# CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

- ## Problem / objective

- Label ambiguity issue in Partial-Label Learning

- ## Contribution / Key idea

- Propose novel partial-label learning method called CroSel

CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Partial Label 이란?

- Candidate label set
- 각 이미지마다 candidate label set 할당
- Candidate label set 의 구성 : 하나의 true label + 여러 negative labels
- 주의 : Candidate label set 에는 true label 이 반드시 포함되어 있다는 전제.

## Partial Label Learning 이란?

- Task : Multi-class classification
- 학습 목표 : 각 이미지마다 주어진 candidate label set 에서 true label 찾아내는 능력을 학습

전유진

## CroSel: Cross Selection of Confident Pseudo Labels  for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

### Ours 등장 배경

Partial Label Learning 의 한계 : Negative labels 들이 multi-class classifier 의 학습을 방해

Ours :

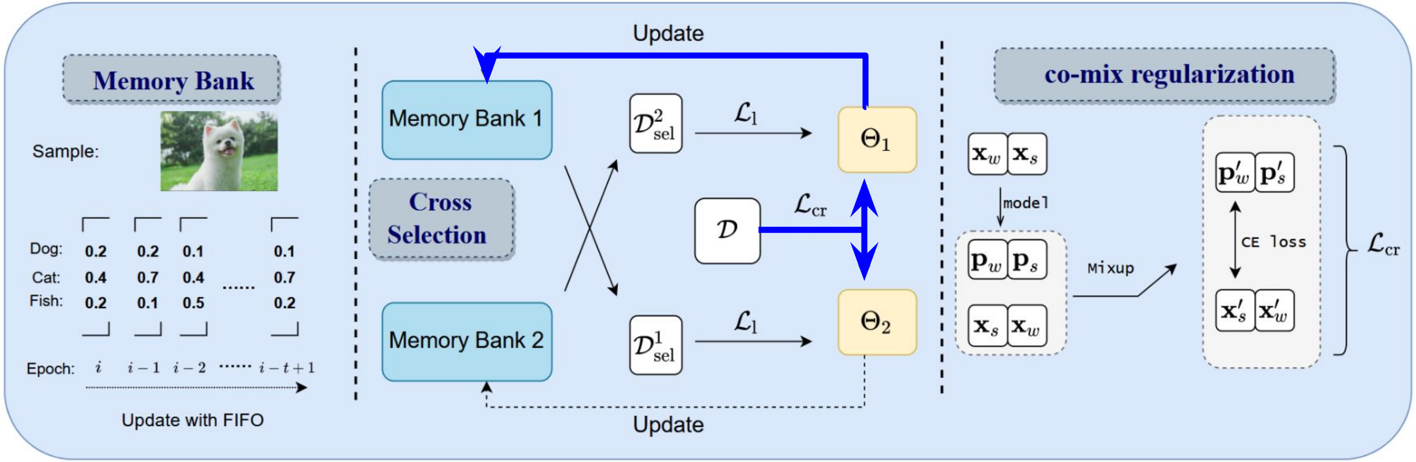
- Candidate label set 에서 true label 찾아내서, 모델 학습 중에 다른 negative labels 의 방해를 최소화시키겠다.
- 선택한 true labels 들을 사용하여 supervised learning 을 하겠다.

# CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Warm up




- 두 모델에 대하여, 10 에포크만큼 학습 및 메모리뱅크 업데이트
- 학습 목표 : 이미지의 true label 은 candidate label set 안에 반드시 존재한다.

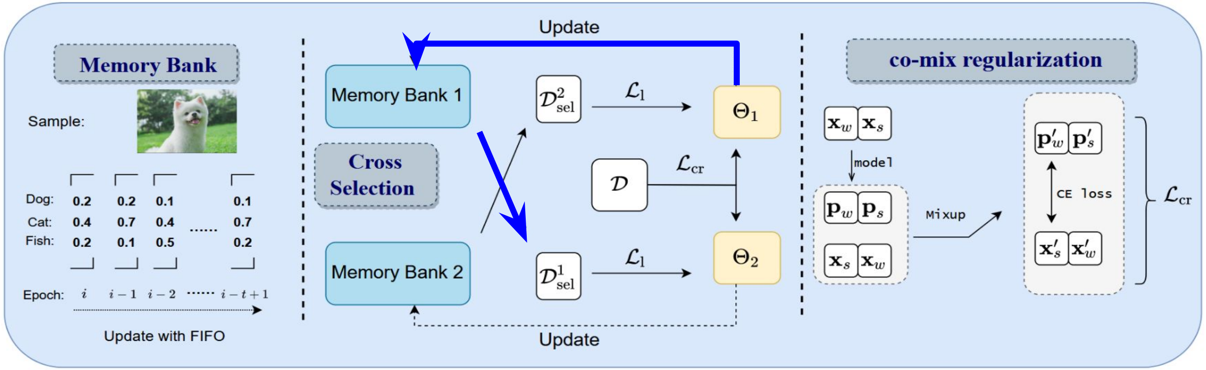
전유진

# [CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning](#)

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## High-confident data selection




