Anomaly Detection via Deep One-Class Classification Using Vision Transformers

Berker Eriş, Enes Yemenli, Emircan Sarı
Istanbul Technical University
Emails: {berkereris, enesyemenli, emircan.sari}@itu.edu.tr

***Abstract*—In many real-world scenarios, anomalous examples are rare or unavailable during training. To address this, one-class classification (OCC) approaches are employed to model only the "normal" class and detect deviations as anomalies. In this project, we define the *ship* class from the CIFAR-10 dataset as normal and consider the remaining classes as anomalies. We propose a solution that uses a Vision Transformer (ViT) model for feature extraction and Deep Support Vector Data Description (Deep SVDD) for anomaly detection. Deep SVDD minimizes the volume of a hypersphere that encloses normal data representations in the latent space. This method is evaluated against classical techniques: One-Class SVM, Isolation Forest, and PCA. Our results show that ViT-based Deep SVDD achieves superior performance in detecting anomalies.**

I. Problem Statement – Hypothesis – Literature Survey

A. Problem Statement

In settings such as medical diagnosis or manufacturing quality control, training data often contains only normal examples. This leads to a one-class classification (OCC) problem, where the model must learn the representation of the normal class and detect significant deviations during inference as anomalies.

B. Hypothesis

We hypothesize that Vision Transformers (ViTs), with their ability to learn rich and high-level semantic representations, can effectively encode "normal" data in a compact latent space. When combined with Deep SVDD, the model can construct a minimal-volume hypersphere around these representations, enabling robust detection of anomalous inputs.

C. Literature Survey

- **Deep SVDD** (Yin et al., 2022): Learns a hypersphere in latent space minimizing distance from the center.
- **Vision Transformers** (Vaswani et al., 2017): Self-attention-based models capturing global dependencies.
- **One-Class SVM** (Schölkopf et al., 2001): Separates normal data from the origin in a kernel-transformed space.
- **PCA** (Bro et al., 2014): Projects data to low-dimensional space, detects anomalies via reconstruction error.
- **Isolation Forest** (Liu et al., 2008): Uses random partitioning to isolate outliers.

II. Method – Data – Results – Discussion and Conclusions

A. Methodology

We used a pretrained ViT-Base model to extract latent representations from CIFAR-10 *ship* class images. Anomaly detection is performed using:

- **Deep SVDD:** Trains a hypersphere in feature space to enclose normal points.
- **One-Class SVM, Isolation Forest, PCA:** Applied to ViT features as baselines.

B. Data

- Dataset: CIFAR-10 (https://www.cs.toronto.edu/~kriz/cifar.html)
- Normal class: *Ship*
- Anomalous classes: All others (airplane, cat, truck, etc.)
- Preprocessing: Normalization and resizing according to ViT input.

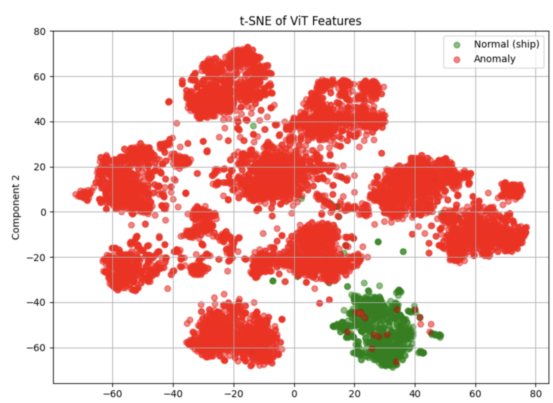
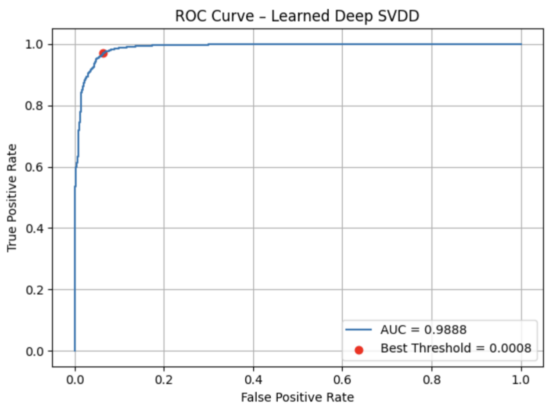
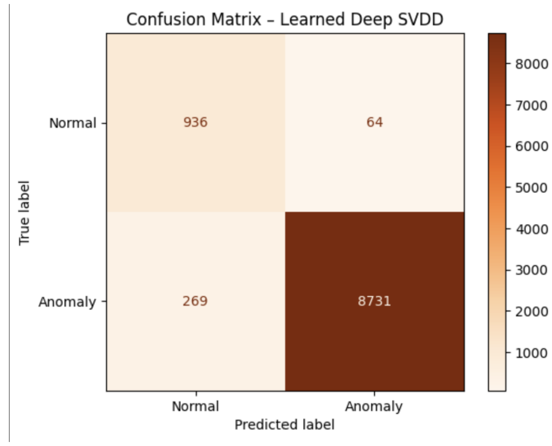


