G3C: Graph consistency based on contrastive clustering for semi-supervised classification

Wei Liu[1], Jiangtao Song[1], Yuanbo Li [1], Cong Hu [1]*, Xiaojun Wu [1],
[1]School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, China
6233110026@stu.jiangnan.edu.cn, jtsong_sd@163.com, liyuanbo12138@163.com,
conghu@jiangnan.edu.cn, wu_xiaojun@jiangnan.edu.cn

***Abstract*—Graph Semi-Supervised Learning exploits topological structures to leverage unlabeled data. However, in open-set scenarios, existing methods typically rely on model predictions to construct adjacency graphs. This dependency makes them susceptible to confirmation bias, where incorrect pseudo-labels from unlabeled samples propagate noise through the graph structure. Furthermore, these approaches often neglect global cluster compactness, focusing primarily on local pairwise consistency. To address these limitations, we propose a novel framework named Graph consistency based on contrastive clustering (G3C). Unlike traditional approaches, G3C integrates contrastive clustering to jointly optimize sample similarity and global cluster structures. Specifically, we perform dynamic neighbor searching within the augmented feature space to construct reliable weight matrices. It also constrains the representation of data in label space and feature space to distinguish known classes from outliers through graph consistency regularization. Extensive experiments on CIFAR-10, CIFAR-100, and ImageNet-30 demonstrate that G3C outperforms state-of-the-art approaches in both closed-set and open-set settings, validating the efficacy of combining contrastive clustering with graph consistency.**

***Index Terms*—contrastive clustering, graph consistency, semi-supervised classification**

I. Introduction

The rapid advancement of artificial neural networks has propelled Semi-Supervised Learning (SSL) to the forefront of various computer vision tasks, including image classification [1], video understanding [2], and face recognition [3]. By leveraging limited labeled data alongside massive unlabeled data, SSL not only alleviates the exorbitant costs associated with manual annotation but also empowers models to uncover latent data patterns, thereby significantly enhancing generalization capabilities.

There are many methods involved in SSL, such as contrastive learning [4] and graph semi-supervised learning (GSSL) [5]. GSSL exploits the topological structure of data to propagate label information based on instance correlations. Miller et al. [6] combined active learning with GSSL, which uses the Laplacian matrix spectrum to truncate the matrix, reducing computational and storage costs without affecting model performance. Conversely, Contrastive Learning, a dominant approach in unsupervised representation learning, learns data representation by bringing similar samples (positive sample pairs) closer and pushing dissimilar samples (negative sample pairs) apart without manual labeling [7]. Recently, Zhu et al. [8] analyzed the trajectory-based optimization and signal-to-noise ratio in contrastive paradigms, providing a theoretical foundation for understanding generalization in downstream tasks.

However, both paradigms face distinct limitations [9]. GSSL relies heavily on the quality of the affinity graph, making it susceptible to noise during construction. Specifically, when the graph representation is singular and lacks robust supervisory guidance, such noise can dominate the learning process, inevitably leading to degraded classification performance. On the other hand, standard contrastive learning focuses primarily on instance-level discrimination, often overlooking the global semantic structure of the data. To bridge this gap, Contrastive Clustering has emerged as a promising direction, optimizing both sample similarity and intra-cluster consistency simultaneously [10], [11]. This enables the utilization of global structural information for more accurate semantic grouping in unsupervised scenarios.

In this work, we propose a novel framework named Graph Consistency based on contrastive clustering (G3C). Unlike traditional approaches that rely on prediction-dependent graphs, G3C performs dynamic neighbor searching within the augmented feature space to construct reliable weight matrices. To ensure the reliability of this feature-based graph, we integrate contrastive clustering to jointly optimize sample similarity and global cluster structures. This dual optimization strategy drives the model to learn compact semantic groupings, effectively mitigating the noise propagation caused by incorrect pseudo-labels. Additionally, we introduce an unknown-aware graph consistency regularization that aligns the pseudo-label graph with the embedding graph, constraining data representations to distinguish known classes from correlated outliers.

Our contributions are summarized as follows:

- We propose G3C, a unified semi-supervised learning framework that combines contrastive clustering with graph consistency to enhance model robustness, particularly in open-set scenarios.
- We introduce a dual optimization strategy that jointly optimizes sample similarity and global cluster structures. This ensures feature-space compactness, effectively mitigating the noise propagation caused by prediction bias.
- Extensive experiments demonstrate that G3C outperforms state-of-the-art methods, achieving superior pseudo-label quality, classification accuracy, and AUROC scores on standard benchmarks.

