SKIN LESION SEGMENTATION USING LOCAL BINARY CONVOLUTION-DECONVOLUTION ARCHITECTURE

OMRAN SALIH AND SERESTINA VIRIRI✉

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa
e-mail: omran@aims.ac.za, viriris@ukzn.ac.za

ABSTRACT

Deep learning techniques such as Deep Convolutional Networks have achieved great success in skin lesion segmentation towards melanoma detection. The performance is however restrained by distinctive and challenging features of skin lesions such as irregular and fuzzy border, noise and artefacts presence and low contrast between lesions. The methods are also restricted with scarcity of annotated lesion images training dataset and limited computing resources. Recent research in convolutional neural network (CNN) has provided a variety of new architectures for deep learning. One interesting new architecture is the local binary convolutional neural network (LBCNN), which can reduce the workload of CNNs and improve the classification accuracy. The proposed framework employs the local binary convolution on U-net architecture instead of the standard convolution in order to reduced-size deep convolutional encoder-decoder network that adopts loss function for robust segmentation. The proposed framework replaced the encoder part in U-net by LBCNN layers. The approach automatically learns and segments complex features of skin lesion images. The encoder stage learns the contextual information by extracting discriminative features while the decoder stage captures the lesion boundaries of the skin images. This addresses the issues with encoder-decoder network producing coarse segmented output with challenging skin lesions appearances such as low contrast between healthy and unhealthy tissues and fine grained variability. It also addresses issues with multi-size, multi-scale and multi-resolution skin lesion images. The deep convolutional network also adopts a reduced-size network with just five levels of encoding-decoding network. This reduces greatly the consumption of computational processing resources. The system was evaluated on publicly available dataset of ISIC and PH2. The proposed system outperforms most of the existing state-of-art.

Keywords: convolution-deconvolution architecture, deep learning network, local binary convolution melanoma, segmentation, skin Lesion.

INTRODUCTION

Melanoma skin lesion has been reported as the deadliest skin cancer with high mortality (Codella *et al.*, 2019). Segmentation is an important task in automating disease diagnosis. It aids in identification and recognition of disease pattern. Early detection of this disease can reduce the mortality rate (Celebi *et al.*, 2007). In the past few years, with the advantages in technology, various computer aided techniques have evolved for the analysis and segmentation of medical images. Automatic medical image analysis methods have been successful in medical image analysis over the past two decades (Nguyen and Lee, 2019). Automated analysis of skin lesion has assisted clinicians in making quick and accurate decisions in melanoma detection.

Some of these techniques have been categorized into threshold methods that use threshold values to differentiate the skin lesions from the surrounding healthy tissues (Silveira *et al.*, 2009; Emre *et al.*, 2013; Al-Masni *et al.*, 2018), clustering methods that utilize color space features to determine cluster

regions (Sathya and Manavalan, 2011; Xie and Bovik, 2013; Dhanachandra *et al.*, 2015), edge-based and region-based techniques that focus on edge and region analysis models (Rajab *et al.*, 2004; Schaefer *et al.*, 2009; Wong *et al.*, 2011), active contour methods which adopt evolution algorithm to segment skin lesions (Yuan *et al.*, 2009; Kasmi *et al.*, 2016) and supervised segmentation methods for skin lesions segmentation such as SVM and artificial neural networks (ANN) (Wighton *et al.*, 2009; Bi *et al.*, 2016; Li and Shen, 2018). Segmentation of skin lesions still remains challenging and most of the existing techniques have not been able to give accurate and reliable results (Okuboyejo *et al.*, 2014; Pennisi *et al.*, 2016; Premaladha and Ravichandran, 2016). These challenges are due to unique features and the peculiarity in the fine grained appearances of skin lesion images. Skin lesion images are sometimes characterised with hairs, oils, marks, skin lines and blood vessels, variegated colouring, weak edges, fuzzy and irregular borders as depicted in Fig. 1. There is also low contrast between the appearance of the health lesion and the affected lesion (Eltayef, 2017).
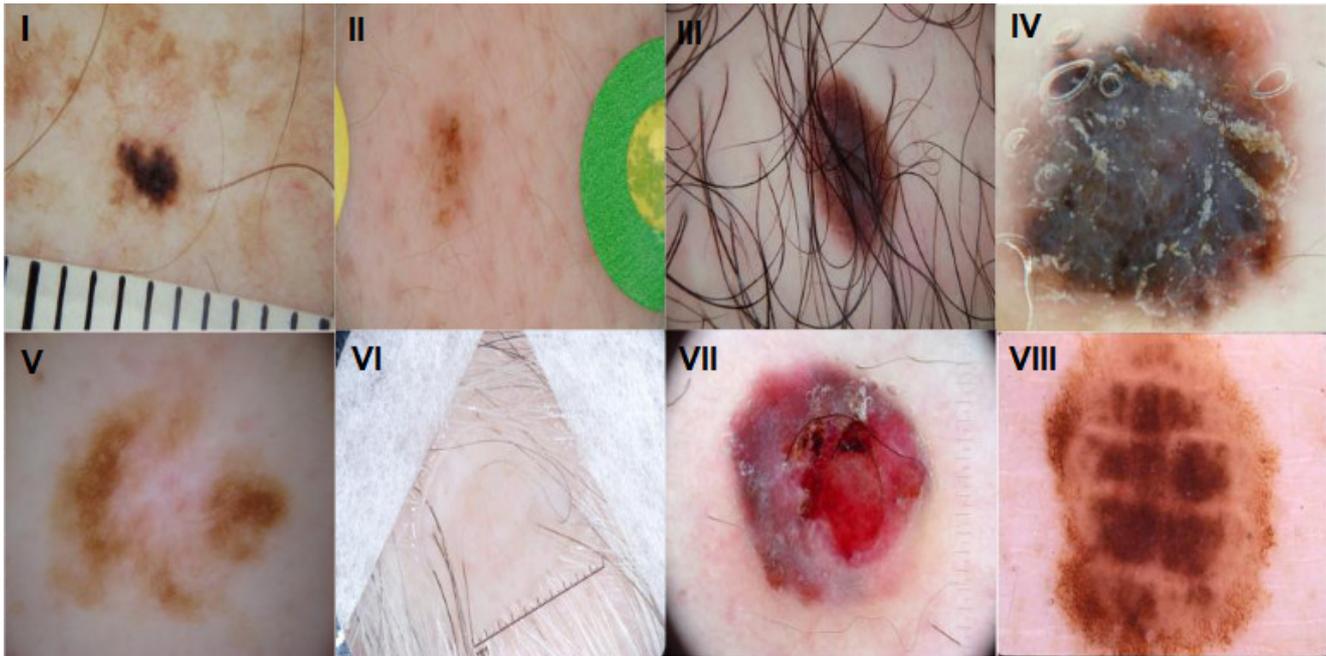
Fig. 1. *(I) ruler mark artefact, (II) color illumination, (III) hair artefact, (IV) bubbles, (V) low contrast, (VI) frame artefact, (VII) blood vessel, (VIII) irregular boundaries..*

Deep learning methods have achieved enhanced performances in medical imaging tasks in the recent years (Lee *et al.*, 2017). Their performance is leveraged in their capacity to learn and extract deep and hierarchical features from complex image dataset (Lee *et al.*, 2017). More recently, deep learning methods such as deep convolutional models have achieved great success for medical image segmentation (Litjens *et al.*, 2017). Deep convolutional neural networks (CNNs) possess capacity to process general and highly variable tasks in fine-grained objects (Sainath *et al.*, 2013). They can also learn hierarchically features that correspond to appearance and semantics of images in large labelled dataset.

The performance of deep learning methods is however restricted with scarcity of annotated medical training data. The segmentation output of these techniques is coarse with poor boundary (Bi *et al.*, 2017), as a result of insufficient training data set. They apply heavy tuning of large number of parameters and pre-processing techniques to minimize the effect of the limitations. This increases computational resources consumption (Bi *et al.*, 2017). This work aims to develop an efficient deep learning framework with a reduced cost for robust segmentation of skin lesions.

The paper identifies some major obstacles that limit the performance of deep learning approach on skin lesion segmentation. First, existing skin lesion dataset are limited in volume and not sufficient to

effectively train deep learning systems. Second, skin lesion images are characterized with low contrast, fuzzy borders and noise presence such as hairs, bubbles, oil, air etc. Third, skin lesion images are multi-size, multi-scale and multi-resolution in nature. Fourth, some training data sets may contain missing or incorrect annotations among the skin lesion images set. Lastly, most of the existing deep learning techniques require expensive computational processing resources which substantially limits them in practical scenarios for medical application.

The proposed model is built from two different deep learning architectures, in order to overcome the challenges of skin lesion segmentation which have been motioned earlier on. These architectures are the local binary convolution neural network (LBCNN) architecture (Juefei *et al.*, 2017) and U-net architecture (Ronneberger *et al.*, 2015). The first part of the LBCNN is used as encoder and the second U-net part is used as decoder in our proposed system. Juefei-Xuetal (Ronneberger *et al.*, 2015) proposed a hybrid combination of fixed and learned weights introducing a local binary convolutional layer (LBC). The authors state that CNN with LBC layers, called local binary convolutional networks (LBCNN) has lower model complexity and are less prone to over-fitting.

