2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

An efficient CNN based algorithm for detecting melanoma cancer regions in H&E-stained images

Salah Alheejawi[1], Richard Berendt[2], Naresh Jha[3], Santi P. Maity[4] and Mrinal Mandal[5]*

***Abstract*— Histopathological images are widely used to diagnose diseases such as skin cancer. As digital histopathological images are typically of very large size, in the order of several billion pixels, automated identification of abnormal cell nuclei and their distribution within multiple tissue sections would enable rapid comprehensive diagnostic assessment. In this paper, we propose a deep learning-based technique to segment the melanoma regions in Hematoxylin and Eosin-stained histopathological images. In this technique, the nuclei in an image are first segmented using a deep learning neural network. The segmented nuclei are then used to generate the melanoma region masks. Experimental results show that the proposed method can provide nuclei segmentation accuracy of around 90% and the melanoma region segmentation accuracy of around 98%. The proposed technique also has a low computational complexity.**

***Keywords*— *Histopathological image analysis, Nuclei segmentation, Melanoma detection, Deep learning*.**

I. Introduction

The Cutaneous Malignant Melanoma (CMM) is an aggressive type of skin cancer. The early diagnosis of CMM is very important as it helps increasing the chances of successful treatment and survival rate [1]. The gold standard for the diagnosis is the examination of histopathological images. The digitized histopathological slides, which are typically obtained by scanning the stained biopsy slides of the skin tissue, can provide the cell morphological features with high resolution. The digitized slides, also known as Whole Slide Images (WSIs), can enable computer aided diagnosis and assist the pathologist to achieve accurate and faster diagnosis [2,3]. The Hematoxylin and Eosin (H&E) stain is widely used in histopathology as the morphological features of the cells become vividly clear. Fig. 1 shows an H&E-stained histopathological image of a CMM (contoured in green) within a cross section of a skin tissue slide.

The deep learning algorithms based on Convolutional Neural Networks (CNN) have recently been used successfully in medical image analysis. Badrinarayanan et al. [4] proposed the SegNet architecture for object segmentation. Ronneberger et al. [5] proposed the U-Net architecture for biomedical image segmentation. Alheejawi et al. [6] proposed a technique (henceforth referred to as the NS-Net+FC technique) to segment and classify the cell nuclei on the H&E-stained images. The nuclei segmentation (NS) is performed using the NS-Net architecture that contains 5 convolutional layers with different filter sizes. From each segmented nucleus, 54 hand-crafted features are extracted, and the feature vector is fed to a Support Vector Machine (SVM) for the nuclei classification (melanoma or non-melanoma nuclei). Sabol et al. [7] proposed a patch-based segmentation technique (henceforth referred to as the PBS-ResNet technique) to detect colorectal cancer regions in H&E-stained WSIs. This technique divides a WSI into nonoverlapping patches of size 64x64 pixels. The ResNet-50 architecture is then used to classify the patches. Although, the patch-based segmentation is computationally fast [8], the technique might generate many false positive patches.

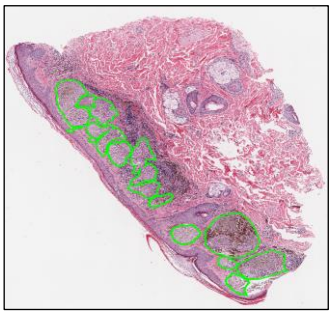

Figure 1. Example of an H&E-stained skin tissue WSI.

In this paper, we propose an automated technique to generate melanoma region masks in H&E-stained images. In the proposed technique, the nuclei are first segmented into melanoma and non-melanoma classes using a CNN. The CMM regions are then generated by applying morphological operations on the melanoma cells.

The organization of the paper is as follows. Section II describes the proposed technique in detail. Section III presents the performance evaluation, followed by the conclusion in Section IV.

II. Proposed Technique

The schematic of the proposed technique is shown in Fig. 2, which consists of two modules: CNN-based nuclei segmentation and melanoma region detection. The details of these two modules are presented in the following.

Figure 2. Schematic of the proposed melanoma detection technique.

* Corresponding author.
S. Alheejawi and M. Mandal are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada (e-mail: alheejaw@ualberta.ca and mmandal@ualberta.ca).
R. Berendt and N. Jha are with the Department of Oncology, Cross Cancer Institute, University of Alberta, Edmonton, Canada (e-mail: {richard.berendt, naresh.jha}@albertahealthservices.ca).
S. P. Maity is with Indian Institute of Engineering Science and Technology (IIEST), Shibpur, India (email: santipmaity@it.iiests.ac.in).