Fig. 1

C. Results



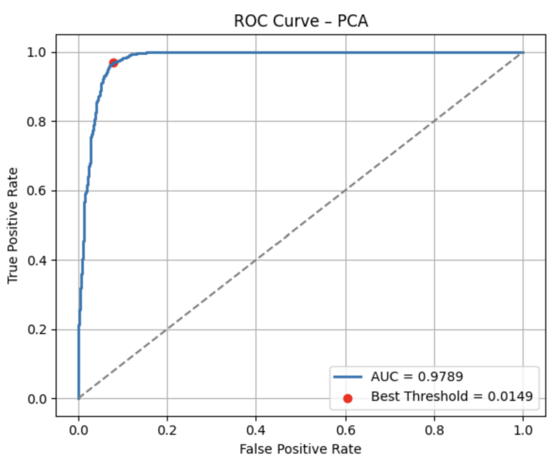
(a) ROC Curve (AUC = 0.9888)



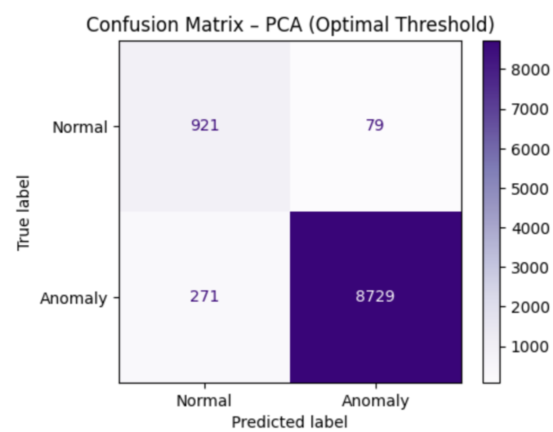
(b) Confusion Matrix

Fig. 2: Deep SVDD Performance

1) Deep SVDD: Deep SVDD achieved the highest AUC among all methods, correctly identifying most normal and anomalous samples with minimal false positives.



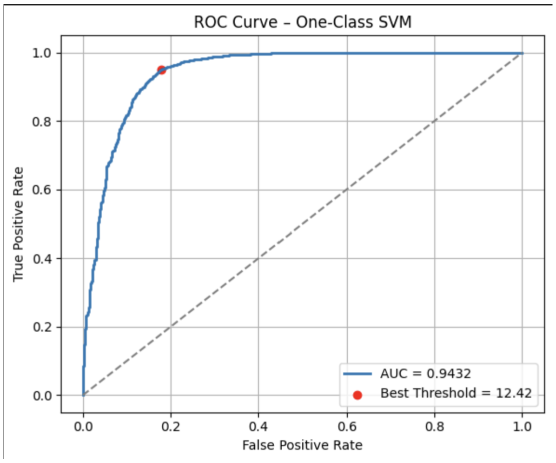
(a) ROC Curve (AUC = 0.9789)



