Towards Robust Pseudo-Label Learning in Semantic Segmentation: An Encoding Perspective

Wangkai Li[1], Rui Sun[1], Zhaoyang Li[1], Tianzhu Zhang[1, 2]

[1]University of Science and Technology of China, [2]Deep Space Exploration Laboratory

中国科学技术大学 University of Science and Technology of China

DSEL DEEP SPACE EXPLORATION LAB 深空探测实验室

NEURAL INFORMATION PROCESSING SYSTEMS

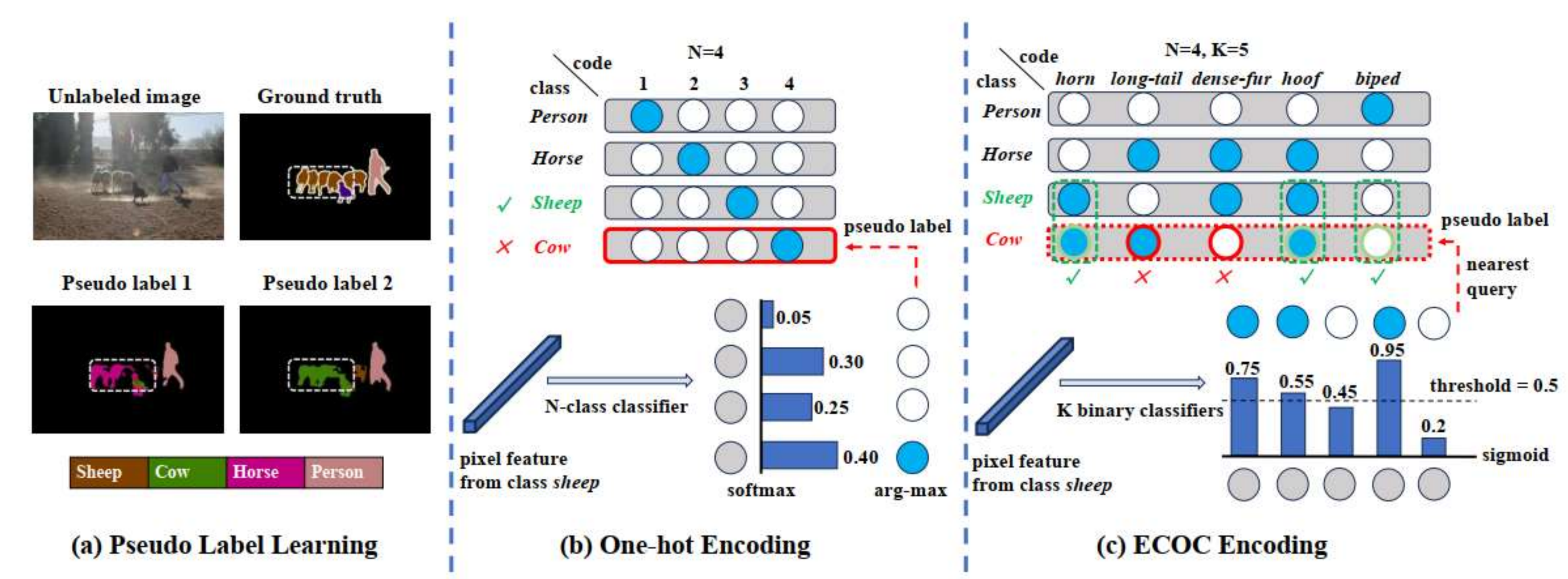
TLDR. Utilize ECOC encoding to denoise pseudo-labels for label-scarce semantic segmentation at the bit level

Introduction

Goal. Domain-adaptive / semi-supervised semantic segmentation aims to avoid laborious pixel-wise annotation using annotated and unlabeled data simultaneously.

Mainstream Paradigm. Self-training and consistency regularization generates pseudo labels for unlabeled data as supervision, summarized as pseudo-label learning.

