Towards Real-world Event-guided Low-light Video Enhancement and Deblurring

Taewoo Kim, Jaeseok Jeong, Hoonhee Cho, Yuhwan Jeong, and Kuk-Jin Yoon

Korea Advanced Institute of Science and Technology

ECCV 2024

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

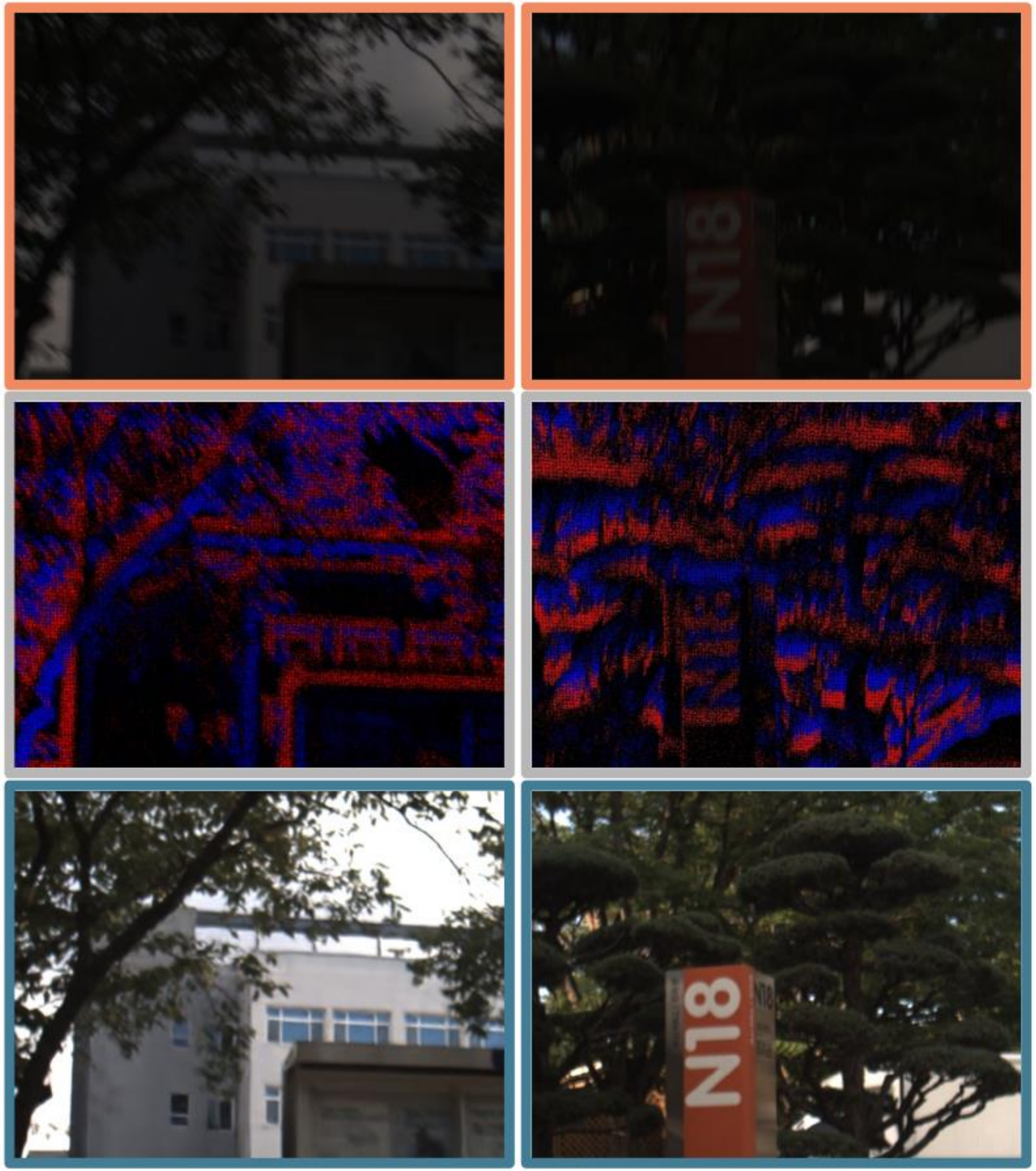
Introduction

- Challenges
  - In low-light conditions, capturing videos with frame-based cameras often requires long exposure times, resulting in motion blur and reduced visibility.
  - While frame-based motion deblurring and low-light enhancement have been studied, they still pose significant challenges.

