TorchAdapt: Towards Light-Agnostic Real-Time Visual Perception

Khurram Azeem Hashmi[1,2] Karthik Palyakere Suresh[1,2] Didier Stricker[1,2] Muhammad Zeshan Afzal[1,2]

[1]DFKI [2]RPTU Kaiserslautern-Landau

ICCV OCT 19-23, 2025 HONOLULU HAWAII

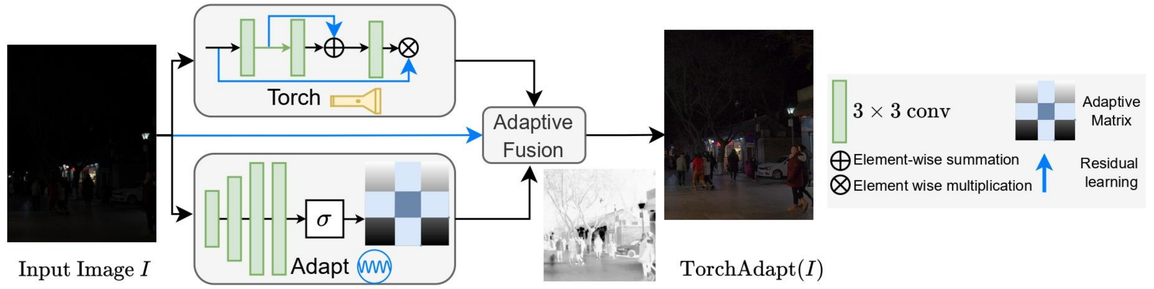
Introduction

- We introduce TorchAdapt, a real-time, general purpose, light-agnostic adaptive feature enhancement framework that boosts the performance of high-level vision tasks under varying illumination conditions.

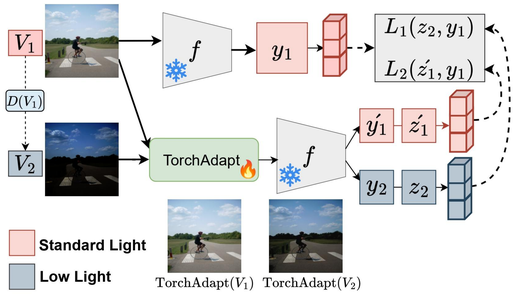
Contributions

- **Method:** We propose TorchAdapt, a realtime, light-agnostic adaptive feature enhancement framework that improves the performance of any high-level vision task under varying illumination without compromising accuracy in well-lit conditions.
- **Simplicity:** Thanks to its architectural innovations and self-supervised objective, TorchAdapt significantly simplifies the training strategy for any high-level vision task pipeline, circumventing corresponding paired images/videos.
- **Flexibility:** TorchAdapt is independent of vision encoders and downstream task architectures, demonstrating model-agnostic generality. Any zero-reference LLIE method can serve as the Torch module in our framework to achieve light-agnostic performance.
- **Results:** We conduct extensive experiments on five visual perception tasks involving images and videos under both low-light and lightagnostic settings using curated datasets. TorchAdapt outperforms state-of-the-art task-specific and LLIE methods across all settings when integrated with the same baselines.

Proposed Pipeline



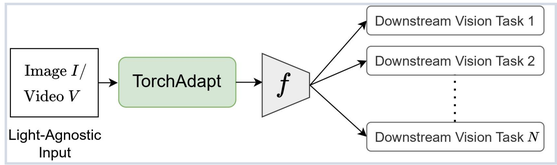
TorchAdapt architectural overview



Light-Agnostic learning pipeline

$$\mathbf{E} = \mathbf{I} + ((\gamma \cdot \mathbf{M}) \odot \mathbf{T}). \quad (4)$$

Explaining adaptive fusion to generate final enhanced image E.



TorchAdapt's downstream integration pipeline

Ablation Studies

#	Torch(§ 3.1)	Adapt(§ 3.2)	Scaling(§ 3.2)	Object Det. LL	Object Det. LA	Instance Seg. LL	Instance Seg. LA
1	×	×	×	78.8	62.4	46.0	33.9
2	✓	×	×	79.1	62.9	51.1	34.8
3	✓	✓	×	79.9	63.5	53.3	36.2
4	✓	×	✓	79.5	62.7	51.4	34.1
5	✓	✓	✓	**80.1**	**64.0**	**53.6**	**36.6**

Contribution of each component. Highlighted settings are set as default.

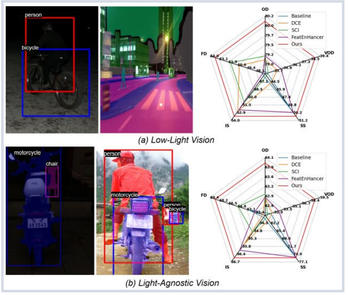
Scale factor (κ)	OD LL	OD LA	SS LL	SS LA
2	79.9	63.7	53.4	36.4
3	80.1	64.0	53.6	36.6
4	79.9	63.3	53.2	36.2
5	79.5	61.1	53.3	34.5

scale factor in TorchAdapt. Highlighted settings are set as default. OD and SS are object detection and semantic segmentation.

