VBLC: Visibility Boosting and Logit-Constraint Learning for Domain Adaptive Semantic Segmentation under Adverse Conditions

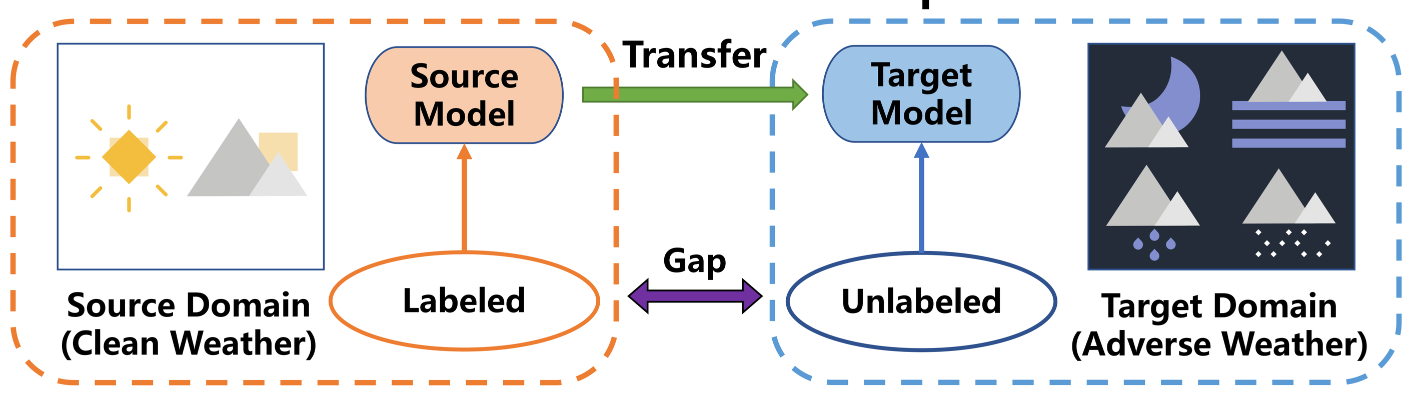
Mingjia Li[1,*] Binhui Xie[1,*] Shuang Li[1,✉] Chi Harold Liu[1] Xinjing Cheng[2]

[1] Beijing Institute of Technology, Beijing, China [2] Inceptio Technology, Shanghai, China

Contact: mingjiali@bit.edu.cn

Paper GitHub

Hightlights



Problem: Normal-to-Adverse Domain Adaptation

- **Input:** Labeled source domain $\mathcal{S}$ in clean weather, Unlabeled adverse target domain $\mathcal{T}$ in adverse weather
- **Core:** Close the gap between clean weather and adverse weather conditions; avoid overconfidence caused by large domain gap.

Contributions:

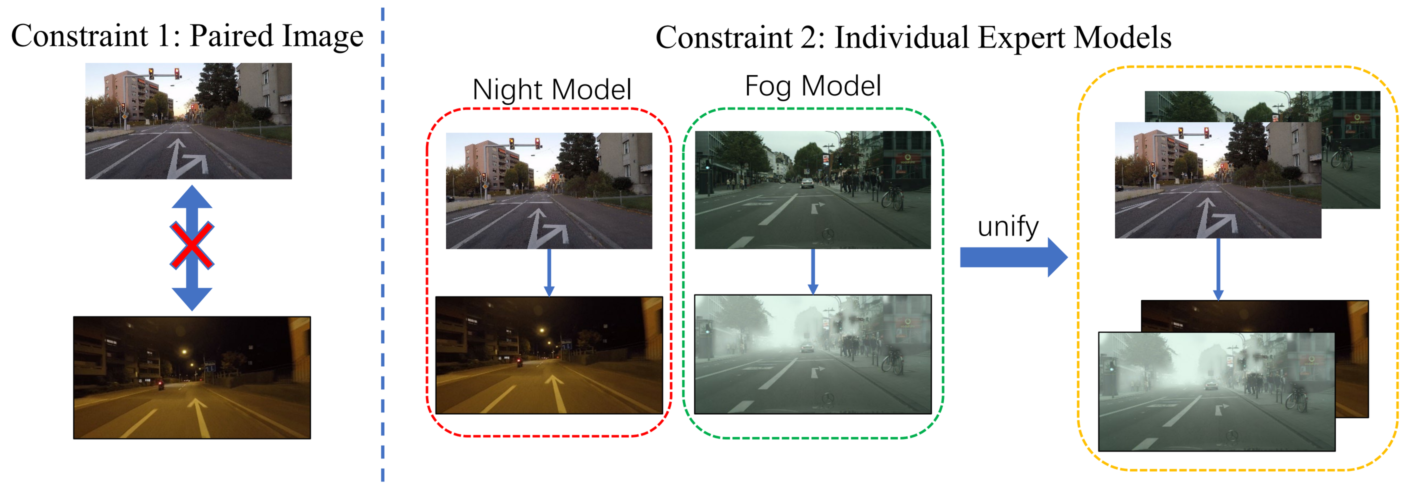
- We propose **VBLC**, a *fully unsupervised* framework that enables adaptation to *hybrid* adverse weather conditions.
- **VBLC** yields good results *without* clean-target image pair as correspondence and achieves SOTA in Normal-to-Adverse DA.