- Method
  - First establish real-world datasets for event-guided video-based low-light enhancement and deblurring using a hybrid camera system based on beam splitters.
  - Introduce an end-to-end framework to effectively handle these tasks.

Introduction

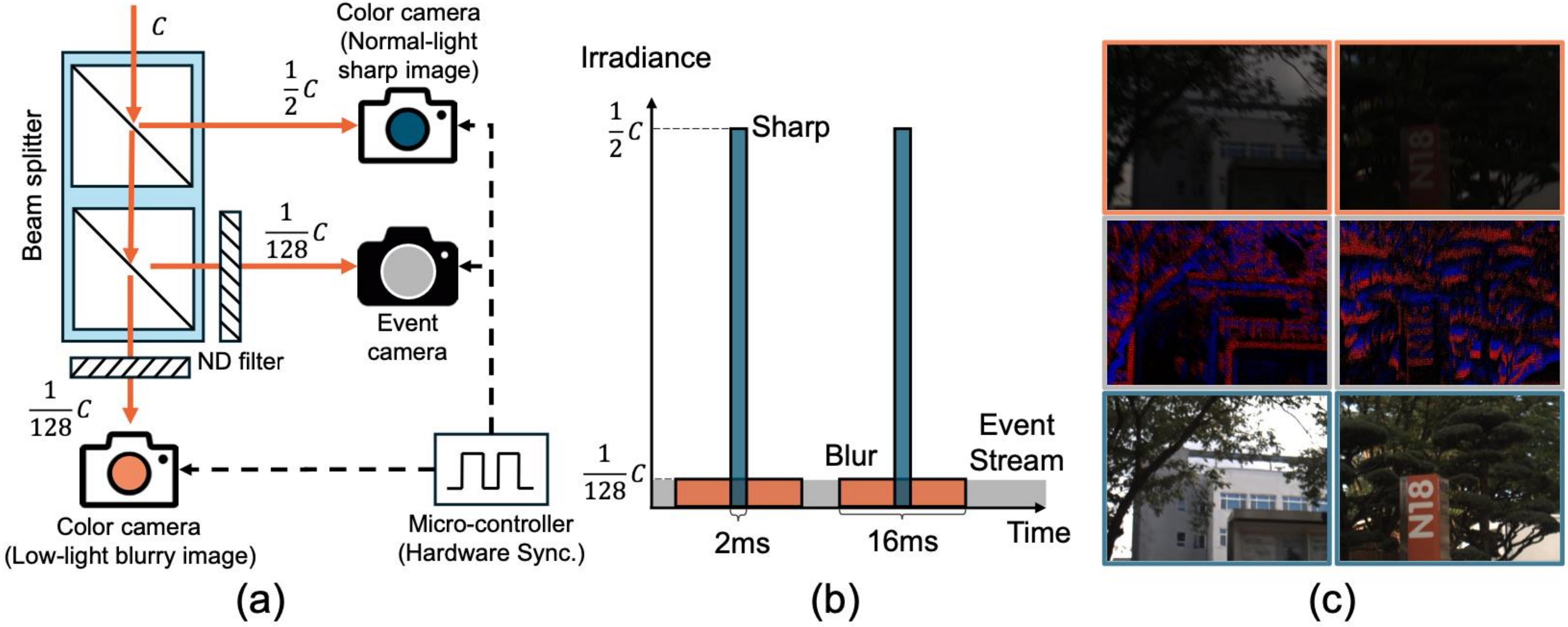
- Why use event camera?
  - **high temporal resolution**
  - **high dynamic range**
  - **low latency, and low power consumption**



Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

Framework



C
Beam splitter
$\frac{1}{2}C$
Color camera
(Normal-light
sharp image)
$\frac{1}{128}C$
Event
camera
ND filter
$\frac{1}{128}C$
Color camera
(Low-light blurry image)
Micro-controller
(Hardware Sync.)
(a)
Irradiance
$\frac{1}{2}C$
Sharp
Blur
Event
Stream
$\frac{1}{128}C$
2ms
16ms
Time
(b)
N18
(c)

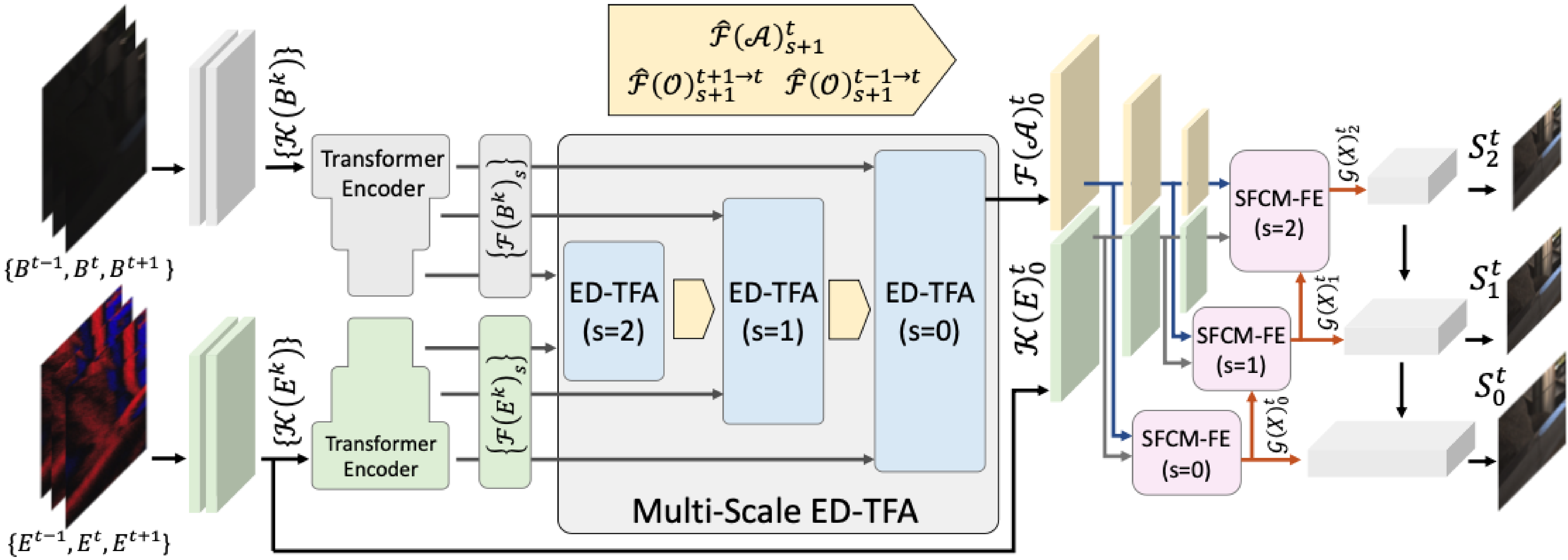
Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

