

Research Article

A Novel Region-Extreme Convolutional Neural Network for Melanoma Malignancy Recognition

Nudrat Nida (),¹ Aun Irtaza (),² and Muhammad Haroon Yousaf (),^{1,3}

¹Department of Computer Engineering, University of Engineering and Technology-Taxila, Taxila, Pakistan ²Department of Computer Science, University of Engineering and Technology-Taxila, Taxila, Pakistan ³Swarm Robotics Lab, National Centre of Robotics and Automation, Islamabad, Pakistan

Correspondence should be addressed to Muhammad Haroon Yousaf; haroon.yousaf@uettaxila.edu.pk

Received 1 January 2021; Accepted 5 July 2021; Published 13 July 2021

Academic Editor: Erivelton Geraldo Nepomuceno

Copyright © 2021 Nudrat Nida et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Melanoma malignancy recognition is a challenging task due to the existence of intraclass similarity, natural or clinical artefacts, skin contrast variation, and higher visual similarity among the normal or melanoma-affected skin. To overcome these problems, we propose a novel solution by leveraging "region-extreme convolutional neural network" for melanoma malignancy recognition as malignant or benign. Recent works on melanoma malignancy recognition employed the traditional machine learning techniques based on various handcrafted features or the recently introduced CNN network. However, the efficient training of these models is possible, if they localize the melanoma affected region and learn high-level feature representation from melanoma lesion to predict melanoma malignancy. In this paper, we incorporate this observation and propose a novel "region-extreme convolutional neural network" for melanoma malignancy recognition. Our proposed region-extreme convolutional neural network refines dermoscopy images to eliminate natural or clinical artefacts, localizes melanoma affected region, and defines precise boundary around the melanoma lesion. The defined melanoma lesion is used to generate deep feature maps for model learning using the extreme learning machine (ELM) classifier. The proposed model is evaluated on two challenge datasets (ISIC-2016 and ISIC-2017) and performs better than ISIC challenge winners. Our region-extreme convolutional neural network recognizes the melanoma malignancy 85% on ISIC-2016 and 93% on ISIC-2017 datasets. Our region-extreme convolutional neural network precisely segments the melanoma lesion with an average Jaccard index of 0.93 and Dice score of 0.94. The regionextreme convolutional neural network has several advantages: it eliminates the clinical and natural artefacts from dermoscopic images, precisely localizes and segments the melanoma lesion, and improves the melanoma malignancy recognition through feedforward model learning. The region-extreme convolutional neural network achieves significant performance improvement over existing methods that makes it adaptable for solving complex medical image analysis problems.

1. Introduction

Melanoma is a dangerous form of skin cancer and it is difficult to identify it at an earlier stage due to visual similarity with normal skin. The morality rate of melanoma is more than 75% every year [1]. In the year 2020, 76,380 new cases of melanoma are suspected to be diagnosed and around 10,130 deaths are reported in the USA [2]. Fortunately, melanoma is curable if it is diagnosed and medicated at an earlier stage. Therefore, earlier diagnosis is desirable to improve the patient's survival rate [3–5]. Dermoscopy images are used for the diagnosis of melanoma and it is reliable than a visual melanoma inspection procedure through the naked eye. However, the manual melanoma diagnosis through the naked eye is error-prone and time-consuming, and different dermatologists' diagnosis predictions often vary that adversely affect the earlier treatment of disease [6]. Therefore, the automatic diagnosis of melanoma facilitates the dermatologists and supports them to validate their predictions even at earlier stages.

Melanoma recognition is a challenging domain due to visual similarity among the types of melanoma such as color, size, texture, location, and shape. The boundaries are also not obscure, and variation makes the task of recognition more complicated at an earlier stage. Moreover, some artefacts such as hair, blood vessels, gel bubble, and clinical ruler marks also degrade the recognition performance [1]. To support the earlier diagnosis and to solve this challenging problem, computer vision-based computer-aided diagnosis (CAD) tools are capable of detecting and recognizing melanoma to assist the dermatologists [7]. Recently, much effort has been invested in the recognition of melanoma automatically [8-14]. Earlier studies employ the low-level handcrafted features to classify normal and melanoma skin regions, including texture [15, 16], shape [17], and color [17, 18] attributes. Some traditional investigations suggest combining handcrafted features along with selection of features to improve the recognition of melanoma [19, 20]. However, the handcrafted features hold low-level representation of melanoma which is not effective to inculcate higher visual similarity among normal and melanoma lesions. Moreover, the artefacts within the dermoscopic images lead to a higher misclassification rate. The researchers also used the segmentation of melanoma region and extracted features from the segmented region to recognize melanoma [16, 19-21]. The feature representation from segmented melanoma region results in better recognition due to a well-defined area of interest. Still, these techniques employ low-level handcrafted feature representation for segmentation and classification with limited recognition capabilities.

Recently, the deep learning techniques have revolutionized the biomedical image analysis domain, including segmentation [22, 23], detection [24-27], and classification [28-30]. The multilayer high-dimensional CNN feature representations capture discriminative attributes for the accurate recognition of melanoma. Codella et al. [31] suggested deep features representation to classify melanoma using support vector machine (SVM). Kawahara et al. [32] presented the benefit of transfer learning of pretrained model including AlexNet [31] and tuned the model for melanoma representation. Yu et al. employed the deep residual network to classify benign and malignant melanoma with a deeper model of CNN [33]. Zhen et al. introduced the features extraction from pretrained ResNet50 and local descriptor fisher vector encoding scheme to train the SVM model for melanoma malignancy recognition [34]. Melanoma malignancy recognition from deep learning techniques indicated good performance as compared to handcrafted features due to high-level feature representation of melanoma.

This research work aims to propose a novel approach for melanoma malignancy recognition and to overcome all the challenges present in the dermoscopic images (Figure 1). The proposed "region-extreme convolutional neural network" differs from the existing melanoma recognition methods in the following ways. First, while existing CNN approaches for melanoma recognition make deterministic diagnosis decision based on features extracted from the entire dermoscopy image, the proposed "region-extreme convolutional neural network" improves the recognition performance by introducing the localization block that allows precise melanoma segmentation and removes artefacts. Then, decision boundaries are established based on these segmented lesions to predict the melanoma malignancy. The malignancy model generated from melanoma lesion plays an important role in accurate prediction as it improves the melanoma malignancy recognition performance through considering only melanoma affected lesion unlike entire dermoscopy image. Therefore, the model obtained from the proposed method improves the recognition performance as compared to existing methods. The proposed approach is evaluated on the ISIC-2016 and ISIC-2017 challenge datasets and experimental finding portrays good recognition performance.

Our contributions in this paper are summarized as follows:

- (1) We propose region-extreme convolutional neural network for localization and recognition of melanoma malignancy at an earlier stage
- (2) Our experimental finding reveals that deep features extracted from the penultimate layers are more powerful than the features extracted from higherorder layers of CNN because the receptive fields of penultimate layers hold higher semantics and contextual information of melanoma malignancy
- (3) Rigorous experimentation against the state-of-theart methods is performed to establish the effectiveness of the proposed approach

2. The Proposed Region-Extreme Convolutional Neural Network for Melanoma Malignancy Recognition

Dermoscopic images used for the recognition of melanoma are occluded with clinical and natural artefacts. Moreover, a huge variation and intraclass visual similarity also exist among the melanoma and non-melanoma regions that adversely degrade the recognition performance [1]. To overcome these challenges, we propose a novel scheme to mitigate these challenges and recognize the melanoma malignancy with higher recognition rate, as shown in Figure 2. In the first step, we refine the images and localize the melanoma affected region. Then, the localized region is used to segment the melanoma region. The localization and segmentation block use the training pair of images including RGB dermoscopic image and corresponding ground truth mask images to generate melanoma localization model during training phase. The location coordinates of melanoma are obtained through binarization of the ground truth image. After binarization, we apply the region split and merged segmentation on binary ground truth image to acquire the location information (x, y, w, h) of melanoma by extracting the largest connected component. These coordinates' information from ground truth image I_G is used to map the location of melanoma in the RGB dermoscopy image I. These coordinates (x, y, w, h) along with RGB dermoscopic images are presented to RCNN to generate a model for the melanoma localization. The RCNN employs selective search to produce region proposal, and deep convolutional



FIGURE 1: (a–e) Sample of ISIC-2016, and ISIC-2017 dataset illustrating the clinical artefacts and natural artefacts. These artefacts consist of (a) color and illumination variation, (b) variation in contrast, (c) black frame and clinical rule marks, (d) natural hair, and (e) clinical color swatches.



