Similarity Min-Max: Zero-Shot Day-Night Domain Adaptation

Rundong Luo Wenjing Wang Wenhan Yang Jiaying Liu

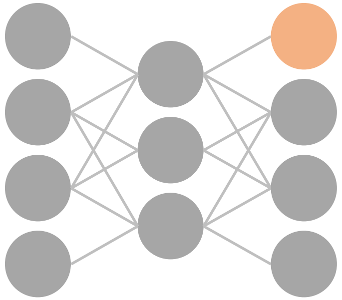
Presenter: Rundong Luo

Wangxuan Institute of Computer Technology, Peking University *Peng Cheng Laboratory*

Background

Low light hinders both human perception and **model performance**

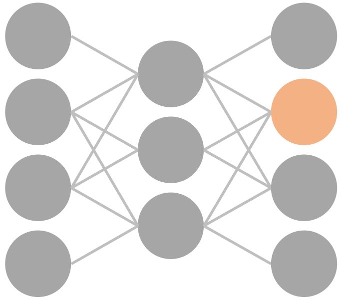
Normal light



Cat (1.00)

Performance drop

Low light



Chair (0.98) ×

Background

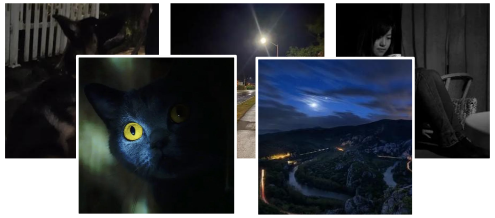
Solutions to improve the model's performance in nighttime scenarios:

Day-Night Domain Adaptation:

- Data: labeled daytime + unlabeled nighttime

Labeled daytime data

+



Unlabeled nighttime data

adapt

Nighttime conditions

Background

Solutions to improve the model's performance in nighttime scenarios:

Day-Night Domain Adaptation:

- Data: labeled daytime + unlabeled nighttime

Our task, Zero-Shot Day-Night Domain Adaptation:

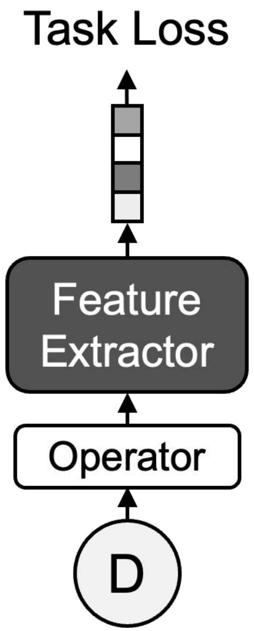
- Data: labeled daytime data **only**

Labeled daytime data

adapt

Nighttime conditions

Motivation

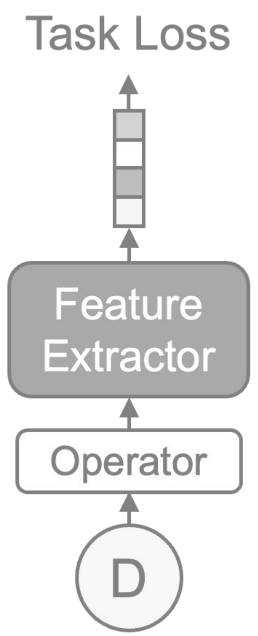


Operator-based [1]

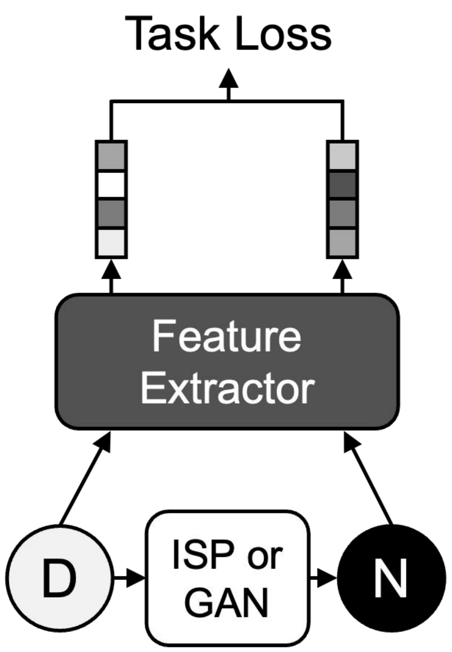
Model-Level: Not robust to complex real-world scenes

[1] Lengyel et al. Zero-Shot Day-Night Domain Adaptation with a Physics Prior. In ICCV, 2021.

