Abstract: How Big is Big Enough?
A Large-Scale Histological Dataset of Mitotic Figures

Christof A. Bertram[1], Marc Aubreville[2], Christian Marzahl[2], Andreas Maier[2], Robert Klopfleisch[1]

[1]Institute of Veterinary Pathology, Freie Universität Berlin, Germany
[2]Pattern Recognition Lab, F.-Alexander-Universität Erlangen-Nürnberg, Germany
`christof.bertram@fu-berlin.de`

Quantification of mitotic figures (MF) within the tumor areas of highest mitotic density is the most important prognostic parameter for outcome assessment of many tumor types. However, high intra- and inter-rater variability results from difficulties in individual MF identification and region of interest (ROI) selection due to uneven MF distribution. Deep learning-based algorithms for MF detection and ROI selection are very promising methods to overcome these limitations. As of today, few datasets of human mammary carcinoma are available. They provide labels only in small image sections of the whole slide image (WSI) and include up to 1,552 MF annotations [1].

Our research group has developed a large-scale, open access dataset with annotations for MF in 32 cases of canine cutaneous mast cell tumors [1]. Entire WSI were completely labeled by two pathologists resulting in 44,800 MF annotations. Of those, 5.5% were initially missed by expert WSI screening and added through a deep learning-based pipeline for identification of potential candidates.

For algorithmic validation, we used a two-stage approach (RetinaNet followed by cell classificator), which yielded a F1 score of 0.820. Through the algorithm-aided completion of the dataset we were able to increase the F1 score by 3.4 percentage points. Influence of the size of the dataset was assessed by stepwise reduction of the number of WSI and size (in high power fields, HPF) of the image sections used for training. With the number of included images, the F1 score moderately increased (3 WSI: 0.772; 6 WSI: 0.804; 12 WSI: 0.817; 21 WSI: 0.820). The size of the tumor area in training (ROI selected by an expert) had significant effects on the F1 score (5 HPF: 0.583; 10 HPF: 0.676; 50 HPF: 0.770; complete WSI: 0.820), which was determined in entire WSI of the test set. We emphasize the benefit of appropriate dataset size and complete WSI labeling.

References

1. Bertram CA, Aubreville M, Marzahl C, et al. A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor. Sci Data. 2019;6(274):1–9.