FIGURE 2: The architecture diagram of the proposed region-extreme convolutional neural network for malignancy recognition melanoma.

features are extracted from each proposal. The softmax layer classifies the test sample proposal into the melanoma affected or normal skin region, while the classification softmax layer recognizes the chosen proposals as melanoma-affected region, and the suspected proposals are passed through the regression layer of RCNN [35]. The regression layer computes the intersection over union score (IoU) using greedy suppression algorithm to detect the part of skin as melanoma or normal region. In our case, if the region's regression score is higher than 0.5, then the proposal is localized as melanoma lesion to reduce the false positive localization. Therefore, only true positive proposals with regression IoU scores greater than 0.5 are considered as melanoma area; otherwise, it is considered as normal skin area. In the third step, the segmented melanoma regions are used to extract deep features for the recognition purpose. The architecture of proposed regionextreme convolutional neural network is shown in Figure 2 and elaborated in detail in the following sections.

2.1. Mathematical Formulation of Melanoma Recognition. Given an input dermoscopy image D^i from the *i* training samples, the skin refinement block of region-extreme convolutional neural network enhances the perceptual quality of dermoscopic image and processed image *I* is processed further for the melanoma localization.

Training of the melanoma localization block requires predicted melanoma region I_P and corresponding ground truth location of melanoma region I_G . I_G is presented as pairs (I_P^i, I_G^i) where i = 1, 2, 3, ..., N are the training samples and N represents the total number of training samples. Here, I_G and I_P are the rectangular regions as (x, y, w, h) with pixel location (x, y), weight as w, and height as h. To localize the melanoma lesion precisely, optimal transformation functions are established by the region-extreme convolutional neural network to reduce the localization error and exactly map the predicted bounding box I_P location (x, y), width w, and height h of melanoma region over ground truth bounding box I_G . The four transformation functions $tx(I_P), ty(I_P), tw(I_P)$, and $th(I_P)$ are mathematically represented as

$$I_{G_{x}} = I_{P_{w}} t_{x} (I_{P}) + I_{P_{x}}.$$
 (1)

Here, I_{G_x} is the centroid x pixel location of actual melanoma lesion, I_{P_x} , I_{P_w} are horizontal and width-wise predicted melanoma lesions, and t_w is a logarithmic representation of targeted horizontal coordinate of melanoma lesion.

$$I_{G_{v}} = I_{P_{h}} t_{y} (I_{P}) + I_{P_{v}},$$
(2)

where I_{G_y} is the centroid y pixel location of actual melanoma lesion, I_{P_y} and I_{P_h} are vertical and height-wise predicted melanoma lesions, and t_y is a logarithmic representation of targeted vertical coordinate of melanoma lesion.

$$I_{G_{w}} = I_{P_{w}} e^{t_{w} (I_{P})}.$$
 (3)

Similarly, I_{G_w} is the width of the actual melanoma lesion, I_{P_w} and I_P represent the width-wise predicted melanoma lesion and overall predicted melanoma lesion, and t_w is a logarithmic representation of targeted width of melanoma lesion.

$$I_{G_{h}} = I_{P_{h}} e^{t_{h} (I_{P})}.$$
(4)

 I_{G_h} represents height, I_{P_h} and I_P represent the heightwise predicted melanoma lesion and overall predicted melanoma lesion, and t_h is a logarithmic representation of targeted height of melanoma lesion.

$$t_* I_p = \omega_*^T \phi_3 I_p. \tag{5}$$

Here, $t_*I_p\varepsilon[t_x(I_P), t_y(I_P), t_w(I_P), t_h(I_P)]$ is the target location of melanoma lesion to map predicted melanoma lesion I_P on actual melanoma lesion I_G . $t_x(I_P), t_y(I_P),$ $t_w(I_P)$, and $t_h(I_P)$ are transformation functions to linearly predict the region proposal I_P using deep features of pooling layer 3 (ϕ_3) where ϕ is 2D compressed receptive field. For precise prediction of melanoma lesion, ω_* are the weights adapted during training phase through gradient decent algorithm and t_* is respective mapping function for location (x, y), width w, and height h, where $*\varepsilon[x, y, w, h]$ and ω_* is 1D weight matrix. $t_x(I_P)$ and $t_y(I_P)$ specify mapping of I_P centroid keeping the characteristic of scale in-variance. $t_w(I_P)$ and $t_h(I_P)$ are a logarithmic representation of width and height of I_P .

The regression targets $R_* \varepsilon[R_x, R_y, R_w, R_h]$ for accurate localization of melanoma region are listed numerically as follows:

$$R_{x} = \frac{I_{G_{x}} - I_{P_{x}}}{I_{P_{x}}}.$$
 (6)

Here, R_x is the regression target function at horizontal axis. I_{G_x} is horizontal location of actual melanoma lesion. I_{P_x} , I_{P_w} represent the horizontal location and width-wise predicted melanoma lesions.

$$R_{y} = \frac{I_{G_{y}} - I_{P_{y}}}{I_{P_{h}}},$$
(7)

 R_y is the regression target function at vertical axis. I_{G_y} is vertical location of actual melanoma lesion. I_{P_y} , I_{P_h} represent the vertical location and height-wise predicted melanoma lesions.

$$R_w = \log \frac{I_{G_x}}{I_{P_w}},\tag{8}$$

 R_w is the regression target function for width mapping. I_{G_x} is horizontal location of actual melanoma lesion. I_{P_w} represents the width-wise predicted melanoma lesion.

$$R_h = \log \frac{I_{G_y}}{I_{P_h}},\tag{9}$$

 R_h is the regression target function for height mapping. I_{G_y} is vertical location of actual melanoma lesion. I_{P_h} represents the height-wise predicted melanoma lesion. For simplicity, we represent the regression target functions R_x , R_y , R_w , R_h as R(x, y, w, h).

Region-extreme convolutional neural network performs melanoma lesion localization through greedy overlapping criteria of ground truth bounding boxes and predicted boxes, known as intersection-over-union (IoU). The information extracted from the proposal I_p improves learning, if I_p is close to one of the ground truth boxes I_G [35]. The IoU threshold chosen for melanoma detection is greater than 0.5 and score lower than 0.5 is rejected by the localization block to improve localization. The acceptable range of melanoma lesion localization is opted within the IoU of 0.5–1, where 0.3–0 is chosen for background normal skin region.

The melanoma localization block defines a bounding box across the melanoma lesion and then crops the infected region. The cropped region is segmented into melanoma defined region and normal skin region. The melanoma defined lesion is passed to deep feature extraction block for discriminative feature maps generation and melanoma malignancy recognition.

The problem to recognize benign and malignant melanoma from training dermoscopy melanoma lesion R(x, y, w, h) can be written as forecasting class probability problem:

$$\Re_E(z_i, R(x, y, w, h)) = P(\text{class}|R(x, y, w, h)), \quad (10)$$

where \mathfrak{R}_E is the predicted class recognized by the proposed region-extreme convolutional neural network and P(class|R(x, y, w, h)) represents the probabilities of benign b and malignant m melanoma class, where class $\varepsilon(b, m)$, bstands for benign and m stands for malignant melanoma class. The test sample z_j is fed to region-extreme convolutional neural network \mathfrak{R}_E to predict the class probability of benign b and malignant m melanoma. Region-extreme convolutional neural network \mathfrak{R}_E recognizes the benign and malignant melanoma classes as shown in (10).

2.2. Skin Refinement Block. Dermoscopy imaging techniques are used for diagnosis of melanoma because these images examine the skin lesion at deeper level and ensure maximum perceptual ability for melanoma diagnosis [3]. Although dermoscopic images enhance the visual clarity, still there is a need to improve the manual prediction through advanced deep learning techniques.

