Graph-Attention-Regularized Deep Support Vector Data Description for Semi-Supervised Anomaly Detection: A Case Study in Automotive Quality Control

Taha J. Alhindi *

Posted Date: 22 October 2025

doi: 10.20944/preprints202510.1674.v1

Keywords: acoustic anomaly detection; attention-weighted graphs; automotive quality control; GAR-DSVDD; semi-supervised anomaly detection; SVDD; windshield-wiper fault detection

 doi:10.20944/preprints202510.1674.v1

Article

Graph-Attention-Regularized Deep Support Vector Data Description for Semi-Supervised Anomaly Detection: A Case Study in Automotive Quality Control

Taha J. Alhindi

Department of Industrial Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia; talhindi@kau.edu.sa

Abstract

This paper presents a graph-attention-regularized deep anomaly detection framework for semi-supervised settings where normal labeled data is scarce and unlabeled data is abundant and contaminated (i.e., includes normal and anomaly points). Unlike conventional one-class and graph-based semi-supervised anomaly detection approaches, which either ignore unlabeled structure or do not utilize the structure effectively through data density breaks, the proposed method constructs an attention-weighted latent k-NN graph to capture unlabeled geometry while reducing the influence of contaminated neighbors. We formulate a new deep support vector data description variant, graph-attention-regularized deep support vector data description, that embeds this attention-weighted graph directly into the one-class objective incorporated with a center-pull regularization for unlabeled data to avoid decision boundary over smoothness, thereby effectively leveraging unlabeled data geometry and information. Experiments on simulated and an industrial windshield-wiper acoustics datasets validate the effectiveness of the proposed approach under scarce-label conditions relative to existing deep and shallow methods.

Keywords: acoustic anomaly detection; attention-weighted graphs; automotive quality control; GAR-DSVDD; semi-supervised anomaly detection; SVDD; windshield-wiper fault detection

MSC: 68T05; 68T07; 62H30

1. Introduction

Anomaly detection is a critical issue in machine learning, with applications in manufacturing, industrial process control, medical diagnosis, cybersecurity, and disaster management [1–5]. The primary objective is to learn a representation of normal operating conditions and identify deviations as anomalies. Unlike standard supervised classification, anomaly detection is characterized by severe class imbalance, where anomalies are rare, diverse, and often scarcely labeled, while normal samples are abundant. However, despite this abundance, labeling a sample as truly normal is often costly and challenging. In many operational settings, normal labeling requires expert review against process specifications, additional quality-control checks (sometimes destructive or time-consuming), or prolonged observation to rule out hidden faults. These steps consume engineer time and test resources and can cause production downtime, making exhaustive normal labeling economically impractical at scale. For instance, in automotive quality management, current practice for windshield-wiper noise relies on expert assessment: after controlled recordings, mel frequency cepstral coefficients/spectrogram features are computed, and an expert labels the noise types [6].

 doi:10.20944/preprints202510.1674.v1

Early efforts in one-class classification introduced support vector data description (SVDD) and the one-class support vector machine (OC-SVM) as core frameworks. SVDD learns a minimum-radius hypersphere that encloses the normal training data, predicting observations outside the decision boundary as anomalies [7]. The OC-SVM learns a maximum-margin separator from the origin in feature space to capture the support of the normal distribution [8]. While these models produce compact, interpretable decision boundaries, they generally require a clean set of normal samples and do not utilize unlabeled data, thereby limiting their effectiveness when normal labeling is costly and unlabeled observations are abundant.

Graph-based semi-supervised learning offers a systematic approach to utilize unlabeled data by representing adjacency and manifold structure. The basic concepts of spectral graphs and Laplacian regularization define the cluster assumption through the concept of smoothness across neighborhoods [9,10]. Recent studies and applications in graph anomaly detection indicate extensive utilization across nodes, edges, subgraphs, and whole graphs [11–13]. SVDD has been extended using semi-supervised methods that utilize the structure of unlabeled data while employing a small number of labeled normal observations to guide the boundary. In order to leverage large unlabeled data without hard pseudo-labels, a common practice is to enhance the SVDD objective with manifold/graph regularization so that points connected on a similarity graph remain close in the learned representation [9,10]. One SVDD variant, graph-based semi-supervised SVDD (S3SVDD), leverages global/local geometric structure with limited labels by utilizing a k-nearest neighbor (k-NN) spectral graph and a Laplacian smoothness term in the SVDD objective to exploit unlabeled data [14]. Similarly, semi-supervised subclass SVDD proposes a new formulation that utilizes global/local geometric structure with limited labels [15]. Separately, a manifold-regularized SVDD for noisy label detection was introduced, showing that incorporating a graph Laplacian into the SVDD objective improves robustness to label noise while still leveraging abundant unlabeled data [16]. In addition, a semi-supervised convolutional neural network (CNN) with SVDD has demonstrated practical viability in industrial monitoring [17]. While traditional graph-based SVDD variants have shown potential, these methods typically rely on fixed, pseudo-labeled neighbor graphs (e.g., k-NN with a chosen metric and k) and uniform edge weights, which can over-connect across data density breaks, propagate contamination from suspect (anomalous) neighbors, and require careful tuning [9,14,18–20].

The introduction of deep learning has led to significant advances in deep anomaly detection. Deep SVDD (DeepSVDD) extended the SVDD framework by combining it with a deep encoder that maps data into a latent representation enclosed by a hypersphere [21]. Autoencoders and their variants have been widely adopted to learn low-dimensional embeddings and reconstruction-based anomaly scores in industrial and time-series contexts [22,23]. Additionally, self-supervised anomaly detection has emerged as a promising direction by utilizing contrastive learning and related pretext objectives that enhance representation quality under limited labels [24,25]. However, common deep anomaly detection methods often use objectives that are indirectly aligned with the decision boundary (e.g., reconstruction error), leverage unlabeled data using proxy tasks rather than task-aware constraints in the decision boundary learner, and offer limited interpretability of decisions [26–31].

Motivated by the aforementioned challenges, this paper presents Graph-Attention-Regularized Deep SVDD (GAR-DSVDD). Specifically, we develop a semi-supervised Deep SVDD that trains a deep encoder and a one-class hypersphere end-to-end, while exploiting unlabeled structures with a center-pull regularizer on a learned graph. To build this graph, we introduce an attention-weighted latent k-NN mechanism that assigns label and score-aware importance to neighbors, emphasizing prototypical normal observations and down-weighting suspect or anomalous samples, thereby mitigating label scarcity and limiting error propagation from contaminated edges. Unlike iterative pseudo-labeling methods, our training uses unlabeled data directly through the center-pull and graph smoothness regularizers, which stabilize optimization and reduce computation. For safety-critical deployment, per-instance attention weights provide transparent explanations over the top-k

 doi:10.20944/preprints202510.1674.v1

neighbors. Comprehensive experiments on a simulated dataset and an industrial windshield-wiper acoustics case study show that GAR-DSVDD achieves superior detection of subtle anomalies compared with classical and deep learning benchmarks, while reducing labeling effort and operator subjectivity.

The remainder of the paper is structured as follows. Section 2 reviews preliminaries and notation for semi-supervised anomaly detection. Section 3 presents the proposed GAR-DSVDD method. Section 4 describes the experimental setup and reports results on simulated datasets and the windshield-wiper acoustics case study, along with sensitivity analyses. Section 5 discusses findings, conclusions, and potential future work directions.

2. Preliminaries

2.1. Data and Basic Notation

Given a training dataset $\mathcal{X} = \{x_i \mid x_i \in \mathbb{R}^p, i = 1, \dots, n\}$ that decomposes into two disjoint subsets: a small set of labeled normal samples $\mathcal{X}_\ell^+$ and a large unlabeled set $\mathcal{X}_u$, *with*

$$\mathcal{X} = \mathcal{X}_\ell^+ \cup \mathcal{X}_u, \quad \mathcal{X}_\ell^+ \cap \mathcal{X}_u = \emptyset.$$

Let $L \subset \{1, \dots, n\}$ be the index set of labeled normal observations and $U = \{1, \dots, n\} \setminus L$ the unlabeled indices, with $|L| = \ell$, $|U| = u$, and label rate $\rho = \ell/n$. The unlabeled pool is assumed predominantly normal and may contain a small contamination of anomalies, upper-bounded by $\varepsilon \in [0,1)$. Features are standardized by coordinate (z-scoring). A learnable encoder $\phi_\theta: \mathbb{R}^p \to \mathbb{R}^s$ maps each input to a latent representation, where θ denotes the encoder's trainable parameters; we write $z_i = \phi_\theta(x_i)$ and collect latents as $\mathcal{Z} = \{z_i\}_{i=1}^n$. We stack them row-wise in $Z = [z_1, \dots, z_n]^\top \in \mathbb{R}^{n \times s}$. We use $\|\cdot\|_2$ for the Euclidean norm and $\langle \cdot, \cdot \rangle$ for the standard inner product in $\mathbb{R}^s$. This setting follows the deep one-class paradigm: the normal class is to be enclosed within a compact latent region while avoiding pseudo-labels for the unlabeled pool. Instead, $\mathcal{X}_u$ informs the latent geometry via a graph defined over z_i (details in §3).