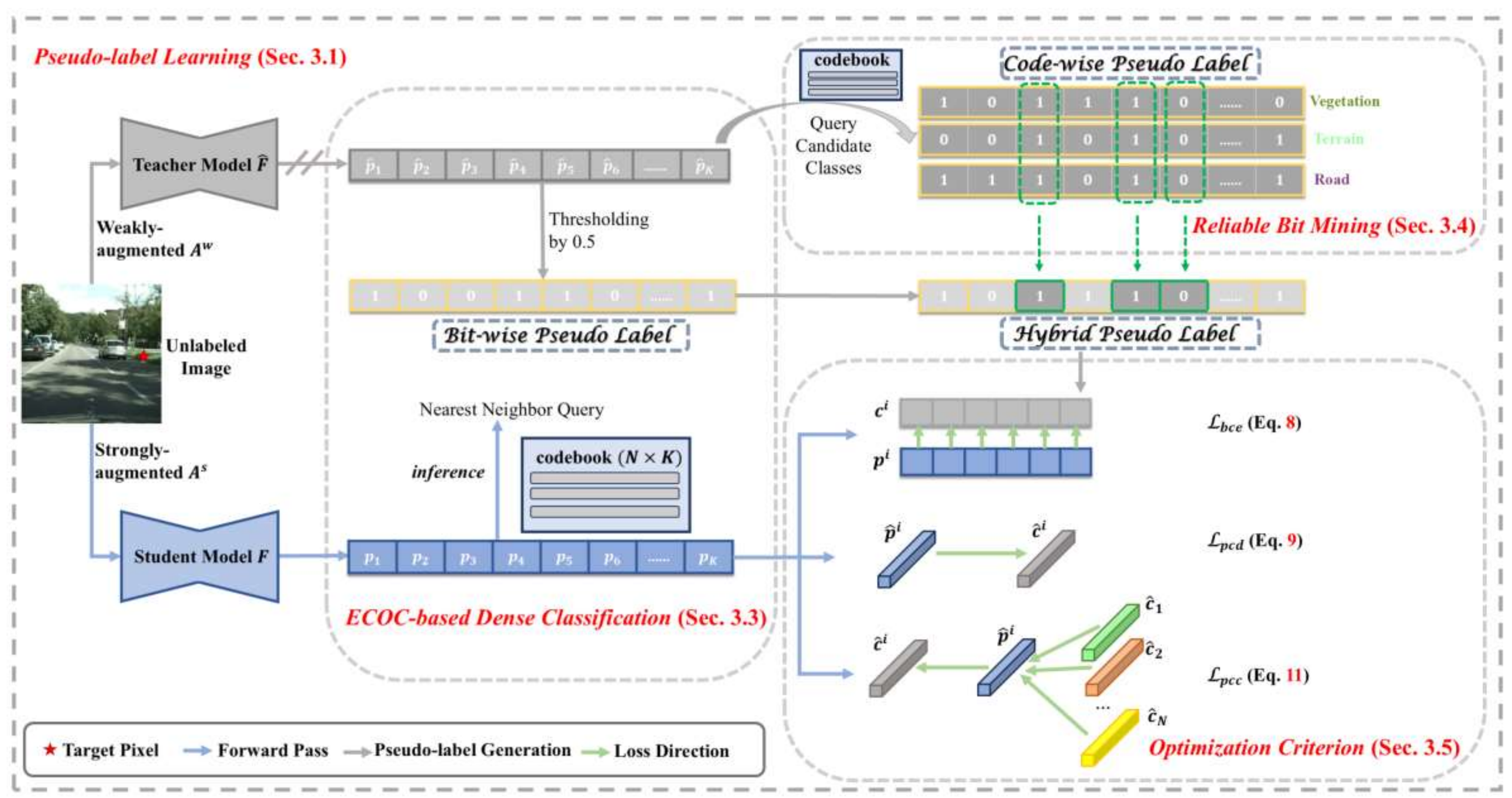
Problem. Pseudo-label learning under UDA/SSL improves segmentation but suffers from noisy pseudo-labels; one-hot encoding amplifies these errors during training.



