Epithelial segmentation from in situ hybridisation histological samples using a deep central attention learning approach

Song, Tzu-Hsi; Landini, Gabriel; Fouad, Shereen; Mehanna, Hisham

DOI:
10.1109/ISBI.2019.8759384

EPITHELIAL SEGMENTATION FROM *IN SITU* HYBRIDISATION HISTOLOGICAL SAMPLES USING A DEEP CENTRAL ATTENTION LEARNING APPROACH

Tzu-Hsi Song[1], *Gabriel Landini*[1], *Shereen Fouad*[1,2] *Hisham Mehanna*[3]

[1]School of Dentistry, University of Birmingham, UK
[2]Department of Computer Science, Birmingham City University, UK
[3]Institute of Cancer and Genomic Sciences, University of Birmingham, UK

ABSTRACT

The assessment of pathological samples by molecular techniques, such as *in situ* hybridization (ISH) and immunohistochemistry (IHC), has revolutionised modern Histopathology. Most often it is important to detect ISH/IHC reaction products in certain cells or tissue types. For instance, detection of human papilloma virus (HPV) in oropharyngeal cancer samples by ISH products is difficult and remains a tedious and time consuming task for experts. Here we introduce a proposed framework to segment epithelial regions in oropharyngeal tissue images with ISH staining. First, we use colour deconvolution to obtain a counterstain channel and generate input patches based on superpixels and their neighbouring areas. Then, a novel deep attention residual network is applied to identify the epithelial regions to produce an epithelium segmentation mask. In the experimental results, comparing the proposed network with other state-of-the-art deep learning approaches, our network provides a better performance than region-based and pixel-based segmentations.

Index Terms— Oropharyngeal Cancer, Tumor Segmentation, Deep Learning, In Situ Hybridisation, Histology

1. INTRODUCTION

Oropharyngeal cancer (OPC) is a type of epithelial head and neck cancer occurring in the oropharynx, tonsils and base of the tongue. The incidence of OPC has seen a significant increase in recent times in the western world. Certain human papillomavirus (HPV) strains (called high-risk) have been identified as risk factors for the development of OPC, although not all OPCs are HPV related. HPV-positve (HPV+) OPC cases appear to have a better prognosis than HPV-negative (HPV-) cases, so there is a well justified diagnostic interest in detecting HPV infection in these tumours. Currently, correctly diagnosing HPV-associated OPC is a major challenge faced by pathologists [1, 2].

In situ hybridization (ISH) is a standard laboratory technique suitable for detecting viral genomes in tissues [1, 2, 3]. We analysed images of OPC tissue microarrays processed with the Ventana INFORM HPV III system (Roche), which consists of a cocktail of HPV genomic probes, enzyme-labelled, used to precipitate a chromogen (nitroblue tretrazolium or NBT/BCIP) therefore enabling the visualisation of the hybridised genomes (deep blue navy colour) in the nucleus of the infected cells. A counterstain is used (in this case Red Counterstain II, pink in colour) to facilitate identifying the general tissue morphology of the sample. For assessing the HPV status of a tumour, however, it is essential to confirm the blue staining localised in the nucleus of the epithelial (tumour) cells, while ignoring a variety of possible artefacts (chromogen non-specific precipitation, drying artefacts, non-specific leukocyte cytoplasm staining) which could potentially lead to false positive readings. Figure 1 shows samples of OPC ISH images.

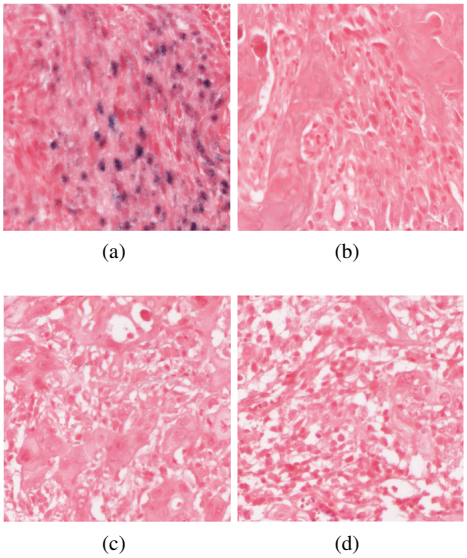


Fig. 1: Samples of ISH technique in: (a) HPV-positive tumour case (note the blue stain in the epithelial tissue); (b) and (c) are HPV-negative tumour cases; (d) stromal tissue. Note the difficulty in distinguishing tumour tissue regions between (c) and (d) because the morphology and texture information are quite similar and confusing. Field width 220 μm.