A. CNN-based Nuclei Segmentation

In this module, the H&E-stained image is segmented into three regions: melanoma nuclei, non-melanoma nuclei and background regions. The segmentation is done by using the proposed CNN architecture shown in Fig. 3. The CNN architecture has 25 convolution layers (compared to 5 layers in the NS-Net [6] architecture) used for nuclei segmentation. The proposed architecture, henceforth referred to as the improved NS-Net (INS-Net), consists of two paths: path A and path B. Out of 25 layers, 11 layers are in the path A, and 12 layers are in path B. Note that there are five Skip Connections (Skip Conn-*m,* where *m*=1,2,3,4,5) in the proposed INS-Net to reduce the gradient vanishing impact [9]. There are 6 types of layers in INS-Net architecture and the details of these layers are explained below.

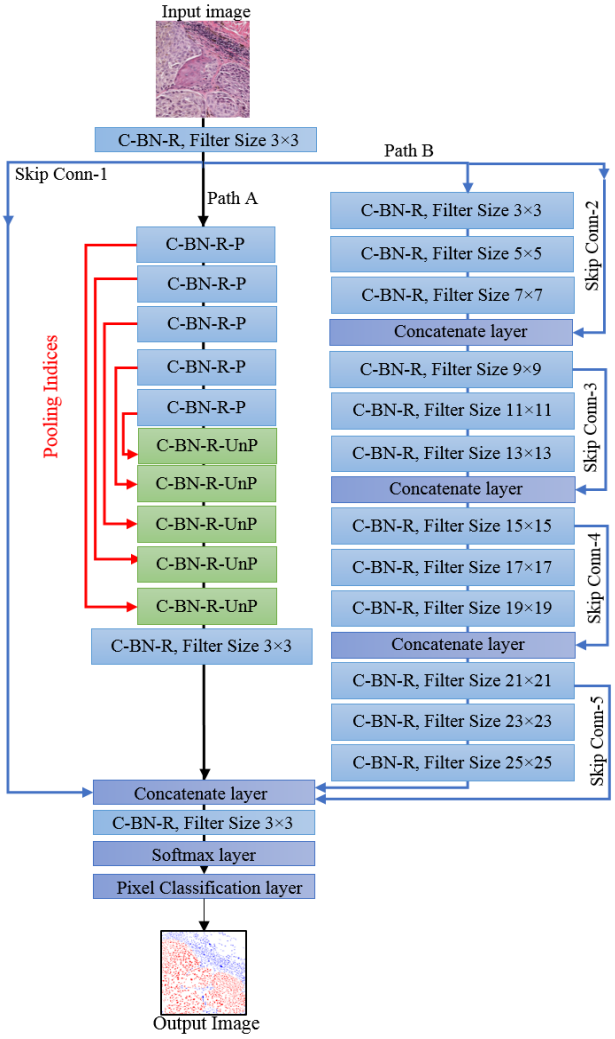

Figure 3. The proposed INS-Net CNN architecture. Note that the output image pixels are classified into 3 types: melanoma (red), non-melanoma nuclei (blue), and background(white) pixels.

1. C-BN-R layer: This layer includes three operations: Convolution, Batch Normalization, and nonlinear activation using Rectified Linear Units (ReLU).
2. C-BN-R-P layer: This layer consists of a C-BN-R layer followed by a Pooling layer to reduce the size (width x height) of the feature maps by selecting a pooling window and a stride value. In this paper, we have used MAX pooling, with a window size of 2x2 and a stride of 2x2. Note that the index of the maximum value is sent to the corresponding C-BN-R-UnP layer.
3. C-BN-R-UnP layer: This layer consists of a C-BN-R layer followed by an UnPooling layer to upsample the feature map size (width x height). In this paper, an upsampling factor of 2x2 is used. The upsampling is done using the bilinear interpolation. The corresponding index from the C-BN-R-P layer is used to determine the location of the pixel that needs to be upsampled.
4. Concatenate layer: This layer combines the feature maps of same size (width x height) from two or more input layers. The depth of the output layer is equal to the sum of the depths of the input layers. In Fig. 3, the last concatenate layer concatenates the output of paths A and B.
5. SoftMax layer: This layer finds the probability values of each pixel corresponding to *K* classes. Note that each pixel will have *K* probability values (In this paper, *K*=3).
6. Pixel classification layer: This layer selects the class with the highest value of SoftMax output for each pixel.

