Unknown-Aware Graph Regularization for Robust Semi-supervised Learning from Uncurated Data

Heejo Kong[1], Suneung Kim[2], Ho-Joong Kim[2], and Seong-Whan Lee[2]

[1]Department of Brain and Cognitive Engineering , Korea University [2]Dept. of Artificial Intelligence, Korea University

Abstract

Recent advances in semi-supervised learning (SSL) have relied on the optimistic assumption that labeled and unlabeled data share the same class distribution. However, this assumption is often violated in real-world scenarios, where unlabeled data may contain out-of-class samples. SSL with such uncurated unlabeled data leads training models to be corrupted. In this paper, we propose a robust SSL method for learning from uncurated real-world data within the context of open-set semi-supervised learning (OSSL). Unlike previous works that rely on feature similarity distance, our method exploits uncertainty in logits. By leveraging task-dependent predictions of logits, our method is capable of robust learning even in the presence of highly correlated outliers. Our key contribution is to present an unknown-aware graph regularization (UAG), a novel technique that enhances the performance of uncertainty-based OSSL frameworks. The technique addresses not only the conflict between training objectives for inliers and outliers but also the limitation of applying the same training rule for all outlier classes, which are existed on previous uncertainty-based approaches. Extensive experiments demonstrate that UAG surpasses state-of-the-art OSSL methods by a large margin across various protocols.

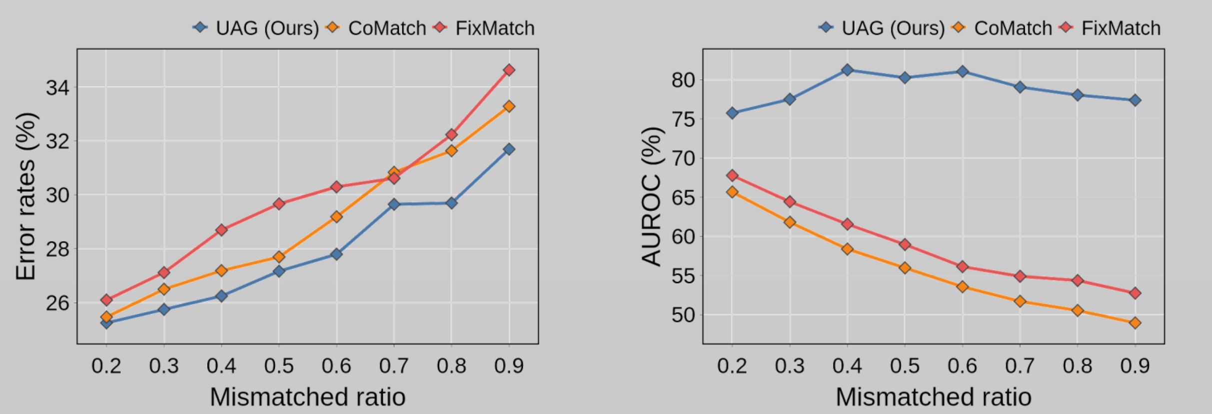
Background

A need of semi-supervised Learning (SSL)

Recent advances in deep supervised learning have been driven by the availability of large-scale annotated datasets. However, the process of constructing such training data is labor-intensive and time-consuming due to the data labeling process. As a remedy, significant efforts have been dedicated to the field of SSL. They have provided effective solutions to leverage abundant unlabeled data with only a fraction of manual annotations, and shown their promising performances.

An optimistic assumption of existing SSL algorithms

All the positive results observed in SSL are typically based on the optimistic assumption that both labeled and unlabeled data are drawn from identical class distribution. However, in practical scenarios, the unlabeled dataset often includes out-of-class data, *i.e.*, outlier, which easily violates this assumption. This uncurated unlabeled data severely degrade the performance of SSL, as shown in Figure1 below. Hence, it is desirable that the training models not only classify samples from known categories, *i.e.*, inliers, but also identify samples from novel classes as outliers. This task is known as <u>**open-set semi-supervised learning (OSSL)**</u>. While OSSL is more realistic and practical for various applications, it has been rarely considered in previous literature.



