Deep One-Class Classification for Anomaly Detection in Image Datasets

Işıl Altınışık
Faculty of Computer and Informatics
Department of Artificial Intelligence and Data Engineering
Istanbul Technical University
altinisik22@itu.edu.tr

June 25, 2025

Abstract

Anomaly detection is crucial in domains where anomalous samples are rare, sensitive, or expensive to collect. In this project, I propose a deep one-class classification approach based on Deep Support Vector Data Description (Deep SVDD) to detect visual anomalies. The hypothesis is that deep models trained only on normal data can learn a compact latent representation, and that deviations from this representation can signal anomalies. I evaluated many approaches on both the CIFAR-10 and MVTec AD datasets. Deep SVDD, in combination with different backbone networks (ResNet18 and MobileNetV2), was trained and tested on normal-vs-anomaly splits. I achieved AUC scores up to 0.86 on MVTec AD and observed class-dependent performance variation on CIFAR-10. Experiments also involved hyperparameter analysis (epoch, resize, threshold), feature visualization via t-SNE, and evaluation metrics such as F1, Precision, and Recall.

1 Introduction

Anomaly detection is the task of identifying samples that deviate significantly from the norm. In fields like industrial inspection, fraud detection, or medical imaging, anomalous samples are often scarce or unavailable during training, making supervised learning approaches impractical. One-class classification (OCC) addresses this by learning only from normal samples.

In visual domains, this problem becomes more challenging due to high-dimensional input and subtle variations. Traditional methods like One-Class SVM or Isolation Forest struggle with image data. Deep learning models, on the other hand, can learn hierarchical representations and capture complex features from high-resolution inputs.

The goal is to develop a deep OCC system using Deep SVDD. The method maps normal samples into a compact hypersphere in latent space. We explore how well this framework performs across two types of datasets: general-purpose (CIFAR-10) and domain-specific (MVTec AD). Additionally, we analyze how hyperparameters affect the performance and visually interpret feature space behavior via t-SNE plots.

2 Problem Statement and Literature Review

2.1 Problem Statement

Given a dataset with only normal-class training samples and a test set containing unknown anomalies, detect and flag anomalous inputs.

2.2 Hypothesis

A deep neural network trained to minimize the distance of normal samples to a central point in feature space will push anomalies further away from this center, allowing for detection via a threshold.

2.3 Literature Survey

- **OC-SVM (Schölkopf et al.):** Learns a boundary enclosing normal data in kernel space.
- **Isolation Forest (Liu et al.):** Randomly isolates instances to detect outliers.
- **Deep SVDD (Ruff et al., 2018):** Trains a deep network to map inputs into a hypersphere with minimal volume.
- **PatchCore, DRAEM, CutPaste (2020–2022):** Advanced image-level anomaly detection methods using augmentation or reconstruction.
- **Vision Transformers:** Achieve strong results in recent deep OCC works due to global attention.

3 Methodology

3.1 Deep SVDD Framework

Deep SVDD aims to find a compact representation of the normal data in latent space. A neural network $f(\cdot)$ is trained such that all normal samples x_i are mapped close to a center c:

$$L = \frac{1}{N} \sum_{i=1}^{N} \|f(x_i) - c\|^2 \tag{1}$$

During inference, test samples are scored by their distance to c. A threshold determines whether they are normal or anomalous.

3.2 Backbone Networks

We experimented with:

- **ResNet18:** Deeper network with skip connections, known for good feature abstraction.
- **MobileNetV2:** Lightweight and efficient, tests whether a lighter model suffices.

3.3 Evaluation Metrics

We use:

- AUC (Area Under ROC Curve)
- F1-score
- Precision, Recall

4 Datasets

4.1 CIFAR-10

CIFAR-10 contains 10 object classes with 32×32 resolution. We resized to 64×64 or 128×128 for better feature extraction. One class was chosen as normal (e.g., airplane) and the rest as anomalies during testing.

4.2 MVTec AD

The MVTec AD dataset includes 15 categories of industrial objects and textures. Each category has hundreds of normal and defective high-resolution images (e.g., bottle, capsule, zipper). I selected the `bottle` category for experiments.