Motivation.

- One-hot limitations: Enforces hard class assignments, ignores shared attributes among confusing classes.
- ECOC advantages: Enables bit error correction by large Hamming distances, stabilizes learning via shared bits.
- Theoretical guarantee:
  - In fully supervised settings, ECOC can serve as an effective equivalent to one-hot encoding.
  - In pseudo-label learning, ECOC exhibits greater robustness with a tighter classification error bound.

Method



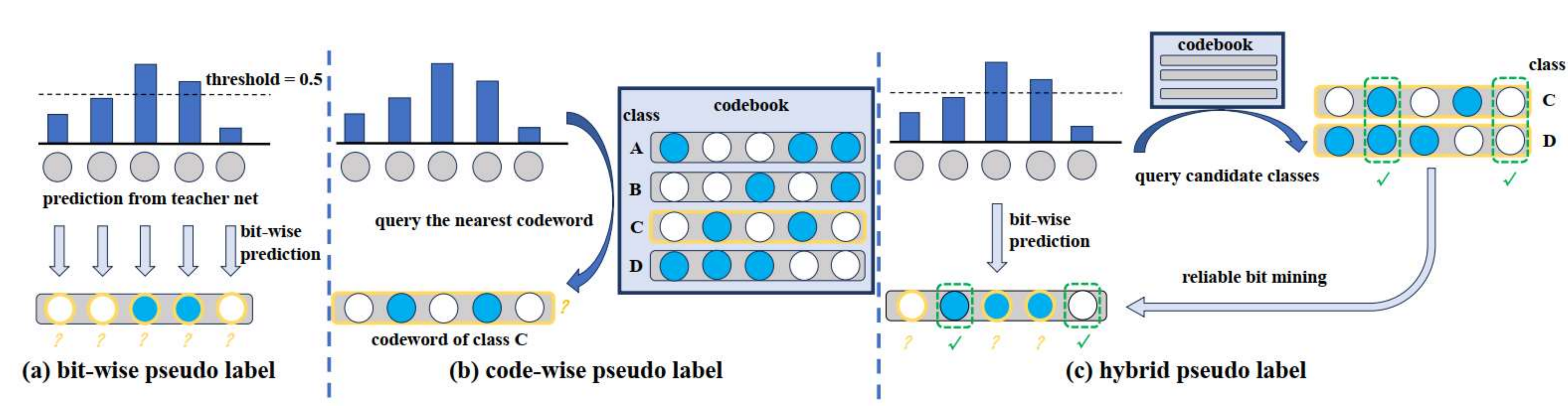
ECOC-based Dense Classification. Replace N-way softmax with K independent sigmoid heads (K bits).

$$d_{SH}(\boldsymbol{c}_n, \boldsymbol{p}^i) = \frac{1}{K}\sum_{k=1}^{1} \| p(k|\boldsymbol{z}_i) - \boldsymbol{c}_{nk} \|_1 ,$$

$$\hat{n}^i = \underset{n}{\operatorname{argmin}}\{d_{SH}(\boldsymbol{c}_n, \boldsymbol{p}^i)\}.$$

Reliable Bit Mining. Extract reliable bits from code-wise labels within candidates and fuse them with bit-wise labels, obtain more robust pseudo-labels in a hybrid way:

$$\boldsymbol{c}^i_{hyb.} = \mathcal{M}^i \odot \boldsymbol{c}^i_{code} + (1\text{-}\mathcal{M}^i) \odot \boldsymbol{c}^i_{bit}.$$



Optimization Criterion. Typically binary cross-entropy:

$$\mathcal{L}^i_{bce} = -\frac{1}{K}\sum_{k=1}^{1}[\boldsymbol{c}^i_{(k)} \log p(k|\boldsymbol{z}_i) + (1-\boldsymbol{c}^i_{(k)})\log(1-p(k|\boldsymbol{z}_i))]$$

