Adaptive Superpixel for Active Learning in Semantic Segmentation

ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

- Problem / objective
  - Difficulty in obtaining label for semantic segmentation
- Contribution / Key idea
  - Superpixel-based active learning framework

[Adaptive Superpixel for Active Learning in Semantic Segmentation](#)

ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Overview

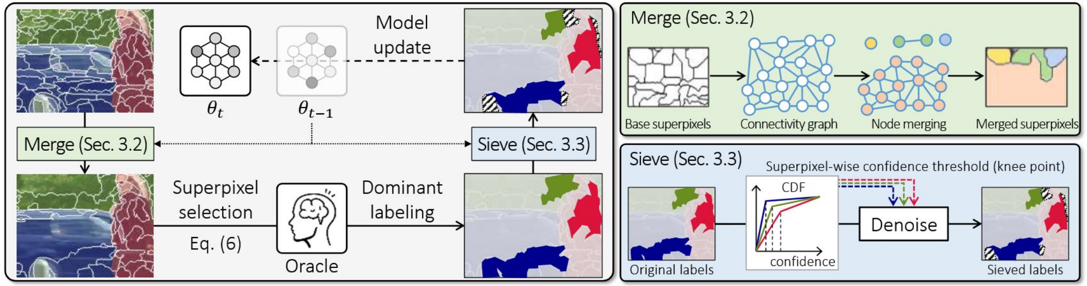


Figure 2: *An overview of the proposed framework.* In each round t, we merge superpixels with a graph using the latest model, and obtain dominant labels for selected superpixels. The dominant labels are selectively propagated to pixels with confidence above the detected knee point, resulting in the creation of a sieved dataset. Finally, we train a model with the sieved one.

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

Produce base superpixels $\mathcal{S}_0 := \bigcup_{i \in \mathcal{I}} \mathcal{S}_0(i)$
Obtain model θ_0 training with $\mathcal{D}_0$
for $t = 1, 2, \ldots, T$ **do**
    Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i \in \mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$
    Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)
    Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)
    Obtain model θ_t training with the sieved $\mathcal{D}_t$
return θ_T

전유진

Adaptive Superpixel for Active Learning in Semantic Segmentation

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

Produce base superpixels $\mathcal{S}_0 := \bigcup_{i\in\mathcal{I}} \mathcal{S}_0(i)$

Obtain model θ_0 training with $\mathcal{D}_0$

for $t = 1, 2, \ldots, T$ **do**

  Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i\in\mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$

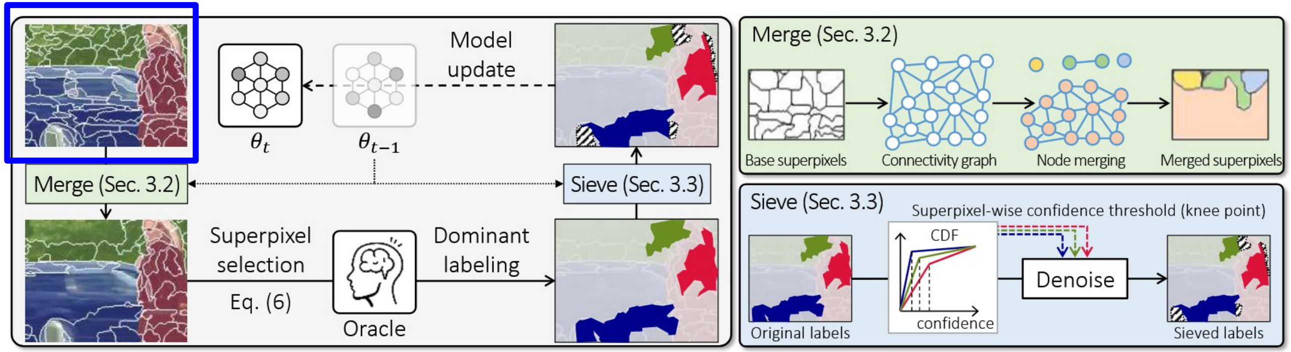
  Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)

  Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)

  Obtain model θ_t training with the sieved $\mathcal{D}_t$

return θ_T

Warm-up



1. 각 이미지에서 슈퍼픽셀 생성 (SEEDS 알고리즘 사용)

$$\text{Image } i \in \mathcal{I} \longrightarrow \text{Superpixels } S_0(i)$$

$$\text{Base superpixels } \mathcal{S}_0 := \bigcup_{i\in\mathcal{I}} S_0(i)$$

2. 슈퍼픽셀들 중 B개 랜덤 선택하여 레이블 사람에게 요청

3. 모델 학습

$$\hat{\mathbb{E}}_{(x,y)\sim\mathcal{D}_0}[\mathrm{CE}(y, f_\theta(x))] \;, \qquad (1)$$

$$\mathcal{D}_0 := \{(x,y) : \exists s \in \mathcal{B}_0, x \in s, y(c) = \mathbb{1}[c = \mathrm{D}(s)], \forall c \in \mathcal{C}\}$$

