# Multi-Task Curriculum Framework for Open-Set Semi-Supervised Learning

School of Industrial and Management Engineering, Korea University

Hansam Cho

KOREA UNIVERSITY

DMQA

hcai

Human-Centered Artificial Intelligence Center

# Contents

# Research Purpose

❖ **Multi-Task Curriculum Framework for Open-Set Semi-Supervised Learning (2020)**

- 도쿄대 연구, 38회 인용 (2022.09.30 기준)
- 학습 데이터에 out-of-distribution (OOD) 데이터가 존재하는 상황에서 semi-supervised learning (SSL) 방법론 제안

## Multi-Task Curriculum Framework for Open-Set Semi-Supervised Learning

Qing Yu[1], Daiki Ikami[1,2], Go Irie[2], and Kiyoharu Aizawa[1]

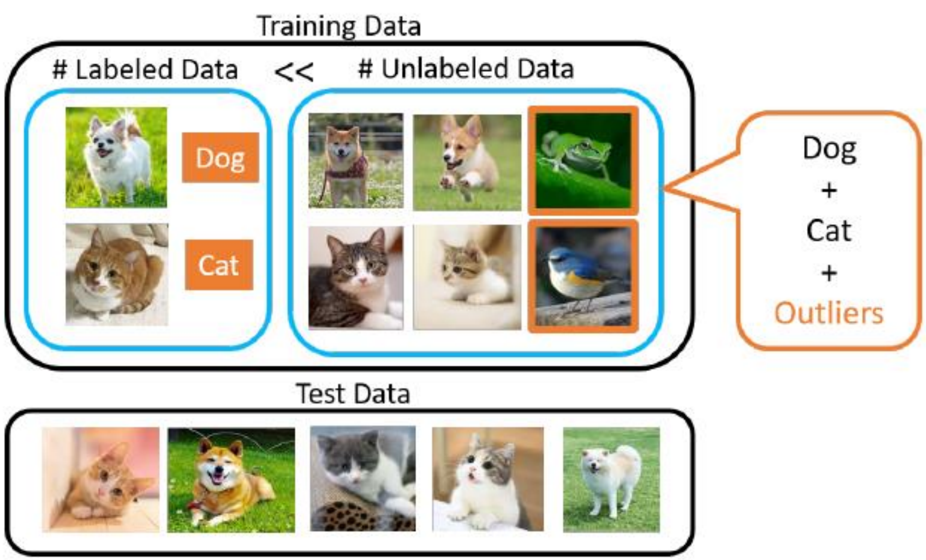
[1] The University of Tokyo, Japan
{yu,ikami,aizawa}@hal.t.u-tokyo.ac.jp
[2] NTT Corporation, Japan
goirie@ieee.org

**Abstract.** Semi-supervised learning (SSL) has been proposed to leverage unlabeled data for training powerful models when only limited labeled data is available. While existing SSL methods assume that samples in the labeled and unlabeled data share the classes of their samples, we address a more complex novel scenario named open-set SSL, where out-of-distribution (OOD) samples are contained in unlabeled data. Instead of training an OOD detector and SSL separately, we propose a multi-task curriculum learning framework. First, to detect the OOD samples in unlabeled data, we estimate the probability of the sample belonging to OOD. We use a joint optimization framework, which updates the network parameters and the OOD score alternately. Simultaneously, to achieve high performance on the classification of in-distribution (ID) data, we select ID samples in unlabeled data having small OOD scores, and use these data with labeled data for training the deep neural networks to classify ID samples in a semi-supervised manner. We conduct several experiments, and our method achieves state-of-the-art results by successfully eliminating the effect of OOD samples.

