



Abstract: the MIDOG Challenge 2021

Mitosis Domain Generalization in Histopathology Images

Marc Aubreville¹, Nikolas Stathonikos², Christof A. Bertram³, Robert Klopffleisch⁴,
Natalie ter Hoeve², Francesco Ciompi⁵, Frauke Wilm⁶, Christian Marzahl⁶,
Taryn A. Donovan⁷, Andreas Maier⁶, Mitko Veta⁸, Katharina Breininger⁶,
and the MIDOG contributors

¹Technische Hochschule Ingolstadt, Ingolstadt, Germany

²UMC Utrecht, Utrecht, The Netherlands

³University of Veterinary Medicine, Vienna, Austria

⁴Freie Universität Berlin, Berlin, Germany

⁵Radboud UMC, Nijmegen, The Netherlands

⁶Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

⁷Schwarzman Animal Medical Center, New York, USA

⁸TU Eindhoven, Eindhoven, The Netherlands

marc.aubreville@thi.de

The density of mitotic figures (MF) within tumor tissue is known to be highly correlated with tumor malignancy and thus is an important marker in tumor grading. Recognition of MF by pathologists is subject to a strong inter-rater bias, limiting its prognostic value. State-of-the-art deep learning methods can support experts but have been observed to strongly deteriorate in a different clinical environment. The variability caused by using different whole slide scanners has been identified as one decisive component in the underlying domain shift. The goal of the MICCAI Mitosis Domain Generalization (MIDOG) 2021 challenge was the creation of scanner-agnostic MF detection algorithms. It was the first challenge to be held on the topic of histology domain generalization. The challenge used a training set of 200 cases, split across four scanning systems. The test set comprised an additional 100 cases split across four scanning systems, including two previously unseen scanners. We evaluated and compared the approaches that were submitted to the challenge and identified methodological factors contributing to better performance. The winning algorithm yielded an F_1 score of 0.748 (CI95: 0.704-0.781). Additionally, we compared the performance of the algorithms to six pathologists. Irrespective of whether the original challenge ground or a newly generated object-level consensus of five experts was used as the basis for evaluation, we found that the best algorithms achieved higher performance levels than all experts. Further, the best algorithm reached the highest F_1 values on either definition of the ground truth, reaching values of 0.742 and 0.748 on the five-expert consensus and the original challenge ground truth, respectively [1].

References

1. Aubreville M, Stathonikos N et al. Mitosis domain generalization in histopathology images—The MIDOG challenge. *Med Image Anal.* 2023;84:102699.