# Uncertainty Driven Pooling Network for Microvessel Segmentation in Routine Histology Images

M. M. Fraz[1,2,4(✉)], M. Shaban[1], S. Graham[1], S. A. Khurram[5], and N. M. Rajpoot[1,2,3]

[1] Department of Computer Science, University of Warwick, Coventry, UK
moazam.fraz@warwick.ac.uk
[2] The Alan Turing Institute, London, UK
[3] University Hospitals Coventry and Warwickshire, NHS Trust, Coventry, UK
[4] National University of Sciences and Technology, Islamabad, Pakistan
[5] The University of Sheffield, Sheffield, UK

**Abstract.** Lymphovascular invasion (LVI) and tumor angiogenesis are correlated with metastasis, cancer recurrence and poor patient survival. In most of the cases, the LVI quantification and angiogenic analysis is based on microvessel segmentation and density estimation in immunohistochemically (IHC) stained tissues. However, in routine H&E stained images, the microvessels display a high level of heterogeneity in terms of size, shape, morphology and texture which makes microvessel segmentation a non-trivial task. Manual delineation of microvessels for biomarker analysis is labor-intensive, time consuming, irreproducible and can suffer from subjectivity among pathologists. Moreover, it is often beneficial to account for the uncertainty of a prediction when making a diagnosis. To address these challenges, we proposed a framework for microvessel segmentation in H&E stained histology images. The framework extends DeepLabV3+ by using an improved dice coefficient based custom loss function and also incorporating an uncertainty prediction mechanism. The proposed method uses an aligned Xception model, followed by atrous spatial pyramid pooling for feature extraction at multiple scales. This architecture counters the challenge of segmenting blood vessels of varying morphological appearance. To incorporate uncertainty, random transformations are introduced at test time for a superior segmentation result and simultaneous uncertainty map generation, highlighting ambiguous regions. The method is evaluated using 1167 images of size $512 \times 512$ pixels, extracted from 13 WSIs of oral squamous cell carcinoma (OSCC) tissue at 20x magnification. The proposed net-work achieves state-of-the-art performance compared to current semantic segmentation deep neural networks (FCN-8, U-Net, SegNet and DeepLabV3+).

**Keywords:** Microvessel detection · Tumor angiogenesis
Lymphovascular invasion · Separable convolution
Pyramid pooling based neural network · Uncertainty quantification


D. Stoyanov et al. (Eds.): COMPAY 2018/OMIA 2018, LNCS 11039, pp. 156–164, 2018.
https://doi.org/10.1007/978-3-030-00949-6_19

## 1 Introduction

The ability for cancer to spread to distant or adjacent tissues is a key indicator of poor patient prognosis. The tumor cells can gain access to blood or lymphatic vessels by intravasation allowing them to circulate through the intravascular stream. This lamphovascular invasion (LVI) can lead to the proliferation of tumor cells at another site in the body. This phenomena is more commonly referred to as metastasis. The lymphatic or vascular invasion by the primary tumor is considered as a sign of aggressive disease and is usually accompanied by metastases to the regional lymph nodes and to distant sites. The formation of new blood vessels is important for growth, survival and metastatic spread of tumor cells [2]. Tumor neoangiogenesis leads to formation of new blood vessels in the tumor tissues, with an initial purpose to facilitate the transport of nutrients and oxygen to help the tumor cells survive [2]. In diagnostic clinical pathology, most of the currently existing tissue datasets and pathways recommend commenting on the presence or absence of LVI however the degree of angiogenesis is not routinely examined or reported. In the existing research literature, there is strong evidence that microvessel density in tumor tissue is directly correlated with an increased risk of cancer spread, an increased incidence of disease recurrence and poor patient survival [10]. Recent studies have re-ported LVI detection and quantification as an important risk factor in disease progression particularly in breast, cervical and lung cancers [10]. However, most of the results in these studies are based on manual localization of microvessels in the subjectively defined regions of tissue whole slide image (WSI). The subjective identification of LVI and angiogenic regions in the tissues is irreproducible, time consuming and often requires clinical knowledge. In routine pathological practice, accurate segmentation of microvessels can assist pathologists in identification of LVI which would otherwise be very time-consuming. Furthermore, objective quantification of LVI and tumor angiogenesis from multi giga pixel histopathology images will provide extremely valuable 'big data' aiding prediction of tumor behavior and prognosis.