Prior work motivating this formulation includes Deep SVDD–style one-class modeling and graph-based semi-supervised learning that leverages unlabeled geometry [9,10,21].

2.2. Latent One-Class Score

We summarize the normal set by a latent center $\boldsymbol{c} \in \mathbb{R}^s$ and a soft margin (squared radius) used for score calibration

$$m = \mathrm{softplus}(\eta), \quad \eta \in \mathbb{R}.$$

Given latents $\mathcal{Z} = \{z_i\}_{i=1}^n$ with $z_i = \phi_\theta(x_i)$, the per-sample anomaly score is

$$f_i = \| z_i - \boldsymbol{c} \|_2^2 - m. \tag{1}$$

A sample is more anomalous as f_i increases; the decision boundary is $f = 0$. Eq. (1) gives the deep "hypersphere" view of SVDD in latent space, where $\boldsymbol{c}$ summarizes the normal set and ϕ_θ maps inputs to latents. The offset $m > 0$ specifies the boundary tolerance (squared radius) via a softplus parameterization, ensuring positivity and smooth gradients near the boundary. The center $\boldsymbol{c}$ is typically initialized as the mean of $\{z_i : x_i \in \mathcal{X}_\ell^+\}$ to stabilize early training. The score is scale-aware: increasing m relaxes the boundary, whereas decreasing m tightens the enclosure, preserving the boundary-level interpretation (distance to $\boldsymbol{c}$ versus tolerance m). During training, m is treated as a calibration constant, so choosing a train-quantile threshold on f is equivalent to thresholding the same quantile of $d^2 = \| z - \boldsymbol{c} \|_2^2$.

 doi:10.20944/preprints202510.1674.v1

3. Graph-Attention-Regularized Deep SVDD

We now detail GAR-DSVDD. The objective has three parts: (i) a one-class enclosure on labeled normal observations, (ii) an unlabeled neighbor-smoothness on squared distances over an attention-weighted directed, row-normalized k-NN graph, and (iii) standard parameter regularization.

3.1. Latent Attention-Weighted k-NN Graph

We construct a directed row-normalized graph $G = (\mathcal{V}, \mathcal{E})$ over the latent set $\{z_i\}_{i=1}^n$, where each latent is produced by the encoder $z_i = \phi_\theta(x_i)$, and we stack them row-wise as $Z = [z_1, \dots, z_n]^\top \in \mathbb{R}^{n \times s}$. For each node i, define the directed neighborhood $\mathcal{N}_k(i)$ as the indices of the k-NN of z_i under a base metric (Euclidean by default). Each candidate edge (i, j) carries a base affinity $\kappa_{\text{base}}(z_i, z_j) \in [0,1]$, chosen as either a Gaussian kernel or a constant:

$$\kappa_{\text{base}}(z_i, z_j) = \exp\left(-\| z_i - z_j \|_2^2 / \sigma^2\right) \quad \text{or} \quad \kappa_{\text{base}}(z_i, z_j) \equiv 1.$$

To emphasize true normal observations and suppress suspect connections (possible anomalies), we compute score-aware attention on each edge. With H heads, the per-head logits are

$$e_{ij}^{(h)} = \frac{q_h(z_i)^\top v_h(z_j)}{\sqrt{s_{att}}} - \gamma\big(\max(0, f_i) + \max(0, f_j)\big), \qquad h = 1, \dots, H,$$

where $\boldsymbol{q}_h(\cdot), \boldsymbol{v}_h(\cdot): \mathbb{R}^s \to \mathbb{R}^{s_{\text{att}}}$ are linear projections (shared across pairs for head h), s_{att} is the attention width, $\gamma \geq 0$ is a contamination reducing coefficient, and $f_i = \| z_i - \boldsymbol{c} \|_2^2 - m$ is the current one-class score (Eq. (1)). We normalize within $\mathcal{N}_k(i)$ using a temperature τ_a-scaled softmax:

$$\alpha_{ij}^{(h)} = \frac{\exp\left(e_{ij}^{(h)} / \tau_a\right)}{\sum_{u \in \mathcal{N}_k(i)} \exp\left(e_{iu}^{(h)} / \tau_a\right)}, \qquad j \in \mathcal{N}_k(i),\ \tau_a > 0$$

and aggregate heads by averaging

$$\tilde{a}_{ij} = \frac{1}{H} \sum_{h=1}^H \alpha_{ij}^{(h)}.$$

We integrate attention with the base affinity using $\mu \in [0,1]$ and then row-normalize the outgoing weights:

$$\widetilde{w}_{ij} = (1 - \mu)\, \kappa_{\text{base}}(z_i, z_j) + \mu \tilde{a}_{ij} \quad (j \in \mathcal{N}_k(i)), \quad \widehat{w}_{ij} = \frac{\widetilde{w}_{ij}}{\sum_{u \in \mathcal{N}_k(i)} \widetilde{w}_{iu}}.$$

Edges outside $\mathcal{N}_k(i)$ have weight 0 ($\widetilde{w}_{ij} = 0$ if $j \notin \mathcal{N}_k(i)$). Row-normalization preserves scale across heterogeneous densities (outgoing weights sum to 1) and reduces error propagation by down-weighting high-score (potentially anomalous) edges during training. We compute attention weights without back-propagation and keep them fixed between graph refreshes. Prior work motivating the attention graphs to interpret unlabeled data geometry includes [9,10,32].

To keep training tractable, we do not materialize a full graph at every step. Instead, we rebuild a latent-space k-NN index periodically and, for each mini-batch, form edges as the union of (i) cached neighbors of the batch points and (ii) within-batch k-NN edges—preserving local geometry for the smoother while keeping computation practical.

3.2. Unlabeled Geometry via Neighbor-Smoothness on Squared Distances

Let

$$d_i^2 = \| z_i - \boldsymbol{c} \|_2^2, \qquad \boldsymbol{d} = [d_1^2, \dots, d_n^2]^\top.$$

Unlabeled samples help create a smooth decision boundary by discouraging sharp variations of d_i^2 along high weight edges of the attention-weighted graph (described in §3.1):

$$\mathcal{L}_{\text{graph}} = \frac{1}{n} \sum_{i=1}^n \sum_{j \in \mathcal{N}_k(i)} \widehat{w}_{ij} \left(d_i^2 - d_j^2\right)^2. \tag{3}$$

 doi:10.20944/preprints202510.1674.v1

Penalizing differences of d_i^2 (rather than feature vectors) aligns the boundary with high-density regions without collapsing representations and preserves the simple one-class test rule on f.

The derivatives of $\mathcal{L}_{\text{graph}}$ are

$$\frac{\partial \mathcal{L}_{\text{graph}}}{\partial d_i^2} = \frac{2}{n}\left(\sum_{j \in \mathcal{N}_k(i)} \widehat{w}_{ij}\left(d_i^2 - d_j^2\right) + \sum_{u:\, i \in \mathcal{N}_k(u)} \widehat{w}_{ui}(d_i^2 - d_u^2) \right),$$

and by the chain rule

$$\frac{\partial d_i^2}{\partial z_i} = 2(z_i - \boldsymbol{c}), \qquad \frac{\partial d_i^2}{\partial \boldsymbol{c}} = -2(z_i - \boldsymbol{c}), \quad \frac{\partial z_i}{\partial \theta} = \nabla_\theta \phi_\theta(x_i),$$

so

$$\frac{\partial \mathcal{L}_{\text{graph}}}{\partial z_i} = \frac{4}{n}(z_i - \boldsymbol{c})\left(\sum_{j \in \mathcal{N}_k(i)} \widehat{w}_{ij}\left(d_i^2 - d_j^2\right) + \sum_{u:\, i \in \mathcal{N}_k(u)} \widehat{w}_{ui}(d_i^2 - d_u^2) \right),$$

$$\frac{\partial \mathcal{L}_{\text{graph}}}{\partial \boldsymbol{c}} = -\frac{4}{n}\sum_{i=1}^{n}(z_i - \boldsymbol{c})\left(\sum_{j \in \mathcal{N}_k(i)} \widehat{w}_{ij}\left(d_i^2 - d_j^2\right) + \sum_{u:\, i \in \mathcal{N}_k(u)} \widehat{w}_{ui}(d_i^2 - d_u^2) \right),$$

and

$$\frac{\partial \mathcal{L}_{\text{graph}}}{\partial \theta} = \sum_{i=1}^{n} \left(\nabla_\theta \phi_\theta(x_i)\right)^{\mathrm{T}} \frac{\partial \mathcal{L}_{\text{graph}}}{\partial z_i}.$$

During training, we rebuild the graph periodically and hold $\widehat{w}_{ij}$ fixed within each interval, so we do not back-propagate through the weights in this term.

3.3. Labeled-Normal Enclosure

Labeled normal observations $\mathcal{X}_\ell^+$ anchor the hypersphere in latent space through a distance-based objective:

$$\mathcal{L}_{\text{enc}} = \frac{1}{\left|\mathcal{X}_\ell^+\right|} \sum_{x_i \in \mathcal{X}_\ell^+} \| z_i - \boldsymbol{c} \|_2^2 + \beta \| \boldsymbol{c} \|_2^2, \tag{4}$$

where $z_i = \phi_\theta(x_i)$, $\boldsymbol{c} \in \mathbb{R}^s$ is a latent center summarizing the normal set. and $\beta \geq 0$ weakly centers the hypersphere. This formulation directly penalizes the squared distance of labeled normal observations from $\boldsymbol{c}$ yielding stable gradients and a simple coupling to the encoder.