Motivation

Normal-to-Adverse DA methods often adopt paired images as weak supervision, and train individual models for each adverse condition.

- How to remove image correspondence to simplify data preparation?
- How to adapt model to a hybrid of adverse weather conditions?
- How to avoid model overconfidence under large domain gap?

Observation: the *veiling effect* can be the major obstacle to a clear vision, which is tackable via a designed physical model.



Approach

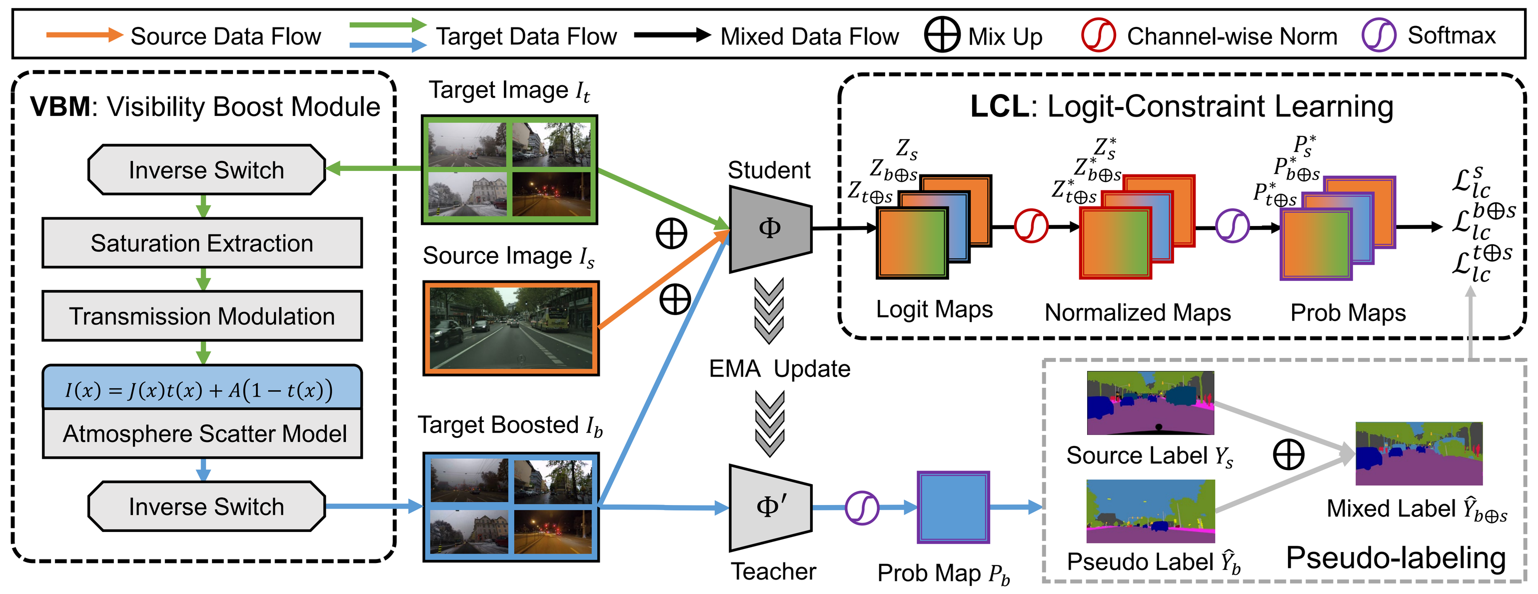


Figure 1: VBLC for Normal-to-Adverse Domain Adaptive Semantic Segmentation.

