

# Towards High-resolution 3D Anomaly Detection via Group-Level Feature Contrastive Learning

Anonymous Authors

## ABSTRACT

High-resolution point clouds (HRPCD) anomaly detection (AD) plays a critical role in precision machining and high-end equipment manufacturing. Despite considerable 3D-AD methods that have been proposed recently, they still cannot meet the requirements of the HRPCD-AD task. There are several challenges: i) It is difficult to directly capture HRPCD information due to large amounts of points at the sample level; ii) The advanced transformer-based methods usually obtain anisotropic features, leading to degradation of the representation; iii) The proportion of abnormal areas is very small, which makes it difficult to characterize. To address these challenges, we propose a novel group-level feature-based network, called Group3AD, which has a significantly efficient representation ability. First, we design an Intercluster Uniformity Network (IUN) to present the mapping of different groups in the feature space as several clusters, and obtain a more uniform distribution between clusters representing different parts of the point clouds in the feature space. Then, an Intracluster Alignment Network (IAN) is designed to encourage groups within the cluster to be distributed tightly in the feature space. In addition, we propose an Adaptive Group-Center Selection (AGCS) based on geometric information to improve the pixel density of potential anomalous regions during inference. The experimental results verify the effectiveness of our proposed Group3AD, which surpasses Reg3D-AD by the margin of 5% in terms of object-level AUROC on Real3D-AD.

## CCS CONCEPTS

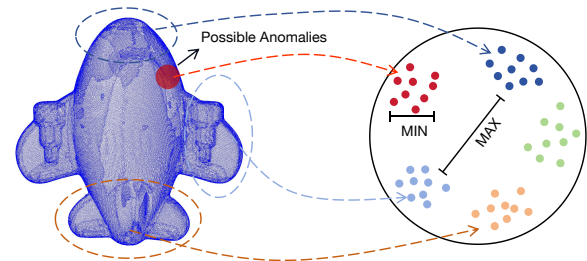
• **Computing methodologies** → **Visual inspection; Anomaly detection.**

## KEYWORDS

Anomaly Detection, 3D Point Clouds, Contrastive Learning, Feature Representation

## 1 INTRODUCTION

Anomaly Detection (AD) is a critical field in machine learning aimed at identifying unusual patterns or abnormalities that do not conform to expected behavior. Traditionally, AD has been extensively applied in 2D image analysis, where methods primarily focus on identifying anomalies through pixel-level discrepancies [25, 39]. However, these 2D techniques come with inherent limitations, such



**Figure 1: Ideal feature distribution of normal point clouds and anomalies for the high-resolution 3D anomaly detection task. Group-level features are used to express structural information.**

as the fixed perspectives and inability to capture complex geometries, which often result in loss of important spatial information and context [3, 8, 10, 22, 25, 39, 40]. Transitioning from 2D images to 3D Point Clouds (PCD) effectively overcomes these limitations. 3D PCDs, which are sets of data points in space, provide a more comprehensive representation of objects by capturing their shape and spatial hierarchy in real environments [4, 16].

Despite the advantages, the field of 3D-AD faces many challenges [9, 24]. The High-Resolution (HR) of 3D PCDs introduces computational and analytical complexities due to the sheer increase in data volume and the intricacies involved in 3D space analysis [21, 24]. Research based on HRPCD-AD has just emerged, and how to construct efficient representations for AD tasks in HRPCD has become a major challenge. There is an urgent need to improve the precision of HRPCD-AD to meet the needs of industrial manufacturing (IM) [21, 24]. The obstacle on the road to establishing efficient representation is threefold. i) Existing HRPCD networks necessitate downsampling for large datasets, risking the loss of crucial anomaly information. ii) Recent 3D representation methods yield embeddings with insufficient distinction in AD. iii) The proportion of anomalies in HRPCD is small and obscure, making it easy to overlook [18, 21, 24, 25, 39].

To make the training of the network free from the constraints of HR PCDs and too few training samples, we introduce the idea of group-level feature enhancement. Existing HRPCD networks clearly require downsampling operations when faced with items composed of millions of points during training or inference [21, 24]. Among them, although Reg3D-AD [24] has a higher group density in HRPCD reasoning in a way similar to PatchCore [29], it still cannot be directly trained by HRPCD. There is a gap in the middle, which affects the performance. IMRNet also realized the drawbacks of downsampling and proposed adaptive sampling, but still did not fundamentally solve the problem. We need to study a solution that is not constrained by point cloud resolution. Unlike previous 3D contrastive learning studies [1, 23, 26], we build contrastive

Unpublished working draft. Not for distribution.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted by ACM, provided that the copies are not made for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

ACM MM, 2024, Melbourne, Australia

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM

<https://doi.org/10.1145/nnnnnnn.nnnnnn>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116

learning at group level in a single sample to makes our network training unconstrained by point cloud resolution and scale. The divide and conquer design makes our network's group batchsize infinitely large in theory, although its batchsize is 1.

To achieve better representation, we pose a key question: **how to create an ideal distribution required by HRPCD-AD in the feature space?** Currently, most network backbones that perform well on HRPCD use the Point Transformer architecture [21, 24]. The embeddings represented by this encoder exhibit anisotropic distribution in the feature space. Research [12, 13, 37] has shown that transformer-based structures place more emphasis on local contextual information in representation, while ignoring global semantic information. This results in high-frequency groups being distributed in a narrow area and close to the center in the feature space. While low-frequency groups are relatively sparse and far from the origin. Commonly, calculating feature similarity is a widely-used method, but this anisotropic spatial distribution will bring problems to the measurement of feature similarity. Figure 1 shows an ideal feature space distribution. Specifically, excellent spatial distribution of features should have both uniformity and alignment [7, 14]. Uniformity requires that the vectors should be distributed as widely as possible and isotropic in the space, while alignment means that the distance between similar vectors in the feature space should be small. Meanwhile, the excellent encoder should be capable of higher anomaly sensitivity to easily distinguish abnormal representation from normal representation in the feature space. Considering above-mentioned distribution characteristics: **intercluster dispersion** and **intracluster compactness**, we propose to construct the group-level contrastive learning architecture. The proposed Intercluster Uniformity Network (IUN) pushes clusters representing similar features far away in the feature space to enhance the uniformity of feature distribution. On the basis of IUN, the Intracluster Alignment Network (IAN) further tightens the features in the same cluster to strengthen the alignment of feature distribution.

To better capture subtle features, we design the Adaptive Group-Center Selection (AGCS). Drawing inspiration from the practices of quality inspection engineers, who often focus more intensively on areas they suspect to be problematic to enhance the detection of abnormalities, AGCS similarly prioritizes regions within 3D point clouds that exhibit potential anomalies. Leveraging the Fast Point Feature Histogram (FPFH) [32], AGCS intensifies the examination of areas with significant local geometric variations. Finally, AGCS enables our network to preserve more potential anomaly groups during inference, thereby enhancing its AD capability.

Our contributions are succinctly outlined as follows:

- A novel Group3AD framework for the HRPCD-AD task is designed, which optimizes the spatial information of 3D point clouds, enhancing the precision of anomaly detection.
- We propose the Intercluster Uniformity Network (IUN) and the Intracluster Alignment Network (IAN), which can disperse features across clusters and tighten features within clusters in the feature space, respectively. Both two networks enhance feature uniformity and alignment, improving the coherence of feature representations for anomaly identification and detection accuracy.

- An efficient Adaptive Group-Center Selection (AGCS) is designed, which focuses on regions with potential anomalies, enhancing model sensitivity and detection efficiency.
- The proposed Group3AD is flexible and scalable. Group3AD can be directly integrated with other network architectures, which promotes the wider application of various anomaly detection tasks.