Related work

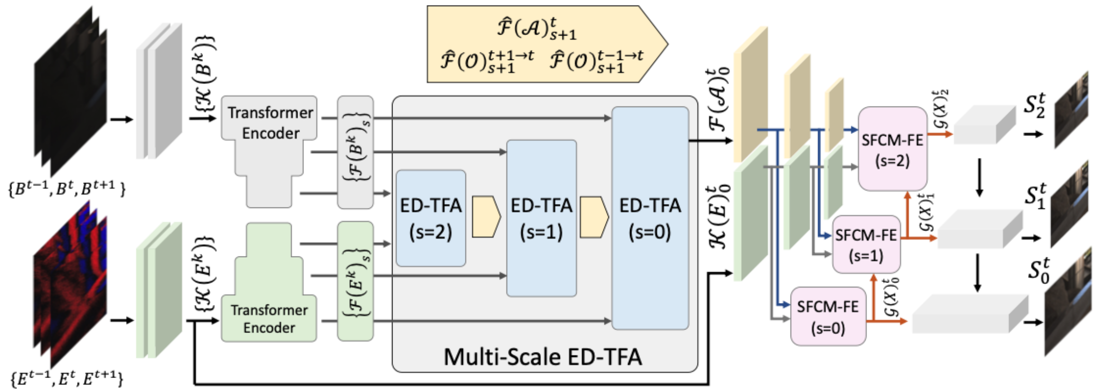
Category	Methods	PSNR	SSIM	Params(MB)
Image-based LLE	SNRNet [47]	26.47	0.851	40.08
	SDSDNet [38]	28.47	0.887	4.43
	LLFormer [40]	26.62	0.862	13.15
	RetinexFormer [1]	26.66	0.865	1.61
Image-based beblur	MPRNet [51]	26.89	0.867	20.13
	MIMOUNet+ [7]	26.52	0.866	16.11
	NAFNet [3]	26.77	0.862	67.91
	RNN-MBP [2]	29.52	0.902	14.16
	DSTNet [25]	29.59	0.903	7.53
Video-based LLE+blur	LEDNet [58]	26.47	0.856	7.41
Event-guided Image-based beblur	e-SLNet† [37]	19.45	0.663	0.17
	REDNet† [46]	29.19	0.903	9.7
	EFNet† [32]	29.85	0.905	8.47
Event-guided Video-based beblur	UEVD† [13]	29.93	0.905	27.88
	GEM† [53]	26.04	0.810	2.36
	REFID† [33]	30.10	0.913	15.9
	Ours-s†	<u>30.98</u>	<u>0.919</u>	5.3
	Ours†	**31.30**	**0.925**	12.8