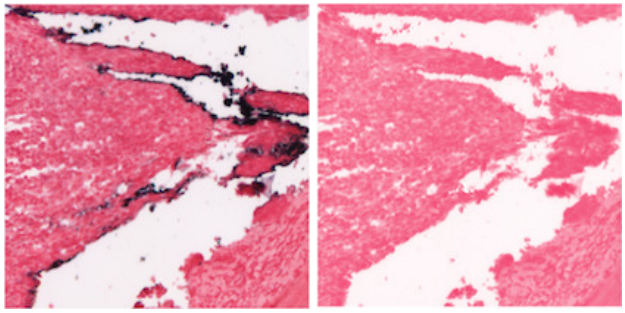
Interpreting ISH samples is difficult to automate and remains a tedious and time consuming task for microscopists as the blue hybridisation staining is fine patterned and Red Counterstain II is not a tissue specific dye. In order for ISH products to be of diagnostic value, these products need to be identified within epithelial cells, to avoid confusion with, e.g. artefacts in other tissues in the same sample. Therefore, our objective was to investigate to what extent tumour epithelial regions can be automatically identified from non-epithelial tissues (e.g. connective stroma) when using information only from the counterstain dye. Recently, several deep learning methods for segmentation have shown better performance than traditional machine learning approaches [4, 5]. For instance, the Unet learning model [4] uses fully convolutional network to take into account biomedical image segmentation. Xu *et al.* [5] proposed a framework that used convolutional neural network (CNN) and superpixels to identify epithelial regions in breast cancer histopathological images. However, a superpixel carries only information on local features and consequently the classification precision of the learning network is limited. In this paper, we present a framework with a novel deep learning method using images of tissues exploiting only the Red counterstain II dye to segment the epithelial regions in OPC samples.

2. METHOD

In order to efficiently and precisely segment epithelial regions, we propose a framework using a novel residual network. First, an image containing information contributed by the counterstain is segmented into superpixels. Then our novel deep learning network is trained based on a gold standard produced by an experienced microscopist to discriminate superpixels belonging to epithelium from those belonging to non-epithelial regions by considering the features of each superpixel and its surrounding area.

2.1. Preprocessing

Tissue microarray sections of OPC processed by ISH staining use NBT/BCIP and Red Counterstain II to localize the HPV genomes in cell nuclei and reveal tissue morphology, respectively. We determined a set of colour vectors to perform colour deconvolution of the RGB images [6, 7] with the purpose of obtaining three separate stain channels: NBT/BCIP (blue), Red Counterstain II (pink) and a residual channel. We found the pink channel preserves tissue morphological information in comparison with original ISH staining colour image reglardless of the tissues being HPV+ or HPV- and avoids the influence of artefacts (shown in Figure 2). The pink channel is further processed with the SLIC superpixel method [8] to segment the image data into coherent regions. The size of the resulting superpixels is around 2500 pixels each. We also consider the features of neighbouring areas surrounding a superpixel because superpixels alone do not provide enough feature information to allow the network to precisely identify superpixel types. To this end, we capture a region of interest consisting of the superpixel region plus surrounding pixels within a square patch of size 100X100 pixels and used these as input images for network training.



(a) (b)

Fig. 2: (a): A tissue sample processed with ISH. Note some dark regions artefacts which are not due to ISH products. (b): Red counterstain channel obtained after colour deconvolution. Note the reduction of the artefacts while maintaining the tissue morphological information. Field width 330 μm.

2.2. Deep Central Attention Residual Network

After collecting the input images, a novel deep learning network is used to identify the epithelial (tumour) and non-epithelial (e.g. stroma) superpixel regions. We adopt the concept of a residual network, that is, make the network concentrate on specific residual features of the input images to increase the accuracy of classification [9]. We designed a central attention residual (CAR) block which considers the correlated features between the central and the neighbouring regions, to build the proposed central attention residual (CAR) network. Figure 3 shows the architecture of the proposed CAR network and the whole framework for epithelium segmentation based on the counterstain (pink) channel.

The proposed CAR network consists of a) four convolution layers that generate a number of feature maps and reduce their dimensions, b) four CAR blocks, c) an average pooling layer and d) a softmax classifier. Each CAR block utilizes the concept of residual network to efficiently learn specific features to achieve more accurate identifications than traditional CNNs [9]. In addition, the CAR blocks include three convolution layers, a central attention (CA) unit, which emphasizes the features information of the central area of the input image, batch normalization and a rectified linear unit (ReLu) activation function. The first two convolution layers of the CAR blocks generate residual feature maps. Then the CA unit uses a convolution layer with two-pixel sliding to obtain high-level feature maps and then a deconvolution layer is