Three publicly available dataset of skin lesion images are used to evaluate the proposed system. The dataset includes ISIC 2018 skin lesion images

dataset of 2594 samples of training images and 260 samples of validation images, ISIC 2017 skin lesion images dataset of 2000 images of training images and 300 images of validation images, and PH2 skin lesion images dataset 200 sample of training images and 50 samples of validation images. An experimental analysis was carried out and the performance compared with the existing state-of-the-art methods. The proposed model outperformed all of these methods in evaluated metrics such as accuracy, Jaccard index and DICE coefficients. The remaining part of this work is organized as follow: Section "Related Works" discusses the related works and LBCNNs architecture is described in Section "Local Binary Convolutional Neural Networks (LBCNN)". Section "The Proposed Model" the experiments and results. The proposed method performance is discussed in section "Experiments and Results". The paper is concluded with sections "Discussion" and "Conclusion".

RELATED WORKS

Deep Convolutional Network (DCN) methods have achieved great success in biomedical image segmentation. Segmentation of biomedical images is one of the challenging tasks in computer vision. Various researchers have proposed and developed various architectures based on DCN for skin lesion segmentation. Some of the architectures which have been critically reviewed in this paper include *U*-net, Fully Convolutional Residual Network (FCRN), deep convolutional-deconvolutional neural network (CDNN) and deep Encoder-decoder networks.

Ronneberger *et al.* (2015) proposed a *U*-net architecture to provide a probability estimate for each pixel in the original image while segmenting skin lesion images. Vesal *et al.* (2018) extended the *U*-net architecture by employing dilated convolutions in the lowest layer for encoder branch of the *U*-net system. This is to provide a more global context for the features extracted from skin lesion image. The system popularly known as SkinNet was an improved version of *U*-net architecture however still possess some challenges for example some image details still get lost through the *U*-net shortcut skip connection, making decoder section weaker in fully recovering feature maps when they are passed through skip connections from the encoder section (Sharma *et al.*, 2019).

Yuan (2017b) deployed a deep convolutional-deconvolutional neural networks on different color spaces of dermoscopy images for skin lesion segmentation. He *et al.* (2017) extended deep dense convolutional layer with a generic multi-path deep network known as RefineNet for segmentation of dermoscopic images. The deep representation of all available layer was then aggregated to form global feature maps using skip connection. The system was able to capture diverse appearance feature in the contextual information extracted from images. The deconvolutional layer was applied to smoothen the segmentation maps and obtain final high-resolution output. Nguyen and Lee (2018) proposed a poly segmentation method based on architecture of a multiple deep Encoder-decoder networks called CDED-net. The system captures object boundaries using multi-scale decoders which is integrated with both boundaries using multi-scale decoders which is integrated with both boundary-emphasization data augmentation method and a novel loss function. System with deep convolutional-deconvolutional neural network require high computational cost. The segmentation results of DCNN-based methods are still coarse and limited most especially with insufficient training data set. Furthermore, medical images boundaries most especially skin lesions images are hardly defined due to the low contrast of lesions between the appearance of the unhealthy lesions and the healthy tissue (Vesal *et al.*, 2018).

Lately, various approaches and techniques of deep learning systems have been employed in the past to tackle skin lesion problem. These include methods such as: Li and Shen (2018) developed a dense deconvolutional network (DDN) for skin lesion segmentation based on residual learning. The DDN specifically consists of dense deconvolutional layers (DDLs), chained residual pooling (CRP), and hierarchical supervision (HS) which they play significant role in learning the discriminative feature representations and integrate multi-level contextual information, effectively. Hasan *et al.* (2020) presented automatic semantic segmentation method called dermoscopy skin network (DSNet) to minimize the parameters number of network (lightweight). Feng *et al.* (2020) proposed Context Pyramid Fusion Network (named CPFNet) to solve context information extraction capability of a single stage that is insufficient in the deep learning, due to the problems such as imbalanced class and blurred boundary. Zafar *et al.* (2020) adapted an automated technique for segmenting lesion boundaries that combines two architectures, the *U*-net and the ResNet, collectively called Res-Unet. It also used image inpainting for hair removal to improve the segmentation results significantly. Al-Masni *et al.* (2020) developed an integrated diagnostic framework that combines a skin lesion boundary segmentation stage and a multiple

skin lesions classification stage. Xie *et al.* (2020b) designed the mutual bootstrapping deep convolutional neural networks (MB-DCNN) model for simultaneous skin lesion segmentation and classification, where both segmentation and classification networks mutually transfer knowledge between each other and facilitate each other in a bootstrapping way. Xie *et al.* (2020a) proposed deep learning network endowed with high-resolution feature blocks preserves spatial details during feature extraction, and its attention mechanism enhances representative features while suppressing noise.

Tang *et al.* (2019) developed skin lesion segmentation (SLS) method based on separable-Unet framework with stochastic weight that takes advantage of the separable convolutional block and *U*-net architectures, which have extremely captured the context feature channel correlation and higher semantic feature information to enhance the pixel-level discriminative representation capability of fully convolutional networks (FCN). Wang *et al.* (2019) presented bi-directional dermoscopic feature learning (biDFL) framework to model the complex correlation between skin lesions and their informative context. Nasr-Esfahani *et al.* (2019) proposed a new class of fully convolutional network (FCNs) with new dense pooling layers for segmentation of lesion regions in skin images called dense pooling convolutional network (DPFCN). The dense pooling layers which preserve their features rather than losing them by greater-than-one strides. Zhang *et al.* (2019) designed a deep learning framework to couple the FCN derived data driven features with hand-crafted texton features from a shallow network by introducing an integrating block trained in an end-to-end manner. The framework performed well on the skin lesion segmentation without the need of complicated data augmentation or comprehensive parameter tuning.

Lei *et al.* (2020) proposed a novel and effective generative adversarial network (GAN) to meet skin lesion challenges such as the large variations in terms of views and scales of lesion areas. Specifically, the GAN network architecture integrates two modules: a skip connection and dense convolution *U*-net (UNet-SCDC) based segmentation module and a dual discrimination (DD) module. Baig *et al.* (2020) presented a survey of more than 100 papers and comparative analysis of the state of the art techniques, model and methodologies. Hajabdollahi *et al.* (2020) developed a pruning framework to reduce the burden of the network for feature extraction through the selection of most informative color channels and simplification of the network. Ozturk and Ozkaya (2020) proposed improved FCN (iFCN) architecture

for the segmentation of full-resolution skin lesion images without any pre- or post-processing. It is to support the residual structure of the FCN architecture with spatial information.

Moreover, the authors have carried out intensive work in the field of skin lesion segmentation and classification. Several methods have been proposed based on Markov random field (MRF) and deep learning techniques. The MRF based techniques are a probabilistic graphical which extracts the image segmentation features as the prior information. The MRF based techniques attended to combine more features in their accounts to overcome all the shortcomings of the pixel-based MRF model, the region based MRF model, and the edge-based MRF model. For instance, some of the pitfalls include: the pixel-based MRF model which can use the macro texture pattern description to have interactions in a large neighborhood, the stochastic region-merging approach use the regular structural context to capture the fundamental relationship between regions more efficiently. For the reasons above, MRF-based techniques work combined the benefits of two or three models based MRF in one model (Salih and Viriri, 2018a;b; Salih *et al.*, 2019; Salih and Viriri, 2020).

Regarding deep learning techniques, (Adegun and Viriri, 2019a; 2020a) proposed a new framework that performs both segmentation and classification of skin lesions for automated detection of skin cancer. The network also integrates the Conditional Random Field (CRF) module which employs a linear combination of Gaussian kernels for its pair wise edge potentials for contour refinement and lesion boundaries localization. Adegun and Viriri (2020b) designed method performs pixel-wise classification of skin lesion images to identify melanoma pixels. The framework employs an end-to-end and pixel by pixels learning approach using Deep Convolutional Networks with softmax classifier. In Adegun and Viriri (2019b), a deep learning model has been adapted for the segmentation of skin lesions. The model demonstrates the segmentation of skin lesions using fully convolutional networks (FCNs) that train skin lesion images from end-to-end using only the images pixels and disease ground truth labels as inputs. The fully convolutional network adapted is based on *U*-net architecture. Adegun and Viriri (2018) presented a detailed and robust survey of the state-of-the-art algorithms and techniques for performing skin lesion segmentation. The approach used is the comparative analysis of the existing methods for skin lesion analysis, critical review of the performance evaluation of some recently developed algorithms for skin lesion images segmentation.

The proposed architecture aims to lower trainable parameters to reduce computational costs, making the system feasible for real-time medical diagnosis. The local binary convolution (LBC) layer has been used instead of the standard convolution to achieve these aims. Introducing the LBC layer to the proposed system also shows the difference between our previous work. Most of the techniques discussed above employ larger and more complicated deep learning architecture. Our proposed system is able to perform segmentation of skin lesion pixels using a moderate-size deep convolutional network. Some of these methods are also too slow and require huge computing processing resources for real-time medical analysis and diagnosis.

LOCAL BINARY CONVOLUTIONAL NEURAL NETWORKS (LBCNN)

A new research field is emerging, using binary weights in CNNs to reduce the network complexity (Courbariaux *et al.*, 2015; 2016; Rastegari *et al.*, 2016; Juefei *et al.*, 2017).

Juefei *et al.* (2017) has developed a network that was inspired by the famous texture descriptor local binary pattern (LBP) and presents a good approximation of a standard learnable convolutional layer. The idea of the method is to use LBC layers in convolutional neural networks as follows:

1. The first layer is a set of fixed sparsed pre-defined binary convolutional layer,

2. a non-linear activation function layer (sigmoid or ReLU, 2×2) is used,

3. the last layer is a set of learnable 1×1 convolutional weights.