The gradients of $\mathcal{L}_{\text{enc}}$ are

$$\frac{\partial \mathcal{L}_{\text{enc}}}{\partial z_i} = \frac{2}{\left|\mathcal{X}_\ell^+\right|}(z_i - \boldsymbol{c}) \quad \text{for } x_i \in \mathcal{X}_\ell^+, \quad \frac{\partial \mathcal{L}_{\text{enc}}}{\partial z_i} = 0 \; otherwise,$$

$$\frac{\partial \mathcal{L}_{\text{enc}}}{\partial \boldsymbol{c}} = -\frac{2}{\left|\mathcal{X}_\ell^+\right|} \sum_{x_i \in \mathcal{X}_\ell^+} (z_i - \boldsymbol{c}) + 2\beta \boldsymbol{c}.$$

Since $z_i = \phi_\theta(x_i)$, by the chain rule

$$\frac{\partial \mathcal{L}_{\text{enc}}}{\partial \theta} = \frac{2}{\left|\mathcal{X}_\ell^+\right|} \sum_{x_i \in \mathcal{X}_\ell^+} \left(\nabla_\theta \phi_\theta(x_i)\right)^{\mathrm{T}} (z_i - \boldsymbol{c}).$$

3.4. Unlabeled Center-Pull

To stabilize training when the labeled normal set $\mathcal{X}_\ell^+$ is small, we add a label-free pull of the unlabeled latents toward the center $\boldsymbol{c}$. This acts on raw squared distances $d_i^2 = \| z_i - \boldsymbol{c} \|_2^2$ and complements the neighbor smoothness. The graph term equalizes differences of d_i^2 along edges and is largely insensitive to the global level of d_i^2; with few labelled observations this can cause scale drift and a poorly calibrated score $f = d^2 - m$. Moreover, early graphs may include contaminated edges

 doi:10.20944/preprints202510.1674.v1

(unlabeled anomalies can appear inside or near the hypersphere), so smoothing alone can propagate their influence. The center-pull softly anchors the mean of d^2 for the normal unlabeled pool, reducing drift and reducing contamination while the graph term aligns local variations.

We define

$$\mathcal{L}_{cp} = \frac{1}{|\mathcal{X}_u|} \sum_{x_i \in \mathcal{X}_u} \| z_i - \boldsymbol{c} \|_2^2. \tag{5}$$

This imposes a global constraint by shrinking the mean of d_i^2 over $x_i \in \mathcal{X}_u$ toward $\boldsymbol{c}$. Unlike $\mathcal{L}_{\text{graph}}$—which is local and relative—$\mathcal{L}_{cp}$ fixes the overall scale and prevents drift when labels are scarce, improving calibration of $f = d^2 - m$.

The gradients are

$$\frac{\partial \mathcal{L}_{cp}}{\partial z_i} = \frac{2}{|\mathcal{X}_u|}(z_i - \boldsymbol{c}) \quad \text{for } x_i \in \mathcal{X}_u, \quad \frac{\partial \mathcal{L}_{cp}}{\partial z_i} = 0 \; otherwise,$$

$$\frac{\partial \mathcal{L}_{cp}}{\partial \boldsymbol{c}} = -\frac{2}{|\mathcal{X}_u|} \sum_{x_i \in \mathcal{X}_u} (z_i - \boldsymbol{c}).$$

Since $z_i = \phi_\theta(x_i)$, by the chain rule

$$\frac{\partial \mathcal{L}_{cp}}{\partial \theta} = \frac{2}{|\mathcal{X}_u|} \sum_{x_i \in \mathcal{X}_u} \left(\nabla_\theta \phi_\theta(x_i)\right)^{\mathrm{T}} (z_i - \boldsymbol{c}).$$

3.5. GAR-DSVDD Objective Function

The total loss combines unlabeled geometry, labeled-normal enclosure, and unlabeled center pull:

$$\min_{\theta, \boldsymbol{c}} \mathcal{L}(\theta, \boldsymbol{c}) = \lambda_u \underbrace{\mathcal{L}_{\text{graph}}}_{\text{unlabeled geometry}} + \underbrace{\mathcal{L}_{\text{enc}}}_{\text{labeled normals}} + \lambda_{\text{cp}} \underbrace{\mathcal{L}_{\text{cp}}}_{\text{unlabeled center-pull}}, \quad \lambda_u, \lambda_{\text{cp}} \geq 0 \tag{6}$$

with $\mathcal{L}_{\text{graph}}$, $\mathcal{L}_{\text{enc}}$, and $\mathcal{L}_{\text{cp}}$ as in Eqs. (3), (4), and (5), respectively, the scalars $\lambda_u, \lambda_{\text{cp}}$ control the strengths of the graph regularizer and unlabeled center pull.

Together, $\mathcal{L}_{\text{enc}}$, $\mathcal{L}_{\text{graph}}$, and $\mathcal{L}_{\text{cp}}$ play complementary roles: $\mathcal{L}_{\text{enc}}$ anchors the center $\boldsymbol{c}$ using trusted labeled normals, preventing center drift; $\mathcal{L}_{\text{graph}}$ enforces local consistency of squared distances on the attention-weighted, row-normalized k-NN graph, transferring structure from the unlabeled pool without collapsing features; and $\mathcal{L}_{\text{cp}}$ imposes a global moment constraint on unlabeled data, fixing the overall scale of d^2 and improving calibration of the score $f = d^2 - m$.

We train $(\theta, \boldsymbol{c})$ jointly using the AdamW optimizer. The attention-weighted k-NN graph is recomputed every T epochs and held fixed within those intervals. No gradients flow through the attention computation or through m. At deployment, a testing observation x_{test} can be labeled as an anomaly if its f_{test} score exceeds a train-quantile threshold τ_{thr} otherwise it will be considered as normal.

In summary, GAR-DSVDD combines a deep one-class boundary with graph-based semi-supervision in a way that is both robust and label-efficient. By regularizing squared distance (scores)—not features—over an attention-weighted latent k-NN graph, it aligns the decision boundary with the data manifold without collapsing representations and preserves the simple test-time rule $f(x) > \tau_{\text{thr}}$. Score-aware attention selectively down-weights suspicious neighbors, curbing over-smoothing from contaminated edges and density breaks. In addition to this, local, edge-wise smoothing of d^2, an unlabeled center-pull imposes a global moment constraint that stabilizes the overall scale of the distance field, improving score calibration when labeled normal observations are scarce. Because unlabeled data enter only through these geometric regularizers (no pseudo-labels), a small labeled set is enough to anchor the hypersphere while the unlabeled pool shapes the boundary.

 doi:10.20944/preprints202510.1674.v1

Thus, the proposed method effectively utilizes information from labeled and unlabeled observations for enhanced anomaly detection in semi-supervised settings.

Algorithm: GAR-DSVDD (training)

Inputs: $\mathcal{X}_\ell^+$, $\mathcal{X}_u$; ϕ_θ; graph params $(k, \sigma, \mu, H, \tau_a)$; weights $(\lambda_u, \lambda_{cp}, \beta)$; refresh period T; total epochs T_{epochs}.

Outputs: $\theta, \boldsymbol{c}$, decision threshold τ_{thr}.

1. Initiation: Compute $z_i = \phi_\theta(x_i)$ for $x_i \in \mathcal{X}_\ell^+$; set $\boldsymbol{c} \leftarrow \frac{1}{|\mathcal{X}_\ell^+|}\sum_{x_i \in \mathcal{X}_\ell^+} z_i$.
2. Build graph ($t = 0$): Compute latents for all data; form k-NN; compute multi-head attentions $\alpha_{ij}^{(h)}$ with temperature τ_a; average to $\tilde{a}_{ij}$; integrate via μ; row-normalize to obtain $\hat{w}_{ij}$.
3. For $t = 1, \ldots, T_{\text{epochs}}$:
   a) Forward on mini-batch $\mathcal{B}$: $z_i = \phi_\theta(x_i)$, $d_i^2 = \| z_i - \boldsymbol{c} \|_2^2$, and for attention/monitoring $f_i = \| z_i - \boldsymbol{c} \|_2^2 - m$.
   b) Loss: $\mathcal{L} = \lambda_u \mathcal{L}_{\text{graph}} + \mathcal{L}_{\text{enc}} + \lambda_{\text{cp}} \mathcal{L}_{\text{cp}}$.
   c) Backpropagation + AdamW updates for $(\theta, \boldsymbol{c})$.
   d) Graph refresh: if $t \bmod T = 0$: recompute latents; rebuild k-NN, attentions and row normalize $\hat{w}_{ij}$

Threshold selection: $\tau_{thr} \leftarrow$ Quantile of $(f_i : x_i \in \mathcal{X}_\ell^+)$ or via validation.