## 2 RELATED WORK

### 2.1 2D Anomaly Detection

Since the introduction of the MVTec AD dataset [2], 2D image anomaly detection (2D-AD) has garnered increased focus [25, 39]. Predominantly, research in this domain utilizes this 2D dataset for exploring unsupervised AD techniques [3, 10, 11, 15, 15, 20, 27, 29, 30, 30, 34, 36, 41, 42]. The intention of image reconstruction-based 2D-AD methods is to reconstruct abnormal images into approximate normal images and achieve anomaly localization through pixel-level comparison. The feature extraction based 2D-AD methods strives to provide more informative embeddings, with more significant differences between normal and anomalous features. Due to the fact that many networks in the former use the method of training from scratch, their performance may be inferior to that of the latter, which uses robust pre-training models.

Notable examples include DRAEM [41], SimpleNet [27], etc. They create fake samples and identify anomalies through reconstruction comparison, building supervised tasks in unsupervised datasets. The representation of pre-trained networks has been proven to be more powerful and effective. PatchCore [29] employs a memory-efficient representation of normal data distribution through sparse sampling of feature space patches for identifying outliers. Cut-Paste [20] markedly enhances AD capabilities by artificially generating anomalies through cutting and pasting segments within images. Self-Taught AD [3] leverages a student-teacher framework where features from a pre-trained teacher network guide a student network trained on normal data to detect anomalies based on feature discrepancies. STPM [36] and MKD [33] both employ a teacher-student architecture for AD, with STPM harnessing multi-scale features directly, while MKD concentrates on distilling knowledge from these multi-scale features through a more efficient network architecture to enhance performance. Normalizing flow [19] methods have created a feature distribution with anomaly sample deviation, expanding its potential. Although 2D-AD methods cannot be directly used for 3D-AD, they highlight the importance of building an easily distinguishable feature distribution and enhancing the representational ability of the network [25, 39].

### 2.2 RGB-D Anomaly Detection

Despite significant advancements in 2D image anomaly detection, the exploration of anomaly detection in 3D PCDs remains relatively underdeveloped [5, 6, 9]. This field saw a notable surge in interest following the release of the pioneering real-world 3D anomaly detection dataset, known as MVTec 3D-AD [4]. This dataset has sparked new investigations into the complexities and potentialities of detecting anomalies within 3D environments. The key challenge in 3D-AD lies in the effective harnessing of depth information, which can significantly enhance detection capabilities in certain

scenarios. This complexity, alongside the potential for improved accuracy, positions 3D-AD as both a challenging and promising avenue for future research.

While AST [31] achieves good results on the MVTEC 3D-AD by utilizing depth information to separate the background, it primarily depends on 2D-AD techniques for anomaly detection, neglecting the depth characteristics of objects. Designing the feature extraction module for detecting anomalies in PCD demands fresh approaches. Recent studies have made efforts to design networks with stronger point clouds representation capabilities [8]. 3D-ST [5] introduces a self-supervised learning strategy for representation learning in PCD and employs a knowledge distillation-based modeling module. While CPMF [8] streamlines anomaly detection by projecting point clouds into 2D images from various angles to reduce feature extraction complexity and computational costs, this approach does not fully leverage the advantages inherent in the 3D nature of point clouds data. BTF [17] underscores the efficacy of traditional handcrafted PCD descriptors, yet notes the underperformance of learned features. BTF attributes this paradox to the inadequate transferability of existing pretrained features on the current small-scale PCD datasets. The use of spatial information for AD tasks still deserves further exploration [8, 21, 24, 38].

## 2.3 High-resolution 3D Anomaly Detection

Some methods have made efforts in the direction of PCD anomaly representation. However, MVTEC 3D-AD [4] is an RGB-D dataset with low resolution, which cannot further explore the value of spatial information in AD tasks. The proposal of the Real3D-AD [24] dataset containing HR multi-view information of objects brings more development space to AD. The point resolution and precision of Real3D-AD are 4.28 and 9 times higher than MVTEC 3D-AD, respectively. The ultra-high accuracy brings more potential and hope to AD tasks but poses significant challenges to establishing representations. CPMF [8] aims to achieve PCD AD by merging handcrafted PCD descriptions with the capabilities of pre-trained 2D neural networks. CPMF performs well on the RGBD dataset, but poorly on Real3D-AD. Reg3D-AD [24] is a benchmark method built on Real3D-AD, which is built on the memory bank and achieves significant performance improvements using pre-trained models. However, the pre-training of Reg3D-AD is difficult to establish directly on HRPCD datasets. IMRNet [21] introduces a self-supervised, scalable framework for 3D PCD AD, leveraging iterative mask reconstruction and geometry-aware sampling to identify and localize anomalies with high accuracy. However, IMRNet has also been affected by the negative performance impact of downsampling.

In view of this, our work aims to design a method that fully utilizes the HRPCD spatial information to enhance the model's generalization ability with small samples in unsupervised manner.

## 3 APPROACH

### 3.1 Problem Definition

Our approach to 3D-AD aligns with the settings defined by Real3D-AD [24], emphasizing the intricacies of handling HRPCD. Defined formally, our task involves a set of training examples  $T = \{t_i\}_{i=1}^N$ , where each  $t_i$  is a normal point clouds sample, belonging to a specific category  $c_j$ , with  $c_j \in C$  where  $C$  represents the entire set

of categories. In testing, determine whether a sample in a certain category contains anomalies, and if so, define the entire sample as an anomaly and locate the anomaly area. The HR 3D-AD dataset is unique in that it exclusively comprises point clouds, offering a more nuanced and comprehensive representation of objects compared to traditional RGB-D (treated as 2.5D) datasets, which may not meet the precision requirements of industrial manufacturing due to potential blind spots from single-view scanning.

### 3.2 Group3AD

This section delves into the detailed architecture of Group3AD, as shown in Figure 2, a model specifically designed to enhance the resolution and accuracy of 3D anomaly detection through group-level feature contrastive learning. Group3AD is composed of three principal components: Intercluster Uniformity Network (IUN), Intracluster Alignment Network (IAN), and Adaptive Group-Center Selection (AGCS). These modules work in synergy to optimize feature representation, thereby improving the precision and efficiency of anomaly detection.

**3.2.1 Intercluster Uniformity Network (IUN).** The IUN establishes an intercluster contrastive learning task. Each cluster is composed of group-level feature vectors with similarity. This network sets each cluster as negative samples in contrastive learning, and widens the distance between each cluster to obtain a uniform structured sample space.