전유진

Adaptive Superpixel for Active Learning in Semantic Segmentation

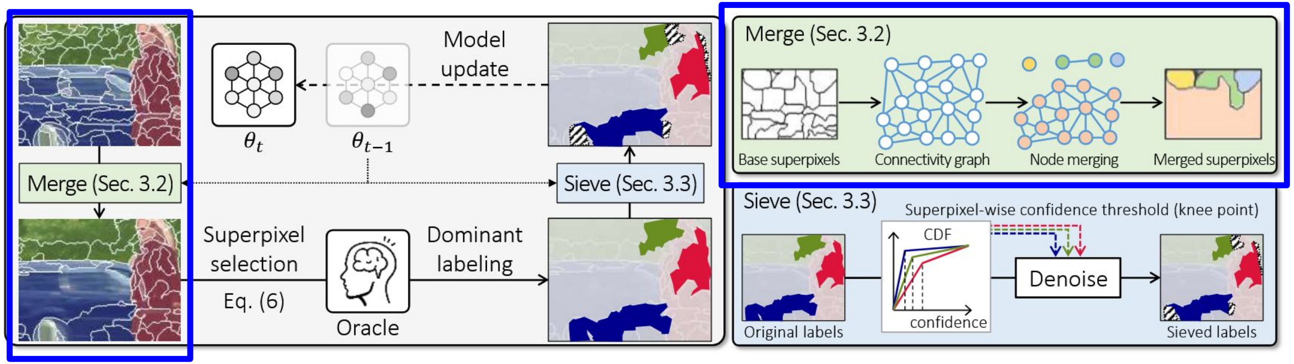
Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

Produce base superpixels $\mathcal{S}_0 := \bigcup_{i \in \mathcal{I}} \mathcal{S}_0(i)$
Obtain model θ_0 training with $\mathcal{D}_0$
for $t = 1, 2, \ldots, T$ **do**
  Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i \in \mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$
  Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)
  Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)
  Obtain model θ_t training with the sieved $\mathcal{D}_t$
return θ_T

Adaptive merging



1. 슈퍼픽셀들을 그래프로 연결하고

$$\text{Superpixels } S \rightarrow \text{Connected graph } \mathcal{G}(S) = (S, \mathcal{E}(S))$$

$$\text{Node set } S \Leftrightarrow \text{Base superpixels } s \in S_0$$

$$\text{Edge set } \mathcal{E}(S) \Leftrightarrow \text{adjacent superpixels } (s, n) \in \mathcal{E}(S)$$

2. 근접한 슈퍼픽셀들을 병합 (메트릭으로 JS divergence 사용)

$$d_{\mathrm{JS}}\big(f_\theta(s) \,\|\, f_\theta(n)\big) < \epsilon \,, \qquad (2)$$

$$d_{\mathrm{JS}}(p \,\|\, q) := \sqrt{\frac{d_{\mathrm{KL}}(p \,\|\, \frac{p+q}{2}) + d_{\mathrm{KL}}(q \,\|\, \frac{p+q}{2})}{2}} \,, \qquad (3)$$

전유진

Adaptive Superpixel for Active Learning in Semantic Segmentation

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

1: Produce base superpixels $\mathcal{S}_0 := \bigcup_{i \in \mathcal{I}} \mathcal{S}_0(i)$

2: Obtain model θ_0 training with $\mathcal{D}_0$

3: **for** $t = 1, 2, \ldots, T$ **do**

4: Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i \in \mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$

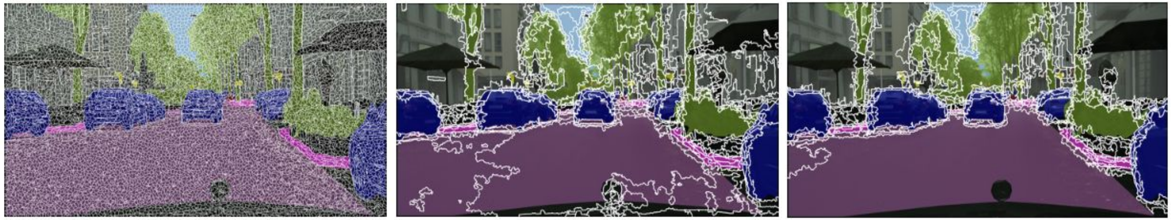
5: Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)

6: Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)

