2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

T-NACD: A Tactile-friendly Novel Anomaly Class Discovery Framework for Black Rubber Products

Long Xiao[1,2], Kailin Lyu[1,2], Jianing Zeng[1], Xuexin Liu[1,2], Zhuojun Zou[1], Lin Shu[1], Jie Hao[1,*]

***Abstract*— Novel Anomaly Class Discovery (NACD) has recently gained attention in industrial anomaly detection, aiming not only to recognize defects but also to recognize fine-grained and previously unseen anomaly types. However, existing methods, primarily based on visual images, struggle to handle common but challenging materials such as black rubber products. Such products often exhibit hardness anomalies and visually indistinguishable surface defects during production due to process instability and high material absorbance. To address this, we present T-NACD, the first tactile-friendly framework for NACD, leveraging tactile sensing to detect subtle geometric anomalies that are invisible to vision. We introduce two tactile-specific data augmentation methods and design a lightweight feature adapter to transfer visual pre-trained models to the tactile domain without the need for large-scale training. Moreover, we propose a dual-margin enhanced entropy regularization method to mitigate forgetting of known categories and reduce feature stocking across similar anomaly classes. To support this study, we collect BRD, the first real-world tactile dataset for black rubber anomaly detection, including BRD-A (surface and hardness defects) and BRD-H (diverse hardness anomalies). Experimental results show that T-NACD outperforms visual SOTA methods by 6.9% and generalizes well to other non-industrial tactile detection tasks.**

I. INTRODUCTION

Novel Anomaly Class Discovery (NACD) [1] has recently emerged as a new paradigm for industrial anomaly detection [2], [3]. Unlike traditional approaches that merely localize visual anomaly regions without identifying fine-grained types [4], NACD aims to recognize specific categories of anomalies such as cracks and dents, and even discover entirely unseen classes. While several anomaly clustering methods [5], [6] have shown promising results in visual anomaly detection tasks, they still face substantial challenges when applied to specific materials and defect types. For example, **black rubber products** are particularly challenging for visual inspection due to their low reflectance and matte surface texture [7], [8]. As illustrated in Figure 1(a), conventional visual methods suffer severe performance degradation on such materials. Specifically, these products often exhibit hardness anomalies (e.g., overly soft) due to manufacturing defects, yet such differences are entirely undetectable through visual inspection [9]. Moreover, the uniformly dark and matte surface makes it extremely difficult to identify fine-grained surface defects like cracks, impurities, dents, and rough edges using optical imaging [10], [11].

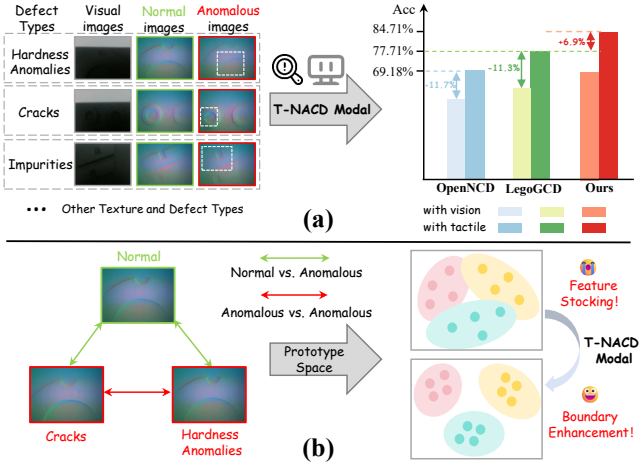


Fig. 1: **(a)** Visual NACD methods perform poorly on anomaly detection tasks for black rubber products, and incorporating tactile images provides only limited improvement. **(b)**The low variance between normal and anomalous samples in black rubber leads to feature space overlap, which is effectively mitigated by T-NACD.

Due to the inherent limitations of visual modalities in handling such cases, we propose the first attempt to detect anomalies in black rubber products using tactile sensing instead of vision. Tactile images, acquired via vision-based tactile sensors [12], [13], record fine-grained surface topology and pressure distribution at the microscale. These sensors offer high geometric sensitivity and structural discriminability without damaging the target object [14], making them especially suitable for detecting texture variations and micro defects that are visually imperceptible [15].

Despite being represented in RGB form, tactile images differ significantly from visual ones and introduce several unique challenges: (1) Compared to visual images, tactile images contain denser and finer-scale geometric deformation information, with less emphasis on color and global context [12], [13]. Existing methods trained solely on visual data often fail to capture the essential features of tactile inputs [16]. (2) The same type of defect can appear differently across black rubber products, while different defect types on the same products may exhibit only subtle local differences in tactile form, leading to significant **feature stocking** between different classes, as shown in Figure 1(b).

* Corresponding to: jie.hao@ia.ac.cn
[1] Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.
[2] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China.
E-mails: {xiaolong2022, lvkailin2024, zengjianing2023, liuxuexin2022, zouzhuojun2018, lin.shu, jie.hao}@ia.ac.cn

Existing methods on fine-grained datasets primarily rely on the rich semantic information in visual images [17], [18]. For example, different bird species may share similar shapes but vary in feather color and organ structures [19]. However, in the case of black rubber products, tactile images from different classes are almost identical in color and across most regions. This makes anomaly detection very challenging.

To address these issues, we propose T-NACD, the first tactile-friendly framework for novel anomaly class discovery. Specifically, we introduce two tactile-specific data augmentation methods and a lightweight feature adapter to bridge the gap between pre-trained visual backbones and tactile inputs, enhancing the model's ability to extract features from tactile images with limited training samples. We further propose a Dual-Margin Enhanced Entropy Regularization Method (DMER) to alleviate feature stocking across similar anomaly classes and mitigate the forgetting of known information. To evaluate our approach, we construct the first real-world tactile dataset **BRD** for black rubber anomaly detection. This includes two subsets: **BRD-A**, covering all surface and hardness-related defects, and **BRD-H**, which provides more diverse hardness anomaly cases. Experimental results demonstrate that our method outperforms state-of-the-art visual baselines on both datasets. **In addition, we demonstrate that T-NACD can serve as a general tactile anomaly detection framework and generalize to other non-industrial tactile datasets for texture detection tasks.**

Our main contributions are summarized as follows:

1) We propose T-NACD, the **first tactile-friendly framework for novel anomaly class discovery**, enabling reliable detection of defects in black rubber products.
2) We present two tactile-specific augmentation methods and a lightweight feature adaptor to transfer visual pre-trained backbone to the tactile domain without large-scale training.
3) We introduce a dual-margin enhanced entropy regularization (DMER) method to mitigate forgetting of known categories and reduce feature stocking between low-variance classes.
4) We build BRD, the **first tactile dataset for black rubber surface and hardness anomaly detection**. Experiments show that T-NACD surpasses visual SOTA methods by 6.9% on BRD, and achieve best performance on two mainstream tactile datasets, demonstrating its **generality**.

