YONSEI University
CI Computational Intelligence Laboratory

AAAI-26 / IAAI-26 / EAAI-26
JANUARY 20–27, 2026 | SINGAPORE

A²LC: Active and Automated Label Correction for Semantic Segmentation

Youjin Jeon*, Kyusik Cho*, Suhan Woo, Euntai Kim†
Yonsei University

Introduction

- Task: **Label acquisition for semantic segmentation**
  - Challenge: High cost and labor-intensive nature of pixel-wise annotation
  - Promising solution: Active Label Correction with foundation models (ALC)

- Two key challenges of **ALC**
  1. Redundant queries for similar patterns
  2. Class imbalance challenge

- Main contributions of **A²LC**
  1. Label Correction Module (**LCM**) automatically corrects noisy labels exhibiting similar features with human-corrected masks.
  2. Adaptively Balanced Confidence in label (**ABC**) increases the sampling frequency of tail classes.

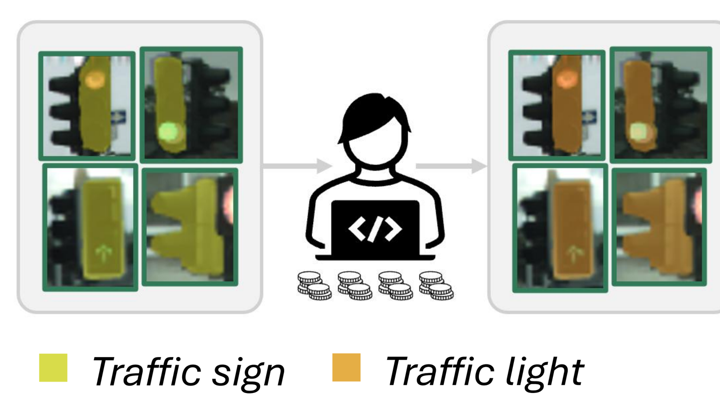
Preliminaries: ALC Framework

Overview of ALC framework

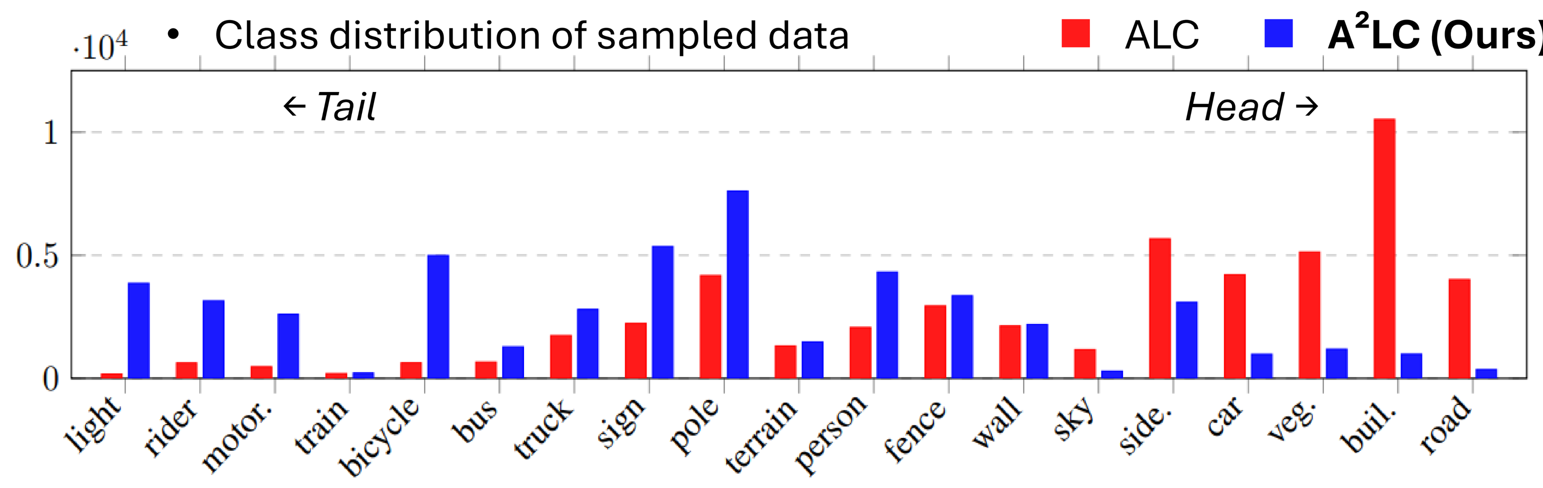
1. Initial pseudo-labels are generated by the foundation model.
2. In each round, top-*B* informative pixels are sampled by acquisition function using deep learning model and queried to annotator.
3. After manual corrections, both the pseudo-labels and model are updated, completing a single round of the correction cycle.