II. Related Work

A. Semi-Supervised Learning

Semi-supervised learning (SSL) mitigates the dependency on massive annotated datasets by combining labeled and unlabeled data. Representative methods like FixMatch [12] effectively integrate pseudo-labeling and consistency regularization, significantly enhancing generalization in closed-set environments. Furthermore, CoMatch [13] introduces contrastive graph regularization to strengthen feature robustness. However, these methods primarily focus on instance-level consistency, often overlooking the global manifold structure of the data. To better capture spatial correlations, Graph Semi-Supervised Learning (GSSL) [5], [14], [15] models topological dependencies via graph structures. Some works utilize techniques like Laplacian spectrum truncation to reduce computational overhead [6]. While capturing structural information, GSSL relies heavily on the quality of the initial graph and is susceptible to noise, especially in the presence of outliers.

Addressing open-set challenges, Kong et al. [9] proposed Unknown-Aware Graph Regularization (UAG), which isolates outliers by maximizing their entropy and constructing pseudo-label graphs. Building upon this, our G3C framework integrates contrastive clustering into the graph consistency paradigm. By performing dynamic neighborhood searching to construct the weight matrix, G3C simultaneously optimizes sample similarity and clustering structure, ensuring robustness against open-set noise.

B. Open-set semi-supervised learning

Open-Set Semi-Supervised Learning (OSSL) handles uncurated unlabeled data containing out-of-class samples. Early approaches primarily aimed to preserve closed-set accuracy by employing adaptive weighting [16] or filtering techniques [17] to exclude potential outliers. However, these methods generally lack explicit optimization objectives for separating in-distribution data from unknown samples. To enhance discriminative capacity, subsequent research pivoted toward distance-based [18], [19] and uncertainty-based detection [20], [21]. These frameworks utilize feature space distances or predictive confidence to identify outliers. Despite advancements, such methods often struggle with correlated outliers (unknown samples sharing high semantic similarity with known classes), where simple distance or single-head uncertainty measures lack precision. Recently, UAG [9] addressed this by adopting a multi-head architecture to model inliers and outliers independently, mitigating conflicts between entropy minimization and maximization. However, it still relies heavily on prediction-based graphs, rendering it vulnerable to confirmation bias. In contrast, our G3C integrates contrastive clustering to optimize intrinsic feature structure, enforcing cluster compactness via dynamic neighbor searching to mitigate such noise.

III. Method

A. Problem Definition

Given a batch of labeled data $\mathcal{X} = \{(x_b, y_b)\}_{b=1}^{B}$, where y_b is the label corresponding to the data x_b, and a batch of unlabeled data $\mathcal{U} = \{u_b\}_{b=1}^{\mu B}$, the parameter μ determines the ratio of labeled data to unlabeled data. Our goal is to classify inliers into known categories and distinguish outliers. The proposed model consists of a shared encoder $f(\cdot)$ with three heads: a closed-set classifier $h(\cdot)$, an open-set classifier $\widetilde{h}(\cdot)$, and a projection head $g(\cdot)$.

B. Overview

Fig. 1 illustrates the proposed G3C framework. We adopt a dual-branch strategy to process labeled data $\mathcal{X}$ and unlabeled data $\mathcal{U}$. For labeled data, the model is trained using the supervised loss $\mathcal{L}_s$. Specifically, the encoder $f(\cdot)$ shares weights across all heads. Let p and $\tilde{p}$ denote the softmax probabilities from the closed-set classifier h and open-set classifier $\tilde{h}$, respectively. The supervised objective is formulated as:

$$\mathcal{L}_s = \frac{1}{B}\sum_{b=1}^{B}\left(H\left(y_b, p\left(A_w\left(x_b\right)\right)\right) + H\left(y_b, \widetilde{p}\left(A_w\left(x_b\right)\right)\right)\right), \tag{1}$$

where $H(\cdot,\cdot)$ denotes the cross-entropy between two distributions. For unlabeled data, we perform three distinct regularization: contrastive clustering, uncertainty-based consistency, and unknown-aware graph regularization, detailed below.

C. Graph Contrastive Clustering

Unlike previous uncertainty-based methods that treat outliers as a single unknown class, we propose a graph contrastive clustering module to capture the intrinsic manifold structure of the data. This allows the model to form compact semantic clusters in the feature space, distinguishing effectively between inliers and correlated outliers.