- Pixel-code distance: $\mathcal{L}^i_{pcd} = 1 - \cos(\hat{\boldsymbol{p}}^i, \hat{\boldsymbol{c}}^i)$
- Pixel-code contrast:

$$\mathcal{L}^i_{pcc} = \log(1 + \sum_{\hat{c}^- \in \hat{\mathcal{C}}^-} \exp(\langle \hat{\boldsymbol{p}}^i, \widehat{\hat{\boldsymbol{c}}^- - \boldsymbol{c}^i} \rangle / \tau))$$

Experiments

Quantitative Results on GTAv→ Cityscapes.

Method	Arch.	Road	Sidewalk	Building	Wall	Fence	Pole	Light	Sign	Veg	Terrain	Sky	Person	Rider	Car	Truck	Bus	Train	Motor	Bike	mIoU
GTAv→Cityscapes(Val.)																					
ProDA [100]	C	87.8	56.0	79.7	46.3	44.8	45.6	53.5	53.5	88.6	45.2	82.1	70.7	39.2	88.8	45.5	50.4	1.0	48.9	56.4	57.5
CPSL [45]	C	92.3	59.5	84.9	45.7	29.7	52.8	61.5	59.5	87.9	41.6	85.0	73.0	35.5	90.4	48.7	73.9	26.3	53.8	53.9	60.8
TransDA [11]	T	94.7	64.2	89.2	48.1	45.8	50.1	60.2	40.8	90.4	50.2	93.7	76.7	47.6	92.5	56.8	60.1	47.6	49.6	55.4	63.9
ADFormer [30]	T	96.7	75.1	88.8	57.5	45.9	45.6	55.4	59.8	90.2	45.6	92.1	70.8	43.0	91.0	78.9	79.3	68.7	52.7	65.0	69.2
CDAC [83]	T	97.1	78.7	91.8	59.6	57.1	59.1	66.1	72.2	91.8	53.1	94.5	79.4	51.6	94.6	84.9	87.8	78.7	64.9	67.6	75.3
DACS [79]	C	89.9	39.7	87.9	39.7	39.5	38.5	46.4	52.8	88.0	44.0	88.8	67.2	35.8	84.5	45.7	50.2	0.2	27.3	34.0	52.1
+ECOCSeg	C	**95.6**	**71.8**	**90.2**	37.8	31.4	**44.8**	**50.8**	**58.8**	**90.4**	**50.3**	**91.3**	**68.6**	23.5	**91.2**	**49.8**	**55.4**	**8.8**	15.2	9.8	**54.5**$_{\uparrow 2.4}$
DAFormer [34]	T	95.7	70.2	89.4	53.5	48.1	49.6	55.8	59.4	89.9	47.9	92.5	72.2	44.7	92.3	74.5	78.2	65.1	55.9	61.8	68.3
+ECOCSeg	T	**96.7**	**75.6**	**89.4**	**54.0**	**51.4**	**55.1**	**59.4**	**61.9**	**90.1**	46.6	90.0	71.5	42.4	**92.8**	**79.7**	**85.4**	**79.1**	**60.0**	58.2	**70.5**$_{\uparrow 2.2}$
MIC [35]	T	97.4	80.1	91.7	61.2	56.9	59.7	66.0	71.3	91.7	51.4	94.3	79.8	56.1	94.6	85.4	90.3	80.4	64.5	68.5	75.9
+ECOCSeg	T	**97.9**	**81.4**	**91.9**	**62.2**	54.3	**64.2**	**67.4**	**76.1**	**92.9**	**54.4**	94.2	**82.1**	53.0	**95.2**	**89.6**	**90.8**	**82.3**	61.9	**69.4**	**76.9**$_{\uparrow 1.0}$

Visualization Results of Reliable Bits.