Melanoma localization and classification through faster region-based convolutional neural network and SVM

Marriam Nawaz[1] · Momina Masood[1] · Ali Javed[1] · Javed Iqbal[1] · Tahira Nazir[1] ·
Awais Mehmood[1] · Rehan Ashraf[2] ⬤

Ⓒ The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Melanoma is a lethal skin cancer disease affecting millions of people around the globe and has a high mortality rate. Dermatologists perform the manual inspection through visual analysis of pigmented skin lesions for melanoma identification at the early stage. However, manual inspection for melanoma detection is limited due to variable accuracy and lesser availability of dermatologists. Therefore, there exists an urgent need to develop automated melanoma detection methods that can effectively localize and classify skin lesions. Accurate localization and classification of the melanoma lesions is a challenging task due to the presence of low contrast information between the moles and skin part, the massive color similarity between the infected and non-infected skin portions, presence of noise, hairs, and tiny blood vessels, variations in color, texture, illumination, contrast, blurring, and melanoma size. To address these afore-mentioned challenges, we propose an effective and efficient melanoma detection method. The proposed method consists of three steps: i) image preprocessing, ii) employing Faster Region-based Convolutional Neural Network (Faster-RCNN) for melanoma localization, and iii) application of Support Vector Machine (SVM) for the classification of localized melanoma region into benign and malignant classes. Performance of the proposed method is evaluated on the benchmark ISIC-2016 dataset launched by ISBI challenge-2016 that is diverse in terms of variations in illumination, color, texture, and size of melanoma, and presence of blurring, noise, hairs, and tiny blood vessels, etc. Moreover, we have also performed a cross-dataset validation over the ISIC-2017 dataset to show the efficacy of our method in real-world scenarios. Our experimental results illustrate that the proposed framework is efficient and able to effectively localize and classify the melanoma lesion than state-of-the-art techniques.

Keywords Deep-learning · Faster-RCNN · Melanoma · Medical imaging · SVM · Skin lesion

✉ Rehan Ashraf
  rehan@ntu.edu.pk

[1] Department of Computer Science, University of Engineering and Technology, Taxila 47050, Pakistan

[2] Department of Computer Science, National Textile University, Faisalabad 37610, Pakistan

1 Introduction

Melanoma is the most serious type of skin cancer that develops in the skin cells called melanocytes [43]. A study [41] reported about 10,000 annual deaths because of melanoma skin cancer in the US only. The melanoma skin lesion is caused by the abnormal production of skin particles which in turn produce the body cells [46]. Based on the adversity of melanoma, these skin moles vary in texture and colors like brown, pink, red and black, etc. If the size of moles is beyond 6 mm with abnormal color, then there is a need for a detailed checkup by the dermatologist for possible melanoma containment. Melanoma lesions are divided into two classes named *Benign* and *Malignant*. The benign melanoma lesions depict the initial level of skin cancer, which is curable. Whereas, malignant melanoma is a dangerous type of skin cancer that can eventually cause the death of the patient. Initially, the dermatologists used to analyze such abnormal moles with naked eyes, by examining their texture, color, and size [34]. However, we often experience delays in the screening process of melanoma because of the limited availability of dermatologists. Melanoma identification at an early stage is crucial as not only it increases the survival chances of the patient but can also save them from the tough surgical processes [37]. The availability of sophisticated computer vision algorithms these days has motivated researchers to develop effective automated methods for melanoma detection [10, 37].

Existing automated approaches for melanoma detection can be broadly categorized into handcrafted features and deep learning (DL) based approaches. The handcrafted features use the key points of extraction-based methods for the recognition of skin moles [6, 13, 47]. However, these techniques are unable to accurately detect the skin lesions because of deviations in size, texture, and color of the skin moles. To increase the accuracy of melanoma identification systems, classification is performed after the segmentation of the melanoma region from the normal skin, as done by the expert dermatologists [8]. Region-of-interest (ROI) based techniques are employed in [40, 44] for the segmentation of melanoma regions. These techniques provide a better representation of the melanoma attributes during feature extraction and precise classification of melanoma affected regions. Thus, effective segmentation is a mandatory requirement for developing accurate melanoma detection systems [19, 20, 42] However, the performance of the ROI or threshold-based methods degrades significantly in case of low-resolution images, low illumination conditions, and variations in contrast, illumination conditions, and chrominance. For real-world problems, it is practically difficult to capture the image with uniform properties, therefore, we need to develop effective melanoma detection systems to overcome the above-mentioned limitations.

In the last few years, we witnessed the effectiveness of DL based approaches in various domains including medical image processing [3, 32, 36, 48]. The deep convolutional neural network (CNN) automatically learns complex keypoints directly from the input samples and provides improved recognition ability of melanoma-affected regions. Based on these advantages, DL based skin lesion detection has attracted the researcher's attention [49]. However, most of the DL based approaches require preprocessing of the input images to overcome the problem of feature map saturation [4]. To avoid the preprocessing step, in [5] a SegNet based network is used which maps input image to pixel-wise semantic label using feature learning. In [4] recurrent and CNNs are employed, whereas, in [9] a semi-automated fully CNN is used for melanoma detection.

Timely and precise automated identification and classification of melanoma lesion is still a challenging task because of the presence of low-intensity information among melanoma lesion and skin portion, and massive similarity between the melanoma affected and non-affected

body parts. In this paper, we have tried to overcome these challenges by employing a faster regional convolutional neural network (Faster-RCNN) to compute the deep features of input images and to localize the melanoma moles. These deep features are then used to train the SVM for classification. The proposed method is robust to variations in chrominance, intensity, contrast, illumination conditions, hair, and tiny blood vessels, blurring, and high-dense noisy images. The major contributions of the proposed work are as follows:

- Accurate localization of melanoma regions due to the region proposal network of Faster-RCNN.
- Efficient classification of melanoma because of the competence of the SVM classifier to deal with the over-fitted training data.
- Rigorous experimentation was performed against several latest melanoma recognition methods on a standard ISIC-2016 database containing different distortions like blurring, chrominance and intensity variations, the existence of hair and tiny blood vessels, and high-dense noisy images to show the efficacy of the introduced framework. Moreover, we have also performed cross-validation over the ISIC-2017 dataset to show the applicability of our method in real-world scenarios.
- To the best of our knowledge, it is the first time in medical analysis when Faster-RCNN has been employed for skin lesion detection. Reported results exhibit the efficacy of Faster-RCNN to detect the melanoma moles and computation of a deep and discriminative set of features with improved performance.

The rest of the paper has the following structure: Section II presents the related work, while the proposed framework is discussed in detail in Section III. Performance evaluation of our framework is presented in Section IV, and finally, Section V concludes the proposed work.

2 Related work

This section presents a critical investigation of the existing latest melanoma recognition methods. Existing works have used either conventional machine learning (ML)-based techniques or DL-based techniques for melanoma detection.

