Neural Computing and Applications (2020) 32:15929–15948
https://doi.org/10.1007/s00521-019-04514-0

RECENT ADVANCES IN DEEP LEARNING FOR MEDICAL IMAGE PROCESSING

An integrated framework of skin lesion detection and recognition through saliency method and optimal deep neural network features selection

M. Attique Khan[1,3] · Tallha Akram[2] · Muhammad Sharif[3] · Kashif Javed[4] · Muhammad Rashid[3] · Syed Ahmad Chan Bukhari[5]

Received: 17 December 2018 / Accepted: 9 October 2019 / Published online: 8 November 2019

Abstract

Malignant melanoma, not belongs to a common type of skin cancers but most serious because of its growth—affecting large number of people worldwide. Recent studies proclaimed that risk factors can be substantially reduced by making it almost treatable, if detected at its early stages. This timely detection and classification demand an automated system, though procedure is quite complex. In this article, a novel strategy is adopted, which not only diagnoses the skin cancer but also assigns a proper class label. The proposed technique is principally built on saliency valuation and the selection of most discriminant deep features selection. The lesion contrast is being enhanced using proposed Gaussian method, followed by color space transformation from RGB to HSV. The new color space facilitates the saliency map construction process, utilizing inner and outer disjoint windows, by making the foreground and background maximally differentiable. From the segmented images, deep features are extracted by utilizing inception CNN model on two basic output layers. These extracted set of features are later fused using proposed decision-controlled parallel fusion method, prior to feature selection using proposed window distance-controlled entropy features selection method. The most discriminant features are later subjected to classification step. To demonstrate the efficiency of the proposed methods, three freely available datasets are utilized such as PH2, ISBI 2016, and ISBI 2017 with achieve accuracy is 97.74%, 96.1%, and 97%, respectively. Simulation results clearly reveal the improved performance of proposed method on all three datasets compared to existing methods.

Keywords Melanoma · Saliency segmentation · CNN features · Fusion · Optimal features · Neural network

1 Introduction

Malignant melanoma, being a minority of skin cancers but most deadly, develops from melanocytes cells. The occurrence of melanoma has increased drastically in the last decade, especially among males [1, 2]. In the USA only, an estimated number of new cases reported, only in 2018, are 178,560—comprising 87,290 noninvasive and 91,270 invasive cases. Estimated number of death from melanoma in the USA is 9320, which includes 5990 men and 3330 women [3, 4]. From the recent studies, it is quite apparent that 86% of melanoma cases are caused by ultraviolet radiations (UVR) from the sun [5]. In case of non-melanoma, basal cell carcinoma (BCC) is a common type of skin cancer, whose approximate number of cases, only in US, are 4.3 million since 2018 [6, 7]. In Australia,

✉ Muhammad Sharif
sharif@ciitwah.edu.pk

Syed Ahmad Chan Bukhari
ahmadchan@gmail.com

1 Department of Computer Science and Engineering, HITEC University, Museum Road, Taxila, Pakistan

2 Department of ECE, COMSATS University Islamabad, Wah Campus, Islamabad, Pakistan

3 Department of CS, COMSATS University Islamabad, Wah Campus, Islamabad, Pakistan

4 Department of Robotics, SMME NUST, Islamabad, Pakistan

5 Division of Computer Science, Mathematics and Science, College of Professional Studies, St. John's University, New York, USA

an estimated 14,320 number of new cases are diagnosed since 2018, which include 8653 men and 5667 women, whereas the number of deaths due to melanoma is 1905—including 1331 men and 574 women [7, 8]. In UK, since 2015, the reported melanoma cases are 15,906 [9].

Lately, computer-based methods have been introduced by the machine learning community for automatic diagnosis of skin cancer at the early stages [10, 11]. These skin lesion commonly categorized into malignant and benign are caused by abnormal production of cells in the human body. The malignant lesions are growths of irregular cells that break without control and order; however, the benign lesions are normal cells and do not conflict around the other healthy cells [12]. The sample images of malignant and benign are shown in Fig. 1.

In machine learning, the automatic analysis of medical images is a multidisciplinary domain that associates the relevant class labels to the images [13, 14]. Under the umbrella of artificial intelligence (AI), computer vision is a subfield which automates the systems based on their visual information [15]. Standard tasks of automated image analysis are segmentation, detection, and recognition [16]. In the segmentation task, abnormalities are separated from given images, whereas in the detection step, the abnormality is identified in the image through features [17]. In the recognition task, multi-level features are extracted and classified through supervised learning methods [18]. The example of a classification task is to categories the benign and melanoma according to their class labels using dermoscopic images, recognition of motor imagery (MI)-related EEG recognition of brain–computer interface (BCI) [19, 20], analysis of mild cognitive impairment (MCI) using function connectivity learning toward GSR bases network configuration [21], and computation of most common features from EEG for SSVEP recognition in BCI using multi-set canonical correlation analysis (MccA) which further improved through multilayer correlation maximization method [22]. These methods outperforms for strong set of relevant features which shows the importance of features for good accuracy computation.

Recent studies show improved performance of deep methods for various computer vision applications [16, 23], which work on the principle of producing discriminant feature vectors—later utilized by the classifiers to achieve best accuracies. A large amount of training data are required of these systems for making any decision—also called data hungry methods. Typically, a CNN is a family of deep architectures including many layers, convolutional, pooling, ReLU, and fully connected. Through these layers, data are transformed into a meaningful information features.

1.1 Motivations

Computer-aided diagnosis (CAD) techniques were proposed by Fernandes et al. [24]. The two state-of-the-art CAD techniques proposed by them include skin lesion analysis and color constancy approach. Their approaches are considered contemporary and fundamental in the field of early skin cancer detection. Fernandes et al. [24] findings inspired our research, and we developed a deep learning-based approach using saliency estimation. Authors in [24] performed in-depth analysis of early skin cancer detection. They used an exhaustive dataset EDRA database which contains images from different sources like the University of Graz (Austria), University Federico II of Naples (Italy), and University of Florence (Italy). We attribute the success of our research to the novel innovations and insights provided by Fernandes et al. [24].

1.2 Problem statement and contribution

In general, a CAD system consists of the following major steps includes lesion preprocessing, segmentation, extraction of feature descriptors, and classification. In medical imaging, a good preprocessing approach make lesion

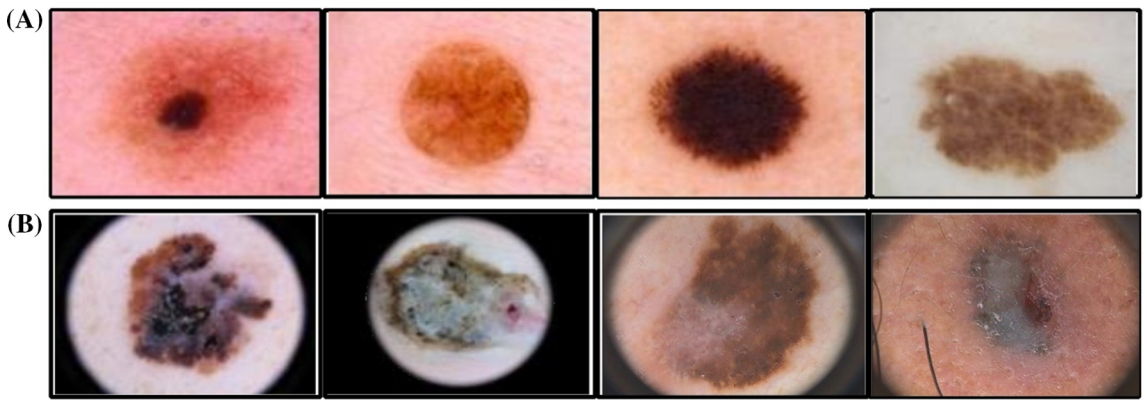

Fig. 1 Sample dermoscopy images from ISBI and Ph2 datasets: **a** benign; **b** malignant

visible as original image which is later helpful for accurate lesion segmentation. The better segmentation shows better representation of lesion in terms of shape and relevant features. The feature extraction is third and more essential step for classification of skin lesions. In this step, the following problems exist which effects on the classification accuracy and system computation time such as number of sufficient features for defined the classification problem, number of selected features, and problems of redundant and irrelevant features.

A new method is proposed which follows four fundamental steps including lesion preprocessing, lesion detection through saliency-based segmentation, features extraction through deep learning, and optimal features selection and final classification using NN. Our major contributions in this work are the following:

- A new contrast stretching procedure is defined using Gaussian relation. Initially, Gaussian distribution of a given image is computed, which is later plug into a Gaussian fitness formulation—based on mean, standard deviation, and a distribution value.
- A novel histogram-based saliency method is proposed, which relies on the best HSV color channel. Two disjoint windows are implemented on the selected channel to generate a saliency histogram.
- A new parallel feature fusion methodology is proposed, decision-controlled parallel (DCdP) features, which combines the set of independent deep features to generate a unified vector.
- An optimal feature selection methodology is exploited, distance-controlled entropy features selection (DcEFS), which has a tendency to mitigate the randomness among selected features.

2 Related work

The performance of skin lesion segmentation and classification truly depends upon the selection of most discriminant features [25, 26]. A good segmentation method traces the precise boundaries, whereas the complex background makes the segmentation process difficult—which directly effects the extracted features. In machine learning community, deep convolutional neural networks (DCNN) performed up to the mark, even on complex and large datasets. The DCNN widely used in medical imaging such as skin lesion recognition and brain tumor classification [27]. Features are extracted from different layers to handle a large number of data and variation between them. Saptarshi et al. [28] extracted texture and morphological features that are later selected through SVM recursive feature elimination (SVM-RFE) approach. Codella et al. [29] fused hand-coded features (LBP) with the deep learning descriptors for melanoma classification. The introduced method reported sensitivity rate of 95% using ISBI 2016 dataset. Manu Goyal et al. [30] performed semantic segmentation using ISBI 2017 dataset. The partial transfer learning and full transfer learning is employed to solve the problem of data deficiency. They achieved similarity rate 81.9% on ISBI 2017 dataset along sensitivity rate 78.9%.

Sara et al. [31] described a melanoma detection technique which is worked on three steps including lesion enhancement, vignette correction, and artifacts removal. In their work, they used wavelet-based decomposition instead of pre- and post-processing approaches. The wavelet decomposition gives an improved performance on ISBI 2017 dataset. Mostafa et al. [32] introduced an encoder–decoder network by dilated residual systems for lesion segmentation using dermoscopy images. In the introduced approach, features are extracted through ResNet pre-trained model and employed NLL and EPE approach for features reduction instead of traditional reduction techniques. The ISBI 2016 and ISBI 2017 datasets are utilized for evaluation and achieved accuracy of 0.984 and 0.936, respectively. Adria et al. [33] utilized pre-trained VGG-19 for CNN features extraction. An augmentation is performed for data deficiency. The transfer learning is performed in the presented method for features mapping and achieved sensitivity rate of 78.66% on ISBI 2016 dataset.

Kawahara et al. [34] proved that the DCNN features could correctly classify the lesions without segmentation and the preprocessing problems such as artifacts or low contrast. The features are extracted through AlexNet, and Dermofit Image Library dataset is utilized for evaluation of a system which includes a total of 1300 dermoscopic images. The presented method gives an accuracy of 85.8%, which is good as compared to existing techniques.

Emre et al. [35] introduced a distinct color-based ML technique for lesion classification using dermoscopy images. The colors are divided into clusters by employing a *K*-means clustering approach, and late regression algorithm is trained for classification of lesions into benign and melanoma. The experiments are performed on Interactive Atlas of Dermoscopy (EDRA) dataset and achieved 72% classification accuracy. Chen et al. [36] implemented two major modules such as SegNet and ClsNet, simultaneously. ResNet 101 was used for features extraction and performed an evaluation on ISBI 2017 dataset. The results show significant performance as compared to the base model. Khan et al. [2] presented a pixel distribution technique for lesion segmentation and later hand-crafted (Harlick, HOG, and Color) features were extracted. The extracted features are fused by a parallel method and employed entropy-controlled approach for significant features selection.