In the proposed INS-Net, the Path B is used to extract the fine features of the cells and the background, whereas the Path A is used to extract the coarse features of the cells and background regions. The INS-Net architecture is used to segment the H&E-stained image pixels into melanoma nuclei, non-melanoma nuclei, and background pixels. The nuclei segmentation results are very important to determine the melanoma region masks in H&E-stained images. These masks can help doctors in determining the depth of melanoma invasion and in grading the melanoma. In the next section, the nuclei segmentation results will be used to determine the melanoma region masks.

A. Melanoma Region Detection (MRD)

This module is used to determine the melanoma regions (MR) from the nuclei segmentation masks obtained using the INS-Net architecture. The MR detection consists of several morphological operations applied on the melanoma nuclei mask. Note that the melanoma nuclei mask is a binary mask representing the melanoma nuclei class obtained from the nuclei mask. The MR detection is done using the following steps.

1. The isolated melanoma nuclei (considered as foreground) in a neighborhood are merged using the binary morphological dilation operation. In this work, a disc structuring element with 5-15 pixels radius is used.
2. Sometimes, there are small, isolated background pixels inside the merged melanoma regions, which can be considered as noise. An image fill operation is performed to change these isolated background pixels to the foreground pixels.
3. During the dilation operation (step 1), the size of an object typically increases. To nullify the increase in size, an erosion operation is performed. The structuring element is same as that used in step 1.
4. Connected melanoma regions with small areas are removed by using an area threshold. In this paper, an area

threshold of 600 pixels (equivalent to 150 μm^2 area) has been used.

Note that a variant of the MR generation technique was also implemented by applying the *image closing* operation followed by image fill and area threshold. But the overall performance is not as good as the proposed MRD module.

III. Results and Discussions

In this section, performance of the proposed technique is presented and compared with the state-of-the-art techniques. The nuclei segmentation performance is presented first followed by the melanoma region detection performance.

The performance is evaluated using an image dataset consisting of 100 H&E-stained 960 x 960 RGB images extracted from 4 WSIs of skin tissue. Four-fold cross validation is used. For each fold, one WSI is arbitrarily chosen as the testing WSI, and 25 H&E-stained images from this WSI are used as the testing dataset. The remaining 75 ROI windows (from the other 3 WSIs) are used as the training dataset. This process is repeated four times for the four-fold cross validation, and four sets of performance measures are obtained. The average performance is then calculated.

A. Nuclei segmentation performance

The nuclei segmentation performance of the INS-Net is evaluated and compared with SegNet, U-Net and NS-Net architectures. The PBS-ResNet technique is not compared as it does not perform nuclei segmentation. The segmentation performance is evaluated using Accuracy, and Dice Coefficient measures defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$Dice\ \ Coefficient = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\%; \quad Recall = \frac{TP}{TP + FN} \times 100\%$$

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

TABLE I. Nuclei segmentation performance

Techniques	Accuracy	Dice Coefficient
SegNet [4]	87.84	85.81
U-Net [5]	78.79	69.63
NS-Net [6]	90.21	88.52
INS-Net	94.12	89.17

Table I shows the segmentation performance of different techniques. In this work, the SegNet, U-Net and NS-Net and the proposed INS-Net architectures are trained with the same number of training images (i.e., 75 H&E-stained images). Fig. 4 shows visual examples of melanoma nuclei classification results using the NS-Net+FC, and the proposed INS-Net compared with the ground truth image. Note that for the SegNet, U-Net and NS-Net techniques, the nuclei are first segmented using the CNNs and detected nuclei are then classified using an SVM classifier [6]. The SVM uses a feature vector of length 54 (18 first-order features, 9 Histogram of Oriented Gradient features, 24 Haralick texture features and 3 Morphological features) calculated for each nucleus. In the proposed (INS-Net) technique, the nuclei segmentation and classification are integrated in a single CNN architecture. In Fig. 4(c), it is observed that the INS-Net has superior nuclei classification results compared to the NS-Net+FC outputs shown in Fig. 4 (b). The objective nuclei classification performance is shown in Table II. It is observed that the proposed INS-Net provides a superior classification performance compared to the NS-Net+FC architecture. The mis-classification results of the melanoma nuclei can impact negatively on the obtained melanoma regions detected by the MRD module. Note that accurate melanoma nuclei classification results are important to obtain a good MR detection performance by the MRD module.

