Tau Protein Discrete Aggregates in Alzheimer's Disease: Neuritic Plaques and Tangles Detection and Segmentation using Computational Histopathology

Kristyna Maňoušková, Valentin Abadie, Mehdi Ounissi, Gabriel Jimenez, Lev Stimmer, Benoit Delatour, Stanley Durrleman, Daniel Racoceanu

Tau Protein Discrete Aggregates in Alzheimer's Disease: Neuritic Plaques and Tangles Detection and Segmentation using Computational Histopathology

K. Maňoušková, V. Abadie, M. Ounissi, G. Jimenez,
L. Stimmer, B. Delatour, S. Durrleman, D. Racoceanu

Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Inria, Hôpital de la Pitié-Salpêtrière, 47 bd de l'Hôpital, 75013 Paris, France

ABSTRACT

Tau proteins in the gray matter are widely known to be a part of Alzheimer's disease symptoms. They can aggregate in three different structures within the brain: neurites, tangles, and neuritic plaques. The morphology and the spatial disposition of these three aggregates are hypothesised to be correlated to the advancement of the disease. In order to establish a behavioural disease model related to the Tau proteins aggregates, it is necessary to develop algorithms to detect and segment them automatically. We present a 5-folded pipeline aiming to perform with clinically operational results. This pipeline is composed of a non-linear colour normalisation, a CNN-based image classifier, an Unet-based image segmentation stage, and a morphological analysis of the segmented objects. The tangle detection and segmentation algorithms improve state-of-the-art performances (75.8% and 91.1% F1-score, respectively), and create a reference for neuritic plaques detection and segmentation (81.3% and 78.2% F1-score, respectively). These results constitute an initial baseline in an area where no prior results exist, as far as we know. The pipeline is complete and based on a promising state-of-the-art architecture. Therefore, we consider this study a handy baseline of an impactful extension to support new advances in Alzheimer's disease. Moreover, building a fully operational pipeline will be crucial to create a 3D histology map for a deeper understanding of clinico-pathological associations in Alzheimer's disease and the histology-based evidence of disease stratification among different sub-types.

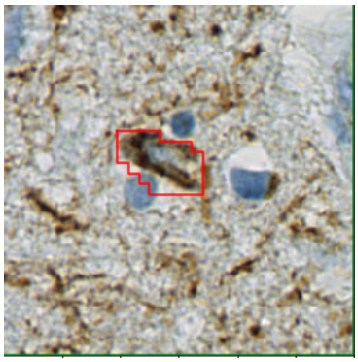
Keywords: Alzheimer's Disease, Tau Proteins, Tangles, Neuritic Plaques, Deep Learning, Computational Pathology, Whole Slide Images, Detection, Segmentation.

1. INTRODUCTION

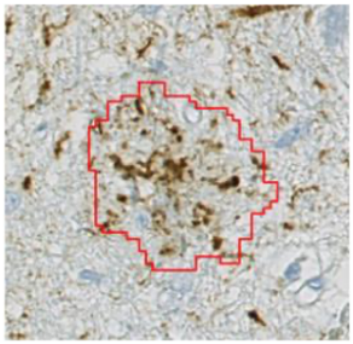
Tau proteins are essential markers of Alzheimer's Disease (AD) and probably present the best histopathological correlation with clinical symptoms.[1] However, it is yet unclear how these markers affect AD, qualitatively. Tau protein aggregates can be found in the gray matter within three types of structures: neurites, randomly and sparsely distributed proteins in neuronal processes; tangles, tau accumulations within the neuronal cell body; and neuritic plaques, spherical structures with tau positive dystrophic neurites surrounding Aβ positive core. We can find examples of these three different shapes in Figure 1.

