CAUG-$\mathscr{A}_r$: Curriculum & Uncertainty-Guided Augmentation for Semi-Supervised Semantic Segmentation

Trần Vạn Tấn[1], Lại Khánh Hoàng[2], Lê Đình Duy[3]
[1]University of Information Technology

What?

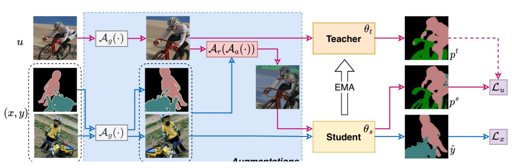
- Improve AugSeg by redesigning the unlabeled intensity augmentation $\mathscr{A}_r$.
- CAUG-$\mathscr{A}_r$ adapts (i) #ops k and (ii) magnitude m over epochs (curriculum).
- It also scales augmentation strength per image using teacher confidence ρ_i.

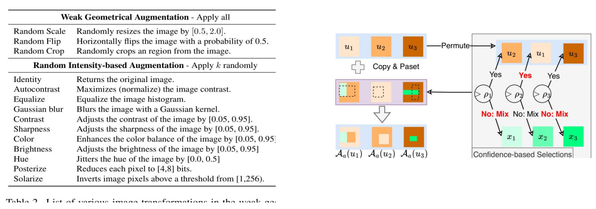
Why?

- Pixel-level annotation is costly, leading to limited labeled data in practice.
- Early pseudo-labels can be noisy, increasing the risk of confirmation bias.
- Adaptive augmentation is needed: epoch-aware and confidence-aware.

Overview

Teacher–Student → **AugSeg Augmentations** → **Ours: CAUG-$\mathscr{A}_r$**





Curriculum over epochs

Early	Mid	Late
k=1–2, m ↓	k=2–3, m ↗	k=3–4, m ↑

training progress

Uncertainty-guided strength

low ρ_i (hard) → mild aug | high ρ_i (easy) → strong aug

Description

1. AugSeg baseline & $\mathscr{A}_r$

- Mean-Teacher framework: EMA teacher generates pseudo-labels for unlabeled data.
- Core unlabeled perturbation: $T(u) = \mathscr{A}_r(\mathscr{A}_a(\mathscr{A}_g(u)))$.
- $\mathscr{A}_r$ (Random Intensity Aug.): randomly sample k ops and continuous strength (Fig. 3).
- Augmentation pool is simplified (Table 2) to avoid over-distortion.

Weak Geometrical Augmentation - Apply all	
Random Scale	Randomly resizes the image by [0.5, 2.0].
Random Flip	Horizontally flips the image with a probability of 0.5.
Random Crop	Randomly crops an region from the image.
Random Intensity-based Augmentation - Apply k randomly	
Identity	Returns the original image.
Autocontrast	Maximizes (normalize) the image contrast.
Equalize	Equalize the image histogram.
Gaussian blur	Blurs the image with a Gaussian kernel.
Contrast	Adjusts the contrast of the image by [0.05, 0.95].
Sharpness	Adjusts the sharpness of the image by [0.05, 0.95].
Color	Enhances the color balance of the image by [0.05, 0.95]
Brightness	Adjusts the brightness of the image by [0.05, 0.95]
Hue	Jitters the hue of the image by [0.0, 0.5]
Posterize	Reduces each pixel to [4,8] bits.
Solarize	Inverts image pixels above a threshold from [1,256).

Table 2. Augmentation pool (Zhao et al., 2022).

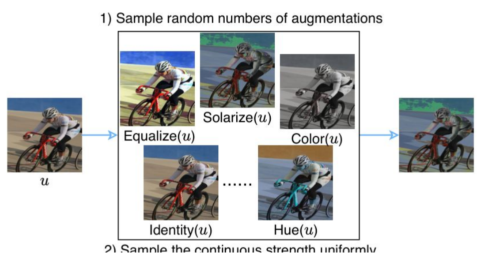


Fig. 3. Random intensity augmentation (Zhao et al., 2022).

2. CAUG-$\mathscr{A}_r$ (Curriculum + Uncertainty)

- Limitation: fixed k / magnitude does not reflect training stage or sample difficulty.
- Curriculum: increase k and magnitude from early → mid → late training.
- Uncertainty-guided: use teacher confidence ρ_i to modulate strength per image.
- High ρ_i (easy) → stronger aug; Low ρ_i (hard) → milder aug to reduce noise.
- Drop-in replacement: only change $\mathscr{A}_r$ in the

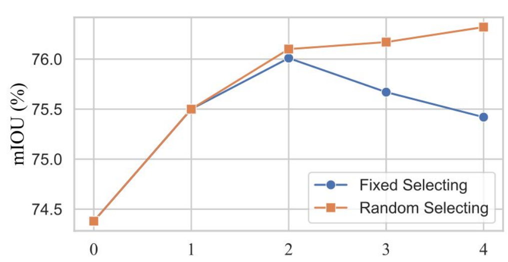


Fig. 6. Selecting strategies vs #selected ops (Zhao et al., 2022).