Experiments are performed on PH2, ISCI, ISBI 2016, and ISBI2017 datasets and achieved significant performance.

3 Proposed method

The proposed lesion segmentation and recognition is an integration of two fundamental modules of lesion detection and recognition. In the former, lesion detection is performed through saliency estimation method, while latter classifies the lesion based on extracted deep features, as shown in Fig. 2.

3.1 Contrast-controlled saliency estimation

Contrast stretching is an essential step, especially if the foreground and background has maximum correlation. Utilizing this procedure, a local contrast of an image is stretched, so as to make foreground sharp and visually distinct. In this work, a new contrast-controlled saliency estimation approach is proposed for dermoscopic images. Initially, lesion contrast is enhanced through a fitness function which is controlled by a Gaussian distribution and probability value. Later, HSV transformation is applied to select the high-intensity value channel.

3.1.1 Contrast enhancement

Suppose Δ denotes database of dermoscopic images, let $O(i,j)$ denotes original RGB dermoscopic image, having dimension $(N \times M)$, where $N \in 642$ and $M \in 1024$, and then resize each image into a fixed dimension of (512×512). In the first step, Gaussian formulation is performed through Eq. 1:

$$G(O) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{O_{ij}-\mu}{\sigma}\right)^2} \tag{1}$$

where σ denotes standard deviation and μ denotes mean value of original image $O(i,j)$. The mean value is computed as $\mu = \sum_{i,j=1}^{N} \frac{O_{ij}}{N}$ and $\sigma = \sqrt{\frac{O_{ij}^2}{N^2} - \left(\frac{O_{ij}}{N}\right)^2}$. The Gaussian value $G(O)$ is utilized in the fitness function $F(i,j)$ which improves the local contrast of lesion region. Because the lesion color is brown and black, it is easy to find out the abnormal region. The function $F(i,j)$ makes a clear difference between lesion and healthy regions. The formulation of fitness function is defined by Eqs. 2–3:

$$F(i,j) = \frac{1 + E_{(\mathrm{uq})}}{\delta} \tag{2}$$

$$\delta = \frac{(1 + G(O))}{\beta} \tag{3}$$

where $E_{(\mathrm{uq})}$ denotes unique value which is formulated as $E_{(\mathrm{uq})} = \sum_{i,j=1}^{N} P\left(\frac{1-O_{ij}}{N}\right)$. The $P(\cdot)$ denotes pixels probability values, N denotes total pixels values, and β is a scaling parameter which is fixed as 0.2 for δ. The output enhanced image is denoted by $E_{\mathrm{O}}(i,j)$ and effects are shown in Fig. 3.

3.2 Saliency estimation

In the domain of computer vision, the foreground regions are extracted through well-known segmentation techniques such as K-means and watershed. Although, these methods

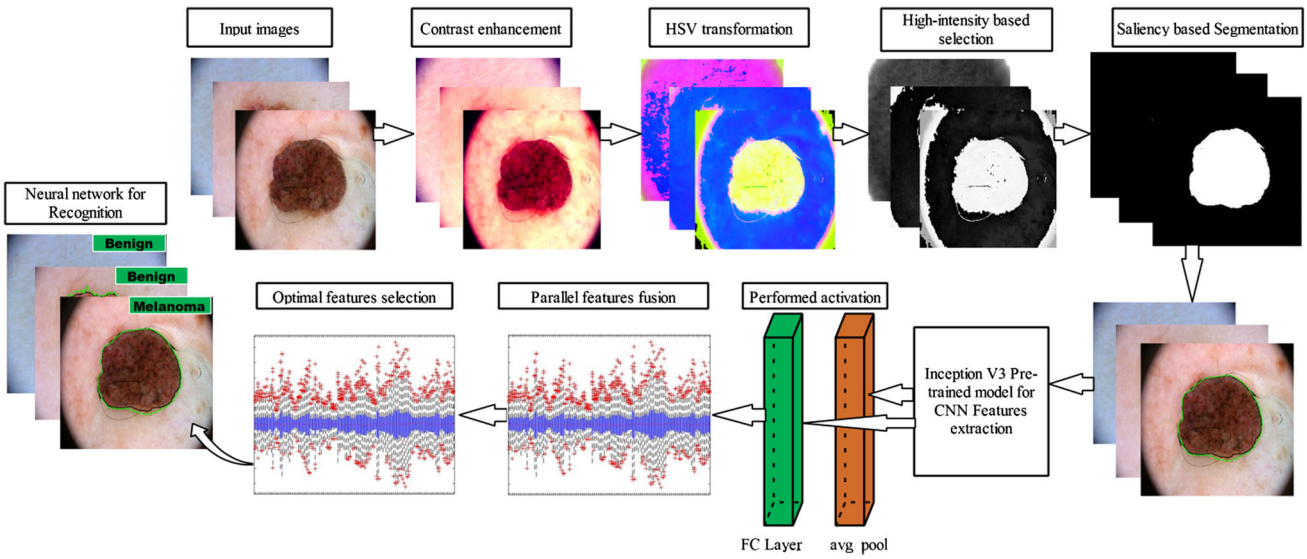

Fig. 2 System architecture of proposed skin lesion segmentation and recognition

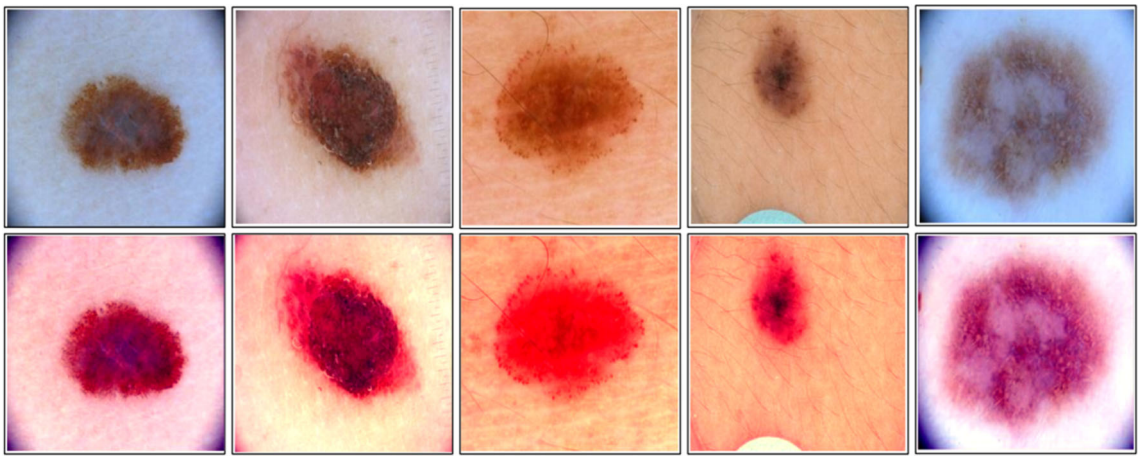

Fig. 3 Lesion preprocessing—contrast stretching. Top row: original RGB dermoscopic images. Bottom row: images after applying contrast stretching

don't produce exemplary results especially with a case when the normal and abnormal pixels have similar contrast. The saliency methods are most important and useful for this type of scenario. Several saliency methods are introduced in the literature which performs efficiently for low-contrast pixels. The saliency detection indicates to identify the most informative part of the image. It has been applied in many domains such as segmentation, object classification, and image retrieval. In this work, a saliency method is proposed. The saliency estimation is performed in three steps. In the first step, HSV color transformation is performed and selects the best channel based on their high-intensity value. In the second step, a saliency map is constructed through distance formulation. Finally, thresholding is performed for binary lesion extraction and mapped on the original image. The detailed description of each step is given below:

As we have enhanced RGB image $E_{\mathrm{O}}(i,j)$ of size 512×512 which is obtained after performing Eqs. 1–3. Let $\xi_{\mathrm{hsv}}(i,j)$ denotes HSV transformed image of three separate channels as hue, saturation, and variation denoted by ξ_{h}, ξ_{s}, and ξ_{v}, respectively. The mathematical formulation of HSV transformation is performed through Eqs. 4–6:

$$\xi_{\mathrm{h}} = \cos^{-1}\frac{\left(\frac{1}{2}(2\xi_{\mathrm{r}}-\xi_{\mathrm{g}}-\xi_{\mathrm{b}})\right)}{(\xi_{\mathrm{r}}-\xi_{\mathrm{g}})^2-(\xi_{\mathrm{r}}-\xi_{\mathrm{b}})(\xi_{\mathrm{g}}-\xi_{\mathrm{b}})} \tag{4}$$

$$\xi_{\mathrm{s}} = \frac{\max\left(\xi_{\mathrm{r}},\xi_{\mathrm{g}},\xi_{\mathrm{b}}\right)-\min\left(\xi_{\mathrm{r}},\xi_{\mathrm{g}},\xi_{\mathrm{b}}\right)}{\max\left(\xi_{\mathrm{r}},\xi_{\mathrm{g}},\xi_{\mathrm{b}}\right)} \tag{5}$$

$$\xi_{v} = \max\left(\xi_{\mathrm{r}},\xi_{\mathrm{g}},\xi_{\mathrm{b}}\right) \tag{6}$$

where ξ_{r}, ξ_{g}, and ξ_{b} denotes red, green, and blue channels that are extracted from enhanced image $E_{\mathrm{O}}(i,j)$. These channels are formulated through following expression: $\xi_{\mathrm{r}} = \frac{(r)}{(r+g+b)}$, $\xi_{\mathrm{g}} = \frac{(g)}{(r+g+b)}$, and $\xi_{\mathrm{b}} = \frac{(b)}{(r+g+b)}$. After extraction of ξ_{h}, ξ_{s}, and ξ_{v} channels, normalized the pixel values through Doane's formulation which is defined through Eqs. 7–10

$$\xi_{\mathrm{h}} = \phi_k(\xi_{\mathrm{h}}), \quad \xi_{\mathrm{s}} = \phi_k(\xi_{\mathrm{s}}), \quad \xi_{\mathrm{v}} = \phi_k(\xi_{\mathrm{v}}) \tag{7}$$

$$\phi_k = 1 + \log_2(\xi_l) + \log_2\left(\frac{|S_1|}{\sigma_{\xi_l}}\right), \quad l \in (h,s,v) \tag{8}$$

$$S_l = \frac{E\left[\xi_l^3\right] - 3\mu\sigma^2 - \mu^3}{\sigma^3} \tag{9}$$

$$\sigma_{\xi_l} = \sqrt{\frac{6(\xi_l - 2)}{(n+1)(n+3)}} \tag{10}$$

where ϕ_k denotes the Doane's function, S_l denotes the third moment skewness, μ described the mean value, and σ_{ξ_l} explains the standard deviation of each channel. These formulations normalize the pixel values of each channel and increase the local contrast of an image. Thereafter, range is defined in three clusters as < 110 or 0.429, < 180 or 0.703, and < 256 or 1 for each channel. From these, select the cluster of < 256 and find its probability value for all. The major purpose of this cluster is to select the channel of brighter region. The probability value for each channel is computed through Eq. 11:

$$P(i,j) = \frac{\xi_{i,j}}{N}, \tag{11}$$

where N denotes total number of pixels for each channel and $\xi_{i,j}$ denotes the favorable pixels. Then, compute the average of three probabilities values for three channels as $\mu_{P(i,j)}$. Through probability value $\mu_{P(i,j)}$, set a threshold function which select the best intensity channel through Eq. 12:

$$\xi_{sl}(i,j) = \begin{cases} \xi_{\mathrm{h}} & \text{if} \quad E(\xi_{\mathrm{h}}(i,j)) < \mu_{P(i,j)} \\ \xi_{\mathrm{s}} & \text{if} \quad E(\xi_{\mathrm{s}}(i,j)) < \mu_{P(i,j)} \\ \xi_{\mathrm{v}} & \text{if} \quad E(\xi_{\mathrm{v}}(i,j)) \leq \mu_{P(i,j)} \end{cases} \tag{12}$$

The above equation explains that if expected value of hue channel is less than mean probability value $\mu_{P(i,j)}$, then select ξ_h channel, otherwise check condition for second channel ξ_s. This process is continuing for all three channels, respectively, and visual results after employing Eq. 12 are shown in Fig. 4.