II. RELATED WORKS

A. Tactile Sensing in Open World

To address the limitations of visual sensing, tactile perception has been widely applied to tasks such as unknown-object recognition [20] and texture classification [21]. Conventional tactile systems based on tactile array sensors capture only part of the rich interactions that occur during manipulation and therefore often fail to represent the full spectrum of tactile information [20]. In contrast, vision-based tactile sensors provide high-resolution, fine-grained descriptions of object surfaces and currently constitute the dominant approach to tactile sensing [12], [13]. Accordingly, in this work we use the GelSight Mini to acquire tactile images.

B. Industrial Anomaly Detection

Industrial anomaly detection [2], [3] initially focused solely on distinguishing between normal and defective products, without identifying the fine-grained types of anomalies. With the advancement of multi-class anomaly detection [22], several anomaly clustering methods [5], [6] have been proposed to localize anomaly regions and group them based on shared characteristics. However, these methods are inherently limited in their ability to recognize previously unseen anomalies. AnomalyNCD [1] made the first attempt to leverage prior knowledge from known anomaly classes to identify novel ones. Nevertheless, its reliance on labeled anomalous samples hinders its applicability in industrial settings. In this work, we achieve anomaly detection for black rubber products without requiring any labeled anomalous samples.

C. Novel Class Discovery

Novel class discovery (NCD) [23], [24] has been widely applied in fields such as medical image classification [25], [26] and semantic segmentation [27]. It trains models using a small set of labeled data alongside a large amount of unlabeled data, enabling the recognition of previously unseen classes. When the unlabeled dataset contains instances from both known and unknown classes, the task becomes generalized category discovery (GCD) [28], which better reflects real-world scenarios. SimGCD [17] was the first to combine contrastive learning with pseudo-label supervision to enhance the model's ability to recognize unknown categories, but it suffers from severe catastrophic forgetting. LegoGCD [18] alleviates this issue by introducing local entropy regularization, but struggles to distinguish between visually similar classes. In this work, we adopt the GCD setting to better align with industrial anomaly detection scenarios and specifically address the performance degradation of visual methods on black rubber tactile images.

III. METHODOLOGY

In this section, we first present the problem definition of NACD (Section III-A) along with the preliminaries (Section III-B). We then introduce the Tactile-specific Feature Extractor (Section III-C), followed by the Dual-Margin Enhanced Entropy Regularization (DMER) (Section III-D), and the Total Loss (Section III-E). Figure 2 illustrates the proposed T-NACD framework.

A. Problem Definition

We divide the entire dataset into two parts: $\mathcal{D}_l$ and $\mathcal{D}_u$. $\mathcal{D}_l = \{(x_i, y_i)\}_{i=1}^{M} \in \mathcal{X} \times \mathcal{Y}_l$ represents the labeled portion, where $\mathcal{Y}_l$ is the set of classes for the labeled samples. $\mathcal{D}_u = \{(x_i, y_i)\}_{i=1}^{N} \in \mathcal{X} \times \mathcal{Y}_u$ represents the unlabeled portion, where $\mathcal{Y}_u$ is the set of classes for the unlabeled samples. In the NACD setting, the unlabeled samples may contain novel classes that are not present in the labeled