7: Obtain model θ_t training with the sieved $\mathcal{D}_t$

8: **return** θ_T

Adaptive merging



(a) Over-segmented $(t = 0)$ (b) Adaptive merged $(t = 2)$ (c) Adaptive merged $(t = 4)$

3. Merged superpixels 획득

Base superpixels $S_0(i) \rightarrow$ Merged superpixels $S_t(i)$

Merged Superpixels $\mathcal{S}_t := \bigcup_{i \in \mathcal{I}} S_t(i)$

Adaptive Superpixel for Active Learning in Semantic Segmentation

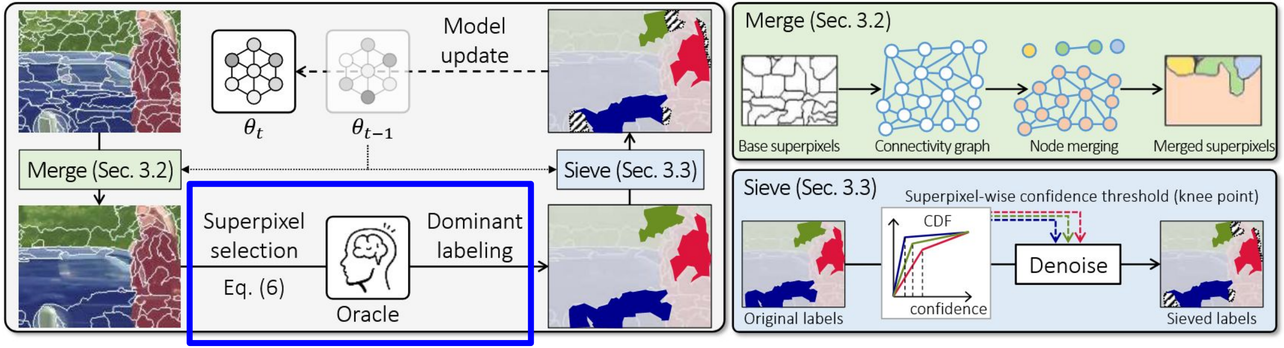
Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

Produce base superpixels $\mathcal{S}_0 := \bigcup_{i \in \mathcal{I}} \mathcal{S}_0(i)$
Obtain model θ_0 training with $\mathcal{D}_0$
for $t = 1, 2, \ldots, T$ **do**
  Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i \in \mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$
  Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)
  Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)
  Obtain model θ_t training with the sieved $\mathcal{D}_t$
return θ_T

Query *top-B* superpixels via Acquisition function



1. Uncertainty 고려

$$u_\theta(x) := \frac{\max_{c \in \mathcal{C} \setminus \{y_\theta(x)\}} f_\theta(c; x)}{\max_{c \in \mathcal{C}} f_\theta(c; x)}, \tag{4}$$

$$u_\theta(s) := \frac{\sum_{x \in s} u_\theta(x)}{|\{x : x \in s\}|}, \tag{5}$$

2. Diversity 고려

$$p(c; \theta) := \frac{|\{x : \exists s \in \mathcal{S}_t, \mathrm{D}_\theta(s) = c, x \in s\}|}{|\{x : \exists s \in \mathcal{S}_t, x \in s\}|}, \tag{6}$$

3. 최종 Acquisition Function

$$a(s; \theta) := u_\theta(s) \exp\left(-p(\mathrm{D}_\theta(s); \theta)\right). \tag{7}$$

전유진

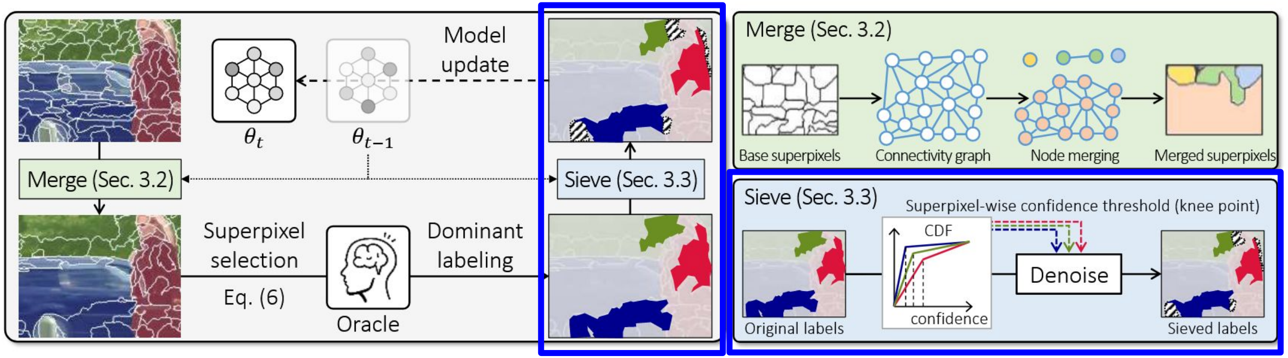
[Adaptive Superpixel for Active Learning in Semantic Segmentation]