The state-of-the-art technology used to assess tau proteins is based on histological whole slide image (WSI) analysis. WSI consists of a slice of the (post-mortem) gray matter of an AD patient that is stained, observed under a microscope, and digitized with several magnification levels (e.g. x20, and x40). A study presented by Signaevsky M. et al.[2] focused on neurofibrillary tangle segmentation in WSI. The authors reported a test F1-score of 0.81 using a fully convolutional neural network (SegNet) trained on the histopathological material from 22 brain autopsies with tauopathies. Wurts, A., et al., extended this initial study by comparing three deep learning models used to segment neurofibrillary tangles (NFT) and dystrophic neurites (DN). The authors combined color-based and manual annotations in a smaller dataset, and reported comparable performances of SegNet and FCN deep learning networks (0.611 best IoU for SegNet and tau segmentation, and 0.542 best IoU

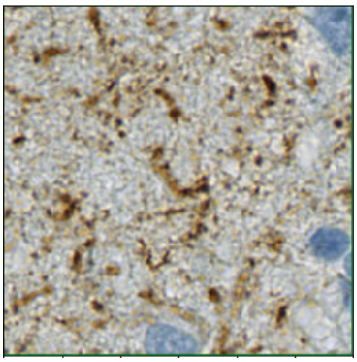
Further authors information: (Send correspondence to Daniel Racoceanu)
Daniel Racoceanu: E-mail: daniel.racoceanu@sorbonne-universite.fr, Telephone: +33 (0)6 33 77 42 67



(a) Example of *tangle* object (surrounded in red): the other brown stained small objects are *neurites*.



(b) *Plaque* object: brown stained little objects form the plaque neuritic crown surrounding the optically empty plaque core.



(c) Region only displaying tau-positive *neurites* (dark brown). Clearer objects form the background.

Figure 1: Examples of the 3 types of Tau proteins objects that can be found in the gray matter.

for FCN and nuclei segmentation). They finally suggested using a pre-trained UNet to increase its performance, as well as a more comprehensive labeling process.[3]

In this article, we present a novel end-to-end approach to process brain whole slide images based on convolutional neural network architectures and an Unet-based segmentation pipeline. In our study, the pipeline presented addresses not only the generation of an operational dataset suitable for a deep learning framework, but also the models deployed to perform detection and segmentation of the different tau protein aggregates.

2. METHODOLOGY

2.1 Data preparation

a. *Raw data:* comprises six post mortem human whole slide images, labeled by pathologists, with essentially four classes: background, tangles, neuritic plaques, and other (gray matter and neurites). The background is very fast to identify and annotate by a specialist; therefore, we focus on the automatic separation between pixels corresponding to tangles, neuritic plaques, and gray matter/neurites. The differentiation between gray matter and neurites is also essential, being easily performed using a simple stain classifier.

b. *Sampling:* to train and validate deep learning algorithms, we randomly sample fixed-size sub-images (i.e. patches) from WSI, controlling the fraction of images with tangles/neuritic plaques on them. The dataset generated is called Balanced Random Sampling (BRS). We computed our results with a balance of 50/50 positive/negative sub-images. Once trained, we use the deep learning algorithm to retrieve and segment every tangle and neuritic plaque of a given WSI, by exhaustively sampling it and feeding every sample to the model. We name this process Exhaustive Deterministic Sampling (EDS) and use it as a test dataset.

c. *Stain normalisation:* tau proteins' staining process is heterogeneous, hence a stain normalisation has been performed for the deep learning algorithm not to cope with scaling issues. We tested various normalisation methods (i.e, histogram normalization, Macenko,[4] Macenko-LS, Vahadane,[5] Vahadane-LS, Reinhard,[6] and Reinhard-LS) and ultimately, we retained the best 4 approaches. A different normalisation method was used for detecting and segmenting tangles and neuritic plaques. Additionally, we applied a pre-processing technique to datasets used for segmentation. It consists of standardising the brightness, i.e., modifying the luminosity channel so that 5% of pixels are fully luminous and thus white. We call this technique Luminosity Standardization (LS).