The conventional ML classifiers have been employed in [7, 9, 14, 39, 45] for melanoma detection. In [9, 14], hand-crafted key-point extraction-based approaches were used for melanoma detection. Codella et al. [14] employed the edge and color histogram along with the local binary patterns (LBP) for melanoma lesion identification. Finally, the SVM classifier was trained for melanoma classification. Similarly, Barata et al. [7] introduced a framework for skin lesion localization using the local and global features extracted from the input images. More specifically, Laplacian pyramids and gradient histogram were used for global key-points extraction, whereas, Bag of Features (BoF) was employed for local keypoints extraction. Based on the obtained features, KNN, SVM, and AdaBoost classifiers were trained for skin lesions classification. It is concluded in [7] that global features are more efficient in melanoma detection over local features. Rehman et al. [39] presented a method for melanoma localization. Speeded-up robust feature (SURF) descriptor along-with histogram-oriented gradient (HOG) was employed for feature extraction that was then used to train the SVM for classifying the melanoma moles. This technique is robust to melanoma detection for low contrast skin lesion images at the expense of increased features computation cost. Singh et al. [45] employed

the Zernike Moments (ZM) and Pseudo Zernike Moments (PZM) features to train the SVM for the classification of skin lesions. Alcon et al. [1] presented a method based on using information about the medical history of the patient before performing the diagnosing procedure. Otsu's thresholding algorithm [31] was used for the lesion segmentation, and the ABCD rule of dermatoscopy was employed for feature extraction. Finally, a hybrid classification model consisting of Decision Tree Learner (J48), Decision Stump, Logistic Model Trees, and Bayesian Networks were trained to classify the skin moles. This approach [1] exhibits better accuracy for melanoma classification, however, it achieves poor specificity. Cavalcanti et al. [12] proposed a two-stage technique to classify skin lesions. After performing the segmentation using the 3-channel image representation-based technique [11], ABCD rules of dermatoscopy along with melanin variation features were employed to compute the key points. Finally, the KNN classifier was trained to discriminate against the skin moles. This method [12] is unable to achieve a better classification accuracy of melanoma lesions due to the morphological variations of skin moles. Giotis et al. [21] employed both automated and manual features to develop a technique for melanoma detection. Automated features were computed using the color and texture information of the skin moles, whereas, manual features like size, color, shape, and body part where moles exist, etc. were decided by the dermatologist after performing the examining process. Finally, feature selection between the manual and automated features was performed using a voting process. This method provides better performance for melanoma lesion detection, however, the hand-crafted key-features extraction-based methods in general exhibit lower performance for melanoma localization because of change in size, shape, and chrominance of the skin moles. Hu et al. [28] presented an approach for the automated classification of melanoma lesions. Initially, [28] introduced an efficient codebook learning technique based on FSM, which employed the linearly independent and linear prediction (LP) approaches to compute keypoints similarity. Secondly, RGB color histogram and SIFT were adopted to calculate the final feature vector. Finally, the SVM classifier was trained over the computed keypoints to classify the input samples into various classes. The approach in [28] is economically efficient, however, it may not performs well for the images with intense variations in color.

Recently, deep-learning methods are commonly used in various computer vision applications due to their high accuracy rate [52]. Nida et al. [35] introduced an automated framework for the identification and segmentation of skin lesions. RCNN was employed for detecting the melanoma moles followed by applying the Fuzzy C-mean (FCM) for lesion segmentation. Although, this method provides superior melanoma segmentation accuracy, however, at the overhead of increased computational cost. Gulati et al. [24] proposed a DL network for automated classification of a skin lesion by employing two different architectures of CNN named AlexNet and VGG16 for feature extraction. The extracted key-points were used to train the ECOC-SVM classifier to differentiate the melanocytic moles. A hybrid CNN model [33] based on AlexNet, ResNet-18, and VGG16 was employed for feature extraction. Later, these features were used to train the SVM for melanoma classification. Similarly, Harangi et al. [26] presented a hybrid CNN model consisting of GoogLeNet, AlexNet, ResNet, and VGGNet to classify the skin moles. The output of each network was fused to generate the output. These methods [24, 26, 33] exhibit better melanoma classification performance, however, these techniques are computationally complex. Li et al. [30] presented a methodology for the automated classification of melanoma lesions. Lesion Feature Network (LFN) was used for computing the image features. A DL architecture containing two fully convolutional residual networks (FCRN) was introduced to consecutively generate the segmentation and

classification results. Finally, the lesion index calculation unit (LICU) was employed to improve the classification output by computing the distance heat-map. Hosny et al. [27] introduced a CNN-based framework for skin lesion classification. Data augmentation and transfer learning along-with AlexNet was applied to classify the skin moles into three classes. Yap et al. [54] presented a methodology for melanoma classification by employing a ResNet DL model for extracting the image features. Afterward, the extracted key-points were combined with the patient's medical history. This method [26] provides reasonable melanoma classification accuracy, however, unable to perform well for those scenarios where the patient's metadata is missing. Al-Masni et al. [2] utilized the VGGNet model for the classification of skin cancer moles. This approach [2] exhibits better classification performance, however suffers from the over-fitting problem. Yang et al. [53] proposed a DL-based framework for the automated segmentation and classification of melanoma moles. Initially, a region average pooling (RAPooling) approach was used which computed the keypoints from the region of interest. Secondly, an end-to-end classification network together with the segmentation process was designed, which employed the localized lesion area to guide the classification by RAPooling. Finally, the RankOpt-based classifier was used for melanoma classification. The approach in [53] performs well for skin moles classification, however, it is an economically inefficient method. Existing literature on melanoma detection is unable to achieve better classification performance under variations in illumination conditions, color and texture of skin moles, contrast, blurring, and high-dense noisy images. Additionally, DL-based approaches are computational complex. Thus, to address the limitations, there exists a dire need to develop a more effective and efficient melanoma detection method.

3 Materials and methods

This work presents a novel technique for the automatic detection and classification of melanoma lesions by employing the Faster-RCNN deep learning model and SVM classifier. In the first step, the input image is preprocessed to eliminate the unnecessary objects that can degrade the classification results of the presented methodology. We employ the Faster-RCNN technique on the processed images to compute the deep features and to localize the melanoma moles. Finally, these deep features are used to train a binary SVM for melanoma lesion classification into benign and malignant classes. The architecture of the proposed framework is shown in Fig. 1.

3.1 Preprocessing

For real-life scenarios, it is practically impossible to obtain images without lightning or chrominance variations that can occur due to the sudden change in illumination source, or reflection of the light produced by the skin. The presence of unnecessary information in the training images can affect the detection performance of the proposed network. Additionally, the presence of hair and tiny blood vessels in skin mole images also affects the classification performance. Therefore, a light adjustment step is performed on the input images to address these problems. For this purpose, the morphological closing operation is applied to remove the undesirable information from the images. To restore the quality of the images, an un-sharp filter [38] is employed to reduce the impact of smoothness caused by the morphological closing operation. We applied the morphological closing operator as follows:
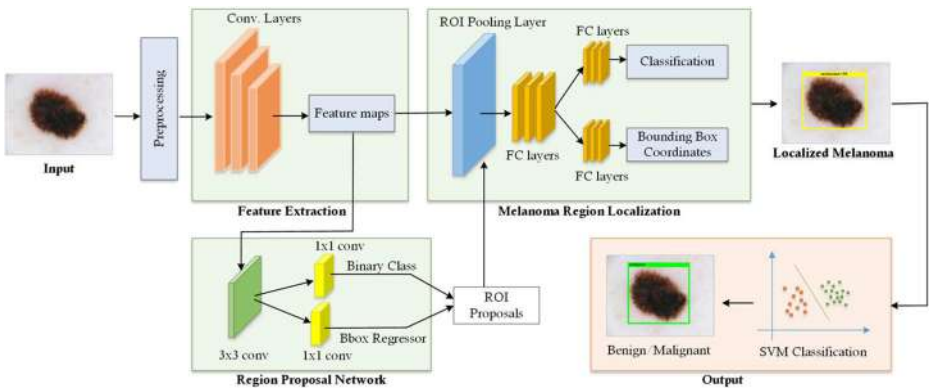
Fig. 1 Architecture of proposed method

$$I_m(x,y) = (I(x,y) \oplus S) \ominus S \qquad (1)$$

where $I(x,y)$ shows the input image, x and y are pixels location, whereas, S is the structuring elements of size 10 with square shape along the angle of 90^0 and 180^0 against every pixel. The $I_m(x,y)$ represents the processed image, without hair or tiny blood vessels. Although, this operation effectively extracts the skin mole region without hair or tiny blood vessels, however, the closing operation resulted in blurriness in the processed image. To reduce the blurring effect from the image, we applied the unsharp filter as follows:

$$I_{us}(x,y) = I_m(x,y) \times \varpi(x,y) \qquad (2)$$

Where

$$\varpi(x,y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (3)$$

The final processed image $I_f(x,y)$ is obtained using Eq. (4) that only represents the necessary information required for melanoma lesion classification.

$$I_f(x,y) = I(x,y) - I_{us}(x,y) \qquad (4)$$

After the image enhancement, the processed images are resized to 227×227 resolution using bicubic interpolation to minimize the computational complexity of the proposed approach. Finally, the augmentation step is employed to increase the image samples to eight times the current images (900 samples) in the dataset. For this purpose, the resized images are rotated at the angles of 0°, 90°, 180°, and 270° degrees and flipped horizontally.

