Measurement 163 (2020) 107922

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier

V.R. Balaji [a,*], S.T. Suganthi [b], R. Rajadevi [c], V. Krishna Kumar [d], B. Saravana Balaji [e], Sanjeevi Pandiyan [f]

[a] Department of ECE, Sri Krishna College of Engineering and Technology, Coimbatore, India
[b] Dept of Computer Engineering, Lebanese French University, Iraq
[c] Department of Information Technology, Kongu Engineering College, Perundurai, India
[d] Dept of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, Coimbatore, Tamil Nadu, India
[e] Department of IT, Lebanese French University, Iraq
[f] Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Article history:
Received 13 February 2020
Received in revised form 17 April 2020
Accepted 29 April 2020
Available online 1 June 2020

Keywords:
Skin lesion
Filtering
Transformation
Symptoms
Computed tomography
Colour and texture features
Graphs
Sensitivity
Specificity

ABSTRACT

The largest organ and the outer covering of the human body is the skin. With seven layers of it covering the other organs inside, skin is one of the important part to take care of. A skin condition is one which affects the integumentary system and that includes a wide variety of diseases including dermatoses. Classifications of these skin conditions are always a challenge for any medical practitioner and they look at the machine learning systems to assist them in predicting and classifying the skin conditions. This in turn will help to cure or at least reduce the effect. If the skin symptoms such as acne, cellulitis, candidiasis, varicella, scleroderma, fungal skin, psoriasis, inflamed skin condition, etc. are left without treatment in its initial stage, then they can effect in different health impediments leading to even death. Image partitioning is a method which supports with the skin disease detection. Any abnormal skin growth is referred to as skin lesion which could either be primary or secondary. Graph cut algorithms are debated and used in the literature for variety of purposes including image smoothing, image segmentation and other problems involving energy minimization as objective. In this work, we intend to use a novel dynamic graph cut algorithm for skin lesion segmentation followed by a probabilistic classifier called as Naïve Bayes classifier for skin disease classification purposes. We have used ISIC 2017 dataset for testing our proposed method and found that the results outperform many state of the art methods including FCN and SegNet by 6.5% and 8.7% respectively. This dataset is available at the International Skin Imaging Collaboration (ISIC) website for public study and experimentation. In terms of accuracy, we could achieve 94.3% for benign cases, 91.2% for melanoma and 92.9% for keratosis on this data set.

1. Introduction

Skin being the largest organ, the various viral, bacterial and inflammatory disease effects gets spread all around and lead to a wide variety of health related issues. The skin diseases include acne, atopic dermatitis, alopecia, morphea, melanoma, photo aging, wounds, psoriasis, wrinkles, vitiligo, etc. as shown in Figure 1. People can be cured from many of these diseases if it's identified at the early stages when the spread has not occurred. Dermoscopy is a method used by the specialists to examine the skin changes with the help of a bright light and employs polarisation in order to reduce the surface reflection.

More than 100 million people are living with different kinds of skin infections all over the world. The skin cancer is also rapidly increasing without much solutions to cure. Among them, melanoma is the most deadliest as well as the diversifying type skin cancer. Early stage investigations can help to reduce the skin disease spreading. In this work, we aim to provide a new dynamic graph cut algorithm for segmenting the affected regions in the skin followed by a probabilistic classifier for classifying the skin disease type. The graph cut method brings in the advantages of accuracy and performance as compared to other image segmenting methods [1]. On the other hand, Naïve Bayes classifier are super simple and they also tend to converge quicker than other equivalent discriminative models like logistic regression which means it requires only less training data.

* Corresponding author.
E-mail addresses: balajivr@skcet.ac.in (V.R. Balaji), suganthi.sb@gmail.com (S.T. Suganthi), rajdevi.it@kongu.edu (R. Rajadevi), krish232007@gmail.com (V. Krishna Kumar), saravanabalaji.b@gmail.com (B. Saravana Balaji), gpsanjeevi@jiangnan.edu.cn (S. Pandiyan).

https://doi.org/10.1016/j.measurement.2020.107922
0263-2241/

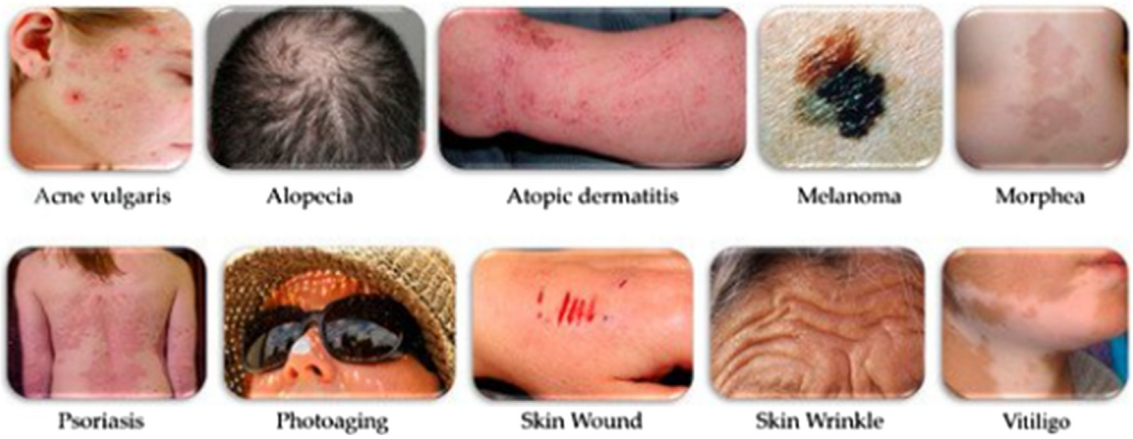

Fig. 1. A wide variety of skin diseases.

The skin lesions could be either primary or secondary. While the primary skin lesions include spots, pimples, plaque, discoloration, nodule, tumour, vesicle, pustule, cyst and bulla, the secondary ones comprise crust, erosion, excoriation, scale, ulcer, fissure, induration, atrophy, maceration, umbilication and phyma. The two terms that are related to the skin lesions are configuration and distribution. Here configuration corresponds to the grouping of the lesions and the distribution tells about the localization of lesions. While in the literature, it is mostly discussed about the major three types of skin infections like melanoma, squamous cell and seborrheic keratosis, in this work we have covered more skin related image segmentation and classification. Similarly, supervised grouping algorithms like SVM are very common but comes with several key parameters to be identified and set properly for better results, we preferred to go with Naïve Bayes classifier which is much simpler and accurate in classification.

While image is a collection of different pixels, image partitioning refers to the process of separating this digital image in to multiple regions or objects and then process them independently based on the need and application. It simplifies the image representation in to more meaningful way for easy analysis and classification [2]. Pixels in the same segment carry the same characteristics while they differ across clusters. A set of outlines that are mined from the image is further processed to label or predict based on the requirement. Though the modest method of image partitioning is through thresholding, there are various other methods discussed in the literature which are much more effective than a simple thresholding in terms of efficiency and effectiveness. This include data clustering algorithms like K-means, colour or intensity based histogram methods, region growing methods, edge detection methods, PDE based methods or graph portioning methods to name a few.

Spectral based image segmentation is a graph partitioning problem which makes use of the homogeneity in image regions and are effective as well [3]. The pixels in the images are generally associated with the edges and nodes of the graph which helps to define the neighbourhood pixels similarity or dissimilarity. The algorithm outputs partition of the nodes and each of them are considered to be the image object segments. Minimum cut, random walker, normalized cuts, isoperimetric partitioning, etc. are some commonly used algorithms under graph cut based methods for image segmentation.

Graph cut methods are used to divide a graph. Here graph is a set of data represented with pairs that re connected to each other via some links. The elements here are termed as nodes or vertices and the links corresponds to arcs or edges. A graph is normally represented with the term G = (V, E) where V signifies the vertices set and E to the edges set, found between the vertices. Though traditional graph based algorithms carry the advantages of having good spatial continuity property, they fail to detect long range cues [4]. This pushes the need for developing an extension to the existing graph based algorithms that can smooth local regions and also capture the texture details with the help of neighbourhood cue. A robust framework is essential to build a dynamic graph cut algorithm and then classify the regions for the identified problem statement.

In this work, we propose a modified graph cut method which is effective than the traditional graph cut algorithms due to its flexibility in minimizing the groups of alike functions that are decipherable in polynomial time as well. New instances are solved by taking references from the past experiences instead of a fresh computation thereby improving the running time considerably. One of the reliable probabilistic framework for labelling uses random fields which can handle and model the complex hidden variables interactions in a simple and precise way. Though minimizing is a NP-hard problem in nature, some families of energy functions can be used to solve it in polynomial time. This is discussed and proposed in this work along with support vector machines finally for classification of the identified segments. It is also observed that the proposed solutions are better in terms of accuracy than the ones obtained through other inference methods.