2.2 Deep learning pipeline architecture

a. *Tangles/neuritic plaques detection:* Eliminating all samples showing no tangle nor neuritic plaques is crucial to further segment these aggregates. Since this part aims to erase all samples with no tangles/neuritic plaques,

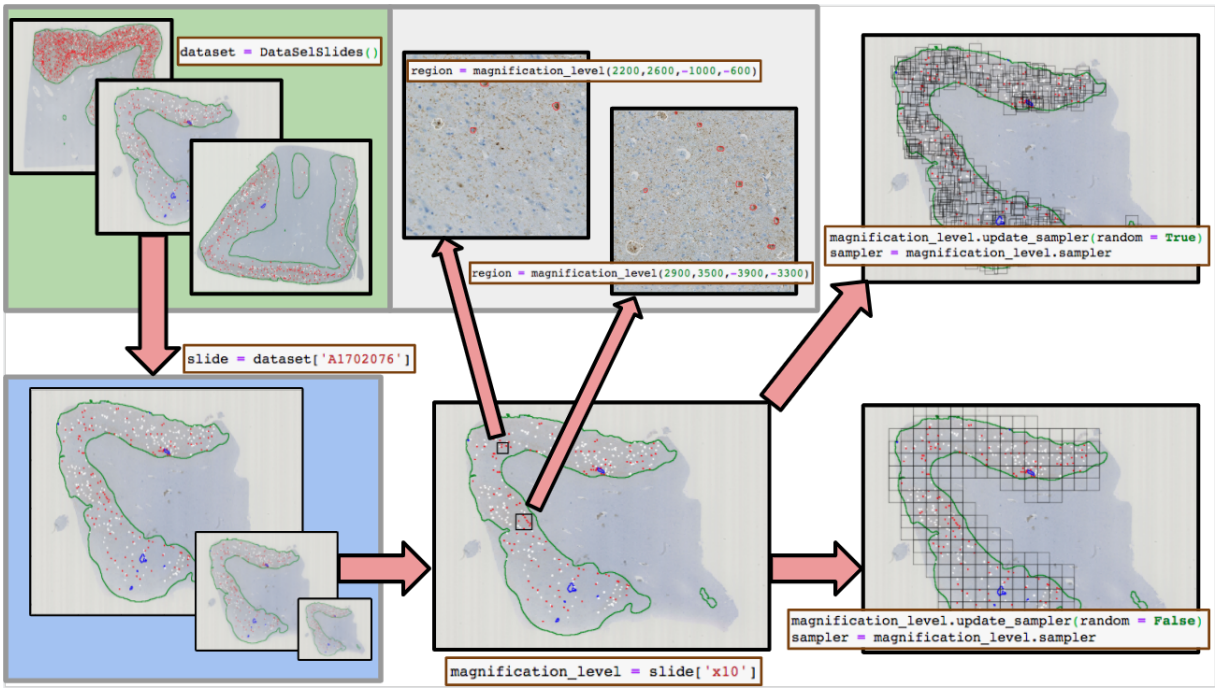


Figure 2: Whole pipeline for handling slides and their labels.

one should raise precision while keeping recall extremely high. We used a convolutional-based deep learning architecture, 5 WSI for training and validation (80%/20% split), and 1 WSI for testing. We also implemented 5-fold cross-validation to assess the CNN generalisation to an independent dataset (see Figure 3).

b. *Tangles/neuritic plaques segmentation:* The segmentation algorithm is trained on 5 WSI using 5-fold cross-validation and tested on 1 WSI. The input test dataset, called Objects Only Sampling (OOS), provides only samples with tangles or neuritic plaque. We used an Unet-based architecture with a focal loss function. A quantitative study combined with an F1-score measure (computed assuming an intersection between a ground-truth segmentation and model segmentation to be a true positive) was used to evaluate the performances (see Figure 4).

c. *Merging segmented regions:* During the test phase of segmentation, every region assessed by the classifier to be positive is segmented by the UNet model. We used this per-region segmentation to compute contour positions in meters corresponding to the slide at a global level. Furthermore, since some segmented objects may overlap or appear in two consecutive regions, we merged them by joining their corresponding contours. This stage allows for the reconstruction of the WSI and the computation of morphological features of the segmented objects (see Figure 5).

3. RESULTS

Table 1 shows the first results related to the study of different normalisation techniques for the detection and segmentation of both regions: neurofibrillary tangles and neuritic plaques. We reported the F1-score for testing as this metric is used to select the best algorithm for the desired task. We selected the histogram and Macenko normalisation for detection (in blue on the table), and Reinhard and Vahadane normalisation with LS for segmentation (in red on the table). The training process for these results was performed using patches of 128×128 pixels and a balanced dataset with 50/50 positives/negatives patches. However, for the testing phase we used the EDS dataset describe above with a stride of 64 pixels (i.e. overlap of 64 pixels between patches). The normalisation was applied in the trainig, validation and testing stages.