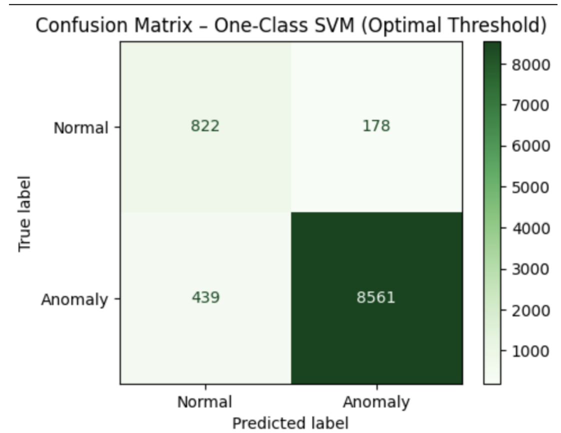
(b) Confusion Matrix

Fig. 3: PCA Performance

2) PCA: PCA also performed well, but had slightly higher false negatives compared to Deep SVDD.



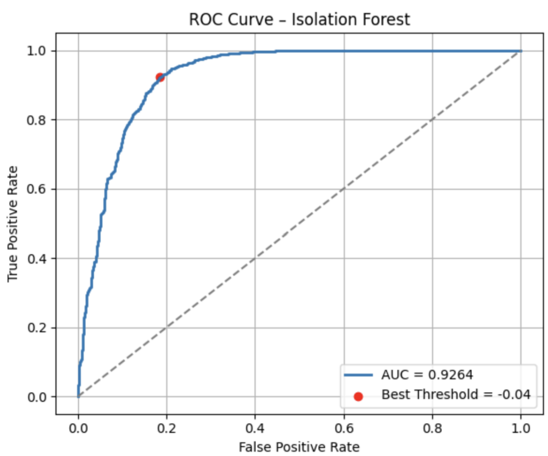
(a) ROC Curve (AUC = 0.9432)



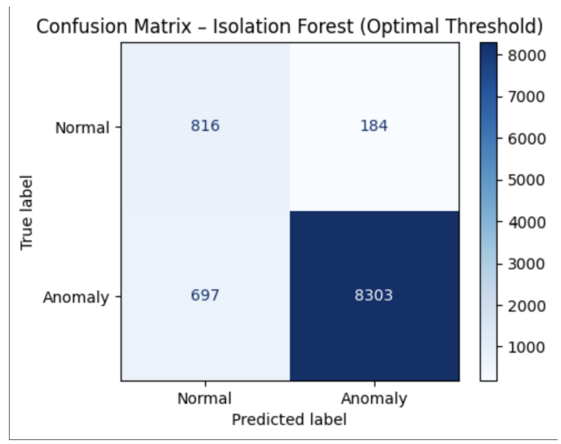
(b) Confusion Matrix

Fig. 4: One-Class SVM Performance

3) One-Class SVM: *One-Class SVM had a balanced performance, though it misclassified more anomalies than PCA and SVDD.*



(a) ROC Curve (AUC = 0.9264)



(b) Confusion Matrix

Fig. 5: Isolation Forest Performance

4) Isolation Forest: *Isolation Forest produced the lowest AUC and had the most false positives among all methods.*

D. Discussion and Conclusion

Comparison Table:

TABLE I: AUC and F1 Scores for All Methods

Method	AUC	F1 Score
Learned Deep SVDD	0.9888	0.9813
One-Class SVM	0.8866	0.9652
Isolation Forest	0.8693	0.9496
PCA	0.9454	0.9803
Basic Deep SVDD	0.5272	0.1050

Deep SVDD outperforms classical OCC methods in AUC and classification accuracy, validating the hypothesis that ViT-based features combined with hypersphere optimization enable robust anomaly detection. Classical methods like PCA and SVM perform reasonably well, but lack the feature richness offered by transformers. These results confirm that learning compact and expressive latent representations is key in one-class scenarios.

Code and Presentation Repository

The full source code and a presentation of this project can be accessed at:

https://github.com/itu-itis23-erisb22/BLG454PROJECT