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

Contact: benjamin.bergner@hpi.de

Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Master Thesis: Self-supervised Multiple Instance Learning

Self-supervised learning (SSL) is an unsupervised representation learning technique and a hot topic in deep learning. It involves solving an artificial task that allows a network to learn the semantics of a dataset. The resulting feature extractor can then be used for transfer learning to reduce the number of labeled examples needed to solve the actual downstream task. This is of great practical value for computer-aided diagnosis, since medical experts are required for labeling, which is expensive [1]. SSL methods are often applied on image patches (e.g. puzzle solving [2], context prediction [3], contrastive learning [4] or vision transformers [5]), whereas the downstream task commonly works with whole images. This discrepancy requires the implementation of SSL and transfer learning in two separate and sequential steps. One technique that inherently uses patches is Multiple Instance Learning [6], and thus might be a more natural choice for many SSL methods.

In this thesis, you will design and evaluate SSL methods in the MIL setting, which raises several interesting research questions: (1) Can SSL MIL outperform standard SSL methods? (2) Can an SSL task and a downstream task be learned in parallel? (3) What are the implications for training time and performance?

You may use existing ideas from the chair to get started, and you may use both natural or medical imaging datasets for your experiments. To be successful, you should have a solid understanding of deep learning and hands-on experience with a major framework, such as PyTorch (both of which you can, for example, acquire in our deep learning course). If you are interested in this topic and some of the references listed below, please contact benjamin.bergner@hpi.de

References

[1] Taleb, Aiham, et al. "3d self-supervised methods for medical imaging." arXiv preprint arXiv:2006.03829 (2020).
[2] Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." European conference on computer vision. Springer, Cham, 2016.
[3] Doersch, Carl, Abhinav Gupta, and Alexei A. Efros. "Unsupervised visual representation learning by context prediction." Proceedings of the IEEE international conference on computer vision. 2015.
[4] Oord, Aaron van den, Yazhe Li, and Oriol Vinyals. "Representation learning with contrastive predictive coding." arXiv preprint arXiv:1807.03748 (2018).
[5] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).
[6] https://en.wikipedia.org/wiki/Multiple_instance_learning