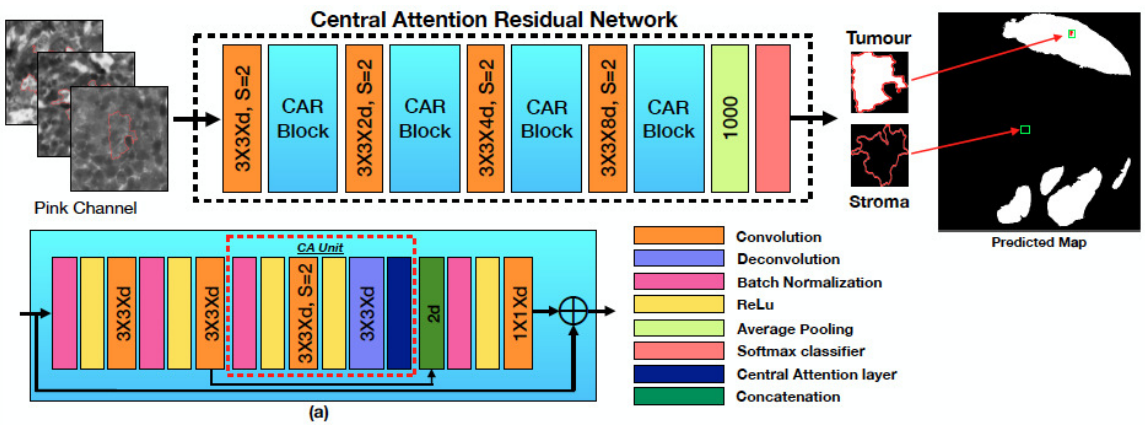


Fig. 3: After computing superpixel regions, these were expanded to 100x100 pixels regions to include neighbouring image data. We then designed a proposed central attention residual network to identify epithelial superpixels. This is then used to construct a predicted epithelium mask on each image. The parameters d and S denote the number of feature maps and the sliding pixel shift, respectively. In (a) is shown the structure of central attention residual (CAR) block.

used to reconstruct the low-level residual feature maps. This reconstructed feature maps are able to identify specific features and decrease the influence of noise. The CA layer is introduced to strengthen the reconstructed residual features of the central area and to 'fade out' the features of the neighbouring area, far from the input image centre to generate the central attention feature maps. The formula of the CA layer is:

$$h_{CA} = h^{'} \bullet G(h^{'}, \sigma^2) \quad (1)$$

where h_{CA} denotes the central attention feature maps and $h^{'}$ is the reconstructed feature maps; $G(h^{'}, \sigma^2)$ is the central attention function (Gaussian) with the variance set to $\sigma^2 = 0.4$. Then these central attention feature maps and the previous residual feature maps are concatenated. After that, the CAR block uses a convolution layer to compress the concatenated feature maps and increase the efficiency of the network learning. The proposed CAR block efficiently utilizes both central (local) and whole (global) feature information to help the whole network learn correlations and similarities to increase the accuracy of the predicted classification. Between two CAR blocks, we use a convolutional layer to increase the number of feature maps and decrease their size. Finally, an average pooling and a softmax classifier are used to identify the superpixel type and generate the predicted mask of epithelium. We use the cross-entropy loss as the loss function to train the proposed network. In the post-processing stage, we use morphological operations and cubic interpolation to smooth the shape of predicted mask boundary.

3. EXPERIMENTAL RESULTS

In the experiment, we used 48 tissue microarray core images of OPC (3300x3300 pixels, inter-pixel distance 0.367 μm), with 38 images for training and 10 for testing, to evaluate the performance of our proposed model. We compared the proposed network with other state-of-the-art deep learning models: Unet [4], CNN [5], residual network (ResNet) [9, 10], with superpixels and different colour types of input images (RGB, greyscale and red Countertain II channel). The superpixel sizes (originally around 2500 pixels) were resized to 50X50 pixels for the training of CNN and ResNet because those methods requires equal sized input images for network training. In the experiment, the CNN network used 4 convolutional layers, 3 maxpooling and a softmax classifier; the ResNet adopted the architecture of ResNet18 [9] to mainly use a convolutional layer, a maximum pooling layer, 3 residual blocks, an average pooling layer and a softmax classifier. The architecture of the Unet is the same as [4] with the input image size of 128X128 for the experiment. All the implementations were done in Python.

For fair comparisons, we used the mean of F1-score and Dice similarity coefficients to evaluate the performance of region-based and pixel-based segmentations, respectively. Table 1 presents the average segmentation accuracy. The proposed method provides the best performance among the approaches, with a region-based accuracy of 86.3% and pixel-based accuracy of 83.8%. Examples of the visual results of the epithelium segmentation achieved with different deep learning approaches are given in Figure 4. The results show that using the counterstain only input images, the predicted map generated with our proposed network is closest to the ground truth. This indicates that colour and greycale input images combine extra features information, such as blue dark spots and artefacts, which introduce confusion to the network learning. The pink channel efficiently maintains the tissue morphological information without the influence of other colour information or artefacts to provide more accurate classification results.

