Evaluating Uncertainty Quantification in Visual Anomaly Detection

This thesis will be done in collaboration with Siemens.

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MATH 4AI — Bavarian AI Chair for Mathematical Foundations of Artificial Intelligence

Supervisor:
Prof. Dr. Gitta Kutyniok

Advisor:
Carlo Kneissl

Contact:
kneissl@math.lmu.de

Type:
MA

Research area:
Uncertainty Quantification; Anomaly Detection

Programming language:
python

Required skills:
probablity/bayesian methods; pytorch

Introduction

In an industrial context, visual anomaly detection aims to detect whether an inspected object differs from a training set of normal (defect-free) images. Approaches can roughly be categorized into one-class classification, probabilistic, and reconstruction-based methods [6]. The latter learn a representation of normality (e.g., a manifold or prototypes) and detect anomalies by *failing* to accurately reconstruct them under the learned model.

A recent reconstruction method is Reverse Distillation (RD) [2] with a teacher–student architecture where a fixed, pretrained teacher encoder extracts multi-scale features and a trainable student (via a one-class bottleneck embedding plus a decoder) learns on normal data to reconstruct those features. At inference, anomalies are indicated by high discrepancies (low cosine similarity) between teacher features and reconstructed features, yielding an anomaly map and an image-level anomaly score.

Uncertainty-aware anomaly detection

Most reconstruction approaches such as RD are typically deterministic and output a single anomaly score (or anomaly map) per input, i.e., a point estimate of reconstruction discrepancy. However, in realistic deployments the model may be *poorly determined* by available normal data (few samples, missing normal modes), affected by nuisance variation (lighting, viewpoint), or confronted with domain shift (new camera, dirty lens). The uncertainty arising from such an incomplete specification of the model by the training data is denoted as *epistemic* uncertainty. Additionally, the image data itself may include inherent randomness (sensor noise, ambiguities), described as *aleatoric* uncertainty. The thesis should mainly consider epistemic uncertainty; but in practice, it may not always be possible or even required to distinguish between epistemic and aleatoric.[1] Thus, in the following, we subsume both as *predictive* uncertainty, indicating how certain the model is regarding the anomaly score or map.

[1] For example, lighting variation can be aleatoric in case of flicker or specular highlights that change unpredictably, or epistemic in case of a new light source not present during training.

Research questions

The research questions to be addressed in this thesis are:

- **RQ 1: Is Uncertainty Quantification (UQ) useful in RD anomaly detection?** E.g., can uncertainty estimates help to *flag potentially unreliable cases* and improve decision policies (e.g., abstain/review)?
- **RQ 2: What are scalable UQ methods [3, 5] in this task?** Deep ensemble is the baseline. MC dropout or Test-Time-Augmentation (TTA) may be good candidates to compare. For MC dropout, an interesting aspect is where to integrate it in the RD architecture, e.g., the bottleneck network may be a good start.

The questions should be evaluated on the MVTec AD [1] and/or MVTec AD2 [4].

References

[1] Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, and Carsten Steger. The MVTec anomaly detection dataset: A comprehensive real-world dataset for unsupervised anomaly detection. *Int. J. Comput. Vision*, 129(4):1038–1059, April 2021.

[2] Hanqiu Deng and Xingyu Li. Anomaly detection via reverse distillation from one-class embedding, 2022.

[3] Jakob Gawlikowski, Cedrique Rovile Njieutcheu Tassi, Mohsin Ali, Jongseok Lee, Matthias Humt, Jianxiang Feng, Anna Kruspe, Rudolph Triebel, Peter Jung, Ribana Roscher, Muhammad Shahzad, Wen Yang, Richard Bamler, and Xiao Xiang Zhu. A survey of uncertainty in deep neural networks, 2022.

[4] Lars Heckler-Kram, Jan-Hendrik Neudeck, Ulla Scheler, Rebecca König, and Carsten Steger. The mvtec ad 2 dataset: Advanced scenarios for unsupervised anomaly detection, 2025.

[5] Adrien Lafage, Olivier Laurent, Firas Gabetni, and Gianni Franchi. Torch-uncertainty: A deep learning framework for uncertainty quantification. In *NeurIPS Datasets and Benchmarks Track*, 2025.

[6] Lukas Ruff, Jacob R. Kauffmann, Robert A. Vandermeulen, Gregoire Montavon, Wojciech Samek, Marius Kloft, Thomas G. Dietterich, and Klaus-Robert Müller. A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109(5):756–795, May 2021.