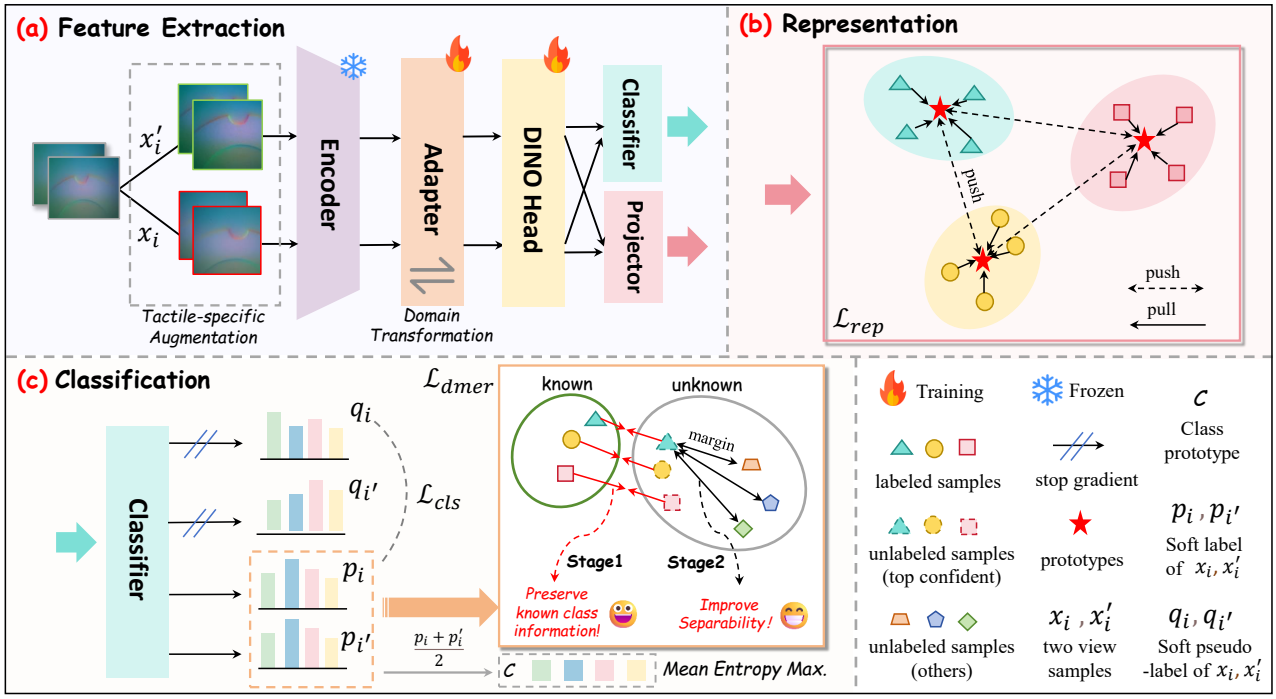


Fig. 2: **Overall pipeline of T-NACD framework**. It employs a Tactile-Specific Feature Extractor to address noisy, domain-misaligned tactile data and transfer representations from the visual to the tactile domain. It further integrates Dual-Margin Enhanced Entropy Regularization (DMER) to stabilize known-class information and reduce feature stocking, enhancing separability and generalization. Together, these components enable generalized novel class discovery in tactile sensing.

samples, meaning $\mathcal{Y}_l \subseteq \mathcal{Y}_u$. The set of novel classes is denoted as $\mathcal{Y}_{\text{novel}} = \mathcal{Y}_u \setminus \mathcal{Y}_l$. In our method, the total number of classes $|\mathcal{Y}_u|$ is treated as prior knowledge, consistent with previous works [17], [18]. More task settings can be found in Section IV-C.

B. Preliminaries

We adopt a parametric clustering framework commonly used in GCD methods [17], [18]. It consists of two modules: *representation learning* and *parametric classification*. In the representation stage, we adopt a ViT-B/16 backbone [29] pretrained with DINO [30] on ImageNet, and perform supervised contrastive learning on labeled data and unsupervised contrastive learning on all data. Formally, given two views (random augmentations) x_i and x_i' of the same image in a mini-batch $\mathcal{B}$, the unsupervised contrastive loss is written as:

$$\mathcal{L}_{\text{rep}}^{u} = \frac{1}{|\mathcal{B}|} \sum_{i \in \mathcal{B}} -\log \frac{\exp\left(z_i^\top \cdot z_i'/\tau_u\right)}{\sum_i^{i \neq n} \exp\left(z_i^\top \cdot z_n'/\tau_u\right)}, \tag{1}$$

where $z_i = g(f(x_i))$, f is the feature backbone, and the projection head g is a multi-layer perceptron (MLP). τ_u is the temperature for the unsupervised loss. The objective of the supervised contrastive loss is to encourage the model to bring samples with the same class label closer in the feature space, formally written as:

$$\mathcal{L}_{\text{rep}}^{s} = \frac{1}{|\mathcal{B}_l|} \sum_{i \in \mathcal{B}_l} \frac{1}{|\mathcal{N}_i|} \sum_{q \in \mathcal{N}_i} -\log \frac{\exp\left(z_i^\top \cdot z_q'/\tau_c\right)}{\sum_{i \neq n} \exp\left(z_i^\top \cdot z_n'/\tau_c\right)}, \tag{2}$$

where $\mathcal{N}_i$ is the set of indices of images that share the same label with x_i in the mini-batch $\mathcal{B}_l$. The overall representation learning loss is balanced with λ: $\mathcal{L}_{\text{rep}} = (1-\lambda)\,\mathcal{L}_{\text{rep}}^{u} + \lambda\,\mathcal{L}_{\text{rep}}^{s}$, where $\mathcal{B}_l$ corresponds to the labelled subset of $\mathcal{B}$.

For classification, we employs a self-distilled prototype classifier with $C = \{c_1, \ldots, c_K\}$, $K = |\mathcal{Y}_l \cup \mathcal{Y}_u|$. For feature $h_i = f(x_i)$, the probability of class k is

$$p_i^{(k)} = \frac{\exp\left(\frac{1}{\tau_s}(h_i/\|h_i\|_2)^\top (c_k/\|c_k\|_2)\right)}{\sum_{k'} \exp\left(\frac{1}{\tau_s}(h_i/\|h_i\|_2)^\top (c_{k'}/\|c_{k'}\|_2)\right)}. \tag{3}$$

Using pseudo labels q_i' for augmented views, the unsupervised and supervised classification losses are

$$L_{\text{cls}}^{u} = \tfrac{1}{|B|} \sum_{i \in B} \ell(q_i', p_i), \quad L_{\text{cls}}^{s} = \tfrac{1}{|B_l|} \sum_{i \in B_l} \ell(y_i, p_i). \tag{4}$$

To stabilize unsupervised learning, we employ the mean-entropy regularizer from SimGCD [17], defined as $H(p) = \sum_k p(k) \log p(k)$, where $p = \frac{1}{2|B|} \sum_{i \in B} (p_i + p_i')$. The classification objective is

$$L_{\text{cls}} = (1-\lambda)(L_{\text{cls}}^{u} - \epsilon H(p)) + \lambda L_{\text{cls}}^{s}. \tag{5}$$

C. Tactile-specific Feature Extractor

Since the visual and tactile domains are not perfectly aligned, this misalignment hinders effective feature extraction from tactile images [16]. In this work, we attempt to bridge the gap between the visual and tactile domains from two complementary perspectives.

Tactile-specific data augmentation. In works such as SimGCD, random augmentations [28], [17] are commonly used to improve the robustness of visual representations. However, in the context of black rubber products defect detection, we are particularly concerned with fine-grained texture patterns and subtle color variations along the image boundaries. To address these challenges, we propose two tactile-specific augmentation strategies: ***i*) Edge color contrast enhancement:** We slightly increase the image contrast and brightness while keeping saturation and hue unchanged. This highlights subtle color variations along the edges of tactile images, making defects more distinguishable. ***ii*) Localized Gaussian noise:** Due to uneven contact forces during acquisition, tactile images may suffer from partial information loss. To simulate this, we divide each image into N patches and randomly add Gaussian noise to selected regions, mimicking local signal degradation or disturbance. This makes the augmented images more representative of real-world scenarios.