The appearance of microvessels in Hematoxylin and Eosin (H&E) stained histology WSIs is characterized by the presence of endothelial cells forming a closed structure that surrounds the red blood cells, as illustrated in Fig. 1. The heterogeneity in the visual appearance of microvessels makes their segmentation a non-trivial task. Immunohistochemical (IHC) staining for markers such as CD31 and CD34 can be used for microvessel identification, which is comparatively expensive than H&E staining. Due to this cost, it is not commonly used in routine clinical practice. Likewise, the IHC-stained histology images are rarely available in public datasets, which is a major hindrance in the development of automated methods to investigate the role of microvessels in tumor prognosis and therapy response.

Deep learning has recently been successfully used for the automated analysis of histopathology images. Specifically, deep learning based architectures have been proposed for the detection, segmentation and classification of histopathological structures in WSI images including deep learning based architectures are

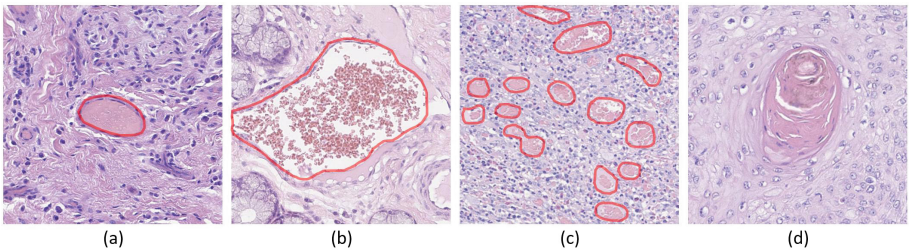


**Fig. 1.** Representative images of microvessels in H&E stained histology images of OSCC tissue, illustrating their shape variability. The microvessel boundary annotation is shown in red color. (a–c) Microvessel with varying red cells density, (c) Microvessels at different sizes, (d) Keratinization, which appears similar to microvessel (Color figure online)

proposed for detection and classification of nuclei [14], mitoses [12], lymphocytes [13], tumor and stromal regions [15] and glandular structure segmentation [7]. A few studies have explored automatic quantification of tumor angiogenic hotspots by the detection of microvessels in IHC stained histology images only [8]. Most recently, a fully convolutional neural network (FCN) based method for microvessel detection in H&E stained images is presented [16]. However, the methods for microvessel segmentation particularly in H&E stained images are limited.

In this paper, we present a framework for precise segmentation of microvessels in H&E stained histology images at multiple scales and resolutions by using an uncertainty aware spatial pyramid pooling deep neural network architecture. The Deeplabv3+ [4] architecture is extended by using an improved dice coefficient minimization based custom loss function and by accounting for the uncertainty of a prediction. The proposed network aims to solve the key challenges posed by automated microvessel segmentation. The method uses a modified Aligned Xception model [5] followed by an atrous spatial pyramid pooling (ASPP) unit [3] for feature extraction at multiple scales. This overcomes a major challenge of segmenting vessels of various sizes. Moreover, despite achieving state-of-the art performance in semantic segmentation, the deep networks typically do not inherently model the segmentation uncertainty. For this purpose, we apply random transformations to the images during test time, as a method to generate the approximate predictive distribution. Taking the average of these predictions of transformed images yields a superior segmentation and enables us to observe ambiguous areas, where the network is uncertain in a decision. The methodology is evaluated on 1167 images of size $512 \times 512$ pixels, extracted from 13 WSIs of oral squamous cell carcinoma (OSCC) tissue at 20x magnification, and demonstrated promising results. Moreover, the proposed network achieves state-of-the-art performance compared to current semantic segmentation deep neural networks (FCN-8 [9], U-Net [11], SegNet [1] and DeepLabv3+ [4]).

## 2 Proposed Method

We propose a framework for the segmentation of microvessels in H&E stained histology images. The framework extends the DeepLabV3+ network architecture by utilizing a modified Aligned Xception with ASPP and a custom loss function for precise segmentation. Moreover, the framework also incorporates random transformations at test time to account for the uncertainty of a prediction, as shown in Fig. 2. A substantially deep network is needed for meaningful feature extraction. Traditional convolutional neural network architectures (AlexNet, VGG, Google Net, ResNet e.t.c.) used for image classification has inherited limitations to model geometric transformations due to fixed geometric structures in their building modules. Moreover, these networks use a hierarchical combination of maxpooling and convolutions to increase the receptive field size. This results in loss of image information which may be very significant for precise object segmentation. In order to deal with these issues, the feature extraction process should be invariant to geometric transformations and retain the low level image information. The Xception model [5] has demonstrated promising performance in image classification task on ImageNet in terms of speed and accuracy. Xception has been modified to incorporate geometric transformations modeling capability for feature extraction. Further to this, Chen *et al.* [4] proposed the replacement of all max pooling operations with depthwise separable convolutions with striding. This allows the application of atrous separable convolutions for multiscale feature extraction at arbitrary resolution. Atrous convolution is an extension of the standard convolution operation, which provide us with the ability to explicitly control the resolution of features computed by deep convolutional neural networks and adjust filter's receptive field for capturing multiscale information. The use of separable convolutions reduces the number of convolutional parameters, hence increasing the computational efficiency. Subsequntly, this is more suitable for processing of multi giga pixel WSIs.