Visibility Boost Module (VBM):

Atmosphere scatter model (ASM) for *veiling effect*:

$$I(x)=J(x)t(x)+A(1-t(x)),$$

where $I(x)$ is the observed image, $J(x)$ is the restored image, $t(x)$ is transmission map, A is global atmospheric light.

Dynamic estimation of transmission map:

$$t(x)=1-\omega_s \min_c\left(\min_{y\in\Omega(x)}\frac{I^c(y)}{A^c}\right),$$

where ω_s is calculated from image saturation for adaptive modulation.

Inverse Switch for nighttime enhancement:

$$1-I(x)=(1-J(x))t(x)+A(1-t(x)).$$

as inverted nighttime image can be viewed as hazy ones.

Logit-Constraint Learning (LCL):

Cross-entropy loss leads to overconfidence: the optimization is driven towards given label regardless of prediction confidence,

$$\frac{\partial\mathcal{L}_{ce}}{\partial z_j}=p_j-y_j\,,\mathcal{L}_{ce}=-\sum_{k=1}^{K}y_k\log(p_k),\ \text{where}\ \ p_i=\frac{e^{z_i}}{\sum_{k=1}^{K}e^{z_k}}.$$

Logit-Constraint Learning loss mitigates overconfidence: the gradient is reduced when encountered with unconfident predictions,

$$\frac{\partial\mathcal{L}_{lc}}{\partial z_j}=\frac{1}{\|z\|}\left((p_j^*-y_j)-\sum_{k=1}^{K}\frac{z_jz_k}{\|z\|^2}(p_k^*-y_k)\right),$$

$$\mathcal{L}_{lc}=-\sum_{k=1}^{K}y_k\log(p_k^*),\ \text{where}\ \ p_i^*=\frac{e^{z_i/\|z\|}}{\sum_{k=1}^{K}e^{z_k/\|z\|}}.$$

Findings

Method	road	side.	buil.	wall	fence	pole	light	sign	veg.	terr.	sky	pers.	rider	car	truck	bus	train	mbike	bike	mIoU
DeepLab-v2	71.9	26.2	51.1	18.8	22.5	19.7	33.0	27.7	67.9	28.6	44.2	43.1	22.1	71.2	29.8	33.3	48.4	26.2	35.8	38.0
AdaptSegNet	69.4	34.0	52.8	13.5	18.0	4.3	14.9	9.7	64.0	23.1	38.2	38.6	20.1	59.3	35.6	30.6	53.9	19.8	33.9	33.4
ADVENT	72.9	14.3	40.5	16.6	21.2	9.3	17.4	21.2	63.8	23.8	18.3	32.6	19.5	69.5	36.2	34.5	46.2	26.9	36.1	32.7
BDL	56.0	32.5	68.1	20.1	17.4	15.8	30.2	28.7	59.9	25.3	37.7	28.7	25.5	70.2	39.6	40.5	52.7	29.2	38.4	37.7
CLAN	**79.1**	29.5	45.9	18.1	21.3	22.1	35.3	40.7	67.4	29.4	32.8	42.7	18.5	73.6	42.0	31.6	55.7	25.4	30.7	39.0
CRST	51.7	24.4	67.8	13.3	9.7	30.2	38.2	34.1	58.0	25.2	**76.8**	39.9	17.1	65.4	3.7	6.6	39.6	11.8	8.6	32.8
FDA	73.2	34.7	59.0	24.8	**29.5**	28.6	43.3	44.9	**70.1**	28.2	54.7	47.0	28.5	74.6	**44.8**	**52.3**	**63.3**	28.3	39.5	45.7
DACS	58.5	34.7	76.4	20.9	22.6	31.7	32.7	46.8	58.7	39.0	36.3	43.7	20.5	72.3	39.6	34.8	51.1	24.6	38.2	41.2
VBLC	49.6	**39.3**	**79.4**	**35.8**	**29.5**	**42.6**	**57.2**	**57.5**	69.1	**42.7**	39.8	**54.5**	**29.3**	**77.8**	43.0	36.2	32.7	**38.7**	**53.4**	**47.8**
SegFormer	66.9	25.8	71.3	20.9	22.2	41.1	47.2	46.6	**74.2**	44.9	75.6	50.4	23.5	73.1	30.3	36.8	55.8	29.4	37.1	45.9
DAFormer	56.9	45.4	84.7	**44.7**	35.1	48.6	44.8	57.4	69.5	52.9	45.8	57.1	28.2	82.8	57.2	63.9	84.0	40.2	50.5	55.3
VBLC	**89.2**	**59.8**	**85.9**	44.0	**37.2**	**53.5**	**64.5**	**63.2**	72.4	**56.3**	**84.1**	**65.5**	**37.7**	**85.1**	**60.1**	**71.8**	**85.2**	**47.7**	**56.3**	**64.2**