The level of sparsity is defined by the user that indicates the percentage of non-zero value weights of binary convolutional layer. Then the first layer is initialized through Bernoulli distribution with $0, 1$ and -1 randomly using the sparsity. Assume that LBC has m pre-defined non-learnable binary filters and p, 1×1 convolution filters. The input image is filtered by m binary filters resulting in m difference maps. These maps are changed to m bit maps by a non-linear activation layer. Lastly, the m bit maps are linearly combined using the p learnable 1×1 weights to approximate the traditional convolutional layer. The weights in first binary convolutional layer are fixed while the weights in the second convolutional layer are learnable. As the method has sparsity in weights and the fixed binary values, it becomes less

representative compared to the typical convolutional layer. To achieve the similar performance with a traditional convolutional layer at each LBC, it needs a great number of local binary filters (512 in Juefei *et al.*, 2017) and 1×1 learnable weight. Fig. 2 shows one block in LBCNN architecture.

LBCNN networks was tested on datasets like: MNIST, Cifar-10 and SVHN for classification and was not tested on image segmentation. Therefore the objective of this paper is to study LBCNN with U-net's decoder for skin lesion segmentation, analyzing its parameters and the behaviour of the network with the insertion of noisy images, and to propose a pre-processing step on images to increase the accuracy of the model.

THE PROPOSED MODEL

This section presents the proposed framework in detail. A robust segmentation approach that overcomes the obstacles outlined above has been proposed through a reduced-size deep learning network using LBC convolution layer. We first introduce image pre-processing and data normalization technique. Second, the proposed system trains on skin lesion images dataset in an end-to-end process using a deep convolution-deconvolution network that uses an efficient alternative to convolutional layers in standard convolutional neural networks (CNN) which is called LBC. It accepts original image pixels and disease ground truth labels as input. The architectural diagram of deep convolutional network and the whole framework of the system is described in Figs. 3 and 4, respectively. The proposed model is composed of the following components as summarised below:

IMAGE SIZE AND NORMALIZATION

The dataset training set contains 2594 skin lesion images of different resolutions. Some lesion images' resolutions are above 1100×800, which needs a high cost of computation. It is essential to rescale the lesion images for the deep learning network. As directly resizing images may distort the skin lesion's shape, we first cropped the lesion image's center area and then proportionally resize the area to a lower resolution. The center square's size was set to be 0.75 of the image's height and automatically cropped concerning the image center. This approach not only enlarges the lesion area for feature detection but also maintains the shape of the skin lesion. We also use a relatively small image size of 256×256. This will reasonably reduce the input feature map size for the network. Finally, the images (X) are normalized by computing the mean
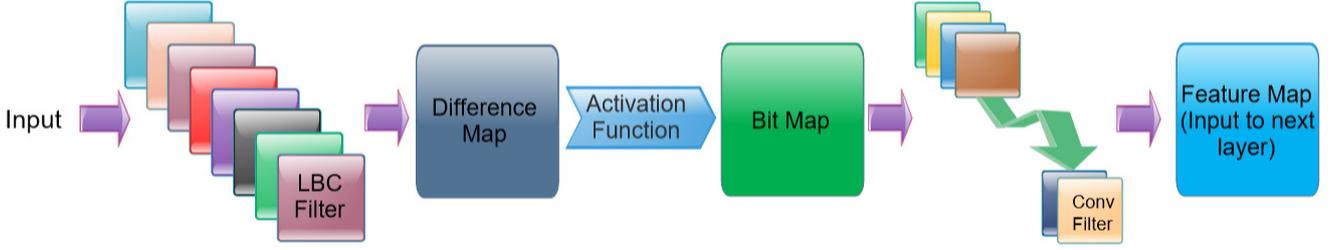
Fig. 2. *Simple block in LBCNNs. LBC filter is non-learnable weights and Conv filter is the learnable weights in the block..*
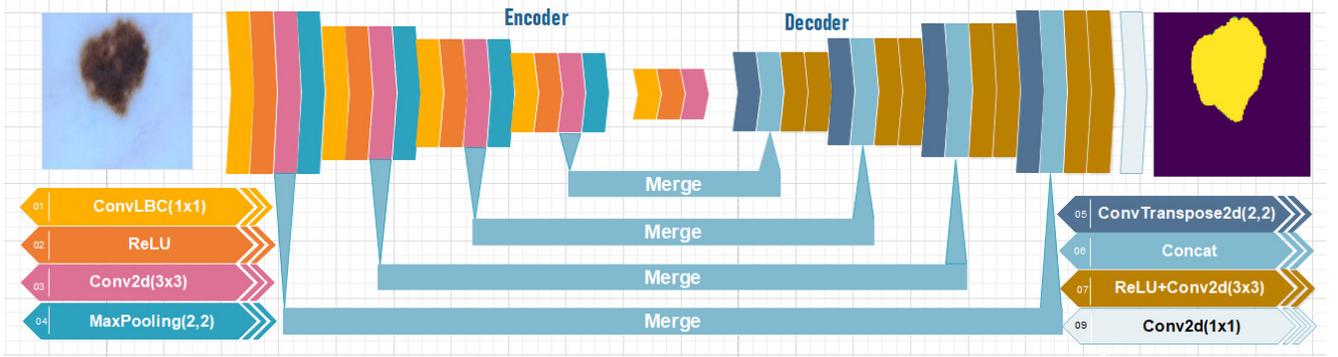

Fig. 3. *The proposed framework and flow of the Local binary convolution-deconvolution Network.*

pixel value $(\hat{X})$ and the standard deviation (σ) for data centering data normalization using the following Eq. 1.

$$X_{\text{Norm}} = \sum_i \frac{X_i - \hat{X}}{\sigma}. \tag{1}$$

DEEP CONVOLUTIONAL ENCODER-DECODER NETWORK

The deep convolution-deconvolution network employs multi-stage segmentation approach that automatically learns and segments complex features of skin lesion images. The encoder stage learns the contextual information by extracting discriminative features while the decoder stage captures the lesion boundaries of the skin images. This addresses issues with encoder-decoder network producing coarse segmented output with challenging skin lesions appearances such as low contrast between healthy and unhealthy tissues and fine grained variability. It also addresses issues with multi-size, multi-scale and multi-resolution skin lesion images. The deep convolution-deconvolution network also adopts a reduced-size network with LBC layers and four levels of encoding-decoding network. This reduces greatly the consumption of computational processing resources. In this paper, a framework that employs a robust convolution-deconvolution network architecture which adopts a multistage approach (with novel loss function with LBC weights to fully recover the missing features

in convolution-deconvolution network) to recover some of the missing feature has been proposed. The proposed deep convolution-deconvolution network architecture is made up of two major section namely the encoder and decoder.

Encoder

In the encoder part, an alternative approach to reducing the computational complexity of CNNs while performing as well as standard CNNs is used. The approach is called the local binary convolution neural network (LBCNN). It is an efficient alternative to convolutional layers in standard convolutional neural networks (CNN). The encoder part is composed of five blocks with each block composing of two convolutional layers (LBC layers [Non-tainable weights], convolutional layer [trainable weights]) and one down-sampling layer. The convolutional layer utilizes ReLU activation function. Convolutional network produces feature map through convolution process. The network consists of two type of filters non-trainable filter kernels and trainable filter kernels and non-linear activation function ReLU. The activation function defined in Eq. 2 is applied on the difference map and the filter kernel. This provides non-linearity to the network. Consecutive feature maps of convolution layer i is thus computes as

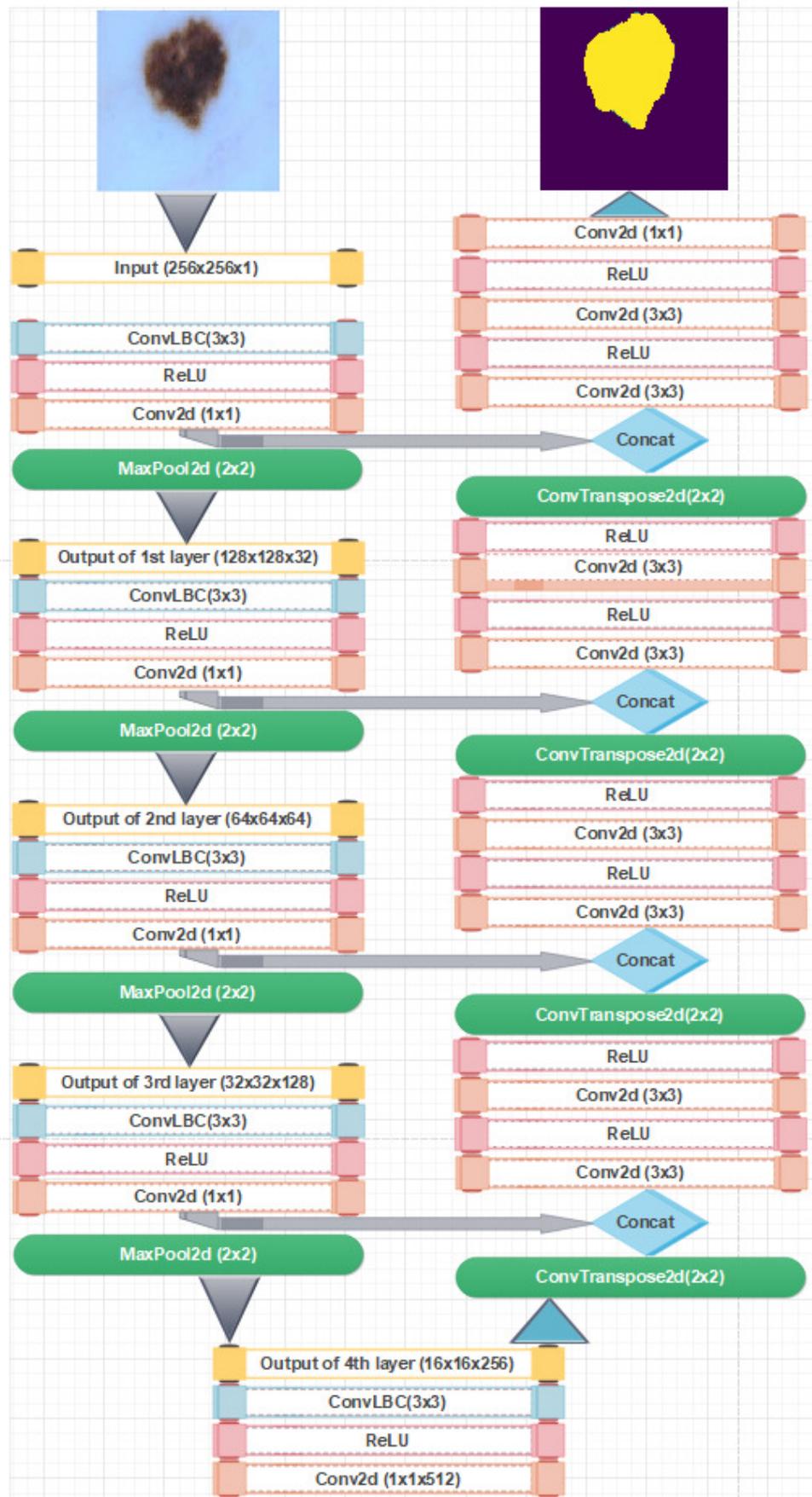$$F_i = \text{ReLU}(W * F_{i-1} + b_i), \tag{2}$$

