NEURAL INFORMATION PROCESSING SYSTEMS

FRBNet: Revisiting Low-Light Vision through Frequency-Domain Radial Basis Network

Fangtong Sun[*], Congyu Li[*], Ke Yang, Yuchen Pan, Hanwen Yu, Xichuan Zhang[†], Yiying Li[†]

Intelligent Game and Decision Lab(IGDL), Beijing, China

{sunfangtong19,liyiying10}@nudt.edu.cn, licongyu@hnu.edu.cn, zhxc@alu.hit.edu.cn

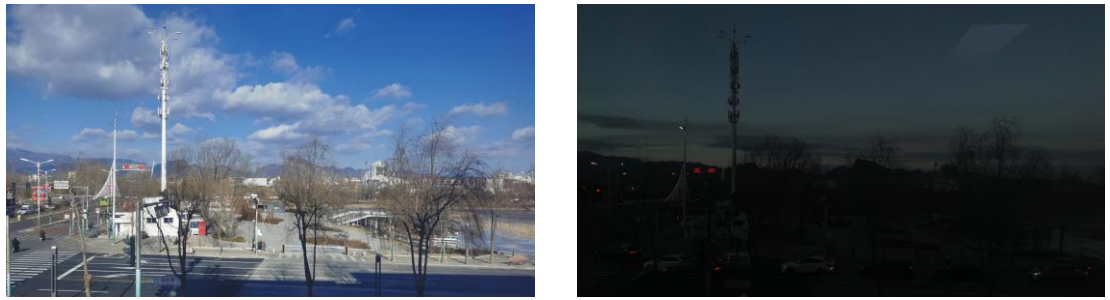
Paper

Code

Background

☐ Low-light Vision

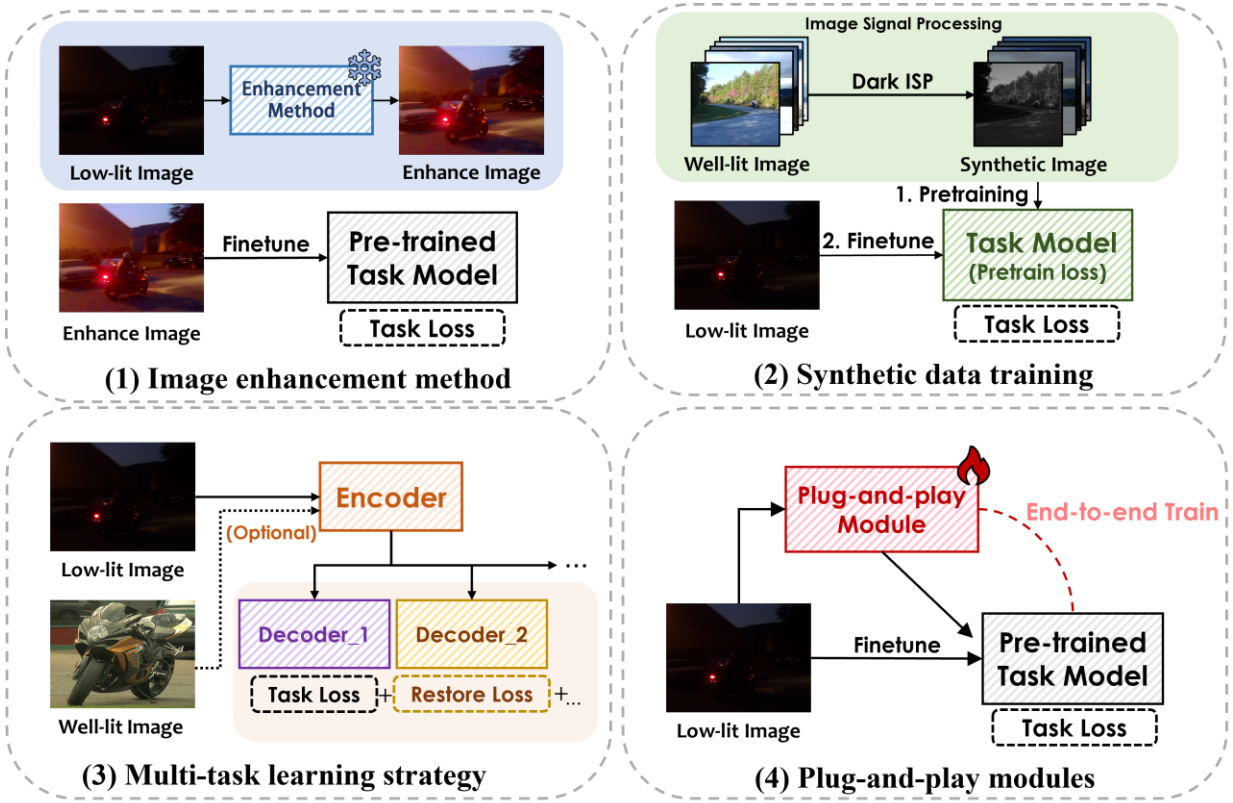
➢ Domain Shift



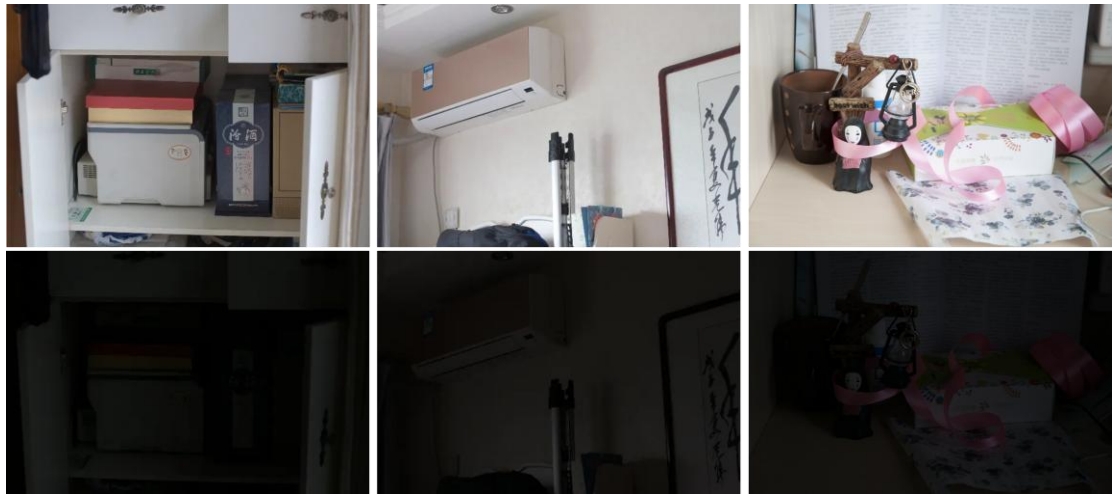
Well-light Image vs. Low-light Image

☐ Four Paradigms for Low-light Vision

- ➢ Image enhancement methods
- ➢ Synthetic data training
- ➢ Multi-task learning strategy
- ➢ Plug-and-play modules