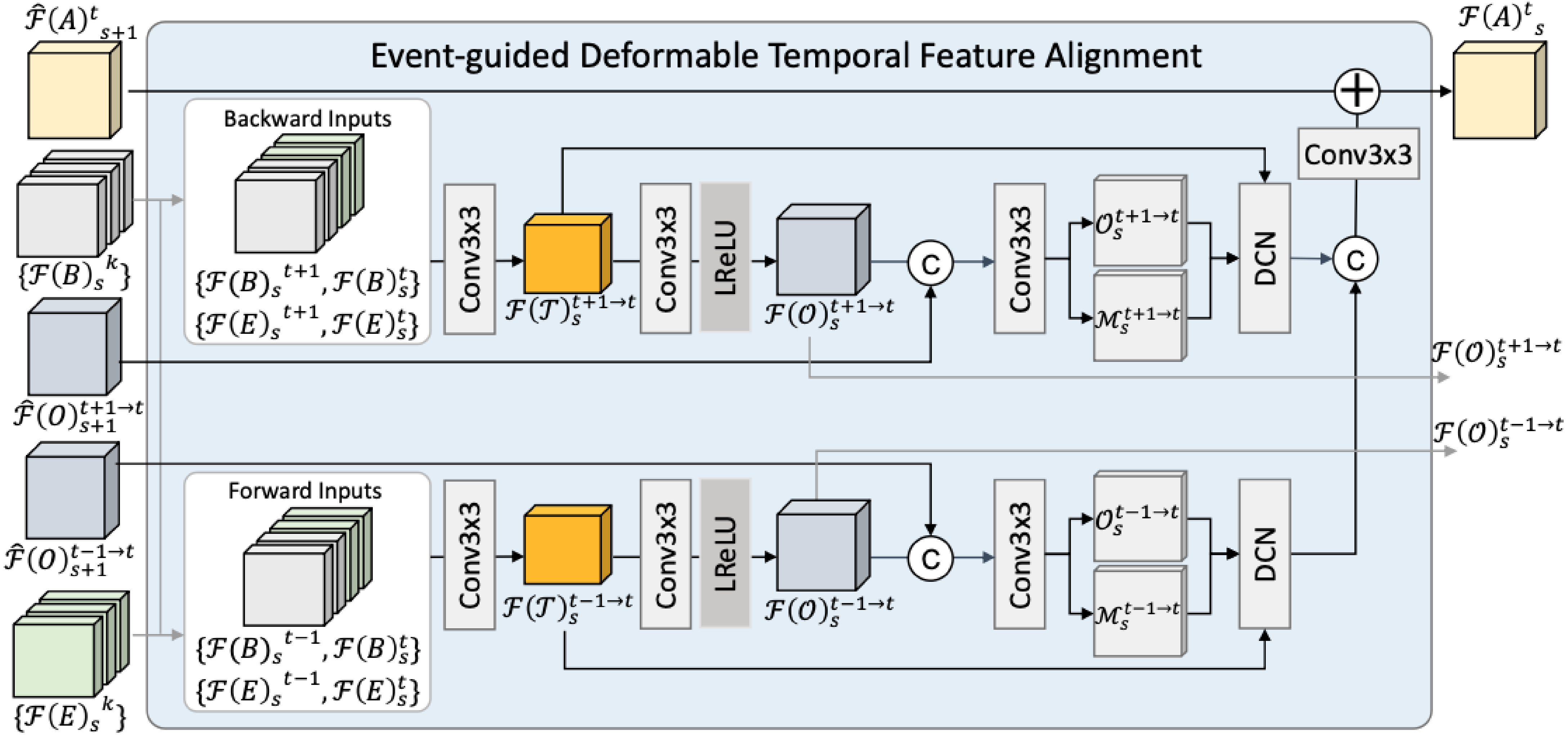
Model



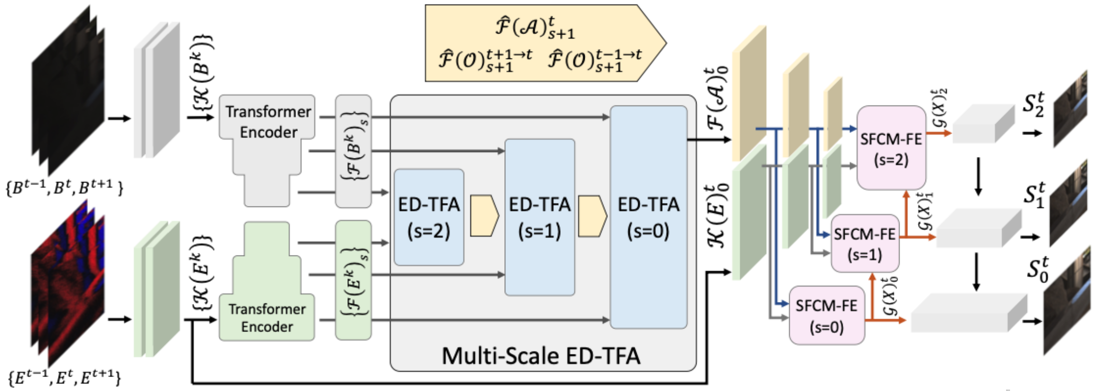
$\{B^{t-1}, B^t, B^{t+1}\}$
$\{E^{t-1}, E^t, E^{t+1}\}$
$\{\mathcal{K}(B^k)\}$
$\{\mathcal{K}(E^k)\}$
Transformer Encoder
Transformer Encoder
$\{\mathcal{F}(B^k)_s\}$
$\{\mathcal{F}(E^k)_s\}$
$\hat{\mathcal{F}}(\mathcal{A})^t_{s+1}$
$\hat{\mathcal{F}}(\mathcal{O})^{t+1\to t}_{s+1}$ $\hat{\mathcal{F}}(\mathcal{O})^{t-1\to t}_{s+1}$
ED-TFA (s=2)
ED-TFA (s=1)
ED-TFA (s=0)
Multi-Scale ED-TFA
$\mathcal{F}(\mathcal{A})^t_0$
$\mathcal{K}(E)^t_0$
SFCM-FE (s=2)
SFCM-FE (s=1)
SFCM-FE (s=0)
$\mathcal{G}(X)^t_2$
$\mathcal{G}(X)^t_1$
$\mathcal{G}(X)^t_0$
S^t_2
S^t_1
S^t_0