3.2 Features extraction using faster-RCNN

Effective feature extraction is mandatory for the accurate classification of melanoma lesions into various classes. However, finding an effective feature descriptor is a complex task due to the following reasons: such as using a large features-set, the model may learn some incoherent characteristics i.e., noise, illumination changes, etc. While using a small features-set, the model may skip some important attributes i.e., texture and chrominance variations which make melanoma indistinguishable from the normal skin. To obtain the discriminative and efficient

set of features, it is necessary to select an automated features extraction technique without requiring to employ the hand-coded key-features extraction approaches. The approaches based on handcrafted features are unable to efficiently detect melanoma moles because of variations in size, texture, and color of the melanoma moles. To overcome the limitations of hand-crafted features, we employed a Faster-RCNN deep-learning model because it obtains the set of deep features learned directly from the images. The convolution filters of Faster-RCNN compute the key-points of the input image effectively by examining the image structure. The motivation of employing the Faster-RCNN over the RCNN [23] and Fast-RCNN [22] techniques for lesion classification is the dependency of RCNN and Fast-RCNN on the input generic object proposals, which is computed through hand-coded techniques like selective search [17] or EdgeBox [50], etc. Thus, RCNN and Fast RCNN approaches are computationally more complex as compared to Faster-RCNN.

The Faster-CNN, better addresses the limitations of RCNN and Fast-RCNN, comprises of two parts: i) Regional Proposal Network (RPN), and ii) Fast-RCNN. The RPN being the fully convolutional module can automatically generate the object proposal of the input image, which is then passed as input and also refined by the Fast-RCNN module. Both modules share the same convolutional layer which allows the input image to pass through CNN only once to produce and refine its object proposal. For melanoma skin lesion classification, it is challenging to obtain the features of interest amid two potential reasons: i) obtaining the actual position of multiple objects (i.e. lesion) from the input image, ii) class associated with each object (i.e. benign, malignant).

The Faster-RCNN technique can efficiently detect and classify the melanoma lesions of different classes by employing its fully convolutional modules RPN and Fast-RCNN which work by replacing the selective search algorithm. From the input image, the main attributes of melanoma moles such as shape, color, size, and texture are used by the RPN module to locate its position in the input image. RPN module works by employing fewer selected windows and achieves higher recall rates, which helps to reduce the features computation cost of the proposed framework.

The Faster R-CNN approach works by performing the following steps:

Convolution layers Faster-RCNN is a fully convolutional layers network consisting of a total of 13 convolutional and ReLU layers along with 4 pooling layers. These convolutional layers help the Faster-RCNN network to calculate the feature map of the input image. The calculated feature map is later shared with the RPN module and related layers.

RPN This step involves the generation of the input object proposal. The RPN module consists of 3×3 fully convolutional layer network, which is used to create the anchors and bounding box regression offsets. This module employs the softmax function to determine whether the computed anchors are part of the foreground or background. Finally, the produced anchors and bounding boxes are employed to compute the object proposals.

RoI pooling This layer works by employing the computed feature map from convolutional layers and proposals from the RPN module to generate the proposal feature maps and share them with all associated layers of the network.

Classification Finally, the classification step is performed to determine the class of the detected objects (skin lesion). It works by using the output of the Roi pooling layer. The bounding box regression is used to exhibit the resultant location of the detected test box.

The convolution layers perform the mapping between the input and output by using Eq. (5).

$$x_{j+1} = I\left(x_j, w_j\right) \tag{5}$$

Here x_j and x_{j+1} are the input and output of the jth layer, respectively, w_j shows the value of weights and biases linked with the jth layer. $I(x_j, w_j)$ shows a dot product between the weights and input regions. Next, the activation function performs transformation by adding the weighted input from the node into the activation of the node or output for that input. Here, we have used the ReLU layer as an activation function which performs pixel-wise (x) activation by using the following equation:

$$I(x) = max\ (0, x) \tag{6}$$

Where

$$I(x) = \begin{cases} 0 & if\ x < 0 \\ x & else \end{cases} \tag{7}$$

Additionally, max-pooling layers are used to perform downsampling operations to reduce the spatial size by using Eq. (8).

$$I(X) = \max(x_1, x_2, x_3, \dots x_n)\ \ where\ x_n \in X \tag{8}$$

Here, $I(X)$ is presenting the optimized feature vector. A detailed description of the employed Faster-RCNN network is given in Table 1. The main workflow of melanoma lesion localization through Faster-RCNN can be categorized into four steps. Firstly, the suspected image is passed to the convolutional layer to extract the feature map, secondly, the computed features are passed as input to the RPN module to acquire the keypoints information of the region proposals. Next, the ROI pooling layer generates the proposal feature maps by employing the computed feature map from convolutional layers and proposals from the RPN module. Finally, the bounding box regression is used to exhibit the resultant location of the detected melanoma lesion.

3.3 Training parameters

To minimize the melanoma classification error, we have used the stochastic gradient descent approach for weight optimization. In the training process, we have employed mini-batch size δ =128, learning rate $\alpha = 0.001$ along with 130 learning drop factor β with the value of 0.1 after performing the thorough investigation. During training, the value of the learning rate is automatically adjusted by using the piecewise learning schedule. To eliminate the lesion's localization error and to have an optimized cost function, we have used 100 epochs for the training phase, which employs that we have repeated Faster-RCNN training 100 times to achieve robust localization results. Furthermore, during the testing phase, to accurately localize the melanoma moles, Faster-RCNN uses greedy overlapping criteria for ground truth box and predicted box called intersection-over-union (IoU) with a value of 0.5. The deep keypoints are fed into the softmax layer of Faster-RCNN to calculate the confidence scores in which only the scored proposals with IoU values greater than 0.5 are selected. Furthermore, a Tensorflow Faster-RCNN is used to determine the number of batch cycles required to achieve an acceptable loss value by utilizing the TensorBoard tool. Fig. 2 presented the graphical view of the training loss graph of our work. It can be visualized from Fig. 2 that the training loss

Table 1 The faster-RCNN architecture

Sr. No	Layers	Filter	Size	Stride	Padding
Feature Extraction Module					
1	Input				
2	1_convolutional layer (ReLU)	64	3×3	1×1	1×1
3	2_convolutional layer (ReLU)	64	3×3	1×1	1×1
4	Max pooling	1	2×2	2×2	0×0
5	3_convolutional layer (ReLU)	128	3×3	1×1	1×1
6	4_convolutional layer (ReLU)	128	3×3	1×1	1×1
7	Max pooling		2×2	2×2	0×0
8	5_convolutional layer (ReLU)	256	3×3	1×1	1×1
9	6_convolutional layer (ReLU)	256	3×3	1×1	1×1
10	7_convolutional layer (ReLU)	256	3×3	1×1	1×1
11	Max pooling		2×2	2×2	0×0
12	8_convolutional layer (ReLU)	512	3×3	1×1	1×1
13	9_convolutional layer (ReLU)	512	3×3	1×1	1×1
14	10_convolutional layer (ReLU)	512	3×3	1×1	1×1
15	Max pooling		2×2	2×2	0×0
16	11_convolutional layer (ReLU)	512	3×3	1×1	1×1
17	12_convolutional layer (ReLU)	512	3×3	1×1	1×1
18	13_convolutional layer (ReLU)	512	3×3	1×1	1×1
RPN Module					
19	[13_convolutionallayer]Input				
20	convolutional layer	512	3×3	1×1	1×1
21	Fully connected layer				
22	Softmax classification	18	1×1	1×1	0×0
23	Regression	36	1×1	1×1	0×0
ROI pooling					
24	[13_convolutional layer + RPN proposal region] input				
25	ROI pooling layer				
Bounding Box generation					
26	[ROI proposed feature map] input				
27	1_Fully connected layer(ReLU)				
28	2_Fully connected layer(ReLU)				
29	Softmax classification				
30	Regression				

reached an optimal value of 0.0021 at the batch size of 3000, which is indicating the robust learning of our model.