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.
1: Produce base superpixels $\mathcal{S}_0 := \bigcup_{i\in\mathcal{I}} \mathcal{S}_0(i)$
2: Obtain model θ_0 training with $\mathcal{D}_0$
3: **for** $t = 1, 2, \dots, T$ **do**
4: Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i\in\mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$
5: Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)
6: Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)
7: Obtain model θ_t training with the sieved $\mathcal{D}_t$
8: **return** θ_T

Sieving



1. 각 슈퍼픽셀에서 잘못된 것 같은 픽셀 제거 (Kneedle 알고리즘 사용)

$$h(s;\theta) := \{x \in s : f_\theta(\mathrm{D}(s); x) \geq \phi(s;\theta)\}\ , \qquad (8)$$

$\phi(s;\theta)$: knee point of the CDF of $f_\theta(\mathrm{D}(s); x)$ in superpixel s

2. Sieved dataset 획득

$$\mathcal{D}_t := \left\{ (x, y) : \begin{array}{l} \exists s \in \cup_{t'=0}^{t} \mathcal{B}_{t'},\ x \in h(s;\theta_{t-1}), \\ y(c) = \mathbb{1}[c = \mathrm{D}(s)]\ \forall c \in \mathcal{C} \end{array} \right\} . \qquad (9)$$

전유진

Adaptive Superpixel for Active Learning in Semantic Segmentation

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Algorithm 1 Proposed Framework

Require: Image set $\mathcal{I}$, batch size B, and final round T.

1: Produce base superpixels $\mathcal{S}_0 := \bigcup_{i \in \mathcal{I}} \mathcal{S}_0(i)$

2: Obtain model θ_0 training with $\mathcal{D}_0$

3: **for** $t = 1, 2, \ldots, T$ **do**

4: Adaptively merge the base superpixels and obtain $\mathcal{S}_t \leftarrow \bigcup_{i \in \mathcal{I}} \mathrm{AM}(\mathcal{S}_0(i), \theta_{t-1})$

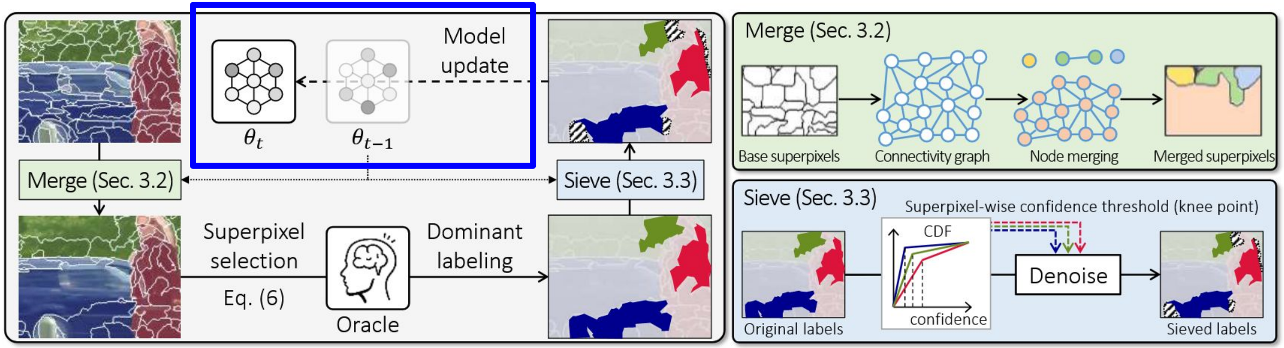
5: Select and query B superpixels $\mathcal{B}_t \subset \mathcal{S}_t$ with (7)

6: Sieve $s \in \bigcup_{t'=0}^{t} \mathcal{B}_{t'}$ and obtain $\mathcal{D}_t$ in (9)

7: Obtain model θ_t training with the sieved $\mathcal{D}_t$

8: **return** θ_T

Model update



1. 업데이트된 데이터셋으로 모델 학습

$$\hat{\mathbb{E}}_{(x,y)\sim\mathcal{D}_t}[\mathrm{CE}(y, f_\theta(x))] \,. \qquad (10)$$