Motivation



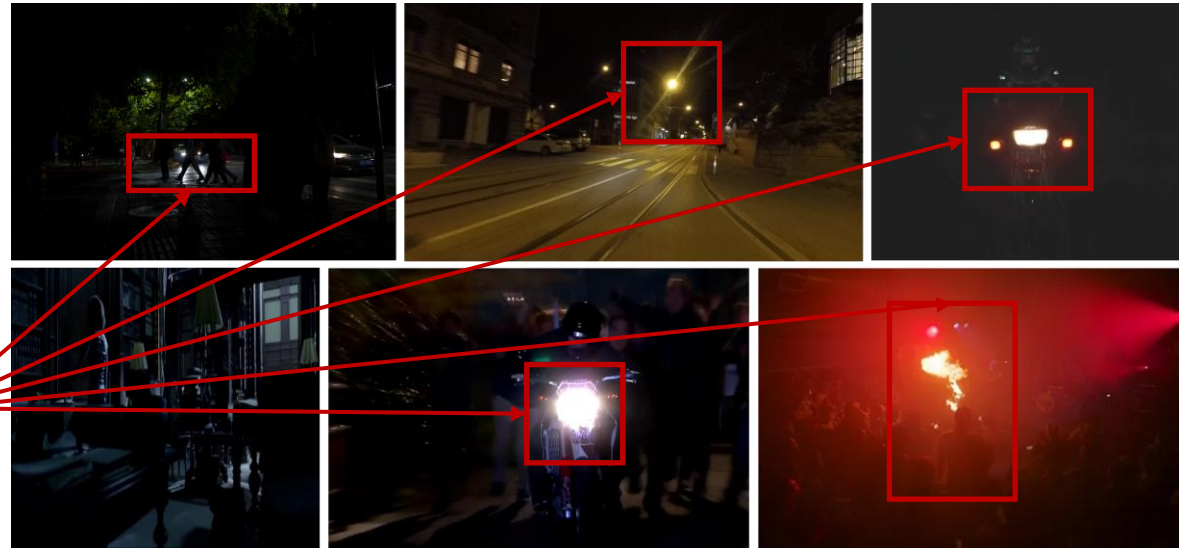
Synthetic low-light data

Classical Lambertian Image Formation Model

$$I_C(x,y) = m[\vec{n}(x,y), \vec{l}(x,y)] \cdot \varphi_C(x,y) \cdot \rho_C(x,y)$$

Assumes idealized diffuse reflection !

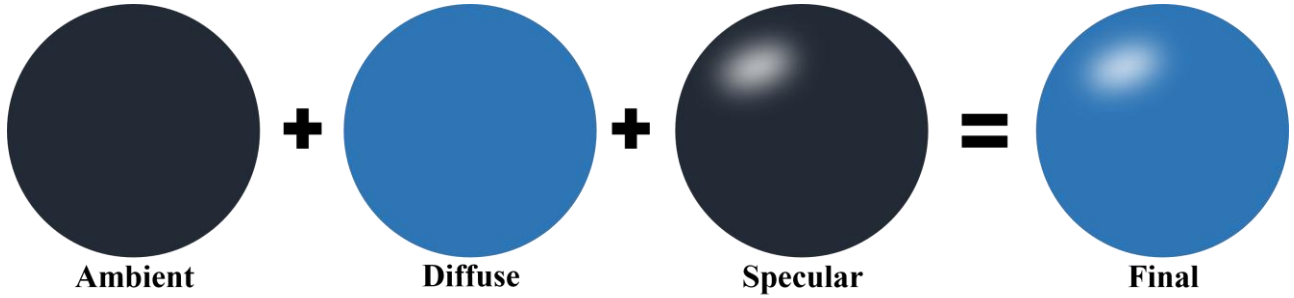
Complex and spatially localized light sources in real-world (streetlights, vehicle headlights, and neon signs)



Real-world low-light data

Motivation

☐ Phong Lighting Model Imaging Mechanism[1]



☐ An extended version of the Lambertian model adapted to real-world low-light scenes:

$$I_C(x,y) = m[\vec{n}(x,y), \vec{l}(x,y)] \cdot \varphi_C(x,y) \cdot \rho_C(x,y) + S_C(x,y)$$

Retain for illumination-invariant features

Spatially irregular highlight component

[1] Bui Tuong Phong. Illumination for computer generated pictures. In Seminal Graphics: Pioneering Efforts that Shaped the Field(SGRAPH), pages 95–101. ACM Press, 1998

Method

☐ Analysis and Derivation

$$\begin{aligned}\mathrm{FCR}_{RG} &= \mathcal{F}[\log(\frac{I_R}{I_G})] \\ &= \mathcal{F}[\log\varphi_R - \log\varphi_G] + \mathcal{F}[\log\rho_R - \log\rho_G] + \mathcal{F}[\log(1+H_R) - \log(1+H_G)]\end{aligned}$$