We randomly apply one of the two aforementioned augmentation strategies to each of the two views based on a predefined probability threshold $\delta_{aug} = 0.5$, in order to enhance the realism of tactile images.

Lightweight Feature Adapter. Transformer models pretrained with DINO self-supervision are widely used in anomaly detection tasks, but they are not fully adapted to tactile images. Given that tactile data is more costly to acquire than visual data, collecting a large-scale dataset specifically for black rubber products would be prohibitively expensive. To address this, we introduce a lightweight feature adapter G_θ between the backbone and the DINO head, which projects the learned features into the tactile domain.

Specifically, The feature adapter G_θ projects the local feature $o_{h,w}$ to the adapted feature $q_{h,w}$ as

$$q_{h,w} = G_\theta(o_{h,w}), \tag{6}$$

From a theoretical perspective, this process can be formulated as a statistical moment alignment problem, in which the parameter θ is learned by minimizing the discrepancies between the source and target domains in terms of class-wise means and covariances:

$$\min_\theta \sum_k \left(\|\mu_s^{(k)} - \mu_t^{(k)}\|_2^2 \; + \; \|\Sigma_s^{(k)} - \Sigma_t^{(k)}\|_F^2 \right), \tag{7}$$

where $\mu_s^{(k)}$ and $\Sigma_s^{(k)}$ denote the mean and covariance of class k in the source domain, and $\mu_t^{(k)}$ and $\Sigma_t^{(k)}$ denote those in the target domain. This optimization objective enables the feature adaptor to alleviate distributional discrepancies during domain adaptation, thereby enhancing the separability and generalization of the prototype classifier. Notably, the adaptor is constructed with simple neural modules such as fully connected layers or multilayer perceptrons. Experimental results show that only two layers are sufficient to deliver strong performance with minimal computational cost, further facilitating the transfer of our method to other pretrained models.

D. Dual-Margin Enhanced Entropy Regularization

Motivation. In the anomaly detection of black rubber products, only subtle differences exist between normal and anomaly samples, and even among different types of anomalies, the distinctions are minimal [7], [8]. Additionally, tactile images inherently exhibit low semantic information [14], making it even more difficult to differentiate between classes and often leading to feature stocking. Therefore, we argue that for black rubber anomaly detection, it is essential to first preserve the information of known classes, and explicitly define boundaries between known and anomalous categories.

Unlike LegoGCD [18], which only focuses on unlabeled samples belonging to known classes, we propose a **Dual-Margin Enhanced Entropy Regularization (DMER)** method that considers all high-confidence unlabeled samples, and then reinforces dual margins around both the known classes and all classes.

We first construct an initial binary mask $\mathbf{M} = [m_1, m_2, \ldots, m_i] \in \{0, 1\}$ to indicate whether each sample is labeled ($m_i = 1$) or unlabeled ($m_i = 0$).

Then, for each sample x_i, we obtain the predicted class probability vector $\boldsymbol{p}_i = [p_i^1, p_i^2, \ldots, p_i^K]$ over K classes. A high-confidence indicator is then defined as: $s_i = \mathbf{1}\left(\max(p_i) \geq \delta\right)$, where δ is a confidence threshold. And the predicted class label y_i for sample x_i is determined by: $y_i = \arg\max_j \left(p_i\right)_j, \quad i = \{1, 2, \ldots, b\}$.

Finally, we construct another two binary mask $\mathbf{U} = [u_1, u_2, \ldots, u_i] \in \{0, 1\}$ and $\mathbf{O} = [o_1, o_2, \ldots, o_i] \in \{0, 1\}$. $\mathbf{U}$ denotes **all high-confidence samples in the unlabeled dataset**, which may belong to either known or unknown classes. And $\mathbf{O}$ identify **high-confidence unlabeled samples predicted as known categories**. Specifically, a sample x_i is selected if it satisfies the following conditions:

$$u_i = \mathbf{1}(m_i = 0) \cdot \mathbf{1}(s_i = 1), \tag{8}$$

$$o_i = \mathbf{1}(m_i = 0) \cdot \mathbf{1}(s_i = 1) \cdot \mathbf{1}(y_i \in \mathcal{Y}_l), \tag{9}$$

where $\mathcal{Y}_l$ is the set of known class labels.

We divide the process into two stages to separately stabilize the information of known classes and strengthen the boundaries among all classes. First, we use the mask $\mathbf{O}$ to select high-confidence samples from the unlabeled dataset that are likely to belong to known classes, and compute their probability distribution $\boldsymbol{p}_i = \text{softmax}(\boldsymbol{z}_i/\tau_o)$. To prevent this distribution from being biased toward majority classes in the unlabeled dataset (i.e., anomaly classes), we calibrate $\boldsymbol{p}_i$ using the global average distribution to obtain the calibrated probability distribution $\boldsymbol{d}_i = \text{softmax}\left(\left(\mathbf{z}_i + \lambda_\text{d} \log(1/\tilde{\boldsymbol{p}}_i)\right)/\tau_o\right)$. where τ_o is a temperature, λ_d is a calibration coefficient, and $\tilde{\boldsymbol{p}}_i$ represents the average predicted probability for each class. Then, we compute the calibration loss as follows:

$$\mathcal{L}_\text{cal} = -\frac{1}{\mathcal{B}} \sum_{i=1}^{\mathcal{B}} \mathbf{1}(o_i = 1) \cdot \sum_k d_i^{(k)} \log p_i^{(k)}, \tag{10}$$

This loss consolidates the feature distribution of known classes and sharpens the boundary between high-confidence samples likely belonging to known classes and other samples in the unlabeled dataset. However, feature stocking remains for other anomalous samples, as their most probable class lacks sufficient confidence, making class distinction challenging.