A probabilistic classifier will be able to predict and output over a set of classes rather than just giving a most likely class with most methods. A naïve Bayes classifier [5] is one such probabilistic method with strong independence between the features. They are more scalable and assume that a particular feature value is self-determining of the other feature values, given the class variable. In a supervised learning setting, they can be trained very efficiently and is a conditional probability model. In this work, we use them for classifying different skin diseases after segmentation.

The rest of this paper is organised as follows: Section 2 discusses the approaches in the literature relevant to this work, Section 3 details the recommended dynamic graph cut algorithm, Section 4 deals with feature mining from the segmented skin regions, Section 5 presents the Naïve Bayes classification method while Section 6 shows the investigational results and lastly Section 7 completes the paper with further scope of research related to this work.

2. Related work

Hundreds of skin conditions affect humans and dermatologists can help to diagnose them for curing and to avoid spreading. Skin

diseases are getting one of the most leading causes of nonfatal disease burden globally. Prevention of these diseases needs to be prioritised and multiple diagnosis schemes related to this issue is discussed in the literature. Skin cancers including carcinoma, melanoma etc. often starts as skin changes and grow rapidly across the body. Their signs can differ in terms of colour, form or borders and they can also be different in size and evolution. So each melanoma is unique in its appearance and the ABCD criteria will help to suspect if they are dangerous or not. To assist the physicians, images containing skin lesions are analysed using Dermoscopy and image processing techniques are generally employed to classify the presence of these dangerous cells in the skin.

Pravin S. Ambad and A. S. Shirsat have used statistical parameters to diagnose multiple skin diseases and it is a combo model to classify skin diseases such as melanoma, psoriasis and dermo [6]. The statistical parameters they have used include texture index, entropy, correlation data and standard deviation. They have presented an image analysis system comprising of image acquisition, noise removal, feature extraction and finally an image classifier. Their database consists of around 130 images with diverse kinds of skin infections present in it.

A prototype that is capable of segmenting and classifying the skin damages present in dermoscopy images is developed by Zaqout [7] and they have used the ABCD rule for the same. His method first starts with the pre-processing which takes care of filtering and improving the contrast of the images followed by region of interest segmentation using thresholding and statistical procedures and then goes for feature extraction where the asymmetry is calculated and finally ends with classification where the images are classified in to benign, malignant or suspicious types. He has implemented the proposed approach using MATLAB libraries and the Hospital Pedro Hispano and Matosinhos datasets. The results are encouraging with an accurateness of 90%, specificity of 92.22% and sensitivity of 85% respectively.

Wei et al. [8] have discussed about developing automatic methods for increasing the diagnosis accuracy for different kinds of skin diseases. Their proposed method starts with image pre-processing in order to remove the noise, followed by GLCM matrix to segment the regions of interest and the features were finally extracted to classify the images using support vector machines. The authors have reduced the irrelevant variables with the help of image filtering and Euclidean distance transformation. Ten different vertical images can be attained by the vertical image division, refining the accuracy of skin infections classification.

A substantial amount of literature review have demonstrated that the tissue lesion features quantification may be of vital prominence, because several of them can be recognized based on quantifiable features mined from an image. Maglogiannis et al. [9] presents intelligent techniques which are helpful in image segmentation followed by classification. They proposed a local thresholding algorithm in their work along with feature extraction on the segmented portions. This includes texture, edges and color features which help them to hypothesis a classification module using SVM. They employ this supervised machine learning algorithm for prediction of dysplastic nevus versus malignant melanoma.

Sumithra et al. [10] have used a novel approach for the purpose of image segmentation and cataloguing of skin lesions. The skin imageries are first filtered for removing the unsolicited regions and noise data followed by the image division process to extract the lesion areas. A region growing method is used by the authors for image segmentation which involves automatic initialization of seed points. They have used different measurement techniques to find the segmentation performance and the results are encouraging. Color and roughness features were used to represent the extracted lesion areas. Support vector machines along with K-NN classifiers were used to classify the skin disease finally. The authors have created their own dataset comprising of 726 examples from 141 pictures with 5 different classes. The results are encouraging with 34% and 46.71% of F-measure using K-NN and SVM classifier correspondingly and 61% for combination of SVM and k-NN.

Skin cancer is a public health as well as an economic issue. That is accounted since skin infections are one of the most common illness, affecting every age, gender and saturating many cultures, summing up to between 30% and 70% of people in the USA. Mendes and da Silva [11] have constructed a classification model for different types of lesions that includes carcinoma and melanoma. They have used the ResNet-152 architecture trained using the ImageNet database. They discuss the importance of skin lesion automated classification in depth with the help of different datasets throughout the paper. They also conclude that clinical images will pose more problems as compared to the experimental datasets due to its variable nature in terms of cameras and environments.

Ali et al. [12] have proposed a fuzzy C-means method for diagnosis of melanoma. Their proposed architecture consists of three stages namely the pre-processing which deals with contrast enlarging, the main treating stage using fuzzy c-means and finally the post processing stage of morphological erosion. The Fuzzy C-means method (FCM) proposed here divides the pre-processed image data in to lesion and skin clusters separately. This approach is evaluated using skin cancer images and the results are satisfactory as well.

A convolution neural network is most commonly applied to visual imagery analysis. Brinker and others [13] made a systematic review of the different algorithms present in this research area and their objective was to study the neural networks based skin lesion classifier techniques in particular. Google scholar, PubMed, Science direct, web of science and medline were used as references for this review article. Most of the methods fall under the steps of training followed by optimization and testing against a new dataset. The authors of this work find that the Convolutional Neural Network (CNN) displays high performance as compared to other techniques but could not actually provide a comparison table as each work used different examples for training and testing thereby creating reproducibility a difficult task.

One of the shift invariant artificial neural network is CNN and is grounded on the shared-weights design. They are regularized versions of the multilayer perceptron's. Harangi [14] have discussed about using deep CNN for skin lesion classification. The author presents that in the past decade, the deep learning-based methods efficiency augmented intensely and their presentations seem to outpace traditional image processing methods in cataloguing tasks. One main drawback with classification methods is to label thousands of images in the training set before testing it. The author created a group of deep convolutional neural networks which helps in classifying the dermoscopy images in to three different classes namely the nevus, melanoma and seborrheic keratosis. They fuse the cataloguing layers output of four different neutral network architectures. The aggregated output is then fed to one single framework which takes care of final classification. Their experimental results prove that fusing strategy outperform the individual networks classification output.

Skin diseases require high-level expertise due to the visual aspects of the same. Liao [15] hence considers that a computer aided diagnostic system would be more effective. A deep CNN is proposed by the author in his work. They train the architecture using 23,000 different skin images using the Dermnet database and tested the same with the OLE dataset. Their classification can reach as high as 73.1% Top-1 accuracy and 91.0% Top-5 accuracy while experimenting on the Dermnet dataset.

Hundreds of skin conditions affect humans and most of them carry the similar symptoms. It is important for the physicians to

understand the differences with these symptoms and for that purpose, a robust image processing and classifier algorithm is required. To develop such an algorithm, there arises a need also to test it with the close to real world dataset samples. BioGPS is one such dataset that contains thousands of images for browsing and research purposes. The HAM10000 is another set with huge collections of skin lesion images containing common pigmented multi-source dermoscopic types. ISIC 2017 is another publicly available dataset for helping researchers in the field of melanoma. It is built on the academia and industry partnership and we have used this in our work for estimating the performance of our proposed methods. The dataset aims at collecting the digital skin lesion images that can be used to educate specialists and the public people as well in melanoma recognition. It also directly aids in the melanoma diagnosis through clinical decision support, teledermatology, and automated diagnosis. In addition to this, the International Skin Imaging Collaboration team is also developing and intensifying an open source public access to skin images archive for experimenting and validating different proposed standards. This online archive serves as a public reserve of skin lesion images for education and for the expansion and validation of robotic analytic systems.

Different methods and technologies are discussed in the literature but a robust skin lesion classifier is still under research. This is the motivation behind this research work and we propose to use dynamic graph cut algorithm in this work for skin lesion segmentation followed by feature extraction and naïve Bayes classifier for classification. This method minimizes the misclassification probability. Accuracy with precise image segmentation along with the ease of use are the advantages of using graph cut algorithms. The proposed dynamism also helps the algorithm to be applied even for low vision clinical image set by minimizing the energy in the cut. Multinomial naïve Bayes, Bernoulli Naïve Bayes, Gaussian are the diverse types of Bayes classifier algorithms available for the purpose of classification.