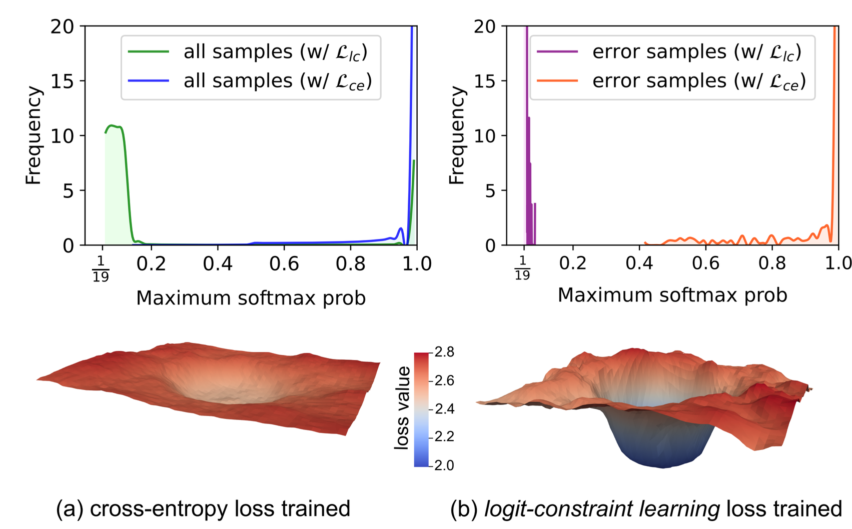
Table 1: Comparison results on **Cityscapes** → **ACDC** segmentation task.

