and accuracy obtained very low. Some researchers have presented various methods for the isolation of epidermal and melanocytic skin diseases. GMM based Mahalanobis classifier utilized for the skin lesion classification strategy has been presented by Gupta C et al. [14]. In this work, eleven different classes of skin lesions have considered as examinations. However, it is complicated to spot the skin lesion hence further utilized effective classification methodology. Wu, Zhe et al. [42] developed a method called the Convolutional neural network (CNN) for anomaly classification. In this work, six different types of skin lesion namely rosacea (ROS), seborrheic keratosis (SK), Squamous cell carcinoma (SCC), lupus erythematosus (LE), basal cell carcinoma (BCC), and actinic keratosis (AK) have classified for skin lesion investigation purposes. The performance obtained by these respective feature-based approach methods is low. Parekh R [2] introduced the GLCM and WDM method for feature extraction. The literature analyzed on skin diseases focuses that various schemes like neural network Multilayer Perceptron (MLP) classifier can be used for classifying the skin lesions for diagnosing purposes. However, the accuracy is stumpy in feature extraction methodology.

Alcon, Jose Fernandez et al. [12] applied the ABCD rule of dermatoscopy algorithm for removing features like asymmetry, border, color, and differential structures of skin lesion images and Bayesian Network (BN) modeling classifier for classifying the skin diseases. Thus the efficiency, sensitivity, and specificity obtained by this lesion classifier method become low.

Table 1 A survey of skin disease feature extraction and classification

Reference	Feature extraction method	Classification method	Identified skin lesion types	Limitation
[10]	Cross correlation technique	Multi-label ensemble multiclass skin lesion classifier	BCC, melanoma, SK, and nevus.	Performance of lesions classifier is less.
[9]	GLCM based FRTA method	Multiclass classification technique	Melanoma, dysplastic nevi, and BCC	Performance is less.
[22]	Gabor method	SVM based MRF classifier method	Psoriasis lesions	i. Sensitivity, specificity, and dice estimation are low. ii. Scaling characteristics are not clear.
[14]	GLCM technique	GMM based Mahalanobis classifier	Discoid Lupus Erythematosus, Discoid Eczema, Molluscum Contagiosum, Tinea Corporis, Acne Blackhead.	Efficiency is low in skin lesion identified method.
[42]	Feature-based approach	CNN	ROS, SK, SCC, LE, BCC, and AK.	Facial skin lesions identification performance is minimum
[2]	GLCM and WDM method	MLP	Acne, Eczema, Urticaria.	Accuracy is low in the feature extraction method.
[12]	ABCD rule of dermatoscopy algorithm.	BN modeling classifier	Benign, malignant.	Efficiency, sensitivity, and specificity are low in the lesion classifier method.
[18]	GLCM, LBP, RGB color channel features.	SVM classifier method.	Melanoma, nevus	Segmentation accuracy is low.
[7]	Wavelet-based texture features extraction and selection.	FFNN, SVM	Breast, retina, skin lesions.	The accuracy of the lesion classifier is low.
[37]	ABCD rule of dermatoscopy algorithm.	KNN classifiers	Malignant, benign.	The performance of skin lesion classifier accuracy is low.
[6]	ABCD rule of dermatoscopy algorithm.	AdaBoost classifier, SVM classifier and KNN classifier.	Melanoma, benign.	Specificity is low in the skin lesion classifier method.
[31]	GRAD-CAM	Deep depth wise separable residual Convolutional Algorithm	Melanoma, BCC SCC and AK.	Classification performance is low
[3]	Deep models	Convolutional Neural Network	SK, BCC, melanoma and nevus	Poor segmentation performance
[16]	Stochastic gradient descent (SGD) algorithms	Convolutional Neural Network	Seborrheic keratosis, nevus and Melanoma	Good accuracy performance with maximum confidence levels
[33]	Gabor method, histogram oriented gradients (HOG), edge histogram (EH) and local binary pattern (LBP)	Support Vector Machine (SVM)	Benign lesions or melanoma	Complex classification frame work

Local binary pattern (LBP) and SVM classifiers have applied for skin lesions feature extraction and classifying techniques have been proposed by Khan, Muhammad Qasim et al. [18]. In this work SVM is a supervised learning technique that has generally applied for classification purposes. This SVM classifies two classes using a hyperplane. Bodasingi N and Balaji N [7] have discussed a method called feed-forward neural network (FFNN) based SVM for precise and computationally effectual disease categorization from the dermoscopic images. However, the overall accuracy is less in classifying images. Valavanis I, et al. [37] have widely described a k-nearest neighbor (KNN) classifier for classification of two types of skin lesion namely malignant and benign dependent on skin color, texture, and shape. Using this method low classification performance has been achieved. Barata, Catarina et al. [6] have discussed three types of classifiers namely AdaBoost classifier, SVM and KNN classifier for precise classification from dermoscopy images. The main benefit of this work is that it highlights the evaluation of local and global characteristics for the widespread extraction of the feature with the making of excellent outcomes for both the evaluation. By this research, specificity obtained very low in the skin lesion classifier method. Sarkar R, Chatterjee C.C et al. [31] have described the Gradient weighted class activation maps (GRAD-CAM) and deep depth wise separable residual Convolutional network model have reliant on the texture and color description acquired from the skin disease imagery, which progresses the accuracy level. Thus this model is difficult to classify the three classes of skin lesions, namely BCC, SCC, and AK.

Akram et al. [3] proposed deep models for feature extraction. Further, the convolutional neural network (CNN) is used to classify the skin lesion from image. The dimensionality reduction process with entropy controlled Discriminant feature selection model effectively segment and select the relevant features with lower dimensionality. The performance efficiency is evaluated using different dermoscopic datasets such as ISIC UDA, ISIC MSK and PH^2 with various experimental analyses. This method reduces the time complexity during experimental analysis. The ensemble of deep convolutional neural network (DCNN) was proposed by Balazs Harangi [16] to enhance the classification accuracy. Here, the DCNN classifies three classes of skin diseases namely Seborrheic keratosis, nevus and Melanoma. The outcomes of classification layers of four different deep neural network structures are fused to accomplish higher classification accuracy. The U-net algorithm based convolutional neural network (CNN) was suggested by Seeja et al. [33] to segment the skin lesion images. The gabor method, histogram oriented gradients (HOG), edge histogram (EH) and local binary pattern (LBP) to extract the shape, color and texture features. The Benign lesions or melanoma based its classification is performed via Support Vector Machine (SVM).