$$\beta_1 = \mathbb{I}(\mathrm{argmax}(\boldsymbol{q}^i) \in S), \quad (1)$$

$$\beta_2 = \mathbb{I}(\mathrm{argmax}(\boldsymbol{q}^i) = \mathrm{argmax}(\boldsymbol{q}^{i+1})), \quad (2)$$

$$\beta_3 = \mathbb{I}(\frac{1}{t}\sum_{i=1}^{t} \max(\boldsymbol{q}^i) > \gamma), \quad (3)$$

High-confident pseudo label 선택 기준 3가지

$$\mathcal{D}_{\mathrm{sel}} = ((\boldsymbol{x}_i, \mathrm{argmax}(\boldsymbol{q}_i^t))|(\beta_1^i \wedge \beta_2^i \wedge \beta_3^i) = 1, \boldsymbol{x}_i \in \mathcal{D}), \quad (4)$$

: 위 3가지 기준 모두 만족해야 선택.

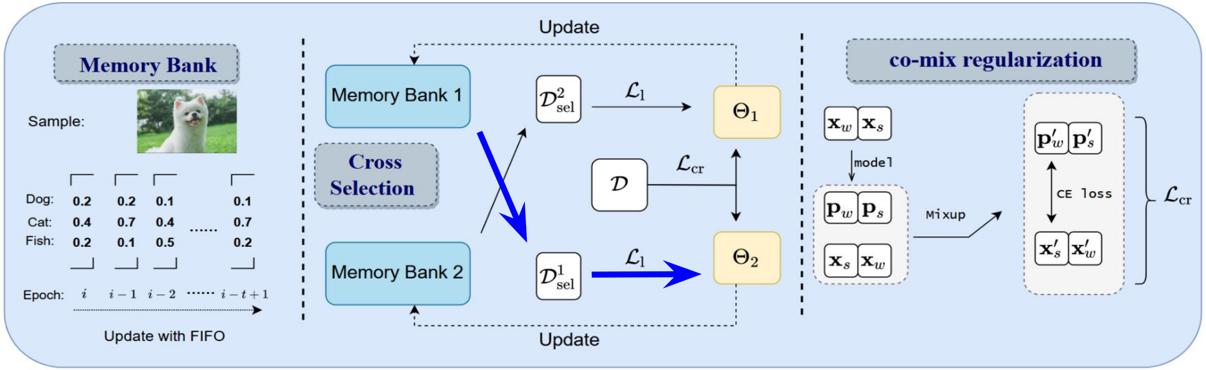
즉, 위 3가지 기준 모두 만족하면 clean label 로 간주. 전유진

CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Cross-supervised training




1. Cross selection : 모델 1 이 선택한 데이터로 모델 2 학습. 마찬가지로, 모델 2 가 선택한 데이터로 모델 1 학습.

- Cross selection 이유 : 두 모델이 서로 다른 decision boundary 를 만들어서, 앞서 선택했던 confident pseudo labels 에서 noisy label 이 있다면 이를 바로 잡아주기 위함.

2. Selected label loss

$$\mathcal{L}_1 = \frac{1}{|\mathcal{D}_{\text{sel}}|}\sum_{\boldsymbol{x}\in\mathcal{D}_{\text{sel}}} \mathcal{L}_{\text{CE}}(f(\boldsymbol{x}_{\text{w}}), \hat{y}), \qquad (5)$$