Motivation



Operator-based [1]



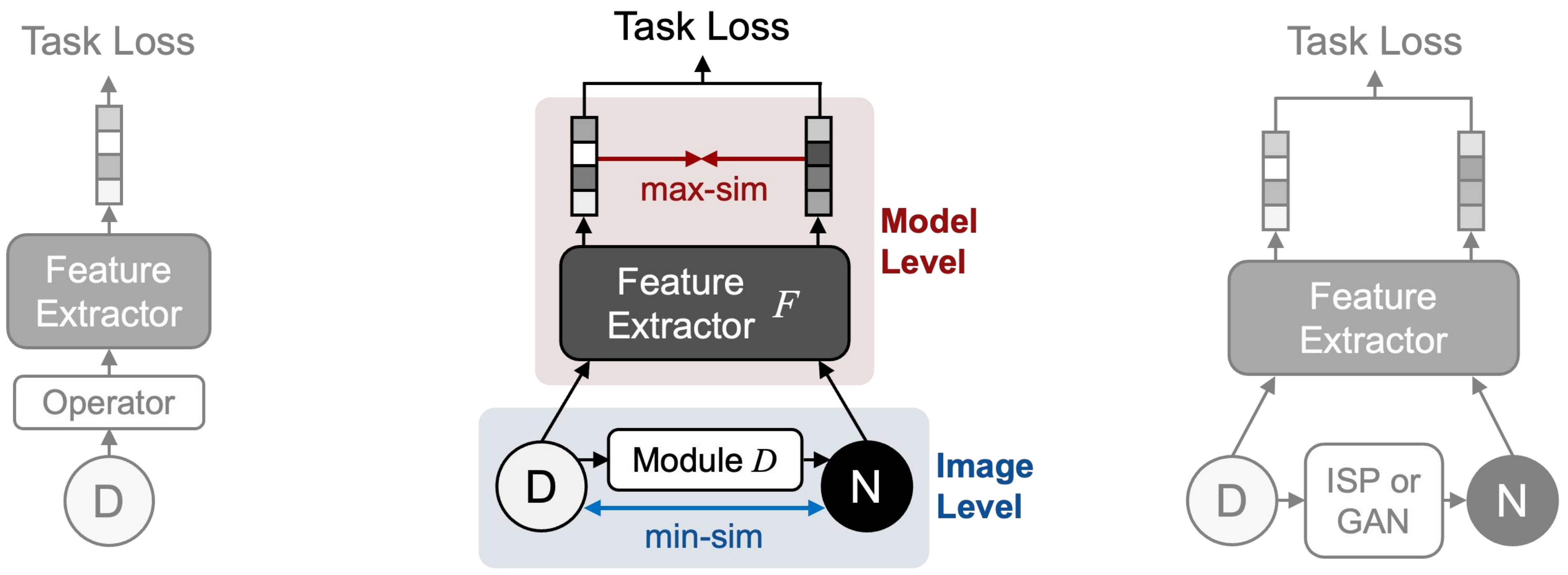
Darkening-based [2,3]

Image-Level
- ISP: device-dependent
- GAN: requires task-specific nighttime data

[1] Lengyel et al. Zero-Shot Day-Night Domain Adaptation with a Physics Prior. In ICCV, 2021.
[2] Cui et al. Multitask AET with Orthogonal Tangent Regularity for Dark Object Detection. In ICCV, 2021.
[3] Sakaridis et al. Guided Curriculum Model Adaptation for semantic nighttime image segmentation. In ICCV 2019.

Motivation

PEKING UNIVERSITY



Operator-based [1] | Our method | Darkening-based [2,3]

Joint image-level & model-level optimization

[1] Lengyel et al. Zero-Shot Day-Night Domain Adaptation with a Physics Prior. In ICCV, 2021.
[2] Cui et al. Multitask AET with Orthogonal Tangent Regularity for Dark Object Detection. In ICCV, 2021.
[3] Sakaridis et al. Guided Curriculum Model Adaptation for semantic nighttime image segmentation. In ICCV 2019.