Therefore, we design a stronger margin-based loss to uniformly improve the separability between classes. Specifically, for each unlabeled sample, we require the predicted probability score of the most likely class is at least a fixed margin δ_m higher than that of any other class. Otherwise, an additional penalty is applied.

$$\mathcal{L}_{\text{mar}} = -\frac{1}{\mathcal{B}} \sum_{i=1}^{\mathcal{B}} \mathbf{1}(u_i = 1) \cdot \sum_{k \neq y_i} \max(0, \delta_m + p_i^{(k)} - p_i^{(y_i)}), \tag{11}$$

where y_i indicates the index of the class with the highest confidence in each probability distribution. Through this loss, we encourage all unlabeled samples to have high confidence, thereby reducing ambiguity.

The final dual-margin enhanced regularization loss is defined as follows. It effectively preserves known class information while reinforcing the distinction between categories.

$$\mathcal{L}_{\text{dmer}} = \mathcal{L}_{\text{cal}} + \lambda_m \cdot \mathcal{L}_{\text{mar}}, \tag{12}$$

where λ_m denotes the dynamic weight of $\mathcal{L}_{\text{mar}}$.

E. Total Loss

We build a new framework based on SimGCD by integrating the Dual-Margin Enhanced Regularization method. The overall loss function for training the model is defined as:

$$\mathcal{L} = \alpha \cdot (\mathcal{L}_{\text{rep}} + \mathcal{L}_{\text{cls}}) + \beta \cdot \mathcal{L}_{\text{dmer}}, \tag{13}$$

where β affects the model's ability to distinguish between low-variance classes. Aligning with SimGCD, we set α to 1 and simultaneously adjust β, as shown in Figure 7.

IV. EXPERIMENTAL SETUP

A. Dataset

We propose the first dataset for black rubber products, named **BRD**, which consists of two subsets: **BRD-A** and **BRD-H**. **BRD-A** contains visual and tactile images of 8 common rubber textures. For each texture, there is one class of normal samples and 5 classes of anomalous samples, representing *hardness anomalies, cracks, impurities, dents*, and *rough edges*,as shown in Figure 4. Each class contains 50 visual images and 50 tactile images, resulting in a total of 4,800 images. **BRD-H** contains 3 common rubber textures. Each texture includes 5 classes of samples with different Shore hardness levels. Each class contains 50 tactile images, totaling 750 images. We use a robotic arm to collect the data, as shown in Figure 3.

We also utilize the classic tactile dataset **FabricVST** [31] and YCB [32] to evaluate the generalization ability of our model. They consist of 50 and 10 categories, making it well-suited for the novel anomaly class discovery task.

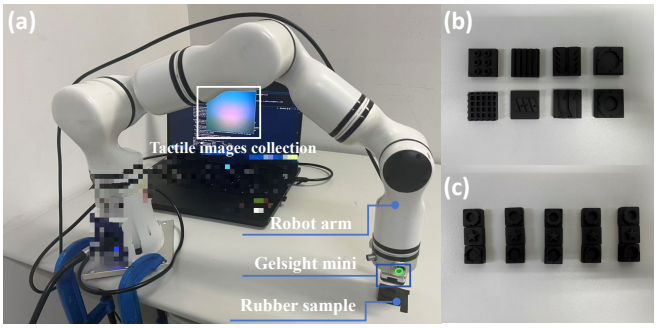


Fig. 3: Dataset collection process. (a) Data collection environment. (b) BRD-A dataset. (c) BRD-H dataset.

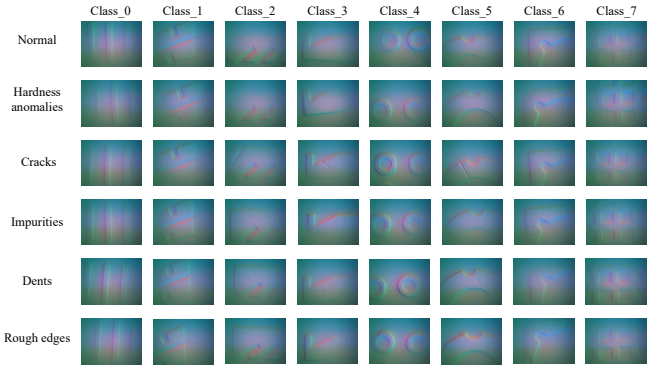


Fig. 4: The 48 normal and anomalous classes in BRD-A.

B. Evaluation Protocol

We use the dataset D, composed of labeled data D_l and unlabeled data D_u, to train the models. For evaluation, we use clustering accuracy (ACC) [28] as the performance metric, defined as $ACC = \frac{1}{M} \sum_{i=1}^{M} \mathbf{1}(y_i^* = p(\hat{y}_i))$, where y_i^* denotes the ground-truth label, $\hat{y}_i$ is the model's prediction, $M = |D_u|$, and p represents the optimal permutation aligning the predicted labels $\hat{y}$ with the ground-truth labels y^*.

C. Implementation details

During model training, we split all data into training, test sets in a 8:2 ratio. To align with the realistic industrial scenario where anomalous classes are largely unknown [2], [3], we designate the 8 normal classes in BRD-A and the 3 normal classes in BRD-H as known classes, while treating all remaining anomalous samples as unknown. The proportion of labeled data within the known classes is set to 50%. This setting presents a challenging task, as previous GCD works [17], [18] typically consider half of all classes as known; we follow this setting in our generalization experiments. All models are trained for 200 epochs. During the representation learning phase, the temperature parameters are set to $\tau_u = 0.07$ and $\tau_c = 1.0$.

V. EXPERIMENTAL RESULTS

A. Motivation Justification

As illustrated in Figure 1, this work is motivated by two fundamental observations: (1) Vision-based NACD methods

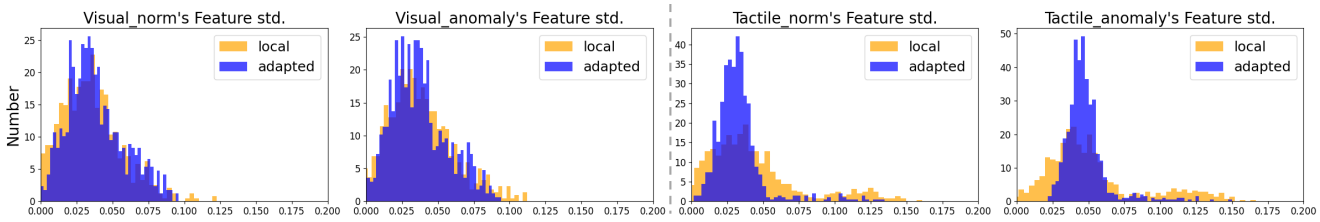


Fig. 5: **Feature space changes of visual and tactile images before and after the adaptor**. For visual images, the features of normal and anomalous samples remain difficult to distinguish even after passing through the adaptor. In contrast, anomalous features in tactile images can be effectively extracted by the adaptor and leveraged for subsequent classification.