We construct two views of the unlabeled data: the original features $z = f(u)$ and the weakly augmented features $z' = f(A_w(u))$. To mitigate the confirmation bias arising from prediction-dependent graphs and capture the evolving manifold structure, we perform dynamic neighbor searching. For each anchor feature z_i, we identify its set of k-nearest neighbors $NN_k(z_i)$ within the augmented view z'. This structural relationship is encoded in a weight matrix W, defined as:

$$W_{ij} = \begin{cases} 1, & z_j = {z'}_i \text{ (self-loop)} \\ 1/k, & z_j \in NN_k(z_i) \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

To enforce cluster compactness, we minimize the contrastive clustering loss $\mathcal{L}_c$. This objective maximizes the similarity between neighbors while pushing apart non-neighbors (where $W_{ij} = 0$):

$$\mathcal{L}_c = -\frac{1}{N}\sum_i \log \frac{\sum_{j:W_{ij}>0} W_{ij} \exp\left(z'_i \cdot z'_j / T_z\right)}{\sum_{k:W_{ik}=0} \exp\left(z'_i \cdot z'_k / T_z\right)}, \tag{3}$$

where T_z is a scalar hyperparameter. By optimizing $\mathcal{L}_c$, G3C ensures that samples with similar semantic content are grouped together, facilitating the separation of outliers in the embedding space.

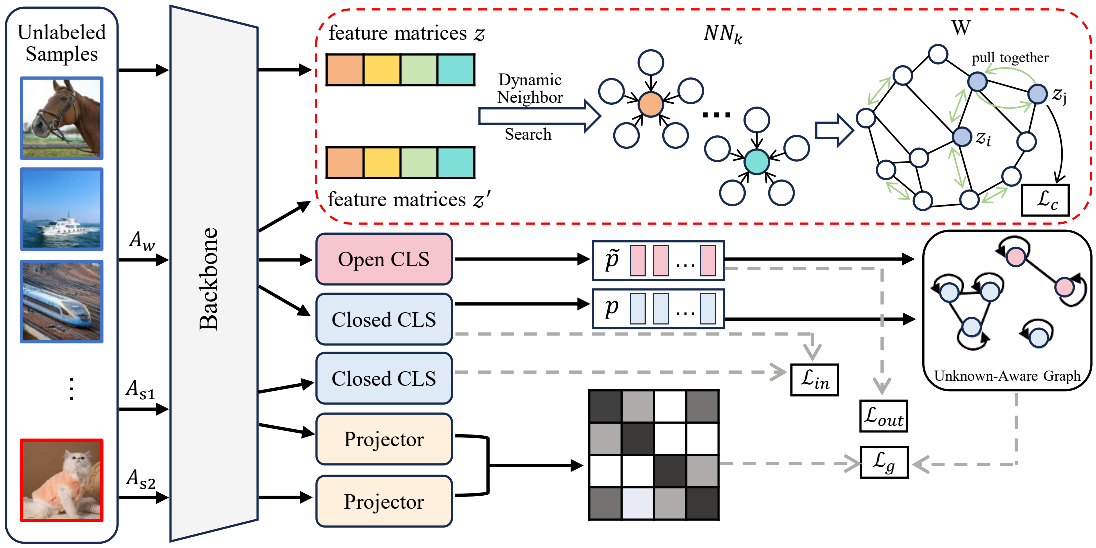


Fig. 1: Overview of the proposed G3C framework. The unlabeled data contains in-distribution samples (blue) and out-of-distribution outliers (red). The framework consists of two objectives: The Graph Contrastive Clustering module constructs a dynamic KNN graph by performing neighbor searching between augmented features z and z', optimizing the contrastive clustering loss $\mathcal{L}_c$ to ensure cluster compactness. The Unknown-Aware Graph mechanism leverages classifier predictions $(p, \tilde{p})$ to build a reliable target graph, guiding the embedding space learning via graph consistency regularization $\mathcal{L}_g$.