We construct fake anomalies on the original HR point clouds so that the encoder can still maintain uniformity when encoding anomaly groups during inference. Given the original point clouds  $P$ ,  $P_i$  represents the  $i$ th point in  $P$ . Firstly, calculate the PPFH features of the point clouds:

$$FPFH(P_i) = \text{ComputeFPFH}(P_i, \text{neighbors}(P_i)). \quad (1)$$

Next, select the point  $c$  with the highest PPFH feature as the center of the local region:

$$c = \text{argmax}_{P_i} FPFH(P_i). \quad (2)$$

Then, based on the center point  $c$ , a random  $r \in [1\%, 10\%]$  of the total point clouds will be selected around  $c$  as the local area:

$$P_{local} = \{P_j \mid \text{distance}(P_j, c) \leq d, j = 1, \dots, N\}, \quad (3)$$

where  $d$  is a distance threshold that ensures  $P_{local}$  contains approximately  $r$  of the points in  $P$  are included. Subsequently, randomly select a value with a variance  $\sigma^2 \in [0.01, 0.05]$  and use  $\sigma^2$  to generate a normally distributed random noise  $\epsilon$  with an average value of 0. Add noise to each dimension of  $P$  to generate anomaly regions  $P_{anomaly}$ :

$$P_{anomaly,i} = P_{local,i} + \epsilon, \quad (4)$$

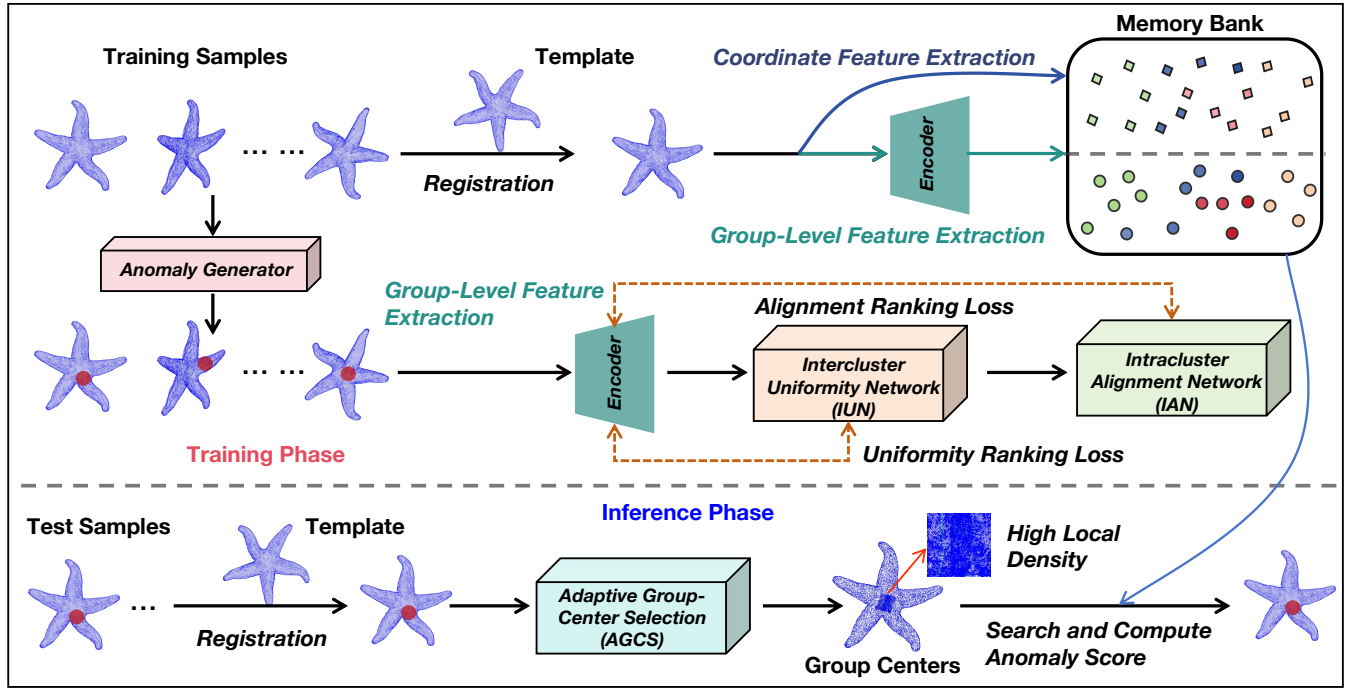
where  $\epsilon = \mathcal{N}(0, \sigma^2)$ . Repeat this process for all  $P$  midpoints to obtain the set of anomalies  $A$ :

$$A = P_{anomaly,i} \mid \forall P_{local,i} \in P_{local}. \quad (5)$$

Combine  $A$  and  $P$  to obtain *FakeAnomalyP*:

$$\text{FakeAnomaly}(P) = P \cup A. \quad (6)$$

We cluster group-level features using the K-means [28] algorithm to determine which cluster ( $P_i$ ) belongs to. K-means determines the number of clusters using the Elbow Method [35]. Determine



**Figure 2: The pipeline of Group3AD, which consists of three main components. (1) Group-Level Feature Extraction extracts group-level features from the input 3D point clouds. (2) Intercluster Uniformity Network (IUN) and Intracluster Alignment Network (IAN) enhance the feature separation between clusters and tighten the distribution within clusters, respectively, for improving anomaly detection accuracy. (3) Adaptive Group-Center Selection (AGCS), used during inference, dynamically focuses on regions with potential anomalies by adjusting the sampling density based on geometric information. This structured approach ensures efficient anomaly detection in complex 3D environments.**

the final size of  $K$  by plotting the relationship between Within Cluster Sum of Squares (WCSS) and the number of clusters ( $K$ ). Find the "Elbow" point in the relationship diagram, which is the position where the rate of WCSS decrease suddenly slows down as  $k$  increases. This point is usually considered the optimal number of clusters, as it represents the best balance between the intracluster compactness and the number of clusters.

The calculation of WCSS is represented as:

$$W_k = \sum_{x_i \in C_k} \|x_i - \mu_k\|_2^2, \quad (7)$$

where  $\mu_k$  is the  $k$ th cluster center,  $C_k$  is the point set in the  $k$ th cluster,  $W_k$  is the WCSS of the  $k$ th cluster, and  $\|x_i - \mu_k\|_2^2$  is the square of the Euclidean distance from point  $x_i$  to the cluster center  $\mu_k$ . The entire dataset  $W$  is the sum of all clustered WCSS:

$$W = \sum_{k=1}^K W_k. \quad (8)$$

Assuming we have  $k$  cluster centers, denoted as  $C = \{C_1, C_2, \dots, C_k\}$ , where each cluster center  $C_i$  is an  $n$ -dimensional vector. For any two different cluster centers  $C_i$  and  $C_j$ , we calculate the Euclidean distance between them:

$$d(C_i, C_j) = \|C_i - C_j\|_2, \quad (9)$$

where  $\|\cdot\|$  represents calculating Euclidean distance. We compare the distances between each pair of cluster centers and select the minimum value among these distances:

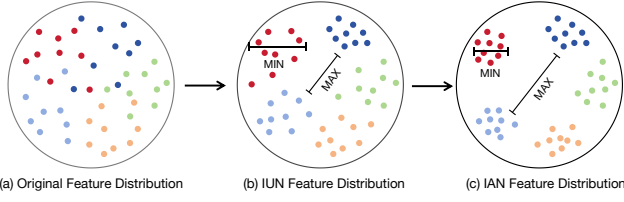
$$\min\_dist = \min_{1 \leq i < j \leq K} \text{distance}(C_i, C_j). \quad (10)$$

Define the uniformity ranking loss function as the reciprocal of this minimum distance:

$$L_{\text{uniformity}} = \frac{1}{\min\_dist}. \quad (11)$$

The larger the value of the uniformity ranking loss function, the smaller the minimum distance between cluster centers, the lower the discrimination between clusters, and the greater the loss. Therefore, our goal is to minimize this loss function, thereby maximizing the minimum distance between cluster centers and improving the quality of clustering. The changes in the distribution of vectors in the feature space represented by the optimized encoder through the uniformity ranking loss are shown in Figure 3(b).