3 Proposed approach

In this work, skin lesions images have identified and classified based on processes like preprocessing, segmentation, feature extraction, and image classification which has illustrated in Fig. 1. Moreover, the attention has been on the detection and classification of three skin diseases, namely melanoma, seborrheic keratosis (SK), Lupus Erythematosus (LE) precisely.

3.1 Preprocessing

The skin images are frequently prone to noise primarily due to hair and very bad light intensity. The preprocessing method has been engaged for noise and artifacts elimination of

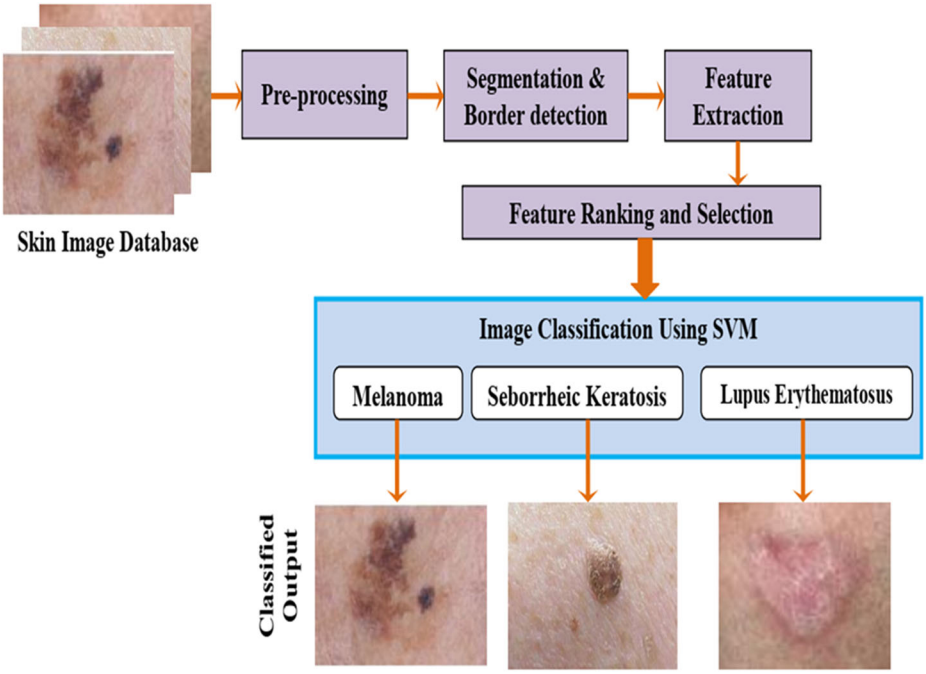

Fig. 1 Block diagram of proposed skin disease classification

image and also maintains improved generality skills. Colour consistency, normalization, and elimination of noise and hair removal are the three steps of the preprocessing method.

- *Colour consistency*

The medical images used in this work have achieved using different dermoscopy with diverse illumination conditions. The color alteration obtained in the image because of irregular light intensity is being rectified and normalized utilizing the gray world constancy method. The method evaluated the color image under unfamiliar illumination source and evaluated the average gray range of red (R), green (G), blue (B) co-ordinates [5].

- *Normalization*

The normalization method alters the range of pixel intensity values. The medical images have collected from different sources for this study. After the color consistency step, we use the normalization method to enhance the dynamic range and image intensity value. At this point, the measured average RGB values have subtracted from the entire training and test image dataset [20].

- *Elimination of Noise and hair removal*

The noise elimination and hair artifacts removal have an important procedure for a precise diagnosis of skin lesions before employing segmentation or feature extraction methods. The outcome of segmentation obtains imprecise results due to some artifacts present in the image. The Gaussian filter has utilized to eliminate the noise during the acquisition procedure. The

thick hair removal is a difficult task because it creates few hair shades in the image. The subsistence of hair shade produces various unwanted expansion of the lesion region. Therefore the image segmentation structure is altered [21]. Some morphological operations employed for the detection and removal of hair from the skin disease area.

Various filtering and thresholding methods have used for the purpose of hair removal. But this method is challenging because of imprecise detection, and information gets lost in the lesion region. The accurate choice of the structuring element (SE) such as size and shape guarantee the removal of proper object from the skin lesion area. The hair and other object from the lesion area have removed by the morphological operations. Morphological bottom hat filtering is employed if the skin area is darker than their other parts [8]. The circular structuring element is significant to remove the hair from the lesion area having a diameter based on the breadth of the hair. The kernel diameter is vital for the elimination of hair shade from the image. The thickness of the hair has evaluated by the ratio of the length of minor and major axis length. The SE has a diameter of four pixels distance effectively eliminates the thin hairs and SE having diameter six pixels distance or eight pixels distance remove thick hairs do not have any hair shade in the original image. Based on the size of the SE, the thin and thick hairs removed effectively.

3.2 Segmentation and border detection

After preprocessing, the segmentation method has employed to segment the skin lesion. Generally, the image segmentation has a difficult task to segment the skin lesion area as a result of alteration and intricacy of skin disease. In the present work, the size of SE has been selected based on the performance of segmentation. The SE uses a circular kernel for morphological closing operations because the majority of skin diseases have a circular shape. The accurate skin lesion region has acquired while subtracting the closed morphological image and complement of the original gray image [13]. By using this method, a slight difference among the lesion region and background has obtained. The segmented skin lesion region has acquired after the morphological closing operation, and then the range of threshold has been taken by decreasing the image variance. Once acquiring the probability and histogram, variance, and mean have been estimated for the range of threshold from minimum to maximum intensity image [3, 24].

The segmented image has covered with a circular binary mask for the elimination of shade corners. The diameter of the circle has altered in accordance with the major axis length of the segmented region of the lesion. Moreover, the diameter of the circle is greater than that of the length of the major axis. Following the creation of this binary circular mask having to alter location and diameter, this has multiplied by the segmented part. The skin lesion border is essential for the detection of diseases. For the border detection process, the morphological gradient method has utilized, but this method expands the lesion edge. In this study SE having a circular kernel with two pixels diameters initially erodes the segmented image. Afterward, the difference value is obtained from the original segmented image and the eroded image. This technique finds the pixel place where the maximum alters in intensity has occurred.

3.3 Feature extraction

Skin lesions are extracted based on their features like shape, texture, and color after the segmentation and border detection method. In this work, we use modified ABCD features and novel cumulative level differences mean (CLDM) based on GLDM for the feature extraction method.