Inference (testing)

Given x_{new}:

1. Compute latent $z_{new} = \phi_\theta(x_{new})$.
2. Score $f(x_{new}) = \| z_{new} - \boldsymbol{c} \|_2^2 - m$.
3. Predict anomaly if $f(x_{new}) > \tau_{thr}$ otherwise normal.

(Note: τ_a is the attention softmax temperature; τ_{thr} is the decision threshold).

4. Experiments

In order to evaluate the performance of the proposed method, we apply it to a semi-supervised setting over a simulated dataset to visualize behavior and an industrial case study in the domain of industrial automotive quality management, where windshield wiper reversal noise is detected.

4.1. Experimental Setup

Experiments are conducted on a simulated two-dimensional dataset and a real industrial windshield-wiper acoustics dataset to compare the proposed GAR-DSVDD against established methods, including DeepSVDD, OCSVM, classical SVDD, and S3SVDD. All deep models share the same encoder capacity and optimization schedule for fairness. Threshold calibration follows each method's policy: for GAR-DSVDD and DeepSVDD, we use a train-quantile rule at a fixed level over labeled-normal scores, whereas OCSVM, SVDD, and S3SVDD use their native decision functions. Across both experiments, training follows a semi-supervised setting: a small subset of labeled normal samples and a large unlabeled pool that may include a small contamination of anomalies (bounded by ε). We construct four disjoint partitions: labeled normal observations, unlabeled, validation, and a held-out test set.

Performance is evaluated on the held-out test split using four measures derived from the confusion counts—true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

Overall accuracy (fraction of correctly classified instances):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}.$$

Detection rate (also called recall or true positive rate), measuring how many anomalies are correctly detected:

 doi:10.20944/preprints202510.1674.v1

$$\text{Detection Rate} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

The F1 score summarizes the trade-off between precision and detection rate under class imbalance. Define precision as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

then

$$\text{F1} = \frac{2\text{Precision} \times \text{Detection Rate}}{\text{Precision} + \text{Detection Rate}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}.$$

Balanced accuracy, averaging sensitivity (detection rate) and specificity (true negative rate). First define specificity

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

then

$$\text{Balanced Accuracy} = \frac{1}{2}(\text{Detection Rate} + \text{Specificity}).$$

For each dataset/seed, we test the directional alternative H_1: GAR-DSVDD > benchmark on the F1 score using paired t-tests and Wilcoxon signed-rank tests across seeds and report their p-values.

4.2. Simulated Data

In the simulated study, we utilize a two-cluster, two-dimensional dataset that has separable centers and evaluate it under a semi-supervised setting. We draw $n_{\text{norm}} = 1000$ normal observations and $n_{\text{anom}} = 500$ anomalies, label a small fraction of normal observations ($\rho = 0.03$), set unlabeled contamination to $\varepsilon = 0.20$, and reserve 20% for validation and 20% for testing. Figure 1 shows the training dataset used for the experiment, where labeled and unlabeled observations are used to train GAR-DSVDD and S3SVDD, while the other methods use labeled observations only. We carried out a total of 10 experiments, using different seeds when generating the data. The training set for one of the experiments is illustrated in Figure 1.

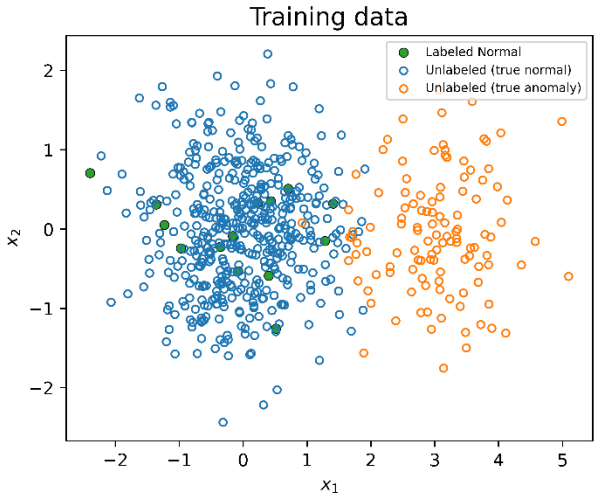


Figure 1. A scatter plot showing the training dataset in one of the experiments.

All deep methods share the same encoder ϕ_θ for fairness: a two-layer MLP with hidden widths $64 - 64$ mapping $x \mapsto z = \phi_\theta(x) \in \mathbb{R}^{16}$ with ReLU activations. Our method constructs a latent k-NN graph G on $\{z_i\}$ with $k = 9$ and attention weights (8 heads, key dimension $s_{att} = 32$). The attention-weighted graph is recomputed every $T = 5$ epochs and held fixed between refreshes; no

 doi:10.20944/preprints202510.1674.v1

gradients flow through the attention computation. We optimize the objective in Eq. (6) using AdamW for 120 epochs with refresh rate $T = 5$ and learning rate 5×10^{-4}, graph regularizer strength $\lambda_u = 1300$, and unlabeled center pull strength $\lambda_{\text{cp}} = 0.9$. Lastly, we calibrate m by setting it to the same train-quantile used for threshold selection over labeled normals.

As for the remaining methods, DeepSVDD is trained on labeled normal using the same applicable parameters of the proposed method, OSCVM uses $\nu = 0.1$, SVDD uses radial basis function (RBF) kernel parameter $\gamma = 1$, and S3SVDD uses RBF kernel parameter $\gamma = 1$ with $k = 9$. The deep methods use a train-quantile rule at $q = 0.95$, and the native decision function is used for the rest.

The performance results of the proposed methods compared to benchmarking methods over all testing performance metrics for an experiment are shown in Table 1, whereas Table 2 compares their F1 score over 10 different testing sets where the data was generated using different seeds.

Table 1. GAR-DSVDD and benchmark methods testing performance over a simulated dataset experiment.

Method	Accuracy	F1	Detection Rate	Precision	Specificity	Balanced Accuracy
GAR-DSVDD	0.92	0.92	0.99	0.86	0.84	0.92
DeepSVDD	0.56	0.69	1.00	0.53	0.11	0.56
OCSVM	0.76	0.81	0.99	0.68	0.53	0.76
SVDD	0.71	0.77	0.99	0.63	0.43	0.71
S3SVDD	0.69	0.76	0.99	0.62	0.38	0.69

Table 2. GAR-DSVDD and benchmark methods testing F1 score over 10 simulated dataset experiments.

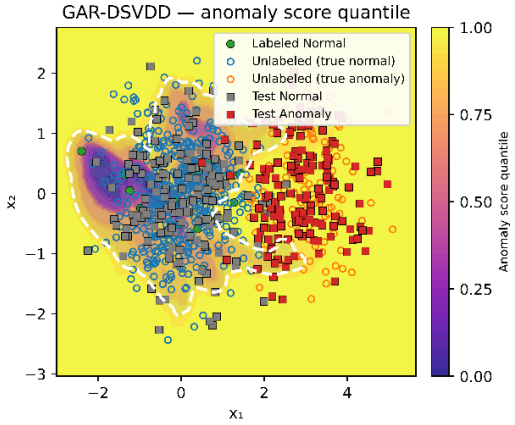
Experiment (over different seeds)	GAR-DSVDD	DeepSVDD	OCSVM	SVDD	S3SVDD
1	0.86	0.68	0.74	0.74	0.74
2	0.83	0.70	0.78	0.78	0.78
3	0.80	0.73	0.80	0.70	0.70
4	0.92	0.69	0.81	0.77	0.76
5	0.88	0.69	0.75	0.67	0.67
6	0.80	0.72	0.80	0.79	0.79
7	0.81	0.69	0.76	0.68	0.69
8	0.86	0.68	0.79	0.75	0.74
9	0.78	0.69	0.76	0.76	0.76
10	0.85	0.68	0.76	0.77	0.77
Mean	0.84	0.70	0.78	0.74	0.74
Standard deviation	0.04	0.02	0.02	0.04	0.04
p-value (Wilcoxon)	-	9.8×10^{-4}	9.8×10^{-4}	9.8×10^{-4}	9.8×10^{-4}
p-value (paired t-test)	-	7.3×10^{-6}	1.3×10^{-3}	4.2×10^{-4}	4.3×10^{-4}

DeepSVDD trains only on labeled normal observations and learns a tight hypersphere where many normal observations on low-density areas fall outside it, thus, increasing false positive rate and reducing overall accuracy, and F1 despite having high detection rate. OCSVM lacks representation learning and geometry-aware regularization, thus constructing a narrow decision boundary where detection rate remains high, but specificity and overall accuracy suffer, leading to a lower F1 score. Since classical SVDD does not utilize unlabeled-geometry guidance, it builds a decision boundary based on labeled observation alone, thus, obtaining a tight hypersphere yielding to a high false positive rate and reduced accuracy and F1 score. Incorporating a fixed-weight graph S3SVDD shows similar performance to SVDD as fixed-weight graph cannot truly suppress influence of suspect (anomalous) neighbors, thus, assumes that most unlabeled observations are anomalous and constructs a narrow hypersphere.