**3.2.2 Intracluster Alignment Network (IAN).** The IAN establishes an intracluster contrastive learning task. Each cluster constitutes one mini-batch. Each vector in the cluster constitutes one positive sample in contrastive learning. Assuming we have  $k$  cluster, denoted as  $C = \{C_1, C_2, \dots, C_k\}$ , where each cluster  $C_k$  contains a certain number of points. We use  $P_k = \{P_{k1}, P_{k2}, \dots, P_{kn}\}$  to represent the



**Figure 3: Flowchat of group-level feature distribution, constrained by IUN and IAN. The basic idea is to minimize the intra-group distance and maximize the inter-group distance.**

set of points  $C_k$  contains. For any two points  $P_{ki}$  and  $P_{kj}$  in cluster  $C_k$ , we calculate the Euclidean distance between them:

$$d(P_{ki}, P_{kj}) = \|P_{ki} - P_{kj}\|, \quad (12)$$

where  $\|\cdot\|$  represents calculating Euclidean distance. For each cluster  $C_k$ , we need to find the maximum distance between all pairs of internal points:

$$\max\_dist_k = \max_{P_{ki}, P_{kj} \in C_k} d(P_{ki}, P_{kj}), \quad (13)$$

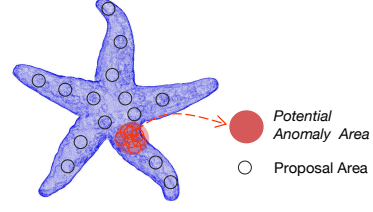
This means that we compare the distances between all possible point pairs within cluster  $C_k$  and select the maximum value among these distances. Finally, we define the alignment ranking loss function as the average of the maximum intracluster distance of all clusters:

$$L_{\text{alignment}} = \frac{T}{k} \sum_{k=1}^K \max\_dist_k^2, \quad (14)$$

where  $T$  is a temperature coefficient that makes the training of the network more stable. A larger intracluster distance usually means that the points within the cluster are more dispersed. The larger the value of this loss function, the more scattered the points within the cluster, the lower the compactness of the cluster, and the greater the loss. Therefore, our goal is to minimize this loss function and improve the compactness of clustering.

**3.2.3 Adaptive Group-Center Selection (AGCS).** Although our method does not require sampling, we still need to use the Farthest Point Sampling (FPS) algorithm to select the center points of the groups. Obtaining more groups in areas with potential anomalies can enable the model to inject more computing power into these areas to achieve stronger perceptual capabilities. In operation, our goal is to allocate higher groups resolution in areas with significant geometric changes, inspiring attention allocation mechanisms for quality inspection engineers in their work.

To evaluate the geometric characteristics of points, we employ the Fast Point Feature Histogram (FPFH) approach. The FPFH descriptor is utilized to analyze variations in surface geometry by comparing a point with the estimated surface normals of its neighboring points. For each point  $p$  within the PCD  $P$ , We calculate the surface normal of this point, denoted as  $n(\cdot)$ . We take  $n$  neighboring points around point  $p$ , denoted as  $\{N_1, N_2, \dots, N_n\}$ . We calculate the surface normal difference between point  $p$  and its  $n$  neighboring points through Euclidean distance, and obtain the surrounding



**Figure 4: Adaptive group-center selection (AGCS). AGCS adaptively selects the points most likely to be within the anomaly region in the selection of group centers via PPFH features.**

variation feature  $F$ . We have:

$$F_p = \sum_{i=1}^N \|f(p) - f(N_i)\|_2. \quad (15)$$

This metric enables us to quantify the dissimilarity in geometric attributes between a point and its neighboring points based on their PPFH descriptors. Such analysis provides crucial insights into the local geometry of the point clouds. In the sampling process of center points, it is assumed that a total of  $on$  center points are required. The center points  $C_{FPFH}$  sampled by the PPFH method are  $a \times n$ , and the remaining  $(1 - a) \times n$  center points  $C_{FPS}$  are obtained through FPS sampling, where  $a$  represents the assigned weight. Ultimately, all center points  $C$  are composed of both:

$$C = C_{FPFH} \cup C_{FPS}. \quad (16)$$

The AGCS is implemented as detailed in Algorithm 1.

---

**Algorithm 1** Adaptive Group-Center Selection (AGCS)

---

- 1: **Input:** Point clouds  $\mathbf{P}$ , number of points  $N$ , attention factor  $\alpha$
  - 2: **Output:** Set of center points  $C_{\text{points}}$
  - 3: **function** AGCS( $\mathbf{P}, N, \alpha$ )
  - 4:    $N \leftarrow \text{ComputeNormals}(\mathbf{P})$   $\triangleright$  Compute surface normals for each point in the point clouds  $\mathbf{P}$
  - 5:    $F_{\text{FPFH}} \leftarrow \text{ComputeFPFH}(\mathbf{P}, N)$   $\triangleright$  Compute PPFH features for the point clouds based on the surface normals
  - 6:    $F_{\text{var}} \leftarrow \text{ComputeLocalVariation}(F_{\text{FPFH}})$   $\triangleright$  Evaluate local geometric variation, indicative of potential anomalies
  - 7:    $P_{\text{high\_var}} \leftarrow \text{SelectHighVariationPoints}(F_{\text{var}}, \alpha)$   $\triangleright$  Select points with high variation in their local geometric features, which are more likely to be near anomalies
  - 8:    $C_{\text{FPFH}} \leftarrow \text{FPS}(P_{\text{high\_var}}, N \times \alpha)$   $\triangleright$  Sample using FPS on high variation points
  - 9:    $C_{\text{FPS}} \leftarrow \text{FPS}(\mathbf{P}, N \times (1 - \alpha))$   $\triangleright$  Sample remaining points uniformly
  - 10:    $C_{\text{points}} \leftarrow C_{\text{FPFH}} \cup C_{\text{FPS}}$
  - 11:   **return**  $C_{\text{points}}$
  - 12: **end function**
- 

**3.2.4 Overall Pipeline.** In the Group3AD framework, the **training phase** starts with a pre-trained encoder to extract features from 3D point clouds initially. The Intercluster Uniformity Network (IUN) subsequently enhances the encoder's capabilities, which increases the separation between different clusters to improve the distinctness

of feature representations. Additionally, the Intracluster Alignment Network (IAN) further refines these features by ensuring they are tightly aligned within each cluster. The enhanced encoder then populates a memory bank with these optimized features. During the **inference phase**, the Adaptive Group-Center Selection (AGCS) assists by selectively targeting potential anomaly regions, comparing these regions against the memory bank to accurately compute anomaly scores.

The training and inference implementation of Group3AD are described in Algorithm 2. We denote  $\mathcal{P}$  as a set of point clouds from the training or testing dataloader, and use  $\mathbf{P}_{\text{mod}}$  for the modified point clouds after generating fake anomalies,  $\mathbf{F}$  for features extracted by the encoder  $\Phi_{\text{enc}}$ .  $\mathcal{C}$  and  $\mathcal{C}_{\text{centers}}$  are for clusters and their centers, respectively.  $\mathcal{L}_{\text{IUN}}$  and  $\mathcal{L}_{\text{IAN}}$  are for the loss functions of the IUN and IAN phases.  $\mathcal{C}_{\text{points}}$  is for the center points selected by AGCS.  $\mathbf{S}$  is for the computed anomaly scores during inference.