3.3.1 Modified ABCD feature vector [41]

The novel ABCD method examines the asymmetry (A), border (B), color (C), Differential structures (D) have mainly utilized for the analysis of skin lesion by a dermatologist. The asymmetry index has computed by obtaining the entire area of the segmented binary mask and located the centroid on the lesion. By using the centroid, we split the lesion region into the upper and lower part along with the right and left part. The upper and lower halves area has differentiated $\Delta R_1$by employing an XOR method. The ratio R_1 among the area difference $\Delta R_1$and then the entire area has computed. The right and left halves have used to compute the ratio R_2. The average R_1 and R_2 are the asymmetry index R which has specified in the below equation.

$$R = \frac{R_1 + R_2}{2} \times 100 \tag{1}$$

The various actions that have employed for computing border irregularity like compact index, fractal index, edge abruptness, and color variation, which have explained below.

- *Compact index (CI):* The compact index is significant to find the border irregularity of skin lesions. The CI has computed as follows,

$$CI = \frac{LP^2}{4\Pi R} \tag{2}$$

Where *LP* is the perimeter of skin lesion, and *R* is the entire area**.**

- *Fractal dimension:* Fractal dimension (FD) is difficult patterns that are estimating how the particulars vary with the scale. This process used the box-counting method in which the edge of the lesion has found and then enclosed by the grid following that the different occupied boxes in the grid are enumerated. This method has repetitive by utilizing an advance grid with tiny boxes until getting the correct fractal pattern. The fractal dimension is employed to get the accurate fractal pattern, which has given in the below equation**.**

$$FD = \frac{\log E}{\log F} \tag{3}$$

Where *E* denotes the tiny number of boxes that enclosed the edge line and *F* denotes the inverse side length of the tiny box.

- *Edge abruptness (EA):* The radial distance variation is known as the edge abruptness. This edge abruptness is distinct where a skin lesion having uneven borders have diverse radial distance; however, which explains the distance among the centroid to the perimeter. This EA has computed by the subsequent procedure.

Step 1 Find the binary skin lesion centroid.

1 By employing a convolutional 2d filter, we scan the image; this helps to find the points on the perimeter of the lesion.
2 Compute the radial distance (RD) for every point which has specified in the below equation,

$$RD = \sqrt{(l - c_a)^2 + (m - c_b)^2} \tag{4}$$

Where c_a and c_b are the centroid coordinates a and b, and l and m are the coordinates of points on the perimeter of the lesion.

Step 4 Compute the edge abruptness by using the mean radial distance.

- *Color variation:* The variation of color has estimated by employing an RGB coordinate**.** The color variation reveals whether any change among the skin lesion and nearby skin is vanishing gradually, and sudden alteration has indicated as a warning. The color transition has computed by the mean(Q_m) and variance (Q_v) of the luminance gradient on the lesion border, which takes place for conversion of RGB segmented lesion into hue, saturation, and value (HSV) method.

Initially, to compute the color transition of skin lesion, convert the RGB segmented lesion into the HSV model. Next, the flags that match every color utilized to signify their existence. At last, every image flag has calculated that signifies the present color, and also the diameter has estimated by calculating each pixel of the minor axis length of the segmented skin lesion.

3.3.2 Novel cumulative level difference means

Gray level difference mean (GLDM) textures have computed by utilizing the two separated pixels of gray level difference by displacement D. Suppose an image *I(a, b)*, separation vector, $D = (\Delta a, \Delta b)$ and then a gray level difference is $G_D(a, b) = |\ G(a, b) - G(a + \Delta a, b + \Delta b)\ |$.

The probability density function of a gray level difference g with separation vector D has considered as follows,

$$P\left(\frac{g}{D}\right) = probability[G_D(a, b) = g] \tag{5}$$

The possibilities of the separation vector D are (0, s), (s, 0), (−s, −s), (−s, s), which denotes vertical axis, horizontal axis and two separations for diagonal where s is the inter-sample distance. Generally, few texture methods have measured $P\left(\frac{g}{D}\right)$ as contrast, energy, angular second moment. The cumulative level differences mean (CLDM) is proposed in this method for extracting the texture features. Let us imagine L probable level difference ranges among a pixel pair; the CLDM has evaluated as follows [16].

Step 1 Initially count the probable level difference range appeared in the image for a given inter-sample distance s. The definite counts of level difference have indicated as K_n because the pixel-pair detached by a distance s that contains a gray level difference n.

The probable level difference counts have specified to $\{K_0, K_1, K_2, .. \ldots\ldots\ldots, K_{L-1}\}$, where L = 256 for the intensity of the image.

1 Calculate the cumulative sum for the level difference counts, which has given by $\{K_0, K_0 + K_1, K_0 + K_1 + K_2, \ldots\ldots\ldots K_0 + . \ldots\ldots + K_{L-1}\}$. This step explains the related level difference which is the sum of its counts and previous level difference counts.

2 The mean of the CLDM is evaluated by

$$CLDM = \frac{K_0 + K_0 + K_1 + K_0 + K_1 + K_2 + \ldots\ldots\ldots + K_0 + \ldots\ldots\ldots + K_{L-1}}{L} \tag{6}$$

$$= \frac{L.K_0 + (L-1)K_1 + (L-2)K_2 + \ldots\ldots\ldots + K_{L-1}}{L} \tag{7}$$

$$= \sum_{n=0}^{L-1} \frac{(L-n)K_n}{L} \tag{8}$$

$$= K_0 + \frac{(L-1)K_1}{L} + \frac{(L-2)K_2}{L} + \ldots\ldots\ldots + \frac{K_{L-1}}{L} \tag{9}$$

The CLDM has shown in Eq. (9), which denotes the texture resemblance of a certain lesion. Here K_0 represents the number of counts at which two-pixel pairs separated by a distance s. Moreover, to improve the texture features the higher-level differences have added but with diverse weights in relation to the level difference range were included for assuring that this feature looks like texture resemblance [2]. Some sets of features utilized for classifying a few classes, but it cannot be effectual for another class. Furthermore cannot get improved classification if the higher number of extracted features. Hence, the feature ranking and selection method are important for selecting the features.

3.4 Feature ranking and selection

Feature ranking is a method for selecting the diverse extracted features to find the important features to be utilized in the classification method. This method gives better classification accurateness and minimizes the errors [23]. Moreover, minimizing the feature space dimensionality would minimize the difficulty of classification, timing of feature extraction, training and testing time, and also storage space used by the classification method. The Eigen-vector Centrality feature ranking and selection (ECFS) has employed in this study to rank each feature as a grade.