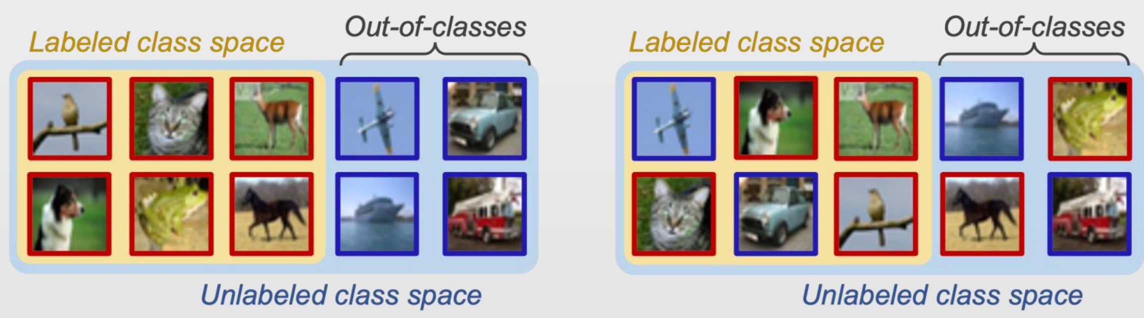
(a) Error rates (b) AUROC

Figure 1. Error rates(left) and AUROC(right) results with various mismatched ratio of outliers. All the experiments in this figure are conducted on CIFAR-100 dataset (with 100 labeled data per class) under the correlated setting. These results show that as the proportion of outliers in unlabeled data increases, the performance of existing SSL algorithms (CoMatch & FixMatch) decreases drastically.

Motivations

Similarity-based OSSL approach

The key difficulty in OSSL is the absence of supervision for distinguishing unknown samples from known ones. Recent notable approaches have tackled this problem by utilizing similarity distances in feature space. Based on intra- or inter-class distances of known categories, they aim to identify unknown samples, which deviate significantly from in-distribution (ID) data, and consider the samples as outliers. Although the studies have substantially improved the performance of OSSL, they are still quite limited for general use. As discussed in OpenMatch, similarity-based measures can fail to detect highly correlated outliers that exhibit similar visual characteristics to ID data. It is evident that this limitation is more pronounced when unknown classes share the same superclass space with inliers, defined as correlated outliers in Figure 2. In a scenario where the scale of unlabeled data is larger, the correlated outliers are more natural, but it has never been considered.

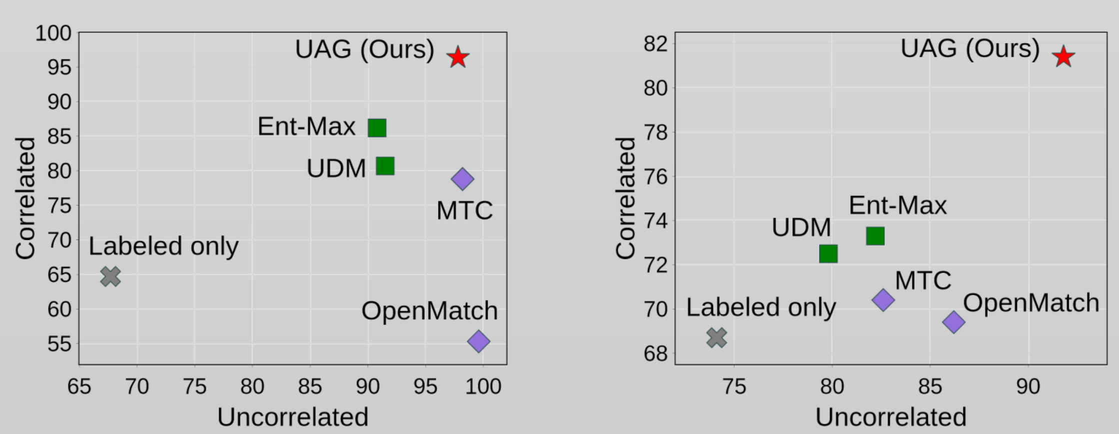


(a) Uncorrelated outliers (b) Correlated outliers

Figure 2. An illustration of the example for uncorrelated and correlated outliers. The classes marked with red and blue box share the same superclass, animal and transportation, respectively. In an uncorrelated setting, out-of-class data do not share a superclass with labeled classes, whereas in a correlated setting, they share the same superclass space.