Problem Formulation

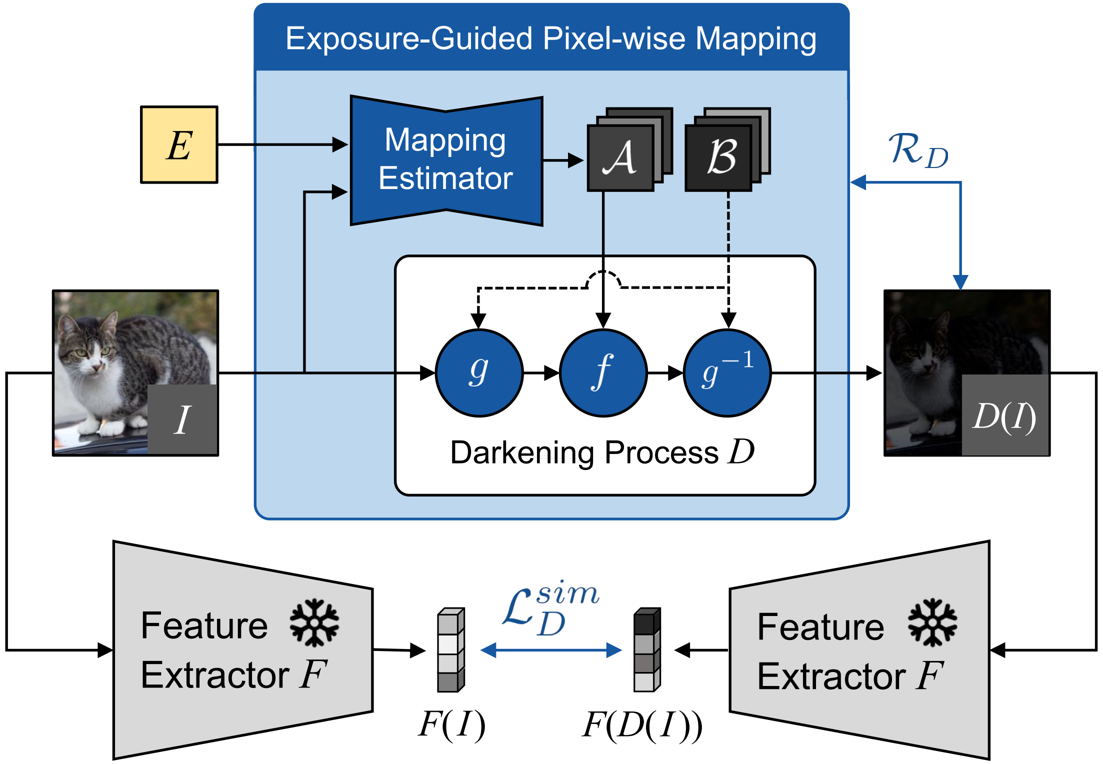
- Given a feature extractor F, we hope it to be robust to illumination changes, *i.e., the* daytime image I and its darkened version $D(I)$ should have similar representations.
- Formulation: $$\max_{\theta_F} \min_{\theta_D} \quad \mathrm{Sim}(F(I), F(D(I))) \quad (1)$$
- Add regularization to avoid trivial solutions:

$$\max_{\theta_F} \min_{\theta_D} \ \mathrm{Sim}(F(I), F(D(I))) + \mathcal{R}_D(\theta_D) - \mathcal{R}_F(\theta_F) \quad (2)$$

- How to design darkening module D and regularizer R_D, R_F?

Training Framework

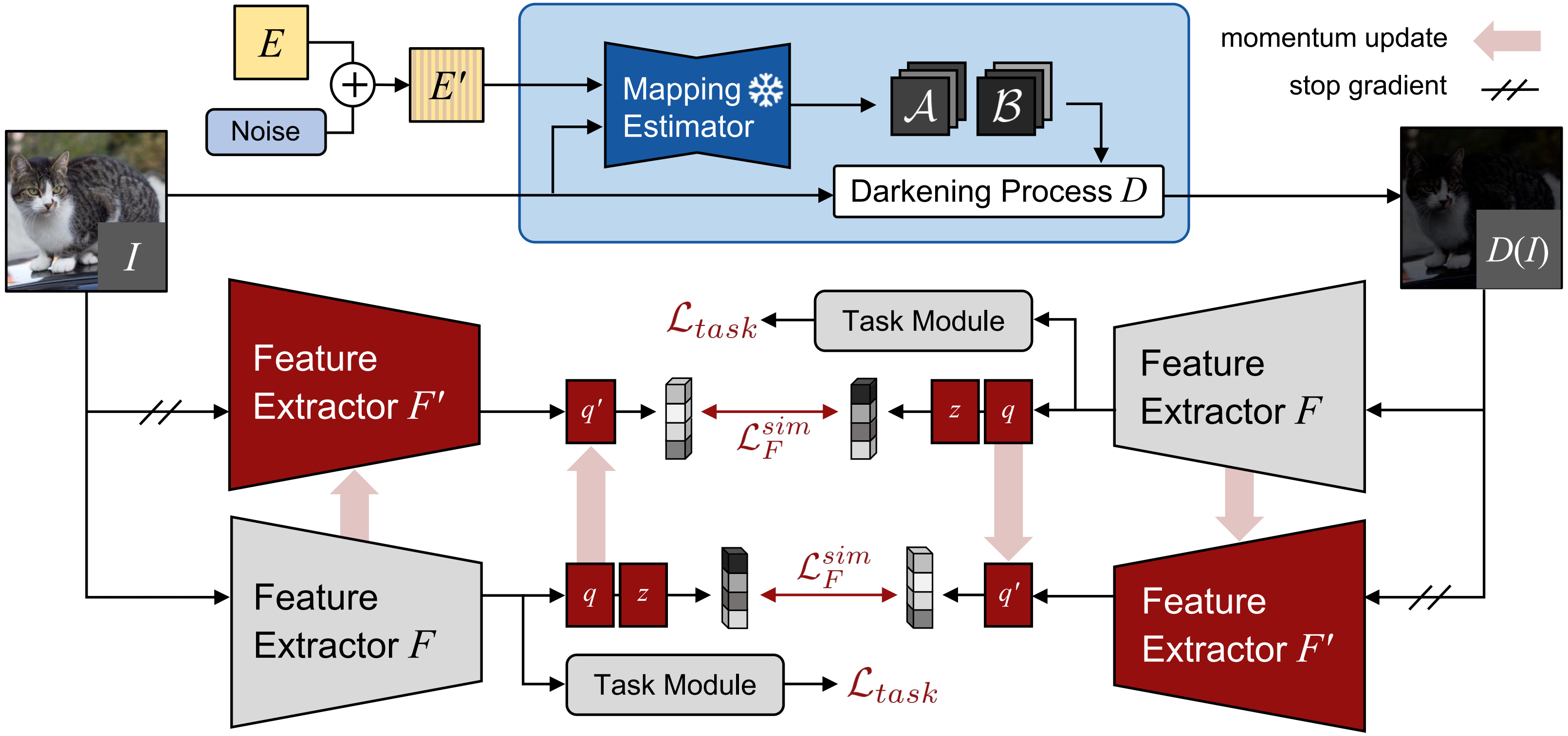
- Image-Level Darkening
  - **Regularization ($\mathcal{R}_D$)**
    - Total variance loss
    - Color consistency loss
    - Exposure alignment loss
  - **Similarity loss ($\mathcal{L}_D^{sim}$)**
    - Cosine similarity