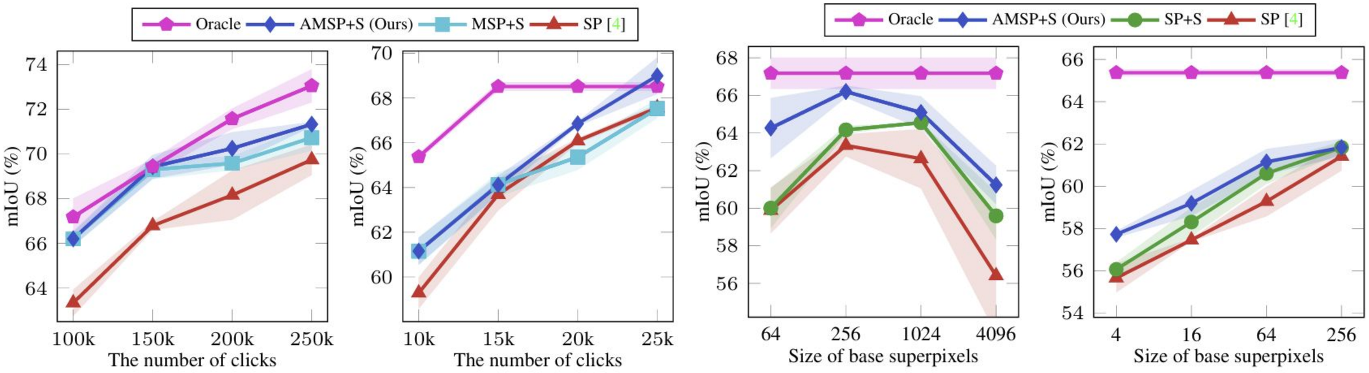
전유진

[Adaptive Superpixel for Active Learning in Semantic Segmentation]

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

ICCV 2023

Experiments



(a) Cityscapes (b) PASCAL (c) Cityscapes (d) PASCAL

Figure 3: *Effect of adaptive superpixels.* (a, b) mIoU versus the number of clicks as budget. (c, d) mIoU versus the size of base superpixels. Each experiment is conducted with three trials and the shaded region indicates ranges.

SP : baseline
SP+S : Ours (Sieving only)
MSP+S : Ours (Merging until 2nd round + Sieving)
AMSP+S : Ours (Merging + Sieving)

전유진

[Adaptive Superpixel for Active Learning in Semantic Segmentation](#)

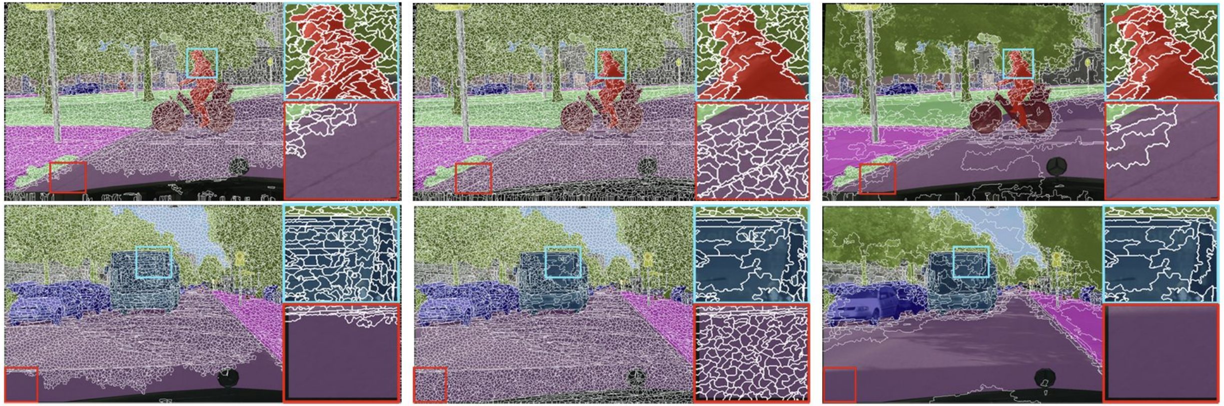
Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

ICCV 2023

Analysis on Adaptive merging

Methods	mIoU
SP [4]	63.77
AMSP+S (bottom 10%)	64.33
AMSP+S (top 10%)	65.99
AMSP+S (complete 100%)	**66.53**

Table 5: *Various levels of partial merging.* Experiments are conducted under the same setting of Figure 3a with 100k clicks (Cityscapes, superpixel size of 256).



(a) Merging superpixels with low 10% uncertainty (b) Merging superpixels with high 10% uncertainty (c) Merging all superpixels

Figure 10: *Qualitative results for partial merging.* The cyan boxes encompass superpixels exhibiting the highest 10% uncertainty, while the red boxes encompass superpixels with the lowest 10% uncertainty. (b) By merging only a portion of superpixels in the order of high uncertainty, we can reduce time complexity, as it creates similar merged superpixels compared with the cyan box in (c).