---

**Algorithm 2** Group3AD with Two-Phase Training
 

---

```

1: Input: Training dataloader  $\mathcal{D}_{\text{train}}$ , Testing dataloader  $\mathcal{D}_{\text{test}}$ ,
   epochs  $E$ 
2: Output: Trained Encoder  $\Phi_{\text{enc}}$ , IUN  $\Phi_{\text{IUN}}$ , IAN  $\Phi_{\text{IAN}}$ , and
   AGCS settings
3: Initialize  $\Phi_{\text{enc}}$ ,  $\Phi_{\text{IUN}}$ , and  $\Phi_{\text{IAN}}$  randomly
4: /*Phase 1: Training IUN*/
5: for  $i = 1$  to  $E/2$  do
6:   for Pointclouds  $\mathbf{P}$  from  $\mathcal{D}_{\text{train}}$  do
7:      $\mathbf{P}_{\text{mod}} \leftarrow \text{generate\_fake\_anomalies}(\mathbf{P})$ 
8:      $\mathbf{F} \leftarrow \Phi_{\text{enc}}(\mathbf{P}_{\text{mod}})$ 
9:      $\mathcal{C}, \mathcal{C}_{\text{centers}} \leftarrow \text{cluster\_features}(\mathbf{F})$ 
10:     $\mathcal{L}_{\text{IUN}} \leftarrow \text{compute\_IUN\_loss}(\mathcal{C}_{\text{centers}})$ 
11:    Perform backpropagation on  $\mathcal{L}_{\text{IUN}}$ 
12:   end for
13: end for
14: /*Phase 2: Training IAN*/
15: for  $i = E/2 + 1$  to  $E$  do
16:   for Pointclouds  $\mathbf{P}$  from  $\mathcal{D}_{\text{train}}$  do
17:      $\mathbf{F} \leftarrow \Phi_{\text{enc}}(\mathbf{P})$ 
18:      $\mathcal{C} \leftarrow \text{cluster\_features}(\mathbf{F})$   $\triangleright$  Assuming the function also
   returns clusters  $\mathcal{C}$ 
19:      $\mathcal{L}_{\text{IAN}} \leftarrow \text{compute\_IAN\_loss}(\mathcal{C})$ 
20:     Perform backpropagation on  $\mathcal{L}_{\text{IAN}}$ 
21:   end for
22: end for
23: /*Inference Phase*/
24: for Pointclouds  $\mathbf{P}$  from  $\mathcal{D}_{\text{test}}$  do
25:    $\mathcal{C}_{\text{points}} \leftarrow \text{AGCS.select\_center\_points}(\mathbf{P})$ 
26:    $\mathbf{F} \leftarrow \Phi_{\text{enc}}(\mathbf{P})$ 
27:    $\mathbf{S} \leftarrow \text{compute\_anomaly\_scores}(\mathbf{F}, \mathcal{C}_{\text{points}})$   $\triangleright$  Further
   processing based on  $\mathbf{S}$ 
28: end for

```

---

## 4 EXPERIMENTS

### 4.1 Experimental Settings

**4.1.1 Datasets.** In our experiments, we utilize the Real3D-AD [4] dataset. Real3D-AD is the representative HR multi-view scanning

3D dataset derived from real-world scenarios. Real3D-AD, comprising 1,254 HR 3D items with each ranging from forty thousand to millions of points, stands as the most extensive dataset for high-precision 3D industrial anomaly detection. Real3D-AD exceeds other available datasets for 3D anomaly detection in terms of point clouds resolution (0.0010mm-0.0015mm), comprehensive 360-degree coverage, and flawless prototype quality.

**4.1.2 Evaluation Metrics.** Our study standardizes the evaluation by employing metrics specifically designed for 3D-AD, which are identical to those detailed in Real3D-AD. We assess the performance of anomaly detection at both the object and point levels using the Area Under the Receiver Operating Characteristic Curve (AUROC) and the Area Under the Precision-Recall Curve (AUPR/AP). Superior anomaly detection capabilities are indicated by higher values of AUROC and AUPR. All tests are performed using a hardware setup comprising a 13th Gen Intel(R) Core(TM) i7-13700K CPU, 32GB DDR5 SDRAM, and a GeForce RTX 3090 graphics card.

## 4.2 Results and Analysis

**4.2.1 Anomaly Detection on Real3D-AD.** Tables 1-4 summarize per-class comparisons between Group3AD and other state-of-the-art methods, namely Reg3D-AD [24], CPMF [8], IMRNet [21], and several benchmarking methods reported in Real3D-AD [24].

(1) **O-AUROC.** Regarding the O-AUROC metric, while Reg3D-AD [24] sets the benchmark among current approaches with an average O-AUROC of 0.704, this mark falls short of being entirely effective. The introduced Group3AD method, however, markedly surpasses these existing standards, registering an impressive O-AUROC of 0.751, illustrating a significant advancement over prior techniques, as shown in Table 1.

(2) **O-AUPR.** In the context of O-AUPR, Reg3D-AD leads among contemporary strategies with an average performance of 0.723, indicating a gap towards optimal precision-recall balance. The newly developed Group3AD method exceeds these precedents, demonstrating a notable O-AUPR of 0.74, showcasing a substantial improvement in accurately identifying anomalies, as shown in Table 2.

(3) **P-AUROC.** Our findings indicate that Group3AD exhibits a marked improvement over Reg3D-AD in the context of P-AUROC scores across various test scenarios. While Reg3D-AD achieved a P-AUROC of 0.705, Group3AD advanced this benchmark to 0.735, indicating a significant enhancement in detection capability, as shown in Table 3. This improvement demonstrates the robustness of Group3AD in navigating the complexities of 3D anomaly detection.

(4) **P-AURR.** Reflecting on the P-AURR metric, we focus on how Group3AD compares to the established benchmark method, Reg3D-AD. Reg3D-AD sets a solid foundation with a P-AURR score of 0.109. Our novel approach, Group3AD, significantly enhances this benchmark by achieving a P-AURR of 0.137, indicating a marked improvement in recall capabilities, as shown in Table 4.

In reviewing the performance across O-AUROC, O-AUPR, P-AUROC, and P-AURR metrics, it is clear that while Reg3D-AD [24] has set solid benchmarks, Group3AD significantly surpasses these, demonstrating exceptional improvements (performance improvements of 6.7%, 2.4%, 4.3%, and 25.7% respectively). This performance solidifies Group3AD's status as a powerful tool for precise and reliable anomaly detection in industrial settings.

**Table 1: O-AUROC score ( $\uparrow$ ) on Real3D-AD. The best and second-best results are marked in red and blue, respectively.**

Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
BTF(Raw)	0.520	0.560	0.462	0.432	0.545	<b>0.784</b>	0.549	0.648	0.779	<b>0.754</b>	<b>0.575</b>	0.630	0.603
BTF(FPFH)	0.730	0.647	0.703	0.789	0.707	<b>0.691</b>	0.602	<b>0.686</b>	0.596	0.396	0.530	0.539	0.635
M3DM(PointBERT)	0.407	0.506	0.442	0.673	0.627	0.466	0.556	0.617	0.494	0.577	0.528	0.562	0.538
M3DM(PointMAE)	0.434	0.541	0.450	0.683	0.602	0.433	0.540	0.644	0.495	<b>0.694</b>	0.551	0.552	0.552
PatchCore(FPFH)	<b>0.882</b>	0.590	0.565	0.837	0.574	0.546	0.675	0.370	0.505	0.589	0.441	0.541	0.593
PatchCore(FPFH+Raw)	<b>0.848</b>	<b>0.777</b>	0.626	<b>0.853</b>	0.784	0.628	0.837	0.359	0.767	0.663	0.471	0.570	0.682
PatchCore(PointMAE)	0.726	0.498	0.585	0.827	0.783	0.489	0.630	0.374	0.539	0.501	0.519	0.663	0.594
CPMF	0.632	0.518	0.718	0.640	0.640	0.554	0.840	0.349	<b>0.843</b>	0.393	0.526	<b>0.845</b>	0.625
IMRNet	0.762	0.711	0.755	0.780	<b>0.905</b>	0.517	0.880	<b>0.674</b>	0.604	0.665	<b>0.674</b>	0.774	<b>0.725</b>
<b>Reg3D-AD</b>	0.716	0.697	<b>0.827</b>	<b>0.852</b>	0.900	0.584	<b>0.915</b>	0.417	0.762	0.583	0.506	0.685	0.704
<b>Group3AD(Ours)</b>	0.744	<b>0.728</b>	<b>0.847</b>	0.786	<b>0.932</b>	0.679	<b>0.976</b>	0.539	<b>0.841</b>	0.585	0.562	<b>0.796</b>	<b>0.751</b>