The artefacts within the dermoscopy images D are responsible for degrading the melanoma segmentation performance. Therefore, the hair, clinical rule marks, and veins are removed through morphological closing operation applied twice across the image using equation (11). The twoline structuring elements S_1 of 10 pixels in 90° and 180° are used to eliminate the hair and clinical rule marks in horizontal and vertical direction. We have selected these structuring elements to remove the line shape objects including the hair and clinical rule marks from the RGB dermoscopy images, as these artefacts degrade the segmentation quality of melanoma.

$$J = (D \oplus S_1) \ominus S_1. \tag{11}$$

In equation (11), J represents the enhanced dermoscopic image through morphological closing operation, where D is input dermoscopic image and S_1 is line structuring element. Then, the resultant image J is processed further to improve the image edges and entropy through unsharp filtering operation.

$$J_s = J \times f_{\text{unsharp}}.$$
 (12)

The unsharp filter f_{unsharp} is convolved with the image J to smooth the image and generate a blur image J_s .

$$f_{\text{unsharp}} = -\frac{i}{\pi\sigma^4} \left[1 - \frac{a^2 + b^2}{2\sigma^2} \right] e^{(a^2 + b^2/2\sigma^2)}.$$
 (13)

Here, f_{unsharp} represents the unsharp kernel to transform dermoscopic image J into enhanced image J_s . a and brepresent the distance from centroid pixel in horizontal and vertical direction, and σ represents the standard deviation of image pixel's probability distribution.

$$I = J - J_s(x, y).$$
 (14)

Then, the smooth image I_s is subtracted from original image J, to obtain the sharped image I as shown in equation (14). After refining the dermoscopy images, I are forwarded to localization and segmentation block for localization of melanoma lesion.

2.3. Localization and Segmentation Block. Melanoma localization is a significant step in the proposed region-extreme convolutional neural network as the entire dermoscopic image consists of normal skin and clinical or natural artefacts along with melanoma lesion. The architecture of CNN network is used as backbone model in learning of regionextreme convolutional neural network model which is shown in Table 1.

Recent deep learning approaches have achieved good results in automatic malignancy recognition of melanoma but still there is a need to improve the recognition performance as difference exists between the expert prediction and the prediction obtained from automated method due to challenges mentioned above. Therefore, proposing the discriminative melanoma malignancy representation still demands more analysis with limited training data available. Recent findings employ very large deep networks for the malignancy analysis and prediction. However, it is very difficult to explore very deep networks and identify optimal solution as compared to shallower networks. Moreover, the problem of vanishing gradient becomes more crucial during training of very deep convolutional neural network and thus, it is more challenging to identify reasonable hyperparameters of large networks. Like other biomedical image analysis tasks, the availability of skin lesion training data is limited which makes the malignancy recognition problem even more challenging. We exploit the significance of melanoma lesion segmentation and investigate the benefits of generating learning model based on melanoma affected skin pixels only, despite the entire dermoscopy image. In this way, inference model becomes more decisive and performs accurate melanoma malignancy predictions.

In the proposed scheme, region based convolutional neural network (RCNN) is used for melanoma affected lesion localization that transforms the melanoma detection task as regression problem to predict the possible location of melanoma lesion [1]. We have selected RCNN because RCNN is able to learn the melanoma representation from the deep features and localize accurate region of melanoma [1]. Therefore, RCNN is applied to the image for localization of melanoma region. During training, region-extreme convolutional neural network considers the melanoma lesion and the rest of the region is considered as background, including the clinical artefacts, e.g., black frame, clinical swatches, clinical scale marks, and healthy skin.

We opt for localization of melanoma lesion R(x, y, w, h)to reduce the search space for precise segmentation of melanoma affected region that is free from normal skin and artefacts. RCNN localizes a rectangular region R(x, y, w, h)around the melanoma infected lesion. For more preciseness, melanoma lesion R(x, y, w, h) is further processed to remove normal skin pixels and define precise boundaries around the melanoma affected region. Therefore, fuzzy c means (FCM) clustering algorithm is applied to segment out melanoma region. The FCM clusters the image pixels and the FCM algorithm requires a predefined number of clusters as an input parameter. In our approach, as localized region consists of melanoma region and normal skin pixels, therefore FCM divides the image into two groups resulting in definite melanoma boundary representation. The FCM clustering refines the segmentation clusters $c_{a\varepsilon(1,2)}$ through evolving the objective function O_d under defined initial conditions.

$$S_{sa} = \frac{1}{\sum_{k=1}^{24} \left(\left\| R(x, y, w, h) - c_a \right\|^{(2/d-1)} / \left\| R(x, y, w, h) - c_k \right\| \right)}.$$
(15)

The variable *d* describes the total number of pixels existing within the localized rectangular melanoma region R(x, y, w, h) and S_{sa} describes the degree of membership of pixels with cluster *a*, having cluster center c_a .

$$c_a = \frac{\sum_{s=1}^{N} S_{sa} R(x, y, w, h)}{\sum_{s=1}^{N} S_{sa}^d}.$$
 (16)

In order to define precise pixel's membership of R(x, y, w, h) with cluster center c_a , the local minima are estimated through computing the objective function O_d :

$$O_d = \sum_{S=1}^d \sum_{a=1}^2 S_{sa}^d |R(x, y, w, h) - c_a|^2.$$
(17)

When the centroids of FCM converge, the segmentation of R(x, y, w, h) generates melanoma affected region with defined boundaries which is separated from the normal skin region.

2.4. Deep Feature Extraction Block. After segmenting the melanoma affected lesion R(x, y, w, h), deep feature representation F_N of benign and malignant melanoma is computed through CNN model. In our algorithm, we have used AlexNet [36], VGG16 [37], VGG19 [37], ResNet18 [38], ResNet101 [38], GoogleNet [39], and DenseNet201

Layers	Functionality	Filters	Size	Stride	Padding
Input layer	Preprocesses the input dermoscopic image		$32 \times 32 \times 3$		
Convolutional layer	Convolves input data with kernels	3	5×5		0
Max pooling	Performs maximum pooling operation on receptive fields		3×3	2	
RELU	Applies RELU activation function on receptive fields				
Convolutional layer	Convolves receptive fields with kernels	32	5×5	1	2
RELU	Applies RELU activation function on receptive fields				
Average pooling	Performs average pooling on receptive fields		3×3	2	0
Convolutional layer	Convolves receptive fields with kernels	64	5×5	1	2
RELU	Applies RELU activation function on receptive fields				
Avg pooling	Performs average pooling on receptive fields		3×3	2	0
Fully connected	Flattens the receptive fields from 2D to 1D	64			
RELU	Applies RELU activation function on receptive fields				
Fully connected	Transforms high level melanoma representation to nonlinear transfer function	2			
Softmax	Assigns class probabilities to region proposals				
Regression	Performs non-maximum suppression and estimates the location of melanoma				

TABLE 1: The region-extreme convolutional neural network architecture used for localization of melanoma.

[40], and the best performing model is chosen from these CNN models to capture a deep visual representation of benign and malignant melanoma. We have explored the recognition capabilities of the deep features at different network depths and the impact of suitable CNN network for melanoma malignancy prediction from the dermoscopic images. In the proposed technique, deep features are extracted from either pooling or fully connected layers and softmax layer is eliminated. The deep feature representation F_N of benign and malignant melanoma is normalized through min-max normalization algorithm and then passed to PCA to optimize the features equations (18) and (19). The PCA reduces the variance among the features of similar class and improves the learning process.

$$F'_N \nu = \tau \nu. \tag{18}$$

The mathematical expression (equation (18)) represents the principal component symmetric matrix F'_N , which is also known as co-variance matrix, where τ represents the diagonal matrix and ν represents the orthogonal projections matrix which is used for PCA computation.

$$F'_{N} = \begin{bmatrix} \nu_{11}F_{N} & \dots & \nu_{1n}F_{N} \\ \vdots & \ddots & \vdots \\ \nu_{nn}F_{N} & \dots & \nu_{nn}F_{N} \end{bmatrix}.$$
 (19)

The idea of using PCA on top of the deep features is to identify significant pattern that is helpful in recognition and improving the learning performance. Recently, PCA was applied in hierarchical layers for feature extraction that validates the performance improvement hypothesis [41]. PCA features were used to train autoencoder, and significant performance improvement was reported [42]. In our work, we have extracted the deep features from CNN network, normalized them, and then used PCA to optimize the deep features without reducing the dimension. Despite [41, 42], we have used PCA after deep features extraction from CNN to get discriminative features. We have observed that PCA significantly improves the recognition of benign and malignant melanoma using the deep convolutional features. The PCA detects the correlation ν among CNN features of

the same class and reduces the correlated features within the different class which is significant to discriminate the benign and malignant melanoma. The experiments and results are discussed in Section 3.4.2.

