Digital Health - Machine Learning, Chair of Prof. Dr. Lippert
Master’s Thesis Proposal

Contact: benjamin.bergner@hpi.de

Master Thesis: Correcting Confounders in Deep Image Classification Networks

Deep neural networks are remarkable image classifiers. However, sometimes they compute the correct class for the wrong reasons. For example, an image might be assigned the label ‘airplane’ even though the network is only looking for the blue sky surrounding it. More critically, a computer-aided diagnosis system could, for example, identify scanner types instead of medical findings relevant for a disease [1]. This is problematic since the model suggests good performance but may fail utterly in practice. In this thesis, you will develop and evaluate a method to correct such confounders by means of strong supervision. For example, marking target objects (such as medical findings) can support the model distinguishing signals from noise in the image. An important research question is then: How many strongly labeled samples are necessary to resolve the confounding variables?

You may use existing ideas from the chair to get started, and you may use both natural or medical imaging datasets for your experiments. It is optionally possible to identify confounders in a medical dataset, for which corresponding domain knowledge is required. To be successful, you should additionally have hands-on experience with a major framework, such as PyTorch (which you can, for example, acquire in our deep learning course). If you are highly motivated and interested in this topic, please contact benjamin.bergner@hpi.de

References

[1] Zech, John R., et al. "Confounding variables can degrade generalization performance of radiological deep learning models." arXiv preprint arXiv:1807.00431 (2018).