Two key challenges of **ALC**

1. Redundant queries lead to inefficient resource utilization.



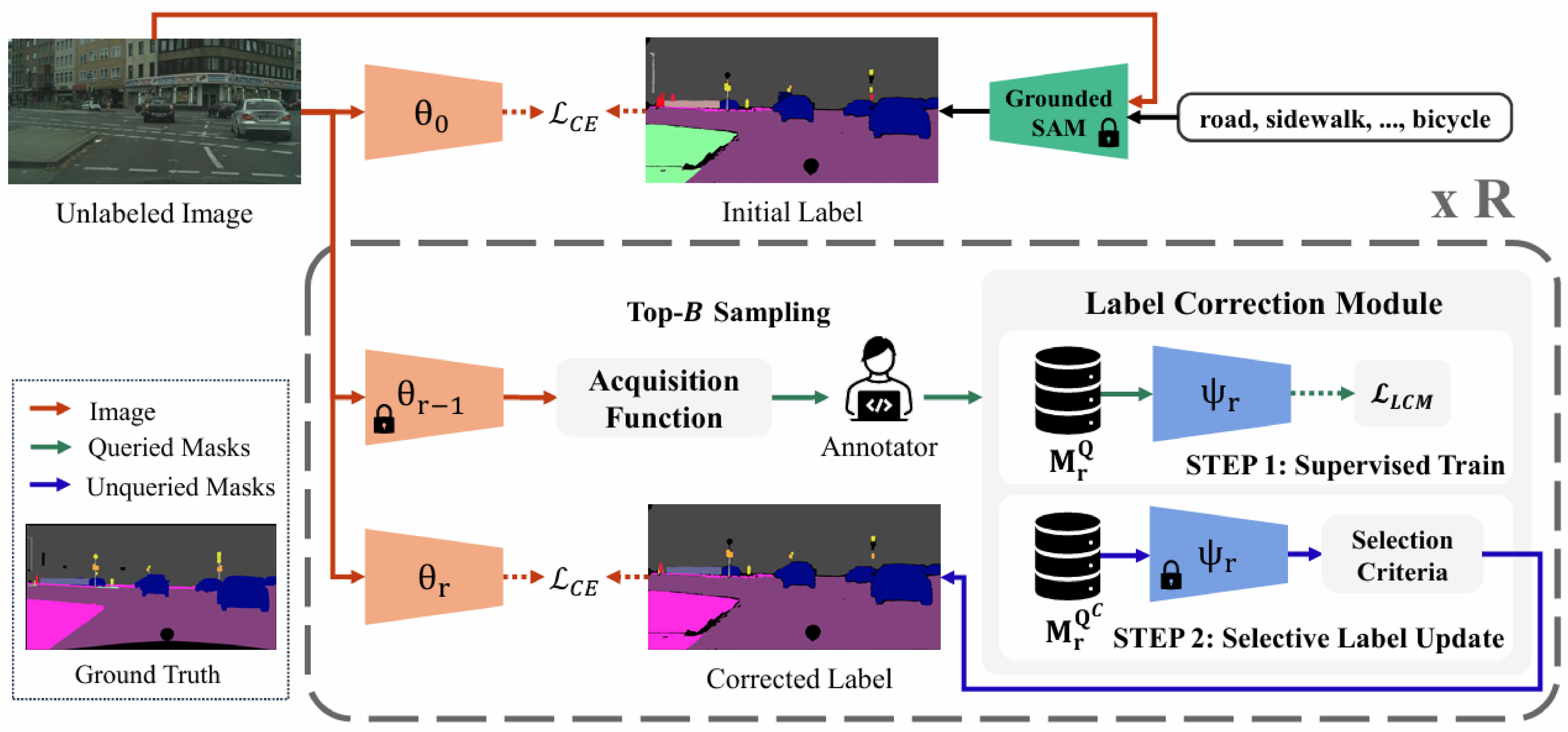
2. Sampling bias toward head classes results in suboptimal corrections.