전유진

Adaptive Superpixel for Active Learning in Semantic Segmentation

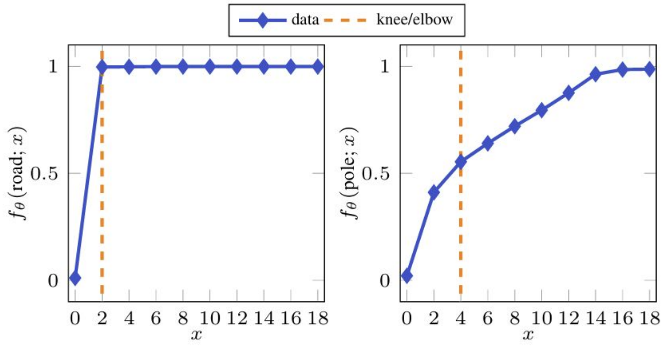
ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Analysis on Sieving

Methods	mIoU
SP [4]	63.77
AMSP+S ($\phi(s;\theta) = 0.0$)	<u>65.35</u>
AMSP+S ($\phi(s;\theta) = 0.2$)	61.80
AMSP+S ($\phi(s;\theta) = 0.4$)	57.77
AMSP+S ($\phi(s;\theta) = 0.6$)	45.84
AMSP+S ($\phi(s;\theta) = 0.8$)	38.99
AMSP+S (Kneedle [30])	**66.53**

Table 6: *Various sieving methods.* Experiments are conducted on Cityscapes dataset with an average superpixel size of 256, using 100k costs for two rounds.



(a) Road (b) Pole

Figure 11: *Examples of knee points on Cityscapes.* We obtain (a) a high knee value for the common road class and (b) a low knee value for the rare pole class.

전유진

[Adaptive Superpixel for Active Learning in Semantic Segmentation]

