Semantic Segmentation with Active Semi-Supervised Learning

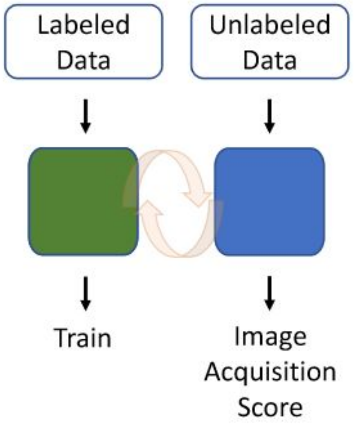
WACV 2023

Aneesh Rangnekar, Christopher Kanan, Matthew Hoffman

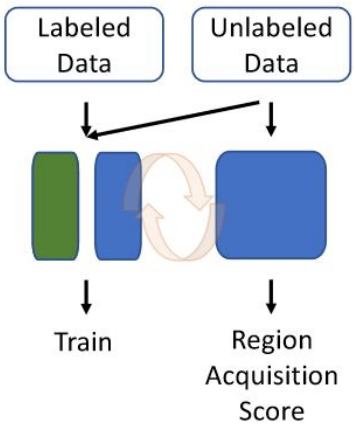
- **Problem/Objective**
  - Desire to reduce labeling costs in semantic segmentation
- **Contribution/Key Idea**
  - Active learning
  - Semi-supervised learning