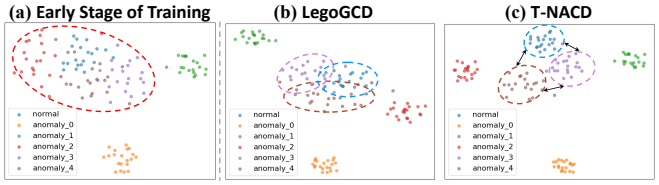


Fig. 6: **Impact of DMER on feature stocking**. (a) In the early stage of training, significant feature overlap occurs among different anomalous samples of the same texture. (b) LegoGCD can effectively cluster known classes but still suffers from feature stocking. (c) DMER encourages each sample to exhibit a single high-confidence class, thereby reducing inter-class confusion.

fail to detect the hardness of black rubber products and perform poorly in tasks involving fine-grained surface defect detection. (2) Replacing visual inputs with images captured by tactile sensors, combined with existing methods, enables hardness detection, but the task accuracy remains suboptimal. Our experiments are primarily designed to verify the following two issues.

- There is a significant domain shift between visual and tactile image features, and our proposed **feature adaptor** effectively compacts the gap.
- In the black rubber dataset, the feature representations of normal classes and unknown anomalous classes heavily overlap, making them hard to separate. Our proposed **dual-margin enhanced entropy regularization** helps preserve known class information while alleviating feature stacking across categories.

B. Tactile Domain Adaptation

To demonstrate the effectiveness of the adaptor, we selecte four types of images with the same textures from BRD-A: normal and anomalous samples in the visual domain, and their corresponding tactile images. We then computed histograms of the feature distributions before and after the feature adaptor during training. As shown in Figure 5, in the visual domain, normal and anomalous samples are difficult to distinguish, even with the use of an adaptor. In contrast, tactile images of normal samples exhibit clearer distributional differences, and the gap between normal and anomalous features becomes more pronounced after the adaptor. This provides strong evidence that the feature adaptor successfully transfers the pretrained model's feature extraction capabilities to the tactile domain.

C. Alleviation of Feature Stocking

We select six classes of tactile samples from the same texture category in BRD-A as the test set, including one normal sample and five anomalous samples. We train for 50 epochs on both LegoGCD and T-NACD, and then perform clustering on the feature space output by the adaptor. Figure 6(a) shows the feature distribution of epoch 3 on T-NACD, where severe feature stacking is observed. As shown in Figure 6(b), LegoGCD effectively separates three sample clusters and maintains the consistency of the `normal` class, but fails to clearly distinguish `anomaly_4` and `anomaly_5`. In contrast, T-NACD utilizes the DMER module to encourage each sample to align with its highest-confidence class, thereby enhancing feature boundaries and successfully separating `anomaly_4`, `anomaly_5`, and the `normal` class.

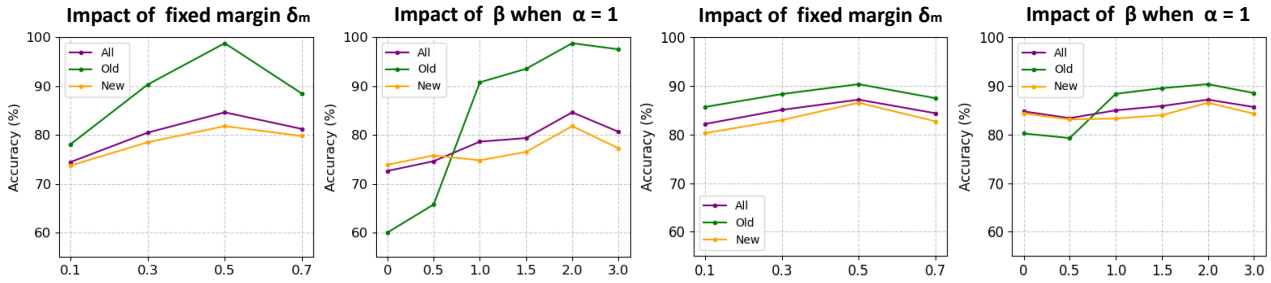
D. Comparison with the Baselines

We compare our method with classic Generalized Category Discovery (GCD) approaches, including k-means [33], RS+ [34], ORCA [35], GCD [28], SimGCD [17], OpenNCD [36], and LegoGCD [18]. Table I presents the evaluation results on our proposed BRD-A and BRD-H datasets, as well as a standard tactile dataset FabricVST [31] and YCB [32].

Specifically, compared to baselines such as OpenNCD [36] and LegoGCD [18], our method achieves the best performance on the black rubber datasets, as shown in Table I. It exhibits high accuracy in identifying old classes (i.e., normal samples) while also demonstrating strong ability in discovering new classes. Specifically, T-NACD outperforms the best baseline by **6.9%/2.41%** on BRD-A and BRD-H, respectively. When compared to LegoGCD on BRD-A, our method improves old-class/new-class accuracy by **17.5%/4.78%**. Notably, T-NACD achieves 98.75% accuracy on known categories, which aligns with the high precision requirements in industrial inspection scenarios, because we first need to ensure the accurate identification of normal samples. Furthermore, under the challenging setting of discovering 40 anomaly classes from 8 known normal classes, T-NACD still achieves over 80% new class accuracy, showing its powerful anomaly recognition capability. We further