3. Dynamic graph cut based affected skin region segmentation

The input digital image can be divided in to multiple segments using a partitioning algorithm. This is necessary because it helps to reduce the processing resources and time as well. In a segmented image, the pixels share similar attributes and they help to solve the problem in a more granular manner. There are a widespread variety of algorithms discussed in the literature that can help to distinct the foreground from the background [16]. The boundary or the regional information are generally used in most of those methods which has certain limitations. Graph cut methods on the other hand brings in the advantage of using both the boundary as well as the regional information. It could either be based on the speed-up principle, interaction based or shape-based graph cut. Figure 2 represents a classic graph cut segmentation algorithm process:

Graph cut optimization problem can help in a wide variety of tasks including image smoothing, object labelling, image segmentation etc. The maximum flow problem in a graph can be resolved for energy minimization problems. For binary segmentation of images, the following notations are followed:

Input: Digital skin image

$X_{\varepsilon\{R,G,B\}}$

Output: Segmented image

$S \epsilon R$

Energy function:

$E_{color} + E_{coherence}$

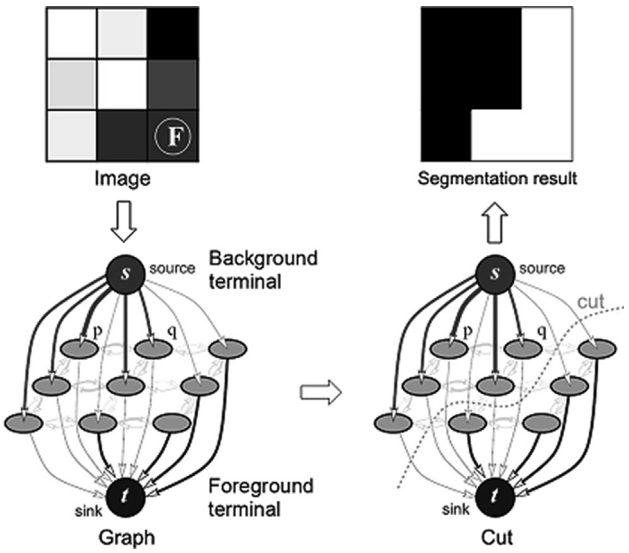

Fig. 2. Graph cut image segmentation algorithm block diagram.

Optimization: Segmented image estimated as a global minimum

If the undirected graph is denoted as

$G = (V, E)$

Here V refers to the nodes set and E to the undirected edges between pixel pairs. Then, $W(V_i, V_j)$ = weight of the edge between two nodes. Every edge (vi, vj) belongs to E and has a conforming weight value w((vi , vj)), which is a non-negative quantity of the change between adjacent pixel elements vi and vj.

S refers to the segmentation of the graph G in such a way that

$G' = \left(V, E'\right)$

Here S takes care of dividing the graph in to G' in such a way that it contains different components of C. The standard graph cut algorithms help to optimize the energy function over the segmentation. In case of repeated graph cuts, the primary step is to optimize the color constraints using K-means and the additional step is to perform the usual graph cut algorithm. Here

E_{color}

Corresponds to the likelihood of each colour and taxon refers to the group of pixels that has certain characteristics repeated in a given image.

The pixels in the input picture are called as the vertex in the graph. A source vertex correspond to the foreground object while the sink vertex correspond to the background object. Similarly, when it comes to edges, we have both the terminal edges as well as the non-terminal edges. There will also be a directed edge and we compute the terminal edge weights using the feature distributions. For non-terminal edge weights, the similarities between the pixel node and the neighbour are considered. Once the underlying graph is found, it is further taken in to processing for computing the minimum cut in the graph and there are several standard algorithms for this purpose [17]. In this work, we suggest to use a novel method of dynamically adjusting the weights of the edges using a thresholding technique.

One of the record cancer today is the Skin cancer which if detected at the early stage can help to cure it in a simpler way and economically as well. An accurate segmentation of the affected skin region is the first step in this process which can help the diagnosis to define well the cancer region. The primary approach of

image segmentation is grounded on classification that is associated to the problem of the thresholds estimation. The objective of this skin disease classification study is to develop an effective algorithm to divide the skin images based on a mixture of different features.

Most of the digitally captured images will have a lot of noise components added to it due to various reasons. This also includes the shading areas which should be differentiated from the skin lesion regions. A shading attenuation is hence a must before image segmentation. This is done as part of the image pre-processing step which includes image smoothing, colour conversion and shading effects attenuation.

3.1. Image restoration through filtering

Suppressing the noise from the images is one of the most important task to be performed before extracting the features or coming to a decision. Digital input images may be corrupted due to variety of reasons including bad focus, illumination effects, camera characteristics or due to imaging system issues [18]. A restoration filter is hence essential to overcome this issue. The overall process of image restoration is shown in Figure 3:

Median filter is one of the efficient filter in terms of preserving the edges while reducing the noise to considerable levels. Adaptive median filter on the other hand helps to overcome the drawbacks of the median filter and will reduce the blurring effect. Mean filter substitutes every pixel with the mean value of its neighbours. Though all these filters are effective, in this work we prefer to go for a Gaussian filter which is best known for suppressing and blurring the noise and produce a high quality output image. The Gaussian filter function is defined as:

$$g(x,y) = \frac{1}{M}\sum f(x,y)exp[-((x-i)^2 + (y-j)^2)/2\sigma^2](i,j)\epsilon S$$

where "S" corresponds to pixel set in the neighbourhood. Also the variable "M" is given by:

$$\mathrm{M} = \sum exp\left[-(\left(x-i)^2\right) + (y-j)^2)/2\sigma^2\right]$$

The above equation helps to define the pixel set along with corresponding weight values.

3.2. Colour space conversion and skin region detection

The translation of the colour representation from one basis to another is called as colour space conversion. After translation, the new image looks as similar to the original one. There are many colour spaces including RGB, CIE, HSV, YUV and CMYK. Differentiating skin pixels from the non-skin ones is part of the skin colour detection process. HSV is the alternative to RGB colour model that helps to align more meticulously with the way human vision sees the colour [19]. The hue (H) here varies from 0 to 1.0 representing colour from red through magenta, saturation (S) diverges from 0 to 1.0 representing unsaturated to fully saturated images and value (V) differs from 0 to 1.0 representing the brightness increase.

The colour tone of the human skin varies a lot from individual to individual. This can be detected either based on the pixel value or based on the region. In the pixel built skin detection method, every pixel is categorised as either skin or non-skin independently from its neighbour pixel value. The skin detection depends on the colour fall. In case of region based skin detection method, the skin pixels present are spatially organised to improve the performance. This region based method requires additional information about the image such as colour intensity, texture, etc. Finally, the percentage of the actual skin colour pixels in each segmented region can be obtained for further processing.

During the process of skin colour detection, it is essential to consider the following factors: (a) The difference of skin pixel and non-skin pixel values in the image. (b) The device used for capturing the digital image as the same object may appear differently for different cameras and settings. (c) The illumination varies drastically across images (d) Skin tones also vary from person to person (e) If there is any movement in the skin region during capture, then it degrades the quality of the image (f) Shadows and lightness has an important role to change the image colour. (g) The colour space that is used for the skin detection or classification.

In this work, we have preferred go with HSV colour space model as the luminance and the hue parameters helps to differentiate the colour and the intensity data even under varying illumination conditions. HSV space is almost similar to the way in which humans look at it. RGB and CMYK are not like that and hence we have used HSV colour space for skin region detection. Since the original pre-processed image is RGB, we have used the following conversion method to get the HSV values.

$$\mathrm{H} = \arccos\frac{\frac{1}{2}(2R-G-B)}{\sqrt{(R-G)^2-(R-B)(G-B)}}$$

$$\mathrm{S} = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$$

$$\mathrm{V} = \max(R,G,B)$$

The algorithm for skin region detection is explained as follows:

Step 1: We input the RGB colour images as input to the colour space conversion algorithm.
Step 2: Above transformation formulas are used to convert from RGB to HSV
Step 3: Histogram of the converted values are formed for 3 regions namely the hue, saturation and value
Step 4: Threshold is now applied to the processed image which is further smoothened and filtered
Step 5: The output image now contains the skin pixel values which are further segmented using the graph cut algorithm.

While the hue represents the actual colour of the image, saturation corresponds to the grey matter in the colour and value indicates the range between grey and white. HSL also called as the hue, saturation and lightness representation is almost similar to

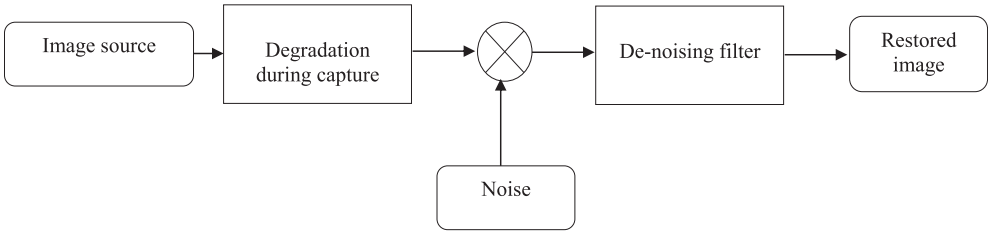

Fig. 3. Image restoration framework.

the HSV model and can also be considered for skin region detection where the fully saturated colours has a lightness value of 1⁄2 around a circle. Here the lightness value of 0 represents a full black while 1 represents full white.