**Table 2: O-AUPR score ( $\uparrow$ ) on Real3D-AD. The best and second-best results are marked in red and blue, respectively.**

Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
BTF(Raw)	0.506	0.523	0.490	0.464	0.535	<b>0.760</b>	0.633	0.598	<b>0.793</b>	<b>0.751</b>	<b>0.579</b>	0.700	0.611
BTF(FPFH)	0.659	0.653	0.638	0.814	0.677	<b>0.620</b>	0.638	0.603	0.567	0.434	0.557	0.505	0.614
M3DM(PointBERT)	0.497	0.517	0.480	0.716	0.661	0.569	0.628	<b>0.628</b>	0.491	0.638	<b>0.573</b>	0.569	0.581
M3DM(PointMAE)	0.479	0.508	0.498	0.739	0.620	0.533	0.525	<b>0.663</b>	0.518	0.616	<b>0.573</b>	0.593	0.572
PatchCore(FPFH)	<b>0.852</b>	0.611	0.553	0.872	0.569	0.506	0.642	0.411	0.508	0.573	0.491	0.506	0.591
PatchCore(FPFH+Raw)	<b>0.807</b>	<b>0.766</b>	0.611	<b>0.885</b>	0.767	0.560	0.844	0.411	0.763	0.553	0.473	0.559	0.667
PatchCore(PointMAE)	0.747	0.555	0.576	0.864	0.801	0.488	0.720	0.444	0.546	0.590	0.561	0.708	0.633
<b>Reg3D-AD</b>	0.703	<b>0.753</b>	<b>0.824</b>	<b>0.884</b>	<b>0.884</b>	0.588	<b>0.939</b>	0.454	0.787	0.646	0.491	<b>0.721</b>	<b>0.723</b>
<b>Group3AD(Ours)</b>	0.757	0.706	<b>0.837</b>	0.674	<b>0.932</b>	0.612	<b>0.981</b>	0.533	<b>0.842</b>	<b>0.648</b>	0.567	<b>0.785</b>	<b>0.740</b>

**Table 3: P-AUROC score ( $\uparrow$ ) on Real3D-AD. The best and second-best results are marked in red and blue, respectively.**

Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
BTF(Raw)	0.564	0.647	0.735	0.608	0.563	0.601	0.514	0.597	0.520	0.489	0.392	0.623	0.571
BTF(FPFH)	<b>0.738</b>	0.708	<b>0.864</b>	0.693	<b>0.882</b>	<b>0.875</b>	0.709	<b>0.891</b>	0.512	0.571	0.501	<b>0.815</b>	0.730
M3DM(PointBERT)	0.523	0.593	0.682	<b>0.790</b>	0.594	0.668	0.589	0.646	0.574	0.732	0.563	0.677	0.636
M3DM(PointMAE)	0.530	0.607	0.683	0.735	0.618	0.678	0.600	0.654	0.561	0.748	0.555	0.679	0.637
PatchCore(FPFH)	0.471	0.643	0.637	0.618	0.760	0.430	0.464	<b>0.830</b>	0.544	0.596	0.522	0.411	0.577
PatchCore(FPFH+Raw)	0.556	0.740	<b>0.749</b>	0.558	0.854	0.658	0.781	0.539	0.808	0.753	0.613	0.549	0.680
PatchCore(PointMAE)	0.579	0.610	0.635	0.683	0.776	0.439	0.714	0.514	0.660	0.725	<b>0.641</b>	0.727	0.642
CPMF	0.618	<b>0.836</b>	0.734	0.559	0.753	<b>0.719</b>	<b>0.988</b>	0.449	<b>0.962</b>	0.725	<b>0.800</b>	<b>0.959</b>	<b>0.758</b>
<b>Reg3D-AD</b>	0.631	0.718	0.724	0.676	0.835	0.503	0.826	0.545	0.817	<b>0.811</b>	0.617	0.759	0.705
<b>Group3AD(Ours)</b>	<b>0.636</b>	<b>0.745</b>	0.738	<b>0.759</b>	<b>0.862</b>	0.631	<b>0.836</b>	0.564	<b>0.827</b>	<b>0.798</b>	0.625	0.803	<b>0.735</b>

**Table 4: P-AUPR score ( $\uparrow$ ) on Real3D-AD. The best and second-best results are marked in red and blue, respectively.**

Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
BTF(Raw)	0.012	0.014	0.025	0.049	0.032	0.020	0.017	0.014	0.031	0.011	0.017	0.016	0.022
BTF(FPFH)	<b>0.027</b>	0.028	0.118	0.044	0.239	<b>0.068</b>	0.036	<b>0.075</b>	0.027	0.018	0.034	0.055	0.064
M3DM(PointBERT)	0.007	0.017	0.016	<b>0.377</b>	0.038	0.011	0.039	0.017	0.028	0.021	0.040	0.018	0.052
M3DM(PointMAE)	0.007	0.018	0.016	<b>0.310</b>	0.033	0.011	0.025	0.018	0.030	0.022	0.040	0.021	0.046
PatchCore(FPFH)	<b>0.027</b>	0.034	<b>0.142</b>	0.040	0.273	<b>0.055</b>	0.052	<b>0.093</b>	0.031	0.031	0.037	0.040	0.071
PatchCore(FPFH+Raw)	0.016	<b>0.160</b>	0.092	0.045	<b>0.363</b>	0.034	0.266	0.066	<b>0.291</b>	0.049	0.035	0.055	0.123
PatchCore(PointMAE)	0.016	0.069	0.020	0.052	0.107	0.008	0.201	0.008	0.071	0.043	0.046	0.055	0.058
CPMF	0.010	0.064	0.050	0.031	0.074	0.018	<b>0.559</b>	0.007	<b>0.636</b>	0.025	<b>0.128</b>	<b>0.391</b>	<b>0.166</b>
<b>Reg3D-AD</b>	0.017	0.135	0.109	0.044	0.191	0.010	0.437	0.016	0.182	<b>0.065</b>	0.039	0.067	0.109
<b>Group3AD(Ours)</b>	<b>0.018</b>	<b>0.174</b>	<b>0.122</b>	0.068	<b>0.287</b>	0.016	<b>0.448</b>	0.009	0.24	<b>0.067</b>	<b>0.056</b>	<b>0.134</b>	<b>0.137</b>

4.2.2 *Evaluating Intercluster Uniformity & Intracluster Alignment Method for Anomaly Detection.* Our experimental analysis prominently features the evaluation of the Intercluster Uniformity Network (IUN) and Intracluster Alignment Network (IAN) through a series of ablation studies. By comparing the performance metrics of models with the IUN&IAN component both enabled and

disabled, From Table 5, we observed a significant enhancement in the model’s ability to discriminate between normal and anomalous features when the IUN&IAN was active. This enhancement directly correlates with IUN&IAN’s primary function: to optimize the separation and uniform distribution of feature clusters in the

**Table 5: Ablation studies on the INU & IAN Methods. The best and second-best results are marked in red and blue, respectively.**