Fig. 4. *Architectural Diagram for the Local binary convolution-deconvolution network.* .

where F_i is the feature map, F_{i-1} is the feature map for the previous layer, W is the filter kernel and b_i is the bias applied to each feature map of each layer. The ReLU activation function can be stated as

$$\text{ReLU}(x) = \max(0,x) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases} \quad (3)$$

where x is the resulting feature map. The down-sampling layer utilize the max-pooling function that reduces the sizes of the extracted feature maps. The max-pooling layer with size 2×2 pool over every pixel within a 2×2 area from the feature map. This results in reduction in the feature map size. A max-pooling layer performs down-sampling by breaking down the input feature map into pooling section, and computing the maximum of each section.

$$\text{Layer} = \text{maxPooling2dLayer}(poolSize). \quad (4)$$

Decoder

The decoder part also consist of five blocks with each block composing of two convolution layers, one up-sampling layer with concatenation function. This concatenates the up-sampling layer input with the corresponding feature map from the convolutional layer in the encoder part. The convolutional layers in this section also utilize ReLU activation function. The feature maps are convolved with the decoder filters in the convolutional layers to produce dense feature map. The decoders work to restore the feature maps to the original size in the network using the up-sampling layers with the function stated below.

$$y = \text{UpSample}(x,n), \quad (5)$$

where x is the feature map from the encoder and n is the up-sampling layer input. The up-sampling layer works in the opposite of the pooling layers which is to only restore image size resolution image by copying the pixel as many times as needed. It is an advanced technique of unpooling that reverts maxpooling operation by using the value and location of the maximum values in the maxpooling layers for feature maps restoration. There exists a skip connection in between the encoder and the decoder. The skip connection allow concatenation of the last convolution layers of the encoder with the first layer of the decoder. The encoded features are merged with the decoded features at a given spatial resolution. The skip connection is built with two convolutional layers.

The Proposed Model Algorithm

The proposed model steps are depicted in Algorithm 1. It starts with an inputs images y, then the encoder-decoder functions generate the skin lesion feature before sent to 1×1 convolution to get the segmented output. In the encoder part, the input image $\text{Feature}_{\text{mapP}}$, is first sent to the convolutional local binary layers as $\text{ConvLBC}(F_{\text{mapP}})$ and second to the ReLU activation function as $\text{ReLU}(\text{Diff}_{\text{map}})$, then to the convolutional layer with $\text{Conv}(\text{Bit}_{\text{map}})$ and finally to the down-sampled with the MaxPooling function as $\text{MaxPool}(\text{Feature}_{\text{map}})$. This process is done using for loop structure. The decoder part use the result from the encoder to achieve its goal. It passes into the for loops structure and goes through features up-sampling using $\text{Upsample}(\text{Feature}_{\text{Rel}_i})$ function. This is merged with corresponding layer from the encoder using $\text{Concat}(\text{Feature}_{\text{Up}}, \text{Feature}_{\text{map}})$. Then send it the convolution layer and the ReLU activation function and two times.

Multi-stage Approach

The convolution-deconvolution network adopts multi-stage approach in which the whole process is broken into units and the whole framework is made up of various stages. At the initial stage, the encoder which is majority made up of max-pooling and convolutional layers learns general appearance and localization information of the input image by capturing the semantic and contextual information. The max-pooling layers down-sample the feature maps from the input image to enable the convolutional layers extract the contextual information. In the later stage, the decoder which is made up of convolutional and up-sampling layers learns the characteristics of the lesion boundaries in recovering spatial information. The up-sampling layers restore the down-sampled feature maps for the lesion boundary localization. This process trains skin lesion images and disease labels to produce pixel-wise prediction. The combination of various components such as max-pooling and convolution layers in the encoder section and transit of features from the encoder section to the decoder and then with the up-sampling and convolutional layers in decoder perform well in achieving invariance but with some trade-off on lesion boundary localization accuracy. This sometimes results in Deep convolutional networks producing coarse segmented output. LBCNN layers is employed to improve the localization of lesions boundaries as well as to reduce the parameters for the network.

Reducing the encoder-decoder netowrk size, both encoding part and the decoding part are made up of five blocks each. Each block in the encoding part is composed of 3×3 non-trainable layers and 1×1 convolutional layers and one max pooling layer. In the decoding part, the previous block output is

up-sampled with the nearest neighbour using 2×2 convolutional layers and concatenated with the output from encoder part at the corresponding level. The 2×2 convolutional layers reduces the up-sampled feature maps number by half. All convolutional layers in each block of both the encoding and the decoding part are followed by rectified exponential linear unit (ReLU) activation function. ReLU increase training rate for large networks as it re-scale the feature maps. LBCNN layers are introduced to reduce the network size.

Algorithm 1 Proposed Model for the Deep Local Binary Convolution-Deconvolution Network

---

1: **procedure** ENCODER(Y) $\triangleright$ $Y : y_i$ is an input image with dimension (L, H, C).
2:     Initialization: $\text{Feature}_{\text{mapP}_0} = y_0$;
3:     **for** $(i = 0 : N - 1)$ **do**
4:         $\text{Diff}_{\text{map}_i} = \text{ConvLBC}(\text{Feature}_{\text{mapP}_i})$;
5:         $\text{Bit}_{\text{map}_i} = \text{ReLU}(\text{Diff}_{\text{map}_i})$;
6:         $\text{Feature}_{\text{map}_i} = \text{Conv}(\text{Bit}_{\text{map}_i})$;
7:         $\text{Feature}_{\text{mapP}_i} = \text{MaxPool}(\text{Feature}_{\text{map}_i})$;
8:         **if** $i <= N$ **then**
9:             $\text{Feature}_{\text{mapP}_{i+1}} = \text{Feature}_{\text{mapP}_i}$;
10:         **else**
11:             **Return** $\text{Feature}_{\text{mapP}_i}$;
12:         **end if**
13:     **end for**
14: **end procedure**
15: **procedure** DECODER($\text{Feature}_{\text{mapP}_i}$) $\triangleright$ $\text{Feature}_{\text{mapP}_i}$ is the downsampled feature maps.
16:     Initialization: $\text{Feature}_{\text{Rel}_{N-1}} = \text{Feature}_{\text{mapP}_i}$;
17:     **for** $(i = N - 1 : 0)$ **do**
18:         $\text{Feature}_{\text{Up}_i} = \text{Upsample}(\text{Feature}_{\text{Rel}_i})$;
19:         $\text{Feature}_{\text{Con}_i} = \text{Concat}(\text{Feature}_{\text{Up}_i}, \text{Feature}_{\text{map}_i})$;
20:         $\text{Feature}_{\text{Map}_i} = \text{Conv}(\text{Feature}_{\text{Con}_i})$;
21:         $\text{Feature}_{\text{Rel}_i} = \text{ReLU}(\text{Feature}_{\text{Map}_i})$;
22:         $\text{Feature}_{\text{Map}_i} = \text{Conv}(\text{Feature}_{\text{Rel}_i})$;
23:         $\text{Feature}_{\text{Rel}_i} = \text{ReLU}(\text{Feature}_{\text{Map}_i})$;
24:         **if** $i <= N$ AND $i! = 0$ **then**
25:             $\text{Feature}_{\text{Rel}_{i+1}} = \text{Feature}_{\text{Rel}_i}$;
26:         **else**
27:             **Return** $\text{Feature}_{\text{Rel}_i}$;
28:         **end if**
29:     **end for**
30: **end procedure**
31: Output= $\text{Conv}_{1 \times 1}(\text{Feature}_{\text{Rel}_i})$
32: Outputs = sigmoid(output) $\triangleright$ Final Segmented Output Display.

EXPERIMENTS AND RESULTS

Various experiments were carried out in this section to evaluate the performance of our proposed segmentation approach. Two publicly available datasets were used for empirical experiments. The results achieved were compared with the existing algorithms as shown in Tables 1 and 3.

DATASETS

The three well-established publicly available dataset used in the evaluation of the proposed segmentation method are from the ISIC 2018 (Tschandl *et al.*, 2018; Codella *et al.*, 2019), ISIC 2017 (Codella *et al.*, 2019) challenge in skin lesion segmentation and PH2 data repository (Mendonca *et al.*, 2013). PH2 contains 200 skin lesion images with highest resolution of 765×574 pixels. They were gotten at Dermatology Service of Hospital Pedro Hispano. The input dataset are skin lesion image in BMP format while the ground truth are mask image in BMP format as well.