3.3. Proposed dynamic graph cut segmentation

The standard graph cut algorithms optimize the energy function defined over the image segmentation. They are iterated in nature. These algorithms starts with the optimization of color parameters using k-means and then perform the graph cut subsequently. Dynamic graph cut algorithms on the other side, will re-run the suggested algorithm after changing the original problem during each iteration. Image segmentation refers to the problem of labelling the pixel values as zero or one. The object is set to 1 while the background is denoted as 0. A minimum graph cut algorithm is used for this purpose. The cut should happen between the object and the background. The different graph cuts output along with energy minimization is represented in Figure 4:

The unary and the pairwise terms takes two variables as input and are combined in the graph as shown in the Figure 5. The edges are merged together in to a single edge with the help of weight factor. The segmented image will be too fine if there is no indication of boundary between them. If S and T represents the segmented regions, then T is the improvement of S when each component of T is found in S components. Here, T is also a proper refinement of S when the condition $T \neq S$ is satisfied. Similarly, a segmentation is considered to be too coarse when there exist a S improvement that is not too fine. The initial segmentation has very less regions in this case. So, the goal of this work is to provide a skin lesion segmentation that is neither too coarse nor too fine.

Proposed Algorithm:

Input: Skin pixels from pre-processing output
Step 1: Let G = (V, E) represents the image graph
Step 2: Each node in the image is assigned a label to it
Step 3: Let minDiff represents the minimal number of disagreement vertices of all iterations
Step 4: Split the graph in to N overlapped subgraphs
Step 5: while min Diff greater than 0 do
(i) If the peaks v_i and v_j are in disjoint, then merge the components; Otherwise do nothing.
(ii) Count the nodes
(iii)
Update all the subgraphs as per the maxflow value
end while
Step 6: Using the distance transform, find the distance between the segmented regions
Step 7: If the distance value is greater than the threshold, then the regions are definitely not related; otherwise merge them and go back to step 2, update the weight vectors and repeat the segmentation operation again.
Output: Segmentation of V into components $S = (C_1, \ldots, C_r)$

This proposed method of image segmentation using dynamic graph cuts moves from global to local segmentation in an iterative process. On natural skin images from the database, our discussed method effectively segments skin lesion regions having similar col-

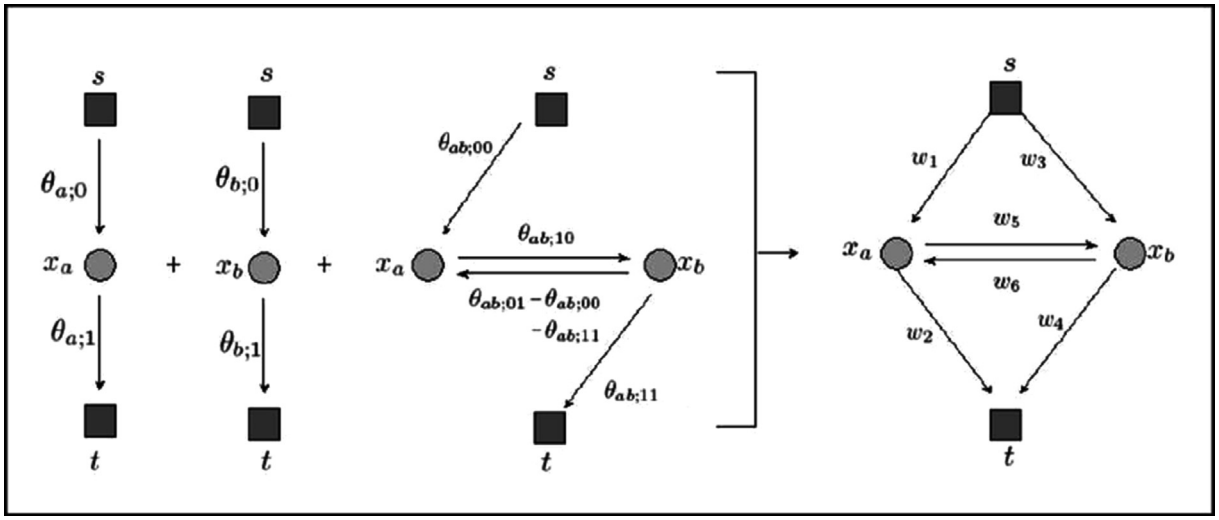

Fig. 4. Graph cut energy minimization.

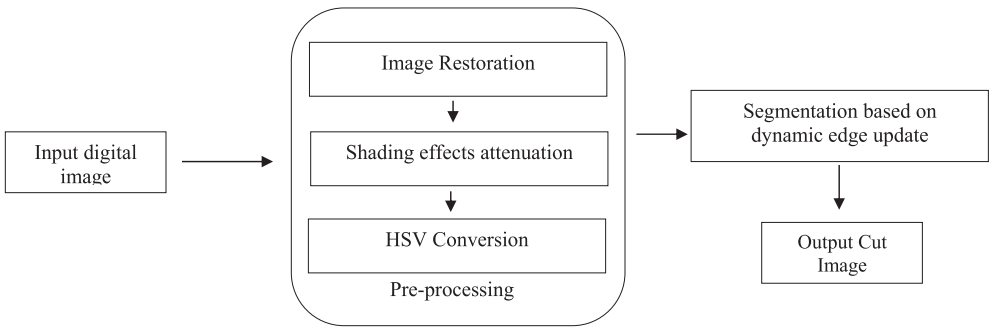

Fig. 5. Proposed dynamic graph cut method framework.

our but dissimilar texture. The overall proposed method is represented in the Figure 5:

The graph theory discussed here uses a integration method that can contract with the alleged neighbouring subgraphs instigating non-convergence and finally give a precision and efficacy study of the method. When the pixels in the skin image are placed under two different clusters, the term E_{data} will calculate the difference between the map 'm' and the assigned region while the term $E_{smoothness}$ will evaluate the smoothness value for segmentation. The output of the discussed method is shown in Figure 6.

The merging of two subgraphs into a solitary graph is accomplished as follows:

(a) Let the two neighbouring subgraphs be denoted as G1(V1, C1) and G2(V2, C2),
(b) The compound graph is represented as G12(V12, C12), where the vertex

$$V1 = \{a, b\} \cup V_1$$

$$V2 = \{a, b\} \cup v2$$

$$V1 \cap V2 \neq \varnothing \ \ and$$

$$V12 = \{a, b\} \cup V12$$

(c) The new merged vertex set $V12 = V1 \cup V2$. This is satisfied only if the two endpoints are all inside V1 ∩ V2. This represents the edge volume of the compound graph G12 which is actually the summary of the volumes of the similar edge in G1 and G2.
(d) If the condition is not satisfied, then the edge volume of the newly compound graph G12 is just that from only one of the identified two subgraphs.

4. Feature extraction from affected skin regions

Skin diseases could be of different types including the dermatitis, psoriasis, herpes, melanoma etc. Once the affected regions were pre-processed, filtered and segmented, we then extract the features from the affected region to classify the type of data. We extract the color and texture details from the image which are fed as input to the naïve Bayes classifier further for classification.

4.1. Texture feature extraction from the segmented skin region

A GLCM (Gray Level Co-occurrence Matrix) is a histogram of co-occurring grey scale pixel values at a given offset [20] over a segmented input skin image. In this work, we extract the following statistical parameters from the GLCM matrix:

a. Contrast: This parameter helps us to measure the local variations that are present in the co-occurrence conditions.

$$\text{Contrast} = \sum_{i,j} |i-j|^2 p(i,j)$$

b. Correlation: This typically presents the combined probability incidence of the identified pixel pairs.

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j))}{\sigma_i \sigma_j}$$

c. Energy: This corresponds to the sum of squared features in the GLCM matrix. This parameter is also called as the angular second moment.

$$\text{Energy} = \sum_{i,j} p(i,j)^2$$

d. Homogeneity: This parameter helps to find the closeness of the elements distribution in the matrix.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

The GLCM is one of the influential tool for extracting the image feature from the segmented image region and Figure 7 shows the sample GLCM matrix for a small segmented skin lesion region.

Entrophy is another interesting feature that can be extracted from this image and that represents the information amount present and it is to be noted that a perfectly flat input image will have a zero entrophy value. If pi represents the probability that the differences between two nearby pixels are equal to I, then the entrophy is given by:

$$\text{Entropy} = -\sum_i p_i (log_2 p_i)$$

Fig. 6. Dermoscopic skin lesion image segmentation (a) Original image (b) Graph cut (c) Segmented pixels from the original image.