Deep features F'_N are normalized within minimum feature $F_{N_{\min}}$ ' and maximum feature $F_{N_{\max}}$ ' range to optimize the decision boundaries of recognition model and reduce the search space with benefits of lower computational time. The extracted deep principle component F'_N strengthens the decision boundaries of model and supports in precise prediction of melanoma benign and malignant class.

2.5. Feed-Forward Model Learning Block. The extreme learning machine (ELM) performs fast learning using single-layer feed-forward neural network with defined number of hidden nodes. Our previous work indicated that ELM holds better generalization performance as compared to back-propagation algorithms [43]. The weights and biases of ELM's first layer are randomly assigned and remain constant, while the weights and biases of hidden layer are opted to minimize the least square error between actual and predicted class, thus establishing stronger decision bound-aries [44, 45].

The feed-forward model learning block constitutes ELM classifier to learn model for benign and malignant melanoma recognition using deep principle components features F'_N . ELM consists of training parameters and these parameters are responsible for generating decision boundaries among benign and malignant melanoma. The selection of number of training parameters is a significant process and reduces the classification loss. A single output unit is used to predict benign and malignant melanoma through ELM output function $y_{L(F_n)}$ described as

$$y_L(F'_N) = \sum_{j=1}^L \omega_i h_i(F'_N),$$
 (20)

where $\omega_i = \omega_1, \omega_2, \omega_3, \dots, \omega_L$ describes the output weights and $h_{i(F'_N)} = h_{1(F'_N)}, h_{2(F'_N)}, \dots, h_{L(F'_N)}$ describes the weight existing between output vector and *L* hidden nodes. The classification decision function of ELM is Mathematical Problems in Engineering

$$y_L(F'_N) = \operatorname{sign}(h_i\omega_i). \tag{21}$$

The objective function of ELM optimizes the decision boundaries to generalize the prediction performance with minimum classification loss. The mathematical representation of objective function is shown in the following expression:

$$f^{*} = \text{minimize} \begin{cases} \|H\omega_{i} - T\|^{2}, \\ \omega_{i}. \end{cases}$$
(22)

Here, H^{\dagger} represents the matrix of hidden layer's output as

$$H = \begin{vmatrix} h_1 F_{N'_1} & \dots & h_L F_{N'_1} \\ \vdots & \ddots & \vdots \\ h_1 F_{N'_N} & \dots & h_L F_{N'_N} \end{vmatrix}.$$
 (23)

The ELM decision boundaries are established through maximizing the margins among the benign and malignant melanoma within feature representation $||2/\omega_i||$ and minimal least square method is used to minimize the norm of output weights $||\omega_i||$.

$$\omega_i = H^{\dagger} T. \tag{24}$$

The parameter H^{\dagger} is Moore–Penrose inverse matrix which is computed from orthogonal projection along with single value decomposition method. At test time, the ELM performs the prediction of class labels using decision function and predicts the melanoma malignancy through estimating the highest value of output node. The detailed explanation about ELM can be found in [44].

3. Experimental Results

3.1. Datasets, Hardware Specification, and Evaluation Parameters. The proposed technique is evaluated using ISIC-2016 and ISIC-2017 challenge dataset [46]. The ISIC-2016 consists of 900 training and 350 testing samples. The ISIC-2016 dataset constitutes training data from 727 nonmelanoma patients and 173 melanoma patients, while test sample contains data of 304 melanoma patients and 75 melanoma patients. The ISIC-2017 dataset consists of 2000 training samples, 150 validation, and 600 test samples of dermoscopy images. The training set constitutes data of 374 melanoma patients, 254 seborrheic keratosis patients, and 1372 benign nevi patients samples, while the test set contains data of 117 melanoma patients, 97 seborrheic patients, and 393 benign nevi patients.

The experiments are conducted on system with 2.59 GHz Core i7 CPU, 12 GB RAM, and NVIDIA GeForce GTX-950 Graphic Processing Unit.

Our proposed technique constitutes skin refinement step, melanoma localization, segmentation, and recognition of melanoma benign and malignant class. Each part of the proposed technique is evaluated with different evaluation measures. As the skin refinement phase is an enhancement process which was applied to improve the visual representation of image, therefore we have used peak signal to noise ratio (PSNR), root mean square error (RMSE), and universal image quality index (UIQI) for the evaluation of skin refinement step.

For the evaluation of melanoma localization, mean average precision (mAP) is used to estimate the precise localization. The intersection over union (IoU) measure is chosen within the range of 0.5-1 for melanoma detection. Jaccard index (Ja), dice score (Di), sensitivity (SE), specificity (SP), and accuracy (Ac) are used for evaluation of segmentation step. At ISIC-2016 and ISIC-2017 challenges, the contestants were ranked based on Ja index for the segmentation task. In the case of recognition step, four parameters are used to evaluate the classification task including, accuracy (AC), average precision (AP), sensitivity (SE), specificity (SP), and F_1 – score, respectively. Ac, SP, and SE are a similar metric as in segmentation. However, in the case of classification, they are recorded at the image level. In classification problems, F_1 -score is interpreted as a better choice to examine the performance of classifier using weighted average of sensitivity and specificity information. In the dataset, the testing set constitutes imbalance samples around 75 melanoma and 304 non-melanoma samples. Therefore, the false positive rate is smaller and true negative rate is larger comparatively. The organizers of the challenge employ average precision AP to rank the classification algorithms of participants [46].

3.2. Performance Evaluation of Skin Refinement. In the skin refinement step, the noisy artefacts are removed and visual information of dermoscopic images is refined to enhance the perceptual quality of the images. The skin refinement step is applied to all the dermoscopy images of ISIC-2016 and ISIC-2017 datasets. The resultant output samples are represented in Figure 3, and it can be observed that morphological operations and noise removal step resulted in dermoscopic images free from the clinical artefacts, e.g., rule marks and natural artefacts like the hair and blood vessels.

To further examine the quality of resultant image obtained from the skin refinement step, the image quality was estimated using RMSE, PSNR, and UIQI, and results are represented in Table 2. The PSNR measure is higher than 37.8 dB that signifies the higher ratio of information within the resultant refined image and minimum RMSE error. Moreover, the UIQI observation portrays the lower information loss in terms of contrast, illumination, and structural information on application of skin refinement step. As UIQI value is within the range of 0.4–0.7, and this measure represents the good perceptual quality of the preprocessed image. Therefore, these evaluation measures indicate that skin refinement step removed the natural and clinical artefacts without degrading the perceptual information of dermoscopic images.

3.3. Segmentation Phase. The next step after refining the dermoscopic images is localization and segmentation of melanoma lesion and the resulting segmented lesion is considered as region of interest (ROI) for the learning of melanoma malignancy model. In this experiment, localization



FIGURE 3: The resultant output sample of skin refinement phase. (a) The clinical and natural artefacts including thick and thin hair, black frame, clinical rule marks, gel bubbles, and color swatches. (b) The resultant intermediate output of morphological operation. (c) The resultant output of morphological closing operation and smooth operation. (d) The sharp image obtained after unsharped filter's convolution.

and segmentation performance of region-extreme convolutional neural network is evaluated. For generation of localization model, we trained the localization block using training images of datasets and evaluated the performance of trained localization model using test samples. The internal representation of the feature map to discriminate the melanoma region through localization and segmentation block is shown in Figure 4. The CNN softmax layer was responsible for predicting the probabilities at pixel level for categorizing each pixel into normal

TABLE 2: The performance of skin refinement step is estimated through computing PSNR, UIQI, and RMSE, considering each type of artefact.