Figure 4 (a) represents the contrast stretching image which is obtained through Eqs. 1–3, (b–e) represents the HSV transformed image, normalized hue, saturated, and variation channels, respectively, which are obtained through Eqs. 4–10. The final selected channel is shown in Fig. 4f which are gained through Eqs. 11 and 12. The final selected channels as shown in Fig. 4 (f) are further utilized for saliency estimation.

In the next step, saliency map is defined by employing number of active pixels in the selected frame. The active pixels define those pixels whose values are greater than 0.709 and equal and less than 1. For his process, histogram formulation is performed and defined by Eq. 13

$$\xi_{\text{sal}}(p_i, p_j) = \sum_{p_i \neq p_j} w(p_i) D_{\text{p}}(p_i, p_j) \tag{13}$$

where $w(p_i)$ denotes the pixel values in the region p_j, D_{p} represents the histogram distance between healthy and lesion pixels. The distance between healthy and abnormal pixels is formulated by equation

$$\overrightarrow{D_{\text{p}}}(p_i, p_j) = \left| h_{\text{r}(i)}(\xi_{\text{sal}}(p)) - h_{\text{r}(j)}(\xi_{\text{sal}}(p)) \right| \tag{14}$$

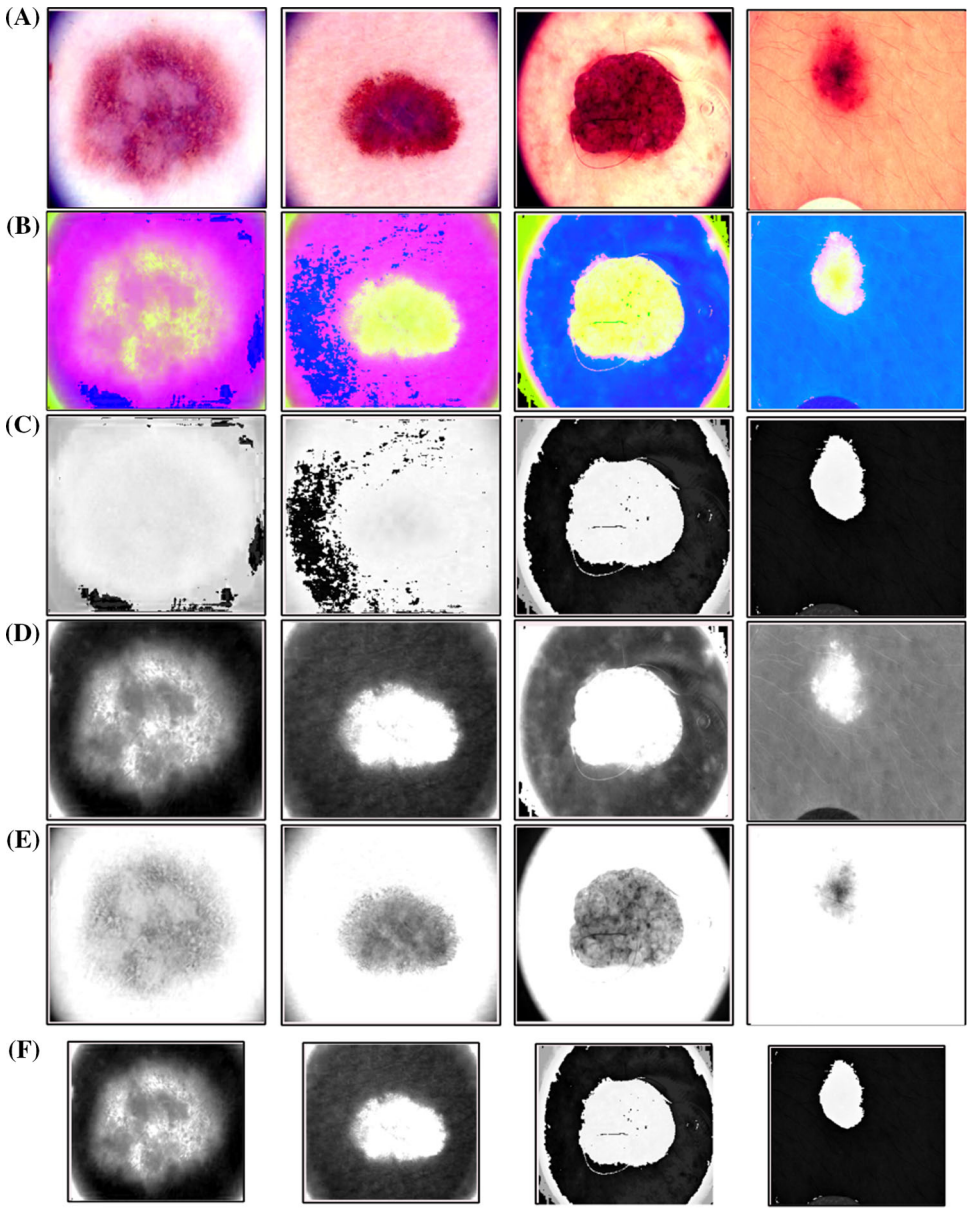

Fig. 4 HSV color transformation and selection of best channel based on finest intensity value. **a** Enhanced image samples, **b** HSV color transformation, **c** extracted hue channel from HSV, **d** extracted saturated channel from HSV image, **e** V channel, **f** finest selected channel

where $h_{r(i)}$ denotes histogram of rows pixels and $h_{r(j)}$ depicts histogram of columns pixels. The extra noise in the image during saliency map construction is a difficult task for accurate lesion segmentation; therefore, we utilized two-dimensional Gaussian filter of scale parameter $\alpha = 0.4$. The gaussian filter is formulated through Eq. 1. After that, divide filtered image into two disjoint parts R_1 and R_2. The R_1 represents the inner part of image and v describe the outer part of filter image. Finally, through inner and outer parts, spatial saliency map is constructed through Eqs. 15–16

$$\xi_{R1}(p) = \mu h_{R1}\left(\xi_G(p_{i,j})\right) \tag{15}$$

$$\xi_{R2}(p) = \mu h_{R2}\left(\xi_G(p_{i,j})\right) \tag{16}$$

where ξ_{R1} and ξ_{R2} denotes the spatial map of inner and outer regions, h_{R1} and h_{R2} described histogram of both disjoint, respectively. The μ denotes mean value for each disjoint and it ranges between 0.4 and 0.6, whereas the best results are achieved on $\mu = 0.542$. Hence, the final saliency map is defined through equation

$$\xi_{FS}(p_i, p_j) = \overrightarrow{D_p}\left[\frac{1}{N1}\sum_{p=R_1}\xi_{R1}(p), \quad \frac{1}{N2}\sum_{p=R_2}\xi_{R2}(p)\right] \tag{17}$$

where $N1$ and $N2$ denote total number of pixels of both disjoint parts $R1$ and $R2$, and $\overrightarrow{D_p}$ is a Euclidean distance between pixels of $R1$ and $R2$. Finally, Otsu thresholding is performed for binary images and results are shown in Fig. 5. In Fig. 5 (c), the proposed saliency results after thresholding are shown, whereas in (d), mapped RGB images are shown which are obtained after multiplication of original and segmented image. Finally, in (e), the ground truth images are presented. Moreover, the comparison of ground truth and proposed segmented images are shown in Fig. 6. In Fig. 6, the green boundary describes the ground truth image, whereas the black boundary represents the proposed segmented image.

3.3 Features extraction and recognition

To classify set of images, various types of features can be utilized such as point [37], shape [23], color, and texture [38, 39]. These set of features produce an impressive performance results for small datasets; however, for large datasets, deep learning-based methods are more effective [40]. Several deep learning methods are introduced such as VGG, AlexNet, ResNet, and Inception. In this work, we used Inception V3 pre-trained deep learning model which includes hundreds of layers. We perform activation on average pool and fully connected layer and extract deep CNN features. The features extraction process includes two major phases including training and testing which are described as follows:

3.3.1 Training and testing phase

In this phase, initially, we divide the segmented input dataset into 60:40 as training and testing. The activation is performed on two layers separately named average pool and fully connected (FC). The activation on average pool layer returns a vector of size N $\times$ 2048 whereas for FC the output vector size of N $\times$ 1000. Later on, both vectors are fused by employing parallel fusion approach. The fusion algorithm is explained in Sect. 3.3.2. The fusion process returns a vector of size 1 $\times$ 2048 which is later utilized for selection of strongest features through the proposed selection process explained in Sect. 3.3. This process is similarly conducted for testing images and obtained an optimal vector which is later matching with a trained model and perform classification. The proposed parallel fusion and optimal feature selection process are shown in Fig. 7.

3.3.2 Features fusion

Features fusion is a process of combining multiple features in one feature matrix to obtain better classification accuracy. The major objective of the fusion process is to raise the performance of systems in the related field [41–43]. In the field of medical imaging, analysis and evaluation of medical obstacles are a core problem from last few years. Through features fusion, this problem can minimize and increase the classification accuracy of different modalities. Two major techniques exist in the literature for features fusion such as serial-based and parallel-based fusion [44]. The serial-based approach is an easy and simple method, whereas the parallel approach is complex because the length of the vectors is not the same to each other. Mostly researchers used zero padding to make the equal length of fused feature vectors (FV) which is not a good idea for this process. In this work, we proposed a decision-controlled parallel (DCdP) approach for features fusion. Let φ_{f^a} denote a feature vector (FV) of dimension $N \times 2048$ that extracted from employing activation on average pooling layer, φ_{f^b} denote FV of dimension $N \times 1000$ which is computed after employing activation on FC layer. Suppose φ_{f^s} denote the fused FV of length $N \times 2048$. A major rule of parallel features fusion is the equal dimension of all fused vectors, if the dimension of FV's are not equal dimension; then, it must be equal through padding operation. In the very first step, we calculate the size of all vectors and set a padding dimension through following formulation:

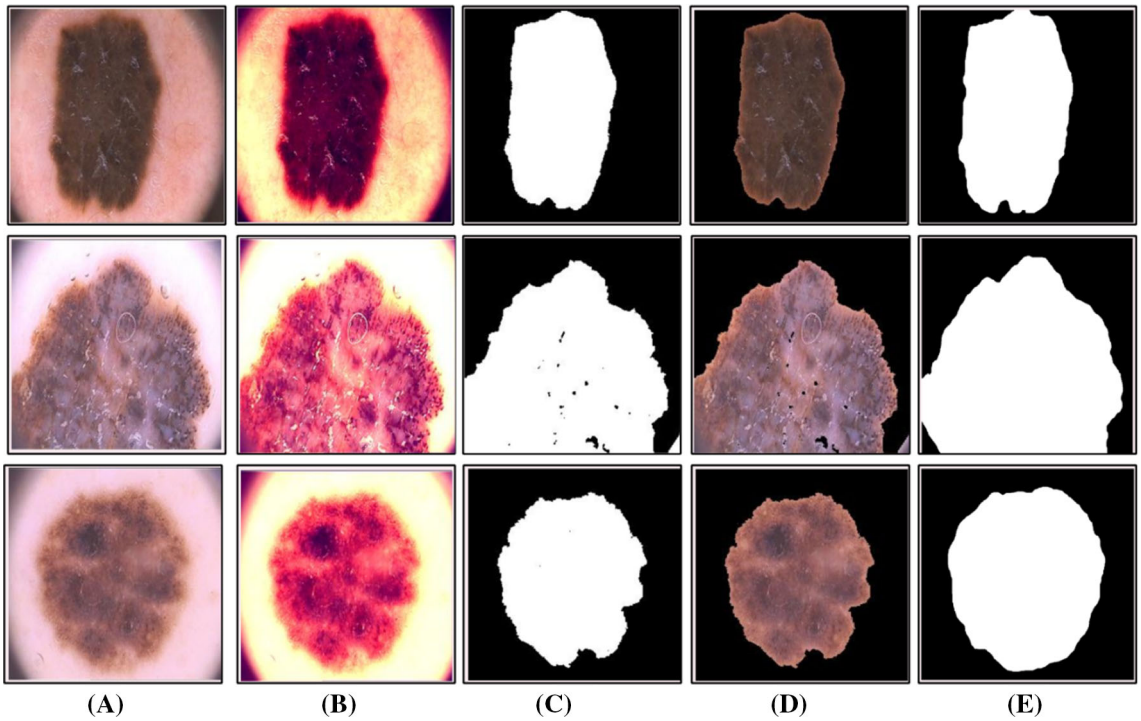

Fig. 5 Proposed lesion segmentation results on PH2 dataset. **a** Original image, **b** enhanced, **c** proposed, **d** mapped on RGB, **e** ground truth

$$P = \frac{(F_l - 1)}{2}, \quad F_l \in \mathrm{MAX}(\varphi_{\mathrm{f}^a}, \varphi_{\mathrm{f}^b}) \quad \text{and} \quad \mathrm{MIN}(\varphi_{\mathrm{f}^a}, \varphi_{\mathrm{f}^b}) \tag{18}$$

where P denote the padding size and F_l denote the length of maximum and minimum FV's. Later, entropy is computed from minimum length vector φ_{f^b}. The output entropy value is used as padding for equal length of minimum dimensional vector

$$H(\varphi_{\mathrm{f}^a}) = -CE_i(\log p_i), \quad \text{where} \quad C = \frac{(\bar{f^a})}{N_{\mathrm{a}}} \tag{19}$$