D. Uncertainty-based Consistency

To separate known classes from outliers, we exploit the uncertainty in the logits. We utilize the Maximum Softmax Probability (MSP) score, defined as:

$$s(x;T) = \max_i \frac{\exp\left([\tilde{h} \circ f(x)]_i / T\right)}{\sum_{j=1}^{K} \exp\left([\tilde{h} \circ f(x)]_j / T\right)}, \tag{4}$$

where $T > 1$ amplifies the distribution gap. $[\tilde{h} \circ f(x)]_i$ represents the logits of category i. Adaptive thresholds $\hat{\tau}_t^{in}$ and $\hat{\tau}_t^{out}$ are maintained via Exponential Moving Average (EMA) on a two-component Gaussian Mixture Model (GMM) fit to the scores at each epoch. Their initial values $\hat{\tau}_0^{in}$ and $\hat{\tau}_0^{out}$ are set to 1. The update rules are as follows:

$$\hat{\tau}_t^{in} \leftarrow m\hat{\tau}_{t-1}^{in} + (1-m)\,\tau_t^{in}, \tag{5}$$

$$\hat{\tau}_t^{out} \leftarrow m\hat{\tau}_{t-1}^{out} + (1-m)\,\tau_t^{out}. \tag{6}$$

Based on these thresholds, we apply exclusive consistency regularization. For samples identified as inliers: $s_b \geq \hat{\tau}_t^{in}$, where $s_b = s(u_b;T)$ represents the MSP score of the unlabeled data u_b, we minimize the standard consistency loss $\mathcal{L}_{in}$ using pseudo-labels $\hat{y}_b$ [12]:

$$\mathcal{L}_{in} = \frac{1}{\mu B} \sum_{b=1}^{\mu B} \mathbb{I}\left(s_b \geq \hat{\tau}_t^{in}\right) \cdot H\left(\hat{y}_b, p\left(A_{s1}\left(u_b\right)\right)\right). \tag{7}$$

Conversely, for detected outliers: $s_b < \hat{\tau}_t^{out}$, we maximize the entropy of the open-set classifier output to prevent it from overconfidently assigning outliers to any specific class. This outlier loss $\mathcal{L}_{out}$ is defined as:

$$\mathcal{L}_{out} = -\frac{1}{\mu B} \sum_{b=1}^{\mu B} \mathbb{I}\left(s_b < \hat{\tau}_t^{out}\right) \cdot H\left(\tilde{p}\left(A_w\left(u_b\right)\right)\right), \tag{8}$$

where $H(\cdot)$ represents the average entropy. Finally, we integrate the inlier consistency and outlier entropy maximization into a unified unsupervised loss $\mathcal{L}_u = \mathcal{L}_{in} + \mathcal{L}_{out}$.

E. Unknown-Aware Graph Regularization

To further constrain the embedding space, we adopt the Unknown-Aware Graph strategy [9]. We construct a target graph W^{uag} that integrates knowledge from both classifiers. The pseudo-label graph $W^{pg} \in \mathbb{R}^{\mu B \times \mu B}$ connects samples with high closed-set prediction similarity:

$$W_{bj}^{pg} = \begin{cases} 1, & b = j \\ p_b \cdot p_j, & b \neq j \text{ and } p_b \cdot p_j \geq \tau_g \\ 0, & \text{otherwise} \end{cases}, \tag{9}$$

where τ_g is a predefined threshold, p_b and p_j represent the softmax probabilities of u_b and u_j predicted by $h \circ f$, respectively. The outlier graph $W^{og} \in \mathbb{R}^{\mu B \times \mu B}$ connects samples that are both identified as outliers:

$$W_{bj}^{og} = \begin{cases} 1, & \eta_b = \eta_j \\ 0, & \text{otherwise} \end{cases}, \tag{10}$$

where $\eta_b = \mathbb{I}(s_b < \hat{\tau}_t^{out})$, returning 1 for values predicted to be outliers and 0 for other values. The final target graph is their intersection: $W^{uag} = W^{pg} \times W^{og}$. That is, only samples with the same prediction in the unknown context and similar

TABLE I: Error rates (%) on CIFAR-10 and CIFAR-100 on 3 different folds. The best results are highlighted in bold.

Dataset	CIFAR-10				CIFAR-100			
	Uncorr.		Corr.		Uncorr.		Corr.	
Label	50	100	50	100	50	100	50	100
FixMatch	16.8	10.7	17.5	12.9	33.6	29.4	30.8	28.8
CoMatch	12.7	9.5	14.8	10.3	28.5	26.4	28.8	25.8
MTC	20.4	13.5	21.8	14.3	36.7	30.9	36.9	29.6
OpenMatch	10.2	7.1	11.7	9.2	30.5	26.7	30.1	25.3
OSP	12.1	9.2	11.1	9.5	29.5	26.5	30.7	26.9
UAG	9.6	**5.8**	8.1	**6.8**	26.6	23.6	**26.4**	23.8
Ours	**9.28**	6.86	**8.01**	7.38	**26.37**	**22.90**	26.97	**23.75**

TABLE II: AUROC (%) on CIFAR-10 and CIFAR-100 on 3 different folds. The best results are highlighted in bold.