Step 1 An undirected graph has drawn wherein every node or vertex signifies an assured feature. Following the edges between the features will be evaluated. The set of features $F = \{f^{(1)}, \ldots., f^{(h)}\}$ where h represents the total feature. Here initially evaluate the ranking grade, then create an undirected graph G = {V, W} where V denotes the vertices, and W indicates the weighted edges among features. The adjacent matrix M

related with G is a $\{h \times h\}$ matrix containing the pair-wise possible elements c_{jk} where $1 \leq j, k \leq h$ explains among every feature pair-wise weighted edges and pair-wise possible elements that can be specified as,

$$c_{jk} = \varphi\left(f^{(j)}, f^{(k)}\right) \tag{10}$$

Here φ is the function which denotes the feature grade based on their differential range.

Step 2 Calculate the feature estimation metrics, such as fisher criterion (f_c) and mutual information (m_i). By employing f_c and m_i the kernel N occupies supervised and unsupervised metrics which has mentioned below,

$$f_c = \sum_{k=1}^{k} \frac{\left(\mu_{j,k} - \mu_j\right)^2}{{\sigma_j}^2}, {\sigma_j}^2 = \sum_{k=1}^{k} \left(\sigma_{j,k}\right)^2 \tag{11}$$

Here $\mu_{j,k}$ and $\sigma_{j,k}$ are the mean and standard deviation that indicating set k in the j^{th} feature whereas the entire j^{th} feature dataset has μ_j and σ_j.

The mutual information (m_i) is an important step to estimate every feature. Its samples specified as u and the known related class labels indicated as v.

$$m_i = \sum_{v \in V} \sum_{U \in t^{(j)}} P(u, v).\log\left(\frac{P(u, v)}{P(u).P(v)}\right) \tag{12}$$

Here $P(u, v)$ is the joint probability distribution. For each feature, we calculate the fisher criterion and mutual information.

Step 3 The kernel R is a $h \times h$ matrix computed by using the dot product.

$$R = \left(f_c.{m_i}^T\right) \tag{13}$$

Step 4 The σ for entire feature pair (j, k) has given by $\sigma(j, k) = \max(\sigma^{(j)}, \sigma^{(k)})$. The adjacency matrix D related to the graph G has constructed as follows,

$$D = \gamma R + (1 - \gamma)\sigma \tag{14}$$

Here $\gamma$is a loading coefficient that has value among 0 and 1. At last, the Eigen-values and vectors of D have computed, and the ranking vector would be the Eigen-vector together with the greatest Eigen-values.

3.5 Image classification using support vector machine

The support vector machine (SVM) is an important image classification method. The major target of the SVM classification is to find the hyperplane that maximizes the margin among the two classes. This classifier performs better accuracy and good classification results compared to other classification methods. Moreover, the SVM evade the trouble of overfitting confronted the neural networks [44].

In this proposed work, the classifier used to identify and classify the skin lesions. The accuracy of classification must be better to make sure the correct classification in every step of the classifier. Here we classify the three skin diseases, namely melanoma, seborrheic keratosis, and Lupus Erythematosus which has given in Fig. 2.

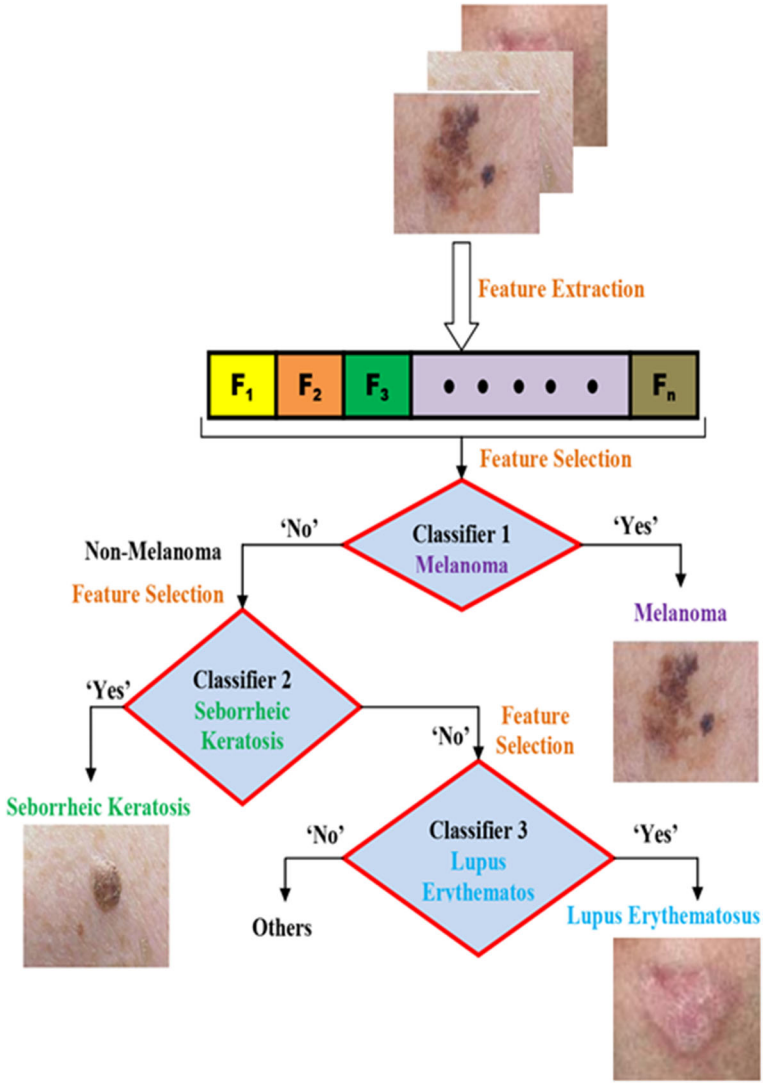

Fig. 2 Flowchart for Classification of three skin diseases

The first classifier is being used to discriminate the images of melanoma from other skin diseases and then consider positive class is a melanoma, and other diseases consider a negative class. Before the skin lesion classification by employing the SVM selection method, we greatly discriminate the features of melanoma from other skin deformities. The second classifier is being used to detach the seborrheic keratoses from lupus Erythematosus. But in the binary classification phase, categorize the images as seborrheic keratoses and others. Here others represent the misclassified images of melanoma, seborrheic keratoses, and lupus Erythematosus. Before the classification, another SVM-RFE feature selection method has presented for better classification accuracy of seborrheic keratoses from Lupus Erythematosus and misclassified images. In the end, the third classifier has employed for the separation of Lupus Erythematosus from the other misclassified images. In this study, the SVM classification method classifies the three skin diseases.