The C denote the Boltzmann constant value, $\bar{f^a}$ denote the mean value of FV φ_{f^a}, p_i is probability value of FV φ_{f^a}, and N_{a} represent total number of features of φ_{f^a}. After that perform padding through the following formula:

$$\varphi_{\mathrm{pd}} = \sum_{i=1}^{N} \varphi_{\min}(i)(L, H(\varphi_{\mathrm{f}^a})) \tag{20}$$

where L denotes the length of fused vector of dimension $N \times 2048$. After padding, a rule is defined which make a decision that which feature is selected for fused vector at ith index. The decision function determines the importance of features in both FV's. The rule described that strong features are selected from both ith and jth indexes based on maximum value. Here, ith index denotes the features of FV φ_{f^a} and ith denotes the features of FV φ_{f^b}. Mathematically, the rule is formulated as follows:

$$D(\varphi_{\mathrm{f}^a}(i)) = \frac{\varphi_{\mathrm{f}^a}(i)}{\sum_{i=1}^{N} \varphi_{\mathrm{f}^a}(i)} \tag{21}$$

$$D(\varphi_{\mathrm{f}^a}(i)) = \begin{cases} \varphi_i & \text{if} \quad \mathrm{MAX}([\varphi_{\mathrm{f}^a}(1), \varphi_{\mathrm{f}^a}(2), \ldots, \varphi_{\mathrm{f}^a}(N)]) = \varphi_{\mathrm{f}^a}(i) \\ 0 & \text{Otherwise} \end{cases} \tag{22}$$

Similar, for second matrix the formulation is performed as follows:

$$D(\varphi_{\mathrm{f}^b}(j)) = \frac{\varphi_{\mathrm{f}^b}(j)}{\sum_{j=1}^{N} \varphi_{\mathrm{f}^b}(j)} \tag{23}$$

$$D(\varphi_{\mathrm{f}^b}(j)) = \begin{cases} \varphi_j & \text{if} \quad \mathrm{MAX}\left([\varphi_{\mathrm{f}^b}(1), \varphi_{f^b}(2), \ldots, \varphi_{\mathrm{f}^b}(N)]\right) = \varphi_{\mathrm{f}^b}(j) \\ 0 & \text{Otherwise} \end{cases} \tag{24}$$

The output of $D(\varphi_{\mathrm{f}^a}(i))$ and $D(\varphi_{\mathrm{f}^b}(j))$ matrix is finally combined and the resultant vector is obtained of size $N \times 2048$.

3.3.3 Features selection

The optimal features selection algorithms not only increase the recognition accuracy but also minimize the redundant information from original vectors [45, 46]. When redundant information is removed, the execution of proposed system is also improved [47]. In this work, we proposed a new selection approach named window distance-controlled entropy features selection (DcEFS). The proposed

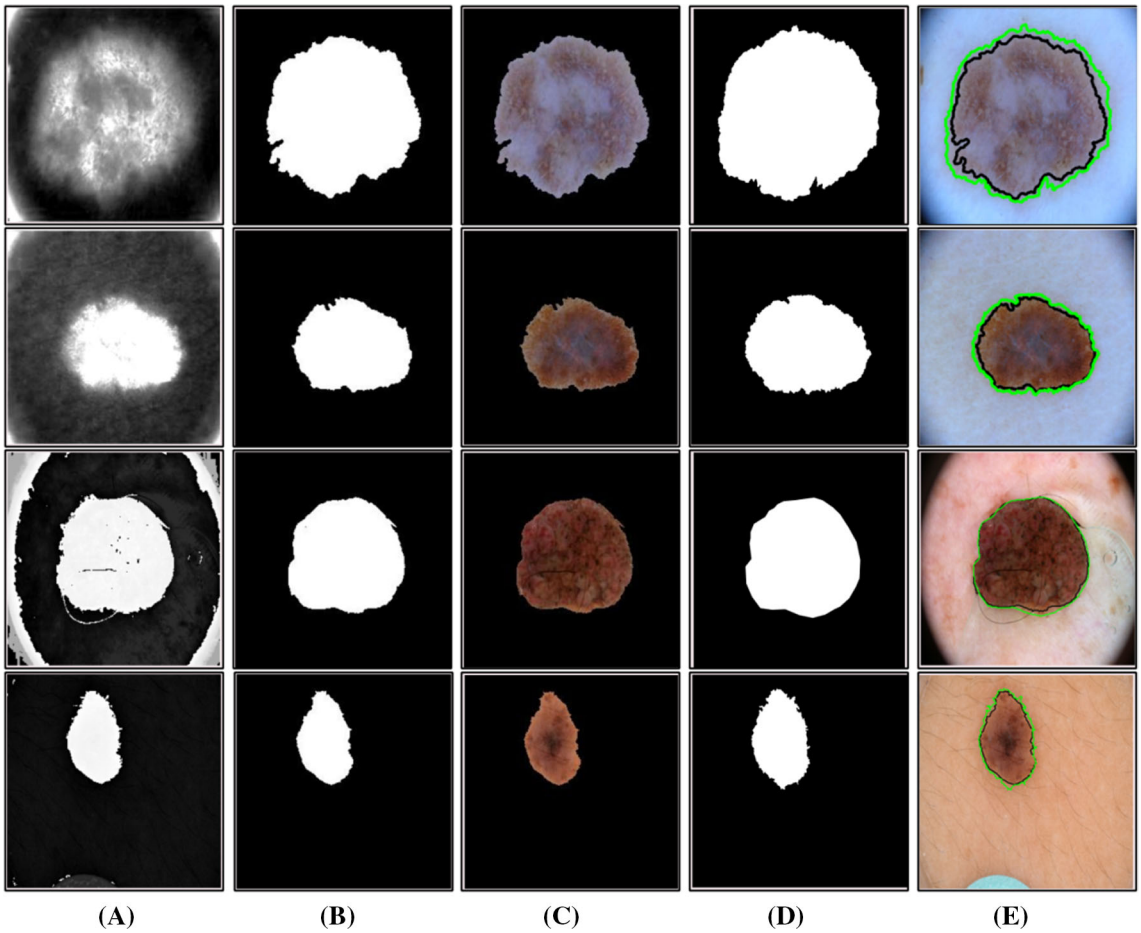

Fig. 6 Proposed detection results. **a** Selected channel, **b** proposed saliency estimation-based segmentation, **c** mapped image, **d** ground truth image, **e** lesion detection and comparison

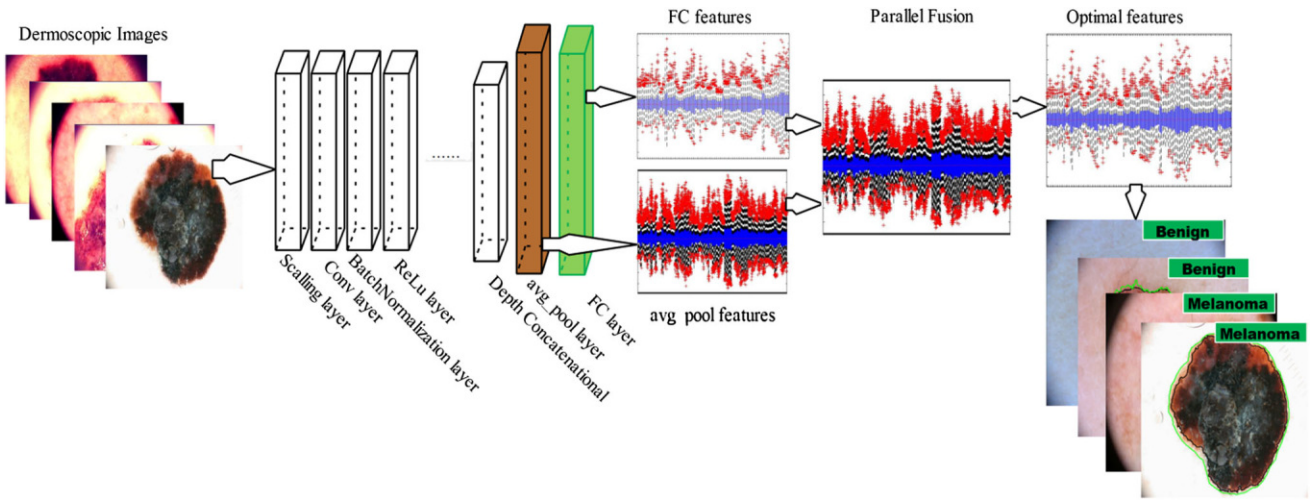

Fig. 7 Proposed CNN features extraction and recognition through parallel features fusion and optimal features selection

approach initially employs a window across fused vector of size 2×2 and computes the Mahalanobis distance (MsD). The MsD is computed through utilization of both horizontal and vertical values of each window. A value is computed after each window operation which is formulated as follows:

$$\bar{M}_{\mathrm{D}}(k) = \sqrt{\frac{(x_i - y_i)^2}{\sigma^2} - \frac{(x_j - y_j)^2}{\sigma^2}} \tag{25}$$

where $\bar{M}_{\mathrm{D}}(k)$ is a MsD vector, x_i, y_i denotes horizontal feature values of selected window, y_i, y_j denotes vertical values in selected window, and σ^2 is a standard deviation. The x_i, y_i, x_j, $y_j \in \varphi_{\mathrm{F}^s}$. Next, a threshold is defined which divide the $\bar{M}_{\mathrm{D}}(k)$ vector into two equal parts as follows:

$$\bar{M}_{\mathrm{D}}(V1, V2) = \begin{cases} V1(k) & \text{for} \quad \bar{M}_{\mathrm{D}}(k)[(i, i+2, i+4), \ldots, N-1] \\ V2(k) & \text{for} \quad \bar{M}_{\mathrm{D}}(k)[(i+1, i+3, i+5, \ldots, N)] \end{cases} \tag{26}$$

The $V1(k)$ and $V2(k)$ denote the two equal length vectors which are extracted through Eq. 26. The entropy is performed on both vectors simultaneously and selects the top 50% features from both of them and later simple concatenation based combined in one vector. The mathematical formulation of entropy calculation is defined through Eqs. 27–28

$$H1(V1(k1)) = -CE_i(\log p_i(V1(k))), \quad \text{where} \quad C = \frac{(V1(k))}{N_{V1}} \tag{27}$$

$$H2(V2(k2)) = -CE_i(\log p_i(V2(k))), \quad \text{where} \quad C = \frac{(V2(k))}{N_{V2}} \tag{28}$$

$$\psi^{f(H1,H2)}(K1, k2) = \begin{pmatrix} H1(V1(k1)) \\ H2(V2(k2)) \end{pmatrix} \tag{29}$$

where $H1(V1(k1))$ and $H2(V2(k2))$ denote entropy vectors of vectors $V1(k)$ and $V2(k)$. The $\psi^{f(H1,H2)}(K1, k2)$ is final selected vector, and $H1$ and $H2$ are 50% reduced entropy vectors as compared to original entropy vectors. Finally, the selected features are utilized for classification through neural network of method multi-perceptron learner (MLP) [40]. The cost function of MLP is defined by Eq. 29

$$J(W, b; k, y) = \frac{1}{2} h_{\mathrm{w,b}}(k) - y, \quad h_{\mathrm{w,b}} = k^n \tag{30}$$

where $h_{\mathrm{w,b}}$ is weight and bias of layers, k^n is last output, k is initial inputs, and y denotes the class labels. The visual classification outputs are shown in Figs. 8 and 9.