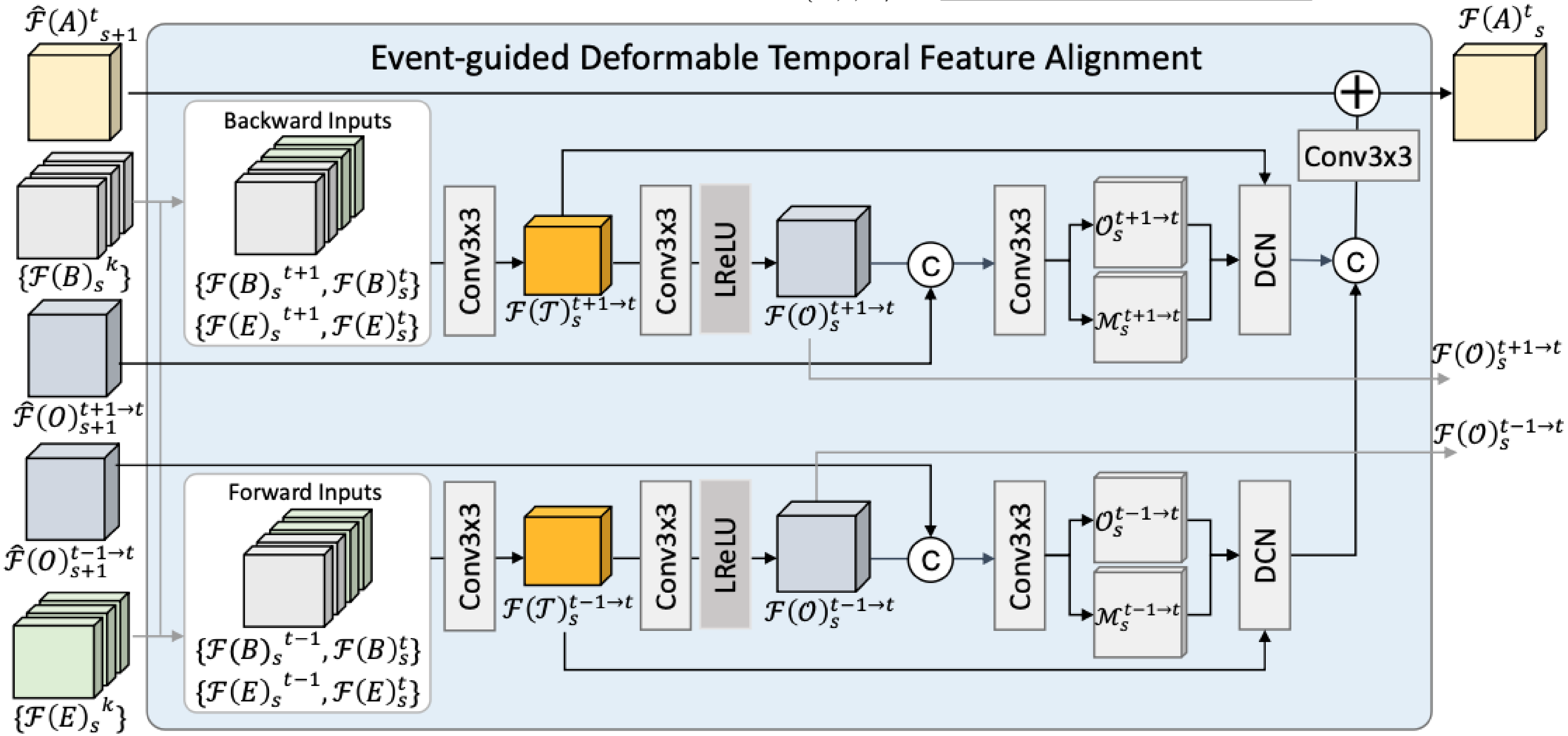
Event-guided Deformable Temporal Feature Alignment (ED-TFA)





