QUANTIFYING STATISTICAL SIGNIFICANCE FOR DEEP SEMI-SUPERVISED ANOMALY DETECTION VIA SELECTIVE INFERENCE

Đặng Quang Vinh [1] Cao Lê Công Thành [1]

[1] University of Information Technology, Vietnam National University - Ho Chi Minh City

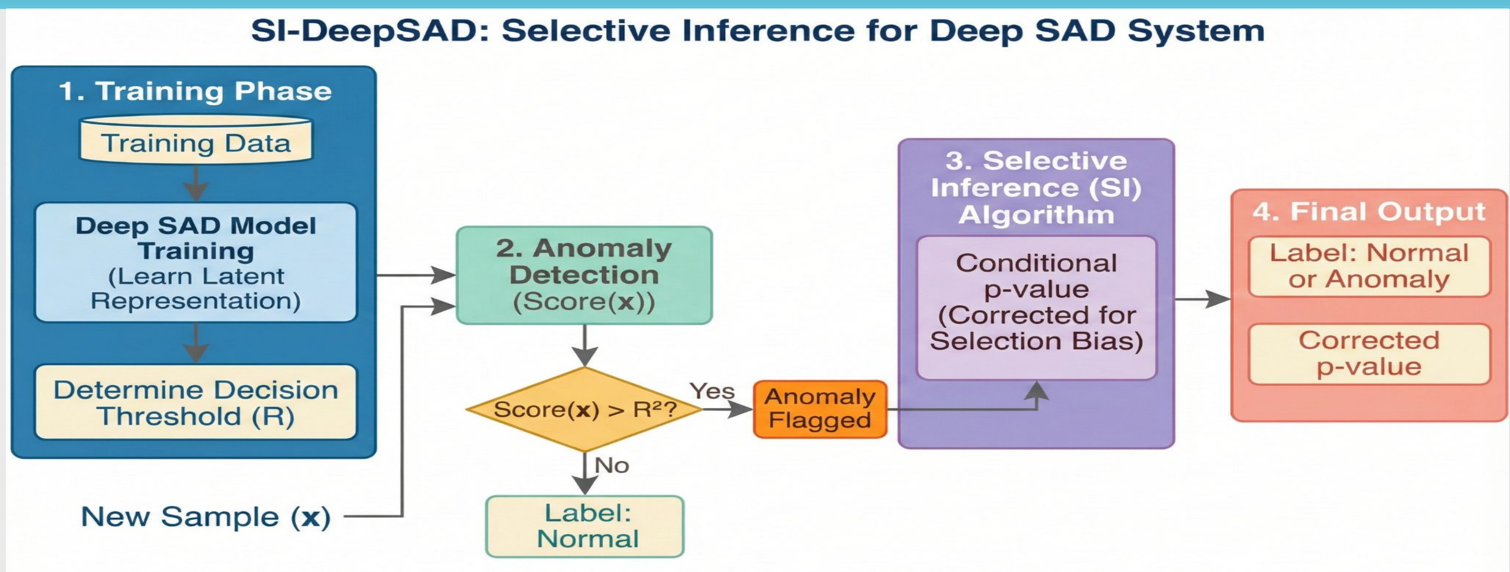
What ?

- **Core Objective:** Develop a statistical inference framework to quantify the reliability of predictions from "black-box" Deep SAD models.
- **Novelty:** Integrate **Selective Inference** to compute **valid p-values**, ensuring rigorous control of the **False Positive Rate** (FPR) at the significance level α.
- **Efficacy:** Demonstrate that the proposed method achieves superior **Statistical Power** compared to traditional correction techniques (e.g., Bonferroni correction) on benchmark datasets.

Why ?

- **Practical Necessity:** In **safety-critical domains** (e.g., healthcare, industry), false alarms (false positives) can lead to severe **financial and psychological consequences.**
- **Research Gap:** Current **SOTA models** like Deep SAD operate as **"black boxes"** providing predictions without **quantitative reliability metrics**, thereby preventing operators from accurately assessing risks.
- **Project Objective:** Integrate **Selective Inference** to **"transparentize"** decisions, providing precise **statistical significance** to **rigorously control** the False Positive Rate.

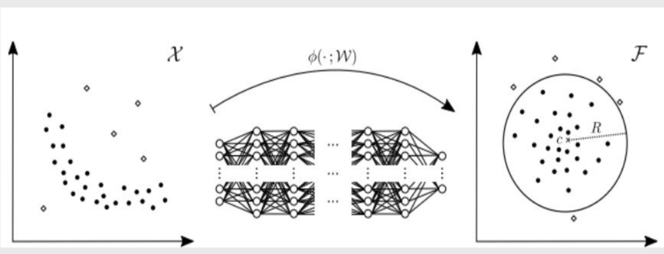
Overview



Description

1. Deep SAD Training

- The Deep SAD model (Ruff et al., 2020) is trained in a semi-supervised manner to learn how to distinguish between normal and anomalous samples with an objective function of minimizing the distance of normal samples to the center c and maximizing the distance of anomalous samples.
- After training is complete, a threshold R is determined as the boundary for detecting anomalies.

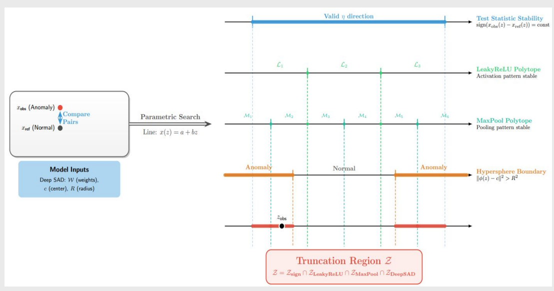


2. Testing New Samples

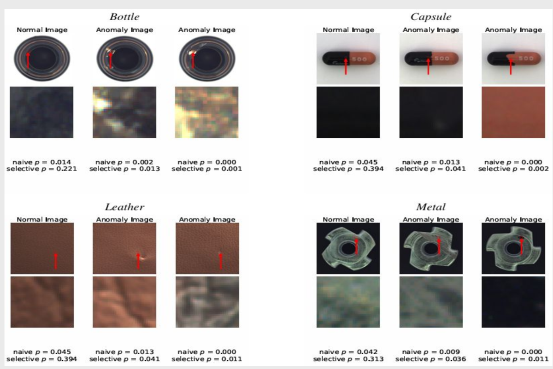
- The model predicts "anomaly" if the score *Score(x) = ||f(x) - c||² > R²*
  - *f(x)*: Representation in the latent space.
  - *R*: Decision radius (calculated from the quantile of the normal set and the unlabeled set in the training set).

3. Selective Inference

- **Technique:** Parametric Selective Inference
- **P-value calculation process:**
  - **Parameterization:** Construct a line $X(z) = a + bz$ passing through the suspected sample (x_{test}).
  - **Constraints Tracking:** Find the set of value intervals of z such that:
    a. **Statistical sign preservation:** Ensure the sign of the Test statistic between x_{test} and x_{ref} does not change so that the test direction vector η is fixed - Solve system of linear inequalities.
    b. **Invariant network structure:** Leaky ReLU and MaxPool neurons keep their activation state - Solve system of linear inequalities.
    c. **The model still predicts as anomaly:** *Score(X(z)) > R²* - solve quadratic inequality.



  - **Statistical calculation:** Calculate the conditional probability (p-value) based on the Truncated Gaussian distribution within the intervals found in step 2.



NII

Cao Lê Công Thành, Đặng Quang Vinh - University of Information Technology, VNU-HCM
Email : 23521437@gm.uit.edu.vn - 23521786@gm.uit.edu.vn