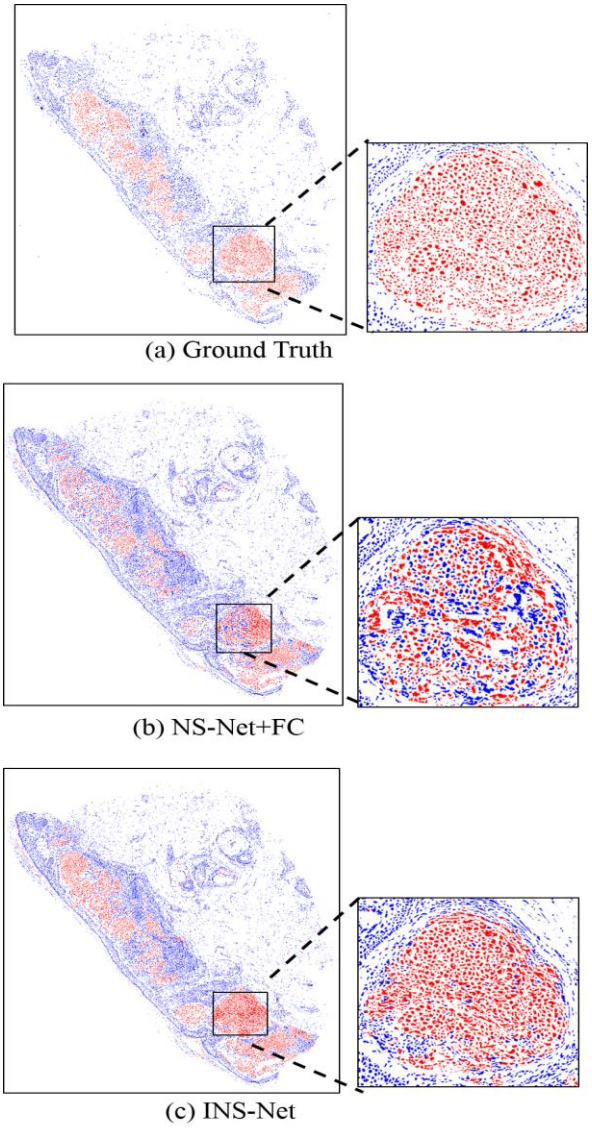

Figure 4. Subjective comparison of nuclei classification results of H&E-stained image shown in Fig. 1. (a) The ground truth classification, (b) classification by NS-Net+FC, and (c) classification by the proposed INS-Net. The melanoma, non-melanoma and background pixels are presented in red, blue and white pixels, respectively.

B. *Melanoma Region Detection performance*

In this section, the MRD is evaluated and compared with the state-of-the-art techniques. The overall performance is shown in Table III. It is observed that the proposed INS-Net+ MRD provides superior melanoma region detection performance compared to that of the PBS-ResNet and NS-Net+FC+MRD techniques. Fig. 5 shows a visual comparison of the regions, for the WSI shown in Fig. 1, detected by the NS-Net+FC+MRD and INS-Net+ MRD. The NS-Net+FC+MRD generates a significant number of FP and FN pixels due to the misclassification of melanoma nuclei. On the other hand, the INS-Net+MRD provides a superior segmentation performance. This is because the technique first classifies the melanoma nuclei and builds the melanoma region by merging the detected nuclei thereby resulting in a superior performance.

TABLE II. NUCLEI CLASSIFICATION PERFORMANCE

Technique	Accuracy	Dice Coefficient
NS-Net+FC [6]	97.27	81.37
INS-Net (proposed)	99.33	86.48

TABLE III. MELANOMA SEGMENTATION PERFORMANCE

Steps	Accuracy	Dice Coefficient
PBS-ResNet [7]	94.66	78.00
NS-Net+FC+MRD[1][6]	95.80	80.88
INS-Net+MRD (proposed)	97.70	85.10

[1]The MRD module has been added to NS-Net+FC to generate the melanoma region mask.