Metric	Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
O-AUROC (↑)	PatchCore(PointMAE)	0.726	0.498	0.585	0.827	0.783	0.489	0.630	0.374	0.539	0.501	0.519	0.663	0.594
	PatchCore(PointMAE) with IUN&IAN	0.748	0.558	0.656	0.655	0.845	0.579	0.741	0.478	0.556	0.522	0.53	0.662	0.628
	Group3AD without IUN&IAN	0.721	0.712	0.842	0.773	0.897	0.656	0.965	0.511	0.838	0.536	0.553	0.775	0.732
	<b>Group3AD</b>	0.744	0.728	0.847	0.786	0.932	0.679	0.976	0.539	0.841	0.585	0.562	0.796	0.751
P-AUPR (↑)	PatchCore(PointMAE)	0.016	0.069	0.020	0.052	0.107	0.008	0.201	0.008	0.071	0.043	0.046	0.055	0.058
	PatchCore(PointMAE) with IUN&IAN	0.013	0.066	0.021	0.047	0.118	0.010	0.210	0.009	0.081	0.037	0.063	0.12	0.066
	Group3AD without IUN&IAN	0.011	0.158	0.100	0.051	0.210	0.014	0.420	0.010	0.240	0.037	0.033	0.101	0.115
	<b>Group3AD</b>	0.018	0.174	0.122	0.068	0.287	0.016	0.448	0.009	0.240	0.067	0.056	0.134	0.137

**Table 6: Ablation studies on the AGCS scheme. The best and second-best results are marked in red and blue, respectively.**

Metric	Method	Airplane	Car	Candybar	Chicken	Diamond	Duck	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Average
O-AUROC (↑)	PatchCore(PointMAE)	0.726	0.498	0.585	0.827	0.783	0.489	0.630	0.374	0.539	0.501	0.519	0.663	0.594
	PatchCore(PointMAE) with AGCS	0.713	0.544	0.614	0.676	0.820	0.544	0.752	0.452	0.590	0.507	0.524	0.778	0.626
	Group3AD without AGCS	0.723	0.708	0.833	0.726	0.906	0.661	0.941	0.523	0.818	0.544	0.531	0.787	0.725
	<b>Group3AD</b>	0.744	0.728	0.847	0.786	0.932	0.679	0.976	0.539	0.841	0.585	0.562	0.796	0.751
P-AUPR (↑)	PatchCore(PointMAE)	0.016	0.069	0.020	0.052	0.107	0.008	0.201	0.008	0.071	0.043	0.046	0.055	0.058
	PatchCore(PointMAE) with AGCS	0.010	0.057	0.022	0.055	0.145	0.013	0.261	0.009	0.092	0.040	0.039	0.137	0.073
	Group3AD without AGCS	0.019	0.176	0.067	0.058	0.253	0.015	0.437	0.008	0.237	0.048	0.042	0.124	0.124
	<b>Group3AD</b>	0.018	0.174	0.122	0.068	0.287	0.016	0.448	0.009	0.240	0.067	0.056	0.134	0.137

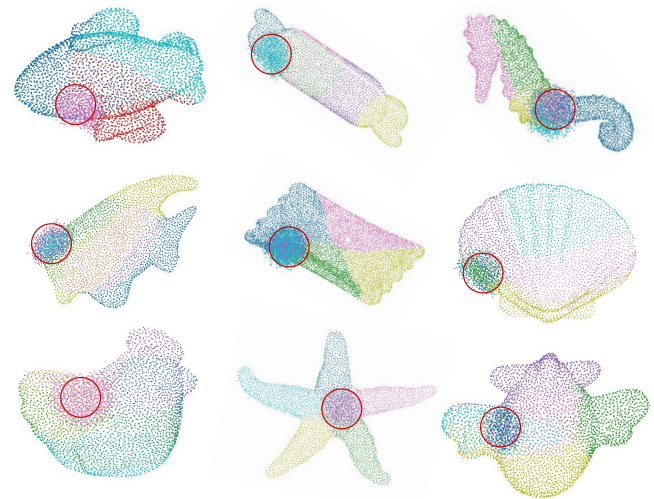
feature space. While the IUN&IAN establishes a foundational separation and distribution of feature clusters, the IAN gently fine-tunes this landscape, promoting tighter and more cohesive clusters. Such optimization evidently aids in reducing ambiguity and overlap between clusters, thereby sharpening the model’s anomaly detection capabilities. The empirical evidence from our ablation studies thus underscores the essential role of IUN&IAN in fortifying the model’s performance, highlighting its value in the context of HR 3D-AD.

**4.2.3 Evaluating Adaptive Group-Center Selection for Anomaly Detection.** In our comprehensive analysis, we further examined the role of the Adaptive Group-Center Selection (AGCS) within our proposed Group3AD. The AGCS enhances the detection process by strategically directing the model’s focus towards regions potentially harboring anomalies, thus adjusting the sampling density in critical areas. Experimental results shown in Table 6 indicate that AGCS substantially improves the model’s effectiveness in detecting anomalies. By concentrating on areas with potential irregularities, AGCS allows the model to identify subtle yet critical anomalies that might be missed under uniform sampling conditions. This prioritization leads to a more efficient detection process, enabling the model to perform more accurately and swiftly, especially in complex point cloud datasets.

### 4.3 Visualization

Figure 5 provides a visual representation of the clustering results achieved with the Group3AD framework on the Real3D-AD dataset, focusing on the effectiveness of the Intercluster Uniformity Network (IUN) and Intracluster Alignment Network (IAN). Each dot signifies a group center, pinpointed through the refined feature space sculpted by the IUN and IAN. This visualization underscores the success of these networks in creating distinct, well-separated clusters while maintaining tight intra-cluster relationships, which are instrumental for the precise localization of anomalies. The clear separation of clusters showcases the IUN’s impact in enhancing feature discrimination across groups, while the density of points within each cluster attests to the IAN’s role in fostering coherence

among features, resulting in improved anomaly detection performance in high-resolution 3D contexts.



**Figure 5: Visualization results obtained by Group3AD. Different colors indicate different groups selected by AGCS. The red circle represents anomaly area.**

## 5 CONCLUSIONS

In this work, we introduced Group3AD, a robust framework tailored for high-resolution 3D-AD in industry, which can effectively enhance anomaly detection and localization accuracy. Demonstrating significant improvements over existing methodologies, especially in terms of reducing false positives and leveraging depth information for clearer anomaly identification, Group3AD emerges as a practical solution for real-world applications. Future endeavors may extend its applicability and efficiency, further solidifying its utility in industrial anomaly detection tasks.



