Master's Thesis: Overcoming Inter-Scanner Variance of Pathological Whole-Slide Images

Ausgleich von scannerinduzierter Varianz in digitalisierten histopathologischen Präparaten

Supervisors: Frauke Wilm, Christian Marzahl, Prof. Katharina Breininger, Prof. Marc Aubreville, Prof. Andreas Maier, Prof. Robert Klopfleisch

With the development of automatic slide scanning systems, the manual evaluation of histopathological samples using light microscopy is more and more accompanied by automatic image analysis algorithms. These methods use high resolution digitizations of microscopic slides to train machine learning models, which are then able to infer learned features onto unseen data. Whilst this works well on data from the same laboratory, these algorithms are oftentimes challenged by imaging data from different laboratories or scanning systems. These inter-scanner variations are attributable to scanning configurations like resolution, illumination and contrast as well as artefacts or blurring [1]. Previous works have mostly focussed on color discrepancies of histological samples [2,3]. Some of these approaches, e.g. the use of a calibration slide to align scanner outputs [4], can be used as a starting point for resolving scanner-dependent variations.

Generally, different approaches for resolving inter-scanner variations can be considered [5]:

1. the use of augmentation methods to artificially generated training samples that represent the variability of input data (e.g. colour augmentation in [2])
2. integration of a preprocessing step to reduce variability before training (e.g. stain normalization in [3])
3. constraining the network to learn a feature representation that is scanner invariant (e.g. [1, 5])

This thesis aims to investigate different strategies to overcome domain shift across laboratories and scanner types for machine learning applications such as tissue segmentation.

The thesis comprises the following items:

- literature review concerning sources and state-of-the-art approaches to tackle device-induced variations in deep learning, with a focus on microscopy image data
- implementation of one or more solutions to overcome common sources of visual variability (starting with domain-adversarial approach)
- integration into existing image analysis pipelines
- evaluation of proposed method and emerging challenges attributable to the image domain (e.g. focus points in cytological data)
- documentation and presentation of the findings, documentation of code

[1] Moyes, Andrew, et al. "A Novel Method For Unsupervised Scanner-Invariance With DCAE Model." (2018).

[2] Tellez, David, et al. "H&E stain augmentation improves generalization of convolutional networks for histopathological mitosis detection." Medical Imaging 2018: Digital Pathology. Vol. 10581. International Society for Optics and Photonics, 2018.

[3] Macenko, Marc, et al. "A method for normalizing histology slides for quantitative analysis." 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, 2009.

[4] https://ffei.ai

[5] Lafarge, Maxime W., et al. "Domain-adversarial neural networks to address the appearance variability of histopathology images." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2017. 83-91.