4 Validation results and analysis

4.1 Experimental setup

The results of propose method are presented in this section in terms of values and graphical analysis. Three publicly available datasets-PH2, ISBI 2016, and 2017 are employed for validation. In the very first phase, segmentation results are simulated on PH2 and ISBI 2016 dataset using ground truth and segmented images. Through both images, a comparison is conducted based on their pixel's similarity values. Two parameters including dice and error rate are computed for performance evaluation on both datasets. In the second phase, recognition results are computed using all three datasets. The neural network (NN) is utilized as a classifier for classification results and compared its performance with coarse tree, linear SVM (LSVM), cosine KNN, Quadratic SVM, and medium Gaussian SVM. The results are validated through seven famous performance measures including recognition time, negative rate (FNR), sensitivity, specificity rate, AUC, precision rate, and accuracy. MATLAB 2018a is use for simulation on a desktop machine of processor Intel (R) Core (TM)-i7 3770 CPU 3.40 GHz, with 16 GB of RAM and 4 GB of a graphics card (Nvidia GeForce GTX 980). Moreover, the deep learning library of MATLAB named MatConvNet is utilized for deep feature extraction through pre-trained CNN models.

4.2 Datasets

The PH2 dataset comprises of a total 200 dermoscopic images including 80 atypical nevi, 80 carcinoma, and 40 melanomas. The ISBI 2016 dataset holds the total of 1279 images which includes 273 malignant (M), and 1006 benign (B). All images are in the RGB format and of different dimensions. Moreover, the ground truth images of this dataset are available for segmentation results. The ISBI2017 dataset includes a total of 2750 dermoscopy RGB images which includes 517 malignant and 2223 benign images.

4.3 Segmentation results

The proposed saliency method is validated on two famous datasets name PH2 and ISBI 2016. The PH2 results are given in Table 1 for maximum dice rate of 96.75% and negative rate 3.24%. The overall average segmentation dice rate is 94.3239%, and average negative rate is 5.6761%, respectively, which is computed for all 200 images of PH2 dataset. The overall results on PH2 dataset are above 90% which clearly show the improved performance of the proposed saliency method. Secondly, the proposed saliency-based segmentation results are computed on ISBI 2016 dataset and obtained maximum dice rate of 97.9448% and an error rate of 2.0552%, respectively. The overall average dice rate is 94.9227% which is computed for all images of ISBI 2016 dataset. In short, these results show the proposed saliency method outperforms on both datasets.

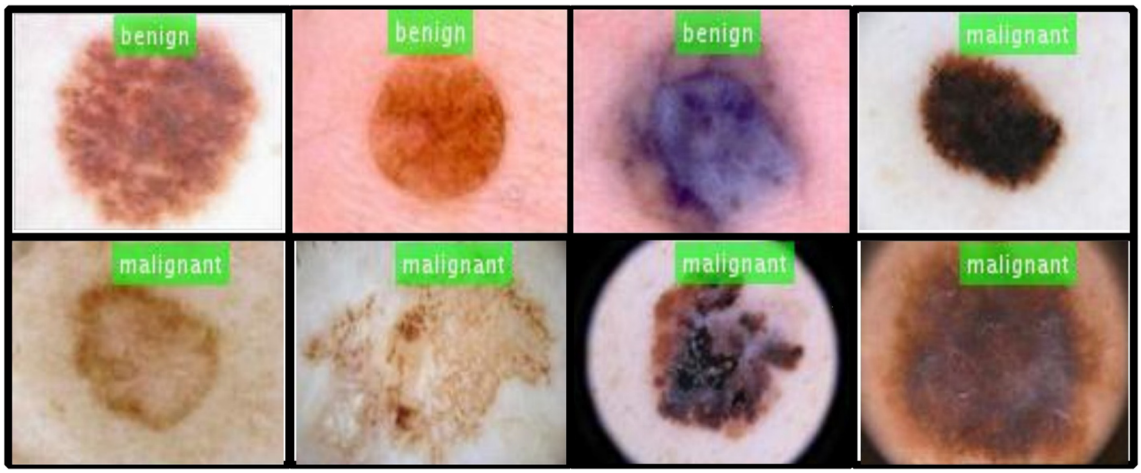

Fig. 8 Proposed labeled results using ISBI 2016 dataset

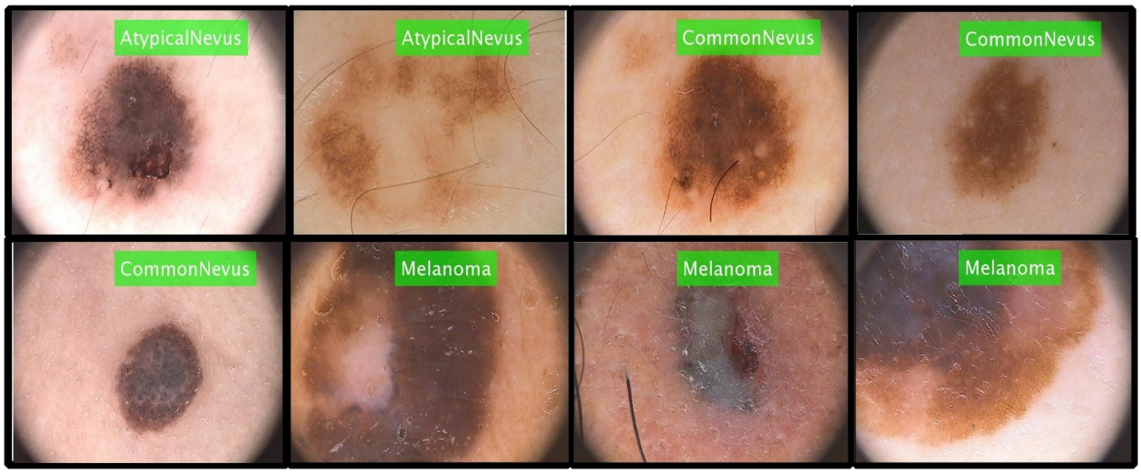

Fig. 9 Proposed labeled results using PH2 dataset

Table 1 Proposed segmentation results using PH2 dataset

Image	Dice (%)	FNR (%)	Image	Dice (%)	FNR (%)
IMD004_lesion	92.0762	7.9238	IMD031_lesion	92.5184	7.4816
IMD015_lesion	93.4803	6.5197	IMD041_lesion	90.2665	9.7335
IMD016_lesion	95.3221	4.6779	IMD042_lesion	91.8763	8.1237
IMD018_lesion	91.9900	8.010	IMD047_lesion	94.1424	5.8576
IMD020_lesion	93.7683	6.2317	IMD057_lesion	91.9757	8.0243
IMD021_lesion	96.2730	3.7270	IMD078_lesion	94.3391	5.6609
IMD027_lesion	90.3243	9.6757	IMD135_lesion	86.8379	13.1621
IMD036_lesion	92.1502	7.8498	IMD149_lesion	96.5717	3.4283
IMD038_lesion	94.2521	5.7479	IMD150_lesion	96.7539	3.2461
IMD039_lesion	89.8617	10.1383	IMD168_lesion	96.1711	3.8289
IMD040_ lesion	95.9971	4.0029	IMD169_lesion	96.5873	3.4127
IMD171_lesion	94.9446	5.0554	IMD142_lesion	92.4603	7.5397
IMD164_lesion	93.2232	6.7768	IMD243_lesion	93.2022	6.7978
IMD312_lesion	92.2997	7.7003	IMD331_lesion	94.3206	5.6794
IMD423_lesion	92.4658	7.5342	IMD437_lesion	94.7521	5.2479
Average (%)				94.3239	5.6761

We also compared the proposed average segmentation accuracy with the recent state-of-the-art methods to show the authenticity of the saliency segmentation. Pathan et al. [48] presented an domain knowledge based approach for lesion segmentation and achieved accuracy 93.4% using PH2 dataset. Penis et al. [49] introduced an Delaunay triangulation approach for melanoma segmentation and reported average accuracy of 89.6% using PH2 dataset. Bi et al. [50] presented a multi-stage fully connected CNN network for skin lesion segmentation and obtained dice rate 90.25% for melanoma studies and 90.77% for non-melanoma studies, whereas our proposed saliency method gives average dice rate of 94.32% and error rate of 5.67% which is exceptional as compared to existing approaches using PH2 dataset. Similarly, a comparison is also conducted for ISBI 2016 dataset with existing techniques. Penis et al. [49] also used ISBI 2016 dataset and reported overall average dice rate of 91.18%. Sedai et al. [51] presented an CNN method for lesion segmentation and achieved average segmentation accuracy of 92.3%. In [48], authors also used ISBI 2016 dataset and reported accuracy of 94.6%. In this method, we achieve a dice rate of 94.9227% with the minimized error rate up to 5.0773%. These results show that our system gives exceptional performance on ISBI 2016 dataset as compared to the listed methods (Table 2).

4.4 Classification results

The proposed classification results computed in three steps: (1) color features are fed to NN and compared accuracy with other methods like LSVM, cosine tree, and few more, (2) geometric features are feed NN for classification, (3) fusion of avg_pool and FC layer deep CNN features, and (4) optimal features selection approach. The 50-by-50 approach is opted in this work for all datasets along tenfold cross-validation. In Table 3, the recognition results using ISBI 2016 dataset are presented. The best achieved accuracy for color features of 91.8%, precision rate 91.5%, AUC of 0.97, specificity of 92%, sensitivity of 92%, and FN rate is 8.0%, respectively, on NN. The execution time of NN for color features is 17.88 s. The best achieved execution time using color features is 6.78 s on LSVM. In the second phase, classification is performed and achieved maximum accuracy of 94.5% on NN. The other calculated performance measures of NN such as precision rate, AUC, specificity, sensitivity, and FN rate are 94%, 0.97, 95%, 93%, and 6.50%, respectively. The best execution time for geometric features is 2.98 s on LSVM, whereas on NN, it is executed in 2.98 s. In the third step, CNN features are extracted from avg_pool and FC layer of Inception V3 model and fusion is performed through parallel approach. The best achieved accuracy on fusion process is 93% which is 1.5% less as compared to geometric features accuracy. The performance of fused CNN features is finally

Table 2 Proposed segmentation results using ISBI 2016 dataset

Image no.	Dice (%)	Negative rate (%)	Image no.	Dice (%)	Negative rate (%)
1	96.3536	3.6464	21	91.678	8.322
2	92.5686	7.4314	22	94.2862	5.7138
3	96.3422	3.6578	23	96.8841	3.1159
4	94.3239	5.6761	24	96.1275	3.8725
5	97.9448	2.0552	25	94.225	5.775
6	90.9322	9.0678	26	93.2442	6.7558
7	93.0414	6.9586	27	92.1744	7.8256
8	96.1587	3.8413	28	90.625	9.375
9	90.2842	9.7158	29	94.4075	5.5925
10	95.2106	4.7894	30	94.2461	5.7539
11	97.1560	2.844	31	94.7472	5.2528
12	91.2158	8.7842	32	91.0949	8.9051
13	90.3740	9.626	33	93.8354	6.1646
14	94.2148	5.7852	34	96.0528	3.9472
15	93.1066	6.8934	35	91.6426	8.3574
16	95.1627	4.8373	36	95.3028	4.6972
17	93.6104	6.3896	37	95.1007	4.8993
18	94.0248	5.9752	38	93.9915	6.0085
19	94.859	5.141	39	96.3885	3.6115
20	97.2522	2.7478	40	94.2151	5.7849
Average (%)				94.9227	5.0773