## REFERENCES

- [1] Mohamed Afham, Isuru Dissanayake, Dinithi Dissanayake, Amaya Dharmasiri, Kanchara Thilakarathna, and Ranga Rodrigo. 2022. Crosspoint: Self-supervised cross-modal contrastive learning for 3d point cloud understanding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9902–9912.
- [2] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. 2019. MVTEC AD—A comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9592–9600.
- [3] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. 2020. Uninformed students: Student-teacher anomaly detection with discriminative latent embeddings. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4183–4192.
- [4] Paul Bergmann, Xin Jin, David Sattlegger, and Carsten Steger. 2021. The mvtec 3d-ad dataset for unsupervised 3d anomaly detection and localization. *arXiv preprint arXiv:2112.09045* (2021).
- [5] Paul Bergmann and David Sattlegger. 2023. Anomaly detection in 3d point clouds using deep geometric descriptors. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2613–2623.
- [6] Luca Bonfiglioli, Marco Toschi, Davide Silvestri, Nicola Fioraio, and Daniele De Gregorio. 2022. The eyecandies dataset for unsupervised multimodal anomaly detection and localization. In *Proceedings of the Asian Conference on Computer Vision*. 3586–3602.
- [7] Xingyu Cai, Jiaji Huang, Yuchen Bian, and Kenneth Church. 2020. Isotropy in the contextual embedding space: Clusters and manifolds. In *International conference on learning representations*.
- [8] Yunkang Cao, Xiaohao Xu, and Weiming Shen. 2023. Complementary pseudo multimodal feature for point cloud anomaly detection. *arXiv preprint arXiv:2303.13194* (2023).
- [9] Ruitao Chen, Guoyang Xie, Jiaqi Liu, Jinbao Wang, Ziqi Luo, Jinfan Wang, and Feng Zheng. 2023. EasyNet: An easy network for 3d industrial anomaly detection. In *Proceedings of the 31st ACM International Conference on Multimedia*. 7038–7046.
- [10] Thomas Defard, Aleksandr Setkov, Angélique Loesch, and Romaric Audigier. 2021. Padim: a patch distribution modeling framework for anomaly detection and localization. In *International Conference on Pattern Recognition*. Springer, 475–489.
- [11] Hanqiu Deng and Xingyu Li. 2022. Anomaly detection via reverse distillation from one-class embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9737–9746.
- [12] Kavin Ethayarajh. 2019. How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings. *arXiv preprint arXiv:1909.00512* (2019).
- [13] Jun Gao, Di He, Xu Tan, Tao Qin, Liwei Wang, and Tie-Yan Liu. 2019. Representation degeneration problem in training natural language generation models. *arXiv preprint arXiv:1907.12009* (2019).
- [14] Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. Simcse: Simple contrastive learning of sentence embeddings. *arXiv preprint arXiv:2104.08821* (2021).
- [15] Denis Gudovskiy, Shun Ishizaka, and Kazuki Kozuka. 2022. Cflow-ad: Real-time unsupervised anomaly detection with localization via conditional normalizing flows. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 98–107.
- [16] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. 2020. Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence* 43, 12 (2020), 4338–4364.
- [17] Eliahu Horowitz and Yedid Hoshen. 2023. Back to the feature: classical 3d features are (almost) all you need for 3d anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2967–2976.
- [18] Xi Jiang, Jianlin Liu, Jinbao Wang, Qiang Nie, Kai Wu, Yong Liu, Chengjie Wang, and Feng Zheng. 2022. Softpatch: Unsupervised anomaly detection with noisy data. *Advances in Neural Information Processing Systems* 35 (2022), 15433–15445.
- [19] Jiarui Lei, Xiaobo Hu, Yue Wang, and Dong Liu. 2023. Pyramidflow: High-resolution defect contrastive localization using pyramid normalizing flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 14143–14152.
- [20] Chun-Liang Li, Kihyuk Sohn, Jinsung Yoon, and Tomas Pfister. 2021. Cutpaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9664–9674.
- [21] Wenqiao Li and Xiaohao Xu. 2023. Towards Scalable 3D Anomaly Detection and Localization: A Benchmark via 3D Anomaly Synthesis and A Self-Supervised Learning Network. *arXiv preprint arXiv:2311.14897* (2023).
- [22] Wujin Li, Jiawei Zhan, Jinbao Wang, Bizhong Xia, Bin-Bin Gao, Jun Liu, Chengjie Wang, and Feng Zheng. 2022. Towards continual adaptation in industrial anomaly detection. In *Proceedings of the 30th ACM International Conference on Multimedia*. 2871–2880.
- [23] Zhenyu Li, Zehui Chen, Ang Li, Liangji Fang, Qinrong Jiang, Xianming Liu, Junjun Jiang, Bolei Zhou, and Hang Zhao. 2022. Simipu: Simple 2d image and 3d point cloud unsupervised pre-training for spatial-aware visual representations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 1500–1508.
- [24] Jiaqi Liu, Guoyang Xie, Ruitao Chen, Xinpeng Li, Jinbao Wang, Yong Liu, Chengjie Wang, and Feng Zheng. 2024. Real3d-ad: A dataset of point cloud anomaly detection. *Advances in Neural Information Processing Systems* 36 (2024).
- [25] Jiaqi Liu, Guoyang Xie, Jinbao Wang, Shangnian Li, Chengjie Wang, Feng Zheng, and Yaochu Jin. 2024. Deep industrial image anomaly detection: A survey. *Machine Intelligence Research* 21, 1 (2024), 104–135.
- [26] Kangcheng Liu, Aoran Xiao, Xiaoqin Zhang, Shijian Lu, and Ling Shao. 2023. Fac: 3d representation learning via foreground aware feature contrast. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9476–9485.
- [27] Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. 2023. Simplenet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 20402–20411.
- [28] James MacQueen et al. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1. Oakland, CA, USA, 281–297.
- [29] Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. 2022. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 14318–14328.
- [30] Marco Rudolph, Bastian Wandt, and Bodo Rosenhahn. 2021. Same same but different: Semi-supervised defect detection with normalizing flows. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 1907–1916.
- [31] Marco Rudolph, Tom Wehrbein, Bodo Rosenhahn, and Bastian Wandt. 2023. Asymmetric student-teacher networks for industrial anomaly detection. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2592–2602.
- [32] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. 2009. Fast point feature histograms (FPFH) for 3D registration. In *2009 IEEE international conference on robotics and automation*. IEEE, 3212–3217.
- [33] Mohammadreza Salehi, Niousha Sadjadi, Soroosh Baselizadeh, Mohammad H Rohban, and Hamid R Rabiee. 2021. Multiresolution knowledge distillation for anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 14902–14912.
- [34] Hannah M Schlüter, Jeremy Tan, Benjamin Hou, and Bernhard Kainz. 2022. Natural synthetic anomalies for self-supervised anomaly detection and localization. In *European Conference on Computer Vision*. Springer, 474–489.
- [35] MA Syakur, B Khusnul Khotimah, EMS Rochman, and Budi Dwi Satoto. 2018. Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering*, Vol. 336. IOP Publishing, 012017.
- [36] Guodong Wang, Shumin Han, Errui Ding, and Di Huang. 2021. Student-teacher feature pyramid matching for anomaly detection. *arXiv preprint arXiv:2103.04257* (2021).
- [37] Lingxiao Wang, Jing Huang, Kevin Huang, Ziniu Hu, Guangtao Wang, and Quanquan Gu. 2019. Improving neural language generation with spectrum control. In *International Conference on Learning Representations*.
- [38] Yue Wang, Jinlong Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, and Chengjie Wang. 2023. Multimodal industrial anomaly detection via hybrid fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8032–8041.
- [39] Guoyang Xie, Jinbao Wang, Jiaqi Liu, Jiayi Lyu, Yong Liu, Chengjie Wang, Feng Zheng, and Yaochu Jin. 2024. Im-ia: Industrial image anomaly detection benchmark in manufacturing. *IEEE Transactions on Cybernetics* (2024).
- [40] Guoyang Xie, Jinbao Wang, Jiaqi Liu, Feng Zheng, and Yaochu Jin. 2023. Pushing the limits of fewshot anomaly detection in industry vision: Graphcore. *arXiv preprint arXiv:2301.12082* (2023).
- [41] Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. 2021. Draem—a discriminatively trained reconstruction embedding for surface anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 8330–8339.
- [42] Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. 2022. Dsr—a dual subspace re-projection network for surface anomaly detection. In *European conference on computer vision*. Springer, 539–554.

929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044