전유진

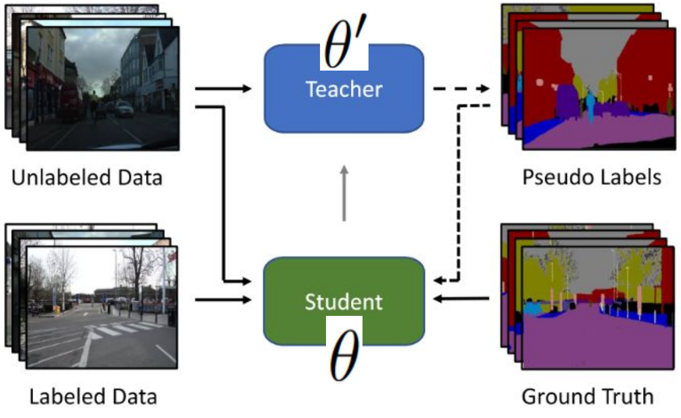
- **Active learning**



- **Active learning + Semi-supervised learning**



- Teacher-Student Framework



$$\theta' = m\theta' + (1-m)\theta$$

Supervised loss for labeled images

$$\mathcal{L}_{sup} = \ell_{ce}(\theta(x_l), y_l)$$

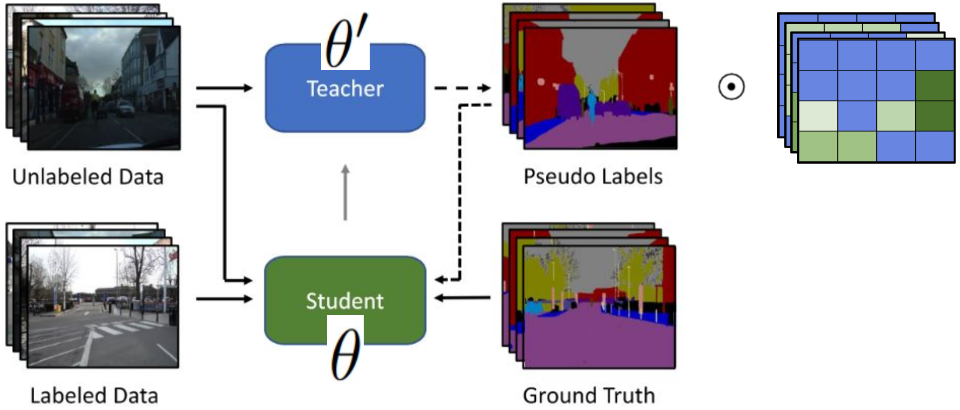
Unsupervised loss for unlabeled images

$$\mathcal{L}_{unsup} = \ell_{ce}(\theta(x_{u-s}), [\theta'(x_{u-w})])$$

Final loss for training

$$\mathcal{L}_{total} = \mathcal{L}_{sup} + \eta \cdot \mathcal{L}_{unsup}$$

- Confidence Weighting



Supervised loss for labeled images

$$\mathcal{L}_{sup} = \ell_{ce}(\theta(x_l), y_l)$$

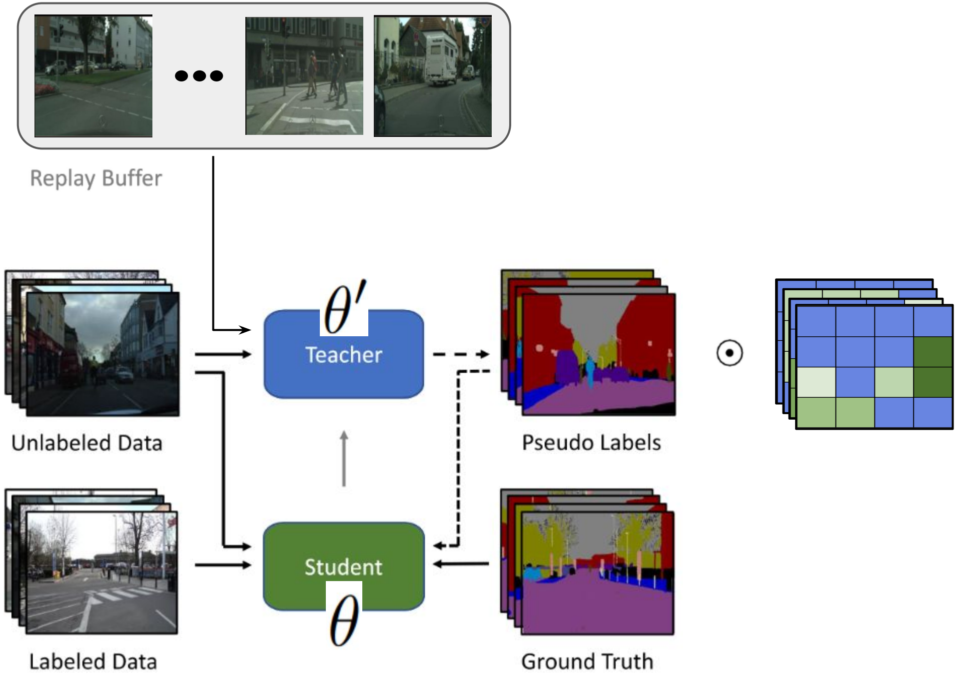
Unsupervised loss for unlabeled images

$$\mathcal{L}_{unsup} = \ell_{ce}(\theta(x_{u-s}), \underline{p \cdot [\theta'(x_{u-w})]})$$