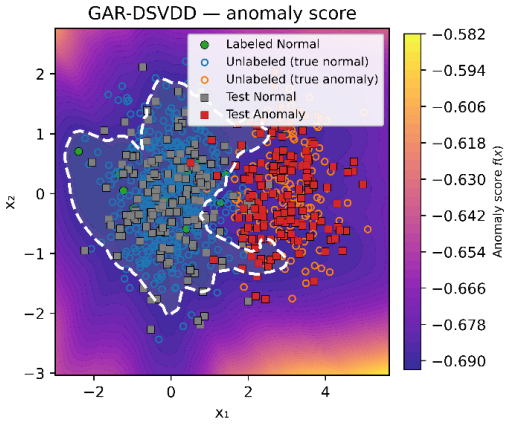
 doi:10.20944/preprints202510.1674.v1

In contrast, the attention-weighted GAR-DSVDD achieves the best F1 and balanced accuracy, among all other methods, by pairing high detection and lower false positives. Attention down weights influence from suspect neighbors in the unlabeled graph, limiting score dispersal from contaminated (anomalous) points and keeping the boundary near density valleys. The center pull regularizer on unlabeled observations prevents over smoothing decision boundary, preserving overall accuracy, while labeled normal observations ensure that the hypersphere is anchored to the normal region. Overall, GAR-DSVDD yields a more robust hypersphere across all experiments.

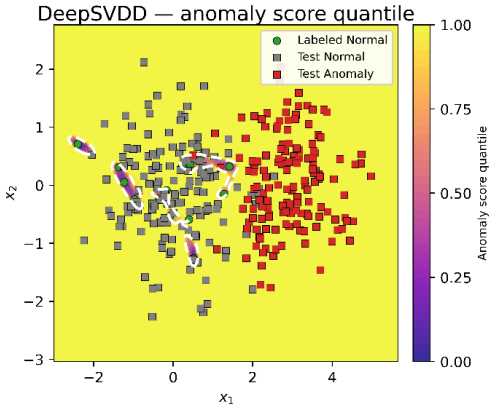
The obtained anomaly score and its quantiles are shown in Figure 2. In the figure, the 95^{th} quantile contour for the anomaly score is drawn to indicate the decision boundary for the deep methods, where the native decision boundary is shown for the shallow methods. Note that DeepSVDD and GAR-DSVDD use a single hypersphere in latent space as its decision region based on the learned quantile. In the figure, after the encoder's nonlinear mapping back to input space, this can look complex or even disconnected, despite being one sphere in the learned latent space.



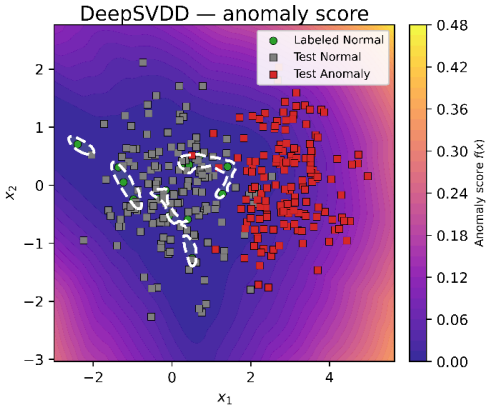
(a) GAR-DSVDD anomaly score quantile



(b) GAR-DSVDD anomaly score

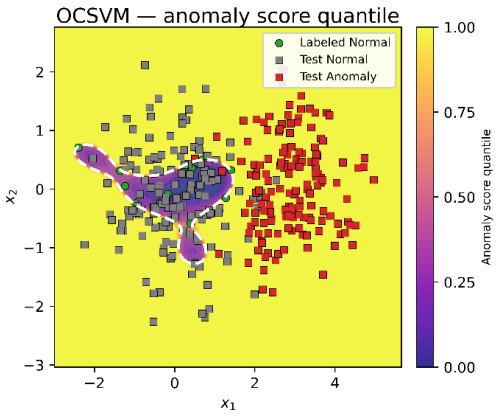


(c) DeepSVDD anomaly score quantile

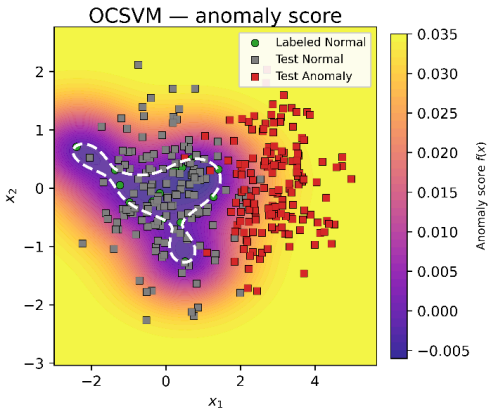


(d) DeepSVDD anomaly score

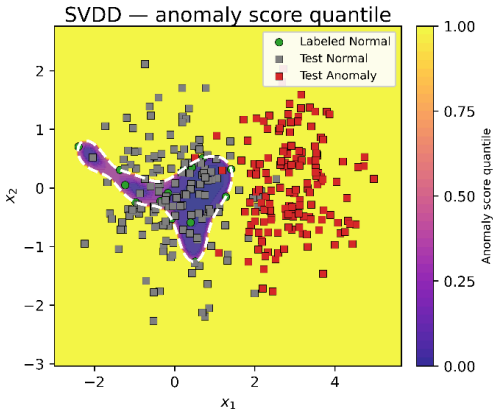
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1674.v1



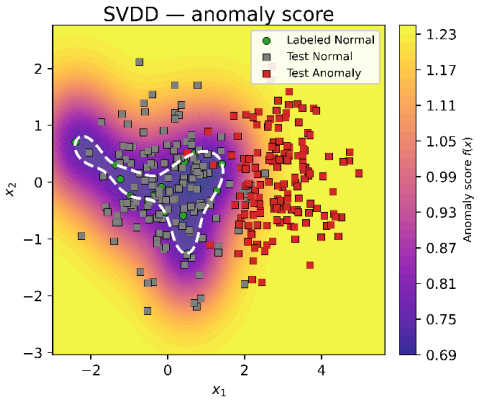
(e) OCSVM anomaly score quantile



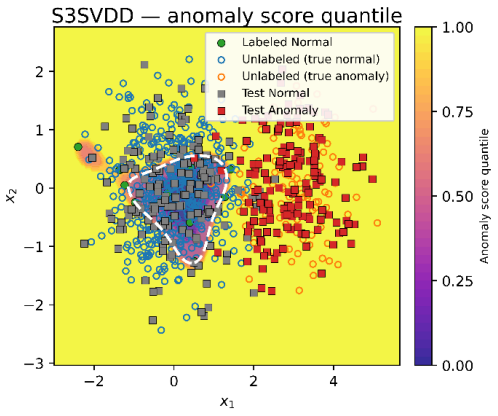
(f) OCSVM anomaly score



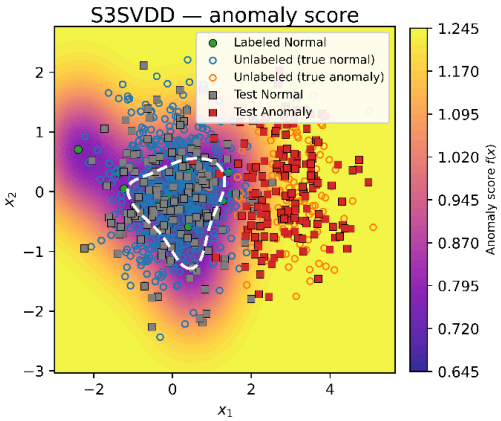
(g) SVDD anomaly score quantile



(h) SVDD anomaly score



(i) S3SVDD anomaly score quantile



(j) S3SVDD anomaly score

Figure 2. GAR-DSVDD and benchmarking methods decision boundaries at testing over the simulated dataset. The proposed method achieves the best decision boundary that well represents the normal region by taking into account unlabeled data information by utilizing its attention-based graph and its deep structure.

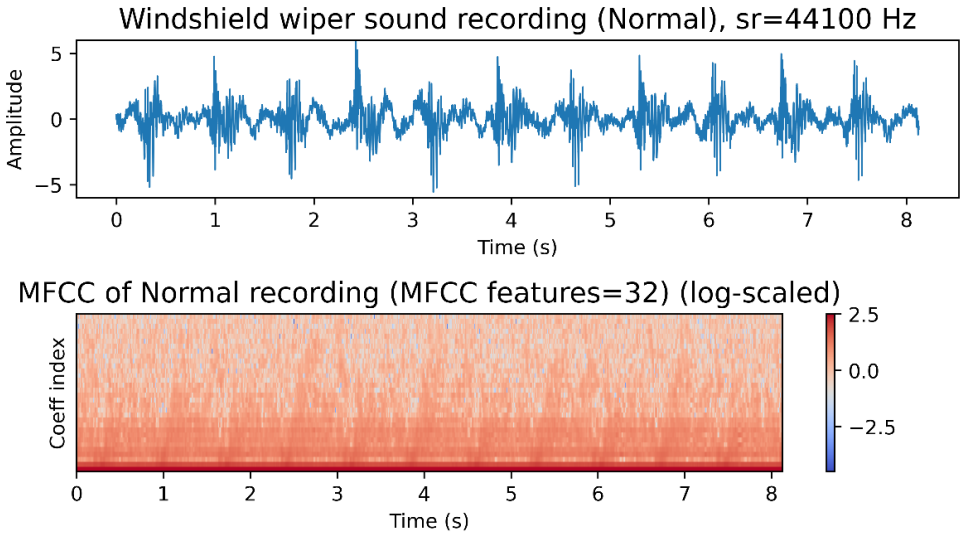
4.3. Case Study: Windshield-Wiper Acoustics