Uncertainty-based OSSL approach

As an alternative to the feature similarity, some OSSL works uncertainty in logit space to resolve the open-set decision problem. In contrast to the features exclusively disregard the classification weights with class-dependent information, the logits effectively activate task-specific information related to high-level semantic attributes. This property enables uncertainty-based measures to serve as a robust detector for highly correlated outliers. However, as illustrated in Figure 3, the performance of uncertainty-based methods is still not competitive in practice.



(a) CIFAR-10 (b) CIFAR-100

Figure 3. AUROC of five OSSL algorithms trained on CIFAR datasets with 100 labels per class. The results are reported for two outlier settings, whether uncorrelated (x-axis) or correlated (y-axis). Methods marked with square (■) use the uncertainty in logits; methods with diamond (◆) use the similarity in features.

Reasons for under-performing of uncertainty-based approach

Aspect 1) Existing methods focus on maximizing the entropy of outliers to improve the discriminative ability for unknown contexts, while they aim to minimize the entropy of inliers for known contexts. <u>**Despite the inherent conflict between the learning objectives, they solely rely on a single classifier.**</u>

Aspect 2) Although the outliers are composed of multiple novel classes, <u>**previous works train the model by assigning them into one generic class, unknown.**</u> This approach leads to the convergence of all outliers into a single representative space, making it difficult to cope with diverse outliers.

Our Approach

Uncertainty-based outlier detection

We exploit the uncertainty of logits to construct a baseline outlier detector. Our works adopt a maximum softmax probability (MSP) as an uncertainty score, and considers the samples with the low scores as outliers.
To detect inliers and outliers with the estimated score $s(x;T)$, we utilize the thresholding, *i.e.*, the samples x is determined as inlier if $s(x;T) < \tau^{in}$ or outlier if $s(x;T) > \tau^{out}$. The thresholds are adaptively decided by two-component GMM with EMA updates

Eq 1. MSP scores

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

Eq 2. EMA thresholds

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

Unknown-Aware Graph regularization (UAG)

To address the two aspects that cause the underperforming of uncertainty-based OSSL approaches, we propose a novel OSSL algorithm, UAG. The key components of the proposed UAG consist of the followings:

- Exclusive multi-head training for addressing an aspect 1
- Contrastive graph regularization for addressing an aspect 2

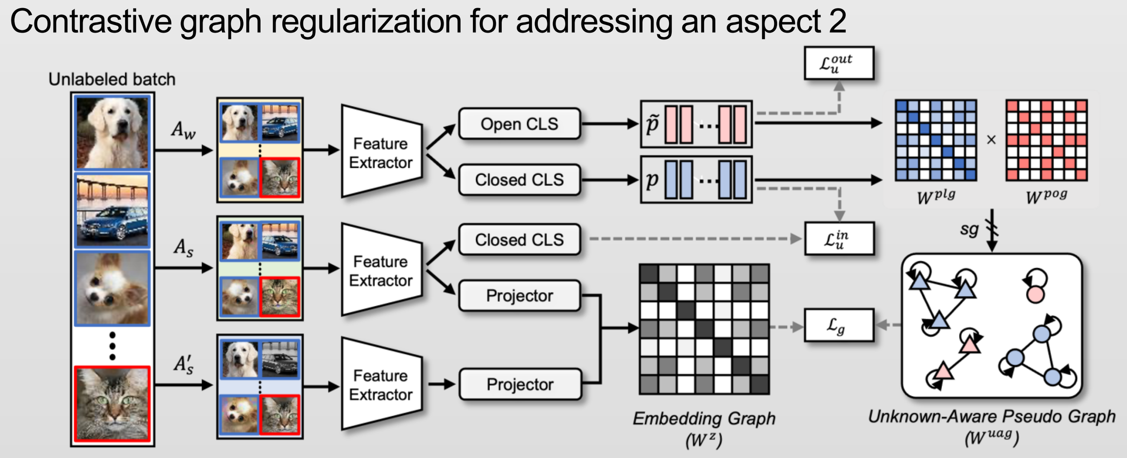


Figure 4. Overall framework of the proposed UAG.