$\hat{y} = \operatorname{argmax}(\boldsymbol{q}^t)$ : selected pseudo-label

전유진

## CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

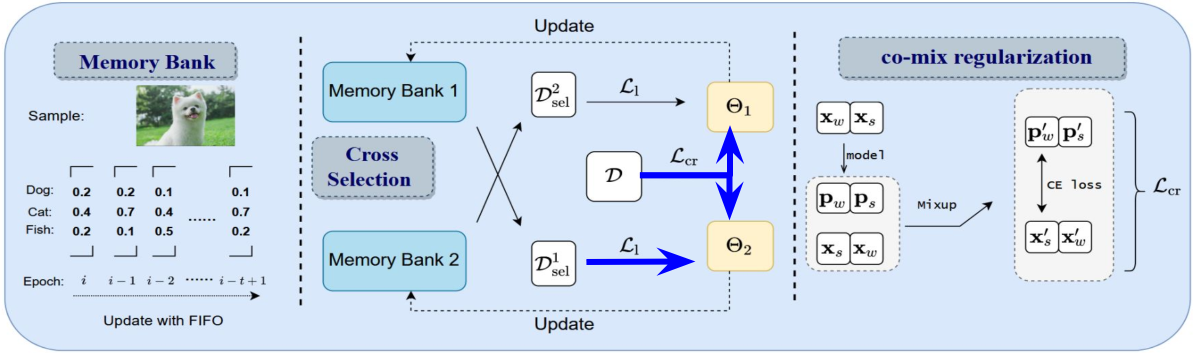
### Co-mix Consistency Regulation

1. pseudo-label 생성

$$\boldsymbol{p}_i = \begin{cases} \dfrac{\exp(f_i(\boldsymbol{x})^{\frac{1}{T}})}{\sum_{i \in S} \exp(f_i(\boldsymbol{x})^{\frac{1}{T}})}, & i \in S, \\ 0, & i \notin S, \end{cases} \tag{7}$$

2. MixUp

$$\lambda \sim \mathrm{Beta}(\alpha, \alpha), \tag{8}$$
$$\lambda' = \max(\lambda, 1 - \lambda), \tag{9}$$
$$\boldsymbol{x}' = \lambda' \boldsymbol{x}_1 + (1 - \lambda') \boldsymbol{x}_2, \tag{10}$$
$$\boldsymbol{p}' = \lambda' \boldsymbol{p}_1 + (1 - \lambda') \boldsymbol{p}_2, \tag{11}$$




3. Consistency regulation loss

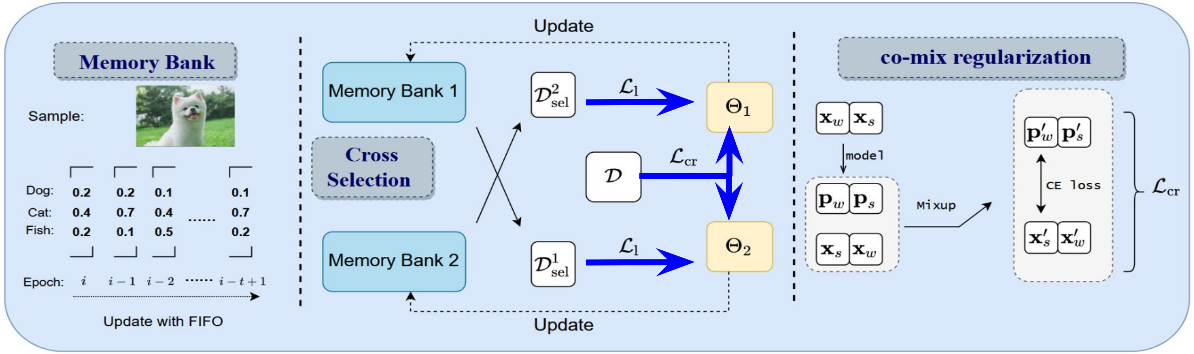
$$\mathcal{L}_{\text{cr}} = \frac{1}{2n} \sum_{i=1}^{2n} \mathcal{L}_{\text{CE}}(f(\boldsymbol{x}'_i), \boldsymbol{p}'_i), \tag{12}$$

전유진

CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Overall loss




$$\mathcal{L}_{\text{all}} = \mathcal{L}_{\text{l}} + \lambda_{\text{d}} * \mathcal{L}_{\text{cr}}, \tag{13}$$

$$\lambda_{\text{d}} = (1 - r_{\text{s}}) * \lambda_{\text{cr}}, \tag{14}$$

$$\mathcal{L}_{\text{l}} = \frac{1}{|\mathcal{D}_{\text{sel}}|}\sum_{\boldsymbol{x}\in\mathcal{D}_{\text{sel}}} \mathcal{L}_{\text{CE}}(f(\boldsymbol{x}_{\text{w}}), \hat{y}), \tag{5}$$

$$\mathcal{L}_{\text{cr}} = \frac{1}{2n}\sum_{i=1}^{2n} \mathcal{L}_{\text{CE}}(f(\boldsymbol{x}'_i), \boldsymbol{p}'_i), \tag{12}$$

전유진

CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Experiments

Table 1. Accuracy (mean±std) comparisons on benchmark datasets.