This section presents our case study on windshield wiper acoustics. We analyzed windshield-wiper sound recordings gathered under tightly controlled laboratory conditions. Each sample was captured from production wiper assemblies inside an anechoic chamber while a sprinkler system supplied water to replicate rainfall. To eliminate confounding noise sources, the vehicle's engine remained off, and the wiper motor was powered by an external supply. A fixed in-cabin microphone recorded the acoustic pressure signal, which was calibrated and expressed in decibels at a constant

 doi:10.20944/preprints202510.1674.v1

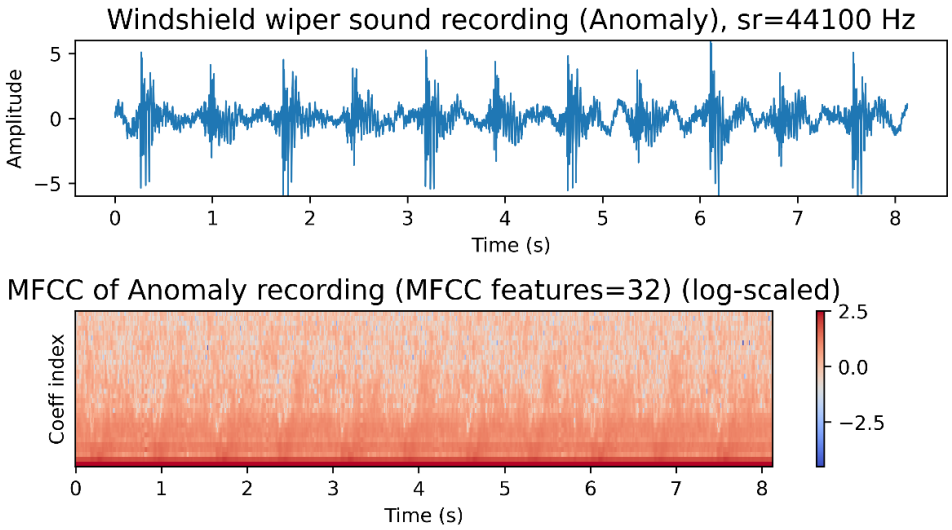
sampling rate. Domain specialists reviewed the recordings and assigned ground-truth labels covering reversal noise fault phenomena alongside normal operation. The dataset consists of a total of 120 windshield wiper recordings, where 61 recordings are for normal operations and 59 for faulty (anomalous) operations. These expert annotations are treated as ground truth for evaluation.

Each observation is a single full recording summarized as a fixed-length vector from recording descriptors, in line with recent acoustic research for sound analysis. We compute 32 MFCCs to capture the spectral envelope on a perceptual scale, together with first- and second-order deltas to encode short-term dynamics; these cepstral families remain standard and effective in contemporary studies. We also include a 12-bin chroma profile to reflect tonal/resonant structure and simple spectral/temporal statistics—spectral centroid, spectral bandwidth, spectral roll-off at 95%, root-mean-square (RMS) energy, and zero-crossing rate (ZCR)—which are routinely used and compared in modern journal work. For every multi-frame descriptor with r coefficients, we pool across time by concatenating the per-coefficient mean and standard deviation, so each such descriptor contributes $2r$ values; scalar descriptors contribute two values (mean and standard deviation). Summing up all parts yields a $226-$dimensional vector per recording of means and standard deviations across time that we use for the reversal-noise dataset. These choices (MFCCs with deltas, chroma, and spectral statistics with mean/std pooling) align with recent papers documenting their efficacy and definitions across environmental-sound, speech-emotion, smart-audio, and urban-acoustic applications [33–35]. Figure 3 shows normal and anomaly operation recordings of windshield wipers with their corresponding MFCC features on the log scale. Table 3 summarizes the feature extraction methods used for acoustic analysis of the windshield wiper sound recordings.



(a) Windshield wipers' normal operation sound recording and its log-scaled MFCC features.

 doi:10.20944/preprints202510.1674.v1



(b) Windshield wipers anomaly operation sound recording and its log-scaled MFCC features.

Figure 3. An illustration of the windshield wipers sound recording and their MFCC features.

Table 3. Extracted features summary for acoustic analysis of windshield wiper operation sound recording.

Feature family	Recording dimension (means & standard deviations)*	Brief description	Reference
MFCC (32)	64	Cepstral summary of spectral envelope on the mel scale, a widely adopted baseline in recent ESC/SER studies.	[33,34]
ΔMFCC (32)	64	First-order derivative of MFCCs capturing short-term spectral dynamics.	[34]
Δ²MFCC (32)	64	Second-order derivative (acceleration) of MFCCs, emphasizing rapid spectral change.	[34]
Chroma (12)	24	Energy folded into 12 pitch-class bins; it reflects tonal/resonant structure seen in mechanical acoustics.	[33,34]
Spectral centroid	2	Power-weighted mean frequency (proxy for "brightness").	[33,34]
Spectral bandwidth	2	Spread around the centroid (spectral dispersion).	[33]
Spectral roll-off (95%)	2	Frequency below which 95% of energy lies (high-frequency content indicator).	[33,34]
RMS energy	2	Framewise signal power (overall loudness proxy).	[33,34]
ZCR	2	Sign-change rate (simple proxy for roughness/high-frequency content).	[33,34]

*Total observation dimension: 1 by $2 \times (32 + 32 + 32 + 12 + 1 + 1 + 1 + 1 + 1) = 226$.

We follow the same semi-supervised protocol as in the simulated study: a small subset of labeled-normal clips is provided, and a large unlabeled pool may include mild anomaly contamination; the remainder is split into validation (for hyperparameter selection) and a held-out test set. Concretely, we use a 60/20/20 split for train/validation/test. Within the training portion, we label 10% of the normal observations ($\rho = 0.10$) and place the rest into the unlabeled set with anomaly contamination ratio $\varepsilon = 0.10$. All deep methods use the shared encoder ϕ_θ: a two-layer MLP 64-64 → 16 with ReLU activations. For our method, GAR-DSVDD, we build a latent k-NN graph with $k = 3$ and multi-head attention (8 heads, key dimension 32) over the encoder representations. The attention-weighted graph is recomputed every $T = 5$ epochs and held fixed between refreshes; no gradients flow through the attention computation. The objective in Eq. (6) for 120 epochs using AdamW (learning rate 5×10^{-4}), graph regularizer strength $\lambda_u = 8$, and unlabeled center pull strength $\lambda_{\text{cp}} = 1000$. m is calibrated by setting it to the same train-quantile used for threshold selection over labeled normals.

Benchmarks are configured to match their best settings from the wiper runs: DeepSVDD is trained on labeled normal observations with the same encoder/optimizer, soft-boundary objective

 doi:10.20944/preprints202510.1674.v1

regularizer value $= 1000$; OCSVM uses $\nu = 0.1$ with an RBF kernel $\gamma = 0.01$; SVDD uses an RBF kernel with $\gamma = 1.0$; and S3SVDD employs a k −NN graph with $k = 3$ and RBF $\gamma = 0.01$. The deep methods apply a train-quantile rule with $q = 0.95$, while OCSVM, SVDD, and S3SVDD use their native decision function.

The performance of GAR-DSVDD and the existing methods overall performance metrics are presented in Table 4. Table 5 shows the F1 score over 10 different testing sets that were prepared using 10 different seeds for statistical validation.

Due to the extreme challenges represented by this dataset and the semi-supervised experiment settings with very few training labeled normal observation, all benchmarking methods perform poorly for anomaly detecting as they fail to utilize unlabeled observations in constructing their decision boundaries as indicated by the F1 and accuracy metrics. The decision boundary obtained by these methods tend to be very narrow, causing most testing observations to be detected as anomalies while increasing the false positive rates, as seen by the specificity metric for these methods.

In contrast, GAR-DSVDD maintains perfect detection while achieving a higher specificity by leveraging an attention-weighted latent k-NN graph that weakens edges in mixed or uncertain neighborhoods, thus, effectively utilizing information from unlabeled observation. This reduces the graph-driven propagation of elevated scores into nearby normal observations, keeps the decision boundary aligned with low-density regions, and prevents the false positive rates that undermine the other methods. Furthermore, the learned neighbor attention weights serve as per-instance attributes, indicating which historical normal observations most influence a observation's score. In an industrial setting, these attributions can streamline expert review and reduce the volume of fully verified normal labels required over time, aligning the method with label-efficiency needs in automotive quality control.

Table 4. GAR-DSVDD and benchmark methods testing performance over the industrial windshield wiper dataset.

Method	Accuracy	F1	Detection Rate	Precision	Specificity	Balanced Accuracy
GAR-DSVDD	0.92	0.91	1	0.83	0.86	0.93
DeepSVDD	0.42	0.59	1	0.42	0	0.5
OCSVM	0.42	0.59	1	0.42	0	0.5
SVDD	0.42	0.59	1	0.42	0	0.5
S3SVDD	0.42	0.59	1	0.42	0	0.5

Table 5. GAR-DSVDD and benchmark methods testing F1 score over 10 industrial windshield wiper experiments where different seeds are used to split the dataset into training, validation, and testing.