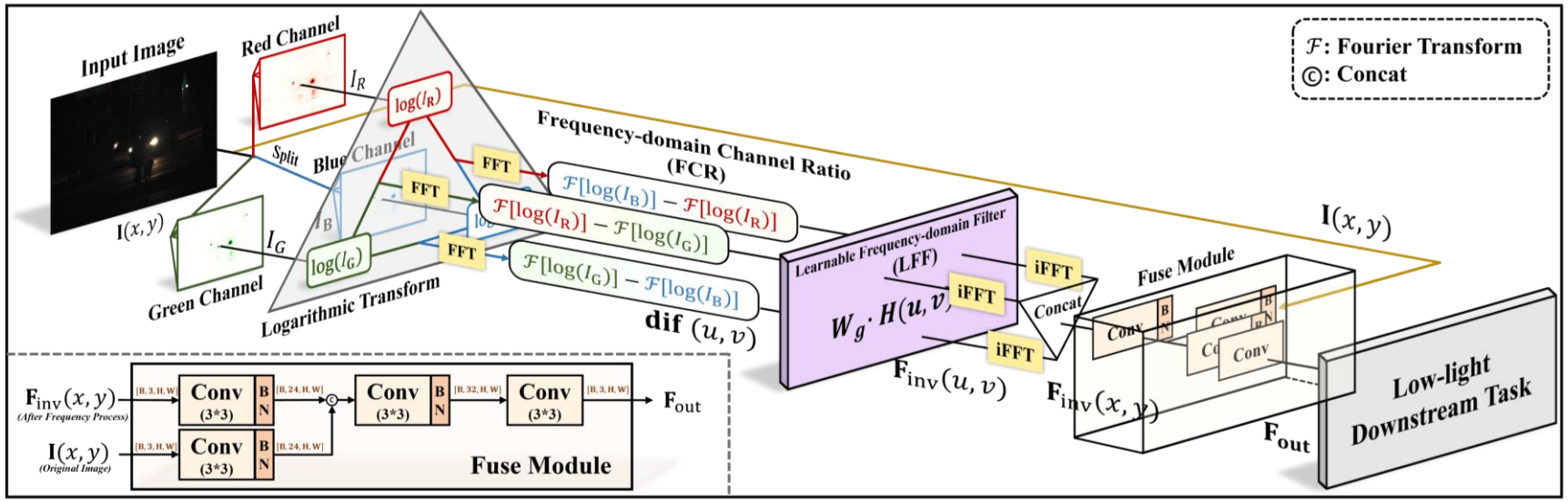
$$\begin{aligned}\Delta &= \mathcal{F}[H_R - H_G] = \mathcal{H}_R - \mathcal{H}_G \\ &= a_R \cdot e^{i\theta_R} - a_G \cdot e^{i\theta_G} \\ &= e^{i\theta_R} \cdot \left(a_R - a_G \cdot e^{i(\theta_G - \theta_R)}\right) \\ &= e^{i\theta_R} \cdot (a_R - a_G \cdot Cor_{RG})\end{aligned}$$

This yields:

$$\mathrm{FCR}_{RG} = \underbrace{\mathcal{F}[\log\varphi_R - \log\varphi_G]}_{\text{illumination}} + \underbrace{\mathcal{F}[\log\rho_R - \log\rho_G]}_{\text{reflectance}} + \underbrace{e^{i\theta_R}(a_R - a_G \cdot Cor_{RG})}_{\text{high-lit residual}}$$

Method

☐ Overall Pipeline of FRBNet



➢ Frequency-domain Channel Ratio

$$\begin{cases} F_{\text{inv}}^{RG}(u,v) = LFF^{RG}(u,v) \cdot \text{dif}^{RG}(u,v) \\ F_{\text{inv}}^{GB}(u,v) = LFF^{GB}(u,v) \cdot \text{dif}^{GB}(u,v) \\ F_{\text{inv}}^{BR}(u,v) = LFF^{BR}(u,v) \cdot \text{dif}^{BR}(u,v) \end{cases}$$

