Deep Learning Approaches for mice glomeruli segmentation

Deep learning approaches for mice glomeruli segmentation

Duccio Meconcelli[1], Simone Bonechi [2], Giovanna Maria Dimitri[1]

1- DIISM, Department of Information Engineering and Mathematics, University of Siena, Siena, Italy

2- DISPOC, Department of Social, Political and Cognitive Sciences, University of Siena, Siena, Italy

Abstract. Deep learning (DL) is widely applied in biomedical image processing nowadays. In this paper, we propose the use of DL architectures for glomerulus segmentation in histopathological images of mouse kidneys. Indeed, in humans, the analysis of the glomeruli is fundamental to decide on the transplantability of the organ. However, no datasets with human samples are publicly available. Therefore, obtaining good segmentation performance on the kidneys of mice could be the first step for a transfer learning approach to humans. We compared the use of two well–known architectures for image segmentation, namely MobileNet and DeepLab V2. Both models showed very promising results.

1 Introduction and Background

Modern deep learning architectures are applied in many fields: from computer vision to biomedicine, from data mining to time series analysis [1, 2, 3, 4, 5]. In this scenario, the application of deep learning techniques to semantic segmentation has seen a rapid increase [6, 7], with several applications in the context of histopathological image analysis [8]. In the present study we propose the application of deep learning techniques to segment glomeruli in mouse renal histopathological images. Glomeruli are composed by a set of capillaries located inside the kidneys. Specifically, glomeruli are the part of the kidney responsible for filtering potentially harmful substances from the blood [9]. The identification of the glomeruli in the histopathological tissue is, in fact, an essential task for the correct diagnosis of numerous pathologies related to the kidneys. An example of glomerulus–related disease is focal segmental glomerulosclerosis, which is often due to drug abuse [10, 11, 12]. In presence of such condition, a biopsy is performed on the patient to identify the percentage of sclerotic and non–sclerotic glomeruli, which is related to the correct functioning of the kidneys. Identifying and counting the number of glomeruli in an image is normally performed visually by a human expert. This procedure takes a long time and, unfortunately, its slowness can drastically affect the possibility of kidney transplantation. This is why the development of automated methods to segment and count healthy glomeruli could be very helpful in speeding up renal tissue analysis. In recent years, many deep learning approaches have been proposed for automatic segmentation of human glomeruli [11, 12, 13, 14]. However, no publicly accessible

datasets are available that can be used to train deep learning architectures. Instead, a reference dataset was released by [19] for the segmentation of mouse glomeruli, which has been used extensively in the literature [16, 17, 18]. In this work we employed such dataset to train and compare the results of two segmentation architectures, MobileNet [21] and DeepLab V2 [20]. Network training on the mouse dataset could be the first pre–training step for a transfer learning approach to perform glomerulus segmentation on human samples, overcoming the lack of available data. The results obtained are very promising and surprisingly similar, opening up the use of MobileNet, designed to work on limited hardware, on smartphones, tablets and electron microscopes. The paper is organized as follows: in Section 2 the dataset and the segmentation architecture are described, while in Section 3 the experimental setup and the results are presented. Finally, Section 3.3 draws the conclusion and shows some future perspectives.

2 Materials and Methods

The dataset of mouse glomeruli is introduced in Section 2.1 while Section 2.2 describes the deep learning architectures employed in our experiments. Finally, the performance indicators used to evaluate the segmentation are described in Section 2.3.

2.1 Dataset

The annotated dataset used in this work is available online [1]. It contains Whole Slide Images (WSIs) specifically designed by the curators with the aim of creating a benchmark for glomerulus segmentation. The dataset is composed by 88 images in tiff format (tiled, jpeg compression) that can be displayed through histopathology software such as Orbit [19]. The histopathological images in the dataset are collected from two species (mouse and rat) with different staining procedures (H and DAB, FastRed, PAS, and three variations of H and E). The manually labelled annotations for each image were released in the SQLite database format and can be opened with the Orbit software, producing a total of 21037 annotated glomeruli. The original images were therefore divided into 512×512 patches to be fed into the deep learning architectures.

