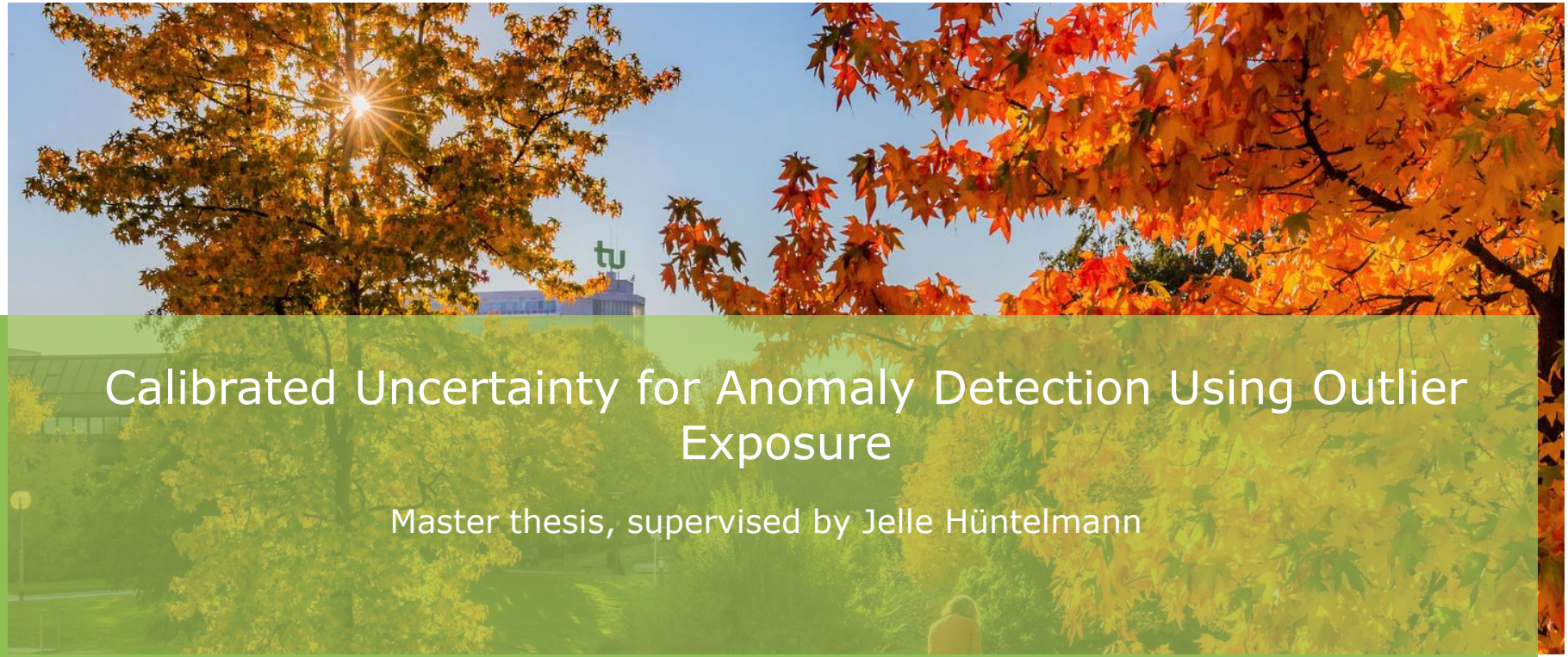
technische universität dortmund

Chair of Data Science and Data Engineering
Prof. Dr. Emmanuel Müller



Calibrated Uncertainty for Anomaly Detection Using Outlier Exposure

Master thesis, supervised by Jelle Hüntelmann

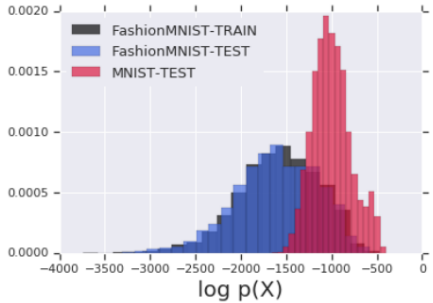
Motivation

- ***Calibrated Uncertainty***: the model has *low* uncertainty for inputs similar to the training data and *high* uncertainty elsewhere. Its predictions are mostly correct where uncertainty is low.
  - *Application*: Use an uncertainty threshold to reject predictions that are less likely to be correct.

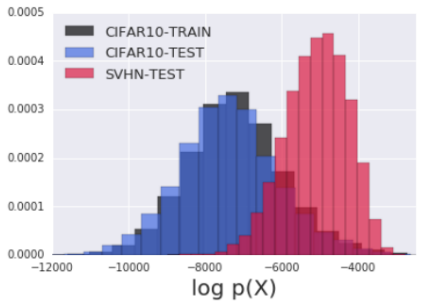
- ***Anomaly Detection***: estimate the density of the training distribution. If the input is from a low-density area, it's probably an anomaly!
  - Model uncertainty should be high and this input is unexpected → raise the alarm to a supervisor!

- ***The Problem***: Several popular deep model architectures fail at this task! → they predict higher likelihood ("normality") for unrelated samples than for ones from the actual training distribution

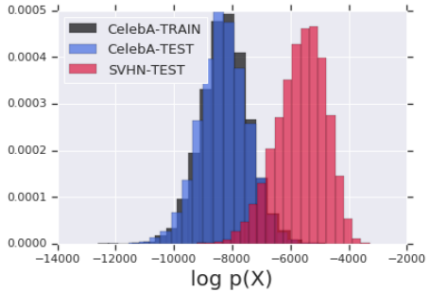
Outliers are more “normal” than training samples!? [1]



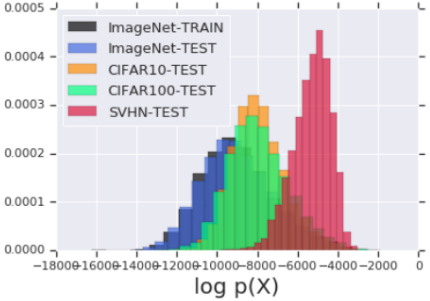
(a) Train on FashionMNIST, Test on MNIST



(b) Train on CIFAR-10, Test on SVHN



(c) Train on CelebA, Test on SVHN



(d) Train on ImageNet,
Test on CIFAR-10 / CIFAR-100 / SVHN

Image from [1] Nalisnick et al., Do Deep Generative Models Know What They Don’t Know?, ICLR 2019

Your Task

- Hendrycks et al. [2] introduced *Outlier Exposure* to improve anomaly detection performance. They also claim that their method mitigates the effect of incorrect density estimates for OOD samples.

- ***Your Task***: Investigate Outlier Exposure as a method to improve uncertainty calibration.
  - Devise your own experiments to ***demonstrate poor calibration***
  - Choose several uncertainty quantification methods and ***measure the improvement (if any)*** Outlier Exposure delivers, identify limitations.

- ***Requirements***:
  - You will need **good programming skills** in Python and preferably experience with **PyTorch** to understand existing code and train and tune new models or implement new methods.

[2] Hendrycks et al., Deep Anomaly Detection with Outlier Exposure, ICLR 2019

Organizational details

- **Master** thesis, available from approx. **November**
- Paper: **Deep Anomaly Detection with Outlier Exposure**, *Hendrycks et al., ICLR 2019* (scan code for github, which also links to the paper)
- Supervised by **Jelle** (jelle.huentelmann@cs.tu-dortmund.de), reach out for more details and information about the qualification task!