Inspired from [4], we have use modified Aligned Xception model for microvessel feature extraction at multiple scales and resolution with geometric deformation invariance. Incorporating multiscale and geometric transformation invariant features allow us to perform accurate microvessel segmentation. Atrous spatial pyramid pooling [3] with varying dilation rates (6, 12 and 18) is applied at the end of the encoder, to aggregate the multilevel features. This pooling module allows our proposed network to segment microvessels of varying shape and sizes. Global average pooling has been used to incorporate the global level context. Moreover, a $1 \times 1$ convolution is performed before each operation, followed by a dropout layer and another $1 \times 1$ convolution for dimensionality reduction. The features from each dilation operation is concatenated to give a powerful representation of high level image contextual information. The low level image information for precise delineation of microvessel boundaries is taken from the shallow layers of the deep network and concatenated with the feature map obtained after bilinear upsampling by a factor of 4. The feature map size is illustrated at each block level. The output is upsampled twice by a factor of 4 to obtain the final output after applying the softmax layer. We have used the loss function $\tau_{DSC}$ based

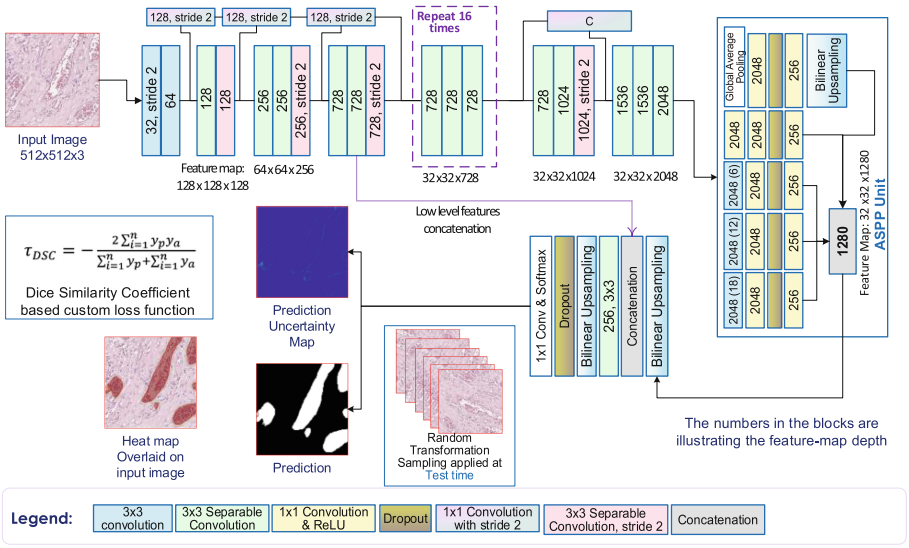


**Fig. 2.** The Proposed Framework: The legend represents the color coding of various types of convolutional layers. The values inside the blocks represents the depth of corresponding feature map. In ASPP Unit, the value in brackets denotes the dilation rates of 6, 12 and 18. To obtain the prediction uncertainty map, Random Transformation Sampling at test time is applied. The network uses custom loss function based on minimization of –ve of Dice Similarity Coefficient.

on minimizing the negative of dice similarity coefficient (DSC) for training the network. The custom loss function is explained in Eq. 1

$$\tau_{DSC} = -\frac{2\sum_{i=1}^{n} y_p y_a}{\sum_{i=1}^{n} y_p + \sum_{i=1}^{n} y_a} \tag{1}$$

where, $y_p$ represents the value of softmax predicted segmentation $map \in [0, \ldots, 1]$, and $y_a$ is the ground at each pixel $i$.