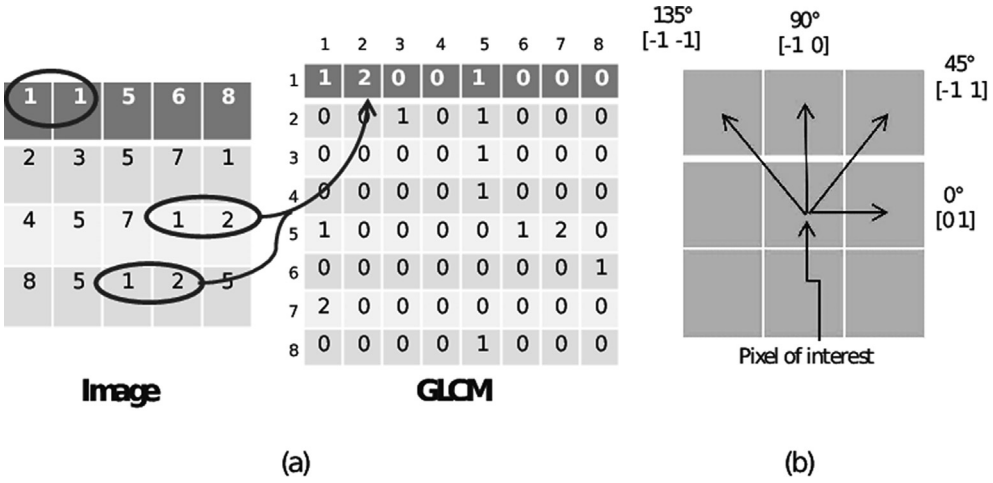

Fig. 7. Original input image (Left) and the GLCM (Middle) matrix representation while pixel of interest (Right).

Once all the GLCM features are extracted, they are stored in the database for further processing and classification. The size of the feature vector will be five multiplied by the number of images as we extract contrast, correlation, energy, homogeneity and entrophy information from each segmented image regions.

4.2. Colour feature from the segmented skin region

A computer based colour analysis technique helps a lot in the classification of skin disease. Certain colors are normally associated with the melanocytic lesions and that includes brown, black, shades of tan and sometimes patches of red or blue as well. The different colour shades and dermoscopic structure repetitive elements help to envisage the in-depth structures of a lesion in high-resolution skin lesion images, which are quite hard to notice through the bare eye inspection by the medical practitioners [21]. Colour descriptors tend to describe the different colour features or properties present inside the lesion.

The HSV colour model that is selected in our work are perceptually uniform and assist to process the dermoscopy skin images more intensely. The colour score that is being computed from the skin lesion region is increased by one, if the distance amid the inspected pixel's value and colour position is equal or below to the pre-computed threshold value. Melanoma has the colour of brown to black depending upon the melanin pigment production at different skin depth. The colour descriptors are usually intended from the different colour channels like mean and standard deviation which are also referred to as the statistical descriptors. One such colour variation is represented in Figure 8.

The statistical colour descriptors extracted from the segmented image region are descripted as follows.

Moment 1: Mean which gives the average value of colour

$$\mu = \frac{1}{N}\sum_{j=1}^{N} p$$

Moment 2: (SD) Standard deviation which is the square root of the variation

$$\sigma = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(p_j - \mu)^2}$$

Moment 3: Skewness helps to find the asymmetry degree

$$S = \sqrt[3]{\frac{1}{N}\left(\sum_{j=1}^{N}(p_j - \mu)^3\right)}$$

Moment 4: Variance

$$V = \frac{1}{N}\sum_{j=1}^{N}(p_j - \mu)^2$$

The above described color moments helps to measure the color distribution characteristics which is similar to the central moments in case of a probability distribution. They serve the pur-

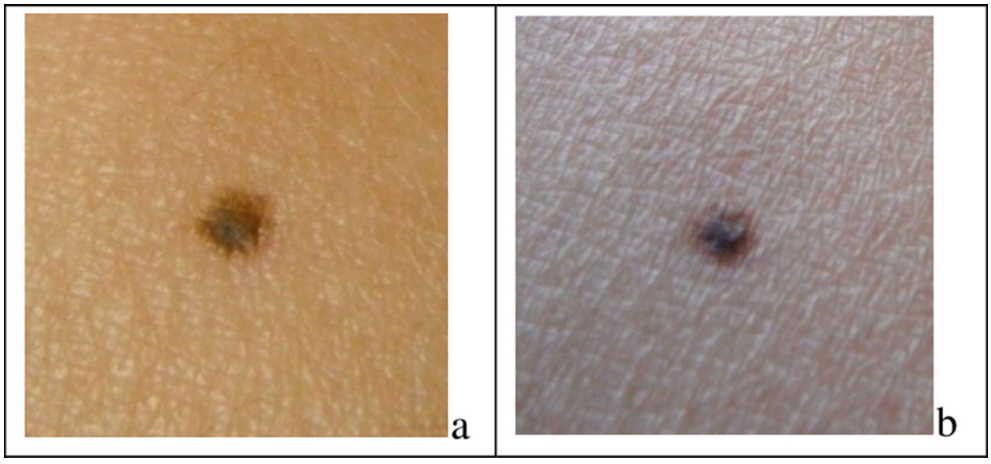

Fig. 8. Colour variation between two regions.

Fig. 9. (Left) Symmetrical (b) Asymmetrical regions.

pose of color indexing in this work. These color descriptors are calculated across channels which provides a total of 72 features and all of them are stored in the database for further training and classification.

4.3. Asymmetry feature from the segmented skin region

One of the important feature that helps in the early detection of skin disease like melanoma is asymmetry. This is because most of the melanomas are asymmetrical. If we divide the region with the help of a line, the two halves of the segmented region does not match. It actually looks like a round shape to an oval one with a symmetrical common mole. The asymmetry is measured and connected to the local origin of the skin lesion (L). The central symmetry of the segmented image can be measured by a rotation of 180° about the centre of gravity. Figure 9represents this.

According to skin physicians, melanomas grow or mature in a radical fashion which means they are mostly asymmetrical. This asymmetry value is designed by subtracting the segmented shape area on one side of the axis from the reflected shape on the other side [22] which results in two area differences as follows:

$$\text{Asymmetry} = \left(\frac{A_{min}}{A_{total}}\right) * 100$$

where ΔAmin is the lowest absolute value difference between different sub regions of the given input image and A_{total} is the identified skin lesion shape region. Asymmetry of a skin lesion region is calculated by detecting the centre of gravity of the pre-processed skin input image. On the other hand, the benign tumours are symmetrical in nature.

Asymmetry is one of the important parameter that helps to differentiate malignant tumours from the healthier ones. It uses the lesion principal axes. An index value is intended from the minimum change amid the lesion image area and the lesion image reflected from the principal axis. Bulkiness value and fractal dimensions also help to measure the symmetricity among which bulkiness is a dimensionless value. It is given by:

$$Bulkiness = \frac{Equivalent\ \ ellipse\ \ area}{Original\ \ area}$$

where equivalent ellipse in the above equation is an ellipse that has the same moment of inertia as the occupied skin lesion image. With the help of the above described bulkiness value, almost 76.6% of skin lesions are classified as malignant. Geometric asymmetry can be calculated by separating the skin lesion into two portions by a conventional line that badges through the centre of mass, post which a assessment is made between these two divided parts by computing the distance present among the size functions. This also governs the qualitative value of asymmetry.

To define the main and ancillary axis of inactivity, we present the space (o, x, y) which signifies all points of the skin lesion. The rate of symmetry for the given input image is calculated through the following steps:

(a) Creation of the object rotation following the two principal axes;
(b) Creating the intersection amongst the two objects namely the original object and the thought object.

At the end of this process, the exact regularity rate of the image is inferred by taking a maximum among these values. Once all the features are extracted, the final operation in this work is to build a robust classifier that can be used for classifying several criteria based on texture, asymmetry, colours and asymmetrical edges allowing the finding to be evaluated.

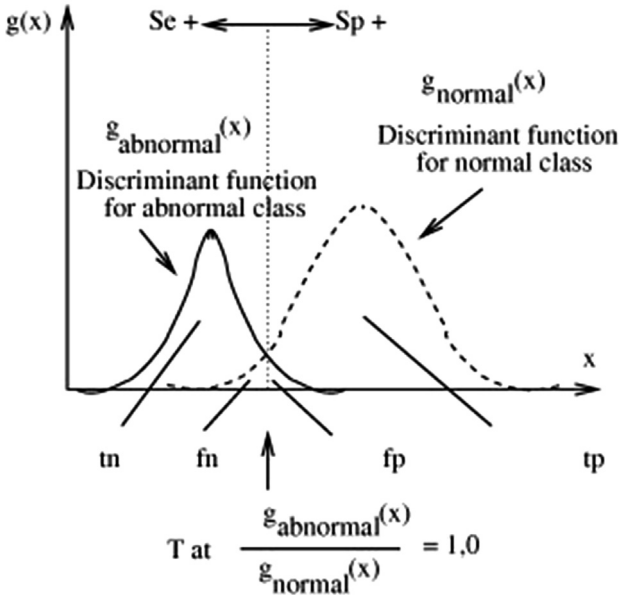

Fig. 10. Bayesian Classifier discriminant function classification.

5. Naïve Bayes probabilistic classifier

A decision support system helps in the grouping of skin disease once the affected region is segmented and the features are extracted from it. Naïve Bayes classifier belongs to the probabilistic classifier type and is independent between different features [23]. The algorithm makes use of the Bayes theorem to classify different objects. The posterior probability equation is represented as follows:

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p^{(x)}}$$

where

P(c|x) signifies the posterior probability assumed predictor (x, attributes) of class (c, target).
P(x|c) is the likelihood which is the predictor given class probability.
P(c) corresponds to the prior probability of class.
and P(x) represents the prior probability of predictor.