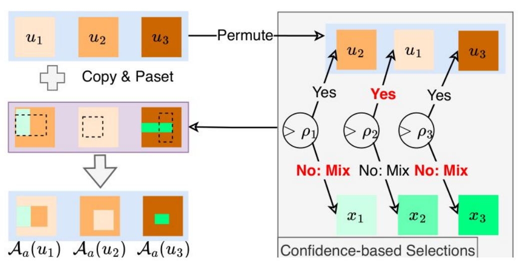


Fig. 4. Confidence score ρ_i in AugSeg (Zhao et al., 2022).

Confidence (teacher) ρ_i can be estimated from pseudo-label probabilities,
e.g., mean max-probability weighted by (1 − normalized entropy).

3. Evaluation & expected impact

- Datasets: Pascal VOC 2012 and Cityscapes under common splits (1/16, 1/8, 1/4, 1/2).
- Metric: mean Intersection-over-Union (mIoU).
- Baseline reference: AugSeg reaches strong SOTA on VOC/Cityscapes (Tables 3 & 5).
- Hypothesis: CAUG-$\mathscr{A}_r$ reduces early noise and improves late-stage generalization → higher mIoU.

Method	Encoder	1/16 (92)	1/8 (183)	1/4 (366)	1/2 (732)	Full (1464)
Supervised	R50	44.03	52.26	61.65	66.72	72.94
PseudoSeg [54]	R50	54.89	61.88	64.85	70.42	71.00
PC^{2}Seg [52]	R50	56.90	64.63	67.62	70.90	72.26
AugSeg	R50	64.22	72.17	76.17	77.40	78.82
Supervised	R101	43.92	59.10	65.88	70.87	74.97
CutMix-Seg [16]	R101	52.16	63.47	69.46	73.73	76.54
PseudoSeg [54]	R101	57.60	65.50	69.14	72.41	73.23
PC^{2}Seg [52]	R101	57.00	66.28	69.78	73.05	74.15
CPS [7]	R101	64.07	67.42	71.71	75.88	-
PS-MT [34]	R101	65.80	69.58	76.57	78.42	80.01
ST++ [47]	R101	65.20	71.00	74.60	77.30	79.10
U^{2}PL [45]	R101	67.98	69.15	73.66	76.16	79.49
AugSeg	R101	71.09	75.45	78.80	80.33	81.36

Table 3. VOC2012 results (Zhao et al., 2022).

Method	ResNet-50				ResNet-101			
	1/16(186)	1/8(372)	1/4(744)	1/2(1488)	1/16(186)	1/8(372)	1/4(744)	1/2(1488)
Supervised	63.34	68.73	74.14	76.62	64.77	71.64	75.24	78.03
MT [44]	66.14	72.03	74.47	77.43	68.08	73.71	76.53	78.59
CCT [42]	66.35	72.46	75.68	76.78	69.64	74.48	76.35	78.29
GCT [27]	65.81	71.33	75.30	77.09	66.90	72.96	76.45	78.58
CPS [7]	69.79	74.39	76.85	78.64	70.50	75.71	77.41	80.08
CPS * [7]	-	-	-	-	69.78	74.31	74.58	76.81
PS-MT† [34]	-	75.76	76.92	77.64	-	76.89	77.60	79.09
U^{2}PL [45]	69.03	73.02	76.31	78.64	70.30	74.37	76.47	79.05
AugSeg	73.73	76.49	78.76	79.33	75.22	77.82	79.56	80.43

Table 5. Cityscapes results (Zhao et al., 2022).

Fig. 7. Qualitative segmentation (Zhao et al., 2022).

AugSeg			mIoU	
MT	$\mathcal{A}_r$	$\mathcal{A}_a$	VOC (366)	Citys (744)
			61.65 (supervised)	74.14 (supervised)
✓			69.06 (7.41↑)	75.96 (1.82↑)
✓	✓		72.41 (10.76↑)	77.29 (3.15↑)
✓		✓	74.33 (12.68↑)	77.44 (3.30↑)
✓	✓	✓	**76.17** (14.52↑)	**78.76**(4.62↑)

Table 6. $\mathscr{A}_r + \mathscr{A}_a$ contribute strongly (Zhao et al., 2022).

Contact: your.email@domain.com