ISIC 2018 (Tschandl *et al.*, 2018; Codella *et al.*, 2019) contains $2,594$ training images with ground truth provided by experts. The image size posses highest resolution of 1022×767. This dataset was categorized into training and testing image set both comprising of images and ground truth labels respectively. The input dataset are skin lesion images in JPG format while the ground truth are mask image in PNG format. The ground truth labels are provided for training and evaluating validation and test phase data using the performance evaluation metric. ISIC 2017 (Codella *et al.*, 2019) contains 2000 training images with the ground truth provided by experts. The image sizes possess highest resolution of 1022×767. This dataset was provided from the ISIC Dermoscopic Archive (Codella *et al.*, 2019). This dataset was categorized into training and testing image set both comprising of images and ground truth labels respectively. The input dataset are skin lesion images in JPEG format while the ground truth are mask image in PNG format. The ground truth labels are provided for training and evaluating validation and test phases data using the performance evaluation metrics.

EVALUATION METRIC CALCULATION

The most common skin lesion segmentation metrics were used for comparison including: Jaccard coefficient, DICE similarity coefficient and accuracy.

The Jaccard similarity coefficient (JI) compares similarity for the pixels in the ground truth and automatic segmentation to see which pixels are shared
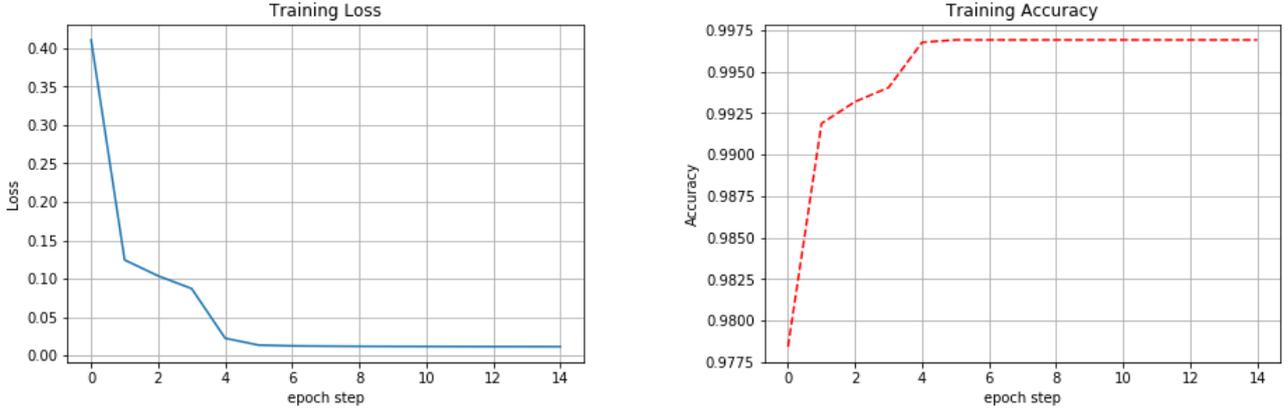
Fig. 5. *The loss and the accuracy training curves on ISIC 2018 dataset.*
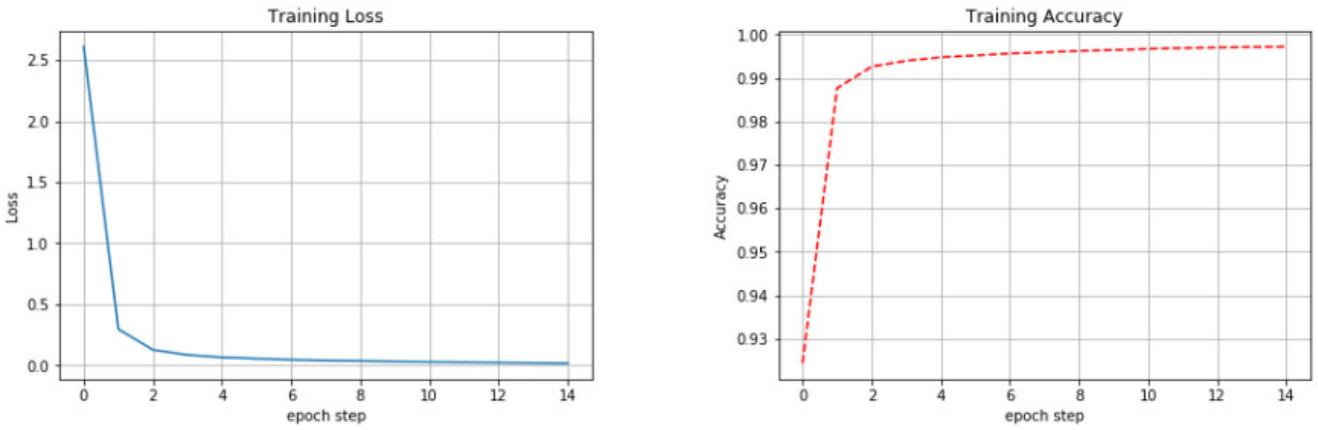

Fig. 6. *The loss and the accuracy training curves on ISIC 2017 dataset.*

and which are distinct. It is a measure of similarity for the two sets of data, with a range from 0% to 100%. The formula to find the index is:

$$\mathbf{JI} = \frac{|X \cap Y|}{|X \cup Y|}, \tag{6}$$

where X, Y are the number of pixels in the ground truth and automatic segmentation respectively.

DICE similarity coefficient (DICE) measure the similarity or overlap between the ground truth and automatic segmentation. It is defined as

$$\mathbf{DICE} = \frac{2TP}{FP \cup 2TP \cup FN}. \tag{7}$$

Accuracy (Acc) measures the proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$\mathbf{Acc} = \frac{TP \cup TN}{TP \cup TN \cup FP \cup FN}, \tag{8}$$

where FP is the number of false positive pixels, FN is the number of false negative pixels, TP is the number of true positive pixels and TN is the number of true negative pixels.

PERFORMANCE EVALUATION

The results show that the proposed system does not only perform well but also outperforms all existing methods. This is attributed mostly to the LBCNN layers segmentation approach in the encoder section (The evaluation metrics in subsection "Evaluation Metric Calculation" used to evaluate the proposed method performance). The performance metrics used are pixel level estimation. The pixel-level accuracy of segmentation is the estimation of how many lesion pixels are correctly identified when compared to the ground truth image. Jaccard index and DICE coefficient estimate the degree of overlapping of ground truth and segmented lesion region.

With regard to lesion segmentation process on the three datasets, first the proposed model was trained on ISIC 2018 dataset with 2594 training skin lesion images. It was tested on 260 skin lesion images. They were resized to 256×256 pixels. The results achieved accuracy, Jaccard index and DICE coefficient of 97.09%, 88.43% and 92.53% respectively even
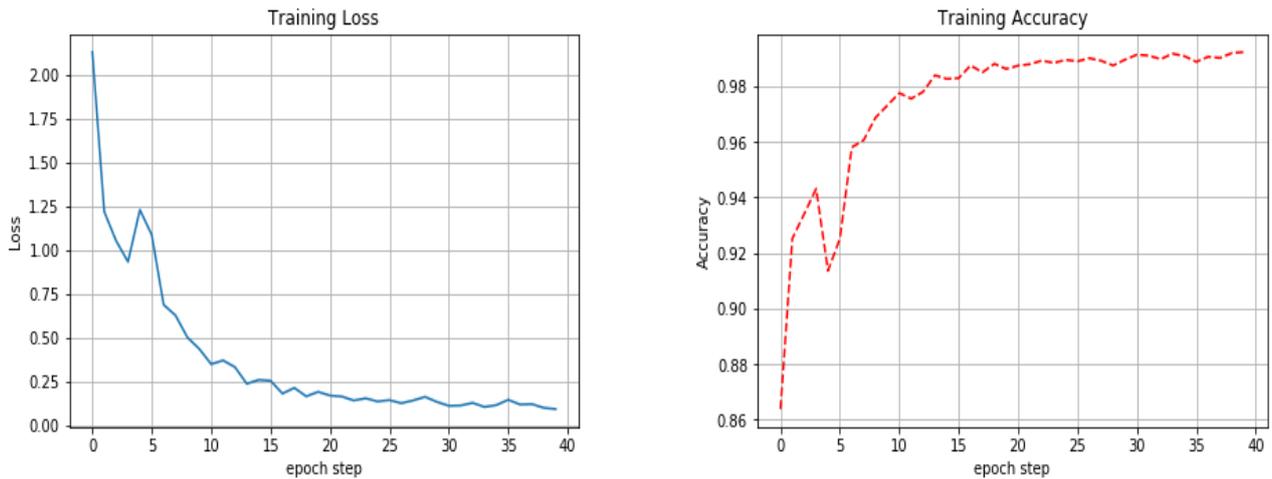
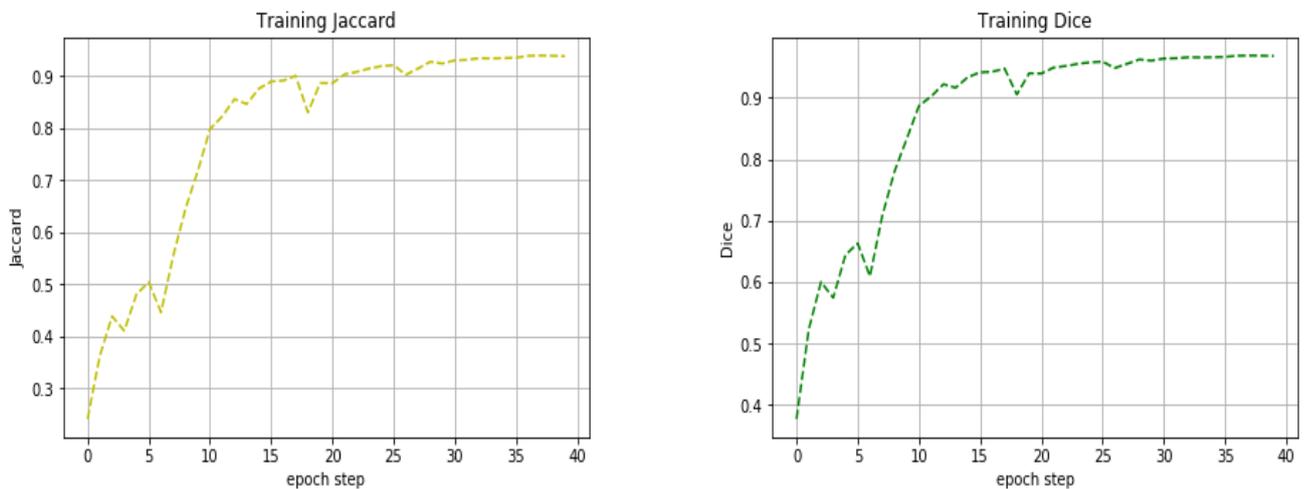Fig. 7. *The loss and the accuracy training curves on PH2 dataset.*