Final loss for training

$$\mathcal{L}_{total} = \mathcal{L}_{sup} + \eta \cdot \mathcal{L}_{unsup}$$

- Balanced ClassMix



Supervised loss for labeled images

$$\mathcal{L}_{sup} = \ell_{ce}(\theta(x_l), y_l)$$

Unsupervised loss for unlabeled images

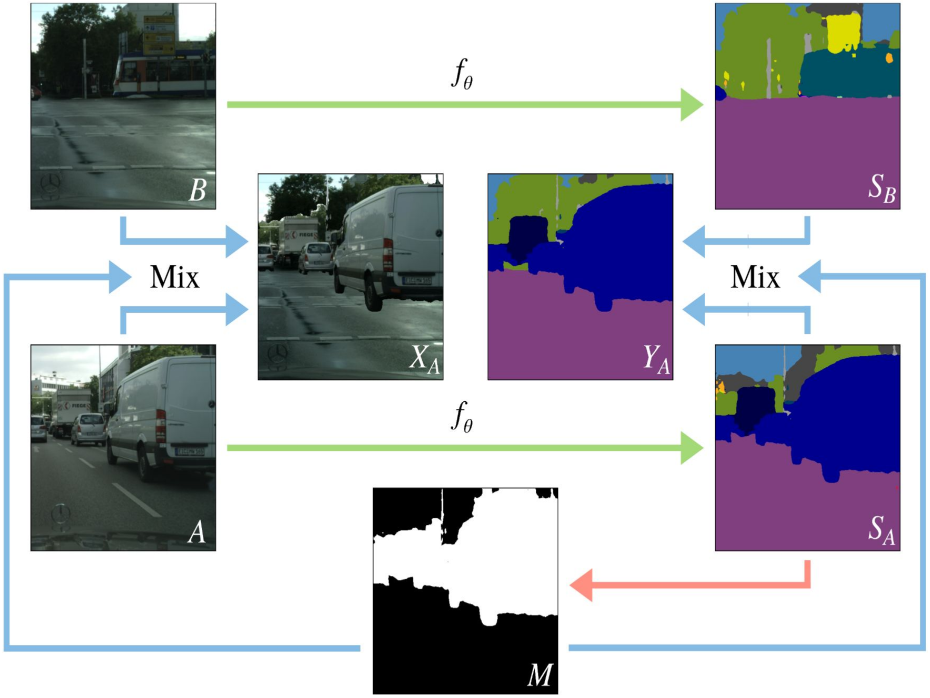
$$\mathcal{L}_{unsup} = \ell_{ce}(\theta(x_{u-s}), p \cdot [\theta'(x_{u-w})])$$

Final loss for training

$$\mathcal{L}_{total} = \mathcal{L}_{sup} + \eta_1 \cdot \mathcal{L}_{unsup1} + \underline{\eta_2 \cdot \mathcal{L}_{unsup2}}$$

- ClassMix

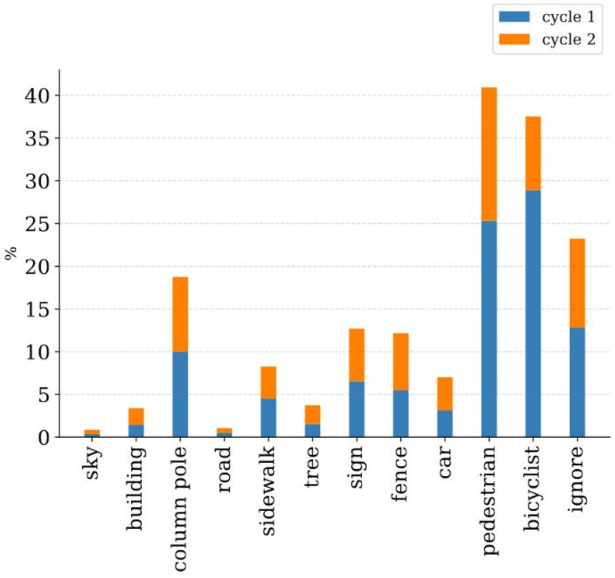
Data augmentation 기법.