Dataset		ISIC-2016			ISIC-2017	
Artefacts	PSNR	UIQI	RMSE	PSNR	UIQI	RMSE
	33.00	0.61	9.98	37.00	0.81	9.01
D11. f	28.52	0.41	16.56	38.52	0.77	7.86
black frame	34.00	0.60	9.24	32.40	0.68	8.42
	30.30	0.54	13.50	33.53	0.74	8.50
	30.67	060	12.92	34.74	070	10.82
Thisle hain	31.75	0.50	11.41	34.95	0.68	09.41
Thick hair	30.02	0.50	13.93	36.22	0.78	08.13
	35.70	0.51	7.25	38.70	0.67	4.98
	34.50	0.43	8.31	38.90	0.63	9.21
Thin hain	32.57	0.64	10.39	33.67	0.70	9.78
Inin nair	33.70	0.58	9.15	37.97	0.78	8.85
	37.78	0.56	5.65	44.78	0.87	3.85
	18.90	0.40	50.12	28.98	0.56	45.22
Derla manlea	29.71	0.57	14.43	32.34	0.67	23.98
Rule marks	32.26	0.55	10.77	34.66	0.54	18.37
	32.35	0.50	10.65	36.79	0.67	14.85
	35.31	0.55	7.57	33.22	0.55	7.57
Cruzatala an	32.35	0.53	10.66	34.75	0.65	9.76
Swatches	33.05	0.67	9.82	38.15	0.77	10.24
	30.25	0.70	13.72	37.67	0.87	12.62



FIGURE 4: The deep feature representation of melanoma region and normal skin region. These deep features are used to establish the model for melanoma malignancy recognition. The green overlay represents the normal skin attributes and purple overlay represents the melanoma affected lesion through localization and segmentation block.

or melanoma affected pixel. The localization block used these labeled pixels and applied selective search to select region proposal and computed the intersection over union (IoU) threshold for estimation of melanoma lesion. We have opted 0.5 IoU threshold for the detection of melanoma lesion; the region is considered as normal skin region below this value as shown in Figure 5. The RCNN localized the melanoma lesion with 0.94 mAP for ISIC-



FIGURE 5: The resultant ISIC samples obtained after melanoma localization step.

2016 and 0.95 mAP for ISIC-2017 dataset, as illustrated in Table 3.

For the comparison with other methods, the melanoma localized regions are further processed to obtain fine boundaries of melanoma region. The FCM clustered the pixels into melanoma lesion and normal skin. The resulting performance portrayed the good melanoma segmentation results as illustrated in Figure 6. The average values of SP as 0.9417, SE as 0.9782, Ac as 0.948, Di as 0.94, and Ja as 0.93 were recorded.

Our results are better than the state-of-the-art methods due to precise localization of melanoma lesion through the proposed method. In the few cases, localization block was unable to detect the melanoma region due to visual similarity with normal skin regions, as shown in Figure 7 [1].

3.4. Performance Evaluation of Recognition Task. To build the recognition model, we trained the recognition block using training images of ISIC-2016 and ISIC-2017 datasets and evaluated the performance of trained model using test samples. The sigmoid activation function of ELM was used for feed-forward model learning with 150000 number of hidden nodes. The backbone CNN models opted for to extract deep discriminative representation of melanoma were AlexNet [36], VGG16 [37], VGG19 [37], ResNet50 [38], ResNet101 [38], GoogleNet [39], and DenseNet201 [40], and their performance was observed based on segmented melanoma regions as listed in Table 4.

3.4.1. Impact of the Features Extracted from Different Network Depths. In this experiment, we explored the recognition capability of region-extreme convolutional neural network using different deep features extracted from the various CNN networks from segmented melanoma lesion. The best network was chosen as a feature extractor of the region-extreme convolutional neural network's recognition stage. The deep features extracted from segmented melanoma lesion were used to feed the ELM classifier for model learning. The generated ELM model was able to discriminate benign and malignant melanoma. The performance of proposed recognition model is reported in Table 4 which illustrates the optimal CNN model and layer for the deep features extraction.

The receptive fields generated from the deep features at different layers of convolutional neural networks represent contextual and semantic variations of dermoscopy images. Therefore, receptive fields at different level of convolutional network perform differently. On inspecting the behavior of melanoma malignancy model at different CNN layer, a common pattern is observed among all kinds of CNN models. The receptive field obtained from initial layers of CNN models was not decisive and is inferior as lower layers of CNN models collect low-level description and these patterns lack in-variance attribute. A similar characteristic was observed from the CNN middle layers, while the receptive fields obtained from the penultimate layers of CNN networks perform better than other layers because the penultimate layers receptive fields hold high level semantics and contextual information of melanoma malignancy.

It is notable that AlexNet fully connected layer Fc6 achieved the highest average precision, because Fc6 holds the higher dimensional (4096) deep features to represent the benign and malignant melanoma for ISIC-2016 and ISIC-2017 dataset, whereas the AlexNet pool 5 layer captured 9216 features but pool 5 feature's recognition capability is lower than fully connected layer Fc6, because higher-dimensional features of pool 5 become repeatable in both benign and malignant classes. This experiment supported us to identify the optimal backbone CNN model with penultimate layer to

Mathematical Problems in Engineering

TABLE 3: The achie	ved mean average	e precision (n	nAP) and	processing	time in 1	milliseconds	(ms)	for melanoma	localization	through	the
region-extreme cor	volutional neura	l network usi	ng ISIC-2	016 and IS	IC-2017 d	dataset.					

ISIC-2016 artefacts	mAP	Time (ms)	ISIC-2017 artefacts	mAP	Time (ms)
Thick hair	0.97	40.1	Thick hair	0.95	40.2
Think hair	0.96	42.1	Think hair	0.98	43.4
Rule marks	0.90	43.3	Rule marks	0.92	54.6
Color charts	0.91	39.4	Color charts	0.85	40.8
Black frame	0.96	41.6	Black frame	0.88	52.4



FIGURE 6: (a) Input images with artefacts. (b) Ground truth melanoma lesion images. (c) Localization of melanoma lesion. (d) Blue overlay of fuzzy segmentation of melanoma boundary and red overlay of ground truth melanoma lesion.

extract the deep features for generation of melanoma malignancy model. The performance of deep features extracted from the ResNet50 was comparable with AlexNet for ISIC-2016. However, the performance of deep features extracted from the ResNet50 was observed to be lower than AlexNet for ISIC-2017. The observed prediction accuracy and precision using AlexNet and ResNet50 features for ISIC-2016 are 0.86 and 0.83, while AlexNet performed better in prediction accuracy (0.93) and precision (0.91) to recognize benign and malignant melanoma.

Another observation that can be derived from Table 4 is that the penultimate layers of shallower sequential CNN model (e.g., AlexNet) performed better than the other deeper sequential and deeper directed acyclic CNN models, e.g., Resent50. Therefore, we have utilized AlexNet as the backbone model for features extraction and fed these extracted deep features from AlexNet to ELM for feedforward model generation. Here, it is important to highlight the essence of feed-forward ELM model that learned the model using deep features without backpropagation and weight optimization like CNN networks. This attribute improves the learning and reduces the computational time as compared to the state-of-the-art CNN models.

This experiment concluded that sequential shallower CNN network (e.g. AlexNet) penultimate layers deep features established the stronger ELM decision boundaries to predict the melanoma malignancy. On the basis of this experiment,



FIGURE 7: The localization and segmentation block fails to identify the melanoma benign cases, due to visual similarity with normal skin attributes.

 TABLE 4: Comparison of different network layers on the classification of benign and malignant melanoma.