Method	BRD-A			BRD-H			FabricVST			YCB		
	All	Old	New	All	Old	New	All	Old	New	All	Old	New
k-means	28.10	35.31	26.66	43.55	51.5	39.58	62.00	65.30	58.17	49.13	53.05	47.09
RS+	31.25	38.47	29.37	44.12	56.39	38.25	61.29	64.88	56.1	50.25	55.12	46.77
ORCA	35.85	42.64	32.59	46.96	50.68	41.02	65.62	70.35	60.36	54.58	60.76	51.89
SimGCD	72.61	60.00	75.13	84.74	80.00	84.33	94.62	93.25	95.60	71.23	69.45	73.22
OpenNCD	69.18	76.94	67.63	83.33	85.33	82.33	87.35	88.42	87.14	67.92	70.24	63.85
LegoGCD	77.71	81.25	77.00	83.67	85.30	82.86	91.25	90.50	92.00	75.36	79.26	72.29
Ours	**84.61**	**98.75**	**81.78**	**87.15**	**90.33**	**86.51**	**95.75**	**95.55**	**95.95**	**82.09**	**87.35**	**79.96**
Δ	**+6.90**	**+17.50**	**+4.78**	**+2.41**	**+10.33**	**+2.18**	**+1.13**	**+2.30**	**+0.35**	**+6.73**	**+8.09**	**+7.67**

TABLE I: **Classification results on BRD datasets and two mainstream tactile datasets.** T-NACD achieves the best performance across all datasets, enhancing the classification of unknown classes while maintaining the accuracy of known classes, making it well-suited for industrial anomaly detection scenarios.



(a) BRD-A **(b) FabricVST**

Fig. 7: Ablation study on the fixed margin δ_m and the regularization coefficient β on BRD-A and FabricVST.

Methods	All	Old	New	Macro-F1
w/o Augment	81.41	92.55	78.99	0.823
w/o Adaptor	77.90	91.95	75.09	0.776
w/o $\mathcal{L}_{cal}$	80.55	89.71	78.71	0.802
w/o $\mathcal{L}_{mar}$	81.05	94.75	78.32	0.809
w/o $\mathcal{L}_{dmer}$	78.3	86.25	76.7	0.789
LegoGCD	77.71	81.25	77.00	0.769
T-NACD	**84.61**	**98.75**	**81.78**	**0.874**

TABLE II: Ablation study on different components of our algorithms on BRD-A.

Methods	BRD-A	BRD-H	FabricVST	YCB
k-means	0.582	0.575	0.621	0.609
RS+	0.603	0.622	0.678	0.638
ORCA	0.614	0.636	0.681	0.649
SimGCD	0.753	0.815	0.882	0.749
OpenNCD	0.718	0.769	0.831	0.701
LegoGCD	0.769	0.803	0.876	0.774
T-NACD	**0.874**	**0.889**	**0.924**	**0.858**

TABLE III: The macro-F1 results on the BRD and two tactile datasets.

evaluate the overall recognition rate of unknown samples using the macro-F1 metric. As shown in Table III, T-NACD achieves the highest performance across all four datasets.

E. Model generalization

Beyond tactile anomaly detection for black rubber products, we also demonstrate that T-NACD, as a general framework for tactile image classification, can generalize to other tactile datasets. As shown in Table I, T-NACD still achieves strong performance on the mainstream tactile datasets FabricVST [31] and YCB [32].

F. Ablation Study

In this section, we conduct ablation studies on the BRD-A and FabricVST to evaluate the effectiveness of the Tactile-specific Feature Extractor, and the DMER module, as shown in Figure 7.

Tactile-specific Feature Extractor. We conduct an ablation study on the Tactile-specific Feature Extractor, as shown in Table II. The introduced tactile-specific data augmentation method and lightweight feature adaptor improve accuracy on the BRD-A dataset by 1.2% and 4.7%, respectively. Furthermore, as illustrated in Figure 5, we successfully aggregate the features of tactile images through the feature extractor and separate the feature spaces of normal and anomalous samples, which is a key factor contributing to the performance improvement of T-NACD on tactile tasks.

Dual-Margin Enhanced Entropy Regularization. We separately evaluate the effects of $\mathcal{L}_{\text{cal}}$ and $\mathcal{L}_{\text{mar}}$ on the BRD-A dataset. As Table II shows, L_{cal} mitigates forgetting of known classes, raising old-class accuracy by 9%. L_{mar} improves inter-class separability on unlabeled data by enforcing confident assignments, yielding larger gains on the more abundant new classes. Figure 7 presents ablation studies on the fixed margin parameter δ_m, due to the high classification difficulty, an excessively large δ_m is suboptimal. We ultimately select $\delta_m = 0.5$ as the default setting throughout our experiments.

For regularization, we fix $\alpha = 1$ to match SimGCD and sweep $\beta \in [0, 3]$, finding $\beta = 2.0$ achieves the best balance on BRD-A and FabricVST.

VI. CONCLUSION

We present T-NACD, a tactile-friendly framework for novel anomaly class discovery, targeting black rubber products where visual methods are insufficient. By introducing two tactile-specific augmentation methods, a lightweight feature adapter, and the DMER module, T-NACD effectively captures fine-grained geometric features and mitigates feature stocking. We also construct BRD, the first tactile anomaly detection dataset for this domain. Experiments show that T-NACD outperforms visual baselines and generalizes well to other tactile detection tasks.

REFERENCES

[1] Z. Huang, X. Li, H. Liu, F. Xue, Y. Wang, and Y. Zhou, "Anomalyncd: Towards novel anomaly class discovery in industrial scenarios," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 4755–4765.

[2] B. Hu, B. Gao, W. L. Woo, L. Ruan, J. Jin, Y. Yang, and Y. Yu, "A lightweight spatial and temporal multi-feature fusion network for defect detection," *IEEE Transactions on Image Processing*, vol. 30, pp. 472–486, 2020.

[3] J. Jeong, Y. Zou, T. Kim, D. Zhang, A. Ravichandran, and O. Dabeer, "Winclip: Zero-/few-shot anomaly classification and segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 19 606–19 616.

[4] Z. Zeng, B. Liu, J. Fu, and H. Chao, "Reference-based defect detection network," *IEEE Transactions on Image Processing*, vol. 30, pp. 6637–6647, 2021.