Traditional deep learning models are capable of learning discriminative features and have the ability to accurately map the high dimensional input data to expected output. However, the models do not quantify that how certain the model is its prediction. A Bayesian approach in machine learning can model the uncertainty, but current deep learning models do not represent the prediction uncertainty. Recently, a number of methods for uncertainty quantification by estimating the posterior distribution have been proposed [6]. We estimate the model uncertainty by applying random transformations [7] to the test input images. With this we are able to capture the noise inherent in the input data and visualize the regions where the segmentation network is uncertain in its prediction. To obtain the predictive distribution, we apply a random transformation $\delta(x)$ to a set of m images, where $\delta$ performs median blur, Gaussian blur, rotation, flipping or Gaussian noise. Taking the mean of this ample gives the

refined prediction and the variance within the sample gives the uncertainty in prediction. The prediction and the uncertainty can be defined as;

$$\mu = -\frac{1}{m}\sum_{i}^{m} f(\delta_i(x); W); \quad \sigma = -\frac{1}{m}\sum_{i}^{m} f(\delta_i(x); W - \mu)^2 \tag{2}$$

where, $\mu$ is the prediction of microvessel segmentation, $\sigma$ is the prediction uncertainty and $m$ in the number of applied transformations. $\delta_i$ denotes the random transformation applied to the input image $x$. Taking the average of the prediction of transformed images give better segmentation.

## 3 Experiments and Results

### 3.1 Materials

We have used a set of 1167 image tiles of size $512 \times 512$ pixels taken from 13 H&E stained WSIs of OSCC tissue at $20\times$ magnification. The ground truth for microvessels is validated by two independent pathologists. The dataset is split into training, validation sets such that 686 training and 226 validation images are obtained from 10 WSIs. The test set is comprises of 255 images taken from remaining 3 WSIs. The training and validation images are augmented with random rotation, elastic distortion, random flip, median blur and Gaussian blurring.

### 3.2 Experimental Settings

The framework is implemented in Keras 2.2 with TensorFlow backend and trained on workstation equipped with Nvidia GeForce GTX 1080 Ti for 175 epochs (35000 iterations). We have used Adam optimizer, the learning rate was initialized at 10-4, the input image to the network is $512 \times 512 \times 3$ and the batch size is 2. As explained in Sect. 2, we have used a custom loss function based on minimizing the dice score.

### 3.3 Evaluation

The model is quantitative evaluated using Jaccard index, Dice Similarity Coefficient (DSC), Accuracy, Sensitivity, Specificity, Precision and Recall. Furthermore, several state-of-the-art segmentation methods including FCN-8 [9], U-Net [11], SegNet [1] and DeepLabv3+ [4] are implemented for comparative analysis, which is presented in Table 1.

**Table 1.** Quantitative performance measures of miscrovessel segmentation, compared with FCN-8 [9], U-Net [11], SegNet [1] and DeepLabv3+ [4].

| | Jaccard index | DSC | Accuracy | Sensitivity | Specificity | Precision | Recall |
|---|---|---|---|---|---|---|---|
| FCN-8 | 0.8562 | 0.9225 | 0.9612 | 0.9086 | 0.9791 | 0.9368 | 0.9086 |
| U-Net | 0.8561 | 0.9225 | 0.9616 | 0.9010 | 0.9818 | 0.9442 | 0.9017 |
| SegNet | 0.8431 | 0.9148 | 0.9569 | 0.9100 | 0.9729 | 0.9524 | 0.8854 |
| DeepLabv3+ | 0.8741 | 0.9329 | 0.9667 | 0.9085 | 0.9834 | 0.9540 | 0.9089 |
| Proposed | **0.8851** | **0.9390** | **0.9694** | **0.9261** | **0.9862** | **0.9558** | **0.9225** |

## 4 Discussion and Conclusion

Microvessels in H&E stained histology images display a high level of heterogeneity with respect to their size, shape, texture, and luminal red cells density. Figure 3 shows the visual results of four challenging cases in microvessel segmentation. Figure 3(a, b) shows a vessel partially filled with red cells. U-Net is unable to segment the microvessel region where red cells are not present whereas SegNet, FCN-8 and DeepLabV3+ manage to segment the microvessel but with low confidence. In contrast, the proposed method segments the complete vessel with high confidence. The misdetections of variable sized microvessels and the segmentation microvessels located in close proximity are the other challenging case, illustrated in Fig. 3c and d respectively. FCN-8, U-Net and DeepLabv3+ are unable to segment small vessels Fig. 3c and the SegNet merged the two closely located by vessels into one large vessels (Fig. 3d). The proposed method successfully segments the vessels belonging to these challenging cases. The proposed method achieves a superior performance for all evaluation measures as illustrated in Table 1. Although, the quantitative performance gain may not appear notably significant, the visual results illustrated in (Fig. 3) show that the proposed framework successfully localizes and segments the microvessels of different shapes and sizes with mercurial density of red cells. The segmentation of microvessels in the histology images is the first step in automated quantification of LVI and estimation of tumor angiogenesis. In routine pathological practice, the microvessels can be identified using IHC stained histology images with associated time and cost implications. We have present a method for the precise segmentation of microvessels in H&E stained histology images. The proposed method uses a modified aligned Xception model, atrous spatial pyramid pooling and a customized dice coefficient minimization based loss function to segment microvessels of various shapes and size. Random transformations at test time are used to incorporate the predictive uncertainty. Taking the average of these predictions gives a superior segmentation. The visual results and quantitative performance measures illustrate that the proposed method is able to precisely segment the microvessels in challenging cases.