➢ Fuse Module

$$\mathbf{F}_{\text{out}} = \text{Conv}\left\{\text{CB}\left[\text{Cat}\left(\text{CB}[\mathbf{F}_{\text{inv}}(x,y)]; \text{CB}[\mathbf{I}(x,y)]\right)\right]\right\}$$

➢ Learnable Frequency-domain Filter

$$\mathbf{LFF}(u,v) = \mathbf{W_g} \cdot \mathbf{H}(u,v)$$

$$\begin{cases} \mathbf{W_g}(u,v) = \exp\left(-\dfrac{\mathbf{r}(u,v)^2}{\sigma_w^2}\right), \quad \mathbf{r}(u,v) = \sqrt{u^2+v^2} \\ \mathbf{H}(u,v) = \Phi(u,v) \cdot M(u,v) \end{cases}$$

$$M(u,v) = 1 + \lambda \cdot \sum_{n=1}^{N} \left[\cos(n\theta(u,v)) + \sin(n\theta(u,v))\right]$$

Result

☐ Low-light Detection Tasks (Object & Dark Face Detection)

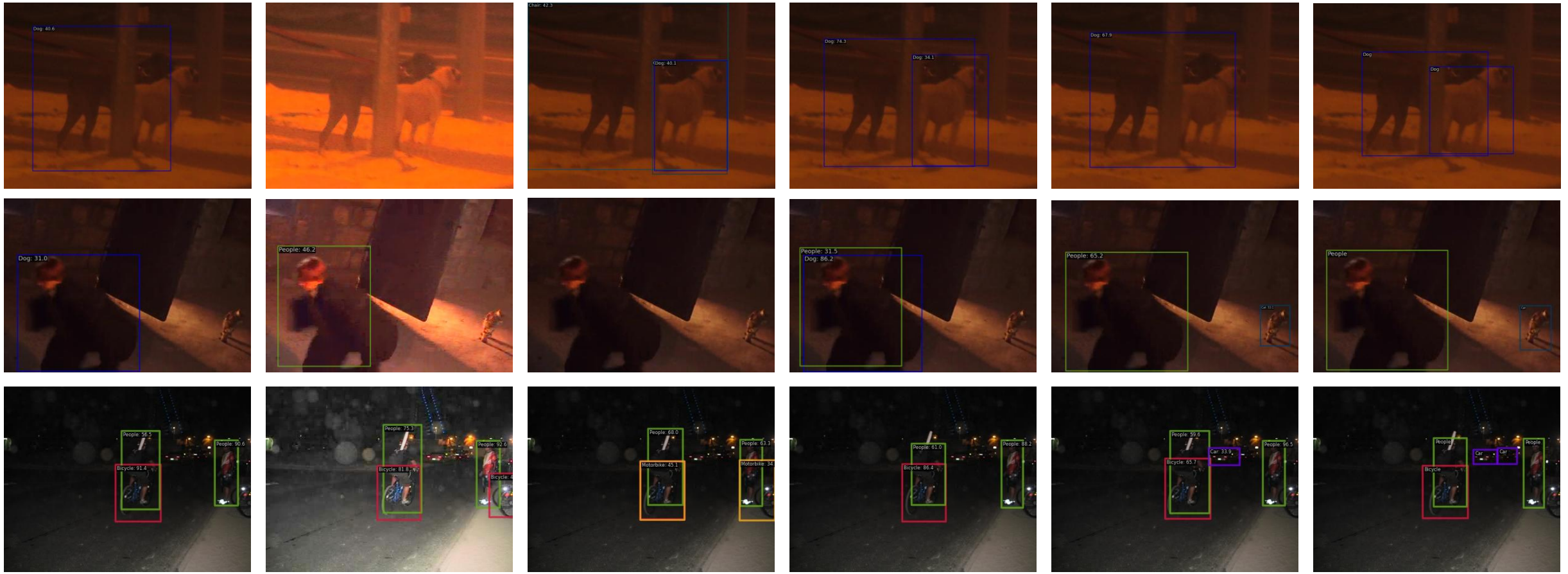
Paradigm	Method	ExDark				DarkFace			
		YOLOv3		TOOD		YOLOv3		TOOD	
		Recall	mAP	Recall	mAP	Recall	mAP	Recall	mAP