5 Experiments and Results

5.1 CIFAR-10 Backbone Comparison

- ResNet18 (airplane class): AUC: **0.6102**, F1: 0.6664
- MobileNetV2 (airplane class): AUC: **0.5607**, F1: 0.6579

Observation: Deeper backbone (ResNet18) performs better in distinguishing normal features.

5.2 Effect of Normal Class

- Normal = **automobile**: AUC: 0.4868, F1: 0.6420
- Normal = **bird**: AUC: 0.6565, F1: 0.6807

Observation: Semantically complex classes like "bird" yielded better results. "Automobile" class underperformed likely due to feature overlap.

5.3 Epoch and Resize Experiments

- Epoch = 3, Resize(128×128): AUC = 0.6083
- Epoch = 5, Resize(128×128): AUC = 0.5566
- Epoch = 10, Resize(64×64): AUC = 0.4855

Observation: Increasing epochs did not improve performance and overfitting occurred especially on low-resolution images.

5.4 Thresholding Strategies

We compared:

- `threshold` = `np.median`(y_scores)
- `threshold` = `np.percentile`($y\_scores, 80$)

Result: The percentile-based threshold provided better precision-recall balance and fewer false positives.

5.5 t-SNE Analysis

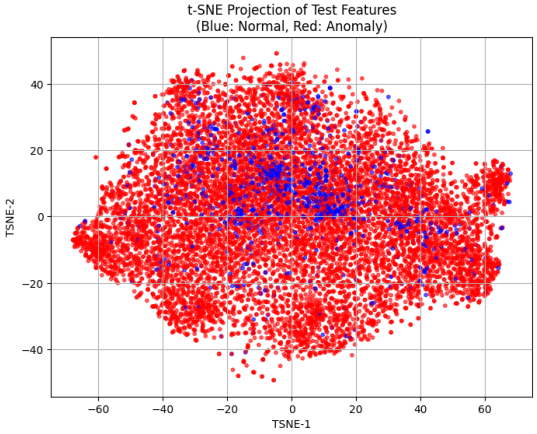

Figure 1: t-SNE projection for 'airplane' class features

Observation: The blue points (normal samples) are scattered among the red (anomalous) points. There is no clear boundary, indicating weak feature separation. This explains lower AUC and higher false positive rates.

5.6 MVTec AD Results

Category: bottle

- AUC: 0.8595
- F1: 0.7692
- Precision: 0.9756
- Recall: 0.6349

Observation: On real-world industrial data, model performance improved significantly due to higher visual consistency and structured anomalies.

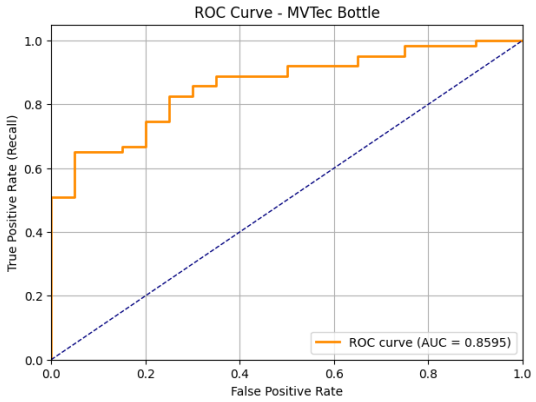

Figure 2: ROC Curve for MVTec Bottle Category

6 Discussion

Performance on CIFAR-10 varied greatly by class. Deep SVDD failed to clearly cluster some categories due to low resolution and high intra-class variance. ResNet18 consistently outperformed MobileNetV2. Resizing images to 128×128 helped moderately, but increasing training epochs had diminishing returns. Threshold tuning significantly affected detection quality.

In contrast, MVTec AD offered clearer visual anomalies, enabling better latent separation and superior performance. This suggests deep OCC models benefit more from structured datasets with high-quality input.

7 Conclusion

This study explored the effectiveness of Deep SVDD on two visual anomaly detection datasets. While CIFAR-10 results were modest, MVTec AD experiments showed strong anomaly separation. t-SNE plots confirmed latent feature overlap on CIFAR, explaining model limitations. Future work could involve training on larger datasets, using attention-based models (ViTs), or applying unsupervised contrastive pretraining to boost latent separability.

Repository

All source code, trained models, and a three-minute project presentation are available at: `https://github.com/YOUR-REPO-HERE`