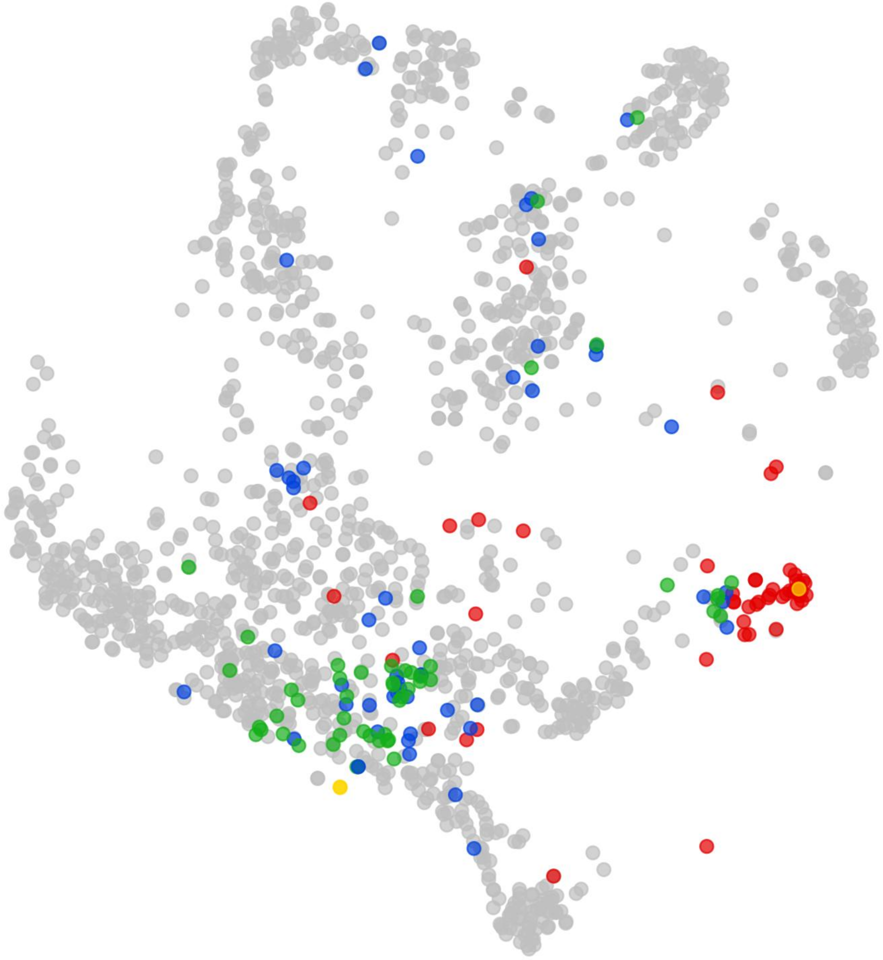
PEKING UNIVERSITY

Training Framework

- Image-Level Darkening
- Model-Level Adaptation
  - **Regularization ($\mathcal{R}_F$)**
    - Task loss
  - **Similarity loss ($\mathcal{L}_F^{sim}$)**
    - BYOL loss