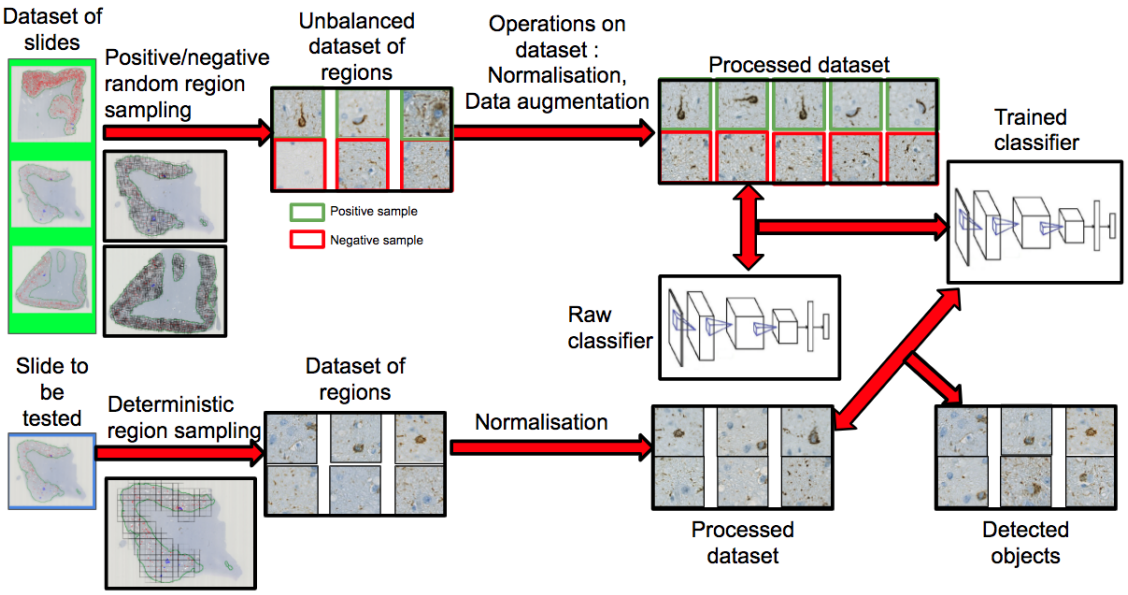


Figure 3: Overall scheme of detection of objects of interest on regions.