	Baseline	84.6	71.0	91.9	72.5	73.8	54.8	80.9	57.0
Enhancement	SMG(*CVPR-23*)	82.3	68.5	91.8	71.5	73.4	52.4	80.2	56.3
	NeRCo(*ICCV-23*)	83.4	68.5	91.8	71.8	73.8	53.0	79.4	56.8
	LightDiff(*ECCV-24*)	84.3	71.3	92.1	72.9	75.5	57.4	81.0	58.7
	DarkIR(*CVPR-25*)	81.9	68.2	90.9	72.0	74.5	55.9	81.4	60.4
Synthetic Data	DAINet*(*CVPR-24*)	86.7	73.4	-	-	74.8	56.9	-	-
	WARLearn(*WACV-25*)	85.6	72.4	92.8	73.4	74.5	56.2	80.8	59.4
Multi-task	MAET(*ICCV-21*)	85.1	72.5	92.5	74.3	74.7	55.7	80.7	59.6
	IAT(*BMVC-22*)	85.0	72.6	92.9	73.0	73.6	55.5	79.7	58.3
Plug-and-play	DENet(*ACCV-22*)	84.2	71.3	92.6	73.5	71.8	52.6	73.6	49.6
	FeatEnHancer(*ICCV-23*)	90.4	71.2	**96.4**	74.6	74.1	55.2	81.7	60.5
	YOLA(*NeurIPS-24*)	86.1	72.7	93.8	75.2	74.9	56.3	**83.1**	63.2
	FRBNet(ours)	**90.6**	**74.9**	93.2	**75.4**	**75.7**	**57.7**	82.7	**65.1**