2.2 DeepLab V2 and MobileNet

DeepLab V2 is one of the most used segmentation networks, proven effective in many applications [20]. Overall, the network is based on a typical encoder–decoder architecture. In DeepLab, the atrous convolutions, which allow to enlarge the receptive field without using pooling operations, have replaced the standard convolutions. This prevents the loss of spatial information due to pooling operations. Instead, MobileNet is a common architecture for semantic segmentation, specifically designed to reduce the number of network parameters

[1] https://datadryad.org/stash/dataset/doi:10.5061/dryad.fqz612jpc

depthwise separable convolutions [21]. Therefore, MobileNet represents a lighter semantic segmentation architecture that can be used on limited hardware.

2.3 Model Evaluation

The Jaccard and Dice indices are two common metrics used to evaluate performance in semantic segmentation. Given two sets A and B, the Jaccard and Dice indices are defined as:

$$Jaccard = \frac{|A \cap B|}{|A \cup B|}$$

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$

Both the Jaccard score and the Dice index are used in statistics for assessing the similarity of two sets [23, 24].

Also, as the main focus of this document is to correctly count the number of glomeruli, we decided to evaluate the performance of the segmentation network using an additional metric. To count the number of glomeruli found by the network we used the OpenCV [22] function cv2, able to detect the number of connected components in an image. Then, we considered the difference between the number of predicted glomeruli and the number of glomeruli present in the ground–truth mask. In particular, the $\Delta_{Glomeruli}$ indicator can be defined as:

$$\Delta_{Glomeruli} = |num_{glomeruliGroundTruthMask} - num_{glomeruliPredictedMask}|$$

3 Experiments and Results

Experiments were performed using the Google Colab platform. In the rest of this section, we will describe the pre–processing steps implemented, together with the results of the application of DeepLab V2 and MobileNet.

3.1 Network Training

Both the architectures were trained with the same procedure. The WSIs were first divided in tiles of dimension 512×512, obtaining a dataset of 10329 image tiles with the corresponding masks, that can be used to train the network models. A data augmentation procedure (i.e. rotation of 90°) has been employed to increase the number of tiles to 20658. The networks were trained with a batch size of 2, using the Adam optimizer with a learning rate of 2.5e-4, momentum of 0.9 and weight decay of 0.0005. Moreover, a 5–fold cross validation strategy was employed over ten epochs, using 4/5 of the augmented tiles ($\sim$ 16525 images) for the training set, and the rest ($\sim$ 4131 images) for the test set.

3.2 Experimental Results

In Table 1, we present the results obtained with the 5–fold cross validation approach with the DeepLab V2 and MobileNet architectures. As we can observe,

on average, the two models achieve a similar Jaccard score even if a lower Dice index is obtained with the MobileNet. However, it is important to note that the number of glomeruli that are found inside each image is calculated more accurately with the MobileNet (difference of only 0.18). This demonstrates that, even if the precision of the segmentation is lower compared to the DeepLab, nonetheless it is possible to use the MobileNet to count the number of glomeruli inside an image in a precise way.

Fold	Dice	Jaccard	Mean $\Delta_{Glomeruli}$
0	0.90 ± 0.16	0.81 ± 0.34	0.35 ± 1.46
1	0.91 ± 0.14	0.79 ± 0.35	0.31 ± 1.55
2	0.92 ± 0.15	0.80 ± 0.34	0.14 ± 1.22
3	0.90 ± 0.15	0.78 ± 0.36	0.46 ± 1.6
4	0.93 ± 0.14	0.86 ± 0.28	0.22 ± 1.17
Folds Average	**0.91 ± 0.01**	**0.80 ± 0.03**	**0.29 ±0.12**

(a) DeepLab

Fold	Dice	Jaccard	Mean $\Delta_{Glomeruli}$
0	0.83 ± 0.33	0.81 ± 0.34	0.19 ± 1.06
1	0.78 ± 0.38	0.77 ± 0.38	0.02 ± 1.02
2	0.83 ± 0.33	0.82 ± 0.33	0.34 ± 1.40
3	0.79 ± 0.36	0.77 ± 0.37	0.24 ± 1.10
4	0.85 ± 0.31	0.83 ± 0.33	0.13 ± 1.26
Folds Average	**0.81 ± 0.02**	**0.80 ± 0.02**	**0.18 ± 0.11**

(b) MobileNet

Table 1: Results for the DeepLab V2 (a) and MobileNet (b) architectures for each validation fold.