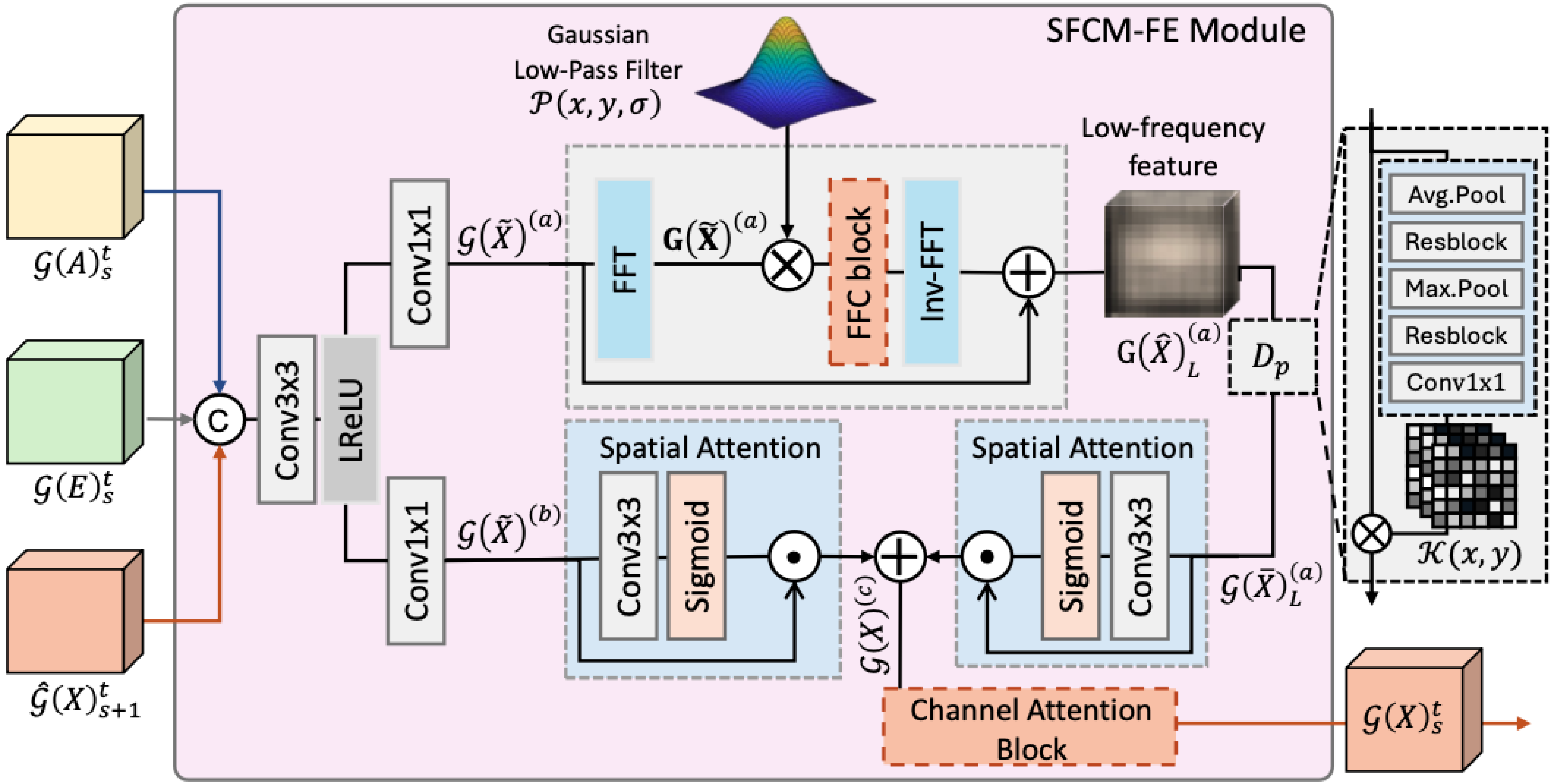
Event-guided Deformable Temporal Feature Alignment (ED-TFA)





Spectral Filtering-based Cross-Modal Feature Enhancement (SFCM-FE)