# Research Purpose

## ❖ Open-Set Semi-Supervised Learning

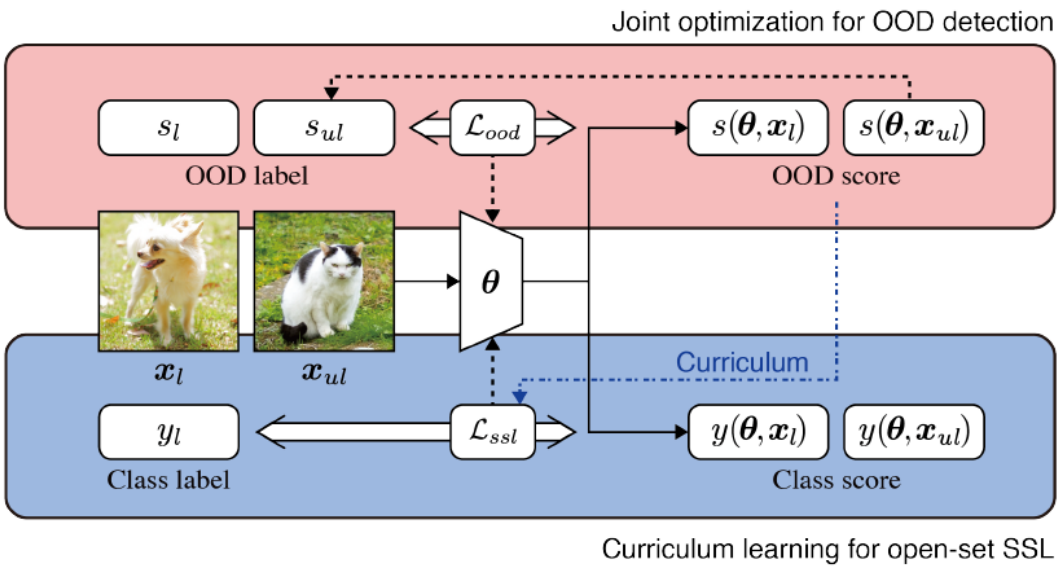
- 기존 SSL 방법론들은 labeled data와 unlabeled data의 label이 모두 동일하다고 가정 후 모델 학습
- 하지만 현실 데이터는 labeled data와 unlabeled data의 label이 일치하지 않는 경우가 많으며 이는 모델의 성능 저하로 이어짐
- OOD detection 모델과 classification 모델을 동시에 학습해서 효율적인 학습 진행



# Proposed Method

❖ Overall Concept

- 하나의 네트워크로 OOD detection과 classification을 동시에 수행 (multi-task learning)



# Proposed Method

❖ Multi-task Curriculum Learning for Open-set SSL

- Labeled data의 OOD score는 0, unlabeled data의 OOD score는 1로 시작
- Unlabeled data 중 OOD score가 일정 threshold 아래인 데이터를 활용해 SSL 사용 (Eq. 9)

**Algorithm 2** Multi-task curriculum learning

**for** $t \leftarrow 0$ to $E$ **do**
  Select ID samples in unlabeled data $X_{ul}^{id}$ by Eq. ($\boxed{9}$)
  update $\boldsymbol{\theta}^{(t+1)}$ by Adam on $\mathcal{L}_{ssl}(\boldsymbol{\theta}^t|X_l, Y_l, X_{ul}^{id}) + \mathcal{L}_{ood}(\boldsymbol{\theta}^t, S_{ul}^t|X_l, S_l, X_{ul})$
  update $S_{ul}^{(t+1)}$ by Eq. ($\boxed{6}$)
**end for**

$$X_{ul}^{id} = \{\boldsymbol{x}_{ul_i} | s(\boldsymbol{\theta}, \boldsymbol{x}_{ul_i}) < th_{otsu}, 1 \le i \le N\}, \quad (9)$$

# Proposed Method

❖ Multi-task Curriculum Learning for Open-set SSL

- OOD는 binary cross entropy loss를 활용 (Eq. 4), SSL은 MixMatch 활용
- 두 loss를 더해서 네트워크 학습 진행
- 학습된 후 unlabeled data에 대한 예측값으로 unlabeled data의 OOD score 업데이트 (Eq. 6)

**Algorithm 2** Multi-task curriculum learning

**for** $t \leftarrow 0$ to $E$ **do**
  Select ID samples in unlabeled data $X_{ul}^{id}$ by Eq. ($\boxed{9}$)
  update $\boldsymbol{\theta}^{(t+1)}$ by Adam on $\mathcal{L}_{ssl}(\boldsymbol{\theta}^t|X_l, Y_l, X_{ul}^{id}) + \mathcal{L}_{ood}(\boldsymbol{\theta}^t, S_{ul}^t|X_l, S_l, X_{ul})$
  update $S_{ul}^{(t+1)}$ by Eq. ($\boxed{6}$)
**end for**

$$\mathcal{L}_{ood} = -\frac{1}{|X_l|}\sum_{i=1}^{|X_l|}(s_{l_i}\log s(\boldsymbol{\theta}, \boldsymbol{x}_{l_i}) + (1 - s_{l_i})\log(1 - s(\boldsymbol{\theta}, \boldsymbol{x}_{l_i}))) \\ -\frac{1}{|X_{ul}|}\sum_{i=1}^{|X_{ul}|}(s_{ul_i}\log s(\boldsymbol{\theta}, \boldsymbol{x}_{ul_i}) + (1 - s_{ul_i})\log(1 - s(\boldsymbol{\theta}, \boldsymbol{x}_{ul_i}))), \tag{4}$$