Table 3 Recognition results on the ISBI 2016 dataset

Classifier	Method				Performance measures						
	Color	Geo	Fused	Selected	Time (s)	FNR (%)	Sensitivity (%)	Specificity (%)	AUC	Precision (%)	Accuracy (%)
Coarse tree	✔				6.87	28.0	70.0	63.0	0.72	70.5	72.1
		✔			3.26	36.0	64.0	46.0	0.66	71.0	68.6
			✔		10.89	31.0	69.0	54.0	0.74	70.5	72.7
				✔	9.29	11.0	89.0	84.0	0.97	89.0	88.9
Linear SVM	✔				6.78	33.0	67.0	50.0	0.75	69.5	71.6
		✔			2.98	40.0	60.5	38.0	0.66	63.0	66.4
			✔		10.64	29.0	71.0	67.0	0.77	70.0	72.1
				✔	9.79	11.5	88.5	84.0	0.96	88.5	88.3
Cosine KNN	✔				8.90	31.5	68.5	58.0	0.76	69.0	71.0
		✔			4.12	39.0	61.0	38.0	0.67	64.0	66.9
			✔		10.73	31.5	65.5	56.0	0.75	70.0	71.9
				✔	8.00	13.5	86.5	84.0	0.94	87.0	86.7
Quadratic SVM	✔				8.95	32.5	67.5	55.0	0.70	69.0	71.0
		✔			3.99	36.0	64.0	43.0	0.68	67.5	69.7
			✔		9.93	30.5	69.5	57.0	0.71	70.5	72.7
				✔	6.32	12.5	87.5	84.0	0.95	88.0	87.8
Medium Gaussian SVM	✔				9.76	33.0	67.0	48.0	0.78	70.0	71.6
		✔			4.66	33.0	67.0	58.0	0.69	67.0	69.1
			✔		10.69	32.0	68.0	51.0	0.74	70.0	72.1
				✔	8.11	11.0	89.0	100.0	0.89	91.0	89.0
Neural network	✔				17.88	**8.20**	**92.0**	**92.0**	**0.97**	**91.5**	**91.8**
		✔			31.00	**5.50**	**93.0**	**95.0**	**0.97**	**94.0**	**94.5**
			✔		25.23	**7.00**	**93.0**	**95.0**	**0.97**	**93.0**	**93.0**
				✔	21.01	**3.90**	**92.0**	**98.0**	**0.95**	**95.0**	**96.1**

Bold values indicate the significant results

improved by proposed optimal selection method which resulted in an accuracy of 96.1% on NN which is better as compared to color, geometric and fused feature sets. The classification performance of NN is also verified by confusion matrices given in Fig. 10. Moreover, the proposed optimal selection approach performs on the average 10% better on supervised learning methods as coarse tree, LSVM, cosine KNN, quadratic SVM, and M-Gaussian SVM, whereas on NN, the 2% better results are achieved. The classification time of each method is also plotted in Fig. 11.

The recognition results of proposed system using ISBI2017 dataset are presented in Table 4. The results are described in four different steps as color features, geometric features, fused features, and proposed selected optimal features. On color features, 92.1% accuracy is achieved through NN with a minimum execution time of 15.43 s. The recognition accuracy is improved by 1.3% on geometric features through NN, whereas the execution time is 28.21 s. The execution time on NN is higher as

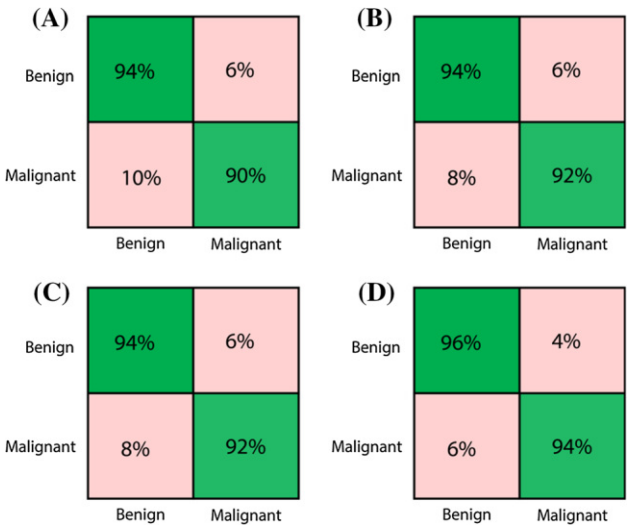

Fig. 10 Confusion matrices of using ISBI 2016 dataset by employing different features extraction techniques. **a** Color features, **b** geometric features, **c** fusion of CNN features, **d** proposed selected features

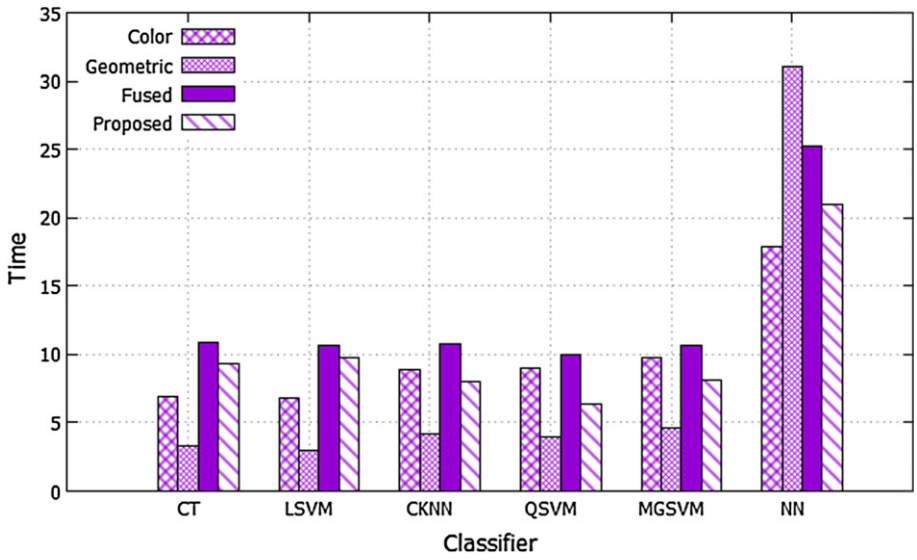

Fig. 11 Comparison of recognition time on different methods using ISBI 2016 dataset

Table 4 Proposed recognition results on ISBI 2017 dataset

Classifier	Method				Performance measures						
	Color	Geo	Fused	Selected	Time (s)	FNR (%)	Sensitivity (%)	Specificity (%)	AUC	Precision (%)	Accuracy (%)
Coarse tree	✔				20.18	9.50	90.5	90.0	0.95	90.5	90.5
		✔			26.38	8.70	91.0	94.0	0.95	91.5	91.3
			✔		26.64	9.40	90.5	90.0	0.91	90.5	90.6
				✔	23.65	3.20	92.5	100	0.93	97.5	96.8
Linear SVM	✔				19.71	8.90	91.5	90.0	0.97	92.0	91.9
		✔			26.32	7.30	93.0	94.0	0.96	93.0	92.7
			✔		31.15	7.30	93.0	93.0	0.96	93.0	92.7
				✔	23.04	7.10	93.0	94.0	0.91	87.5	92.9
Cosine KNN	✔				16.06	10.1	90.0	88.0	0.94	90.0	89.9
		✔			**20.01**	6.70	93.0	94.0	0.96	93.0	93.3
			✔		31.97	9.50	90.5	91.0	0.96	90.5	90.5
				✔	26.15	6.40	87.0	98.0	0.95	92.0	93.6
Quadratic SVM	✔				17.88	8.20	92.0	92.0	0.97	91.5	91.8
		✔			31.00	5.50	93.0	95.0	0.97	94.0	94.5
			✔		**25.23**	7.00	93.0	95.0	0.97	93.0	93.0
				✔	21.01	3.90	92.0	98.0	0.95	95.0	96.1
Medium Gaussian SVM	✔				15.81	8.10	92.0	93.0	0.96	92.0	91.9
		✔			20.02	5.70	94.5	97.0	0.97	94.5	94.3
			✔		29.26	8.20	92.0	93.0	0.97	92.0	91.8
				✔	24.21	5.30	92.0	97.0	0.95	91.5	94.7
Neural network	✔				**15.43**	7.90	92.0	94.0	0.96	92.0	**92.1**
		✔			28.21	6.60	93.0	95.0	0.97	93.5	**93.4**
			✔		31.68	5.50	94.5	96.0	0.97	94.5	**94.5**
				✔	**19.77**	**3.00**	92.5	**100**	0.93	**97.5**	**97.0**

Bold values indicate the significant results

compared to cosine KNN for geometric features. The deep CNN features are extracted in this work and achieve best accuracy of 94.5% through NN which is better as compared to color and geometric features. But on fusion process, the

execution time increases, whereas the overall accuracy is improved as compared to color and geometric vectors. To overcome the problem of higher execution time, a new optimal feature selection technique is proposed in this work which not only decreases the execution time but also improves the recognition accuracy. The best achieved accuracy of 97% on ISBI 2016 through NN along execution time 19.77 s. The recognition accuracy of NN on color, geometric, fused CNN features, and optimal selected features are also shown in Fig. 12. Moreover, the time comparison for each classifier for various feature extraction methods is plotted in Fig. 13. The proposed results explains that optimal selected features provides significant accuracy on ISBI 2017 dataset through NN as compared to other supervised learning methods such as CT, LSVM, CKNN, QSVM, MGSVM, and NN.

In Table 5, the recognition results are presented using PH2 dataset. The experiments are performed on four distinct feature sets as color, geometric, the fusion of deep CNN, and optimal selected features. For color features, a maximum achieved accuracy of 74.50% through NN along computation time 22.36 s. The accuracy of color features is improved when geometric features are extracted and reached up to 89.10% along computational time 18.46 which is also minimized. In this work, deep CNN features are extracted from two layers of Inception V3 model and fused by a parallel approach. The maximum achieved an accuracy of 92.50% which is improved as compared to color and geometric features. But the execution time is increased for the fusion process and reached up to 29.53 s. Finally, to resolve the problem of high computational cost,

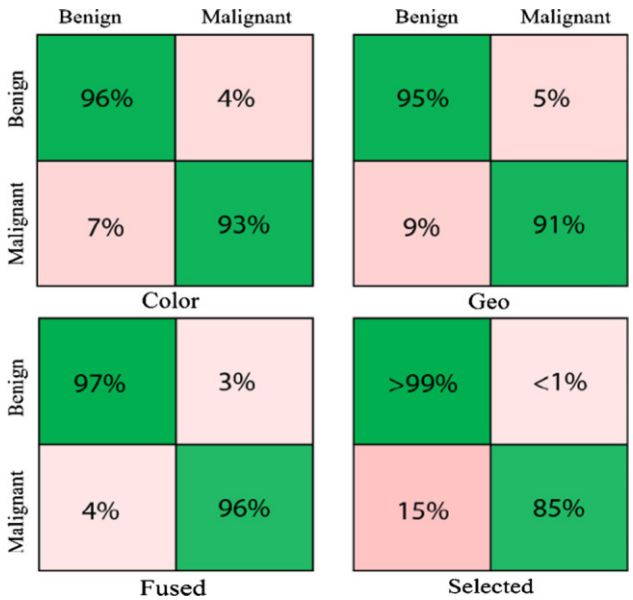

Fig. 12 Confusion matrices of proposed classification results on ISBI 2017 through neural network

the optimal features are selected and obtained maximum accuracy of 97.74 and time 11.86 s on NN which clearly shows that the proposed fusion and selection scheme outperforms on NN using PH2 dataset as compare to other supervised learning methods. The verification of recognition accuracy which is achieved on NN is also verified by confusion matrices given in Fig. 14. The overall comparison of computational time for all classifiers is also plotted in Fig. 15.

4.5 Discussion

In this section, we are conducting a fair comparison with existing algorithms in terms of accuracy, sensitivity, specificity rate, and recognition testing time, but also providing a brief analysis of our proposed algorithm.