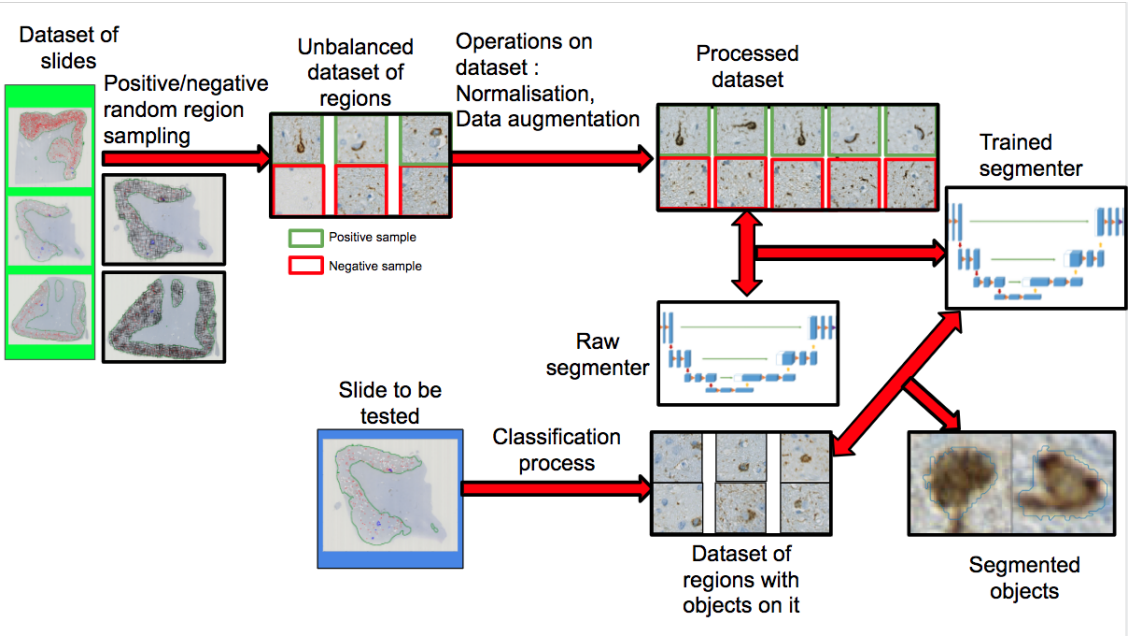


Figure 4: Overall scheme of segmentation of objects of interest on detected regions.

In addition to the cross-validation process implemented to evaluate the performance of the networks, we also applied a cross-testing/cross-validation scheme as shown in Table 2. Therefore, we report the mean performance of the networks after having tested the models with different WSI. As shown in Table 3, the detection task performs effectively for tangle aggregates but needs improvements to achieve clinically operational results. However, as shown in Table 4, the tangle segmentation task provides - according to our knowledge - important state-of-the-art results. On the other hand, the results shown represent an essential baseline for neuritic plaques detection and segmentation as no such study has been reported yet. Neuritic plaques are very diffuse objects, and a pure segmentation may be hard to grasp for a model. Despite these difficulties (see Figure 6b), the segmentation results seem very encouraging.

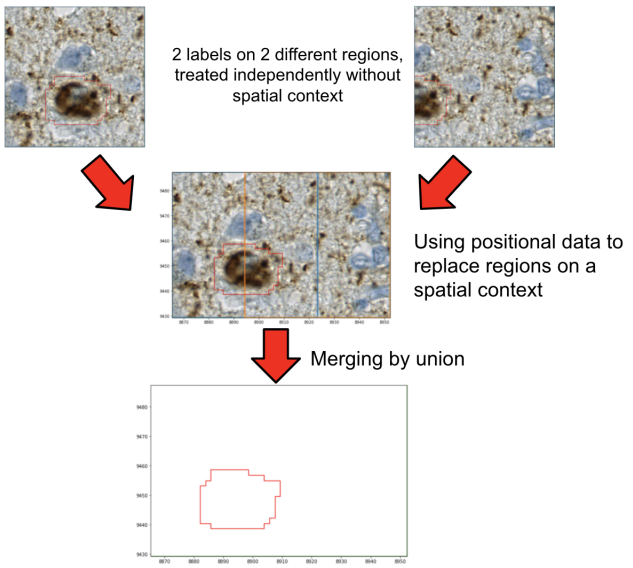


Figure 5: Illustration of merging of labels from different regions.

Table 1: Results for the object detection and segmentation pipeline using different normalization approaches.

Region	Patch size (pixels)	Normalization	Detection F1-score (testing)	Segmentation F1-score (testing)
Neurofibrillary tangle	128×128	**Histogram**	**82.7%**	92.9%
		Macenko	78.2%	92.2%
		Macenko (LS)	78.5%	88.7%
		Vahadane	76.0%	93.7%
		Vahadane (LS)	79.0%	91.4%
		Reinhard	75.7%	91.6%
		Reinhard (LS)	80.6%	**95.1%**
Neuritic plaques		Histogram	79.8%	40.2%
		Macenko	**97.9%**	55.8%
		Macenko (LS)	81.6%	68.5%
		Vahadane	91.2%	41.9%
		Vahadane (LS)	86.0%	**70.8%**
		Reinhard	85.1%	54.6%
		Reinhard (LS)	83.0%	60.9%

Table 2: Cross validation and cross testing procedure for segmentation and detection.