Experiment (over different seeds)	GAR-DSVDD	DeepSVDD	OCSVM	SVDD	S3SVDD
1	0.86	0.55	0.55	0.55	0.55
2	0.93	0.8	0.8	0.82	0.8
3	0.76	0.5	0.5	0.52	0.5
4	0.79	0.60	0.63	0.63	0.63
5	0.67	0.63	0.63	0.63	0.63
6	0.91	0.59	0.59	0.59	0.59
7	0.83	0.45	0.45	0.45	0.45
8	0.97	0.74	0.74	0.74	0.74
9	0.96	0.7	0.7	0.7	0.7
10	0.90	0.63	0.63	0.69	0.63
Mean	0.86	0.62	0.62	0.63	0.62
Standard deviation	0.10	0.11	0.11	0.11	0.11
p-value (Wilcoxon)	-	9.8×10^{-4}	9.8×10^{-4}	9.8×10^{-4}	9.8×10^{-4}
p-value (paired t-test)	-	1.7×10^{-5}	2.2×10^{-5}	3.4×10^{-5}	2.2×10^{-5}

 doi:10.20944/preprints202510.1674.v1

4.4. Sensitivity Analysis

In this section, we perform sensitivity analysis over the simulated data. In Table 6 and Figure 4 we show the effect of increasing the normal labeled observations ratio ρ over the F1 performance of the methods. Across labeled-normal ratios $\rho \in \{0.10, 0.20, 0.50\}$, the results show a clear label-efficiency advantage for our semi-supervised GAR-DSVDD in the low-label setting, with supervised baselines gradually converging as ϱ increases. With scarce labels, GAR-DSVDD leads, consistent with its attention-weighted graph regularization that leverages unlabeled structure while mitigating contamination. As the label budget grows to a moderate level, GAR-DSVDD remains best, but the gap narrows, particularly with DeepSVDD, which benefits directly from richer labeled evidence. OCSVM improves steadily with more labels yet trails the deep methods, while classical SVDD consistently lags despite incremental gains. S3SVDD exhibits non-monotonic behavior, sensitive to label density and split composition, but becomes more competitive with increased labeling ratio. In summary, when only a small portion of data can be labeled, our target operating point, GAR-DSVDD outperforms all alternatives; as labeling increases, purely supervised approaches like DeepSVDD catch up and have similar performance, whereas OCSVM and SVDD continue to benefit but remain behind, and S3SVDD's competitiveness depends on the labeling setting.

Table 6. Effect of Labeled-Normal Ratio (ρ) on F1 for GAR-DSVDD and benchmarks (DeepSVDD, OCSVM, SVDD, S3SVDD).

ρ	GAR-DSVDD	DeepSVDD	OCSVM	SVDD	S3SVDD
0.1	0.91	0.85	0.87	0.68	0.88
0.2	0.91	0.90	0.88	0.71	0.77
0.5	0.93	0.93	0.91	0.77	0.90

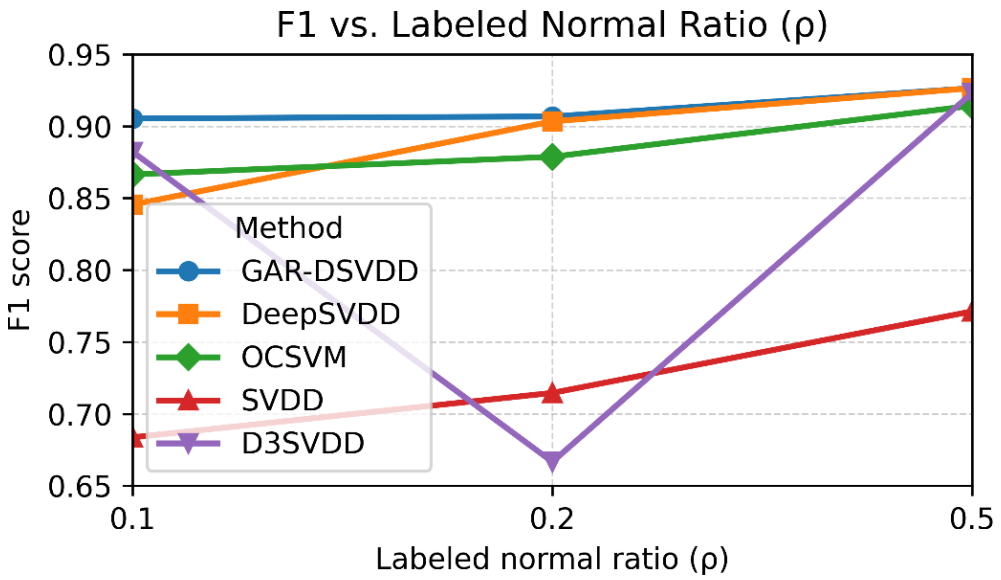


Figure 4. Test F1 vs. ρ. GAR-DSVDD leads when labels are scarce ($\rho = 0.10$) and stays competitive, while DeepSVDD improves with more labels. OCSVM rises steadily but trails; SVDD is lowest; S3SVDD is non-monotonic (dip at $\rho = 0.20$).

Moreover, by setting $\rho = 0.03$, we analyze the effect of λ_u on F1 and accuracy of GAR-DSVDD, as illustrated in Figure 5. For small λ_u ($\approx 1 - 500$), less importance is given to unlabeled geometry, thus, the unlabeled information is underutilized, leading to narrow decision boundary and underperforming results. Increasing λ_u gives more importance to the graph-loss term (unlabeled geometry), which expands the decision boundary with consistent gains in performance. The performance is highest near $\lambda_u \approx 1300$, where the boundary best balances the influence of labeled

 doi:10.20944/preprints202510.1674.v1

data information, unlabeled geometry, and central-pull term on unlabeled observations. At higher λ_u values, excessive emphasis on unlabeled geometry over-smooths the graph structure, widening the decision boundary and allowing anomalous unlabeled observations to be included within the decision boundary, thereby degrading performance. These findings indicate that λ_u should be carefully selected via validation to allow an optimal construction of the decision boundary that keeps anomalous unlabeled observations outside and expands enough to keep normal behavior observations inside.

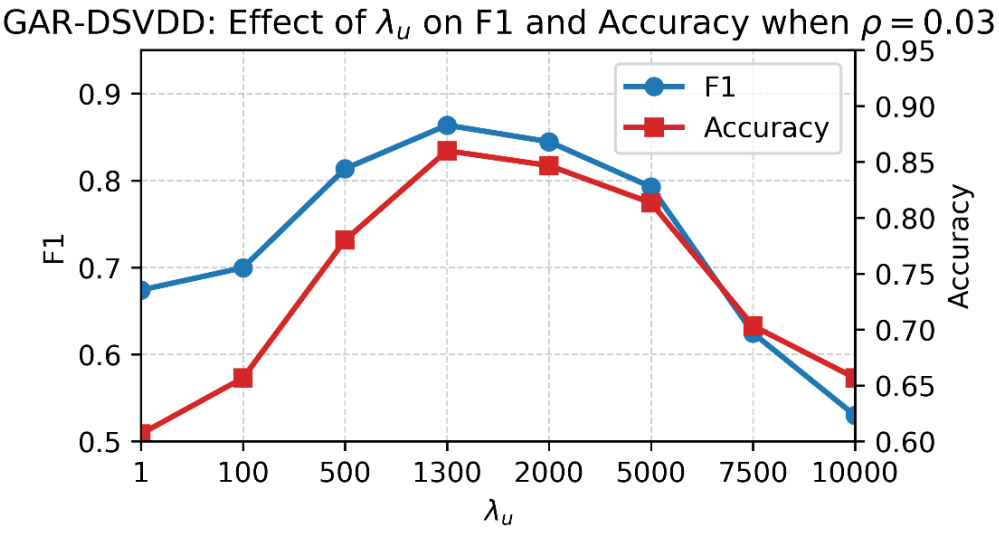


Figure 5. Effect of λ_u on GAR-DSVDD performance at $\rho = 0.03$. Both F1 (left axis) and accuracy (right axis) improve as λ_u increases from 1 to 1300, peaking near $\lambda_u \approx 1300$, then declining for larger values, indicating over-regularization beyond the optimum.

5. Conclusions

Identifying anomalies is a critical issue in applications such as manufacturing, industrial process control, cybersecurity, disaster management, and medical diagnosis, as they provide information about deviations from normal behaviour in critical environments [1–5]. Among available approaches, SVDD-based procedures remain attractive for their compact decision boundaries and test-time simplicity.

In this paper, we presented a GAR-DSVDD for semi-supervised anomaly detection. The method encloses labeled normal observations with a deep one-class boundary while propagating score-level smoothness over an attention-weighted latent k-NN graph to leverage abundant unlabeled data. The attention mechanism reduces over-smoothing from questionable neighbors and respects local density breaks, improving specificity without sacrificing detection. Experiments on simulated data and a windshield-wiper acoustics case study demonstrate that GAR-DSVDD outperforms classical and deep baselines under scarce labels significantly.