For the purpose of classification, we first convert the data set (features extracted) in to a frequency table for finding the repetition of data values. The likelihood table is then created with the help of probabilities like the overcast one. The Bayesian equation is finally applied to calculate each class probability and the one with the highest value will be the output predicted one. The graph that shows the discriminant function for normal and abnormal classes are represented in Figure 10.

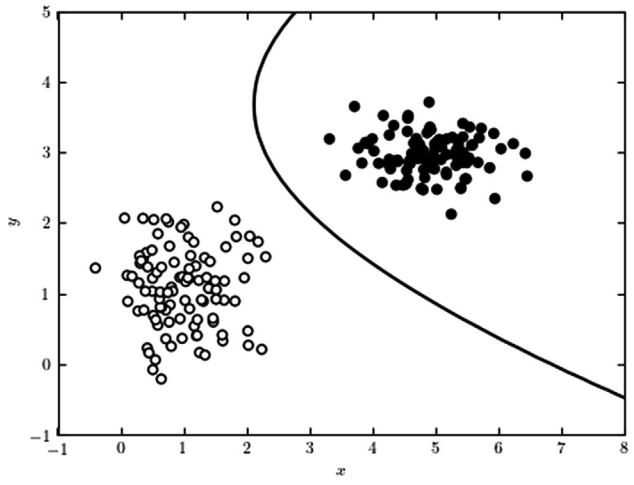

Fig. 11. A Gaussian Naïve Bayes Classifier.

The Bayesian classifier is suitable for both single class as well as multi class classification. The algorithm internally uses Gaussian model, multinomial model or Bernoulli model for training and prediction. Every feature that we have computed from the segmented skin region contributes independently to the overall probabilistic classification.

In our case, we initially use this to classify if the affected region has skin lesion or its normal healthier tissue. So it works as a binary classifier here. Once the preliminary classification is completed, we do the next level of multi class classification with the same probabilistic classifier. Only a less training data is required in case of naïve Bayes classifier as compared to the other methods like logistic regression or support vector machines. This helps us to handle new data training even if the samples are less in number for training. Compared to the numerical variables, this classification method handles well the categorical input variables as like in our research problem. The classifier also helps us with real time prediction, multi class prediction and for recommendation as well.

If the skin segmented feature vector is represented by × = (x1, x2, … xn) where n corresponds to the number of features, then the naïve Bayes algorithm will assign to this instance probabilities as:

$P(C_k|x1, \cdots ., x_n)$for each class that is predicted C_k.

The posterior probability can also be represented as:

$$\text{Posterior} = \frac{prior \times likelihood}{evidence}$$

This method is highly scalable and hence can handle even huge datasets efficiently. The naive Bayes classifiers can be accomplished very competently in a supervised learning environment. To estimate the efficiency of the proposed prototype, a k-fold cross proof technique is used. This method results in a less optimistic estimate of the data model than other similar methods [24]. It also helps in the process of splitting up the training and testing vectors. The k-fold algorithm works as follows:

Step 1: Shuffle the feature set from the skin images randomly.
Step 2: Split the feature set into k different groups
Step 3: For each of the identified unique group:

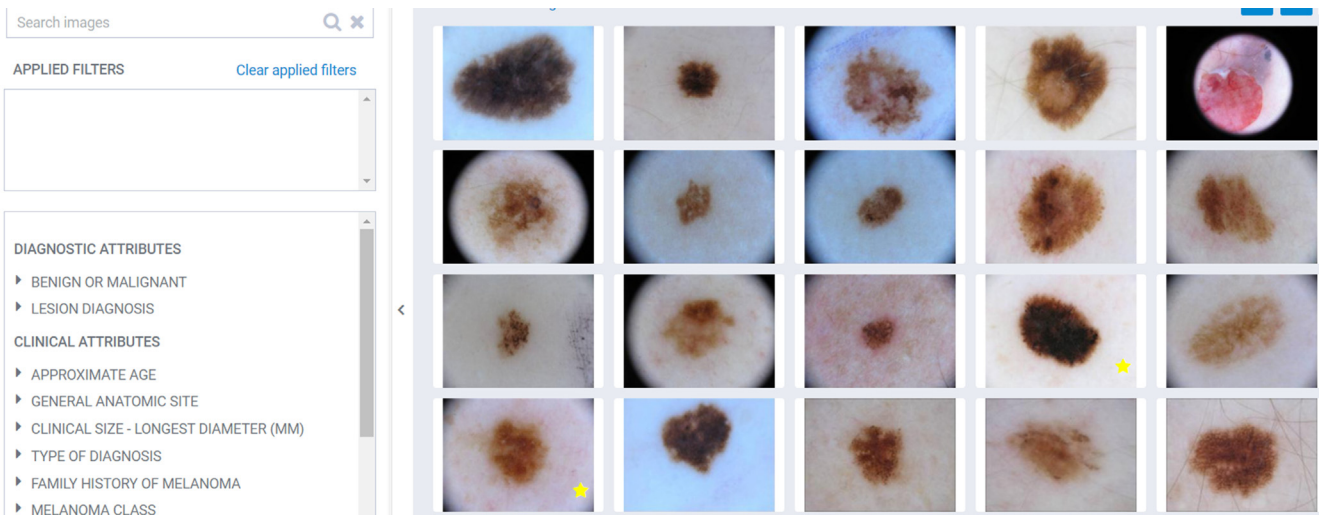

Fig. 12. ISIC dataset with filtering and image searching options.

(a) Take one cluster as a first test data set
(b) Take the left over clusters as a training data set
(c) With the training data, fit a model and also evaluate it on the identified test set
(d) The evaluation score is retained and the model is discarded

Step 4: Summarize the model skill using the model evaluation scores sample.

The value of k is fixed in such a way that each group of samples is large enough to represent the broader dataset. Here, a lower K value means less variance and more bias, while a higher K value means more variance and less bias. A simple Gaussian naïve Bayes classifier output is shown in Figure 11:

A decision boundary is used to separate two classes as in Figure 11. The line here represents the decision boundary where a new point has equal probability of being part of one of the predicted class. The correctness, precision and recall parameters helps to figure the performance of the Bayes classifier algorithm. For each output class, we obtain a predictive posterior distribution during test time by simply leaving the drop out. In similar terms, each iterative pass results in a data from the predictive posterior distributions when the drop out is on. This method of computation requires less resources as it avoids multiple training sessions to cross-validation procedures. The experimental results in the next section shows that this Naive Bayes classifiers gives better accuracy in classification and the proposed method for skin disease detection can successfully improve the detection of skin related diseases.

Table 1
Dataset used for training and testing the proposed algorithm.

Type of skin disease	Training	Testing
Eczema	315	110
Psoriasis	96	55
Vitiligo	214	100
Melanoma	206	105
Carcinoma	110	60
Blue Nevus	185	90

6. Results and discussions

Skin diseases, especially skin cancer is a major health related issue with more than 5 million newly diagnosed cases even in developed countries like United States each year. Among them,

Table 2
Features extracted from the training and testing image set.

Image	Mean R	Mean G	Mean B	Texture 1	Texture 2	Texture 3	Asymmetry
Image 1	0.2511	0.0456	0.0897	0.1687	0.4578	0.2478	1.0456
Image 2	0.2166	0.0578	0.0968	0.1865	0.5528	0.2547	1.2344
Image 3	0.2245	0.0660	0.0998	0.1784	0.5789	0.2589	1.3530
Image 4	0.2124	0.0584	0.0893	0.1896	0.4960	0.2647	1.4021
Image 5	0.2321	0.0559	0.0900	0.1770	0.4875	0.2770	1.1290
Image 6	0.2489	0.0602	0.0955	0.1892	0.5000	0.2500	1.0234
Image 7	0.2578	0.0499	0.0999	0.1995	0.5684	0.2490	1.4995
Image 8	0.2266	0.0502	0.1001	0.1780	0.5127	0.2590	1.2940
Image 9	0.2374	0.0612	0.0884	0.1890	0.4989	0.2571	1.3030
Image 10	0.2390	0.0591	0.0925	0.1925	0.5058	0.2468	1.4145