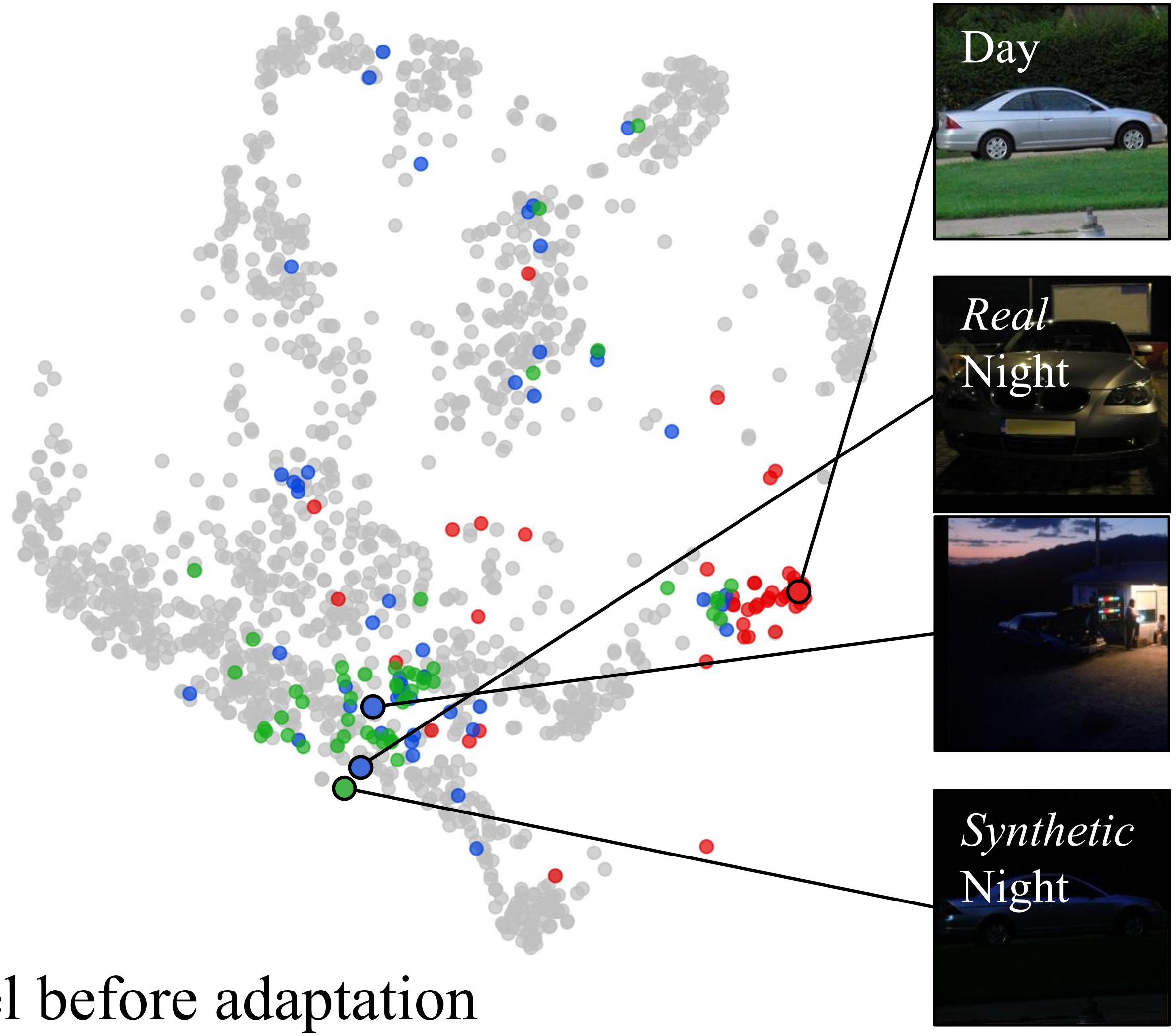


Proof of Concept



Model before adaptation

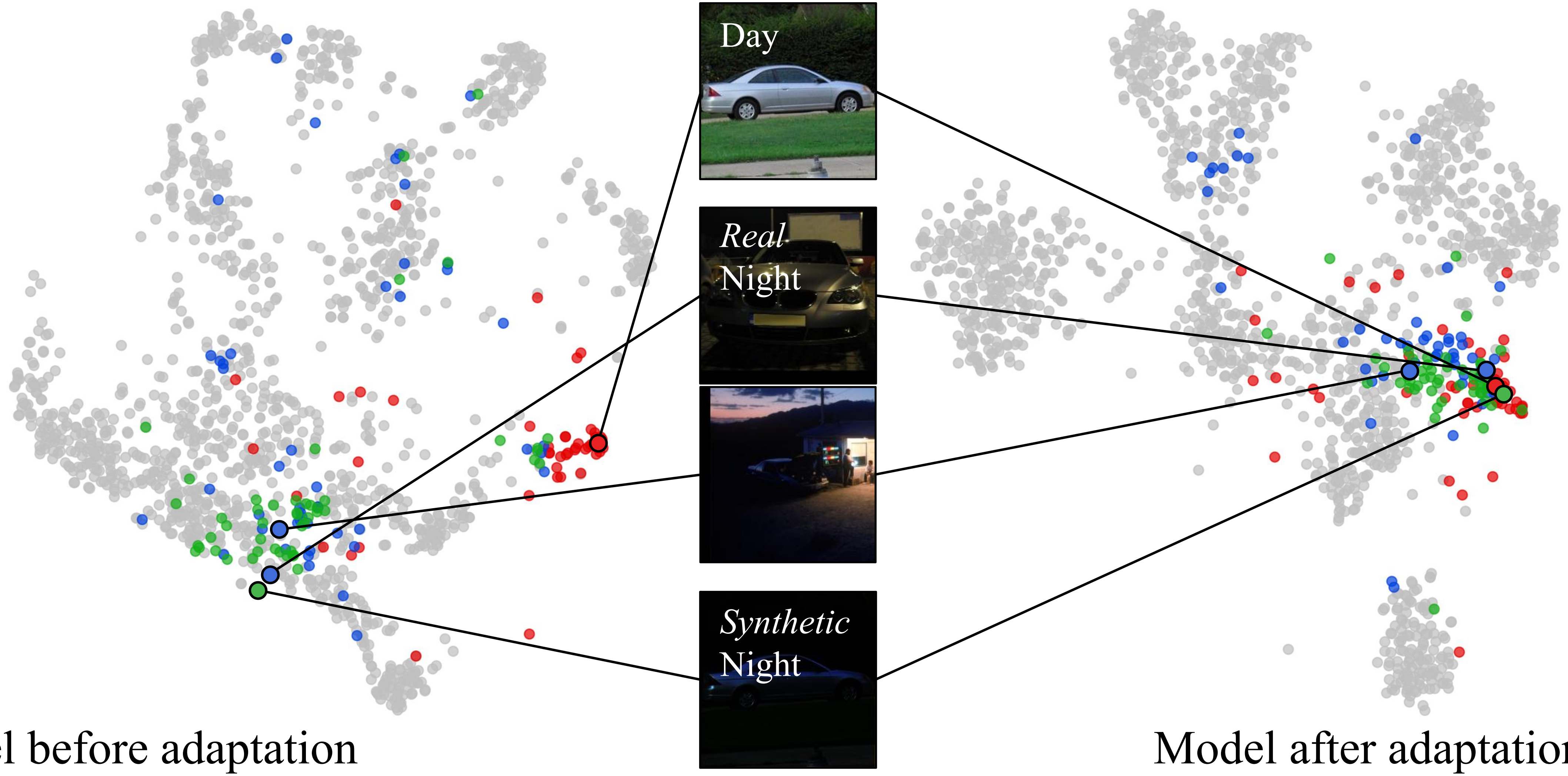
Proof of Concept



Model before adaptation

PEKING UNIVERSITY 北京大學

Proof of Concept



Model before adaptation

Model after adaptation