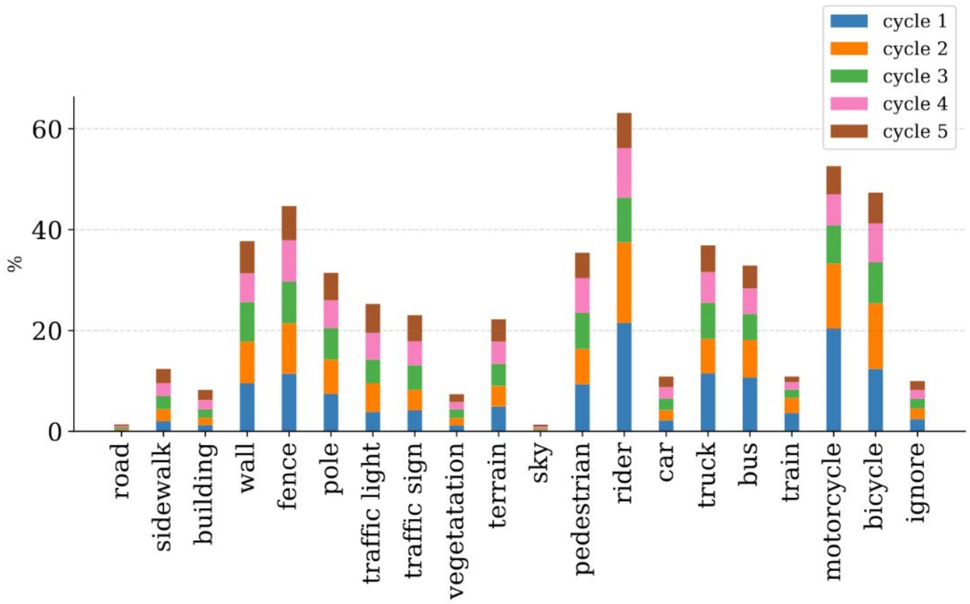
- Sampling Strategy

1. Random sampling
2. Least confidence : 모델의 prediction에서 probability 값이 낮은 샘플부터 선택하는 방식.
3. Softmax entropy : 모델의 prediction distribution에서 entropy 값이 높은 샘플부터 선택하는 방식.
4. Softmax margin : 모델의 prediction distribution에서 margin 값 (= probability 값이 가장 높은 두 클래스 간의 차이) 이 작은 샘플부터 선택하는 방식.