Furthermore, in Figure 1, the comparison of the results obtained with both the DeepLab V2 and MobileNet, across the 5 folds, are shown, using boxplots. Instead, in Figure 2, we show the comparison between the $\Delta_{Glomeruli}$ index for the DeepLab and MobileNet experiments while, in Figure 3, we present an example of the segmentation obtained with the two networks.

3.3 Conclusions

In this work, we presented the application of two DL architectures, DeepLab V2 and MobileNet, to the segmentation of the renal glomeruli of mice. The results obtained open the use of such architectures for the segmentation of human glomeruli. In addition, employing lightweight networks, such as MobileNet, offers the opportunity to develop DL tools for mobile or embedded environments, such as electron microscopes. This could fundamentally help clinicians in the correct identification of glomerular structures, for a timely and accurate diagnosis of kidney health.

References

[1] Bianchini, Monica, et al. *Deep neural networks for structured data.* Computational Intelligence for Pattern Recognition. Springer, Cham, 2018. 29–51.

[2] Dimitri, Giovanna Maria, et al. *"Unsupervised stratification in neuroimaging through deep latent embeddings."* 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC). IEEE, 2020.

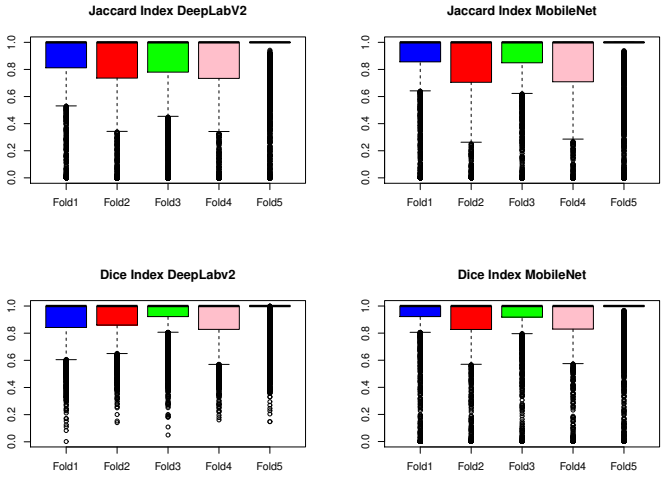


Fig. 1: Jaccard and Dice indices for the DeepLab and MobileNet architectures. In the Figure we can see the distribution of the two performance indicators across the 5–folds.

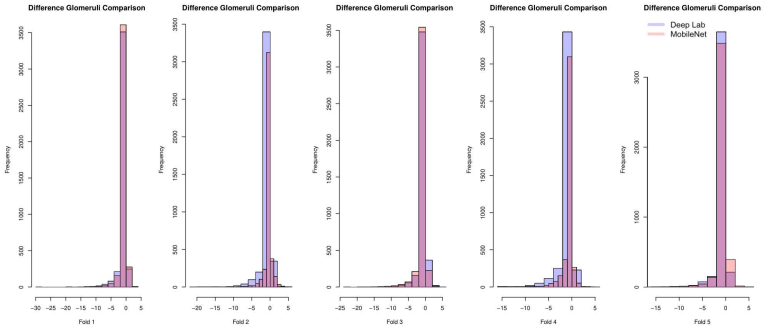


Fig. 2: $\Delta_{Glomeruli}$ index for the DeepLab and MobileNet architectures. Each histogram represents the distribution of the $\Delta_{Glomeruli}$ parameter for each fold. The two colors are used for the two models compared (DeepLab and Mobilenet).

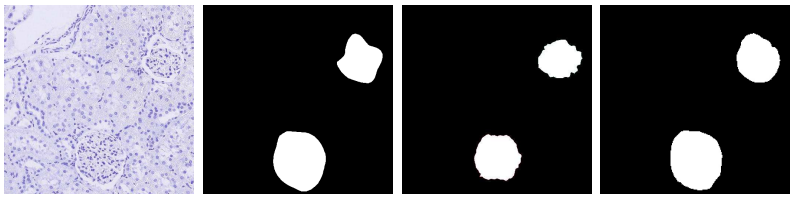
[3] Längkvist, Martin, Lars Karlsson, and Amy Loutfi. *A review of unsupervised feature learning and deep learning for time–series modeling.* Pattern Recognition Letters 42 (2014): 11–24.