4 Results and discussions

In this proposed work, MATLAB R2017a software on Intel core i5-2410M 2.3 GHz processor has employed in this implementation work. Consider images of skin lesions from a diverse database having image sizes varied from 540 × 722 into 4499 × 6748 pixels. Then resize the entire images as a 512 × 512 pixel. In the classification method, three skin disease classes have trained by utilizing sample images. For the training stage, consider 80% training samples take 5263 images, for 72% training samples take 4605 images, for 65% and 55% training samples take 3497, and 3290 images have used for the training of testing samples. The data augmentation method has utilized to decrease the difficulty of overfitting and data disproportion. The image rotation and shifting techniques have operated in the data augmentation method. The image is rotated by 90^0, 180^0 and 270^0 shifting operation performed by shifting eight pixels rightwards and leftwards. To make sure the data balance during the training stage of the classification method, the skin lesion class image has been increased by utilizing the data augmentation method.

4.1 Dataset description

Dermnet Nz database, Xiangya derm dataset, Ham 10,000 dataset, medicine net dataset, PH_2 database, Kaggle dataset are some datasets for images of skin lesion. Among these, we have taken three datasets, which are the Dermnet Nz database, the Xiangya derm dataset, and Ham 10,000 dataset. From these datasets, we have taken three skin diseases namely melanoma, seborrheic keratosis, and lupus Erythematosus which has depicted in Fig. 3.

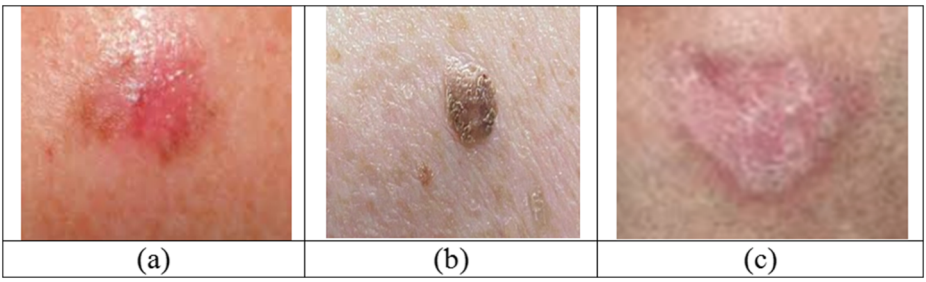

Fig. 3 Sample dataset for skin lesion images (**a**) Melanoma (**b**) Seborrheic keratosis (**c**) Lupus Erythematosus

4.2 Performance measures

The performance of segmentation has been assessed by estimating the accuracy (AC), specificity (SF), sensitivity (SN), and computing image resemblance indexes such as the Jaccard similarity index (JSI) and the Dice similarity coefficient (DSC). The performance of segmentation has been stated in the subsequent formulas.

$$Accuracy\,(AC) = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{15}$$

$$Sensitivity\,(SN) = \frac{T_P}{T_P + F_N} \tag{16}$$

$$Specificity\,(SF) = \frac{T_N}{T_N + F_P} \tag{17}$$

HereF_P,T_P,F_N, and T_N represent pixels of false positive, true positives, false negative, and true negative. The Jaccard similarity index (JSI) of two sets M and N have stated below,

$$JSI(M,N) = \frac{|M \cap N|}{|M \cup N|} \tag{18}$$

Here |M| denote the cardinal set M. The JSI has denoted based on FP, FN, and TP which has given as follows,

$$JSI(M,N) = \frac{TP}{FP + FN + TP} \tag{19}$$

Likewise, the Dice similarity coefficient (DSC) of two sets M and N has stated given below,

$$DSC(M,N) = 2*\frac{|M \cap N|}{|M| + |N|} \tag{20}$$

Here |M| denote the cardinal set M. The DSC has denoted based on FP, FN, and TP gives as follows,

$$DSC(M,N) = 2*\left(\frac{TP}{2*FP + FN + TP}\right) \tag{21}$$

The performance indexes of segmentation have indicated in Table 2. This table indicates three datasets such as the Dermnet Nz database, the Xiangya derm dataset, and Ham 10,000 dataset. Then, this table has shown the performance indexes of segmentation for maximum, minimum, and average values. The performance of segmentation has been measured in terms of sensitivity, specificity, accuracy, JSI, and DSC.

The mean, entropy, and contrast have been estimated for the texture feature extraction method, which has depicted given below.

Table 2 Performance for segmentation images of melanoma, seborrheic keratosis and lupus Erythematosus of three datasets

Dataset	*Values*	*Performances indexes of segmentation*				
		Accuracy	*Sensitivity*	*Specificity*	*JSI*	*DSC*
Dermnet Nz database	Minimum	0.952	0.947	0.967	0.958	0.949
	Maximum	0.992	0.986	0.983	0.990	0.985
	Average	0.966	0.959	0.969	0.974	0.976
Xiangya derm dataset	Minimum	0.959	0.950	0.979	0.968	0.965
	Maximum	0.991	0.984	0.989	0.984	0.982
	Average	0.968	0.965	0.976	0.972	0.981
Ham 10,000 dataset	Minimum	0.967	0.965	0.975	0.973	0.987
	Maximum	0.993	0.987	0.986	0.989	0.992
	Average	0.972	0.962	0.975	0.972	0.987

$$Mean = \frac{1}{V}\sum_c c.P_D(c) \tag{22}$$

$$Entropy = -\sum_c P_D(c).\log(P_D(c)) \tag{23}$$

$$Contrast = \sum_c c^2.P_D(c) \tag{24}$$

Here $P_D(c)$ is the probability density function of a pixel pair detached by a distance D and c component is the probability that $C_D(a, b)$ would have the value one.

The proposed method comprises preprocessing, segmentation and border detection, feature extraction, feature ranking, and selection, and image classification method, which have been used to attain better accurateness for the detection of diverse skin lesion diseases. In the preprocessing step removal of hair and noise performed, which has summarized in Fig. 4.

After the preprocessing step, the segmentation technique is being employed to segment the skin lesions. In the segmentation method, segment the skin disease from the dermoscopic images by the morphological closing operation has used to. Generally, segmentation is the more difficult process because of its uneven structure. The original images and segmented skin lesion images have depicted in Fig. 5.