Network	T	D a standard	ISIC	-2016	ISIC-2017		
Network	Layer	Features	AC	AP	AC	AP	
	pool5	9216	0.84	0.81	0.78	0.77	
AlexNet	Fc6	4096	0.85	0.83	0.93	0.91	
	Fc7	4096	0.84	0.80	0.88	0.87	
	Fc8	1000	0.84	0.77	0.87	0.86	
	Fc6	4096	0.82	0.73	0.86	0.88	
VGG16	FC7	4096	0.82	0.74	0.82	0.81	
	FC8	1000	0.82	0.73	0.82	0.81	
	Fc6	4096	0.84	0.83	0.88	0.87	
VGG19	Fc7	4096	0.81	0.72	0.85	0.84	
	Fc8	1000	0.81	0.70	0.80	0.80	
GoogleNet	Pred	1000	0.81	0.56	0.85	0.84	
DenseNet201	Fc1000	1000	0.82	0.77	0.87	0.86	
ResNet50	Fc1000	1000	0.86	0.83	0.92	0.90	
ResNet101	Fc1000	1000	0.81	0.68	0.86	0.88	

Bold values indicate the best results.

we opted for AlexNet as feature extraction model for recognition block and examined the rest of our experiments using AlexNet as feature extractor for recognition block. 3.4.2. Impact of PCA on Classification Task. To further improve the recognition model, we examined the dimensionality reduction of receptive fields obtained from Alexnet fc6 layer and applied PCA on it. As PCA reduces the covariant features and discovers compact, there is meaningful representation of deep features to recognize intraclass melanoma test samples. It can be noticed in Table 5 that the deep features generated a good model to discriminate benign and malignant melanoma, when represented by PCA. Therefore, using PCA at the top of deep features is beneficial for classification task.

3.4.3. Classification with and without Melanoma Segmentation Phase. To evaluate the performance of our regionextreme convolutional network, we evaluated the performance of classification with and without the segmentation of melanoma region. All the parameters and network architectures remained the same in both experiments. All the recorded experimental results are listed in Table 6. The segmented melanoma ROIs produced a better model in discriminating benign and malignant melanoma as compared to complete dermoscopic images because segmented lesion contains only benign and malignant feature

Dataset	DM	With PCA				Without PCA			
	PIN	AlexNet	VGG16	VGG19	ResNet50	AlexNet	VGG16	VGG19	ResNet50
ISIC-2016	AC	0.86	0.82	0.84	0.86	0.83	0.82	0.84	0.85
	AP	0.83	0.74	0.83	0.83	0.78	0.73	0.79	0.82
1610 2017	AC	0.93	0.86	0.88	0.91	0.92	0.85	0.86	0.89
1SIC-2017	AP	0.91	0.85	0.86	0.89	0.89	0.84	0.85	0.88

TABLE 5: Results of classification with and without PCA.

TABLE 6: Performance of classification task with and without segmentation.

Dataset	DM		Segn	nented		Without segmentation				
	PM	AlexNet	VGG16	VGG19	ResNet50	AlexNet	VGG16	VGG19	ResNet50	
ISIC-2016	AC	0.86	0.82	0.84	0.86	0.84	0.81	0.83	0.84	
	AP	0.83	0.74	0.83	0.83	0.77	0.70	0.73	0.80	
ISIC-2017	AC	0.93	0.86	0.88	0.92	0.91	0.84	0.85	0.88	
	AP	0.91	0.85	0.86	0.90	0.89	0.83	0.83	0.85	

representation despite having feature representation from normal skin or the artefacts. The AP of segmented image was relatively 5% higher than non-segmented dermoscopic images. Defining the melanoma area as a region of interest and the extraction of deep features to generate model is a smart way to train ELM, and to classify precisely without being distracted from the clinical artefacts or healthy skin region. We conclude from this experiment that the segmented melanoma lesion when fed to recognition block for extraction of receptive fields is more decisive in building accurate melanoma malignancy recognition model.

3.4.4. Impact of Classification Model in Melanoma Subclass Recognition. The performance of recognition block is further investigated through the impact of ELM as a classifier against traditional classifiers including support vector machine (SVM) and K-nearest neighbor (KNN). To observe the impact, the entire recognition block remained the same while the classifier layer was changed with ELM, support vector machine (SVM), and K-nearest neighbor (KNN). The region-extreme convolutional neural network segmented the melanoma lesion and extracted the deep features from backbone models such as Alexnet, VGG16, and VGG19 and trained different types of classifiers including ELM, SVM, and KNN as a separate model. It is evident from Table 7 that the ELM classifier outperformed all the classifiers, where K-nearest neighbor performed better that SVM classifier.

The ELM classifier trains the learning model in a twostep process; first it randomizes the deep feature mapping and then it identifies the solution for linear parameters. In the first step, the ELM hidden layer randomly initializes to map the deep features on ELM feature space using nonlinear piece-wise function. The second step of ELM separates it from traditional classifiers including SVM which utilizes the kernel function for feature projection, while the deep learning models require autoencoder or restricted Boltzmann machines to map input features and form the similar deep representation.

Essentially, the ELM generates hidden nodes randomly without explicitly being trained that makes ELM learning

generalized and invariant from input training data. The ELM's hidden nodes are assigned using continuous probability distribution. Therefore, ELM models are stronger than traditional classifiers (SVM, KNN) or back-propagation neural networks. Moreover, the regularization function within the ELM also supports it to avoid the problems of overfitting and underfitting without any computational overhead.

We found that the ELM enhances the performance of the system to predict intraclass label of melanoma, with 83% average precision for ISIC-2016 and 91% for ISIC-2017 dataset. The performance of different classifiers is listed in Table 7. All the performance metrics indicate the effectiveness of ELM within the recognition block of region-extreme convolutional neural network. Therefore, this experiment concludes that ELM is the optimal choice to learn the feed-forward melanoma malignancy model using dermoscopy images.

4. Discussion and Comparison with State-ofthe-Art Methods

4.1. Melanoma Segmentation. At ISIC-2016 challenge of skin lesion segmentation task, 28 contestants participated and the results of top scorers are presented in Figure 8. The final positions were marked based on highest Ja index. Almost all the top scorers utilized the benefit of deep learning methods to segment the melanoma region due to the performance gain of deep learning algorithms. The majority of researchers explored Alexnet, VGG-16, or deep residual network for the segmentation task and performed better than traditional handcrafted features-based segmentation techniques, whereas we have used RCNN [1] with shallower network for the same task and our results validated the good performance of our method as compared to other state-of-the-art traditional handcrafted feature segmentation and deep learning methods. From Figure 8, it can be observed that Ja of our method improved by 0.11 points as compared to ISIC-2016 segmentation task top-scorer ExB. The reason behind this significant improvement of our method is due to precise

			1				0	L L	,	0			
Dataset	DM	ELM			KNN				SVM				
	PIVI	AlexNet	VGG16	VGG19	ResNet50	AlexNet	VGG16	VGG19	ResNet50	AlexNet	VGG16	VGG19	ResNet50
ISIC-2016	AC	0.86	0.82	0.84	0.86	0.80	0.77	0.78	0.78	0.75	0.76	0.75	0.73
	AP	0.83	0.74	0.83	0.83	0.58	0.56	0.58	0.57	0.62	0.63	0.60	0.59
ISIC 2017	AC	0.93	0.86	0.88	0.88	0.83	0.83	0.84	0.80	0.86	0.80	0.81	0.85
15IC-2017	AP	0.91	0.88	0.87	0.86	0.81	0.81	0.81	0.80	0.84	0.80	0.80	0.83

TABLE 7: Comparison of different classifiers for recognition of benign and malignant melanoma class.



FIGURE 8: Comparison of the proposed segmentation technique with ISIC-2016 challenge participants results of skin lesion segmentation task.

localization step before segmentation. Region-extreme convolutional neural network precisely reduced the search space for melanoma lesion and produced the effective melanoma segmentation through considering the artefacts and normal skin as background region.

To further evaluate the performance of our proposed approach for melanoma segmentation, we compared our method with state-of-the art segmentation techniques, as represented in Table 8. Traditional approaches for melanoma segmentation, e.g., active contour, bootstrapping [9], contextual hypergraph [10], clustering [54], region split and merge, region growing [55], sparse coding [48], and thresholding techniques [11] produce lower value of average Ja score than our method. This is due to the fact that these approaches segmented the melanoma lesion based on spatial context information in unsupervised fashion, while our proposed approach initially builds a model based on the high-level color, texture, and spatial representation of melanoma. Therefore, the trained model localized melanoma lesion with high mAP at test time and only the detected region was considered for segmentation.