Fig. 8. *Jaccard Index and DICE coefficient Curves of the proposed method on the PH2 datasets.*

with very little training steps of 15 epochs as shown in Fig. 5. Secondly, the segmentation model was trained and evaluated on ISIC 2017 dataset containing 2000 images and 300 images for both training and testing tasks respectively with small training steps of 15 epochs as shown in Fig. 6. The evaluation was carried out on the ISIC 2017 dataset using metrics such as segmentation accuracy, Jaccard index, and DICE coefficient respectively, and the corresponding results are summarized in Table 2. As shown in Table 2, the proposed model achieved an accuracy of 96.89%, Jaccard index of 89.73%, and DICE Coefficient of 93.07% on the ISIC 2017. The proposed model outperformed all existing methods in Table 2.

Finally, training the proposed model on PH2 skin lesion image dataset requires more number of training steps because of dataset size. The model was trained on 200 images dataset and tested on 50 images as depicted in Figs. 7 and 8. In Table 3, the result

achieved accuracy, Jaccard index and DICE coefficient of 96.87% and 89.32% and 92.84% respectively with 40 training epochs. The results in Table 3 indicate that the proposed model is able to identify and differentiate higher number of affected skin lesions from the healthy tissues on both dataset.

The outputs displayed in Figs. 7 and 8 for PH2 dataset also show that the accuracy, the loss, Jaccard index and DICE score can still improve with increase in training steps and dataset. The robustness of the proposed segmentation method is shown in Figs. 9, 10 and 11 for ISIC 2018, ISIC 2017 and PH2 dataset respectively. The learning ability of the proposed model through experiment with the two datasets was evaluated with the accuracy curve in Figs. 5 and 7. The results from the curve clearly shows that the ISIC dataset with fairly large dataset reached the accuracy percentage of over 99% with lower training steps and epoch number.

179

COMPARISON

We compared the performance of the proposed model with the performance of the state-of-the-art methods from the latest study in literature such as *U*-net, FCN, CDNN, mFCN-PI, FrCN, Ensemble, MS-Unet, MS-Unet, LIN, biDFL, SLS, DCEDN, FCNDN, Res-Unet and MB-DCNN. This was carried out on the three datasets: ISIC 2018, ISIC 2017, and PH2 and the results are stated in Tables 1, 2 and 3, respectively. The outputs show that the proposed model outperform all the state-of-the-arts. From the results in Table 1, the proposed model records higher accuracy percentage and Jaccard index score of 97.09% and 88.43% on ISIC 2018 dataset when compared with other methods. It also shows higher DICE coefficient 92.53% as against the other methods. This output indicates that the proposed system is able to identify and differentiate higher number of affected skin lesions from the healthy tissues on ISIC 2018 skin lesion dataset. Table 2 shows our proposed model segmentation encoder-decoder network's performance and compares its performance with state-of-the-art. The evaluation was carried out on the ISIC 2017 dataset using metrics such as segmentation accuracy, Jaccard index and DICE coefficient, and the corresponding results are summarized in Table 2. As shown in Table 2, the proposed model achieved the highest an accuracy of 96.89%, Jaccard of 89.73% and DICE Coefficient of 93.07% when compared to with other methods. This result shows that the proposed segmentation system can detect and differentiate correctly diseased lesions from the healthy tissues on ISIC 2017 dataset as shown in Fig. 10. From the results in Table 3, the proposed model also gives higher accuracy percentage and Jaccard index score of 96.87% and 89.32% on PH2 dataset when compared with other methods. It also shows high DICE coefficient of 92.84% against some of the methods. These results again indicate that the proposed system is able to identify and differentiate higher number of affected skin lesions from the healthy tissues on PH2 skin lesion dataset.

DISCUSSION

In this paper, a deep convolutional-deconvolution network based architecture has been proposed for robust segmentation of skin lesions towards melanoma detection. This architecture adopts an enhanced deep convolutional network with a reduced size encoder-decoder network. The medium size network aims at minimizing the computational resources consumption. It also allows a multi-stage approach using LBC in a simpler pipeline process.

Table 1. *Segmentation performance (%) of the proposed model compared to the state-of-the-art for the ISIC 2018 test dataset.*

Method	Acc	JI	DICE
U-net (Ronneberger *et al.*, 2015)	90.14	61.64	76.27
FCN (Long *et al.*, 2015)	92.72	72.17	83.83
CDNN (Yuan *et al.*, 2017a)	93.40	76.50	84.90
mFCN-PI (Bi *et al.*, 2017)	95.51	84.64	91.18
FrCN (Al-Masni *et al.*, 2018)	94.03	77.11	87.08
Ensemble (Goyal *et al.*, 2020)	94.10	79.30	87.10
MS-Unet (Tang *et al.*, 2019)	95.87	85.34	91.47
LBCDN (this paper)	*97.09*	*88.43*	*92.53*

Table 2. *Segmentation performance (%) of the proposed model compared to the state-of-the-art for the ISIC 2017 test dataset.*

Method	Acc	JI	DICE
LIN (Li and Shen, 2018)	95.0	75.3	83.9
biDFL (Wang *et al.*, 2019)	94.65	81.47	88.54
SLS (Tang *et al.*, 2019)	94.31	79.26	86.93
FCNs (Zhang *et al.*, 2019)	92.73	72.94	81.81
DCEDN (Adegun and Viriri, 2019a)	95.00	–	92.00
FCNDN (Adegun and Viriri, 2020a)	95.50	–	92.1
Res-Unet (Zafar *et al.*, 2020)	–	77.2	85.8
MB-DCNN (Xie *et al.*, 2020b)	94.7	80.4	87.8
LBCDN (this paper)	*96.89*	*89.73*	*93.07*

Table 3. *Segmentation performance (%) of the proposed model compared to the state-of-the-art for the PH2 test dataset.*

Method	Acc	JI	DICE
U-net (Ronneberger *et al.*, 2015)	92.55	77.95	87.61
FCN (Long *et al.*, 2015)	92.82	80.22	89.03
mFCN-PI (Bi *et al.*, 2017)	94.24	83.99	90.66
FrCN (Al-Masni *et al.*, 2018)	95.08	84.79	91.77
Ensemble (Goyal *et al.*, 2020)	93.80	83.96	90.70
SRMP (Salih and Viriri, 2020)	91.51	78.35	89.65
Res-Unet (Zafar *et al.*, 2020)	––	85.4	92.4
LBCDN (this paper)	*96.87*	*89.32*	*92.84*