3.4 Classification using SVM

SVM [16] is a mathematical framework that constructs hyperplanes to find the decision boundary for classification. SVM can effectively handle the curse of dimensionality as compared to other traditional approaches [18, 29] and reduces the occurrence of empirical error along-with preserving the complexity level of the mapping function. These properties of SVM enable it to better generalize its prediction behavior and perform well for a new data sample. The main reason to select SVM for melanoma classification is its robustness and ability to handle the over-fitted training data.

After extraction of the deep features, we used these features to train the SVM to classify the melanoma into two classes, i.e., benign and malignant. The training data consists of N melanoma feature vectors prepared as: $(x^{(i)}, y^{(i)})$, $i = 1,....N$, where $y^{(i)} \in \{1, -1\}$ represents the benign and malignant melanoma classes. For each feature vector $x^{(i)}$, SVM draws a

TotalLoss
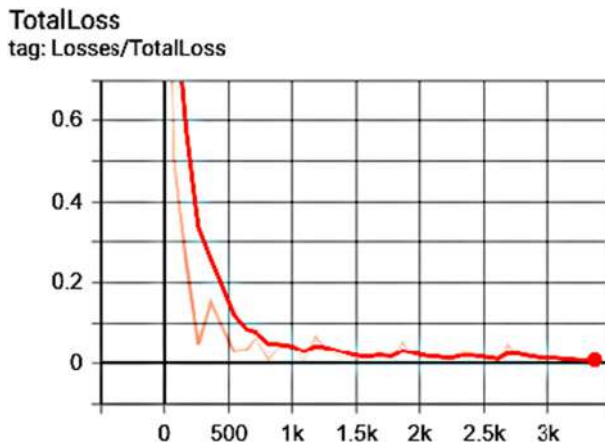tag: Losses/TotalLoss

Fig. 2 Training loss graph

hyperplane that linearly separates the two classes as:

$$w^T.x^{(i)} + \beta \geq 1 \text{ if } y^{(i)} = +1 \tag{9}$$

$$w^T.x^{(i)} + \beta < 1 \text{ if } y^{(i)} = -1 \tag{10}$$

where w is the weight vector and β is the bias. The objective is to maximize the distance between two support vectors by minimizing the norm $||w||$ which can be defined as a quadratic optimization problem as shown in Eq. (11):

$$\min ||w||, \text{ such that } y^{(i)}\left(w^T.x^{(i)} + \beta\right) \geq 1 \tag{11}$$

The two melanoma classes (benign and malignant) can be identified by applying the discriminant function $f(x) = \text{sign } (w^T. x^{(i)} + \beta)$ as follows:

$$\begin{cases} \text{malignant}, f\left(x^{(i)}\right) = +1, \\ \text{benign}, f\left(x^{(i)}\right) = -1 \end{cases} \tag{12}$$

4 Results

This section gives a thorough description of the results obtained after evaluating the performance of the proposed method. Additionally, the details of the dataset are also provided in this section.

4.1 Dataset

The evaluation of the presented technique is performed on the standard dataset named International Skin Imaging Collaboration (ISIC) by "International Symposium on biomedical

images (ISBI) in the challenge of skin lesion analysis towards melanoma detection 2016" [25]. This dataset consists of a total of 1279 samples belonging to two classes of melanoma i.e. malignant and benign. The training collection consists of a total of 900 training images among which 173 images belong to the malignant class and the rest of 727 images belong to the benign class. Whereas, the test collection contains 379 images in which 75 images belong to the malignant class and the remaining 304 images to the benign class. All images in the ISIC-2016 dataset are analyzed by a panel of dermatologists. ISIC-2016 dataset is a pathology-based database for melanoma classification where its images contain various artifacts i.e. variations in lesion size, color, texture, and illumination conditions, presence of hair, and tiny blood vessels, etc., which makes it a challenging dataset for melanoma classification. Shown in Fig. 3 are the sample images of the ISIC-2016 dataset.

4.2 Evaluation metrics

We employed the sensitivity (*sen*), specificity (*spe*), accuracy (*acc*), Mean Average Precision (mAP), and Intersection over union (IOU) metrics to evaluate the results of the introduced technique. We computed the sensitivity, specificity, accuracy, mAP, and IOU metrics as follows:

$$sen = \frac{tp}{tp + fn} \tag{13}$$

$$spe = \frac{tp}{tp + fp} \tag{14}$$

$$acc = \frac{tp + tn}{tp + fp + tn + fn} \tag{15}$$

Fig. 3 Sample images of ISIC-2016 dataset

$$mAP = mean \frac{tp}{tp + fp} \qquad (16)$$

$$IOU = \frac{tp}{tp + fn + fp} \times 2 \qquad (17)$$

where *tp, tn, fp,* and *fn* are representing the true positive, true negative, false positive, and false-negative cases, respectively.

4.3 Melanoma localization

This section provides the discussion of the experiment used to evaluate the performance of the melanoma localization. For this experiment, we used all samples from the ISIC-2016 dataset and reported the visual results of 30 images. Because of the unstable distribution of data samples in the database, we have employed the data augmentation approach already explained in Section 3.1. The data augmentation step improves the diversity of data samples and overcomes the risk of overfitting. Based on this technique, the proposed approach performs the evaluation analysis for all the test samples of the database. Moreover, the localization power of the Faster-RCNN enables it to accurately detect and differentiate the melanoma moles from the skin portion.

We have used the boxplot to show the localization results in terms of mAP and mean IOU as shown in Fig. 4. The regression layer of Faster-RCNN localized the melanoma moles with better mean mAP and IOU. More specifically, we achieved the average mAP and mean IOU of 0.902 and 0.932 respectively as also shown in the red line inside the box (Fig. 4). The qualitative results are shown in Fig. 5. We can see from the resultant images that the presented framework can precisely localize the melanoma moles even under the occurrence of skin marks, hairs, blood vessels, and clinical swatches. Moreover, our approach can accurately diagnose melanoma lesions of varying sizes and orientations.
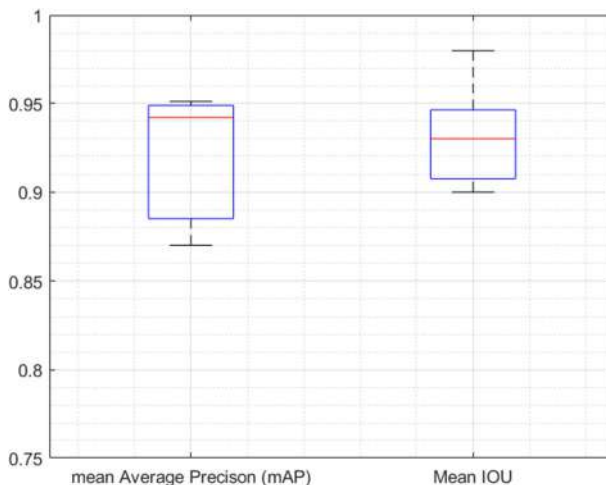