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

ICCV 2023

Analysis on Acquisition function

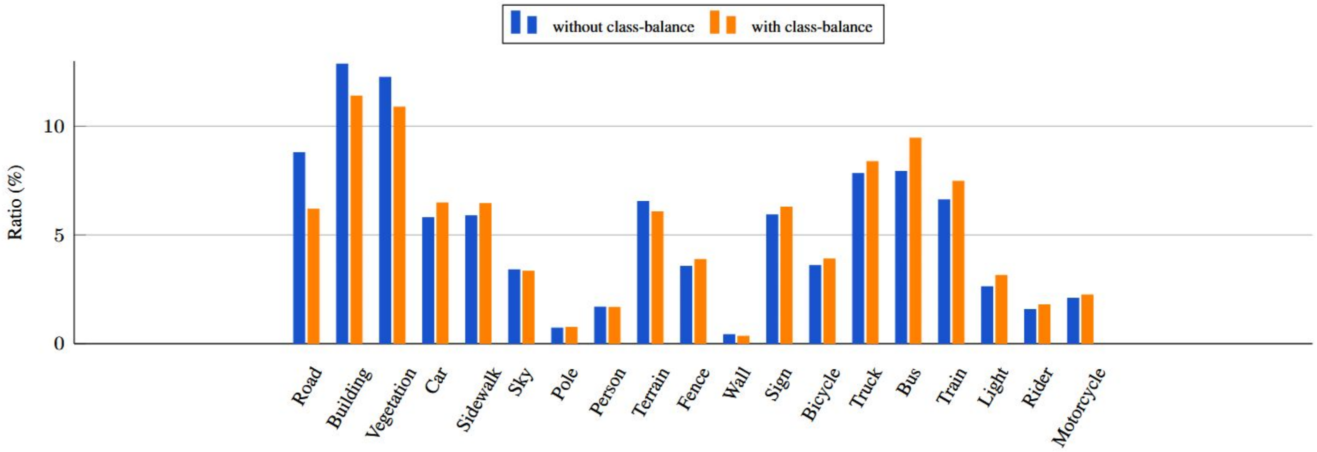


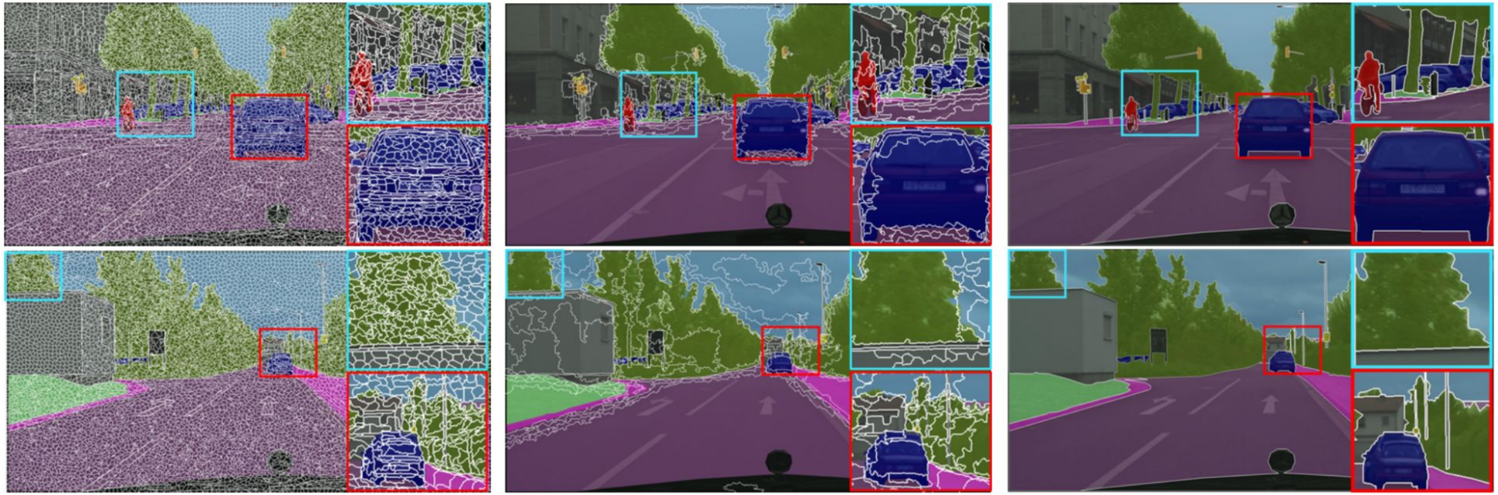
Figure 19: *Effect of class-balanced acquisition function.* According to the ground-truth, class labels are arranged based on the total pixel count for each class, *i.e.* classes become rarer in images as you move from left to right along the x-axis. We observe that classes on the left are selected less with the class-balanced term, while classes on the right are selected more.

[Adaptive Superpixel for Active Learning in Semantic Segmentation]

ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Qualitative results



(a) Base superpixels [35] (b) Merged superpixels (Ours) (c) Oracle superpixels

Figure 4: *Qualitative results of adaptive superpixels.* (a) Base superpixel generated by SEEDS [35] with size 256. (b) Superpixels generated with proposed adaptive merging at round 4. (c) Oracle superpixels generated from the ground truth.