The Local binary convolutional-deconvolutional approach overcomes the limitation of deep convolutional networks in producing coarsely segmented outputs when processing challenging skin lesion images. In this approach, the whole network is divided into stages, with each stage handling a section of the segmentation process. The system also adopts the loss function that learn and compute losses from
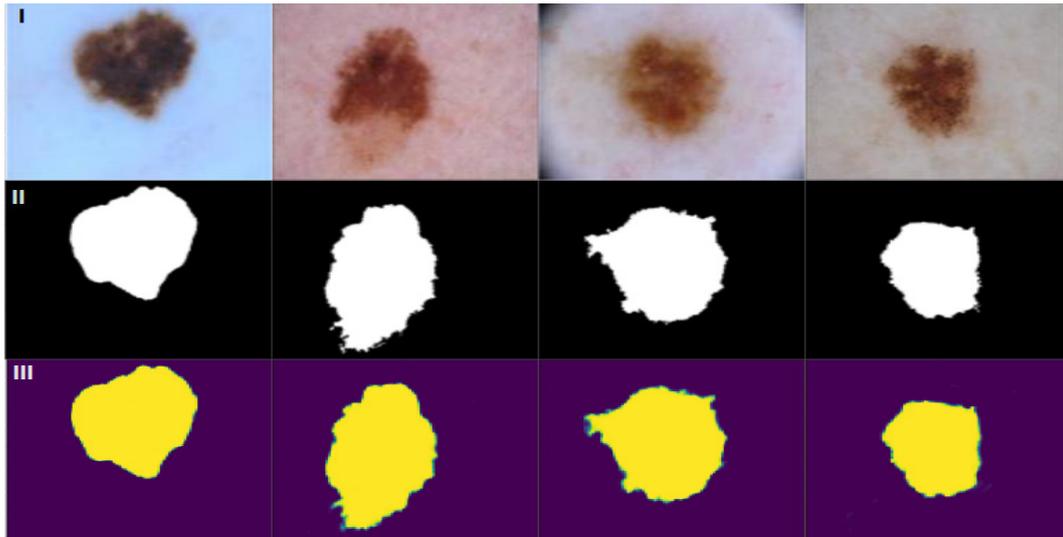
Fig. 9. *Some samples of skin lesion segmentation from ISIC 2018 dataset: I) Original Images, (II) Ground Truth images, (III) Segmented Images.*
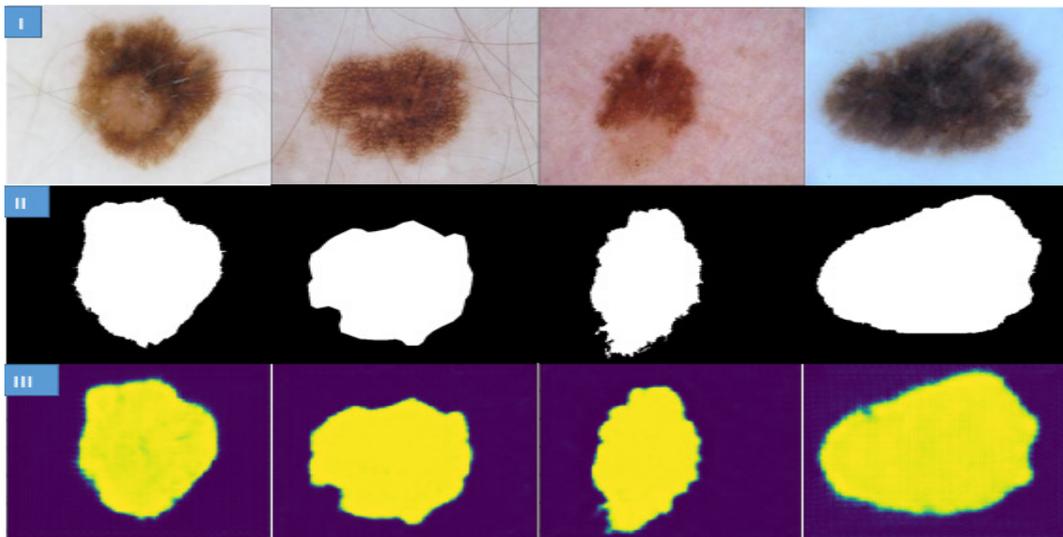

Fig. 10. *Some samples of skin lesion segmentation from ISIC 2017 dataset: I) Original Images, (II) Ground Truth images, (III) Segmented Images.*

the overlap in-between the predicted output and the ground truth label. It consumes lesser system resources since it does not perform sample re-weighting unlike some other loss function like cross entropy. The Local binary convolutional- deconvolutional approach achieves better segmentation by smoothing weak edges, irregular and fuzzy border and well-defined lesion boundary. The proposed model achieved an overall accuracy, Jaccard index and DICE coefficient of 97.09%, 88.43% and 92.53% respectively on ISIC 2018 skin lesion image dataset. It shows the high accuracy, Jaccard index and DICE coefficient of 96.89%, 89.73% and 93.07% respectively on ISIC 2017 skin lesion image dataset. The LBCDN shows an accuracy, Jaccard index and DICE coefficient

of 96.87%, 89.32% and 92.84% respectively on PH2 skin lesion image dataset. It can be inferred from the proposed system evaluation that the techniques is promising and can outperform all existing state-of-the-art.

The proposed method achieved the largest DICE coefficient, which indicates that the proposed method the strongest over all performance when compared to the other methods. The proposed method can be part of a system designed to assist automatic skin lesions in the field of sector health. Skin lesion experimental results using the proposed method indicate that, the proposed method potentially can provide more accurate skin lesions segmentation from images than
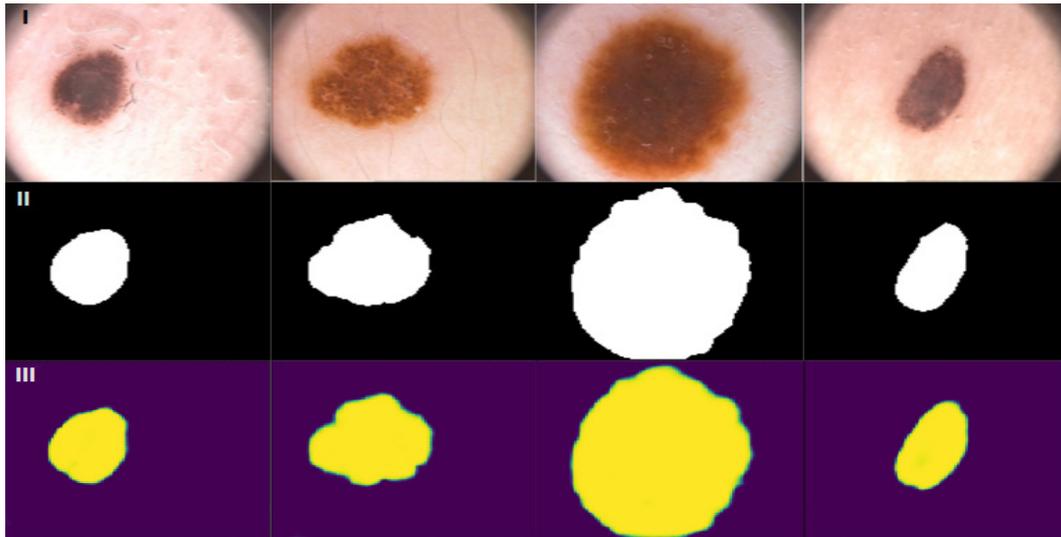
Fig. 11. *Some samples of skin lesion segmentation from PH2 dataset: I) Original Images, (II) Ground Truth images, (III) Segmented Images.*

comparable methods existing in the literature. We plan to employ the method in a full automatic skin lesion detection. This method will assist dermatologists to automatically locate region of skin lesion for more diagnosis.

The proposed system's best performance is mostly due to the local binary convolution descriptor (LBC layer) which has been introduced to the system architecture in the encoder stage. The local binary pattern descriptor (LBC layer) played a significant role in improving our deep learning architectures. This is the first introduction to the LBC layer in deep learning technique for skin lesion segmentation to the best of our knowledge. However, the performance improvement achieved using the local binary descriptor triggers further questions about the suitability of other descriptor methods for deep learning architectures. For instance, the local ternary pattern (LTP) and the completed local binary pattern (CLBP) can be introduced to the deep learning architectures with a promising expectation of having more accurate results.

CONCLUSION

Skin lesion segmentation is an essential step in developing a computer-aided diagnosis system for skin lesions. This paper successfully developed an architecture (LBCDN) based on deep learning for robust skin lesion segmentation by introducing a local binary convolution layer instead of the standard convolution layer, to effectively improving the accuracy considerably. This architecture proposed an enhanced deep convolutional network with a reduced size encoder-decoder system, making the LBCDN architecture minimize computational cost. The reduced size contribution of the LBCDN architecture is mostly from introducing the local binary convolution descriptor (LBC layer) into the LBCDN encoder stage. The proposed model architecture has been trained and evaluated on the ISIC 2018 dataset, ISIC 2017 dataset and PH2 dataset. The proposed architecture achieved an overall accuracy, the Jaccard index, and the DICE coefficient of 97.09%, 88.43%, and 92.53%, respectively on ISIC 2018 skin lesion image dataset. It shows highest results on ISIC 2017 dataset as well. It also achieved accuracy, Jaccard index, and DICE coefficient of 96.87%, 89.32%, and 92.84% respectively on the PH2 skin lesion image dataset. The LBCDN architecture achieves the largest DICE coefficient which indicates that the proposed architecture is the best method most robust overall performance compared to all of the existing methods. It can also be concluded that the LBCDN architecture overcomes the limitation of deep convolutional-deconvolutional networks in producing coarsely segmented outputs when processing challenging skin lesion images. Finally, this architecture can be integrated into a full CAD system for automatic skin lesion detection.

REFERENCES

Adegun A, Viriri S (2018). Skin lesion images segmentation: A survey of the state-of-the-art. Lect Notes Comput Sci 11308:321–30.

Adegun AA, Viriri S (2019a). Deep learning-based system

for automatic melanoma detection. IEEE Access 8:7160–72.

Adegun A, Viriri S (2019b). Deep learning model for skin lesion segmentation: Fully convolutional network. Lect Notes Comput Sci 11663:232–42.

Adegun AA, Viriri S (2020a). FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. IEEE Access 8:150377–96.

Adegun A, Viriri S (2020b). Deep convolutional network-based framework for melanoma lesion detection and segmentation. Lect Notes Comput Sci 12002:51–62.

Al-Masni MA, Al-antari MA, Choi MT, Han SM, Kim TS (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Comput Meth Prog Bio 162:221–31.

Al-Masni MA, Kim DH, Kim TS (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput Meth Prog Bio 190:105351–62.

Baig R, Bibi M, Hamid A, Kausar S, Khalid S (2020). Deep learning approaches towards skin lesion segmentation and classification from dermoscopic images – A review. Curr Med Imaging Rev 16:513–33.

Bi L, Kim J, Ahn E, Kumar A, Fulham M, Feng D (2017). Dermoscopic image segmentation via multistage fully convolutional networks. IEEE T Biomed Eng 64:2065–74.

Bi L, Kim J, Ahn E, Feng D, Fulham M (2016). Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata. In: Proc 13th IEEE Int Symp Biomed Imaging 1059–62.

Celebi ME, Kingravi HA, Uddin B, Iyatomi H, Aslandogan, YA, Stoecker, WV, Moss, RH (2007). A methodological approach to the classification of dermoscopy images. Comput Med Imaging Graph 31:362–73.