Trainig	Validation	Test
WSI_1, WSI_2, WSI_3, WSI_4	WSI_5	WSI_6
WSI_1, WSI_2, WSI_3, WSI_5	WSI_4	
WSI_1, WSI_2, WSI_4, WSI_5	WSI_3	
WSI_1, WSI_3, WSI_4, WSI_5	WSI_2	
WSI_2, WSI_3, WSI_4, WSI_5	WSI_1	

...

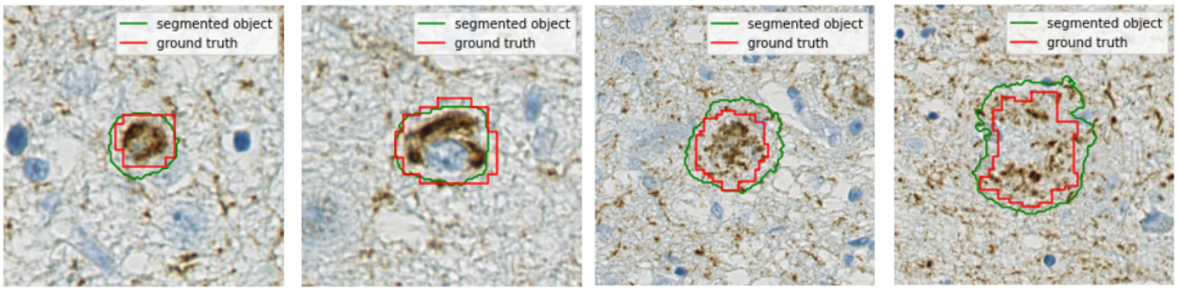
Trainig	Validation	Test
WSI_2, WSI_3, WSI_4, WSI_5	WSI_6	WSI_1
WSI_2, WSI_3, WSI_4, WSI_6	WSI_5	
WSI_2, WSI_3, WSI_5, WSI_6	WSI_4	
WSI_2, WSI_4, WSI_5, WSI_6	WSI_3	
WSI_3, WSI_4, WSI_5, WSI_6	WSI_2	

Table 3: Results for detection of **tangles** and **plaques** using cross-validation and cross-testing.

Region	Patch size (pixels)	Normalization	F1-score (cross-validation)	F1-score (cross-testing)
Neurofibrillary tangle	128×128	**Histogram**	99.6%	75.8%
Neuritic plaques	128×128	**Macenko**	99.8%	81.3%

Table 4: Results for segmentation of **tangles** and **plaques** using cross-validation and cross-testing.

Region	Patch size (pixels)	Normalization	F1-score (cross-validation)	F1-score (cross-testing)
Neurofibrillary tangle	128 × 128	**Reinhard (LS)**	83.8%	91.1%
Neuritic plaques	128 × 128	**Vahadane (LS)**	81.5%	78.2%



(a) Examples of **tangle** segmentation. (b) Examples of **neuritic plaques** segmentation.

Figure 6: Examples of the 3 types of tau proteins objects that can be found in the gray matter.

4. DICUSSION AND CONCLUSION

This study has proven that we can generate clinically operational results for tau aggregates detection and segmentation by providing adequate WSI sampling and the proper normalization method. Future research will focus on bridging the gap between these promising results and real use cases to reduce the number of false negatives samples driven by hyper-parameter tuning. In particular, an extension to this dataset with additional slides from different patients, as well as different staining/scanning procedures, might help robustifying the pipelines presented in this work. Furthermore, from the segmented tau aggregates studied, a morphological analysis could also be implemented to deeper studying Alzheimers' patients stratification and to better understanding this critical brain disorder.

Finally, the pipelines presented in this study could also be the core backend of a semi-supervised WSI annotator. This will be definetly likely to reduce the time spent by pathologists for WSI annotation and improve the quality of this annotations, enabling the elaboration of a more reliable, effective and efficient AI tools.

ACKNOWLEDGMENTS

This research is supported by Mr Jean-Paul Baudecroux and The Big Brain Theory Program - Paris Brain Institute (ICM).

REFERENCES

[1] Duyckaerts, C., Delatour, B., and Potier, M.-C., "Classification and basic pathology of Alzheimer disease," *Acta Neuropathologica* **118**, 5–36 (July 2009).

