

## **Master Thesis: Detecting Difficult and Out-of-Distribution Samples in Image Classification**

Deep learning models can achieve remarkable performance in image classification. However, they also tend to be overconfident, i.e. assign high class probabilities even if the computed class is incorrect [1]. This limits the applicability and trust in deep learning models in critical applications such as computer-aided diagnosis, where wrong decisions can have severe consequences for patients. With a model that can tell when it is uncertain, a radiologist could identify difficult cases that need more careful review. On the other hand, an AI scientist could identify samples that are underrepresented in the training data to iterate the model.

In this thesis, you will develop and evaluate a novel method capable of detecting difficult and out-of-distribution samples. 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, please contact [benjamin.bergner@hpi.de](mailto:benjamin.bergner@hpi.de)

### References

[1] Guo, Chuan, et al. "On calibration of modern neural networks." International Conference on Machine Learning. PMLR, 2017.