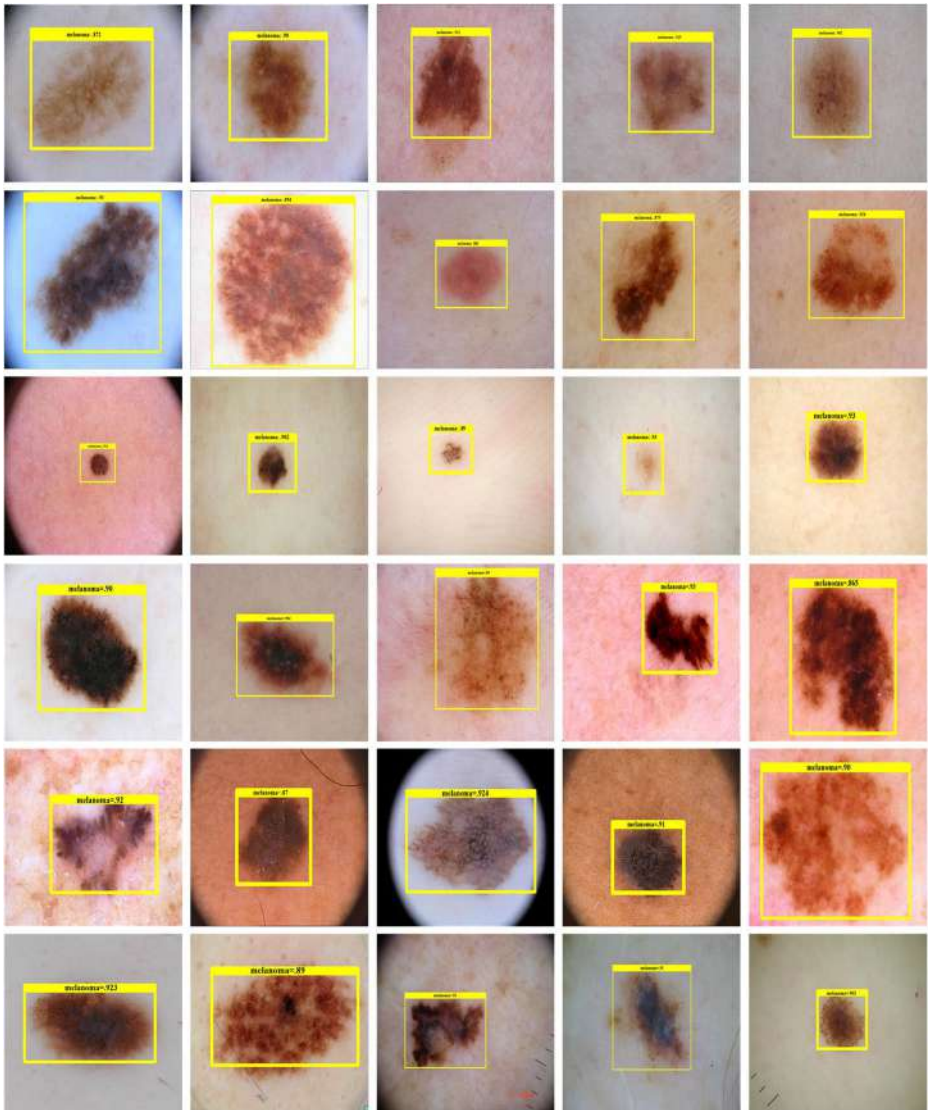**Fig. 4** Melanoma localization results

Fig. 5 Localization of melanoma regions

Although our approach is robust in localizing the skin lesions under the presence of noise, hair, tiny blood vessels, contrast, and chrominance variations, however, still there are some cases in which it may not work accurately. In Fig. 6 we have reported the results on which our method fails to detect the melanoma lesion. The false detection is due to intense variations in light intensity which results in extreme matching between the skin color and melanoma lesions.

4.4 Melanoma classification

In this part, we have discussed the classification results to show the effectiveness of the presented solution for classifying the melanoma moles. To evaluate the classification

Fig. 6 Sample images of inaccurately localized melanoma lesions

performance of our work, several experiments have been performed over the test images of ISIC-2016 datasets. We have employed an SVM classifier for this purpose. We have increased the size of the training database through the data augmentation step as discussed earlier. So, more than 7000 images localized through Faster-RCNN have been used to train the SVM classifier. The ability of the SVM classifier to deal with over-fitted training data has enabled it to better classify the melanoma lesions. Classification results through the SVM classifier are shown in Fig. 7, which have input images, localization of melanoma regions, and then classification results of the localized region into benign and malignant. The third column of the figure exhibits the benign cases and the last column presents the malignant class images. It can be seen from the reported results that the SVM classifier has accurately classified benign and malignant cases. Table 2 demonstrates the class-wise evaluation results of the presented solution in terms of accuracy, sensitivity, and specificity. Our method has achieved an average accuracy of 0.891 along with the average values of 0.859 and 0.870 for sensitivity and specificity respectively that are showing the proficiency of the introduced framework. Moreover, the confusion matrix is reported in Fig. 8.
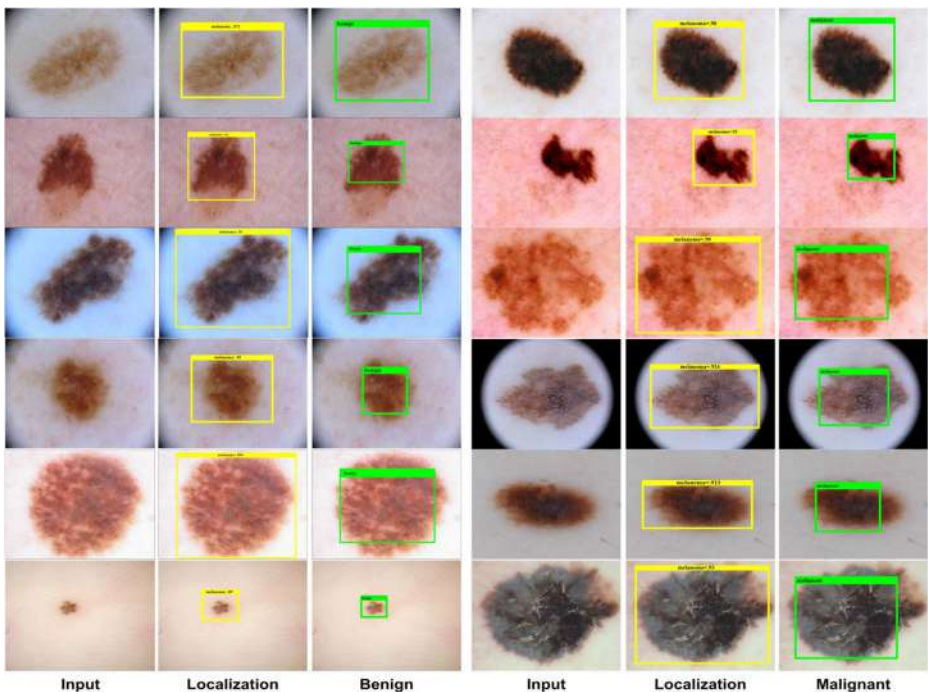