Dataset	CIFAR-10				CIFAR-100			
	Uncorr.		Corr.		Uncorr.		Corr.	
Label	50	100	50	100	50	100	50	100
FixMatch	54.2	58.5	55.9	59.4	64.1	66.7	60.4	62.6
CoMatch	47.6	47.7	48.1	48.9	60.1	61.7	55.6	57.9
MTC	96.5	98.2	73.5	78.2	81.7	82.6	69.3	70.4
OpenMatch	**97.9**	**99.6**	75.6	55.3	86.1	86.2	69.9	69.4
OSP	62.9	66	45.7	46.4	58.5	60.3	56.1	59.5
UAG	95.5	97.8	90.2	96.4	87.9	91.8	**78.7**	81.4
Ours	97.41	97.90	**96.89**	**97.77**	**90.35**	**92.23**	75.69	**81.84**

predictions in the known context will be assigned to the same cluster in the graph.

We align the projected embedding space with this target structure. Let W^z be the similarity matrix of projected features from two strong augmentations (z^{s1}, z^{s2}):

$$W_{bj}^{z} = \begin{cases} \exp\left(z_b^{s1} \cdot z_b^{s2}/T_z\right), & b = j \\ \exp\left(z_b^{s1} \cdot z_j^{s1}/T_z\right), & \text{otherwise} \end{cases}, \tag{11}$$

subsequently, the graph is normalized by $\hat{W}_{bj} = W_{bj}/\sum_j W_{bj}$. So that each row of the similarity matrix sums to 1. The graph consistency loss $\mathcal{L}_g$ minimizes the cross-entropy between the normalized target and embedding graphs:

$$\mathcal{L}_g = \frac{1}{\mu B} \sum_{b=1}^{\mu B} H\left(\hat{W}_{bj}^{uag}, \hat{W}_{bj}^{z}\right), \tag{12}$$

F. Total Objective

The final objective function of G3C is a weighted sum of the aforementioned losses:

$$\mathcal{L}_{total} = \mathcal{L}_s + \lambda_u \mathcal{L}_u + \lambda_g \mathcal{L}_g + \lambda_c \mathcal{L}_c, \tag{13}$$

where λ terms are hyperparameters balancing the contributions of each module.

IV. Experiment

A. Experimental Setup

Datasets and Metrics. We evaluate G3C on three benchmarks: CIFAR-10, CIFAR-100, and ImageNet-30. For the CIFAR datasets, we adopt a standard split where 60% of classes are treated as known and the remaining 40% as unknown. To rigorously evaluate robustness, we consider two distinct open-set configurations: *uncorrelated* (outliers share no superclass with inliers) and *correlated* (outliers and inliers share superclasses, presenting a harder semantic separation task). For ImageNet-30 [22], we employ a subset of ImageNet due to computational constraints. In this setting, the first 20 classes are treated as known (with only 10% labeled) and the remaining 10 as unknown. All images in ImageNet-30 are resized to 224×224 to assess scalability on higher-resolution data. We report the error rate for closed-set accuracy and AUROC for open-set detection performance.

TABLE III: Error rates (%) and AUROC (%) on ImageNet-30 on 3 different folds. The best results are highlighted in bold.

Dataset	ImageNet-30	
Metric	Error Rate	AUROC
FixMatch	12.5	87.9
CoMatch	8.8	65.8
MTC	13.6	93.8
OpenMatch	10.4	96.4
UAG	6.1	97.4
Ours	**4.55**	**98.04**

Baselines and Implementation Details. We benchmark our framework against state-of-the-art semi-supervised learning methods, including FixMatch [12], CoMatch [13], OpenMatch [19], MTC [23], OSP [24], and UAG [9]. Following established protocols, we employ a Wide ResNet (WRN-28) backbone for the CIFAR datasets and a ResNet-18 backbone for ImageNet-30. All models are optimized using SGD with a momentum of 0.9. The learning rate is initialized at 0.03 and follows a cosine decay schedule. We set the weight decay to 5×10^{-4} for both backbones.