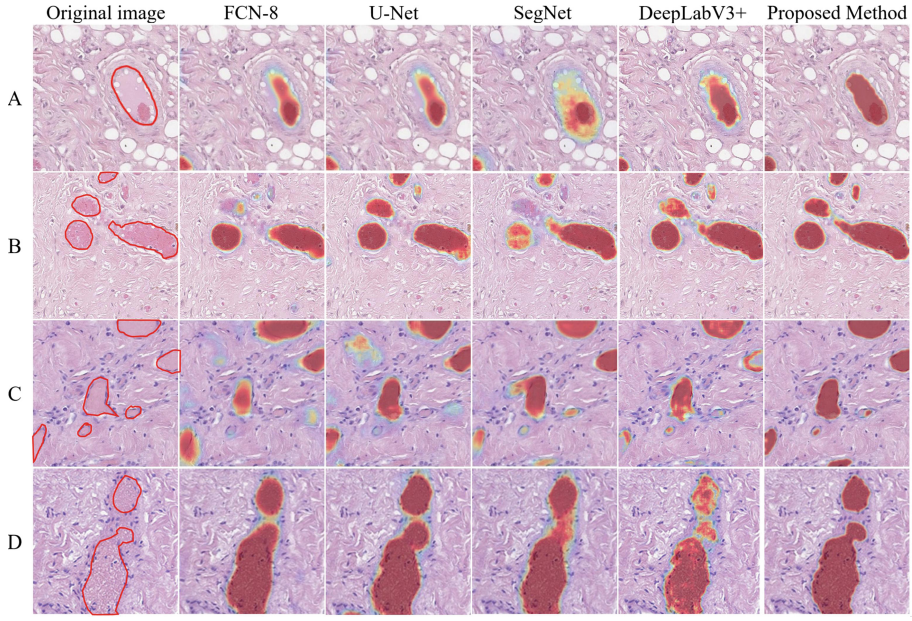


**Fig. 3.** Visual illustration of microvessel segmentation results shown as miscrovessel prediction heatmap obtained by FCN-8 [9], U-Net [11], SegNet [1], DeepLabv3+[4] and the proposed approach, overlaid on the original images. The miscrovessel boundary is marked on the original images in the 1st column. (Color figure online)

## References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. arXiv preprint arXiv:1511.00561 (2015)
2. Carmeliet, P.: Angiogenesis in life, disease and medicine. Nature **438**(7070), 932 (2005)
3. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. **40**(4), 834–848 (2018)
4. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. arXiv preprint arXiv:1802.02611 (2018)
5. Chollet, F.: Xception: deep learning with depthwise separable convolutions. arXiv preprint arXiv:1610–02357 (2017)
6. Gal, Y., Ghahramani, Z.: Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In: International Conference on Machine Learning, pp. 1050–1059 (2016)
7. Graham, S., Chen, H., Dou, Q., Heng, P.A., Rajpoot, N.: MILD-Net: minimal information loss dilated network for gland instance segmentation in colon histology images. arXiv preprint arXiv:1806.01963 (2018)

8. Kather, J.N., Marx, A., Reyes-Aldasoro, C.C., Schad, L.R., Zöllner, F.G., Weis, C.A.: Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. Oncotarget **6**(22), 19163 (2015)
9. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
10. Noma, D., et al.: Prognostic effect of Lymphovascular Invasion on TNM staging in stage i non-small-cell lung cancer. Clin. Lung Cancer **19**(1), e109–e122 (2018)
11. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015 Part III. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
12. Saha, M., Chakraborty, C., Racoceanu, D.: Efficient deep learning model for mitosis detection using breast histopathology images. Comput. Med. Imaging Graph. **64**, 29–40 (2018)
13. Saltz, J., et al.: Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep. **23**(1), 181 (2018)
14. Sirinukunwattana, K., Raza, S.E.A., Tsang, Y.W., Snead, D.R., Cree, I.A., Rajpoot, N.M.: Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. Med. Imaging **35**(5), 1196–1206 (2016)
15. Xu, J., Luo, X., Wang, G., Gilmore, H., Madabhushi, A.: A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing **191**, 214–223 (2016)
16. Yi, F., et al.: Microvessel prediction in H&E stained pathology images using fully convolutional neural networks. BMC Bioinf. **19**(1), 64 (2018)