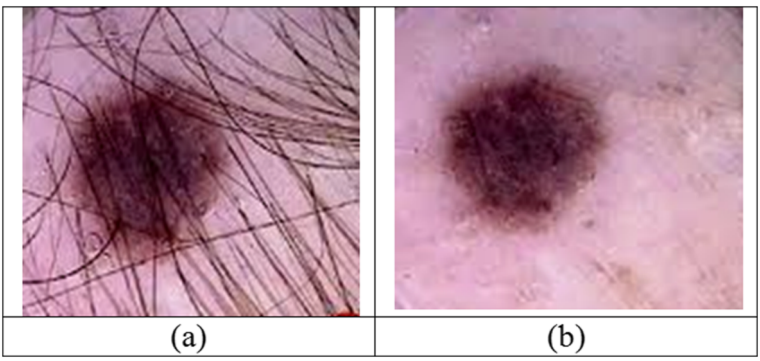

Fig. 4 **a** Original image **b** Preprocessing image

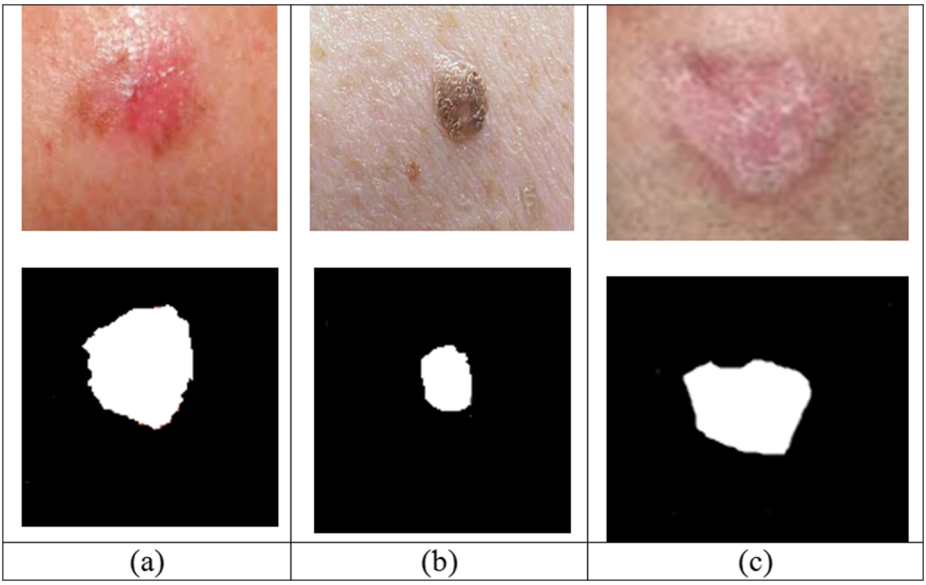

Fig. 5 Skin lesion segmentation results for (**a**) melanoma, (**b**) Seborrheic keratosis, (**c**) Lupus Erythematosus. First row: Original skin lesion image Second row: Segmentation output of skin lesion image

After the segmentation process, the feature has extracted by a method called CLDM based modified ABCD features. Then, the features have selected by the ECFS method and classified the skin disease classes by means of the SVM algorithm. The performance of skin lesion classification has depicted in Table 3. Here we have taken three classifiers namely classifier 1, classifier 2, and classifier 3, with diverse training and testing samples. Each classifier has different classification accuracy, sensitivity, and specificity. The accuracy of classification has been estimated by reducing training samples from 70% to 40% of the whole database.

Table 4 depicts the performances of skin lesion classification using different datasets. In this experiment, we have chosen the Dermnet Nz database, the Xiangya derm dataset, and Ham 10,000 datasets with training, testing samples with accuracy, specificity, and sensitivity measures for skin lesion classification. More number of samples is used for training and few of these samples are utilized for testing purpose. The Ham 10,000 dataset takes 90% training

Table 3 Classification performance of SVM classifier

Classifiers	Training samples	Testing samples	Accuracy	Sensitivity	Specificity
Classifier 1 (Melanoma)	70%	30%	99.7%	99.2%	99.6%
	60%	40%	98.3%	98.5%	98.2%
	50%	50%	97.6%	95.6%	96.4%
	40%	60%	96.5%	95.4%	94.6%
Classifier 2 (Seborrheic keratosis)	70%	30%	99.2%	99.4%	99.3%
	60%	40%	97.5%	98.2%	98.6%
	50%	50%	95.6%	97.9%	97.4%
	40%	60%	95.4%	96.8%	95.9%
Classifier 3 (Lupus Erythematosus)	70%	30%	99.6%	99.4%	99.5%
	60%	40%	98.2%	98.9%	98.4%
	50%	50%	97.7%	97.3%	97.6%
	40%	60%	96.4%	95.9%	96.8%

Table 4 Performance of skin lesion classification using different datasets

Name of the dataset	Training samples	Testing samples	Accuracy	Specificity	Sensitivity
Ham 10,000 dataset	90%	10%	98.91%	99.14%	98.09%
Dermnet Nz database	60%	40%	99.07%	99.79%	99.67%
Xiangya derm dataset	85%	15%	99.56%	98.12%	98.98%

and 10% testing samples with 98.91% accuracy, 99.14% specificity, and 98.09% sensitivity. Further, 60% of training, and 40% testing samples with 99.07% accuracy, 99.79% specificity, and 99.67% sensitivity results are obtained for the Dermnet Nz database. The Xiangya derm dataset yields 85% training and 15% testing samples with 99.56% accuracy, 98.12% specificity, and 98.98% sensitivity results.

4.3 Comparative analysis

The proposed method of segmentation has compared with the different existing methods such as gray-level thresholding [4], k means clustering algorithm [32], watershed-based segmentation [15], and edge detection based segmentation [29]. When the proposed method has compared with existing methods, our proposed segmentation provides the best segmentation outputs. The performance range of accuracy, sensitivity, specificity, JSI, and DSC of proposed segmentation is 0.987, 0.983, 0.953, 0.976, and 0.973 which has described in Table 5.

In the feature extraction method, we use the CLDM based modified ABCD feature algorithm for extracting the skin lesion features from the segmented image. In this work, the proposed feature extraction method has compared with other existing methods like the GLCM method [9], ABCD rule [12], GRAD-CAM method [31], and LBP method [18] which has depicted in Fig. 6. The proposed CLDM based modified ABCD features provides better results when compared with other existing methods**.**

In this paper, skin lesion has classified employing the SVM method. The proposed SVM classifier has compared with other existing methods like KNN classifier [37], FFNN classifier [7] and BN modeling classifier [12]. Here we select 10, 20, 30, 40, 50 and 60 features and provides to the classifier for the detection of diseases. Then using each feature, the proposed SVM classifier has compared with other existing methods, which has shown in Fig. 7. The feature selection is significant for classification performance. From this figure, the proposed classification accuracy produces optimal results.