Fig. 7 Classification results

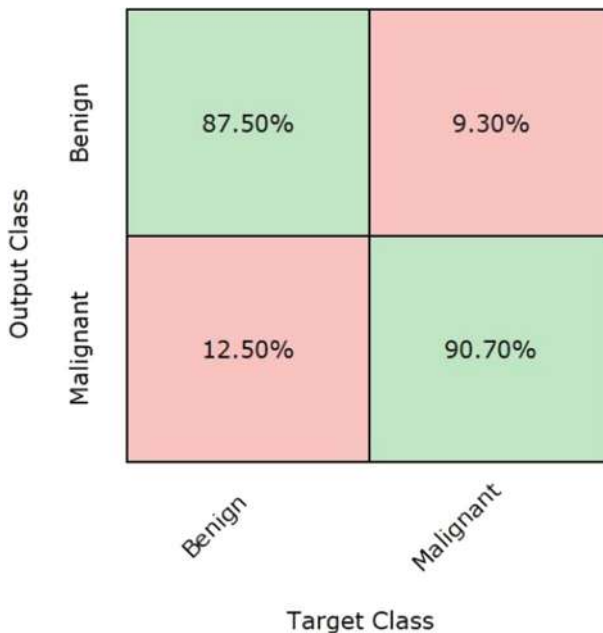**Table 2** Stage wise performance of the proposed method

Stages	Acc	Sen	Spe
Benign	0.8750	0.845	0.851
Malignant	0.9070	0.874	0.890
Average	0.891	0.859	0.870

4.5 Comparative analysis

We have performed an experimental analysis to compare the proposed method against the work of the top eight teams [25] of ISBI challenge-2016 and results are reported in Fig. 9. The teams are ranked according to their obtained specificity score. It can be witnessed from Fig. 9 that our approach attains the highest specificity and average AUC values. More specifically, we achieved the specificity rate of 0.870 and the average AUC value of 0.843, which is greater than the specificity and AUC values achieved by all the works [50]. However, these networks are computationally more complex than our approach, as they have deeper network architectures. From this comparative analysis, we can conclude that the presented methodology exhibits effective and efficient classification performance due to the accurate localization of melanoma lesions using Faster-RCNN.

4.6 Comparison with base models

We have also evaluated the performance of our approach with existing latest techniques like DenseNet-201 [2], ResNet-50 [2], Inception-v3 [2] and Inception-ResNet-v2 [2]. The results are reported in Table 3 and Fig. 10. The presented framework achieves the average specificity

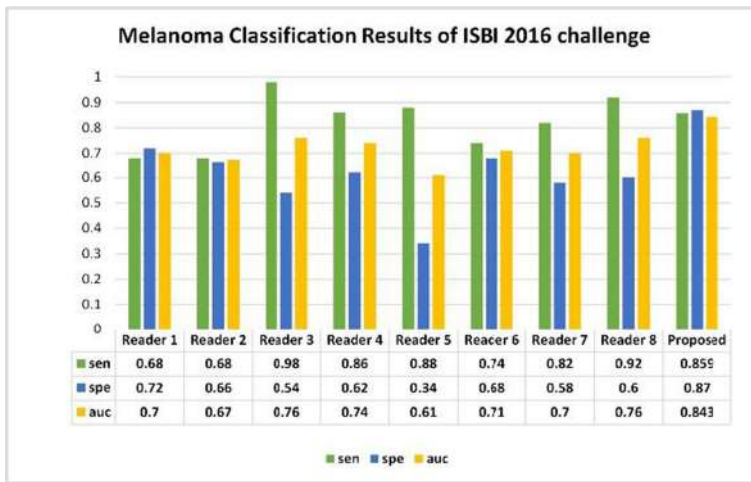**Fig. 8** Sample images of inaccurately localized melanoma lesions

Fig. 9 A comparison of classification results with other technique

score of 0.87, whereas the comparative methods namely DenseNet-201, ResNet-50, Inception-v3, and Inception-ResNet-v2 achieved the specificity score of 0.652, 0.679, 0.662, and 0.714 respectively. The average sensitivity rate of our method is 0.859 while the other methods like DenseNet-201, ResNet-50, Inception-v3, and Inception-ResNet-v2 achieved the sensitivity rate of 0.812, 0.799, 0.770, and 0.818 respectively. Similarly, the average AUC value of our approach is 0.843, while the AUC value of DenseNet-201, ResNet-50, Inception-v3, and Inception-ResNet-v2 frameworks are 0.736, 0.739, 0.716, and 0.766 respectively. From the results, it can be visualized that our method achieved higher results as compared to comparative techniques in terms of AUC, sensitivity, and specificity. The comparative methods employed very dense layered deep networks that can easily encounter the problem of over-fitting. As in our technique, the SVM classifier can deal with an over-fitted model, therefore we can say that our method is more robust in terms of pattern recognition for melanoma classification. Furthermore, the presented framework has potential and applicable for future medical practices, as it requires a minimum time (.004 s) to process a dermoscopic image which is less than from all other approaches (Table 3).

4.7 Performance comparison with state-of-the-art techniques

This experiment has been designed to evaluate the performance of our approach for melanoma lesion classification against the latest state-of-the-art methods. For classification, we have compared our framework against [51, 55–57] techniques and run on Nvidia GTX1070 GPU-

Table 3 Comparative results

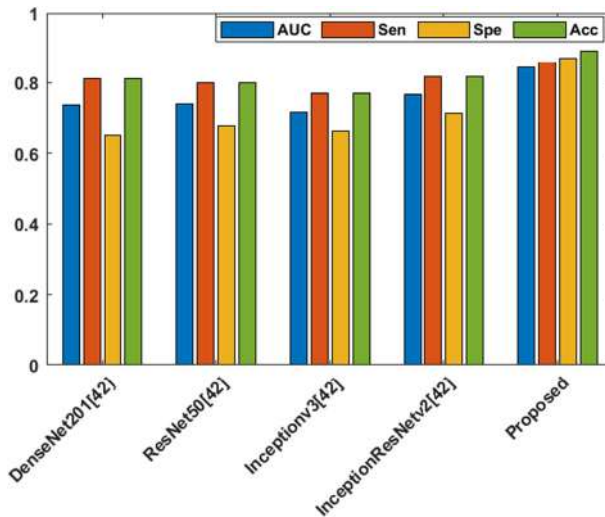| Approach	AUC	Sen	Spe	Acc	Test Time (s)
DenseNet-201 [2]	0.736	0.812	0.652	0.812	0.012
ResNet-50 [2]	0.739	0.799	0.679	0.799	0.005
Inception-v3 [2]	0.716	0.770	0.662	0.770	0.005
Inception-ResNet-v2 [2]	0.766	0.818	0.714	0.817	0.010
Proposed	0.843	0.859	0.870	0.891	0.004

Fig. 10 Graphical representation of comparative results

based system. The comparative results in terms of acc, mAP, and AUC are reported in Fig. 11. From the results in Fig. 11, we can visualize that our approach achieved an accuracy of 89.1%, mAP of 90.2%, and AUC of 84.3%, which is the highest among all the comparative methods. The introduced framework attained the average AUC of 0.843 while the comparative approaches achieved the average AUC of 0.818, hence, our method achieved a 2.47% performance gain. Similarly, the mAP value of our method is 0.902 while the average mAP value of comparative approaches is 0.647, which depicts a 25.4% performance gain. In terms of average acc, our work achieved 0.891, whereas other approaches showed an average acc value of 0.849. Therefore, we can say that our method gave a 4.1% performance gain and more robust to skin lesion classification than the comparative approaches. The stable performance of the presented framework is effective towards the low-resolution feature map generation based on the region proposals, while the methods in [51, 55–57] are applied directly on entire images which result in misclassification due to the presence of complex artifacts (i.e. hair, tiny blood vessels, noise, etc.). However, our method uses a region proposals network that contributes to the precise localization of the melanoma moles even for the input samples suffering from the artifacts. Hence, from the results, we can say that our approach is more accurate for skin lesion classification as compared to the other methods used for the comparison (Fig. 11).