| Dataset | $q$ | Ours | PoP | CRDPLL | PiCO | PRODEN | LWS | CC |
|---|---|---|---|---|---|---|---|---|
| CIFAR-10 | 0.1 | 97.31±.04% | 97.17±.01% | **97.41±.06%** | 96.10±.06% | 95.66±.08% | 91.20±.07% | 90.73±.10% |
| | 0.3 | **97.50±.05%** | 97.08±.01% | 97.38±.04% | 95.74±.10% | 95.21±.07% | 89.20±.09% | 88.04±.06% |
| | 0.5 | **97.34±.05%** | 96.66±.03% | 96.76±.05% | 95.32±.12% | 94.55±.13% | 80.23±.21% | 81.01±.38% |
| SVHN | 0.1 | **97.71±.05%** | 97.55±.06% | 97.63±.06% | 96.58±.04% | 96.20±.07% | 96.42±.09% | 96.99±.17% |
| | 0.3 | **97.96±.05%** | 97.50±.03% | 97.65±.07% | 96.32±.09% | 96.11±.05% | 96.15±.08% | 96.67±.20% |
| | 0.5 | **97.86±.06%** | 97.31±.01% | 97.70±.05% | 95.78±.05% | 95.97±.03% | 95.79±.05% | 95.83±.23% |
| CIFAR-100 | 0.01 | **84.24±.09%** | 83.03±.04% | 82.95±.10% | 74.89±.11% | 72.24±.12% | 62.03±.21% | 66.91±.24% |
| | 0.05 | **83.92±.24%** | 82.79±.02% | 82.38±.09% | 73.26±.09% | 70.03±.18% | 57.10±.17% | 64.51±.37% |
| | 0.10 | **84.07±.16%** | 82.39±.04% | 82.15±.20% | 70.03±.10% | 69.82±.11% | 52.60±.54% | 61.50±.36% |

$q = P(\overline{y} \in S | \overline{y} \neq y)$ : candidate label set 구성 기준. 결국, noise magnitude

전유진

# CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning



*Shiyu Tian1 Hongxin Wei2 Yiqun War ... of Science and Technology*
*3Singapore University of Technology ...*

## Experiments

Table 2. Selection ratio and selection accuracy (mean±std) on benchmark datasets. S-ratio represents the selection ratio and S-acc represents selection accuracy in $\mathcal{D}_{\text{sel}}$.

| Datasets | Setting | Index | Performance |
|---|---|---|---|
| CIFAR-10 | $q = 0.1$ | S-ratio | 99.09±.07% |
| | | S-acc | 99.79±.05% |
| | $q = 0.3$ | S-ratio | 98.10±.10% |
| | | S-acc | 99.55±.03% |
| | $q = 0.5$ | S-ratio | 96.25±.12% |
| | | S-acc | 99.44±.06% |
| SVHN | $q = 0.1$ | S-ratio | 97.25±.14% |
| | | S-acc | 99.84±.06% |
| | $q = 0.3$ | S-ratio | 76.42±.21% |
| | | S-acc | 99.77±.06% |
| | $q = 0.5$ | S-ratio | 73.21±.15% |
| | | S-acc | 99.34±.02% |
| CIFAR-100 | $q = 0.01$ | S-ratio | 96.58±.13% |
| | | S-acc | 99.71±.06% |
| | $q = 0.05$ | S-ratio | 95.45±.21% |
| | | S-acc | 98.29±.15% |
| | $q = 0.10$ | S-ratio | 93.61±.12% |
| | | S-acc | 97.93±.11% |

전유진

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Experiments

Table 3. Results on CIFAR-100 in fine-grained settings.

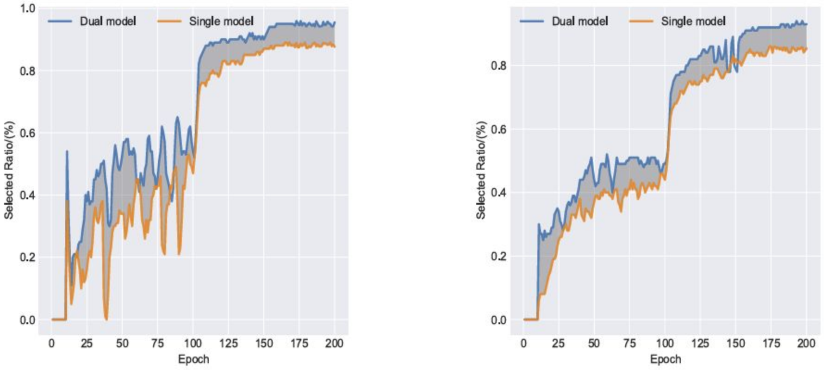
| Method | Accuracy | Method | Accuracy |
|---|---|---|---|
| PoP | 82.04% | CRDPLL | 81.53% |
| PiCO | 73.38% | PRODEN | 71.16% |
| LWS | 54.08% | CC | 64.91% |
| Crosel (ours) | | **83.34%** | |

# CroSel: Cross Selection of Confident Pseudo Labels for Partial-Label Learning

CVPR 2024

*Shiyu Tian1 Hongxin Wei2 Yiqun Wang1 Lei Feng3* 1Chongqing University 2Southern University of Science and Technology 3Singapore University of Technology and Design*

## Experiments




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

Figure 3. Selection ratio comparison between dual model and single model on CIAFR-10 and CIAFR-100.

전유진