A) Exclusive multi-head training

We adopt multi-head classifiers that exclusively learn the objectives for inliers and outliers. It inherently prevents the conflict that arises when a single classifier learns both objectives, since our strategy at least avoids that a classifier head trains to minimize the the entropy of incorrect pseudo-outliers, and vice versa.

Eq 3. Supervised loss for labeled data

$$\mathcal{L}_s = \frac{1}{B}\sum_{b=1}^{B}\left(\mathrm{H}(y_b, p(y \mid \mathrm{A}_w(x_b))) + \mathrm{H}(y_b, \tilde{p}(y \mid \mathrm{A}_w(x_b)))\right)$$

Eq 4. Unsupervised loss for unlabeled data

$$\mathcal{L}_u^{in} = \frac{1}{\mu B}\sum_{b=1}^{\mu B} \mathbb{1}\left(s_b \geq \hat{\tau}_t^{in}\right) \cdot \mathrm{H}\left(\hat{y}_b, p(y \mid \mathrm{A}_s(u_b))\right)$$
$$\mathcal{L}_u^{out} = -\frac{1}{\mu B}\sum_{b=1}^{\mu B} \mathbb{1}\left(s_b < \hat{\tau}_t^{out}\right) \cdot \mathrm{H}\left(\tilde{p}(\mathrm{A}_w(u_b))\right)$$

B) Contrastive graph regularization

Based on the empirical rationale discussed in our manuscript, we derive an unknown-aware graph regularization that allows the outliers to form multiple clusters in their embeddings. By leveraging the batch-wise predictions of closed-set and open-set classifiers, we first construct the unknown-aware pseudo graph as a target; $W^{uag} = W^{plg} \cdot W^{pog}$

Eq 5. Pseudo-label graph

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

Eq 6. Pseudo-outlier graph

$$W_{bj}^{z} = \begin{cases} \exp\left(z_b \cdot z_b' / t_e\right) & \text{if } b = j \\ \exp\left(z_b \cdot z_j / t_e\right) & \text{otherwise} \end{cases}$$

Then, the embedding graph is constructed based on the batch-wise similarity of the projected embedding. Consequently, the graph contrastive loss is derived to minimize the cross-entropy between the two normalized graphs.

Eq 7. Embedding graph

$$W_{bj}^{pog} = \begin{cases} 1 & \text{if } \eta_b = \eta_j \\ 0 & \text{otherwise} \end{cases}$$

Eq 8. Graph contrastive loss

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

C) Overall objectives

$$\mathcal{L} = \mathcal{L}_s + \lambda_u \mathcal{L}_u + \lambda_g \mathcal{L}_g$$

Experiments

Experimental results

Dataset	CIFAR-10				CIFAR-100				ImageNet-30
	Uncorr.		Corr.		Uncorr.		Corr.		
No. of labeled	50	100	50	100	50	100	50	100	10%
Labeled Only	34.3±1.2	29.4±0.8	30.9±1.3	25.8±0.7	38.9±0.8	31.7±0.6	38.1±0.7	30.7±0.4	20.2±1.2
FixMatch	16.8±1.1	10.7±0.9	17.5±0.9	12.9±0.8	33.6±0.8	29.4±0.8	30.8±0.6	28.8±0.7	12.5±0.3
CoMatch	12.7±0.7	9.5±0.5	14.8±0.8	10.3±0.4	28.5±0.6	26.4±0.6	28.8±0.7	25.8±0.5	8.8±0.9
MTC	20.4±0.9	13.5±0.8	21.8±1.2	14.3±0.6	36.7±0.9	30.9±0.5	36.9±1.2	29.6±0.6	13.6±0.7
OpenMatch	10.2±0.9	7.1±0.5	11.7±0.8	9.2±0.6	30.5±0.4	26.7±0.6	30.1±0.5	25.3±0.5	10.4±1.0
OSP	12.1±0.8	9.2±0.6	11.1±0.8	9.5±0.5	29.5±0.5	26.5±0.6	30.7±0.9	26.9±0.5	-
Ours	**9.6±0.7**	**5.8±0.4**	**8.1±0.9**	**6.8±0.5**	**26.6±0.3**	**23.6±0.2**	**26.4±0.6**	**23.8±0.4**	**6.1±0.6**