Besides the traditional techniques for melanoma segmentation, deep learning-based approaches including FCN [14] and Segnet [50] performed melanoma segmentation with Ja score of 0.86. The FCN [14] and Segnet [50] were designed as a deeper network for segmentation of melanoma lesion, therefore training the model requires tuning of thousands of hyperparameters which makes the melanoma segmentation task computationally expensive for real-time application. It is evident from Table 8 that our method exhibits lower computational time with precise melanoma segmentation as compared to state-of-the-art methods. Our Jaccard score is 13% higher than Segnet [50], due to the training of region-extreme convolutional neural network

Mathematical Problems in Engineering

TABLE 8: Performance comparison of region-extreme convolutional neural network's melanoma segmentation with state-of-the-art methods.

Technique	Ac	Di	Ja	Time (ms)
Adaptive thresholding [11]	0.72	0.56	0.45	2
Bootstrap learning [9]	0.78	0.72	0.57	_
Contextual hypergraph [10]	0.83	0.75	0.6	0.3
ISO [47]	0.82	0.68	0.56	_
Level set [12]	0.7	0.58	0.46	0.84
Sparse coding [48]	0.91	0.8	0.67	0.1
Statistical region growing [13]	0.73	0.55	0.43	0.4
Yen's thresholding [49]	0.81	0.67	0.58	_
FCN [14]	0.82	0.82	0.86	0.05
Segnet [50]	0.91	0.92	0.86	0.06
Region-extreme convolutional neural network	0.94	0.94	0.93	0.05

Bold values indicate the best results.

TABLE 9: Performance comparison of region-extreme convolutional neural network with ISIC-2016 challenge results for skin lesion classification task.

Method	AC	AP
CUMED	0.855	0.637
GTDL	0.813	0.619
BF-TB	0.834	0.598
ThrunLab	0.786	0.563
Jordan Yap	0.844	0.559
Haebeom Lee	0.821	0.555
GT-DL1	0.815	0.552
GT-DL2	0.681	0.545
Sebastien PARIS	0.731	0.542
USYD-BMIT	0.599	0.537
Region-extreme convolutional neural network (our)	0.86	0.831

Bold values indicate the best results.

considering the clinical and natural artefacts as background region and only detecting the melanoma lesions.

4.2. Melanoma Malignancy Recognition. In ISIC-2016 challenge, 25 teams participated at melanoma malignancy recognition task and submitted their results. The results of classification task were evaluated based on average precision (AP). In Table 9, top ten results of ten participants are listed. It can be observed that our system's AP value 0.831 is higher than all the participants, validating the effectiveness of the proposed scheme for the classification of melanoma into benign and malignant.

Moreover, our region-extreme convolutional neural network also performed better that existing state-of-the-art melanoma malignancy models, as illustrated in Table 10. To compare the melanoma malignancy recognition performance of region-extreme convolutional neural network against other CNN based methods, we have included a comparative study as illustrated in Table 10. Our region-extreme convolutional neural network first identified the melanoma lesion and segmented it to extract the deep features from penultimate layers of backbone model and further used these deep features to generate feed-forward extreme learning model for melanoma malignancy recognition. The method proposed by [34, 51] also used CNN models to extract the deep features and used fisher encoding scheme to train the SVM for melanoma malignancy recognition. The performance of [34] is comparable to our method, while our method performed 3% higher than [51] due to the difference in mapping function of classifier. The mapping function of the SVM [51] cannot define the decision boundaries accurately among the benign and malignant melanoma. Our region-extreme convolutional neural network improved the recognition performance without adapting transfer learning across domains and effectively generalized the recognition model through the ELM.

Comparing the recognition performance with [33, 52, 53] our region-extreme convolutional neural network outperformed these methods by 1%, 4%, and 1% on ISIC-2016 dataset and 2%, 10%, and 6% on ISIC-2017 dataset, because [33, 52, 53] built the melanoma malignancy model using entire dermoscopic image which also includes the artefacts and more than 50% of normal skin pixels. However, our region-extreme convolutional neural network only considered the melanoma lesion to establish the melanoma malignancy model.

4.3. Computational Time Complexity. The average computational time to recognize the melanoma malignancy using single dermoscopic image is reported in milliseconds (ms) in Table 10. The state-of-the-art methods have higher computational cost as [33] utilized very deep residual convolutional neural network that requires fine-tuning of millions of hyperparameters, while [53] aggregated the handcrafted features, combined three CNN pretrained

Method	Network	ISIC-2016	ISIC-2017	Time (ms)
	Alexnet	0.83	0.80	
CNN SVM [51]	VGG-16	0.84	0.79	0.1
	ResNet50	0.84	0.82	
	Alexnet	0.83	0.81	
CNNaug SVM [51]	VGGNet	0.83	0.82	0.1
	ResNet50	0.83	0.81	
Fisher GL [52]	Alexnet	0.82	0.81	0.13
Transfer learning CNN [38]	ResNet50	0.84	0.83	0.1
CUMED [33]	ResNet50	0.85	0.87	0.2
Codella [53]	HybridNet	0.80	0.91	0.09
DCNN-FV [34]	ResNet50	0.86	0.89	0.05
Region-extreme convolutional neural network (Ours)	Alexnet	0.86	0.93	0.05
Region-extreme convolutional neural network (Ours)	ResNet50	0.86	0.92	0.06

TABLE 10: Performance comparison of region-extreme convolutional neural network's melanoma malignancy recognition with state-of-theart CNN methods.

Bold values indicate the best results.

networks, and used sparse coding scheme to ensemble fully convolutional U-Net for melanoma malignancy recognition. Our proposed method requires 0.05 milliseconds to recognize the malignancy of melanoma. This is due to the feedforward model learning of ELM that generated melanoma malignancy model without back-propagation of the loss for weight optimization. Therefore, our region-extreme convolutional neural network is more efficient and can be easily employed to other applications as well.

5. Conclusion

In this paper, we presented a novel region-extreme convolutional neural network model to predict the malignancy of melanoma from dermoscopic images. The novel regionextreme convolutional neural network is capable of melanoma localization with precise melanoma boundary segmentation along with malignancy recognition of benign and malignant melanoma. Our demonstrated results on ISIC-2016 and ISIC-2017 showed that the localization of melanoma lesion is fruitful to learn accurate model for melanoma recognition as compared to using an entire dermoscopic image. We achieved higher performance gain as compared to state-of-the-art methods for melanoma segmentation and recognition task. We also revealed that the deep features extracted from inner layers of CNN model hold high-level representation of melanoma malignancy information and established stronger model as compared to the features extracted from the last layer of CNN model. Furthermore, our region-extreme convolutional neural network also validates that the deep features with ELM classifier generate accurate model for melanoma recognition. Our region-extreme convolutional neural network highlighted the effective procedure for training and can be utilized in other biomedical image analysis applications. In the future, we will investigate the comprehensive clinical decision support system for medical images and explore our system on other applications.

Data Availability

The authors have used two publicly available datasets (ISIC2016 and ISIC2017).

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Nudrat Nida contributed to conceptualization of this study, methodology, software, writing, and validation. Aun Irtaza contributed to data curation, technical support, investigation, writing—review, and validation. Muhammad Haroon Yousaf contributed to supervision, technical support, validation, writing—review and editing, and project administration.

References

- N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering," *International Journal of Medical Informatics*, vol. 124, pp. 37–48, 2019.
- [2] F. Rashid, A. Irtaza, N. Nida, A. Javed, H. Malik, and K. M. Malik, "Segmenting melanoma lesion using single shot detector (SSD) and level set segmentation technique," in Proceedings of the 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), pp. 1–5, Karachi, Pakistan, December 2019.
- [3] C. Barata, M. E. Celebi, and J. S. Marques, "Development of a clinically oriented system for melanoma diagnosis," *Pattern Recognition*, vol. 69, pp. 270–285, 2017.
- [4] K. A. Freedberg, A. C. Geller, D. R. Miller, R. A. Lew, and H. K. Koh, "Screening for malignant melanoma: a cost-effectiveness analysis," *Journal of the American Academy of Dermatology*, vol. 41, no. 5, pp. 738–745, 1999.
- [5] C. M. Balch, A. C. Buzaid, S.-J. Soong et al., "Final version of the american joint committee on cancer staging system for

cutaneous melanoma," *Journal of Clinical Oncology*, vol. 19, no. 16, pp. 3635–3648, 2001.