- Experiment



(a) CamVid

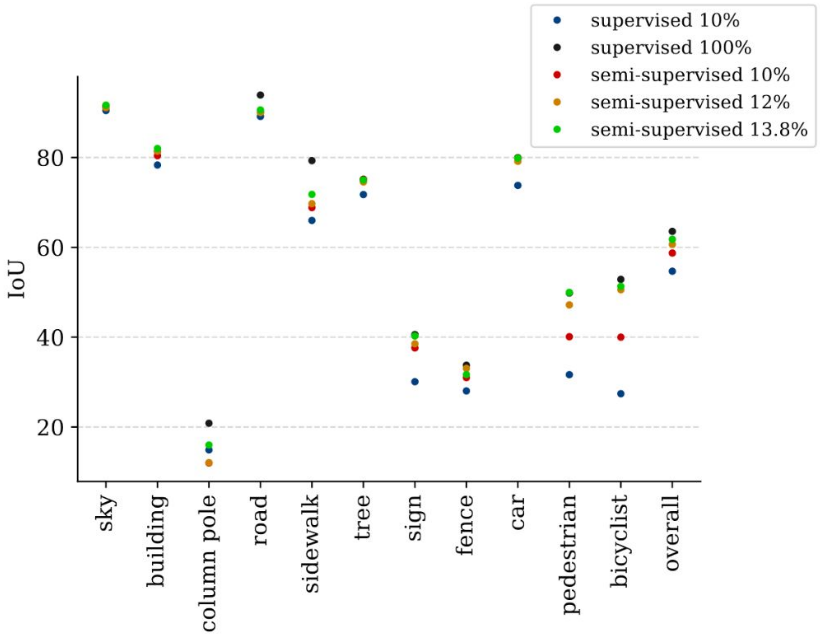


(b) CityScapes

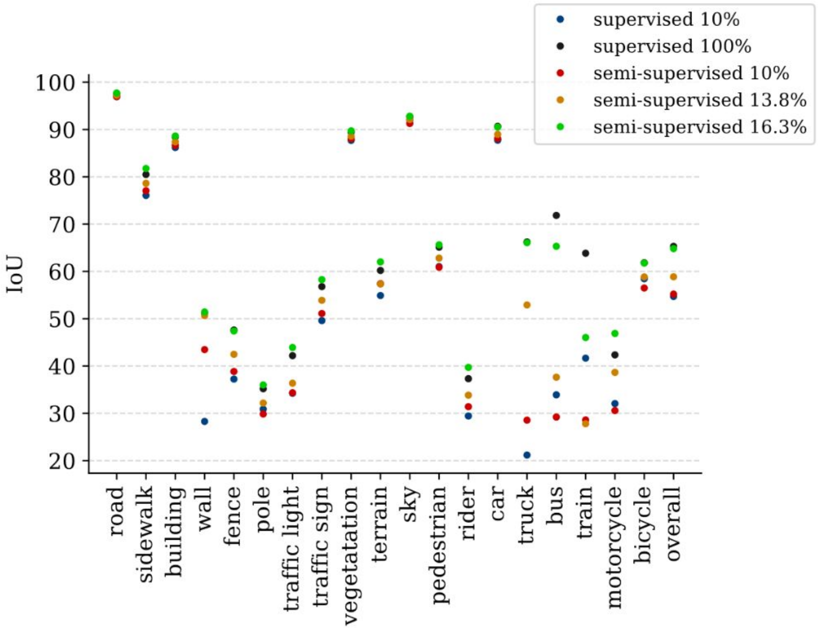
- Experiment

Method	Road	Side walk	Building	Wall	Fence	Pole	Traffic Light	Traffic Sign	Vegetation	Terrain
Supervised	97.58	80.55	88.43	51.22	47.61	35.19	42.19	56.79	89.41	60.22
Random	96.03	72.36	86.79	43.56	44.22	36.99	35.28	53.87	86.91	54.58
Entropy	96.28	73.31	87.13	43.82	43.87	38.10	37.74	55.39	87.52	53.68
Core-Set [71]	96.12	72.76	87.03	44.86	45.86	35.84	34.81	53.07	87.18	53.49
DEAL [84]	95.89	71.69	87.09	45.61	44.94	38.29	36.51	55.47	87.53	56.90
S4AL	97.73	81.76	88.63	51.42	47.40	36.00	43.91	58.27	89.72	62.01
	Sky	Pedes-trian	Rider	Car	Truck	Bus	Train	Motor-Cycle	Bicycle	mIoU
Supervised	92.69	65.12	37.32	90.67	66.24	71.84	63.84	42.35	61.84	65.30
Random	91.47	62.74	37.51	88.05	56.64	61.00	43.69	30.58	55.67	59.00
Entropy	92.05	63.96	34.44	88.38	59.38	64.64	50.80	36.13	57.10	61.46
Core-Set [71]	91.89	62.48	36.28	87.63	57.25	67.02	56.59	29.34	53.56	60.69
DEAL [84]	91.78	64.25	39.77	88.11	56.87	64.46	50.39	38.92	56.59	61.64
S4AL	92.81	65.62	39.71	90.52	66.07	65.31	46.03	46.88	61.77	64.80

- Experiment



(a) CamVid



(b) CityScapes

- Experiment

(a) **IoU:** with respect to different block sampling ratios on CamVid and CityScapes datasets.

CamVid	mIoU	CityScapes	mIoU
$30 \times 30 \times 2$	60.4 ± 1.4	$43 \times 43 \times 2$	61.8 ± 0.8
$30 \times 30 \times 4$	61.4 ± 0.6	$43 \times 43 \times 4$	62.6 ± 2.2
$60 \times 60 \times 1$	60.8 ± 2.4	$86 \times 86 \times 1$	61.4 ± 1.6

(b) **IoU:** with respect to different sampling schemes on CamVid and CityScapes datasets.

	mIoU	
	CamVid	CityScapes
Random	59.1 ± 1.8	59.8 ± 2.5
LS	60.5 ± 0.5	60.3 ± 1.4
Ent	61.2 ± 0.5	62.5 ± 1.8
Margin	60.8 ± 0.6	61.8 ± 0.5