Quantifying Statistical Significance of Deep Nearest Neighbor Anomaly Detection via Selective Inference

Mizuki Niihori, Shuichi Nishino, Teruyuki Katsuoka, Tomohiro Shiraishi, Kouichi Taji, Ichiro Takeuchi

Nagoya University

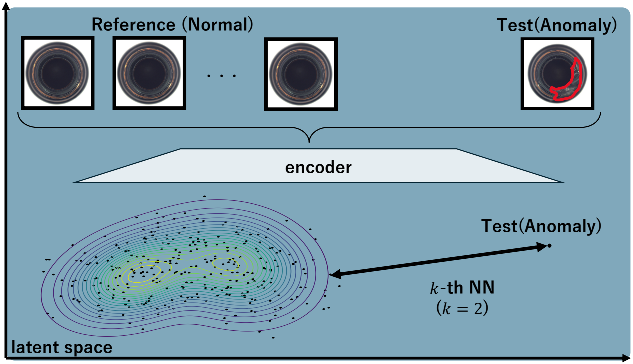
RIKEN

NeurIPS2025

NAGOYA UNIVERSITY

Introduction

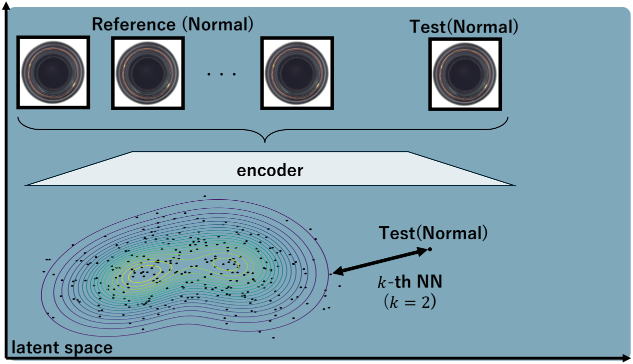
- KNN-based AD(kNNAD) is one of the simplest yet effective methods, which is widely used
  - Performs kNN-based anomaly detection in latent space
  - Effective for high-dimensional and unstructured data such as images



- However, the reliability of detected anomalies is not quantified
- We propose a method to quantify the statistical significance of kNNAD results

Statistical Significance in AD

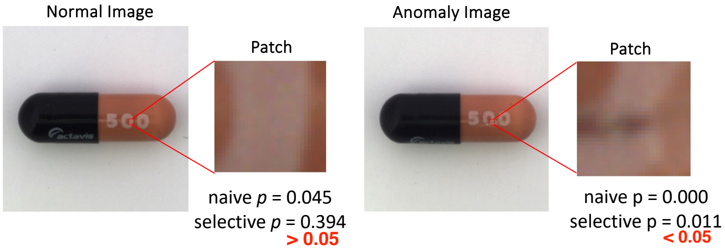
- When a normal image is inputted, kNNAD may **incorrectly** predict normal images as anomaly (**false positive**)



- We quantify the statistical significance of kNNAD as a valid p-value, enabling control of the false positive rate (FPR)

Issue and Our Approach

- We quantify the statistical significance of AD by p-values
- Naive p-values are often underestimated, failing to control the FPR
- This is due to selection bias caused by using the same data for both detection and testing
- We apply **Selective Inference** to rigorously correct this selection bias, enabling proper control of the FPR



MVTec Image Examples