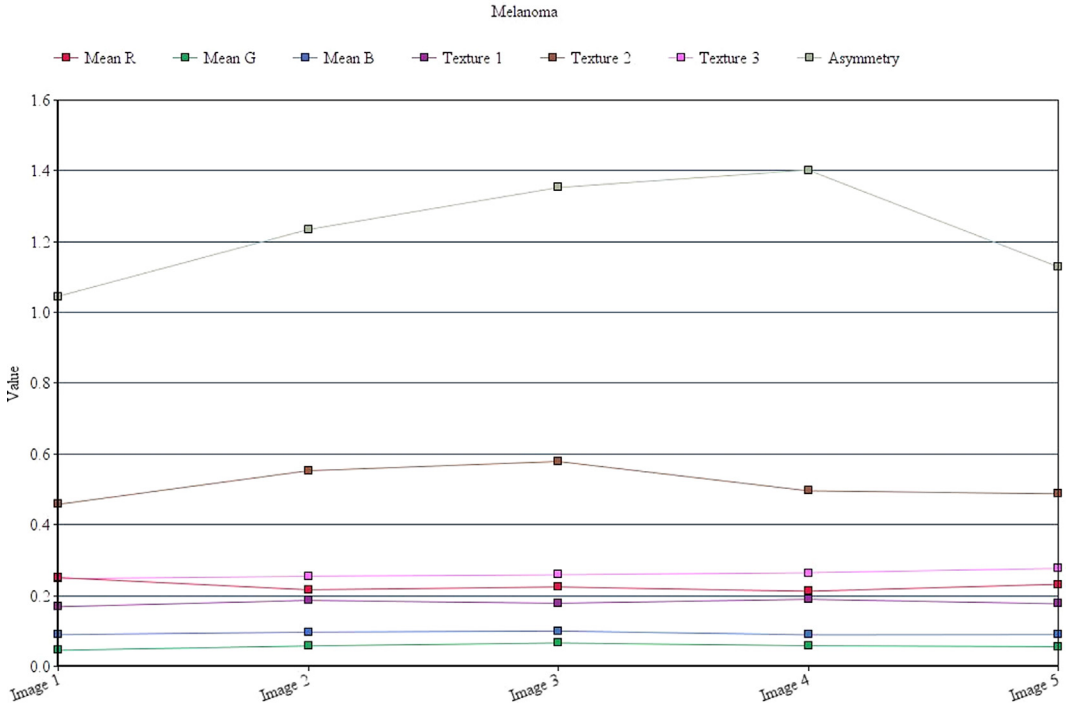

Fig. 13. Melanoma feature set (a).

Table 3
Individual classifier performance comparison.

Feature category	Color	Texture	Asymmetry
Sensitivity	95.45	71.25	80.20
Specificity	92.61	74.79	87.50
Accuracy	91.25	72.80	79.35

melanoma is the deadly form of skin cancer. This is amenable to early discovery by expert visualization. Dermoscopy relates to techniques for eliminating the skin surface reflection and visualizing the deeper skin levels. For this purpose, ISIC provides an online archive of skin data that are helpful for segmentation of lesion, detection and localization of visual dermoscopic features and finally the classification of skin disease [25]. A snapshot of the dataset used in this research is shown in Figure 12:

The diagnostic attributes of the data set can be classified in to benign or malignant type, clinical attributes include age, site, diameter, diagnosis type, family history, melanoma class, mitotic index, thickness, type, ulceration, melanocytic, nevus type and personal history related data. There are around 23,906 images in total for training and testing the proposed algorithms. Although skin lesions are perceptible to the normal human eye, early-stage melanomas may be hard to differentiate from benign skin lesions with comparable presences. This has headed towards many squandered melanomas regardless of an epidemic of skin biopsies. The number of needless biopsies diverges by the expertise of the examiner, clinical setting and the technology use. Table 1 provides the list of images used for training and testing the type of classes they belong to.

The feature set starts with texture data followed by color and asymmetry features. The texture data includes contrast, correlation and energy values. The color set includes the mean and variance at the first level. The asymmetry data is the last one to be present in the database. Physicians who specialize in skin diseases routinely employ total body photography and dermoscopy as computer aided diagnostic tools for the discovery and analysis of melanomas. As it is not possible to present the full feature set in this article, we showcase only the features from the top 10 images in Table 2.

The Intel i7-10510 8 M cache, 4.8 GHz processor is used for training and testing. The feature set were cleaned up for outlier's removal, weightages were multiplied and rakings were done for classification. The graphs following in Figure 13 shows the feature set plotting across different disease categories such as melanoma (a) and carcinoma (b).

The individual classifier performance is tabulated in Table 3. Here each feature category such as color, texture and asymmetry are considered. This table helps us in the study of estimating the generalization error based on the features that are selected for the models. By applying the threshold value of 0 for non-cancer tissues and 1 for cancer ones, we can classify the system accordingly (Figure 14).

The graphs following in Fig. 14 shows the feature set plotting across different disease categories such as melanoma Carcinoma feature set (a). The individual classifier performance is represented graphically in Figure 15. The probabilistic naïve Bayes classifier performance is then calculated in terms of accuracy, sensitivity and specificity. This is equated against the state-of-the-art methods discussed in the literature and is tabulated in Table 4.

The sensitivity or the true positive rate helps to measure the percentage of actual positives that are correctly classified using the machine learning algorithm. On the other side, specificity corresponds to the measurement of the percentage of actual negatives that are correctly recognised. They are represented as:

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{TN} + \text{FN})$$

Once these statistical values of the classifier are calculated, they are compared against the existing methods for robustness and efficiency findings.

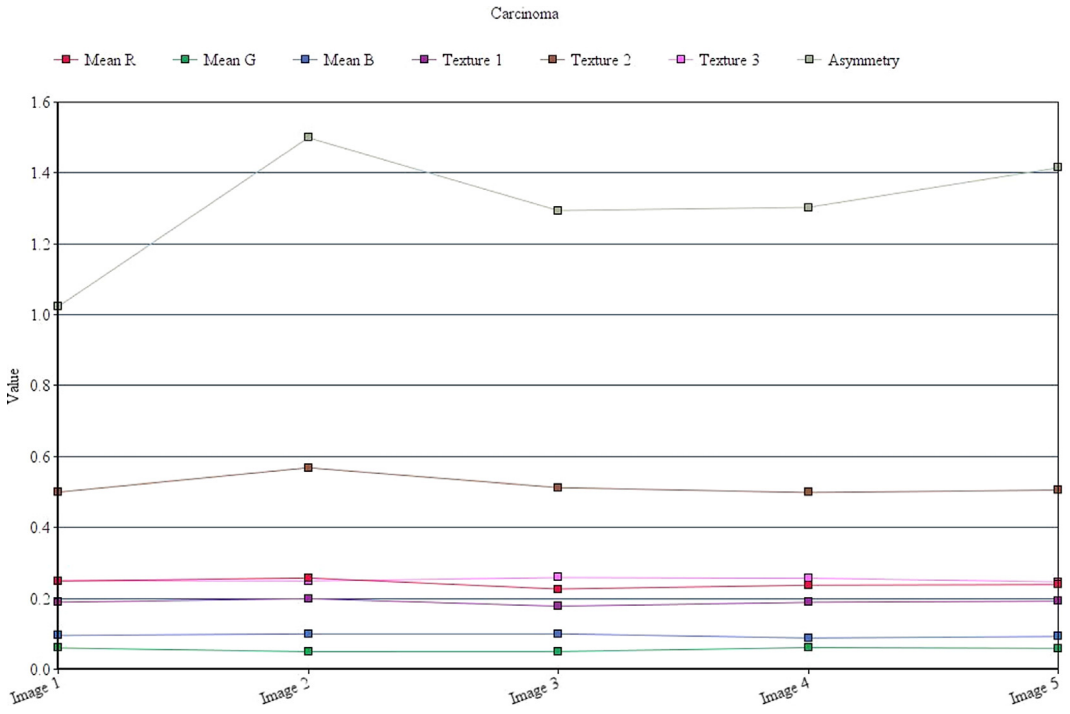

Fig. 14. Carcinoma feature set (a).