The training process spans 200×1024 iterations for CIFAR-10/100 and 100×1024 iterations for ImageNet-30. We use a batch size of $B = 64$ and set the labeled-to-unlabeled ratio μ to 4. For the remaining hyperparameters, we adopt the settings from prior work [9] to ensure stability and avoid redundant tuning. Specifically, the temperature T for the uncertainty score is set to 1.5, and the momentum m for the threshold update is 0.9. Regarding the graph construction, we set the graph threshold τ_g to 0.8. The embedding temperature T_z is fixed at 0.2.

B. Main Results

Table I and Table III report the closed-set error rates. G3C demonstrates superior robustness, particularly in challenging low-label regimes. On CIFAR-10 with only 50 labels, G3C outperforms the strong baseline UAG in both uncorrelated and correlated settings, consistent with the improved pseudo-label accuracy shown in Fig. 2. While UAG remains competitive given abundant labels (100 labels), G3C consistently regains the advantage as dataset complexity increases. This scalability is further validated on ImageNet-30 , where G3C achieves a significant reduction in error rate, marking a 1.55% im-

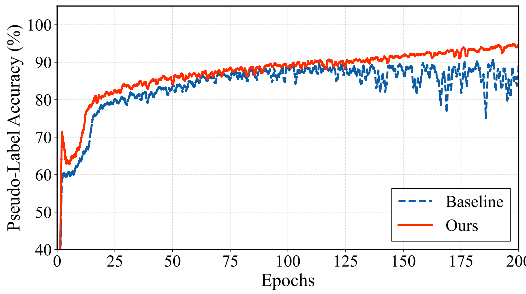
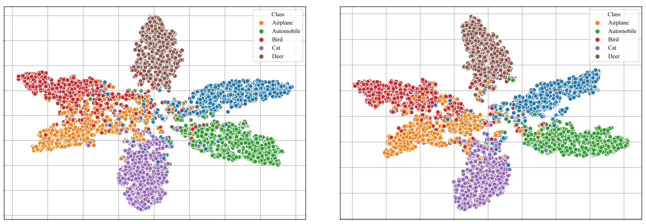


Fig. 2: Comparison of pseudo-label accuracy during training on CIFAR-10 with 50 labeled samples in the uncorrelated setting. Our method (red) demonstrates higher accuracy and stability compared to the baseline (blue).



(a) w/o CC

(b) G3C

Fig. 3: t-SNE visualization of learned features. (a) Features generated by the model without the contrastive clustering module. (b) Features generated by the proposed G3C model.

provement. These results confirm that integrating contrastive clustering enables more effective feature learning when labeled data is scarce or the visual manifold is complex.

Table II and Table III detail the AUROC scores, evaluating the model's ability to distinguish knowns from unknowns. G3C achieves state-of-the-art performance in nearly all configurations. Crucially, in the correlated setting—where outliers are semantically close to inliers—G3C exhibits remarkable stability. For instance, on CIFAR-10 (Corr. 50 labels), G3C achieves 96.89%, surpassing UAG's 90.2%. Similarly, on ImageNet-30, G3C sets a new benchmark with 98.04% AUROC. This validates that our method effectively pushes outliers away from the decision boundaries of known classes, even when semantic ambiguity is high.

V. Ablation Study

A. *Effectiveness of Contrastive Clustering Module*

To investigate the contribution of the Contrastive Clustering (CC) module to semi-supervised classification performance, we utilize t-SNE (t-distributed Stochastic Neighbor Embedding) for a qualitative analysis of the learned feature space. This ablation study is conducted on the CIFAR-10 dataset under the uncorrelated setting, using a WRN backbone with 50 labeled samples. The visualization results are presented in Fig. 3. As illustrated in Fig. 3(a), the model lacking the CC module (w/o CC) exhibits relatively ambiguous decision boundaries. Notably, the model struggles to differentiate between the *Bird* and *Airplane* categories, where samples near the boundary are poorly separated, making the decision boundary nearly indistinguishable. In contrast, as shown in Fig. 3(b), the full G3C framework produces significantly clearer classification boundaries. These observations are more consistent with the low-density separation and clustering hypotheses. This qualitative improvement demonstrates that the joint optimization of sample similarity and global cluster structures effectively bolsters the clustering performance of the G3C model, leading to the formation of more cohesive and discriminative cluster structures.

TABLE IV: Quantitative ablation results (Error Rate (%) and AUROC (%)) on CIFAR-10 under the uncorrelated setting.