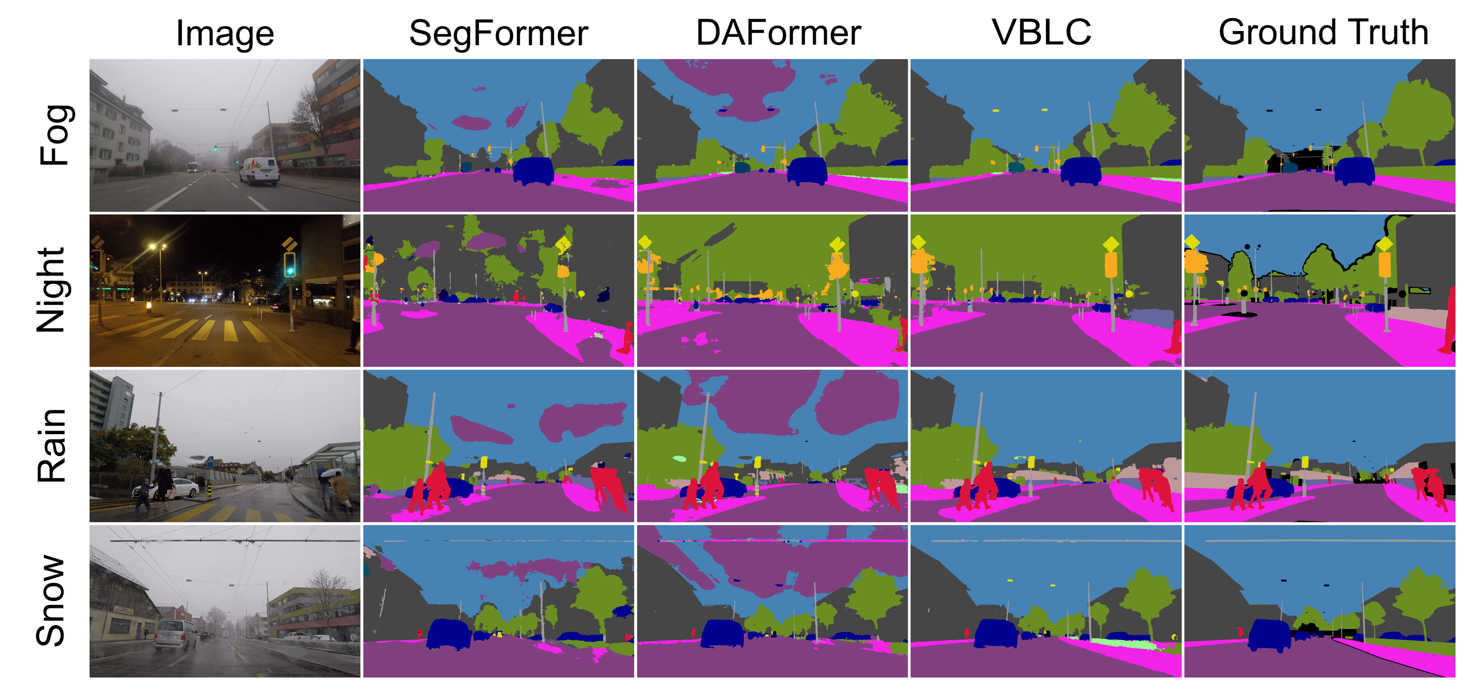
Method	road	side.	buil.	wall	fence	pole	light	sign	veg.	terr.	sky	pers.	rider	car	truck	bus	train	mbike	bike	mIoU
DeepLab-v2	96.7	72.4	74.1	28.6	41.4	42.2	49.8	67.6	72.6	62.5	80.6	70.4	54.4	88.4	56.1	72.4	33.7	42.7	70.1	61.9
FDA	87.0	56.9	82.1	4.3	11.6	36.3	41.8	60.4	80.6	51.6	70.6	66.7	50.3	86.0	46.4	63.7	26.2	41.4	66.3	54.2
DACS	97.9	82.3	**88.7**	40.8	42.4	41.0	53.5	67.3	**89.2**	58.2	**90.8**	70.8	54.4	91.3	62.9	82.5	**56.4**	47.0	72.4	67.9
VBLC	**98.6**	**86.9**	87.2	**62.1**	**55.3**	**54.2**	**65.1**	**77.8**	86.9	**66.8**	90.1	**77.5**	**63.2**	**93.7**	**77.3**	**86.6**	55.0	**59.4**	**79.5**	**74.9**
SegFormer	97.8	81.6	86.9	54.3	48.3	49.2	57.3	71.6	86.9	65.5	83.4	71.9	57.1	91.8	67.9	80.1	73.1	49.9	74.6	71.0
DAFormer	98.5	87.0	90.8	55.1	53.7	56.3	62.8	73.6	91.5	70.7	90.0	75.6	56.8	92.7	65.9	88.3	79.9	56.9	77.6	74.9
VBLC	**98.7**	**88.4**	**91.9**	**66.3**	**65.2**	**62.7**	**69.1**	**79.6**	**92.2**	**72.4**	**92.3**	**80.0**	**66.0**	**94.6**	**79.9**	**90.9**	**81.8**	**64.0**	**80.6**	**79.8**

Table 2: Comparison results on **Cityscapes** → **{Foggy,Rain}Cityscapes**.

Method	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Deformable DETR	13.4	22.7	13.4	3.4	17.0	26.8
SFA	14.3	24.6	14.6	4.2	17.4	28.2
SFA + **VBLC**	**15.6**	**26.0**	**16.4**	**5.2**	**17.8**	**30.9**

# Class	Method	IDD (mIoU)	Mapillary (mIoU)	Average mIoU
7	MTKT	68.3	69.3	68.8
	ADAS	70.4	**75.1**	72.7
	VBLC	**73.9**	71.7	**72.8**
19	CCL	**53.6**	51.4	52.5
	ADAS	48.3	53.6	50.5
	VBLC	52.9	**57.8**	**55.3**





References

- Narasimhan et. al, "Contrast restoration of weather degraded images", TPAMI 2003.
- He et. al, "Single image haze removal using dark channel prior", CVPR 2009.
- Sakaridis et. al, "Map-Guided Curriculum Domain Adaptation and Uncertainty-Aware Evaluation for Semantic Nighttime Image Segmentation", TPAMI 2022.