4.8 Cross dataset-validation

We finalized the evaluation analysis of our framework by applying the cross-dataset validation. By using the cross-dataset validation, we can estimate the robustness of the presented method in terms of dealing with the training and testing complexities and show its suitability for real-world scenarios. We have used the ISIC-2017 dataset [15] to perform the cross dataset-validation process. ISIC-2017 database contains a total of 2750 samples, in which 517 images are of malignant category, while the remaining 2223 samples are of benign category. The images of ISIC-2017 are diverse in terms of angles, lighting conditions, and different artifacts like the presence of dark corners, skin hair, and ruler markers, etc. To accomplish the cross-dataset validation task, we took the following scenarios: (a) training over the ISIC-2016 dataset
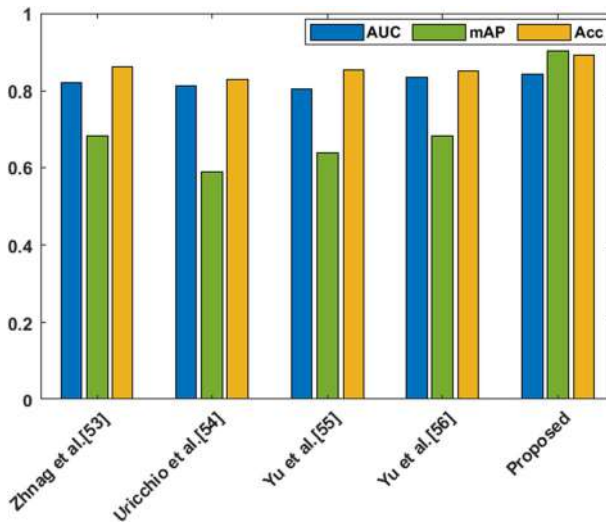
Fig. 11 Comparison of presented model with latest techniques

and perform testing over the ISIC-2017 database (b) training over the ISIC-2017 dataset, and perform testing over the ISIC-2016 database. Figure 12 presented the classified visual results for the ISIC-2017 dataset. In the cross-dataset validation process, our work attained an average accuracy of 0.9125%, 0.9122% for training, and 0.8941%, 0.9070% accuracy on test sets over
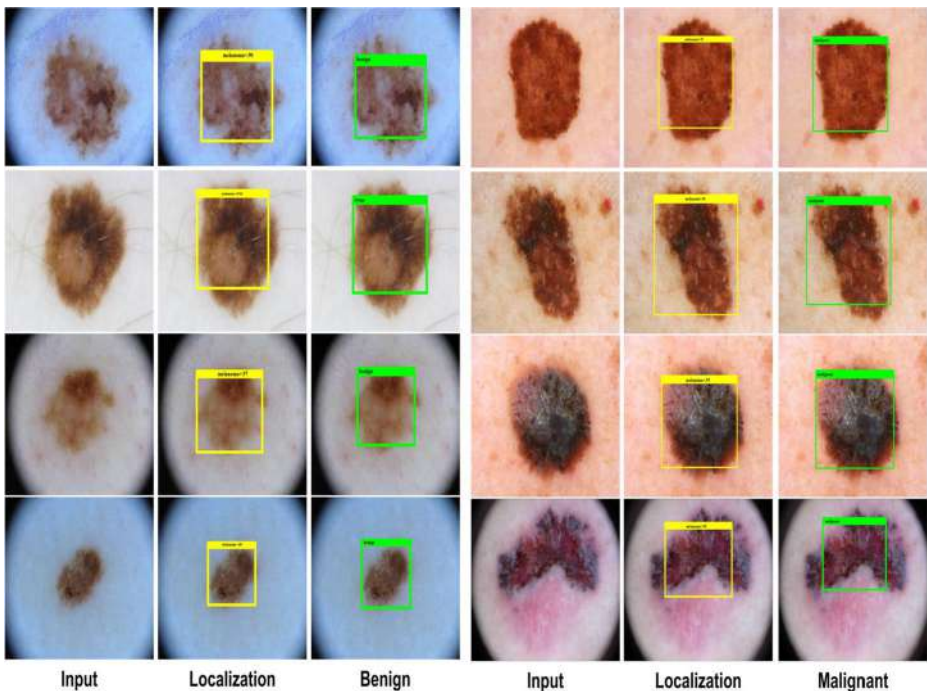

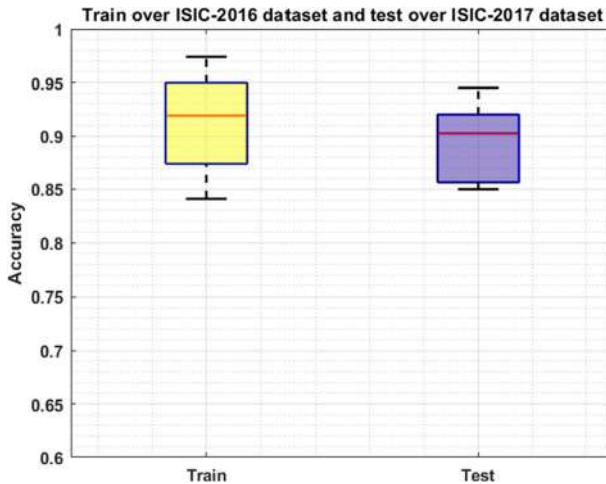**Fig. 12** Classification results over ISIC-2017 dataset

Fig. 13 Cross validation over ISIC-2017 dataset

the ISIC-2017 and ISIC-2016 databases respectively (Figs. 13 and 14). Therefore, through cross-database validation, we can conclude that our approach can be utilized in real-world scenarios to deal with any condition of melanoma lesion to assist the dermatologists.

5 Conclusion and future work

This paper has presented a novel approach for the automated classification of melanoma lesions by employing the Faster-RCNN deep-learning technique along with the SVM classifier. In the presented work, we have also introduced the application of Faster-RCNN for melanoma lesion classification. More specifically, we employed the Faster-RCNN for deep feature extraction and melanoma detection. Finally, we used the deep features to train the SVM for melanoma
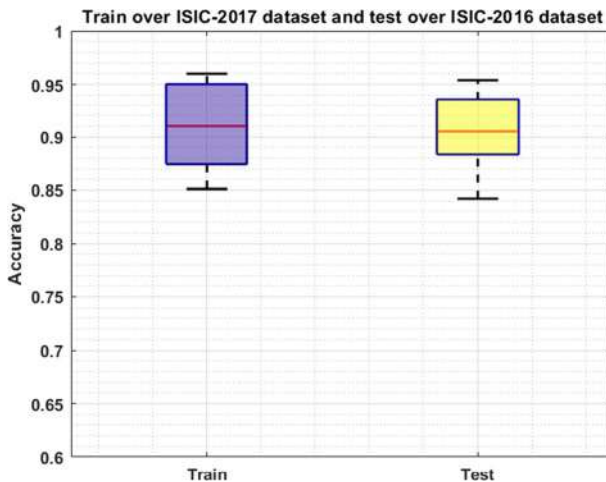

Fig. 14 Cross validation over ISIC-2016 dataset

classification. The proposed method effectively localize the melanoma segment from the input image and classifies the melanoma region into benign and malignant classes. Our approach is robust to various artifacts i.e. noise, blurring, chrominance changes, light variations, melanoma size, and presence of hair and tiny blood vessels. Experimental results have confirmed that the presented framework outperforms the existing latest techniques. In the future, we plan to extend our technique to apply it to other skin diseases as well.

