Abstract: Are Fast Labeling Methods Reliable?
A Case Study of Computer-aided Expert Annotations on Microscopy Slides

Christian Marzahl[1,2], Christof A. Bertram[3], Marc Aubreville[1], Anne Petrick[3], Kristina Weiler[4], Agnes C. Gläsel[4], Marco Fragoso[3], Sophie Merz[3], Florian Bartenschlager[3], Judith Hoppe[3], Alina Langenhagen[3], Anne Katherine Jasensky[5], Jörn Voigt[2], Robert Klopfleisch[3], Andreas Maier[1]

[1]Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany
[2]R & D Projects, EUROIMMUN Medizinische Labordiagnostika AG
[3]Institute of Veterinary Pathology, Freie Universität Berlin, Germany
[4]Department of Veterinary Clinical Sciences, Clinical Pathology and Clinical Pathophysiology, Justus-Liebig-Universität Giessen, Germany
[5]Laboklin GmbH und Co. KG, Bad Kissingen, Germany
c.marzahl@euroimmun.de

Deep-learning-based pipelines have shown the potential to revolutionize microscopy image diagnostics by providing visual augmentations and evaluations to a pathologist. However, to match human performance, the methods rely on the availability of vast amounts of high-quality labeled data, which poses a significant challenge. To circumvent this, augmented labeling methods, also known as expert-algorithm-collaboration, have recently become popular. However, potential biases introduced by this operation mode and their effects on training deep neuronal networks are not entirely understood [1]. This work aimed to evualte this for three pathological pattern of interest. Ten trained pathology experts performed a labeling tasks without and with computer-generated augmentation. To investigate different biasing effects, we intentionally introduced errors to the augmentation. In total, experts annotated 26,015 cells on 1,200 images in this novel annotation study. Backed by this extensive data set, we found that the concordance of multiple experts was significantly increased in the computer-aided setting, versus the unaided annotation. However, a significant percentage of the deliberately introduced false labels was not identified by the experts.

References

1. Marzahl C, Bertram CA, Aubreville M, et al. Are fast labeling methods reliable? A case study of computer-aided expert annotations on microscopy slides. Proc MICCAI. 2020; p. 24–32.