Considering future work, many industrial datasets contain multiple normal classes (e.g., distinct operating conditions or product variants). Our current formulation targets a single normal class; extending it to multi-normal settings is a natural next step. One direction is to learn several hyperspheres using a shared encoder, where each hypersphere retains its corresponding normal class while discouraging overlapping with others; the graph attention can be made class-aware to limit cross-class leakage. Finally, online updates with periodic graph refresh and automatic selection of graph/attention hyperparameters would enhance robustness in evolving production environments.

Data Availability Statement: The industrial windshield wiper acoustic data presented in this study are available on reasonable requests from the corresponding author due to confidentiality. The code for the simulated dataset can be found at https://github.com/alhinditaha/GAR-DSVDD.

 doi:10.20944/preprints202510.1674.v1

Acknowledgments: The project was funded by KAU Endowment (WAQF) at King Abdulaziz University, Jeddah, Saudi Arabia. The authors, therefore, acknowledge with thanks WAQF and the Deanship of Scientific Research (DSR) for technical and financial support.

Conflicts of Interest: The authors declare no conflicts of interest.

GenAI Usage Disclosure: No generative artificial intelligence (GenAI) technology was used for generating text, data or graphics, study design, or data collection, analysis or interpretation. GenAI tools (ChatGPT, OpenAI: GPT5) were solely used, under the authors' supervision, to enhance readability and refine language (grammar, spelling, punctuation, and formatting). The paper has been carefully reviewed by the authors to ensure its accuracy and coherence.

Abbreviations

The following abbreviations are used in this manuscript:

AdamW	Adaptive Moment Estimation with (decoupled) Weight decay
DeepSVDD	Deep Support Vector Data Description
GAR-DSVDD	Graph-Attention-Regularized Deep Support Vector Data Description
k-NN	k-Nearest Neighbors
MFCC	Mel-Frequency Cepstral Coefficients
OCSVM	One-Class Support Vector Machine
ReLU	Rectified Linear Unit
S3SVDD	Graph-based Semi-Supervised Support Vector Data Description
SVDD	Support Vector Data Description

References

1. Li, H.; Boulanger, P. A survey of heart anomaly detection using ambulatory Electrocardiogram (ECG). *Sensors* **2020**, *20*, 1461.
2. Alhindi, T.J.; Alturkistani, O.; Baek, J.; Jeong, M.K. Multi-class support vector data description with dynamic time warping kernel for monitoring fires in diverse non-fire environments. *IEEE Sensors Journal* **2025**.
3. Sakong, W.; Kim, W. An adaptive policy-based anomaly object control system for enhanced cybersecurity. *IEEE Access* **2024**, *12*, 55281-55291.
4. Karsaz, A. A modified convolutional neural network architecture for diabetic retinopathy screening using SVDD. *Applied Soft Computing* **2022**, *125*, 109102.
5. Cai, L.; Yin, H.; Lin, J.; Zhou, H.; Zhao, D. A relevant variable selection and SVDD-based fault detection method for process monitoring. *IEEE Transactions on Automation Science and Engineering* **2022**, *20*, 2855-2865.
6. Alhindi, T.J.; Baek, J.; Jeong, Y.-S.; Jeong, M.K. Orthogonal binary singular value decomposition method for automated windshield wiper fault detection. *International Journal of Production Research* **2024**, *62*, 3383-3397, doi:10.1080/00207543.2023.2232471.
7. Tax, D.M.; Duin, R.P. Support vector data description. *Machine learning* **2004**, *54*, 45-66.
8. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural computation* **2001**, *13*, 1443-1471.
9. Zhu, X.; Ghahramani, Z.; Lafferty, J.D. Semi-supervised learning using gaussian fields and harmonic functions. In Proceedings of the Proceedings of the 20th International Conference on Machine Learning (ICML-03), 2003; pp. 912-919.
10. Belkin, M.; Niyogi, P.; Sindhwani, V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of machine learning research* **2006**, *7*.
11. Luo, X.; Wu, J.; Yang, J.; Xue, S.; Peng, H.; Zhou, C.; Chen, H.; Li, Z.; Sheng, Q.Z. Deep graph level anomaly detection with contrastive learning. *Scientific Reports* **2022**, *12*, 19867.

 doi:10.20944/preprints202510.1674.v1

12. Ma, X.; Wu, J.; Xue, S.; Yang, J.; Zhou, C.; Sheng, Q.Z.; Xiong, H.; Akoglu, L. A comprehensive survey on graph anomaly detection with deep learning. *IEEE transactions on knowledge and data engineering* **2021**, *35*, 12012-12038.
13. Qiao, H.; Tong, H.; An, B.; King, I.; Aggarwal, C.; Pang, G. Deep graph anomaly detection: A survey and new perspectives. *IEEE Transactions on Knowledge and Data Engineering* **2025**.
14. Duong, P.; Nguyen, V.; Dinh, M.; Le, T.; Tran, D.; Ma, W. Graph-based semi-supervised support vector data description for novelty detection. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), 2015; pp. 1-6.
15. Mygdalis, V.; Iosifidis, A.; Tefas, A.; Pitas, I. Semi-supervised subclass support vector data description for image and video classification. *Neurocomputing* **2018**, *278*, 51-61.
16. Wu, X.; Liu, S.; Bai, Y. The manifold regularized SVDD for noisy label detection. *Information Sciences* **2023**, *619*, 235-248.
17. Peng, D.; Liu, C.; Desmet, W.; Gryllias, K. Semi-supervised CNN-based SVDD anomaly detection for condition monitoring of wind turbines. In Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering, 2022; p. V001T001A019.
18. Von Luxburg, U. A tutorial on spectral clustering. *Statistics and computing* **2007**, *17*, 395-416.
19. Zelnik-Manor, L.; Perona, P. Self-tuning spectral clustering. *Advances in Neural Information Processing Systems* **2004**, *17*.
20. Song, Y.; Zhang, J.; Zhang, C. A survey of large-scale graph-based semi-supervised classification algorithms. *International Journal of Cognitive Computing in Engineering* **2022**, *3*, 188-198.
21. Ruff, L.; Vandermeulen, R.; Goernitz, N.; Deecke, L.; Siddiqui, S.A.; Binder, A.; Müller, E.; Kloft, M. Deep one-class classification. In Proceedings of the International Conference on Machine Learning, 2018; pp. 4393-4402.
22. Pota, M.; De Pietro, G.; Esposito, M. Real-time anomaly detection on time series of industrial furnaces: A comparison of autoencoder architectures. *Engineering Applications of Artificial Intelligence* **2023**, *124*, 106597.
23. Neloy, A.A.; Turgeon, M. A comprehensive study of auto-encoders for anomaly detection: Efficiency and trade-offs. *Machine Learning with Applications* **2024**, *17*, 100572.
24. Tack, J.; Mo, S.; Jeong, J.; Shin, J. Csi: Novelty detection via contrastive learning on distributionally shifted instances. *Advances in Neural Information Processing Systems* **2020**, *33*, 11839-11852.
25. Hojjati, H.; Ho, T.K.K.; Armanfard, N. Self-supervised anomaly detection in computer vision and beyond: A survey and outlook. *Neural Networks* **2024**, *172*, 106106.
26. Chalapathy, R.; Chawla, S. Deep learning for anomaly detection: A survey. *ArXiv Preprint arXiv:1901.03407* **2019**.
27. Pang, G.; Shen, C.; Cao, L.; Hengel, A.V.D. Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)* **2021**, *54*, 1-38.
28. Ruff, L.; Kauffmann, J.R.; Vandermeulen, R.A.; Montavon, G.; Samek, W.; Kloft, M.; Dietterich, T.G.; Müller, K.-R. A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE* **2021**, *109*, 756-795.
29. Li, Z.; Zhu, Y.; Van Leeuwen, M. A survey on explainable anomaly detection. *ACM Transactions on Knowledge Discovery from Data* **2023**, *18*, 1-54.
30. Bouman, R.; Heskes, T. Autoencoders for Anomaly Detection are Unreliable. *arXiv preprint arXiv:2501.13864* **2025**.
31. Kim, S.; Lee, S.Y.; Bu, F.; Kang, S.; Kim, K.; Yoo, J.; Shin, K. Rethinking reconstruction-based graph-level anomaly detection: limitations and a simple remedy. *Advances in Neural Information Processing Systems* **2024**, *37*, 95931-95962.
32. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *Stat* **2017**, *1050*, 10-48550.
33. Bansal, A.; Garg, N.K. Environmental Sound Classification: A descriptive review of the literature. *Intelligent Systems with Applications* **2022**, *16*, 200115.

 doi:10.20944/preprints202510.1674.v1

34. Madanian, S.; Chen, T.; Adeleye, O.; Templeton, J.M.; Poellabauer, C.; Parry, D.; Schneider, S.L. Speech emotion recognition using machine learning—A systematic review. *Intelligent Systems with Applications* **2023**, *20*, 200266.
35. Mannem, K.R.; Mengiste, E.; Hasan, S.; de Soto, B.G.; Sacks, R. Smart audio signal classification for tracking of construction tasks. *Automation in Construction* **2024**, *165*, 105485.