[5] A.-T. Ardelean and T. Weyrich, "Blind localization and clustering of anomalies in textures," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 3900–3909.

[6] Y. Lee, H. Lim, S. Jang, and H. Yoon, "Uniformaly: Towards task-agnostic unified framework for visual anomaly detection," *Pattern Recognition*, p. 111820, 2025.

[7] T.-H. Nguyen, H.-L. Nguyen, N.-T. Bui, T.-H. Bui, V.-B. Vu, H.-N. Duong, and H.-H. Hoang, "Vision-based system for black rubber roller surface inspection," *Applied Sciences*, vol. 13, no. 15, p. 8999, 2023.

[8] W. Kitisavetjit, N. Paradee, K. Ounjai, E. Kalkornsurapranee, R. Promsung, J. Johns, and Y. Nakaramontri, "Carbon nanotube/conductive carbon black-filled natural rubber composites for strain sensing," *Materials Chemistry and Physics*, vol. 341, p. 130860, 2025.

[9] E. Preto, P. R. Barboni, J. A. Malmonge, and S. da Silva, "Detection of torque variation in bolted joints structures using a new sensor made of natural rubber filled with carbon black and carbon nanotubes," *Journal of Intelligent Material Systems and Structures*, vol. 36, no. 8, pp. 531–542, 2025.

[10] N. S. Lewis, "Comparisons between mammalian and artificial olfaction based on arrays of carbon black- polymer composite vapor detectors," *Accounts of chemical research*, vol. 37, no. 9, pp. 663–672, 2004.

[11] Y. Chen, L. Wang, Z. Wu, J. Luo, B. Li, X. Huang, H. Xue, and J. Gao, "Super-hydrophobic, durable and cost-effective carbon black/rubber composites for high performance strain sensors," *Composites Part B: Engineering*, vol. 176, p. 107358, 2019.

[12] S. Dong, W. Yuan, and E. H. Adelson, "Improved gelsight tactile sensor for measuring geometry and slip," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 137–144.

[13] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer *et al.*, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.

[14] M. Meribout, N. A. Takele, O. Derege, N. Rifiki, M. El Khalil, V. Tiwari, and J. Zhong, "Tactile sensors: A review," *Measurement*, vol. 238, p. 115332, 2024.

[15] X. Liu, W. Yang, F. Meng, and T. Sun, "Material recognition using robotic hand with capacitive tactile sensor array and machine learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–9, 2024.

[16] V. U. Ludwig and J. Simner, "What colour does that feel? tactile–visual mapping and the development of cross-modality," *Cortex*, vol. 49, no. 4, pp. 1089–1099, 2013.

[17] X. Wen, B. Zhao, and X. Qi, "Parametric classification for generalized category discovery: A baseline study," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 16 590–16 600.

[18] X. Cao, X. Zheng, G. Wang, W. Yu, Y. Shen, K. Li, Y. Lu, and Y. Tian, "Solving the catastrophic forgetting problem in generalized category discovery," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 16 880–16 889.

[19] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The caltech-ucsd birds-200-2011 dataset," 2011.

[20] W. Zheng, H. Liu, D. Guo, and F. Sun, "Robust tactile object recognition in open-set scenarios using gaussian prototype learning," *Frontiers in Neuroscience*, vol. 16, p. 1070645, 2022.

[21] K. Liu, Q. Yang, Y. Xie, and X. Huang, "Towards open-set material recognition using robot tactile sensing," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10 345–10 351.

[22] Z. You, L. Cui, Y. Shen, K. Yang, X. Lu, Y. Zheng, and X. Le, "A unified model for multi-class anomaly detection," *Advances in Neural Information Processing Systems*, vol. 35, pp. 4571–4584, 2022.

[23] K. Han, A. Vedaldi, and A. Zisserman, "Learning to discover novel visual categories via deep transfer clustering," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 8401–8409.

[24] Y.-C. Hsu, Z. Lv, and Z. Kira, "Learning to cluster in order to transfer across domains and tasks," *arXiv preprint arXiv:1711.10125*, 2017.

[25] J. Fan, D. Liu, H. Chang, H. Huang, M. Chen, and W. Cai, "Seeing unseen: Discover novel biomedical concepts via geometry-constrained probabilistic modeling," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 11 524–11 534.

[26] J. Zhou, Y. Liu, and Q. Chen, "Novel class discovery in chest x-rays via paired images and text," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, 2024, pp. 7650–7658.

[27] Y. Zhao, Z. Zhong, N. Sebe, and G. H. Lee, "Novel class discovery in semantic segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 4340–4349.

[28] S. Vaze, K. Han, A. Vedaldi, and A. Zisserman, "Generalized category discovery," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 7492–7501.

[29] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[30] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin, "Emerging properties in self-supervised vision transformers," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 9650–9660.

[31] G. Cao, J. Jiang, D. Bollegala, M. Li, and S. Luo, "Multimodal zero-shot learning for tactile texture recognition," *Robotics and Autonomous Systems*, vol. 176, p. 104688, 2024.

[32] S. Suresh, Z. Si, S. Anderson, M. Kaess, and M. Mukadam, "Midastouch: Monte-carlo inference over distributions across sliding touch," in *Conference on Robot Learning*. PMLR, 2023, pp. 319–331.

[33] J. B. McQueen, "Some methods of classification and analysis of multivariate observations," in *Proc. of 5th Berkeley Symposium on Math. Stat. and Prob.*, 1967, pp. 281–297.

[34] K. Han, S.-A. Rebuffi, S. Ehrhardt, A. Vedaldi, and A. Zisserman, "Autonovel: Automatically discovering and learning novel visual categories," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6767–6781, 2021.

[35] J. Liu, Y. Wang, T. Zhang, Y. Fan, Q. Yang, and J. Shao, "Open-world semi-supervised novel class discovery," *arXiv preprint arXiv:2305.13095*, 2023.

[36] J. Liu, Y. Wang, T. Zhang, Y. Fan, Q. Yang *et al.*, "Open-world semi-supervised novel class discovery," in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 2023, pp. 4002–4010.