Dataset	CIFAR-10		CIFAR-100	
Metric	Error Rate	AUROC	Error Rate	AUROC
w/o CC	11.23	96.81	27.01	88.21
G3C	**9.28**	**97.41**	**26.37**	**90.35**

TABLE V: Sensitivity analysis of the hyperparameter λ_c on CIFAR-100 under the uncorrelated setting.

Dataset	CIFAR-100				
λ_c	0.1	0.5	1	2	5
Error Rate	27.15	26.45	**25.97**	26.19	26.38
AUROC	89.97	**91.22**	91.19	90.89	90.82

Furthermore, we perform a quantitative evaluation of the CC module on both CIFAR-10 and CIFAR-100 datasets, with the results summarized in Table IV. The empirical data reveals that the integration of the CC module leads to a significant reduction in the classification error rate, decreasing from 11.23% to 9.28% on CIFAR-10, and from 27.01% to 26.37% on CIFAR-100. This improvement demonstrates that the CC module effectively bolsters semi-supervised classification performance across datasets with varying complexities. Additionally, the full G3C framework achieves higher AUROC scores, improving from 96.81% to 97.41% on CIFAR-10 and notably increasing from 88.21% to 90.35% on CIFAR-100. This indicates that in open-set scenarios, the CC module facilitates more precise identification of anomalous points beyond the known categories, thereby enhancing the model's capability in binary discriminative tasks between known and unknown data.

B. *Parameter Sensitivity Analysis*

We further investigate the sensitivity of the hyperparameter λ_c, which governs the weight of the contrastive clustering loss in our total objective. We evaluate the model's performance on the CIFAR-100 dataset under the uncorrelated setting with 50 labeled samples by varying λ_c from 0.1 to 5. The results are reported in Table V.

As observed, increasing λ_c from 0.1 initially yields significant performance gains. It is worth noting that while $\lambda_c = 0.5$

yields the marginally highest AUROC, the corresponding error rate is notably higher compared to the setting where $\lambda_c = 1$. Given that the degradation in AUROC is negligible while the improvement in classification accuracy is substantial at $\lambda_c = 1$, we identify $\lambda_c = 1$ as the optimal equilibrium point that maximizes classification performance without compromising outlier detection capability. While the model remains relatively robust within the range of $[0.5, 2]$, setting λ_c too low fails to fully leverage the structural information, whereas setting it too high may dominate the optimization process, leading to a slight performance degradation. Consequently, we set $\lambda_c = 1$ as the default value for all main experiments to ensure optimal stability and discriminative power.

VI. Conclusion

In this paper, we addressed the limitations of existing Open-Set Semi-Supervised Learning methods, particularly their vulnerability to confirmation bias arising from prediction-dependent graph construction. We proposed a novel framework named G3C, which integrates contrastive clustering with graph consistency regularization. Instead of relying solely on potentially unreliable model predictions to build adjacency graphs, G3C introduces a dual optimization of sample similarity and global cluster structures. This strategy drives the model to learn compact semantic groupings, which in turn enables reliable dynamic neighbor searching in the augmented feature space. By iteratively refining the graph structure based on these evolved features, G3C effectively mitigates the noise propagation caused by incorrect pseudo-labels. Furthermore, to tackle the open-set challenge, we introduced an unknown-aware graph mechanism that constrains the embedding space via graph consistency regularization, ensuring a robust separation between in-distribution samples and out-of-distribution outliers. Extensive empirical evaluations on CIFAR-10, CIFAR-100, and ImageNet-30 demonstrate that G3C generates more accurate pseudo-labels and achieves superior classification accuracy and outlier detection performance compared to SOTA methods. This success validates that explicitly enforcing feature-space compactness via contrastive clustering is essential for learning robust representations from uncurated data.

VII. Acknowledgment

This work was supported in part by the National Key R&D Program of China (2023YFE0116300), in part by the National Natural Science Foundation of China (Grant No.62006097, U1836218).

References

[1] X. Gao, K. Wei, S. Niu, H. Zhao, and J. Dong, "Edlrdpl net: A new deep dictionary learning network for sar image classification," *IEEE Signal Processing Letters*, 2024.

[2] Y. Tang, J. Bi, S. Xu, L. Song, S. Liang, T. Wang, D. Zhang, J. An, J. Lin, R. Zhu *et al.*, "Video understanding with large language models: A survey," *IEEE Transactions on Circuits and Systems for Video Technology*, 2025.

[3] L. Laishram, M. Shaheryar, J. T. Lee, and S. K. Jung, "Toward a privacy-preserving face recognition system: A survey of leakages and solutions," *ACM Computing Surveys*, vol. 57, no. 6, pp. 1–38, 2025.