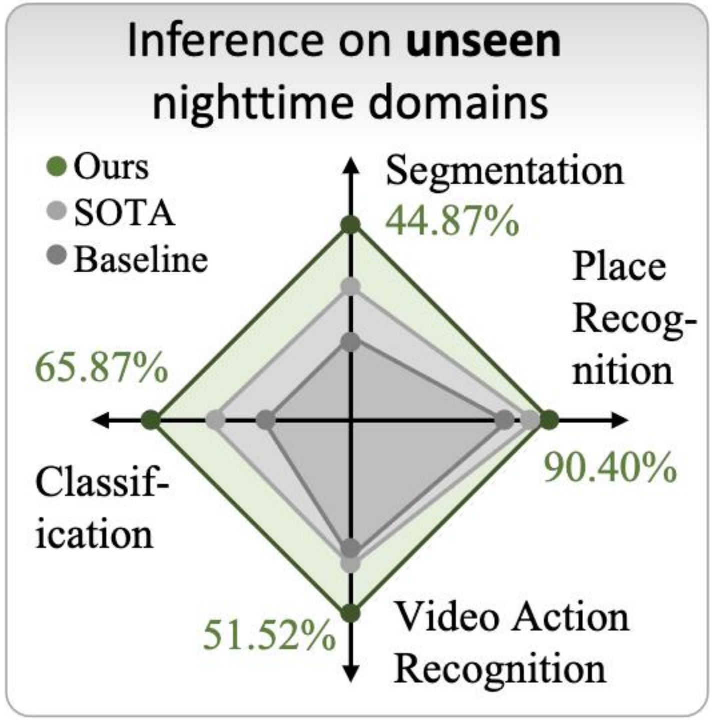
Tasks and datasets for evaluation

- Low-Light image classification: CODaN
- Nighttime semantic segmentation: Nighttime Driving, Dark-Zurich
- Visual place recognition: Tokyo 24/7
- Low-Light video action recognition: ARID