A²LC Framework

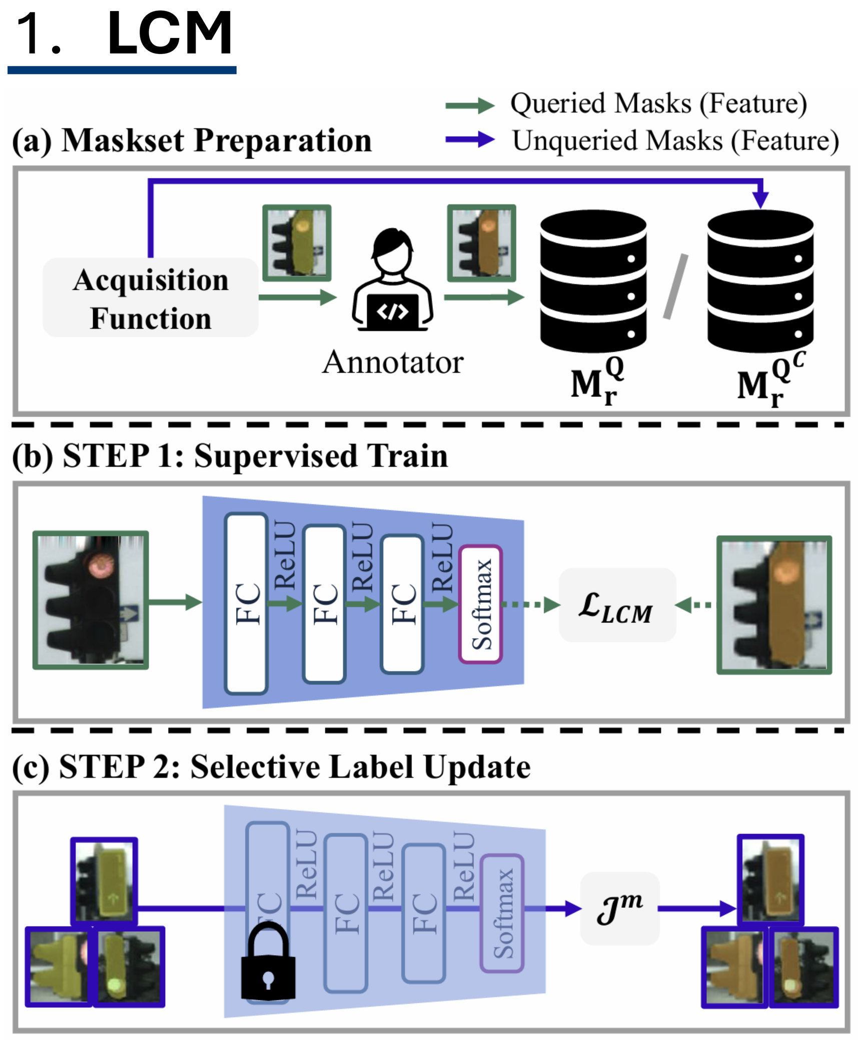
Overview of A²LC framework

- A²LC framework is a semi-automated label correction framework for semantic segmentation built upon cascading stages of manual and automatic correction.



Main contributions of A²LC

1. LCM



- LCM performs automatic correction by propagating human-provided labels beyond the queried samples.
- LCM correction is applied only to the masks that satisfy all selection criteria.
- **Selection criteria** are designed to identify predictions with high reliability, so that automatic label updates are applied exclusively to the subset of confidently inferred labels.

$$\mathcal{J}^m = \mathcal{J}_1^m \wedge \mathcal{J}_2^m \wedge \mathcal{J}_3^m$$
$$\mathcal{J}_1^m = \mathbb{I}\Big(\max_{c\in\mathcal{C}} \psi_r(c;m) \geq \tau\Big)$$
$$\mathcal{J}_2^m = \mathbb{I}\Big(\hat{y}_{\psi_r}(m) \notin \{c \mid \mathrm{rank}(c) \geq (1-\alpha)\cdot|\mathcal{C}|,\ c\in\mathcal{C}\}\Big)$$
$$\mathcal{J}_3^m = \mathbb{I}\Big(\hat{y}(m) \neq \arg\max_{c\in\mathcal{C}} \mathrm{rank}(c)\Big)$$

2. ABC

- ABC guides both correction stages toward improved class balance by incorporating pixel-wise adaptive class weight into the acquisition function.

$$a_{\mathrm{ABC}}(m;\theta) := \sum_{x\in m} \frac{f_\theta(x)\cdot f_\theta(m')}{\|f_\theta(x)\|\|f_\theta(m')\|}\cdot a_{\mathrm{ABC}}(x;\theta) \qquad a_{\mathrm{ABC}}(x;\theta) := w(x)\cdot a_{\mathrm{CIL}}(x;\theta)$$

- **Adaptive class weight** is composed of two components, class rarity score and dataset imbalance score.
  - **Class rarity score** prioritizes pixels belonging to tail classes during sampling.
  - **Dataset imbalance score** adaptively emphasizes the class rarity score according to the pseudo-label statistics at each round.