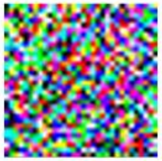
$$s_{ul_i} \leftarrow s(\boldsymbol{\theta}, \boldsymbol{x}_{ul_i}). \tag{6}$$

# Experiment

TIN

LSUN

Gaussian



Uniform

## ❖ Result

- ID datasets: CIFAR-10, SVHN
- OOD datasets: Tiny-ImageNet(TIN), LSUN, Gaussian, Uniform
- Baseline은 MixMatch
- 제안 방법론이 모든 세팅에서 baseline보다 좋은 성능

**CIFAR-10**

| OOD dataset | 250 labeled | | 1000 labeled | | 4000 labeled | |
|---|---|---|---|---|---|---|
| | Baseline | Ours | Baseline | Ours | Baseline | Ours |
| TIN | $82.42 \pm 0.70$ | $\mathbf{86.44 \pm 0.64}$ | $88.03 \pm 0.22$ | $\mathbf{89.85 \pm 0.11}$ | $91.25 \pm 0.13$ | $\mathbf{93.03 \pm 0.05}$ |
| LSUN | $76.32 \pm 4.19$ | $\mathbf{86.65 \pm 0.41}$ | $87.03 \pm 0.41$ | $\mathbf{90.19 \pm 0.47}$ | $91.18 \pm 0.33$ | $\mathbf{92.91 \pm 0.03}$ |
| Gaussian | $75.76 \pm 3.49$ | $\mathbf{87.34 \pm 0.13}$ | $85.71 \pm 1.14$ | $\mathbf{89.80 \pm 0.26}$ | $91.51 \pm 0.35$ | $\mathbf{92.53 \pm 0.08}$ |
| Uniform | $72.90 \pm 0.96$ | $\mathbf{85.54 \pm 0.11}$ | $84.49 \pm 1.06$ | $\mathbf{89.87 \pm 0.08}$ | $90.47 \pm 0.38$ | $\mathbf{92.83 \pm 0.04}$ |
| Clean | $87.65 \pm 0.29$ | | $90.67 \pm 0.29$ | | $93.30 \pm 0.10$ | |

**SVHN**

| OOD dataset | 250 labeled | | 1000 labeled | | 4000 labeled | |
|---|---|---|---|---|---|---|
| | Baseline | Ours | Baseline | Ours | Baseline | Ours |
| TIN | $94.66 \pm 0.14$ | $\mathbf{95.21 \pm 0.27}$ | $95.58 \pm 0.38$ | $\mathbf{96.65 \pm 0.14}$ | $96.73 \pm 0.05$ | $\mathbf{97.01 \pm 0.03}$ |
| LSUN | $94.98 \pm 0.23$ | $\mathbf{95.40 \pm 0.17}$ | $95.46 \pm 0.05$ | $\mathbf{96.51 \pm 0.16}$ | $96.75 \pm 0.01$ | $\mathbf{97.15 \pm 0.02}$ |
| Gaussian | $93.42 \pm 1.09$ | $\mathbf{95.23 \pm 0.04}$ | $95.85 \pm 0.33$ | $\mathbf{96.50 \pm 0.11}$ | $96.97 \pm 0.02$ | $\mathbf{97.07 \pm 0.07}$ |
| Uniform | $94.78 \pm 0.25$ | $\mathbf{95.07 \pm 0.12}$ | $95.62 \pm 0.50$ | $\mathbf{96.47 \pm 0.24}$ | $96.86 \pm 0.12$ | $\mathbf{97.04 \pm 0.02}$ |
| Clean | $96.04 \pm 0.39$ | | $96.84 \pm 0.06$ | | $97.23 \pm 0.05$ | |

# Experiment

## ❖ Result

- Labeled data 개수가 많아질수록 성능이 증가
- Synthetic data(Gaussian, Uniform)가 OOD일 경우 SSL의 성능 하락에 더 큰 영향
- CIFAR-10보다 SVHN이 쉬운 task여서 OOD 데이터에 따른 성능 하락에 덜 영향을 받음