Baselines

- Low-light image (video) enhancement
- Zero-shot domain adaptation
- Domain generalization

Experiments



I. Classification

Method	Top-1 (%)
ResNet-18 [18]	53.32
Low-Light Enhancement	
EnlightenGAN [23]	56.68
LEDNet [63]	57.40
Zero-DCE++ [30]	57.96
RUAS [33]	58.36
SCI [34]	58.68
URetinexNet [56]	58.72
Domain Generalization	
MixStyle [62]	53.12
IRM [1]	54.52
AdaBN [31]	54.25
Zero-Shot Day-Night Domain Adaptation	
MAET† [8]	56.48
CIConv [29]	60.32
Ours	**65.87**

II. Segmentation

Method	Nighttime Driving	Dark-Zurich
RefineNet [32]	34.3	30.6
Low-Light Enhancement		
EnlightenGAN [23]	25.2	24.9
Zero-DCE++ [30]	32.7	28.3
RUAS [33]	25.1	23.4
SCI [34]	28.6	25.7
URetinexNet [56]	28.1	24.0
LEDNet [63]	27.6	26.6
Domain Generalization		
AdaBN [31]	37.2	31.1
RobustNet [6]	33.0	34.5
SAN-SAW [38]	28.1	16.0
Zero-Shot Day-Night Domain Adaptation		
MAET [8]	28.1	26.4
CIConv [29]	41.2	34.5
Ours	**44.9**	**40.2**

III. Visual Place Recognition

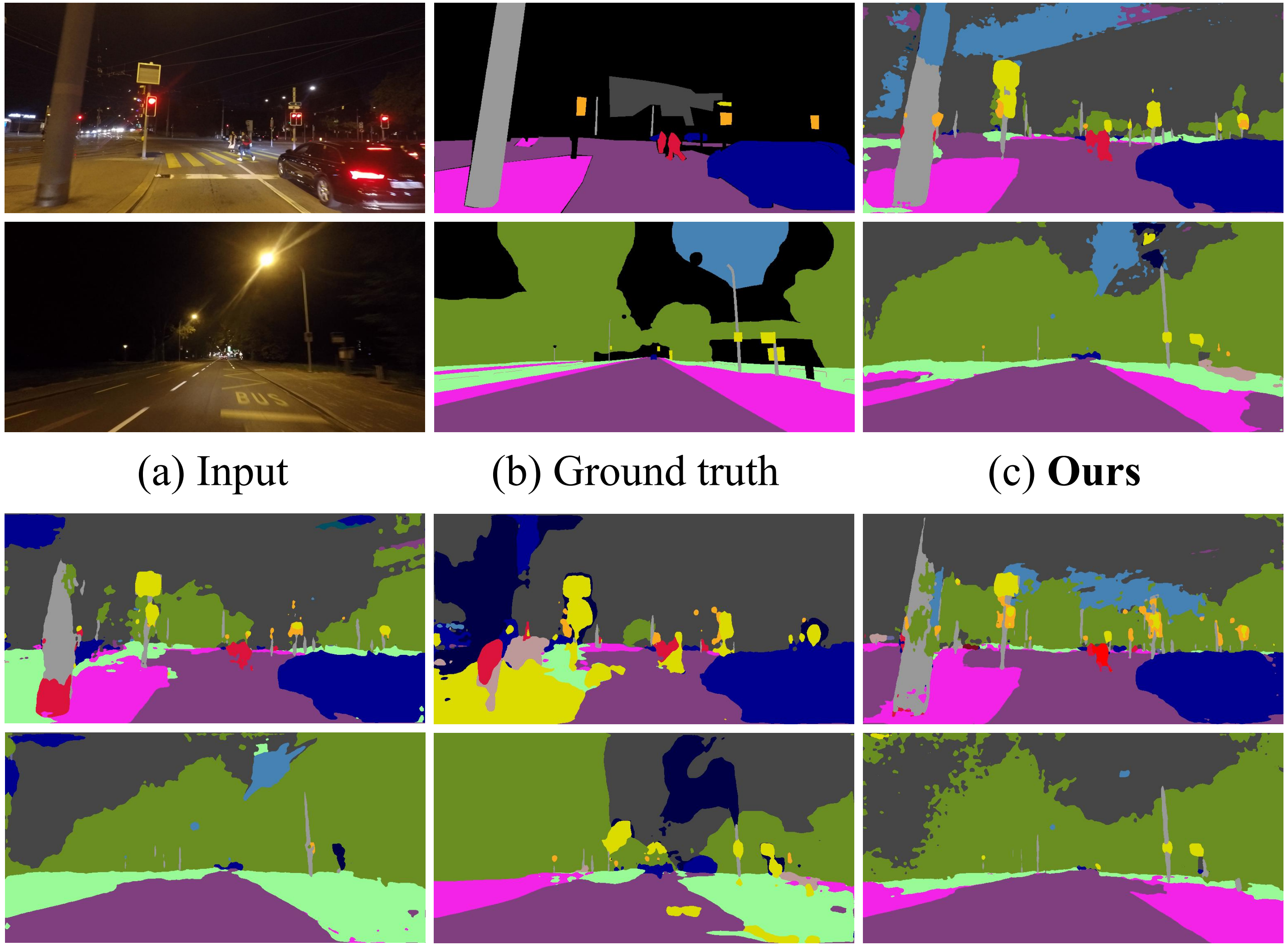
Method	mAP (%)
Zero-Shot Day-Night Domain Adaptation	
EdgeMAC [42]	75.9
U-Net jointly [21]	79.8
GeM [43]	85.0
CIConv-GeM [29]	88.3
Ours-GeM	**90.4**
Day-Night Domain Adaptation (night images are available for training)	
U-Net jointly [21]	86.5
EdgeMAC + CLAHE [21]	90.5
EdgeMAC + U-Net jointly [21]	90.0