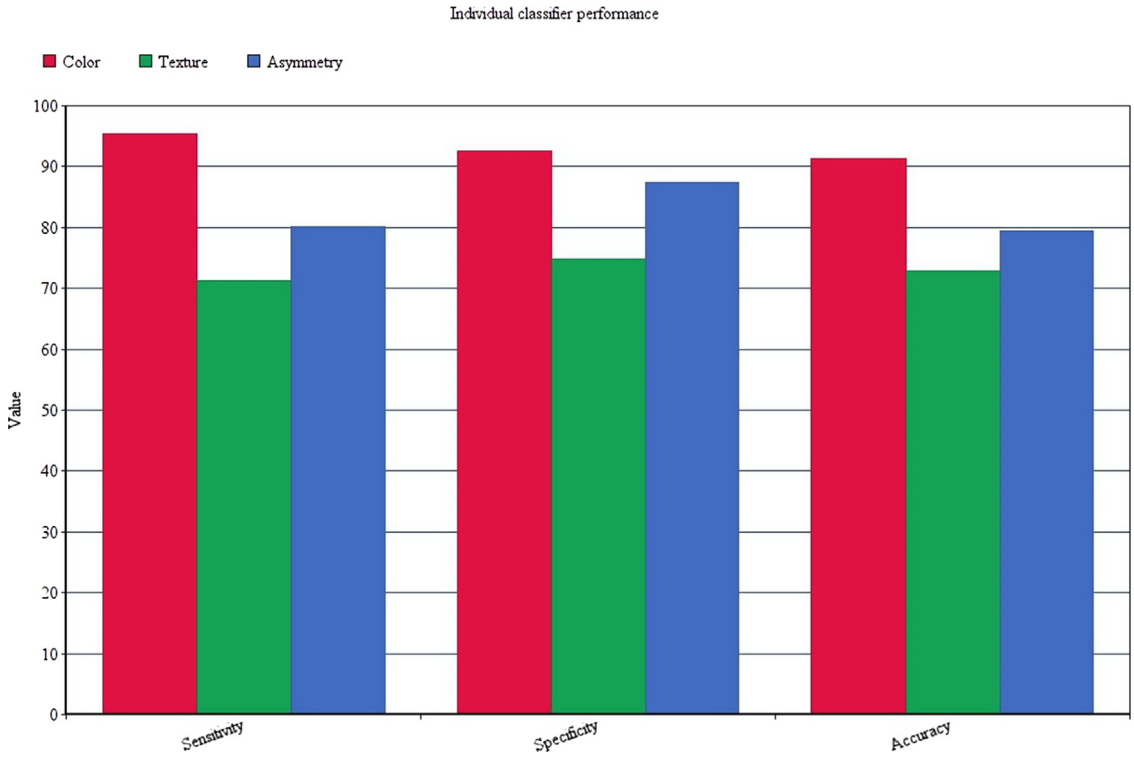

Fig. 15. Individual classifier performance comparison.

Table 4
Proposed dynamic graph cut based segmentation and Naïve Bayes Classifier performance vs other state of the art methods.

Methodology	Sensitivity	Specificity	Diagnostic accuracy
Simple ABCD rule on the skin image	81.6%	61.4%	69.5%
Support Vector Machine based classification	87.5%	66.9%	72.4%
Proposed dynamic graph cut and Naïve Bayes classifier	91.7%	70.1%	72.7%

Different algorithms have been developed to confirm standardization of dermoscopic examinations. Nevertheless, only a few experiments have related the specificity, sensitivity, and diagnostic accuracy of these dermoscopic methods. This includes the simple ABCD rule on the affected skin region and support vector machines based skin lesion classification. We have taken those latest methods in our comparison and found that the overall accuracy is 94.3% for benign cases, 91.2% for melanoma and 92.9% for keratosis on this data set.

7. Conclusion and future directions

Any abnormal growth in the skin region compared to the surrounding regions is called as the skin lesion. While many different conditions can cause skin lesions, some of them needs immediate attention and subsequent treatment. For the purpose of diagnosing a skin lesion, the physician will request for a skin image processing with the help of automated tools and they in turn require algorithms for segmentation and processing. To support this activity, we have proposed a novel skin image segmentation algorithm in this work followed by classification with the help of naïve Bayes classifier. Graph cut segmentation proposed in this work has the advantages of giving practical efficiency, globally optimal labelling, integrate multiple cues and constraints, numerical robustness and unrestricted topological properties of regions. While the traditional methods has the constraints of having the seed pixels that cannot change their labels later, the method proposed here dynamically updates the seeds over iterations to bring in better segmentation results than many other similar methods. A quantitative experimental assessment has been piloted on a publicly available research database, by taking into account many well-known state-of-the-art skin image segmentation approaches for the purpose of comparison. We found that our dynamic graph cut with naïve Bayes classifier method produced better results than most of the literature methods. The classifier was easy to implement and fast to predict the output and performs well even in case of multi class prediction. Moving forward, we would like to experiment this with different color models and also extend the algorithm for other practical applications and measure the performance.

CRediT authorship contribution statement

V.R. Balaji: Writing - review & editing. **S.T. Suganthi:** Conceptualization, Writing - original draft, Formal analysis, Writing - review & editing. **R. Rajadevi:** Software, Writing - review & editing. **V. Krishna Kumar:** Methodology, Writing - review & editing. **B. Saravana Balaji:** Conceptualization, Writing - original draft, Writing - review & editing. **Sanjeevi Pandiyan:** Conceptualization, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Faliu Yi, Inkyu Moon, Image segmentation: a survey of graph-cut methods, Published in the International Conference on Systems and Informatics, 2012.
[2] Liu Xiaolong, Zhidong Deng, Yuhan Yang, Recent progress in semantic image segmentation, Published in the Journal of Artificial Intelligence Review, 2018.
[3] Anders Eriksson, Olof Barr, Kalle Åström, Image segmentation using minimal graph cuts, published in the Journal of SSBA Symposium on Image Analysis, 2006.
[4] Pedro F. Felzenszwalb, Daniel P. Huttenlocher, Efficient graph-based image segmentation, Int. J. Comput. Vision 59 (2) (2004).
[5] Ayman M. Mansour, Texture classification using Naïve Bayes classifier, published in the IJCSNS International Journal of Computer Science and Network Security, vol.18 No.1, January 2018.
[6] Pravin S. Ambad, A.S. Shirsat, A image analysis system to detect skin diseases, published in the IOSR Journal of VLSI and Signal Processing (IOSR-JVSP), Volume 6, 2016.
[7] Ihab Zaqout, Diagnosis of skin lesions based on dermoscopic images using image processing techniques, An article published in the Journal of Pattern Recognition - Selected Methods and Applications, Intechopen, June 2019.
[8] Li-sheng Wei, Quan Gan, Tao Ji, Skin disease recognition method based on image color and texture features, a research article published in the Journal of Computational and Mathematical models in medicine, 2018.
[9] Ilias Maglogiannis, Elias Zafiropoulos, Christos Kyranoudis, Intelligent segmentation and classification of pigmented skin lesions in dermatological images, published in the Springer journal with Hellenic Conference on Artificial Intelligence, 2006.
[10] R. Sumithra, Mahamad Suhil, D.S. Guru, Segmentation and classification of skin lesions for disease diagnosis, published in the Elsevier Journal of Computer Science, Volume 45, 2015.
[11] Danilo Barros Mendes, Nilton Correia da Silva, Skin lesions classification using convolutional neural networks in clinical images, published in ArXiv 2018.
[12] Abder-Rahman Ali, Micael S. Couceiro and Aboul Ella Hassenian, Melanoma detection using fuzzy C-means clustering coupled with mathematical morphology, published in the 14th International Conference on Hybrid Intelligent Systems, 2014.
[13] Titus Josef Brinker, Achim Hekler, Jochen Sven Utika, Niels Grabe, Dirk Schadendorf, Joachim Klode, Carola Berking, Theresa Steeb, Alexander H Enk, Christof von Kalle, Skin cancer classification using convolutional neural networks: systematic review, published in the journal of medical internet research, 2018.
[14] B. Harangi, Skin lesion classification with ensembles of deep convolutional neural networks, published in the Journal of biomedical informatics, 2018.
[15] Haofu Liao, A deep learning approach to universal skin disease classification, published in the project report of CSC400 – Graduate problem, 2018.
[16] Swarnendu Ghosh, Nibaran Das, Ishita Das, Ujjwal Maulik, Understanding deep learning techniques for image segmentation, published in the journal of computer vision and pattern recognition, 2019.
[17] Krzysztof Ciesielski, Jayaram K. Udupa, Paulo André Vechiatto Miranda, A unifying graph-cut image segmentation framework: algorithms it encompasses and equivalences among them, Med. Imaging J. (2012).
[18] Azadeh Noori Hoshyar, Adel Al-Jumaily, Afsaneh Noori Hoshyar, Comparing the performance of various filters on skin cancer images, J. Comput. Sci. 42 (2014) 32–37, https://doi.org/10.1016/j.procs.2014.11.030.
[19] S. Kolkur, D. Kalbande, J. Jatakia, Human skin detection using RGB, HSV and YCbCr color models, published in the journal of ArXiv 2017.
[20] Zeeshan Abbas, Rehman, Mobeen-ur, Najam, Shahzaib Rizvi, S.M., An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach towards Classification of Skin Lesion, published in the Amity International Conference on Artificial Intelligence, 2019.
[21] Y. Cheng, R. Swamisai, S.E. Umbaugh, R.H. Moss, W.V. Stoecker, S. Teegala, S.K. Srinivasan, Skin lesion classification using relative color features, Skin Res. Technol. (2008).
[22] Ihab Zaqout, Diagnosis of Skin Lesions Based on Dermoscopic Images Using Image Processing Techniques, published in the journal of open science and open minds, Intechopen, 2019.
[23] Anurag Kumar Verma, Saurabh Pal, Surjeet Kumar, Comparison of skin disease prediction by feature selection using ensemble data mining techniques, published in the Elsevier Informatics in Medicine unlocked, 2019.
[24] Srujan Kotagiri Raju, M. Ramakrishna Murty, Support vector machine with K-fold cross validation model for software fault prediction, an article published in the International Journal of Pure and Applied Mathematics, July 2018.
[25] ISIC publicly available dataset for training and testing the proposed methodology: https://www.isic-archive.com/#!/topWithHeader/tightContentTop/about/isicArchive.