The comparison results are shown in Table 6. Chen et al. [36] implemented an multi-tasking framework for skin lesion classification. They used ISBI 2017 dataset for evaluation and reported an accuracy of 80.1%. Yang et al. [52] implemented CNN model for lesion classification and achieved accuracy of 83%, sensitivity rate of 60.7%, and specificity 88.4%, respectively, for ISBI 2017 dataset. Khan et al. [2] used three datasets including PH2, ISBI 2016, and ISBI 2017 and achieved accuracy of 97.5%, 83.2%, and 88.2%, respectively. They also computed sensitivity and specificity values for all datasets as presented in Table 6. Harangi et al. [53] improved the recognition rate for lesion classification and achieved accuracy of 89.1% using ISBI 2017 which is good as compared to the existing methods. Oliveira et al. [54] achieved exceptional performance and obtained accuracy of 94.3%, sensitivity rate 91.8%, and specificity 96.7%, respectively. In this work, our method outperforms for ISBI 2017 dataset by achieving accuracy of 97%, sensitivity rate of 92.50%, specificity of 100%, and recognition time of 19.77 s. Moreover, for ISBI 2016, our method attains an accuracy of 96.1%, sensitivity of 92%, specificity of 98%, and recognition time is 21.01 s. Recently, the best classification accuracy is achieved by Khan et al. [2], in which they reported accuracy of 97.5%, sensitivity 96.67%, and specificity 98.7%, whereas our method obtains an accuracy of 97.74%, sensitivity of 97.39%, specificity of 98.10, and recognition time is 11.60 s, which is an exceptional result compared to the existing methods. Moreover, in the testing process, we also compute the individual time of each step and plotted in Fig. 16. The steps which are involved in proposed method are contrast enhancement, HSV conversion, high-intensity-based selection, saliency-based segmentation, parallel feature fusion, optimal feature selection, and neural network. From Fig. 16, it is clearly shown that the NN bases step consume much time as compared to others. In

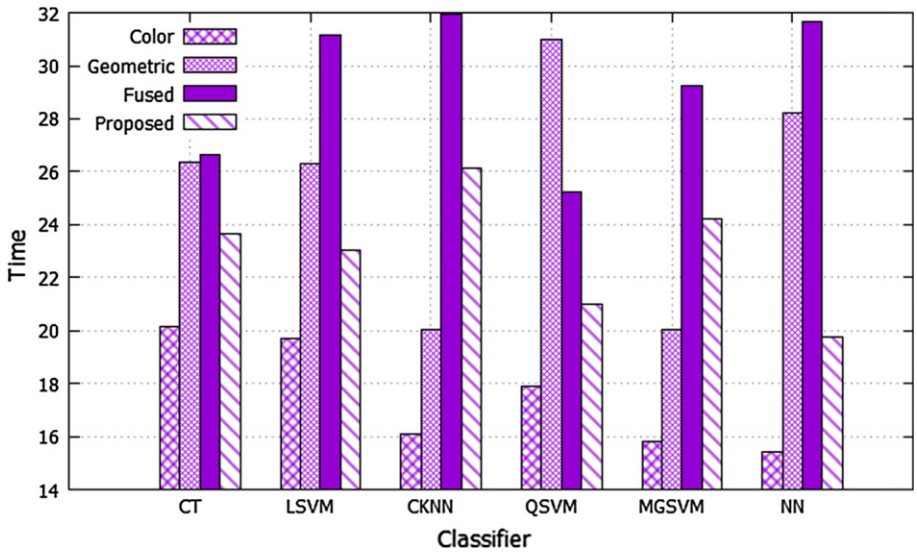

Fig. 13 Time comparison for each classifier on different features extraction methods using ISBI 2017 dataset

Table 5 Lesion recognition results on PH2 dataset

Classifier	Method				Performance measures						
	Color	Geo	Fused	Selected	Time (s)	FPR (%)	Sensitivity (%)	Specificity (%)	AUC	Precision (%)	Accuracy (%)
Coarse tree	✔				**10.14**	15.00	63.66	63.67	0.81	64.00	71.50
		✔			21.89	7.66	78.33	78.34	0.94	86.67	85.50
			✔		26.71	6.66	83.66	83.67	0.94	82.00	86.00
				✔	14.65	3.00	94.00	94.00	0.98	93.67	93.80
Linear SVM	✔				13.21	13.66	66.66	66.67	0.82	66.66	73.00
		✔			19.38	5.00	85.00	85.00	0.93	87.00	89.00
			✔		27.25	4.00	87.33	87.34	0.95	89.00	91.00
				✔	16.97	3.66	92.66	92.67	0.95	93.00	92.70
Cosine KNN	✔				18.15	19.66	64.66	64.67	0.78	58.66	70.50
		✔			19.57	6.00	84.66	84.67	0.94	86.00	88.50
			✔		**23.95**	5.00	86.66	86.67	0.95	89.00	91.00
				✔	12.10	3.00	94.00	94.00	0.95	96.00	93.80
Quadratic SVM	✔				23.05	14.0	66.66	66.67	0.83	67.00	72.50
		✔			17.37	9.66	75.33	75.34	0.88	76.00	86.50
			✔		28.69	2.00	87.00	86.67	0.95	89.34	91.00
				✔	14.93	2.66	95.00	95.00	0.98	96.00	94.80
Medium Gaussian SVM	✔				22.90	15.66	60.00	60.00	0.80	57.00	71.00
		✔			**16.84**	7.00	82.33	82.34	0.94	84.00	86.00
			✔		26.58	5.33	86.00	86.00	0.96	87.34	89.00
				✔	13.95	2.66	95.00	95.00	0.98	91.67	94.80
Neural network	✔				22.36	12.66	67.66	67.67	0.80	67.66	**74.50**
		✔			18.46	5.00	85.00	85.00	0.94	86.67	**89.10**
			✔		29.53	5.33	86.00	86.00	0.97	91.70	**92.50**
				✔	**11.86**	1.66	**97.39**	**98.10**	0.97	91.00	**97.74**

Bold values indicate the significant results

addition, we also calculate the training execution time for all datasets as shown in Fig. 17.

With the proposed algorithm, we are achieving more than 94% segmentation accuracy. In the proposed

Fig. 14 Confusion matrices for verification of recognition results which are achieved through NN

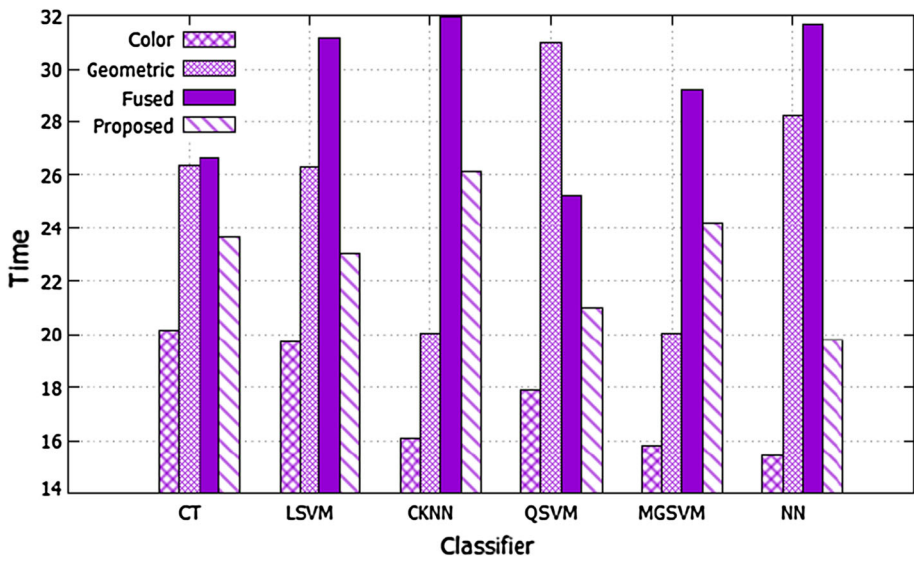

Fig. 15 Time comparison for each classifier on different features extraction methods

framework, improved segmentation accuracy leads to the extraction of most relevant features, which in turn produces good classification results. Therefore, our primary focus here is to extract only the salient foreground, while ignoring the background details. An addition of a preprocessing step enables us to extract lesion area with greater accuracy. In the classification phase, promising results are achieved on all datasets such as 96.1% on ISBI 2016, 97% on ISBI 2017, and 97.74% on PH2, and in minimum computation time.

5 Conclusion

Early detection and recognition of skin lesion is a vital study in the field of medical, which attracted many researchers in the last decade. We proposed a saliency segmentation and deep CNN-based lesion classification approach. A contrast is improved in the first step and later saliency map is computed for lesion segmentation. In the features extraction step, DCNN features are extracted through two different layers and fused by a decision-

Table 6 Comparison with recent state-of-the-art articles of proposed recognition accuracy

Reference	Year	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	Recognition time (s)
Chen et al. [36]	2018	ISBI 2017	80.1	–	–	–
Harangi et al. [53]	2018	ISBI 2017	89.1	–	–	–
Yang et al. [52]	2018	ISBI 2017	83.0	60.7	88.4	–
Oliveira et al. [54]	2017	ISBI 2017	94.3	91.8	96.7	
Khan et al. [2]	2018	ISBI 2017	88.2	88.5	91.0	
Khan et al. [2]	2018	ISBI 2016	83.2	75.5	93.0	–
Yu et al. [55]	2017	ISBI 2016	85.5	85.3	**99.3**	–
Vason et al. [56]	2017	ISBI 2016	83.6	–	–	–
Maia et al. [57]	2018	PH2	92.5	72.5	97.5	–
Bi et al. [58]	2016	PH2	92.0	87.5	93.13	–
Khan et al. [2]	2018	PH2	97.5	96.67	**98.7**	–
Proposed	2018	ISBI 2017	**97.0**	**92.50**	**100**	**19.77**
		ISBI 2016	**96.1**	**92.0**	98.0	**21.01**
		PH2	**97.74**	**97.39**	98.10	**11.60**

Bold values indicate the significant results

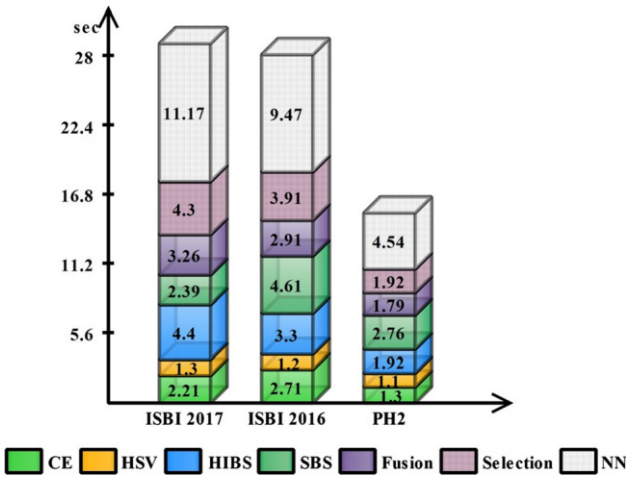

Fig. 16 Individual computational time of each step

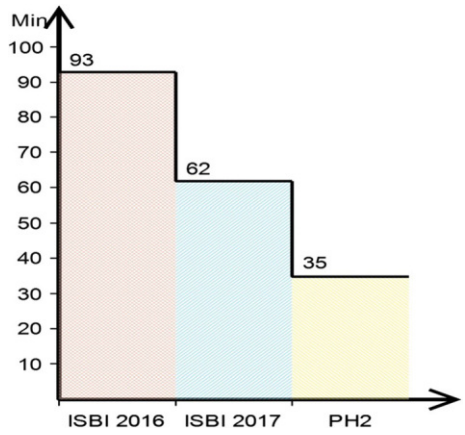

Fig. 17 Training time in minutes for selected datasets

controlled parallel (DCdP) approach. Later, optimal features are selected and provided to ANN for classification.