$$w(x) := \hat{w}(x)^{\mathrm{KL}^3(\mathbb{P}_{\mathrm{dist}}\|\mathbb{U}_{\mathrm{dist}})} \qquad \hat{w}(x) := \frac{\min_{c\in C}|\{x'\in\mathcal{M} : \hat{y}(x') = c\}|}{|\{x'\in\mathcal{M} : \hat{y}(x') = \hat{y}(x)\}|} \qquad \mathrm{KL}(\mathbb{P}_{\mathrm{dist}} \,\|\, \mathbb{U}_{\mathrm{dist}}) = \sum_{c\in C}\mathbb{P}(c)\log\frac{\mathbb{P}(c)}{\mathbb{U}(c)}$$

Experimental Results

Quantitative results

- **High efficiency**: Requires only 20% (Cityscapes) and 60% (PASCAL) of the budget.
- **Strong effectiveness**: Achieves +27.23% (Cityscapes) and +14.30% (PASCAL) performance gains under the same budget.

Dataset	mIoU (%)	Methods	Init.	1R	2R	3R	4R	5R
Cityscapes	Data	ALC	$50.68_{\pm0.00}$	$62.28_{\pm1.25}$	$65.36_{\pm0.56}$	$66.19_{\pm0.60}$	$66.87_{\pm0.29}$	$67.01_{\pm0.30}$
		A²LC (Ours)	$50.68_{\pm0.00}$	$\mathbf{70.01}_{\pm1.05}$	$\mathbf{75.60}_{\pm0.58}$	$\mathbf{80.84}_{\pm0.53}$	$\mathbf{82.69}_{\pm0.92}$	$\mathbf{85.26}_{\pm0.55}$
	Model	ALC	$51.55_{\pm0.71}$	$56.61_{\pm0.42}$	$58.27_{\pm0.63}$	$58.58_{\pm0.08}$	$58.52_{\pm0.26}$	$58.59_{\pm0.05}$
		A²LC (Ours)	$51.55_{\pm0.71}$	$\mathbf{60.83}_{\pm0.66}$	$\mathbf{63.89}_{\pm0.44}$	$\mathbf{67.50}_{\pm0.31}$	$\mathbf{68.87}_{\pm0.78}$	$\mathbf{70.51}_{\pm0.32}$
PASCAL	Data	ALC	$58.63_{\pm0.00}$	$\mathbf{68.19}_{\pm0.34}$	$72.72_{\pm0.16}$	$74.84_{\pm0.29}$	$76.41_{\pm0.70}$	$77.06_{\pm0.73}$
		A²LC (Ours)	$58.63_{\pm0.00}$	$67.49_{\pm2.03}$	$\mathbf{74.88}_{\pm1.45}$	$\mathbf{80.88}_{\pm0.83}$	$\mathbf{84.81}_{\pm0.48}$	$\mathbf{88.08}_{\pm0.44}$
	Model	ALC	$56.94_{\pm0.44}$	$\mathbf{62.11}_{\pm0.61}$	$\mathbf{64.12}_{\pm0.31}$	$64.15_{\pm0.68}$	$65.00_{\pm0.33}$	$65.48_{\pm0.83}$
		A²LC (Ours)	$56.94_{\pm0.44}$	$60.87_{\pm2.93}$	$64.08_{\pm2.43}$	$\mathbf{66.45}_{\pm0.95}$	$\mathbf{67.76}_{\pm0.06}$	$\mathbf{68.42}_{\pm0.87}$

Qualitative results

- Constructed pseudo-labels

Unlabeled Image | Init. | ALC | A²LC (Ours) | Ground Truth

- Visualization of LCM-corrected masks

Ablation study

Methods			Data mIoU (%)					Model mIoU (%)				
LCM	ABC	Mask	1R	2R	3R	4R	5R	1R	2R	3R	4R	5R
			62.60	65.20	66.03	66.85	66.86	56.86	58.93	58.63	58.59	58.59
✓			65.66	71.53	76.00	78.38	80.59	59.02	61.39	64.19	65.58	66.94
	✓		67.31	70.06	74.49	78.32	81.58	59.09	60.86	62.05	65.27	67.75
✓	✓		70.62	**77.29**	80.39	82.80	84.15	60.32	**65.09**	67.36	68.71	70.11
✓	✓	✓	**71.04**	76.08	**81.13**	**83.53**	**85.57**	**61.06**	64.39	**67.83**	**69.63**	**70.88**