IV. Video Action Recognition

Method	Top-1 (%)
I3D [3]	47.02
Low-Light Video Enhancement	
StableLLVE [59]	45.08
SMOID [22]	47.27
SGZ [61]	46.42
Domain Generalization & Zero-Shot Day-Night Domain Adaptation	
AdaBN [31]	46.17
Ours	**51.52**

II. Nighttime Segmentation

Method	Nighttime Driving	Dark-Zurich
RefineNet [32]	34.3	30.6
Low-Light Enhancement		
EnlightenGAN [23]	25.2	24.9
Zero-DCE++ [30]	32.7	28.3
RUAS [33]	25.1	23.4
SCI [34]	28.6	25.7
URetinexNet [56]	28.1	24.0
LEDNet [63]	27.6	26.6
Domain Generalization		
AdaBN [31]	37.2	31.1
RobustNet [6]	33.0	34.5
SAN-SAW [38]	28.1	16.0
Zero-Shot Day-Night Domain Adaptation		
MAET [8]	28.1	26.4
CIConv [29]	41.2	34.5
Ours	**44.9**	**40.2**



(a) Input (b) Ground truth (c) **Ours**

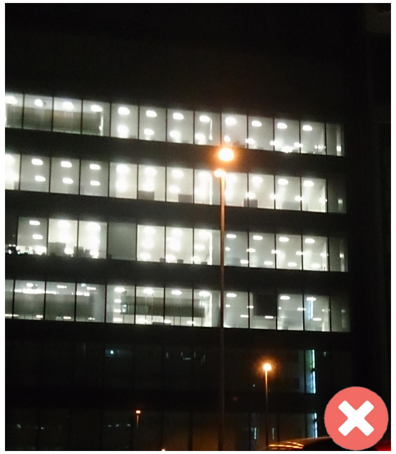
(d) RefineNet (e) MAET (f) CIConv

PEKING UNIVERSITY 北京大學

III. Visual Place Recognition

Method	mAP (%)
Zero-Shot Day-Night Domain Adaptation	
EdgeMAC [42]	75.9
U-Net jointly [21]	79.8
GeM [43]	85.0
CIConv-GeM [29]	88.3
Ours-GeM	**90.4**
Day-Night Domain Adaptation (night images are available for training)	
U-Net jointly [21]	86.5
EdgeMAC + CLAHE [21]	90.5
EdgeMAC + U-Net jointly [21]	90.0

(a) Query



(b) GeM

(c) **Ours**

III. Low-Light Video Action Recognition

Conclusion

- **Task:** zero-shot day-night domain adaptation
- **Framework:** similarity min-max framework
  - A bi-level optimization problem
  - Two-stage training: image-level darkening and model-level adaptation
- **Performance:** state-of-the-art results across multiple downstream tasks.

Thank You

Poster Info:

Room Nord, No. 168

Oct. 5th, 10:30-12:30

Project

Laboratory

Contact: Rundong Luo

rundong_luo@stu.pku.edu.cn / red-fairy.github.io