[4] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. *Deep learning.* Nature 521.7553 (2015): 436–444.

[5] Richards, Blake A., et al. *A deep learning framework for neuroscience.* Nature neuroscience 22.11 (2019): 1761–1770.

[6] Bonechi, Simone, et al. *Segmentation of Aorta 3D CT Images Based on 2D Convolutional Neural Networks.* Electronics 10.20 (2021): 2559.

[7] Yu, Changqian, et al. *BiSeNet: Bilateral segmentation network for real–time semantic segmentation.* Proceedings of the European conference on computer vision (ECCV). 2018.



(a) (b) (c) (d)

Fig. 3: In (a), an image from the dataset, in (b), the corresponding ground–truth mask and, in (c) and (d), the DeepLab and MobileNet segmentation, respectively.

[8] Gupta, Laxmi, et al. *Stain independent segmentation of whole slide images: A case study in renal histology.* 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018.

[9] Stevens, Lesley A., et al. *Assessing kidney function–measured and estimated glomerular filtration rate.* New England Journal of Medicine 354.23 (2006): 2473–2483.

[10] D'Agati, Vivette D., Frederick J. Kaskel, and Ronald J. Falk. *Focal segmental glomerulosclerosis.* New England Journal of Medicine 365.25 (2011): 2398–2411.

[11] Dimitri, Giovanna Maria et al. *Deep Learning Approaches for the Segmentation of Glomeruli in Kidney Histopathological Images*, Mathematics 10.11 (2022): 1934.

[12] Kannan, Shruti, et al. *Segmentation of glomeruli within trichrome images using deep learning.* Kidney international reports 4.7 (2019): 955–962.

[13] Bouteldja, Nassim, et al. *Deep learning–based segmentation and quantification in experimental kidney histopathology.* Journal of the American Society of Nephrology 32.1 (2021): 52–68.

[14] Bueno, Gloria, et al. *Glomerulosclerosis identification in whole slide images using semantic segmentation.* Computer methods and programs in biomedicine 184 (2020): 105273.

[15] Ginley, Brandon G., et al. *Computational analysis of the structural progression of human glomeruli in diabetic nephropathy.* Medical Imaging 2018: Digital Pathology. Vol. 10581. International Society for Optics and Photonics, 2018.

[16] Sheehan, Susan M., and Ron Korstanje. *Automatic glomerular identification and quantification of histological phenotypes using image analysis and machine learning.* American Journal of Physiology–Renal Physiology 315.6 (2018): F1644–F1651.

[17] Qu, Chengqing, et al. *Three–Dimensional Visualization of the Podocyte Actin Network Using Integrated Membrane Extraction, Electron Microscopy, and Machine Learning.* Journal of the American Society of Nephrology 33.1 (2022): 155–173.

[18] Xu, Yanzhe, et al. *Improved small blob detection in 3D images using jointly constrained deep learning and Hessian analysis.* Scientific reports 10.1 (2020): 1–12.

[19] Stritt, Manuel, Anna K. Stalder, and Enrico Vezzali. *Orbit image analysis: an open–source whole slide image analysis tool.* PLoS computational biology 16.2 (2020): e1007313.

[20] Chen, Liang-Chieh, et al. *DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs.* IEEE Transactions on Pattern Analysis and Machine Intelligence 40.4 (2017): 834–848.

[21] Howard, Andrew G., et al. *MobileNets: Efficient convolutional neural networks for mobile vision applications.* arXiv preprint arXiv:1704.04861 (2017).

[22] Bradski, Gary. *The openCV library.* Dr. Dobb's Journal: Software Tools for the Professional Programmer 25.11 (2000): 120-123

[23] Dice, Lee R. (1945). *Measures of the Amount of Ecologic Association Between Species.* Ecology. 26 (3): 297â302. doi:10.2307/1932409. JSTOR 1932409.

[24] Murphy, Allan H. (1996). "The Finley Affair: A Signal Event in the History of Forecast Verification". Weather and Forecasting. 11 (1): 3.