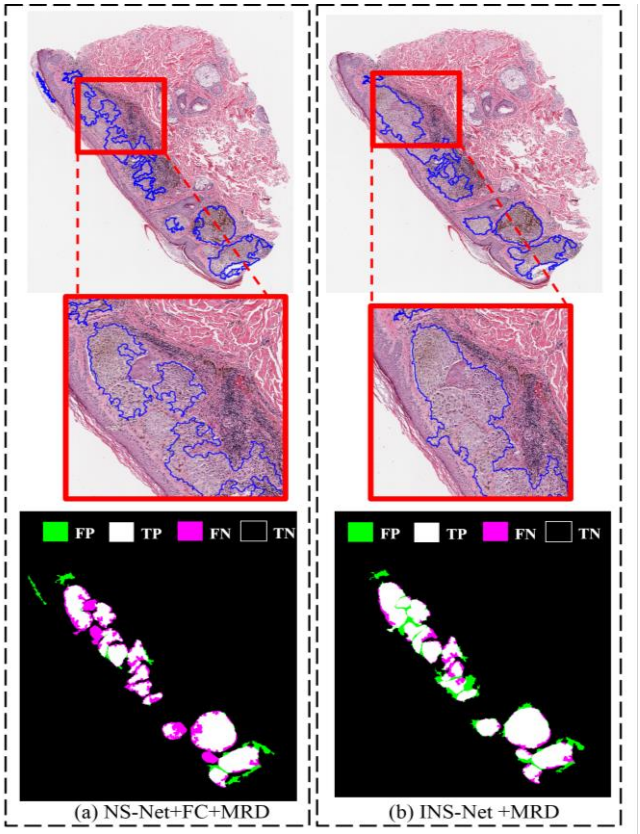

Figure 5. Subjective comparison of the detected melanoma regions. (a) NS-Net+FC+MRD, and (b) INS-Net+MRD. The melanoma regions are shown with blue contour. The last row shows the FP (green), TP (green), FN (red) and TN (black) regions.

IV. CONCLUSION

This paper proposes an automated technique to detect the melanoma regions in a skin tissue slide stained with H&E. The proposed technique segments the cell nuclei in H&E-stained images using a deep learning INS-Net architecture. The INS-Net architecture segments the whole slide image into melanoma nuclei, non-melanoma nuclei and background regions. The segmented melanoma nuclei are then used to generate a melanoma region mask using morphological operations. The proposed technique provides an excellent segmentation performance with a low computational complexity.

ACKNOWLEDGMENT

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) (Grant number RGPIN-2020-05873).

REFERENCES

[1] L. Brochez, E. Verhaeghe et al., "Inter-observer variation in the histopathological diagnosis of clinically suspicious pigmented skin lesions," *Journal of Pathology*, 196(4), pp.459–466, 2002.

[2] M. G. Rojo, A. M. Castro, and L. Gonçalves, "COST action "eurotelepath": digital pathology integration in electronic health record, including primary care centres," *Diagnostic Pathology* 6, S6 (2011), pp. S1–6, 2011.

[3] R. S. Weinstein, A. R. Graham, L. C. Richter et al., "Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future," *Human Pathology*, vol. 40, no. 8, pp. 1057–1069, 2009.

[4] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. of Medical Image Computing and Computer-Assisted Intervention* (MICCAI). Vol. 9351, pp. 234–241, 2015.

[6] S. Alheejawi, R. Berendt, N. Jha and M. Mandal, "Melanoma Cell Detection in Lymph Nodes Histopathological Images using Deep Learning," *Signal & Image Processing: An International Journal* (SIPIJ) Vol.11, No.4, August 2020. DOI: 10.5121/sipij.2020.11401

[7] P. Sabol, P. Sinčák, P. Hartono, P. Kočan, et al., "Explainable classifier for improving the accountability in decision-making for colorectal cancer diagnosis from histopathological images" *Journal of Biomedical Informatics*, Vol. 109, 103523, Sept 2020.

[8] M. Gadermayr, A. -K. Dombrowski, B. M. Klinkhammer, P. Boor, D. Merhof, "CNN cascades for segmenting sparse objects in gigapixel whole slide images," *Computerized Medical Imaging and Graphics*, Vol. 71, pp. 40-48, 2019.

[9] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, *private communication*, May 1995.

[10] H. Kaiming, Z. Xiangyu, R. Shaoqing, and S. Jian, "Deep Residual Learning for Image Recognition" *IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), Las Vegas, NV, USA, pp. 770-778, 2016.