[4] Y. Zhao and L. Bai, "Contrastive clustering with a graph consistency constraint," *Pattern Recognition*, vol. 146, p. 110032, 2024.

[5] Z. Song, X. Yang, Z. Xu, and I. King, "Graph-based semi-supervised learning: A comprehensive review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 11, pp. 8174–8194, 2022.

[6] K. S. Miller and A. L. Bertozzi, "Model change active learning in graph-based semi-supervised learning," *Communications on Applied Mathematics and Computation*, vol. 6, no. 2, pp. 1270–1298, 2024.

[7] Y. Wei and K. Wang, "Multi-domain time-frequency fusion feature contrastive learning for machinery fault diagnosis," *IEEE Signal Processing Letters*, vol. 32, pp. 1116–1120, 2025.

[8] H. Zhu and P. Koniusz, "Transductive few-shot learning with prototype-based label propagation by iterative graph refinement," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 23 996–24 006.

[9] H. Kong, S. Kim, H.-J. Kim, and S.-W. Lee, "Unknown-aware graph regularization for robust semi-supervised learning from uncurated data," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 38, no. 12, 2024, pp. 13 265–13 273.

[10] W. Lan, G. Zhou, Q. Chen, W. Wang, S. Pan, Y. Pan, and S. Zhang, "Contrastive clustering learning for multi-behavior recommendation," *ACM Transactions on Information Systems*, vol. 43, no. 1, pp. 1–23, 2024.

[11] J. Bian, X. Xie, J.-H. Lai, and F. Nie, "Multi-view contrastive clustering via integrating graph aggregation and confidence enhancement," *Information Fusion*, vol. 108, p. 102393, 2024.

[12] K. Sohn, D. Berthelot, N. Carlini, Z. Zhang, H. Zhang, C. A. Raffel, E. D. Cubuk, A. Kurakin, and C.-L. Li, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," *Advances in neural information processing systems*, vol. 33, pp. 596–608, 2020.

[13] J. Li, C. Xiong, and S. C. Hoi, "Comatch: Semi-supervised learning with contrastive graph regularization," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 9475–9484.

[14] W. Lu, Z. Guan, W. Zhao, Y. Yang, Y. Lv, L. Xing, B. Yu, and D. Tao, "Pseudo contrastive learning for graph-based semi-supervised learning," *Neurocomputing*, vol. 624, p. 129375, 2025.

[15] J. Bi and F. Dornaika, "Sample-weighted fused graph-based semi-supervised learning on multi-view data," *Information Fusion*, vol. 104, p. 102175, 2024.

[16] L.-Z. Guo, Z.-Y. Zhang, Y. Jiang, Y.-F. Li, and Z.-H. Zhou, "Safe deep semi-supervised learning for unseen-class unlabeled data," in *International conference on machine learning*. PMLR, 2020, pp. 3897–3906.

[17] Y. Chen, X. Zhu, W. Li, and S. Gong, "Semi-supervised learning under class distribution mismatch," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 04, 2020, pp. 3569–3576.

[18] Q. Yu, D. Ikami, G. Irie, and K. Aizawa, "Multi-task curriculum framework for open-set semi-supervised learning," in *European conference on computer vision*. Springer, 2020, pp. 438–454.

[19] K. Saito, D. Kim, and K. Saenko, "Openmatch: Open-set consistency regularization for semi-supervised learning with outliers," *arXiv preprint arXiv:2105.14148*, 2021.

[20] Y.-X. Sun and W. Wang, "Exploiting mixed unlabeled data for detecting samples of seen and unseen out-of-distribution classes," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, no. 8, 2022, pp. 8386–8394.

[21] R. He, Z. Han, X. Lu, and Y. Yin, "Safe-student for safe deep semi-supervised learning with unseen-class unlabeled data," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 14 585–14 594.

[22] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang *et al.*, "A survey on evaluation of large language models," *ACM transactions on intelligent systems and technology*, vol. 15, no. 3, pp. 1–45, 2024.

[23] Q. Yu, D. Ikami, G. Irie, and K. Aizawa, "Multi-task curriculum framework for open-set semi-supervised learning," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16*. Springer, 2020, pp. 438–454.

[24] Y. Wang, P. Qiao, C. Liu, G. Song, X. Zheng, and J. Chen, "Out-of-distributed semantic pruning for robust semi-supervised learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 23 849–23 858.