☐ Efficiency-Performance Trade-off

Category	Metric	Non-architectural Methods				End-to-End Trained Plug-and-Play Module			
		KinD	Zero-DCE	SMG	MAET	DENet	FeatEnHancer	YOLA	FRBNet
Efficiency	# Params ↓	8.2M	79K	17.9M	40M	40K	138K	**8K**	9K
	Flops(G) ↓	**50.6**				61.7	79.5	55.0	53.1
	FPS(img/s)↑	**95.8**				83.8	33.1	81.1	89.5
Performance	Det(mAP) ↑	69.4	71.1	68.5	72.5	71.3	71.2	72.7	**74.9**
	Seg(mIoU) ↑	48.1	48.7	49.7	-	52.2	56.0	58.7	**61.6**

Result

☐ Low-light Detection Tasks



Baseline DarkIR FeatEnHancer YOLA FRBNet(Ours) GT

Result

Low-light Semantic Segmentation

Method	RO	SI	BU	WA	FE	PO	TL	TS	VE	TE	SK	PE	CA	TR	BI	mIoU
Baseline	90.0	61.4	74.2	32.8	34.4	45.7	49.8	31.2	68.8	14.6	80.4	27.1	62.1	76.3	14.4	50.8
RetinexNet	89.4	61.0	70.6	30.1	28.1	42.4	47.6	25.7	65.8	8.6	77.3	21.5	54.8	67.4	8.2	46.5
DRBN	90.5	61.5	72.8	31.9	32.5	44.5	47.3	27.2	65.7	10.2	76.5	24.2	55.4	71.1	11.9	48.2
FIDE	90.0	60.7	72.8	32.4	34.1	43.3	47.9	26.1	67.0	13.7	78.0	26.5	57.1	71.0	12.4	48.8
KinD	90.0	61.0	73.2	31.9	32.8	43.5	42.7	27.7	65.5	13.3	77.4	22.8	55.1	74.5	11.5	48.1
EnGAN	89.7	58.9	73.7	32.8	31.8	44.7	49.2	26.2	67.3	14.2	77.8	25.0	59.0	71.2	7.8	48.6
ZeroDCE	90.6	59.9	73.9	32.6	31.7	44.3	46.2	25.8	67.2	14.6	79.1	24.7	59.4	66.8	13.9	48.7
SSIENet	89.6	59.3	72.5	29.9	31.7	45.4	43.9	24.5	66.7	10.6	78.3	22.8	52.6	71.1	5.4	46.9
Xue *et al.*	93.2	72.6	78.4	43.8	**46.5**	48.1	51.1	38.8	68.6	**14.9**	79.1	21.9	61.6	85.2	36.1	55.8
FeatEnHancer	93.5	70.6	75.6	41.8	33.4	51.3	55.2	35.9	68.5	13.4	80.6	27.6	61.8	80.0	51.2	56.0
YOLA	93.2	72.1	79.3	41.1	39.1	**53.1**	60.4	44.4	71.5	4.7	83.2	37.8	66.8	85.0	49.2	58.7
FRBNet(ours)	**94.4**	**75.5**	**79.7**	**46.0**	45.4	52.3	**64.9**	**50.8**	**72.2**	9.5	**84.2**	**40.9**	**70.4**	**88.7**	**49.3**	**61.6**

Low-light Instance Segmentation

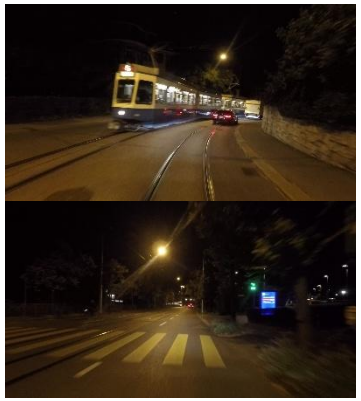
Method	mAP	mAP_{50}	mAP_{75}
Mask RCNN	23.7	41.5	23.3
MBLLEN	22.5	40.7	22.3
DarkIR	27.4	46.3	27.5
YOLA	24.9	44.8	24.2
FeatEnHancer	29.1	48.7	29.7
FRBNet(ours)	**30.2**	**50.5**	**30.4**

Ablation Study

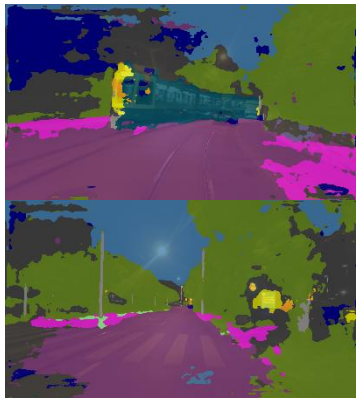
	$\mathbf{H}(u,v)$	$\mathbf{W_g}$	FCR	ExDark	DarkFace
Baseline				71.0	57.0
Ablation Cases	✓			72.5	62.0
	✓	✓		72.9	62.5
	✓		✓	73.5	63.7
FRBNet	✓	✓	✓	**74.9**	**65.1**