Codella N, Rotemberg V, Tschandl P, Celebi ME, Dusza S, Gutman D, Kittler H, others (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC). arXiv preprint arXiv:1902.03368.

Courbariaux M, Hubara I, Soudry D, El-Yaniv R, Bengio Y (2016). Binarized neural networks: Training deep neural networks with weights and activations constrained to $+1$ or -1. arXiv preprint arXiv:1602.02830.

Courbariaux M, Bengio Y, David JP (2015). Binaryconnect: Training deep neural networks with binary weights during propagations. In: Cortes C, Lee DD, Garnett R, Lawrence ND, Sugiyama M, eds. Proc 29th Ann Conf Neural Inf Proc Syst 1:3123–31.

Dhanachandra N, Manglem K, Chanu YJ (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. Procedia Comput Sci 54:764–71.

Eltayef KAA (2017). Segmentation and lesion detection in dermoscopic images. PhD thesis. London: Brunel University.

Emre CM, Wen Q, Hwang S, Iyatomi H, Schaefer G (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods. Skin Res Technol 19:e252–58.

Feng S, Zhao H, Shi F, Cheng X, Wang M, Ma Y, Xiang D, Zhu W, Chen X (2020). CPFNet: Context pyramid fusion network for medical image segmentation. IEEE T Med Imaging 39:3008–18.

Goyal M, Oakley A, Bansal P, Dancey D, Yap MH (2020). Skin lesion segmentation in dermoscopic images with ensemble deep learning methods. IEEE Access 8:4171–81.

Hajabdollahi M, Esfandiarpoor R, Khadivi P, Soroushmehr SM, Karimi N, Samavi S (2020). Simplification of neural networks for skin lesion image segmentation using color channel pruning. Comput Med Imaging Graph 8:1017–29.

Hasan MK, Dahal L, Samarakoon PN, Tushar FI, Marti R (2020). DSNet: Automatic dermoscopic skin lesion segmentation. Comput Biol Med 2:103738–62.

He X, Yu Z, Wang T, Lei B (2017). Skin lesion segmentation via deep RefineNet. Lect Notes Comput Sci 10553:303–11.

Juefei-Xu F, Naresh Boddeti V, Savvides M (2017). Local binary convolutional neural networks. In: Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 4284–93.

Kasmi R, Mokrani K, Rader RK, Cole JG, Stoecker, WV (2016). Biologically inspired skin lesion segmentation using a geodesic active contour technique. Skin Res Tech 22:208–22.

Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N (2017). Deep learning in medical imaging: General overview. Korean J Radiol 18:570–84.

Lei B, Xia Z, Jiang F, Jiang X, Ge Z, Xu Y, Qin J, Chen S, Wang T, Wang S (2020). Skin Lesion segmentation via generative adversarial networks with dual discriminators. Med Image Anal 23:101716–44.

Li H, He X, Zhou F, Yu Z, Ni D, Chen S, Wang T, Lei B (2019). Dense deconvolutional network for skin lesion segmentation. IEEE J Biomed Health Inform 23:527–37.

Li Y, Shen L (2018). Skin lesion analysis towards melanoma detection using deep learning network. J Sens 18:556–70.

Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van DL, Jeroen A, Van GB, Sánchez CI (2017). A survey on deep learning in medical image analysis. Med Image Anal 42:60–88.

Long J, Shelhamer E, Darrell T (2015). Fully convolutional networks for semantic segmentation. In: Proc IEEE Conf Comput Vision Pattern Recogn (CVPR) 3431–40.

Mendonca T, Ferreira PM, Marques JS, Marcal AR, Rozeira J (2013). PH 2-A dermoscopic image database for research and benchmarking. Proc 35th Ann Int Conf IEEE Eng Med Biol Soc (EMBC) 5437–40.

Nasr-Esfahani E, Rafiei S, Jafari MH, Karimi N, Wrobel JS, Samavi S, Soroushmehr SR (2019). Dense pooling layers in fully convolutional network for skin lesion segmentation. Comput Med Imaging Graph 78:1016–58.

Nguyen NQ, Lee SW (2019). Robust boundary segmentation in medical images using a consecutive deep encoder-decoder network. IEEE Access 7:33795–808.

Nguyen Q, Lee SW (2018). Colorectal segmentation using multiple encoder-decoder network in colonoscopy images. In Proc IEEE 1st Int Conf Artif Intel Knowl Eng (AIKE) 208–11.

Okuboyejo D, Olugbara OO, Odunaike S (2014). Unsupervised Restoration of Hair-Occluded Lesion in Dermoscopic Images. In: Proc 18th Conf Medical Image Underst Anal 91–96.

Ozturk S, Ozkaya U (2020). Skin lesion segmentation with improved convolutional neural network. J Digit Imaging 33:958–70.

Pennisi A, Bloisi DD, Nardi D, Giampetruzzi AR, Mondino C, Facchiano A (2016). Skin lesion image segmentation using Delaunay triangulation for melanoma detection. Comput Med Imaging Graph 52:89–103.

Premaladha J, Ravichandran KS (2016). Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms. J Med Syst 40:96–107.

Rajab MI, Woolfson MS, Morgan SP (2004). Application of region-based segmentation and neural network edge detection to skin lesions. Comput Med Imaging Graph 28:61–68.

Rastegari M, Ordonez V, Redmon J, Farhadi A (2016). Xnor-net: Imagenet classification using binary convolutional neural networks. Lect Notes Comput Sci 9908:525–42.

Ronneberger O, Fischer P, Brox T (2015). *U*-net: Convolutional networks for biomedical image segmentation. Lect Notes Comput Sci 9351:234–41.

Salih O, Viriri S (2020). Skin lesion segmentation using

stochastic region-merging and pixel-based Markov random field. Symmetry 12:1224–35.

Salih O, Viriri S, Adegun A (2019). Skin lesion segmentation based on region-edge Markov random field. Lect Notes Comput Sci 11845:407–18.

Salih O, Viriri S (2018a). Skin cancer segmentation using a unified Markov random field. Lect Notes Comput Sci 11241:25-33.

Salih O, Viriri S (2018b). Skin lesion segmentation using enhanced unified Markov random field. Lect Notes Comput Sci 11308:331–40.

Sainath TN, Mohamed AR, Kingsbury B, Ramabhadran B (2013). Deep convolutional neural networks for LVCSR. In: Proc 2013 IEEE Int Conf Acoust Speech Signal Proces 8614–18.

Sathya B, Manavalan R. (2011). Image segmentation by clustering methods: performance analysis. Int J Comput Appl 29:27–32.

Schaefer G, Rajab MI, Celebi ME, Iyatomi H (2009). Skin lesion segmentation using co-operative neural network edge detection and colour normalisation. In: Proc 9th Int Conf Inform Tech Appl Biomed 1–4.

Sharma S, Ball JE, Tang B, Carruth DW, Doude M, Islam MA (2019). Semantic segmentation with transfer learning for off-road autonomous driving. J Sens 19:2577–97.

Silveira M, Nascimento JC, Marques JS, Marçal RS, 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 3:35–45.

Tang Y, Yang F, Yuan S (2019). A multi-stage framework with context information fusion structure for skin lesion segmentation. In: Proc 16th IEEE Int Symp Biomed Imaging (ISBI 2019) 1407–10.

Tang P, Liang Q, Yan X, Xiang S, Sun W, Zhang D, Coppola G (2019). Efficient skin lesion segmentation using separable-Unet with stochastic weight averaging. Comput Methods Programs Biomed 178:289–301.

Tschandl P, Rosendahl C, Kittler H (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5:180–61.

Vesal S, Ravikumar N, Maier A (2018). Skinnet: A deep learning framework for skin lesion segmentation. In: Proc 2018 IEEE Nucl Sci Symp Med Imaging Conf (NSS/MIC) 1–3.

Wang X, Jiang X, Ding H, Liu J (2019). Bi-directional dermoscopic feature learning and multi-scale consistent decision fusion for skin lesion segmentation. IEEE Trans Image Process 29:3039–51.

Wighton P, Sadeghi M, Lee TK, Atkins MS (2009). A fully automatic random walker segmentation for skin lesions in a supervised setting. Lect Notes Comput Sci 5762:1108–15.

Wong A, Scharcanski J, Fieguth P (2011). Automatic skin lesion segmentation via iterative stochastic region merging. IEEE Trans Biomed Eng 15:929–36.

Xie F, Bovik AC (2013). Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. Pattern Recogn 46:1012–19.

Xie F, Yang J, Liu J, Jiang Z, Zheng Y, Wang Y. Skin lesion segmentation using high-resolution convolutional neural network (2020). Comput Meth Program Biomed 186:1052–41.

Xie Y, Zhang J, Xia Y, Shen C (2020). A mutual bootstrapping model for automated skin lesion segmentation and classification. IEEE T Med Imaging 39:2482–93.

Yuan X, Situ N, Zouridakis G (2009). A narrow band graph partitioning method for skin lesion segmentation. Pattern Recogn 42:1017–28.

Yuan Y, Chao M, Lo YC (2017). Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans Med Imaging 36:1876–86.

Yuan Y (2017). Hierarchical convolutional-deconvolutional neural networks for automatic liver and tumor segmentation. arXiv preprint arXiv:1710.04540.

Zafar K, Gilani SO, Waris A, Ahmed A, Jamil M, Khan MN, Sohail Kashif A (2020). Skin lesion segmentation from dermoscopic images using convolutional neural network. J Sens 20:1601–14.

Zhang L, Yang G, Ye X. Automatic skin lesion segmentation by coupling deep fully convolutional networks and shallow network with textons (2019). J Med Imaging 6:024001–12.