Table 6 explains the performance evaluation feature selection process. In this experiment, we have chosen different kinds of features such as 10, 20, 30, 40, and 50 with various state-of-

Table 5 Performance evaluation of the proposed method with conventional techniques

Methods	Performance evaluation				
	Accuracy	Sensitivity	Specificity	JSI	DSC
Gray level thresholding	0.742	0.851	0.934	0.954	0.960
K means clustering	0.870	0.900	0.947	0.898	0.944
Watershed based segmentation	0.923	0.917	0.870	0.912	0.856
Edge detection based segmentation	0.954	0.934	0.869	0.923	0.917
Proposed method	0.987	0.983	0.953	0.976	0.973

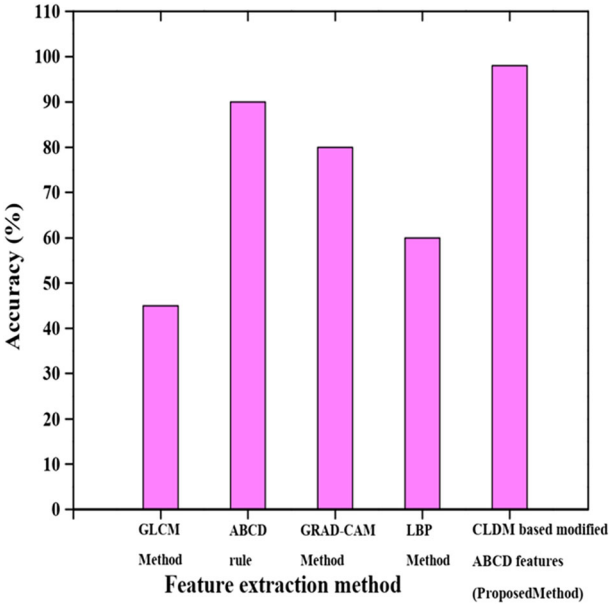

Fig. 6 Comparison of proposed feature extraction accuracy with the existing method

art methods including Wavelet-based methods, Multi-level feature selection model, Eigenvector model, and the proposed method. The proposed feature selection process accomplishes 0.870, 0.934, 0.953, 0.967 and 0.984 results. Finally, the proposed method yields higher feature selection performance than other methods.

The proposed SVM classifier has compared with other existing methods like KNN classifier, FFNN classifier, Naïve Bayes, and BN modeling classifier. The classification algorithm for the proposed SVM provides better results, and it has indicated in Table 7. The

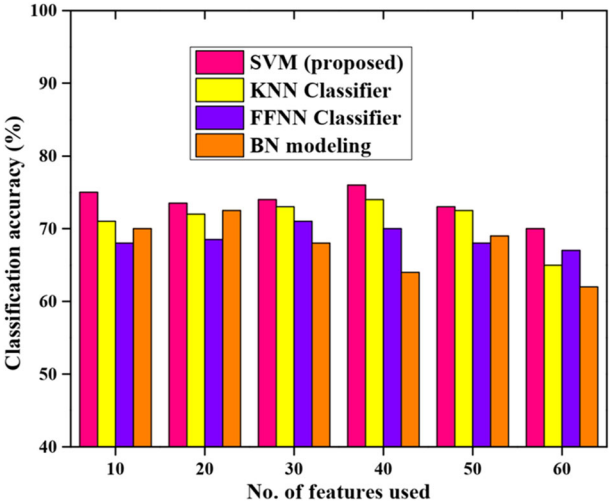

Fig. 7 Comparison of classification accuracy with existing techniques

Table 6 Performance evaluation feature selection

Methods	Number of features used				
	10	20	30	40	50
Wavelet based methods	0.672	0.751	0.799	0.654	0.720
Multi level feature selection model	0.702	0.870	0.807	0.758	0.844
Eigen vector model	0.701	0.917	0.805	0.812	0.896
Edge detection based segmentation	0.854	0.804	0.859	0.723	0.817
Proposed method	0.870	0.934	0.953	0.967	0.984

classification method has compared with different performance evaluations such as accuracy, sensitivity, specificity, JSI, and DSC. Here the proposed SVM classifier with the accuracy is 97%, specificity is 98%, sensitivity is 97%, JSI is 97% and DSC is 98%, respectively.

5 Conclusion

In this research, the combined approach has classified the three skin lesions classes, namely melanoma, seborrheic keratosis, and lupus Erythematosus using the SVM classification method. The CLDM based modified ABCD features have been chosen for the effectual feature extraction of the skin lesion images. The ECFS method was introduced for the feature selection method before the phase of classification technique to progress the performance of classification. The different testing and training samples have employed in each dataset. The performances have been computed by assessing the accuracy, sensitivity, specificity, JSI, and DSC. The proposed method has compared with other existing methods like KNN classifier, FFNN classifier, and BN modeling classifier to prove high performance. The three skin lesion classes have been selected from the Dermnet Nz database, Xiangya derm dataset, and Ham 10,000 dataset. The proposed SVM method provides better classification performance in terms of the accuracy is 97%, specificity is 98%, sensitivity is 97%, JSI is 97% and DSC is 98% has obtained. The proposed method has few restrictions despite better performance. Among the various skin abnormalities but this research work considered the three skin lesion classes for disease classification. In the future, the technique has to be implemented in an embedded system, hence it makes the cost becomes minimum.

Table 7 Comparison of performance metrics for skin lesion classification

Classification methods	Performance metrics				
	Accuracy	Specificity	Sensitivity	JSI	DSC
KNN classifier	95%	91%	89%	92%	88%
FFNN classifier	92%	89%	77%	83%	89%
BN modeling classifier	94%	73%	86%	79%	86%
Naïve bayes	88%	86%	76%	72%	81%
SVM classifier (proposed method)	97%	98%	97%	97%	98%