Result

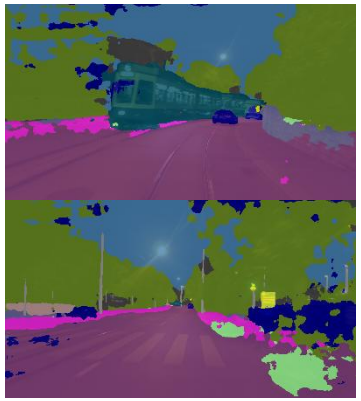
☐ Low-light Segmentation Tasks



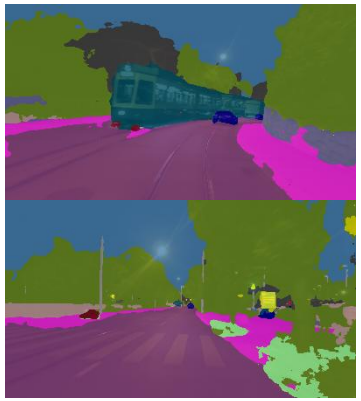
Input



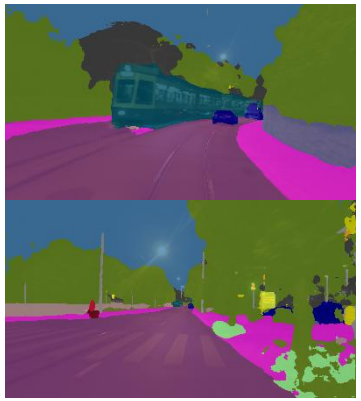
Baseline



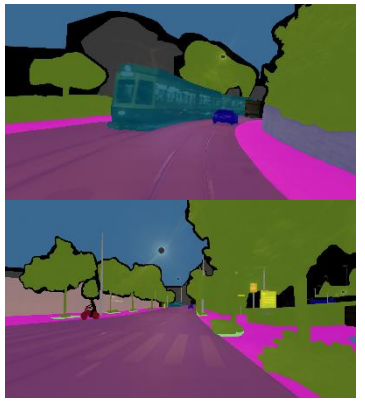
FeatEnHancer



YOLA



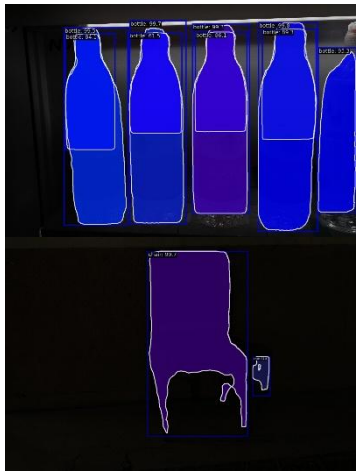
FRBNet(Ours)



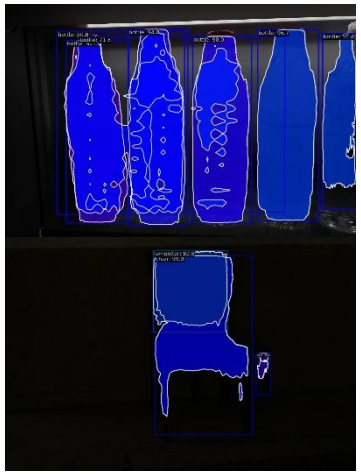
GT

Baseline

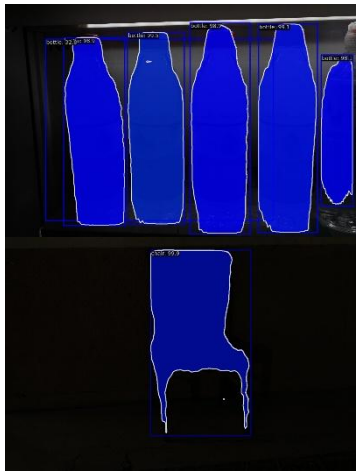
DarkIR



FeatEnHancer



YOLA



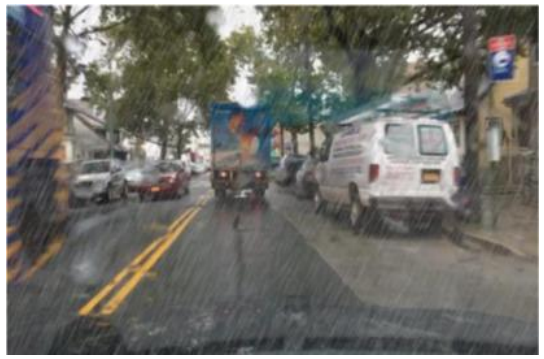
FRBNet(Ours)

GT

Conclusion

- **FRBNet**
  - For extracting illumination-invariant feature
  - Plug-and-play module
  - Superior performance

- **Future research**
  - Exploring broader application



NEURAL INFORMATION PROCESSING SYSTEMS

Thanks for Watching!

Fangtong Sun