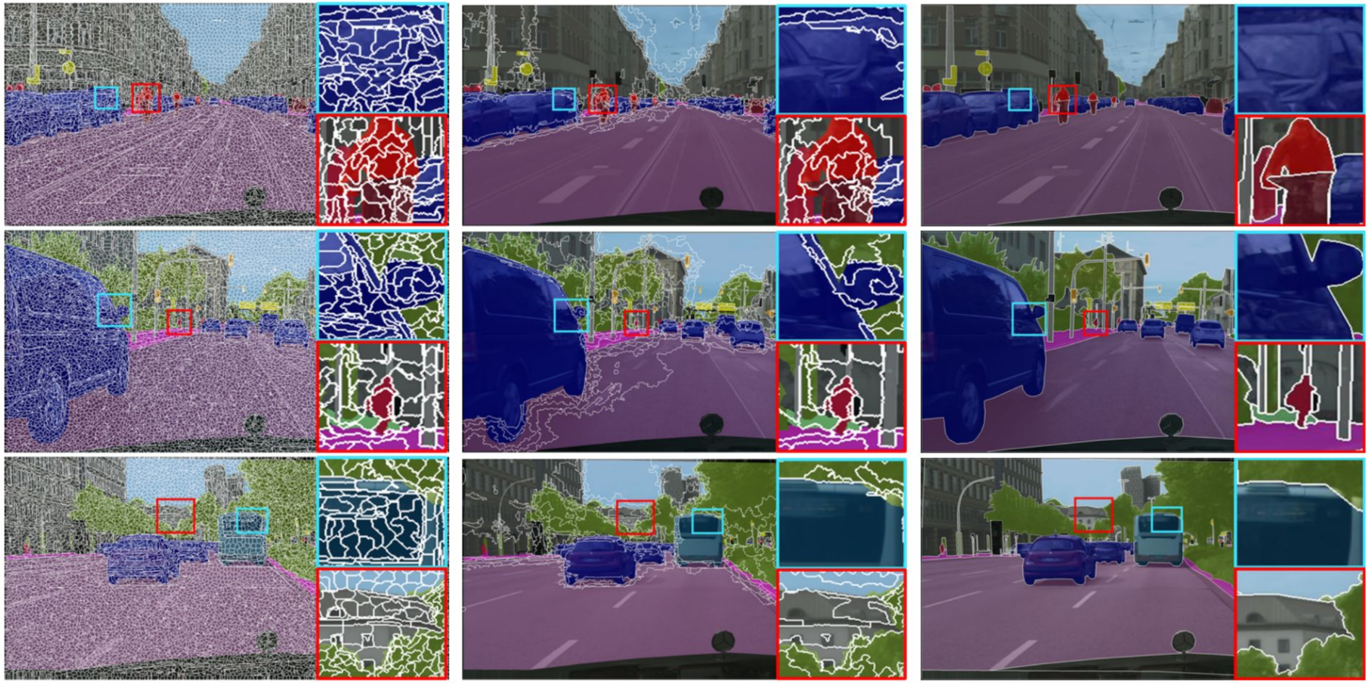
전유진

[Adaptive Superpixel for Active Learning in Semantic Segmentation](#)

ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Qualitat



(a) Base superpixels [35]　　(b) Merged superpixels (Ours)　　(c) Oracle superpixels

Figure 18: *Qualitative results of adaptive superpixels.* (a) Base superpixel generated by SEEDS [35] with size 256. (b) Superpixels generated with proposed adaptive merging at round 4. (c) Oracle superpixels generated from the ground truth.

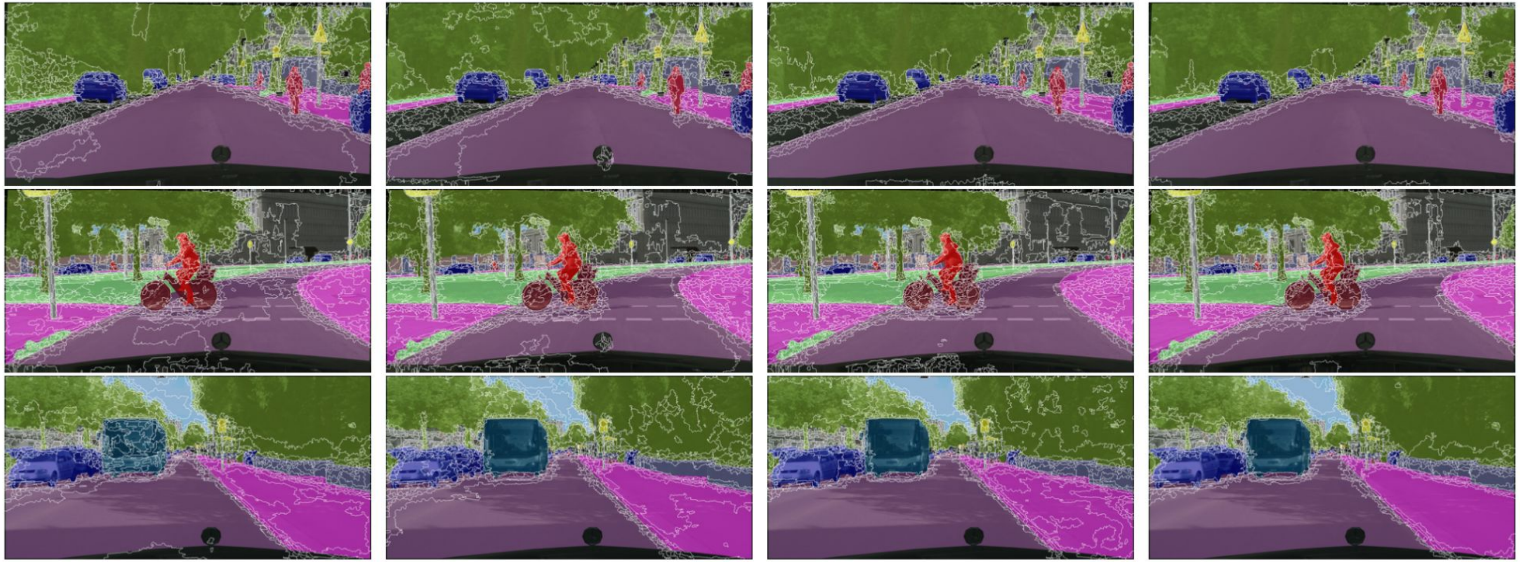
전유진

[Adaptive Superpixel for Active Learning in Semantic Segmentation]

ICCV 2023

Hoyoung Kim Minhyeon Oh Sehyun Hwang Suha Kwak Jungseul Ok

Qualitative results



(a) Adaptive merged ($t = 1$) (b) Adaptive merged ($t = 2$) (c) Adaptive merged ($t = 3$) (d) Adaptive merged ($t = 4$)

Figure 16: *Qualitative results with varying round.* (a-d) Superpixels generated with proposed adaptive merging at rounds 1 to 4. Thanks to the improved model, we observe that the merging becomes more accurate as the round increases. We use the model reported in Figure 3a.

전유진