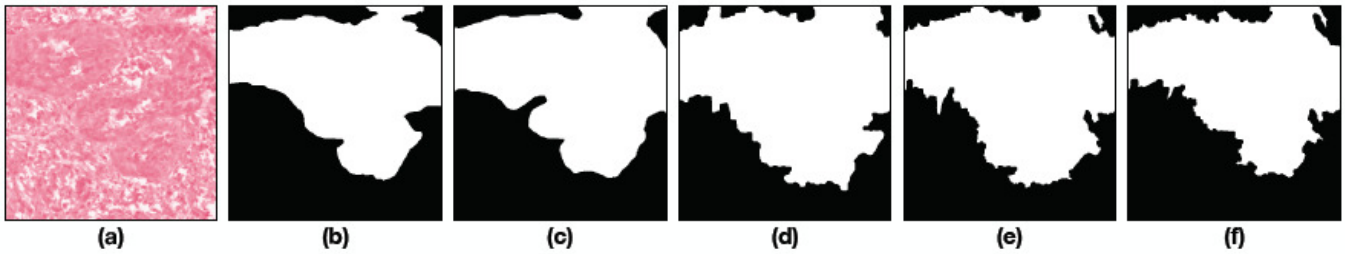


Fig. 4: An example of the results of epithelium segmentation using different methods: (a) pink channel (b) ground truth (c) Unet (d) superpixel+CNN (e) superpixel+ResNet (f) superpixel+proposed network. The predicted map of proposed network shows less difference with ground truth than those of other methods do. Field width 352 μm.

Table 1: Comparing the proposed model with other deep learning approaches

Input Types	Unet		CNN		ResNet		Proposed CAR network	
	F1-score	Dice	F1-score	Dice	F1-score	Dice	F1-score	Dice
RGB	*	74.71%	80.67%	77.46%	83.51%	80.74%	85.34%	82.16%
Gray	*	71.09%	78.63%	73.82%	80.81%	78.15%	84.67%	81.5%
Pink	*	76.23%	82.18%	79.27%	84.35%	81.07%	**86.31%**	**83.77%**

*: not applicable.

4. CONCLUSION

The proposed framework utilises colour deconvolution and the novel CAR network solely based on a single counterstain channel with superpixels to identify the epithelial regions in OPC tissue ISH images. The experimental results show that our proposed network can identify epithelial regions in OPC tissue microarrays based only on Red Counterstain II staining by capturing more neighbouring feature information surrounding superpixels. The CAR network can efficiently learn the correlation between central and neighbouring regions. Compared with other deep learning methods, the proposed deep learning network provides better performance than other state-of-the-art approaches. This framework may help pathologists in automating the identification of the ISH products in given histological compartments. The technique should also have wide applicability to other ISH or IHC analyses. Our next aim it so investigate the ISH products in relation with epithelium regions for the further development of an automated HPV status classifier.

5. ACKNOWLEDGEMENTS

This work was supported by the EPSRC (UK) through funding under grant EP/M023869/1 Novel context-based segmentation algorithms for intelligent microscopy".

6. REFERENCES

[1] A. S. Tsao, V. Papadimitrakopoulou, H. Lin, M. Guo, J. J. Lee, F. C. Holsinger, W. K. Hong, and E. M. Sturgis, "Concordance of oral hpv prevalence between patients with oropharyngeal cancer and their partners," *Infectious agents and cancer*, vol. 11, no. 1, pp. 21–30, 2016.

[2] P. Morbini and M. Benazzo, "Human papillomavirus and head and neck carcinomas: focus on evidence in the babel of published data," *Acta Otorhinolaryngologica Italica*, vol. 36, no. 4, pp. 249–258, 2016.

[3] A. Schache, J. Croud, M. Robinson, and S. Thavaraj, "Human papillomavirus testing in head and neck squamous cell carcinoma: best practice for diagnosis," *Histopathology*, pp. 237–255, 2014.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.

[5] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, 2016.

[6] A. C. Ruifrok, D. A. Johnston, et al., "Quantification of histochemical staining by color deconvolution," *Analytical and quantitative cytology and histology*, vol. 23, no. 4, pp. 291–299, 2001.

[7] G. Landini, *Colour Deconvolution Plugin for*

ImageJ, https://imagej.net/Colour_Deconvolution, accessed 2017-4-29.

[8] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[9] 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*, 2016, pp. 770–778.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*, 2016, pp. 630–645.