[2] Signaevsky, M., Prastawa, M., Farrell, K., Tabish, N., Baldwin, E., Han, N., Iida, M. A., Koll, J., Bryce, C., Purohit, D., Haroutunian, V., McKee, A. C., Stein, T. D., White, C. L., Walker, J., Richardson, T. E., Hanson, R., Donovan, M. J., Cordon-Cardo, C., Zeineh, J., Fernandez, G., and Crary, J. F., "Artificial intelligence in neuropathology: deep learning-based assessment of tauopathy," *Laboratory Investigation* **99**, 1019–1029 (July 2019).

[3] Wurts, A., Oakley, D. H., Hyman, B. T., and Samsi, S., "Segmentation of Tau Stained Alzheimers Brain Tissue Using Convolutional Neural Networks," in [*42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*], 1420–1423, IEEE, Montreal, QC, Canada (July 2020).

[4] Macenko, M., Niethammer, M., Marron, J. S., Borland, D., Woosley, J. T., Guan, X., Schmitt, C., and Thomas, N. E., "A method for normalizing histology slides for quantitative analysis," in [*2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*], 1107–1110, IEEE (2009).

[5] Vahadane, A., Peng, T., Albarqouni, S., Baust, M., Steiger, K., Schlitter, A., Sethi, A., Esposito, I., and Navab, N., "Structure-preserved color normalization for histological images," 1012–1015 (04 2015).

[6] Reinhard, E., Adhikhmin, M., Gooch, B., and Shirley, P., "Color transfer between images," *IEEE Computer graphics and applications* **21**(5), 34–41 (2001).

[7] Calderon-Garcidueñas, A. L. and Duyckaerts, C., "Chapter 23 - alzheimer disease," in [*Neuropathology*], Kovacs, G. G. and Alafuzoff, I., eds., *Handbook of Clinical Neurology* **145**, 325 – 337, Elsevier (2018).

[8] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," *CoRR* **abs/1505.04597** (2015).

[9] "Supplement 122: Specimen module and revised pathology sop classes," *Digital Imaging and Communications in Medicine (DICOM)* .

[10] "Supplement 145: Whole slide microscopic image iod and sop classes," *Digital Imaging and Communications in Medicine (DICOM)* .

[11] Bradski, G. and Kaehler, A., [*Learning OpenCV: Computer vision with the OpenCV library*], " O'Reilly Media, Inc." (2008).

[12] Goode, A., Gilbert, B., Harkes, J., Jukic, D., and Satyanarayanan, M., "Openslide: A vendor-neutral software foundation for digital pathology," *Journal of pathology informatics* **4** (2013).

[13] Gillies, S., "The shapely user manual," *URL https://pypi. org/project/Shapely* (2013).

[14] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P., "Focal loss for dense object detection," in [*Proceedings of the IEEE international conference on computer vision*], 2980–2988 (2017).

[15] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980* (2014).

[16] Tang, Z., Chuang, K. V., DeCarli, C., Jin, L.-W., Beckett, L., Keiser, M. J., and Dugger, B. N., "Interpretable classification of alzheimer's disease pathologies with a convolutional neural network pipeline," *Nature communications* **10**(1), 1–14 (2019).

[17] Khan, A. M. et al., "A nonlinear mapping to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE TBME* **61**(6), 1729–1738 (2014).

[18] Janowczyk, A. and Madabhushi, A., "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *Journal of pathology informatics* **7** (2016).

[19] Song, J., Xiao, L., Molaei, M., and Lian, Z., "Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images," *Knowledge-Based Systems* **176**, 40–53 (2019).

[20] Zemouri, R., Zerhouni, N., and Racoceanu, D., "Deep learning in the biomedical applications: Recent and future status," *Applied Sciences* **9**(8), 1526 (2019).

[21] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q., "Densely connected convolutional networks," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 4700–4708 (2017).

[22] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556* (2014).

[23] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 770–778 (2016).

[24] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., and Sánchez, C. I., "A survey on deep learning in medical image analysis," *Medical image analysis* **42**, 60–88 (2017).

[25] Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., et al., "Qupath: Open source software for digital pathology image analysis," *Scientific reports* **7**(1), 1–7 (2017).