Table 1. Error rates with standard deviation for CIFAR-10/100 and ImageNet-30 on 3 different folds.

Dataset	CIFAR-10				CIFAR-100				ImageNet-30
	Uncorr.		Corr.		Uncorr.		Corr.		
No. of labeled	50	100	50	100	50	100	50	100	10%
Labeled Only	65.8±0.6	67.7±0.5	63.9±0.6	64.7±0.6	71.3±0.8	74.1±0.8	64.4±1.0	68.7±0.9	81.3±1.0
FixMatch	54.2±0.6	58.5±0.4	55.9±0.4	59.4±0.5	64.1±0.8	66.7±0.5	60.4±0.7	62.6±0.5	87.9±0.6
CoMatch	47.6±0.5	47.7±0.6	48.1±0.6	48.9±0.6	60.1±1.3	61.7±1.2	55.6±0.5	57.9±0.9	65.8±1.2
MTC	96.5±0.4	98.2±0.3	73.5±0.6	78.2±0.5	81.7±2.8	82.6±3.4	69.3±2.5	70.4±3.5	93.8±0.8
OpenMatch	**97.9±0.4**	**99.6±0.3**	75.6±0.5	55.3±1.2	86.1±1.3	86.2±2.1	69.9±0.8	69.4±0.5	96.4±0.7
OSP	62.9±0.6	66.0±0.7	45.7±0.8	46.4±0.7	58.5±1.2	60.3±1.4	56.1±0.9	59.5±0.8	-
Ours	95.5±0.4	97.8±0.5	**90.2±0.7**	**96.4±0.3**	**87.9±0.8**	**91.8±0.6**	**78.7±0.7**	**81.4±0.5**	**97.4±0.5**

Table 2. AUROC performance of Table 1.

Ablation studies

				Uncorr.		Corr.	
Ent-Max	MHT	PLG	POG	Error	AUROC	Error	AUROC
				29.4	66.7	28.8	62.6
✓				29.3	82.2	30.2	73.3
✓	✓			27.5	89.1	28.2	77.7
✓		✓		25.8	63.9	25.8	60.1
✓		✓	✓	**23.5**	72.3	**23.6**	68.2
✓	✓	✓	✓	23.6	**91.8**	23.8	**81.4**

Table 3. Ablation studies on the individual modules.

(a) FixMatch (a) w/ Ent-Max (a) w/ UAG (ours)

Table 3. t-SNE visualization of embeddings obtained from the ablated models. Pink points denote outliers, while the other colored points represent distinct known classes. (a) A model trained solely with FixMatch. (b) A model trained with entropy maximization further applied on (a). (c) A model trained using the proposed method UAG.

Conclusion

In this paper, we propose an unknown-aware graph regularization (UAG), a novel approach for open-set semi-supervised learning (OSSL). Our study focuses on highly correlated scenarios, where inlier and outlier classes share the same superclass space, presenting a more challenging and realistic OSSL benchmark. Extensive experiments demonstrate the outstanding performance of UAG across various OSSL scenarios. We believe that our work serves as a valuable baseline and will guide future research directions in both semi-supervised learning (SSL) and OSSL.

Acknowledgements

This research was supported by the Challengeable Future Defense Technology Research and Development Program (912911601) of Agency for Defense Development in 2020 and was partly supported by the Institute of Information \& Communications Technology Planning \& Evaluation (IITP) grant, funded by the Korea government (MSIT) (No. 2019-0-00079, Artificial Intelligence Graduate School Program (Korea University)).