- [6] M. Binder, M. Puespoeck-Schwarz, A. Steiner et al., "Epiluminescence microscopy of small pigmented skin lesions: short-term formal training improves the diagnostic performance of dermatologists," *Journal of the American Academy* of Dermatology, vol. 36, no. 2, pp. 197–202, 1997.
- [7] N. Codella, V. Rotemberg, P. Tschandl et al., "Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC)," 2018, https://arxiv.org/abs/1902.03368.
- [8] M. E. Celebi, N. Codella, A. Halpern, and D. Shen, "Guest editorial skin lesion image analysis for melanoma detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 479-480, 2019.
- [9] N. Tong, H. Lu, X. Ruan, and M.-H. Yang, "Salient object detection via bootstrap learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1884–1892, Boston, MA, USA, June 2015.
- [10] X. Li, Y. Li, C. Shen, A. Dick, and A. Van Den Hengel, "Contextual hypergraph modeling for salient object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3328–3335, Sydney, Australia, December 2013.
- [11] M. Silveira, J. C. Nascimento, J. S. Marques et al., "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 35–45, 2009.
- [12] C. Chunming Li, C.-Y. Chiu-Yen Kao, J. C. Gore, and Z. Zhaohua Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [13] M. Emre Celebi, H. A. Kingravi, H. Iyatomi et al., "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, no. 3, pp. 347–353, 2008.
- [14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.
- [15] L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees, "A color and texture based hierarchical k-nn approach to the classification of non-melanoma skin lesions," in *Color Medical Image Analysis*Springer, Dordrecht, Netherlands, 2013.
- [16] T. Tommasi, E. La Torre, and B. Caputo, "Melanoma recognition using representative and discriminative kernel classifiers," in *Proceedings of the International Workshop on Computer Vision Approaches to Medical Image Analysis*, pp. 1–12, Graz, Austria, May 2006.
- [17] R. J. Stanley, W. V. Stoecker, and R. H. Moss, "A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images," *Skin Research and Technology*, vol. 13, no. 1, pp. 62–72, 2007.
- [18] Y. Cheng, R. Swamisai, S. E. Umbaugh et al., "Skin lesion classification using relative color features," *Skin Research and Technology*, vol. 14, no. 1, pp. 53–64, 2008.
- [19] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler, "Automated melanoma recognition," *IEEE Transactions on Medical Imaging*, vol. 20, no. 3, pp. 233–239, 2001.
- [20] M. E. Celebi, H. A. Kingravi, B. Uddin et al., "A methodological approach to the classification of dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 31, no. 6, pp. 362–373, 2007.

- [21] G. Schaefer, B. Krawczyk, M. E. Celebi, and H. Iyatomi, "An ensemble classification approach for melanoma diagnosis," *Memetic Computing*, vol. 6, no. 4, pp. 233–240, 2014.
- [22] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.
- [23] H. Chen, X. Qi, L. Yu, and P.-A. Heng, "DCAN: deep contouraware networks for accurate gland segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2487–2496, Las Vegas, NV, USA, June 2016.
- [24] H.-C. Shin, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285– 1298, 2016.
- [25] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu et al., "Convolutional neural networks for medical image analysis: full training or fine tuning?" *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.
- [26] H. R. Roth, L. Lu, A. Seff et al., "A new 2.5D representation for Lymph node detection using random sets of deep convolutional neural network observations," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 520–527, Boston, MA, USA, September 2014.
- [27] Q. Dou, H. Chen, L. Yu et al., "Automatic detection of cerebral microbleeds from mr images via 3D convolutional neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1182–1195, 2016.
- [28] A. A. A. Setio, F. Ciompi, G. Litjens et al., "Pulmonary nodule detection in ct images: false positive reduction using multiview convolutional networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.
- [29] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.
- [30] N. Nida and M. Khan, "A framework for building an automatic brain tumor diagnostic tool using magnetic resounance imaging," *Pakistan Journal of Science*, vol. 67, no. 3, p. 276, 2015.
- [31] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images," in *Proceedings* of the International Workshop on Machine Learning in Medical Imaging, pp. 118–126, Munich, Germany, October 2015.
- [32] J. Kawahara, A. BenTaieb, and G. Hamarneh, "Deep features to classify skin lesions," in *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1397–1400, Prague, Czech Republic, April 2016.
- [33] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2016.
- [34] Z. Yu, X. Jiang, F. Zhou et al., "Melanoma recognition in dermoscopy images via aggregated deep convolutional features," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 4, pp. 1006–1016, 2018.

- [35] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Regionbased convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis* and Machine Intelligence, vol. 38, no. 1, pp. 142–158, 2015.
- [36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.
- [37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https:// arxiv.org/abs/1409.1556.
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference* on Computer Vision and Pattern Recognition, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [39] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.
- [40] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, "DenseNet: implementing efficient convnet descriptor pyramids," 2014, https://arxiv.org/abs/1404.1869.
- [41] B. Mitchell and J. Sheppard, "Deep structure learning: beyond connectionist approaches," in *Proceedings of the 2012 11th International Conference on Machine Learning and Applications*, pp. 162–167, Boston, MA, USA, December 2012.
- [42] S.-Z. Li, B. Yu, W. Wu, S.-Z. Su, and R.-R. Ji, "Feature learning based on SAE-PCA network for human gesture recognition in RGBD images," *Neurocomputing*, vol. 151, pp. 565–573, 2015.
- [43] N. Nida, M. H. Yousaf, A. Irtaza, and S. A. Velastin, "Instructor activity recognition through deep spatiotemporal features and feedforward extreme learning machines," *Mathematical Problems in Engineering*, vol. 2019, Article ID 2474865, , 2019.
- [44] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*, pp. 985–990, Budapest, Hungary, July 2004.
- [45] N. Nan Liu and H. Han Wang, "Ensemble based extreme learning machine," *IEEE Signal Processing Letters*, vol. 17, no. 8, pp. 754–757, 2010.
- [46] D. Gutman, N. C. Codella, E. Celebi et al., "Skin lesion analysis toward melanoma detection: a challenge at the International Symposium on Biomedical Imaging (ISBI), 2016, International Skin Imaging Collaboration (ISIC)," 2016, https://arxiv. org/abs/1605.01397.
- [47] T. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems*, *Man, and Cybernetics*, vol. 8, no. 8, pp. 630–632, 1978.
- [48] B. Bozorgtabar, M. Abedini, and R. Garnavi, "Sparse coding based skin lesion segmentation using dynamic rule-based refinement," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, pp. 254–261, Athens, Greece, 2016.
- [49] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal* of *Electronic Imaging*, vol. 13, no. 1, pp. 146–166, 2004.
- [50] V. Badrinarayanan, A. Handa, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling," 2015, https://arxiv.org/abs/1505. 07293.

- [51] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 806–813, Columbus, OH, USA, June 2014.
- [52] T. Uricchio, M. Bertini, L. Seidenari, and A. Bimbo, "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, Santiago, Chile, 2015.
- [53] N. C. Codella, Q.-B. Nguyen, S. Pankanti et al., "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5–1, 2017.
- [54] Z. Zhao, T. Bouwmans, X. Zhang, and Y. Fang, "A fuzzy background modeling approach for motion detection in dynamic backgrounds," in *Proceedings of the International Conference on Multimedia and Signal Processing*, pp. 177–185, Shanghai, China, December 2012.
- [55] A. Wong, J. Scharcanski, and P. Fieguth, "Automatic skin lesion segmentation via iterative stochastic region merging," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 6, pp. 929–936, 2011.