**CIFAR-10**

| OOD dataset | 250 labeled Baseline | 250 labeled Ours | 1000 labeled Baseline | 1000 labeled Ours | 4000 labeled Baseline | 4000 labeled Ours |
|---|---|---|---|---|---|---|
| TIN | 82.42 ± 0.70 | **86.44 ± 0.64** | 88.03 ± 0.22 | **89.85 ± 0.11** | 91.25 ± 0.13 | **93.03 ± 0.05** |
| LSUN | 76.32 ± 4.19 | **86.65 ± 0.41** | 87.03 ± 0.41 | **90.19 ± 0.47** | 91.18 ± 0.33 | **92.91 ± 0.03** |
| Gaussian | 75.76 ± 3.49 | **87.34 ± 0.13** | 85.71 ± 1.14 | **89.80 ± 0.26** | 91.51 ± 0.35 | **92.53 ± 0.08** |
| Uniform | 72.90 ± 0.96 | **85.54 ± 0.11** | 84.49 ± 1.06 | **89.87 ± 0.08** | 90.47 ± 0.38 | **92.83 ± 0.04** |
| Clean | 87.65 ± 0.29 | | 90.67 ± 0.29 | | 93.30 ± 0.10 | |

**SVHN**

| OOD dataset | 250 labeled Baseline | 250 labeled Ours | 1000 labeled Baseline | 1000 labeled Ours | 4000 labeled Baseline | 4000 labeled Ours |
|---|---|---|---|---|---|---|
| TIN | 94.66 ± 0.14 | **95.21 ± 0.27** | 95.58 ± 0.38 | **96.65 ± 0.14** | 96.73 ± 0.05 | **97.01 ± 0.03** |
| LSUN | 94.98 ± 0.23 | **95.40 ± 0.17** | 95.46 ± 0.05 | **96.51 ± 0.16** | 96.75 ± 0.01 | **97.15 ± 0.02** |
| Gaussian | 93.42 ± 1.09 | **95.23 ± 0.04** | 95.85 ± 0.33 | **96.50 ± 0.11** | 96.97 ± 0.02 | **97.07 ± 0.07** |
| Uniform | 94.78 ± 0.25 | **95.07 ± 0.12** | 95.62 ± 0.50 | **96.47 ± 0.24** | 96.86 ± 0.12 | **97.04 ± 0.02** |
| Clean | 96.04 ± 0.39 | | 96.84 ± 0.06 | | 97.23 ± 0.05 | |

# Experiment

## ❖ Result

- OOD data가 많아질수록 성능 하락하지만 제안방법론은 base line보다 강건한 성능

**Table 4.** Accuracy (%) for CIFAR-10 as ID and LSUN as OOD on different numbers of OOD samples when 250 labeled samples are used.

| #OOD samples | 2000 | | 5000 | | 10000 | | 20000 | |
|---|---|---|---|---|---|---|---|---|
| | Baseline | Ours | Baseline | Ours | Baseline | Ours | Baseline | Ours |
| Accuracy (%) | **88.16** | 84.82 | 82.98 | **86.20** | 76.32 | **86.65** | 70.3 | **85.83** |

- 기존 OOD detection 방법론 대비 좋은 성능을 보임

**Table 5.** The comparison of AUROC (%) in the task of OOD Detection.

| ID dataset | OOD dataset | Hendrycks & Gimpel [9] | ODIN [17] | Ours |
|---|---|---|---|---|
| CIFAR-10 | TIN | 50.92 | 54.54 | **98.86** |
| | LSUN | 54.34 | 58.02 | **99.82** |
| | Gaussian | 32.41 | 37.49 | **100.00** |
| | Uniform | 45.43 | 51.05 | **100.00** |
| SVHN | TIN | 50.48 | 57.09 | **99.57** |
| | LSUN | 51.44 | 53.68 | **99.84** |
| | Gaussian | 21.20 | 1.87 | **99.98** |
| | Uniform | 2.79 | 8.31 | **99.97** |

# Conclusion

❖ Conclusion

- Unlabeled data에 mismatch data가 존재하는 Open-Set SSL을 처음으로 제안
- OOD detection과 SSL을 동시에 수행하는 Multi-task 방법론 제안
- OOD score를 활용해 SSL에 활용할 unlabeled data를 선정
- Mismatch data가 존재하는 상황에서 기존 SSL 방법론의 성능이 하락함을 보이고 제안 방법론의 성능은 강건하게 유지됨을 실험적으로 보임

Thank You