References

1. Alcón JF, Ciuhu C, Ten Kate W, Heinrich A, Uzunbajakava N, Krekels G, Siem D, de Haan G (2009) Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J Sel Top Signal Process 3(1):14–25
2. Al-Masni MA, Kim D-H, Kim T-S (2020) Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput Meth Prog Biomed 190:105351
3. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 35(5):1207–1216
4. Attia M, Hossny M, Nahavandi S, Yazdabadi A (2017) Skin melanoma segmentation using recurrent and convolutional neural networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp 292–296: IEEE
5. Badrinarayanan V, Handa A, Cipolla R (2015) Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling, arXiv preprint arXiv:.07293
6. Ballerini L, Fisher R B, Aldridge B, Rees J (2013) A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions. In: Color Medical Image Analysis: Springer, pp 63–86
7. Barata C, Ruela M, Francisco M, Mendonça T, Marques JS (2013) Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Syst J 8(3):965–979
8. Barata C, Celebi ME, Marques JS (2017) Development of a clinically oriented system for melanoma diagnosis. Pattern Recogn 69:270–285
9. Bi L, Kim J, Ahn E, Feng D, Fulham M (2017) Semi-automatic skin lesion segmentation via fully convolutional networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp 561–564: IEEE
10. Burdick J, Marques O, Romero-Lopez A, Giró Nieto X, Weinthal J (2017) The impact of segmentation on the accuracy and sensitivity of a melanoma classifier based on skin lesion images. In: SIIM 2017 scientific program: Pittsburgh, PA, June 1–June 3, 2017, David L. Lawrence Convention Center, pp 1–6
11. Cavalcanti PG, Scharcanski J (2011) Automated prescreening of pigmented skin lesions using standard cameras. Comput Med Imaging Graph 35(6):481–491
12. Cavalcanti PG, Scharcanski J, Baranoski GV (2013) A two-stage approach for discriminating melanocytic skin lesions using standard cameras. Expert Syst Appl 40(10):4054–4064
13. Cheng Y, Swamisai R, Umbaugh SE, Moss RH, Stoecker WV, Teegala S, Srinivasan SK (2008) Skin lesion classification using relative color features. Skin Res Technol 14(1):53–64
14. Codella N C, Nguyen Q-B, Pankanti S, Gutman D A, Helba B, Halpern A C, Smith J R (2017) Deep learning ensembles for melanoma recognition in dermoscopy images. IBM J Res Dev 61, no4/5, pp. 5: 1–5: 15
15. Codella N C, Gutman D, Celebi M E, Helba B, Marchetti M A, Dusza S W, Kalloo A, Liopyris K, Mishra N, Kittler H (2018) Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC), in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp 168–172: IEEE
16. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
17. Dollár P, Zitnick CL (2014) Fast edge detection using structured forests. IEEE Trans Pattern Anal Mach Intell 37(8):1558–1570
18. Fukunaga K, Narendra PM (1975) A branch and bound algorithm for computing k-nearest neighbors. IEEE Trans Comput 100(7):750–753
19. Ganster H, Pinz P, Rohrer R, Wildling E, Binder M, Kittler H (2001) Automated melanoma recognition. IEEE Trans Med Imaging 20(3):233–239
20. Garnavi R, Aldeen M, Celebi ME, Bhuiyan A, Dolianitis C, Varigos G (2010) Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels. Int J Med Med Sci 1(2):126–134

21. Giotis I, Molders N, Land S, Biehl M, Jonkman MF, Petkov N (2015) MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert Syst Appl 42(19):6578–6585
22. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448
23. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587
24. Gulati S, Bhogal R K (2019) Detection of Malignant Melanoma Using Deep Learning. In: International Conference on Advances in Computing and Data Sciences. Springer, pp 312–325
25. Gutman D, Codella N C, Celebi E, Helba B, Marchetti M, Mishra N, Halpern A (2016) Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC), arXiv preprint arXiv:.01397
26. Harangi B (2018) Skin lesion classification with ensembles of deep convolutional neural networks. J Biomed Inform 86:25–32
27. Hosny K M, Kassem M A, Foaud M M (2018) Skin Cancer Classification using Deep Learning and Transfer Learning. In: 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), pp 90–93: IEEE
28. Hu K, Niu X, Liu S, Zhang Y, Cao C, Xiao F, Yang W, Gao X (2019) Classification of melanoma based on feature similarity measurement for codebook learning in the bag-of-features model. Biomed Signal Process Control 51:200–209
29. Lewis DD (1998) Naive (Bayes) at forty: The independence assumption in information retrieval, in European conference on machine learning. Springer, pp 4–15
30. Li Y, Shen L (2018) Skin lesion analysis towards melanoma detection using deep learning network. Sensors 18(2):556
31. Liao P-S, Chen T-S, Chung P-C (2001) A fast algorithm for multilevel thresholding. J Inf Sci Eng 17(5): 713–727
32. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3431–3440
33. Mahbod A, Schaefer G, Wang C, Ecker R, Ellinge I (2019) Skin lesion classification using hybrid deep neural networks, in ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 1229–1233: IEEE
34. Nachbar F, Stolz W, Merkle T, Cognetta AB, Vogt T, Landthaler M, Bilek P, Braun-Falco O, Plewig G (1994) The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. J Am Acad Dermatol 30(4):551–559
35. Nida N, Irtaza A, Javed A, Yousaf MH, Mahmood MT (2019) Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. Int J Med Inform 124: 37–48
36. Nijeweme-d'Hollosy WO, van Velsen L, Poel M, Groothuis-Oudshoorn CG, Soer R, Hermens H (2018) Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. Int J Med Inform 110:31–41
37. Okur E, Turkan M (2018) A survey on automated melanoma detection. Eng Appl Artif Intell 73:50–67
38. Polesel A, Ramponi G, Mathews VJ (2000) Image enhancement via adaptive unsharp masking. IEEE Trans Image Process 9(3):505–510
39. Rehman A, Khan MA, Mehmood Z, Saba T, Sardaraz M, Rashid M (2020) Microscopic melanoma detection and classification: a framework of pixel-based fusion and multilevel features reduction. Microsc Res Tech 83(4):410–423
40. Ridler T, Calvard S (1978) Picture thresholding using an iterative selection method. IEEE Trans Syst Man Cybern 8(8):630–632
41. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM (2015) Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA Dermatol 151(10):1081–1086
42. Schaefer G, Krawczyk B, Celebi ME, Iyatomi H (2014) An ensemble classification approach for melanoma diagnosis. Memet Comput 6(4):233–240
43. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, Jemal A (2017) Colorectal cancer statistics. CA: a Cancer Journal for Clinicians 67(3):177–193
44. Silveira M, Nascimento JC, Marques JS, Marçal AR, Mendonça T, Yamauchi S, Maeda J, Rozeira J (2009) Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. IEEE J Sel Top Signal Process 3(1):35–45
45. Singh S, Alam M, Singh B (2020) Orthogonal moment feature extraction and classification of melanoma images. Journal of Information Optimization Sciences 41(1):195–203
46. Society A C (2016) Cancer facts & figures. American Cancer Society

47. Stanley RJ, Stoecker WV, Moss RH (2007) A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images. Skin Res Technol 13(1):62–72
48. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging 35(5): 1299–1312
49. Tan TY, Zhang L, Lim CP (2020) Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks. Knowl-Based Syst 187:104807
50. Uijlings JR, Van De Sande KE, Gevers T, Smeulders AW (2013) Selective search for object recognition. Int J Comput Vis 104(2):154–171
51. Uricchio T, Bertini M, Seidenari L, Bimbo A (2015) Fisher encoded convolutional bag-of-windows for efficient image retrieval and social image tagging. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp 9–15
52. Walter M (2016) Is this the end? machine learning and 2 other threats to radiologys future, goo. gIIM9X3SF, pp l3
53. 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
54. Yap J, Yolland W, Tschandl P (2018) Multimodal skin lesion classification using deep learning. Exp Dermatol 27(11):1261–1267
55. Yu L, Chen H, Dou Q, Qin J, Heng P-A (2016) Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging 36(4):994–1004
56. Yu Z, Jiang F, Zhou F, He X, Ni D, Chen S, Wang T, Lei B (2020) Convolutional descriptors aggregation via cross-net for skin lesion recognition. Appl Soft Comput:106281
57. Zhang J, Xie Y, Wu Q, Xia Y (2019) Medical image classification using synergic deep learning. Med Image Anal 54:10–19