References

1. Adegun A, Viriri S (2020) Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art. Artificial Intelligence Review pp 1–31
2. Akar M, Sirakov NM (2019) Support vector machine skin lesion classification in Clifford algebra subspaces. Appl Math 64(5):581–598
3. Akram T, Lodhi HMJ, Naqvi SR, Naeem S, Alhaisoni M, Ali M, Haider SA, Qadri NN (2020) A multilevel features selection framework for skin lesion classification. Hum Cent Comput Inf Sci 10:12
4. Ansari MA, Anand RS (2007) Region-based segmentation and image analysis with application to medical imaging. In: IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007), p 724–729
5. Barata C, Celebi ME, Marques JS (2014) Improving dermoscopy image classification using color constancy. IEEE J Biomed Health Inform 19(3):1146–1152
6. Barata C, Ruela M, Francisco M, Mendonca T, Marques JS (2013) Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Syst J 8(3):965–979
7. Bodasingi N, Balaji N (2017) Classification of multiple diseases based on wavelet features. J Eng 4:110–118
8. Chatterjee S , Dey D , Munshi S (2015) Mathematical morphology aided shape, texture and color feature extraction from skin lesion for identification of malignant melanoma. In Proc. IEEE CATCON pp. 200–203.
9. Chatterjee S, Dey D, Munshi S (2019) Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification. Comput Methods Prog Biomed 178:201–218
10. Chatterjee S, Dey D, Munshi S, Gorai S (2019) Extraction of features from cross-correlation in space and frequency domains for classification of skin lesions. Biomed Signal Process Control 53(101581):101581
11. Cula OG, Dana KJ, Murphy FP, Rao BK (2004) Bidirectional imaging and modeling of skin texture. IEEE Trans Biomed Eng 51(12):2148–2159
12. Fernandez AJ, Ciuhu C, Ten Kate W, Heinrich A, Uzunbajakava N, Krekels G, Siem D, De Haan G (2009) Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J Sel Top Sign Proces 3(1):14–25
13. Ferreira A, Gentil F, Tavares JM (2014) Segmentation algorithms for ear image data towards biomechanical studies. Comput Methods Biomech Biomed Engin 17(8):888–904
14. Gupta C, Gondhi NK, Lehana PK (2019) Analysis and identification of dermatological diseases using Gaussian mixture modeling. IEEE Access 7:99407–99427
15. Han B (2015) Watershed segmentation algorithm based on morphological gradient reconstruction. 2nd international conference on information science and control engineering IEEE pp. 533-536.
16. Harangi B (October 2018) Skin lesion classification with ensembles of deep convolutional neural networks. J Biomed Inform Vol 86:25–32
17. Jiji GW, Raj PSJD (2014) Content-based image retrieval in dermatology using the intelligent technique. IET Image Process 9(4):306–317
18. Khan MQ, Hussain A, Rehman SU, Khan U, Maqsood M, Mehmood K, Khan MA (2019) Classification of melanoma and nevus in digital images for diagnosis of skin Cancer. IEEE Access 7:90132–90144
19. Kia S, Setayeshi S, Shamsaei M, Kia M (2013) Computer-aided diagnosis (CAD) of the skin disease based on an intelligent classification of sonogram using neural network. Neural Comput & Applic 22(6):1049–1062
20. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Adv Neural Inf Proces Syst pp1097–1105.
21. Lee I, Du X, Anthony B (2017) Hair segmentation using adaptive threshold from edge and branch length measures. Comput Biol Med 89:314–324
22. Lu J, Kazmierczak E, Manton JH, Sinclair R (2012) Automatic segmentation of scaling in 2-d psoriasis skin images. IEEE Trans Med Imaging 32(4):719–730
23. Mendes DB, da Silva NC (2018) Skin lesions classification using convolutional neural networks in clinical images. arXiv preprint arXiv:1812.02316
24. Ma Z, Tavares J M RS, Jorge RM N (2009) A review on the current segmentation algorithms for medical images. In Proceedings of the 1st International Conference on Imaging Theory and Applications (IMAGAPP), 2009.
25. Mirzaalian H, Lee TK, Hamarneh G (2013) Spatial normalization of human back images for dermatological studies. IEEE J Biomed Health Inform 18(4):1494–1501

26. Mower WR, Kadera SP, Rodriguez AD, Vanderkraan V, Krishna PK, Chiu E, Wilson MJ, Gupta M, Krishnadasan A, Moran GJ, Talan DA (2019) Identification of clinical characteristics associated with high-level care among patients with skin and soft tissue infections. Ann Emerg Med 73(4):366–374
27. Mullangi P, Rao YS, Kotipalli P (2018) Texture and clustering-based skin disease classification. Sensors and Image Processing Springer, Singapore, pp 103–110
28. Oliveira RB, Mercedes Filho E, Ma Z, Papa JP, Pereira AS, Tavares JM (2016) Computational methods for the image segmentation of pigmented skin lesions: a review. Comput Methods Prog Biomed 131:127–141
29. Padmapriya B, Kesavamurthi T, Ferose HW (2012) Edge-based image segmentation technique for detection and estimation of the bladder wall thickness. Proce Eng 30:828–835
30. Rejeesh MR (2019) Interest point based face recognition using adaptive neuro fuzzy inference system. Multimed Tools Appl 78(16):22691–22710
31. Sarkar R, Chatterjee CC, Hazra A (2019) Diagnosis of melanoma from dermoscopic images using a deep depth wise separable residual convolutional network. IET Image Process 13(12):2130–2142
32. Sarpe AI (2010) Image segmentation with clustering K-means and watershed transform. Second international conferences on advances in multimedia IEEE pp.13–17.
33. Seeja RD and Suresh A (2019) Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). Asian Pac J Cancer Prev. 20(5).
34. Sundararaj V (2019) Optimised denoising scheme via opposition-based self-adaptive learning PSO algorithm for wavelet-based ECG signal noise reduction. Int J Biomed Eng Technol 31(4):325
35. Sundararaj V, Muthukumar S, Kumar RS (2018) An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. Comput Secur 77:277–288
36. Sundararaj V, Anoop V, Dixit P, Arjaria A, Chourasia U, Bhambri P, Rejeesh MR, Sundararaj R (2020) CCGPA-MPPT: Cauchy preferential crossover-based global pollination algorithm for MPPT in photovoltaic system. Prog Photovolt 28(11):1128–1145
37. Valavanis I, Maglogiannis I, Chatziioannou AA (2014) Exploring robust diagnostic signatures for cutaneous melanoma utilizing genetic and imaging data. IEEE J Biomed Health Inform 19(1):190–198
38. Verma AK, Pal S, Kumar S (2019) Comparison of skin disease prediction by feature selection using ensemble data mining techniques. Inform Med Unlocked (IMU) 16:100202
39. Vinu S (2016) An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. Int J Intell Eng Syst 9(3):117–126
40. Vinu S (2019) Optimal task assignment in mobile cloud computing by queue based ant-bee algorithm. Wirel Pers Commun 104(1):173–197
41. Wahba MA, Ashour AS, Guo Y, Napoleon SA, Elnaby MMA (2018) A novel cumulative level difference mean based GLDM and modified ABCD features ranked using eigenvector centrality approach for four skin lesion types classification. Comput Methods Prog Biomed 165:163–174
42. Wu Z, Zhao S, Peng Y, He X, Zhao X, Huang K, Wu X, Fan W, Li F, Chen M, Li J (2019) Studies on different CNN algorithms for face skin disease classification based on clinical images. IEEE Access 7: 66505–66511
43. Zhang X, Cui J, Wang W, Lin C (2017) A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. Sensor 17(7): 1474
44. Zien A, Ratsch G, Mika S, Scholkopf B, Lengauer T, Muller KR (2000) Engineering support vector machine kernels that recognize translation initiation sites. Bioinformatics 16(9):799–807