$D(\mathbf{I})$	OD LL	OD LA	SS LL	SS LA
ISP [16]	79.4	61.7	47.6	29.9
$D(\mathbf{I})$ in [48]	79.6	63.1	53.2	36.3
Gamma curve	80.1	64.0	53.6	36.6

darkening function in the light-agnostic learning pipeline.

Experiments and Results



Comparing Low-light and Light-agnostic vision samples and results

Dataset	Task	#Cls	#Train	#Val
Low-light (LL)				
ExDark [47]	Object detection	12	5891	1472
DARK FACE [80]	Face detection	1	5400	600
ACDC [60]	Semantic segmentation	19	400	106
LIS [10]	Instance segmentation	8	1561	669
DarkVision [81]	Video object detection	4	26*	6*
Light-Agnostic (LA)				
COCO [21] ∩ ExDark [47]	Object detection	12	92357	5156
WIDER FACE [78] ∩ DARK FACE [80]	Face detection	1	18280	3822
CityScapes [15] ∩ ACDC [60]	Semantic segmentation	19	3375	606
COCO [40] ∩ LIS [10]	Instance segmentation	8	44321	2550
ImageNet VID [58] ∩ DarkVision [81]	Video object detection	2	401*	51*

Datasets Statistics. *: video samples. Light-Agnostic datasets are obtained through Eq. 6

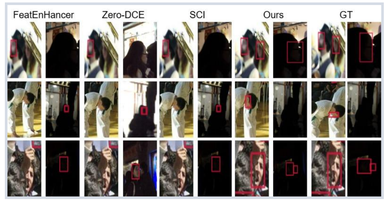
How Light-Agnostic datasets are generated.

$$\mathcal{D}_{\text{LA}} = \{(x, y) \,|\, (x, y) \in \mathcal{D}_{\text{WL}} \cup \mathcal{D}_{\text{LL}},\ y \in \mathcal{C}_{\text{com}}\}, \quad (6)$$

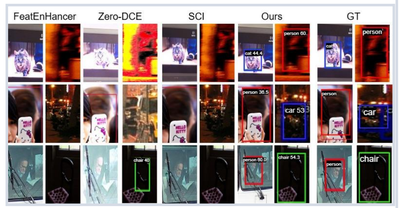
Tasks →	Object Det.				Face Det.				Instance Seg.				Semantic Seg.		Video Object Det.	
Light Settings →	LL		LA		LL		LA		LL		LA		LL	LA	LL	LA
Methods	mAP_{50}	mAP	mAP_{50}	mAP	AP_{50}	AP	AP_{50}	AP	mAP	SegAP	mAP	SegAP	mIoU	mIoU	mAP	mAP
Baseline	78.8	44.2	62.4	33.1	47.3	19.9	52.6	28.0	54.0	46.0	41.5	33.9	49.6	75.5	32.8	51.3
Enhancement																
Zero-DCE [24]	76.2	42.7	57.6	29.5	47.4	20.1	47.6	22.9	51.5	43.5	39.2	32.1	42.5	59.4	7.83	28.3
SCI [36]	75.5	42.2	57.9	29.7	45.1	19.3	45.4	21.7	52.3	44.2	35.1	28.3	42.1	57.4	6.71	25.7
URetinexNet [71]	74.6	42.0	57.3	29.2	44.7	18.3	46.6	22.3	59.5	43.1	37.4	29.7	41.2	58.8	6.68	25.1
IAT [17]	74.1	41.0	52.4	30.5	42.1	15.9	37.1	17.4	49.7	42.3	36.4	30.1	41.7	56.2	6.24	24.5
Task-specific																
Zero-DCE [24]‡	79.1	44.4	62.9	33.2	50.4	21.1	50.1	23.5	59.1	52.7	44.4	35.7	44.7	61.0	34.1	51.5
SCI [50]‡	79.2	44.2	63.0	33.4	51.6	22.2	54.7	29.9	58.4	51.0	43.7	34.9	43.5	59.9	32.9	50.7
MAET [16]‡	79.2	44.3	57.3	27.3	44.3	18.7	47.9	23.1	-	-	-	-	-	-	-	-
Xue *et al.* [76]	-	-	-	-	46.1	17.7	44.2	21.0	-	-	-	-	42.1	69.7	-	-
FeatEnHancer [26]‡	79.2	44.8	62.9	33.3	47.2	19.9	51.6	24.5	59.6	52.4	44.9	36.1	50.1	76.1	34.6	52.7
TorchAdapt	**80.1**	**45.4**	**64.0**	**34.1**	**53.7**	**23.1**	**59.0**	**32.5**	**61.1**	**53.6**	**45.2**	**36.6**	**50.8**	**76.3**	**39.1**	**59.1**
vs. prev. SoTA	+0.9	+0.6	+1.0	+0.7	+2.1	+0.9	+4.3	+2.6	+1.5	+0.9	+0.3	+0.5	+0.7	+0.2	+4.5	+6.4

Quantitative comparison across five tasks under both Low-Light (LL) and Light-Agnostic (LA) settings.

Qualitative Results



Light-agnostic and low-light face detections on the WIDER FACE and DARK FACE datasets.



Light-agnostic and low-light object detections on the COCO and ExDark datasets.