From the results, we can comfortably conclude that an addition of a good preprocessing step not only increases the segmentation accuracy, but also affect the classification accuracy. Additionally, a selection of most discriminant information on one hand decreases the computation time but also improves the classification accuracy.

5.1 Limitations and future work

The limitations of this work include: (a) proposed method is somewhat dependent on the preprocessing step (contrast stretching), (b) addition of preprocessing and segmentation steps make an extra computation, (c) when the number of predictors decreases, there exist a bright chance of performance degradation, especially in the case of complex dermoscopic images.

In the future, we plan to work on CNN-based lesion segmentation using ISBI 2017 dataset. Moreover, the attribute-based lesion recognition is a new research area in the direction of skin cancer. Addition to that, we are planning to implement reinforcement learning for skin lesion detection and classification.

Compliance with ethical standards

Conflict of interest All authors have no conflict of interest and contribute equally in this work for results compilation and other technical support.

References

1. Oliveira RB, Papa JP, Pereira AS, Tavares JMR (2018) Computational methods for pigmented skin lesion classification in images: review and future trends. Neural Comput Appl 29:613–636
2. Khan MA, Akram T, Sharif M, Shahzad A, Aurangzeb K, Alhussein M et al (2018) An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification. BMC Cancer 18:638
3. Akram T, Khan MA, Sharif M, Yasmin M (2018) Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-018-1051-5
4. CFAFACS. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf. Accessed May 3, 2018
5. Parkin D, Mesher D, Sasieni P (2011) 13. Cancers attributable to solar (ultraviolet) radiation exposure in the UK in 2010. Br J Cancer 105:S66
6. Nguyen AH, Detty SQ, Agrawal DK (2017) Clinical implications of high-mobility group box-1 (HMGB1) and the receptor for advanced glycation end-products (RAGE) in cutaneous malignancy: a systematic review. Anticancer Res 37:1–7
7. WABASCSCACS. http://www.cancer.org/cancer/skincancer-basalandsquamouscell/detailedguide/skin-cancer-basal-and-squamous-cell-what-is-basal-and-squamous-cell. Accessed Jan 31, 2018
8. Nasir M, Attique Khan M, Sharif M, Lali IU, Saba T, Iqbal T (2018) An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. Microsc Res Tech 81:528–543
9. Whiteman DC, Green AC, Olsen CM (2016) The growing burden of invasive melanoma: projections of incidence rates and numbers of new cases in six susceptible populations through 2031. J Investig Dermatol 136:1161–1171
10. Ruela M, Barata C, Marques JS, Rozeira J (2017) A system for the detection of melanomas in dermoscopy images using shape and symmetry features. Comput Methods Biomech Biomed Eng Imaging Vis 5:127–137
11. Satapathy SC, Fernandes SL, Lin H (2019) Stroke lesion segmentation and analysis using entropy/Otsu's function—a study with social group optimization. Curr Bioinform 14:305–313
12. Saba T, Khan MA, Rehman A, Marie-Sainte SL (2019) Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. J Med Syst 43:289
13. Liaqat A, Khan MA, Shah JH, Sharif M, Yasmin M, Fernandes SL (2018) Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. J Mech Med Biol 18:1850038
14. Fernandes SL, Tanik UJ, Rajinikanth V, Karthik KA (2019) A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians. Neural Comput Appl 29:1–12
15. Aurangzeb K, Haider I, Khan MA, Saba T, Javed K, Iqbal T et al (2019) Human behavior analysis based on multi-types features fusion and Von Nauman entropy based features reduction. J Med Imaging Health Inform 9:662–669
16. Amin J, Sharif M, Yasmin M, Fernandes SL (2018) Big data analysis for brain tumor detection: deep convolutional neural networks. Future Gener Comput Syst 87:290–297
17. Bokhari F, Syedia T, Sharif M, Yasmin M, Fernandes SL (2018) Fundus image segmentation and feature extraction for the detection of glaucoma: a new approach. Curr Med Imaging Rev 14:77–87
18. Acharya UR, Fernandes SL, WeiKoh JE, Ciaccio EJ, Fabell MKM, Tanik UJ et al (2019) Automated detection of Alzheimer's disease using brain MRI images—a study with various feature extraction techniques. J Med Syst 43:302
19. Jin Z, Zhou G, Gao D, Zhang Y (2018) EEG classification using sparse Bayesian extreme learning machine for brain–computer interface. Neural Comput Appl. https://doi.org/10.1007/s00521-018-3735-3
20. Zhang Y, Nam CS, Zhou G, Jin J, Wang X, Cichocki A (2018) Temporally constrained sparse group spatial patterns for motor imagery BCI. IEEE Trans Cybern 49:1–11
21. Zhang Y, Zhang H, Chen X, Liu M, Zhu X, Lee S-W et al (2019) Strength and similarity guided group-level brain functional network construction for MCI diagnosis. Pattern Recogn 88:421–430
22. Jiao Y, Zhang Y, Wang Y, Wang B, Jin J, Wang X (2018) A novel multilayer correlation maximization model for improving CCA-based frequency recognition in SSVEP brain–computer interface. Int J Neural Syst 28:1750039
23. Sharif M, Khan MA, Faisal M, Yasmin M, Fernandes SL (2018) A framework for offline signature verification system: best features selection approach. Pattern Recognit Lett. https://doi.org/10.1016/j.patrec.2018.01.021
24. Fernandes SL, Chakraborty B, Gurupur VP, Prabhu G (2016) Early skin cancer detection using computer aided diagnosis techniques. J Integr Des Process Sci 20:33–43
25. Khan MA, Akram T, Sharif M, Saba T, Javed K, Lali IU et al (2019) Construction of saliency map and hybrid set of features for efficient segmentation and classification of skin lesion. Microsc Res Tech 82:741–763
26. Afza F, Khan MA, Sharif M, Rehman A (2019) Microscopic skin laceration segmentation and classification: a framework of statistical normal distribution and optimal feature selection. Microsc Res Tech. https://doi.org/10.1002/jemt.23301
27. Khan MA, Javed MY, Sharif M, Saba T, Rehman A (2019) Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. In: 2019 international conference on computer and information sciences (ICCIS), pp 1–7
28. Chatterjee S, Dey D, Munshi S (2018) Optimal selection of features using wavelet fractal descriptors and automatic correlation bias reduction for classifying skin lesions. Biomed Signal Process Control 40:252–262
29. Codella NC, Nguyen Q-B, Pankanti S, Gutman D, Helba B, Halpern A et al (2017) Deep learning ensembles for melanoma recognition in dermoscopy images. IBM J Res Dev 61:5:1–5:15
30. Goyal M, Yap MH (2017) Multi-class semantic segmentation of skin lesions via fully convolutional networks. Preprint arXiv:1711.10449
31. Ross-Howe S, Tizhoosh H (2018) The effects of image pre-and post-processing, wavelet decomposition, and local binary patterns on U-nets for skin lesion segmentation. Preprint arXiv:1805.05239
32. Sarker M, Kamal M, Rashwan HA, Banu SF, Saleh A, Singh VK et al (2018) SLSDeep: skin lesion segmentation based on dilated residual and pyramid pooling networks. Preprint arXiv:1805.10241
33. Lopez AR, Giro-i-Nieto X, Burdick J, Marques O (2017) Skin lesion classification from dermoscopic images using deep learning techniques. In: 2017 13th IASTED international conference on biomedical engineering (BioMed), pp 49–54

34. Kawahara J, BenTaieb A, Hamarneh G (2016) Deep features to classify skin lesions. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pp 1397–1400
35. Celebi ME, Zornberg A (2014) Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification. IEEE Syst J 8:980–984
36. Chen S, Wang Z, Shi J, Liu B, Yu N (2018) A multi-task framework with feature passing module for skin lesion classification and segmentation. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), pp 1126–1129
37. Amin J, Sharif M, Yasmin M, Fernandes SL (2017) A distinctive approach in brain tumor detection and classification using MRI. Pattern Recognit Lett. https://doi.org/10.1016/j.patrec.2017.10.036
38. Klang E (2018) Deep learning and medical imaging. J Thorac Dis 10:1325
39. Naqi S, Sharif M, Yasmin M, Fernandes SL (2018) Lung nodule detection using polygon approximation and hybrid features from CT images. Curr Med Imaging Rev 14:108–117
40. Raza M, Sharif M, Yasmin M, Khan MA, Saba T, Fernandes SL (2018) Appearance based pedestrians' gender recognition by employing stacked auto encoders in deep learning. Future Gener Comput Syst 88:28–39
41. Khan MA, Akram T, Sharif M, Javed MY, Muhammad N, Yasmin M (2018) An implementation of optimized framework for action classification using multilayers neural network on selected fused features. Pattern Anal Appl 22:1–21
42. Sharif M, Khan MA, Akram T, Javed MY, Saba T, Rehman A (2017) A framework of human detection and action recognition based on uniform segmentation and combination of Euclidean distance and joint entropy-based features selection. EURASIP J Image Video Process 2017:89
43. Khan MA, Sharif M, Javed MY, Akram T, Yasmin M, Saba T (2017) License number plate recognition system using entropy-based features selection approach with SVM. IET Image Proc 12:200–209
44. Yang J, Yang J-Y, Zhang D, Lu J-F (2003) Feature fusion: parallel strategy vs. serial strategy. Pattern Recogn 36:1369–1381
45. Adeel A, Khan MA, Sharif M, Azam F, Umer T, Wan S (2019) Diagnosis and recognition of grape leaf diseases: an automated system based on a novel saliency approach and canonical correlation analysis based multiple features fusion. Sustain Comput Inform Syst. https://doi.org/10.1016/j.suscom.2019.08.002
46. Rashid M, Khan MA, Sharif M, Raza M, Sarfraz MM, Afza F (2019) Object detection and classification: a joint selection and fusion strategy of deep convolutional neural network and SIFT point features. Multimed Tools Appl 78:15751–15777
47. Arshad H, Khan MA, Sharif M, Yasmin M, Javed MY (2019) Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution. Int J Mach Learn Cybern. https://doi.org/10.1007/s13042-019-00947-0
48. Pathan S, Prabhu KG, Siddalingaswamy P (2018) Hair detection and lesion segmentation in dermoscopic images using domain knowledge. Med Biol Eng Comput 56:1–15
49. Pennisi A, Bloisi DD, Nardi D, Giampetruzzi AR, Mondino C, Facchiano A (2016) Skin lesion image segmentation using delaunay triangulation for melanoma detection. Comput Med Imaging Graph 52:89–103
50. Bi L, Kim J, Ahn E, Kumar A, Fulham M, Feng D (2017) Dermoscopic image segmentation via multi-stage fully convolutional networks. IEEE Trans Biomed Eng 64:2065–2074
51. Bozorgtabar B, Sedai S, Roy PK, Garnavi R (2017) Skin lesion segmentation using deep convolution networks guided by local unsupervised learning. IBM J Res Dev 61:6:1–6:8
52. Yang J, Xie F, Fan H, Jiang Z, Liu J (2018) Classification for dermoscopy images using convolutional neural networks based on region average pooling. IEEE Access 6:65130–65138
53. Harangi B (2018) Skin lesion classification with ensembles of deep convolutional neural networks. J Biomed Inform 86:25–32
54. Oliveira RB, Pereira AS, Tavares JMR (2017) Skin lesion computational diagnosis of dermoscopic images: ensemble models based on input feature manipulation. Comput Methods Programs Biomed 149:43–53
55. Yu L, Chen H, Dou Q, Qin J, Heng P-A (2017) Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging 36:994–1004
56. Vasconcelos CN, Vasconcelos BN (2017) Experiments using deep learning for dermoscopy image analysis. Pattern Recognit Lett. https://doi.org/10.1016/j.patrec.2017.11.005
57. Maia LB, Lima A, Pereira RMP, Junior GB, de Almeida JDS, de Paiva AC (2018) Evaluation of melanoma diagnosis using deep features. In: 2018 25th international conference on systems, signals and image processing (IWSSIP), pp 1–4
58. Bi L, Kim J, Ahn E, Feng D, Fulham M (2016) Automatic melanoma detection via multi-scale lesion-biased representation and joint reverse classification. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pp 1055–1058