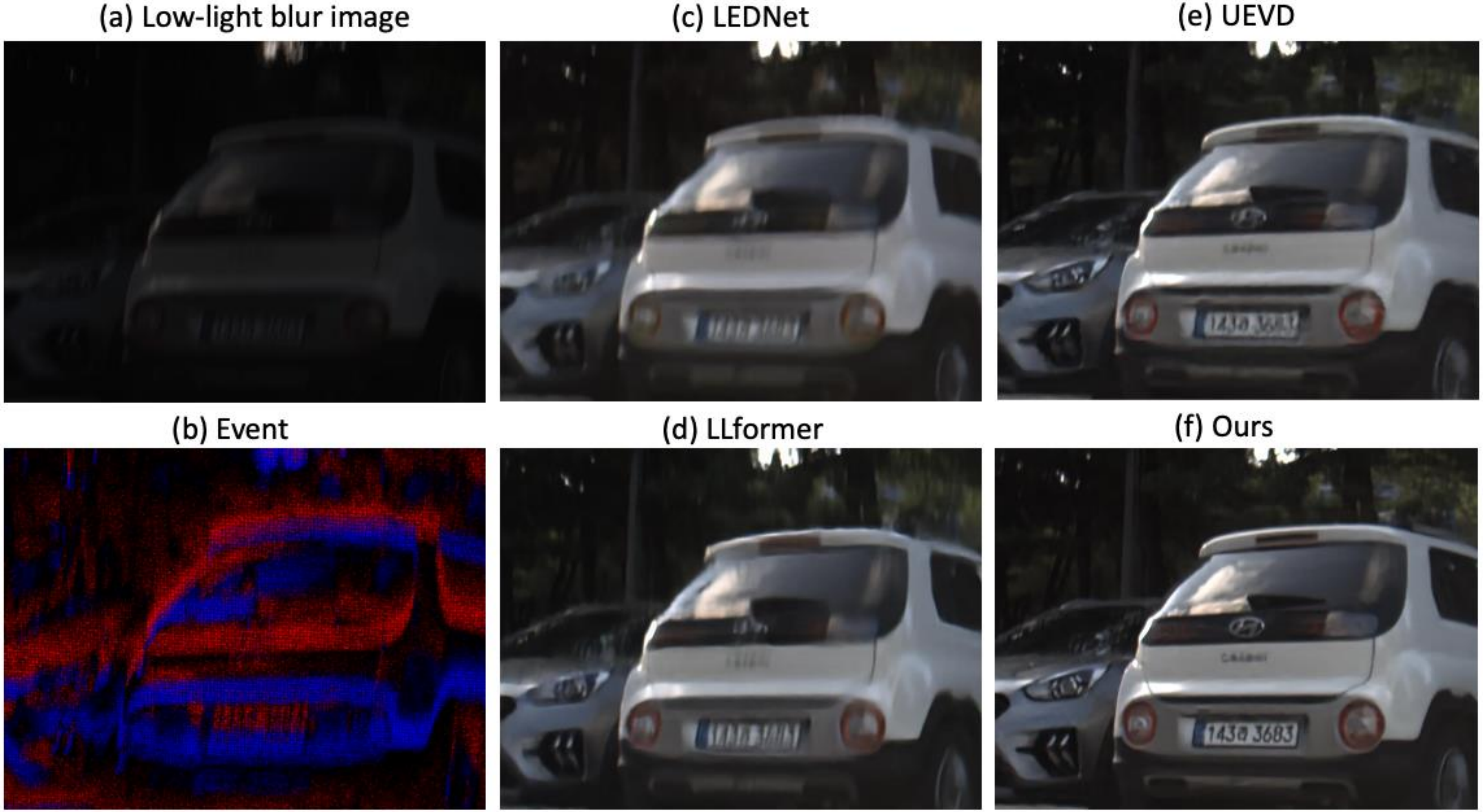
$$\mathcal{P}(x,y,\sigma) = \exp(-\frac{(x-x_c)^2+(y-y_c)^2}{2\sigma^2})$$



Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

Results


(a) Low-light blur image
(c) LEDNet
(e) UEVD
(b) Event
(d) LLformer
(f) Ours

Results

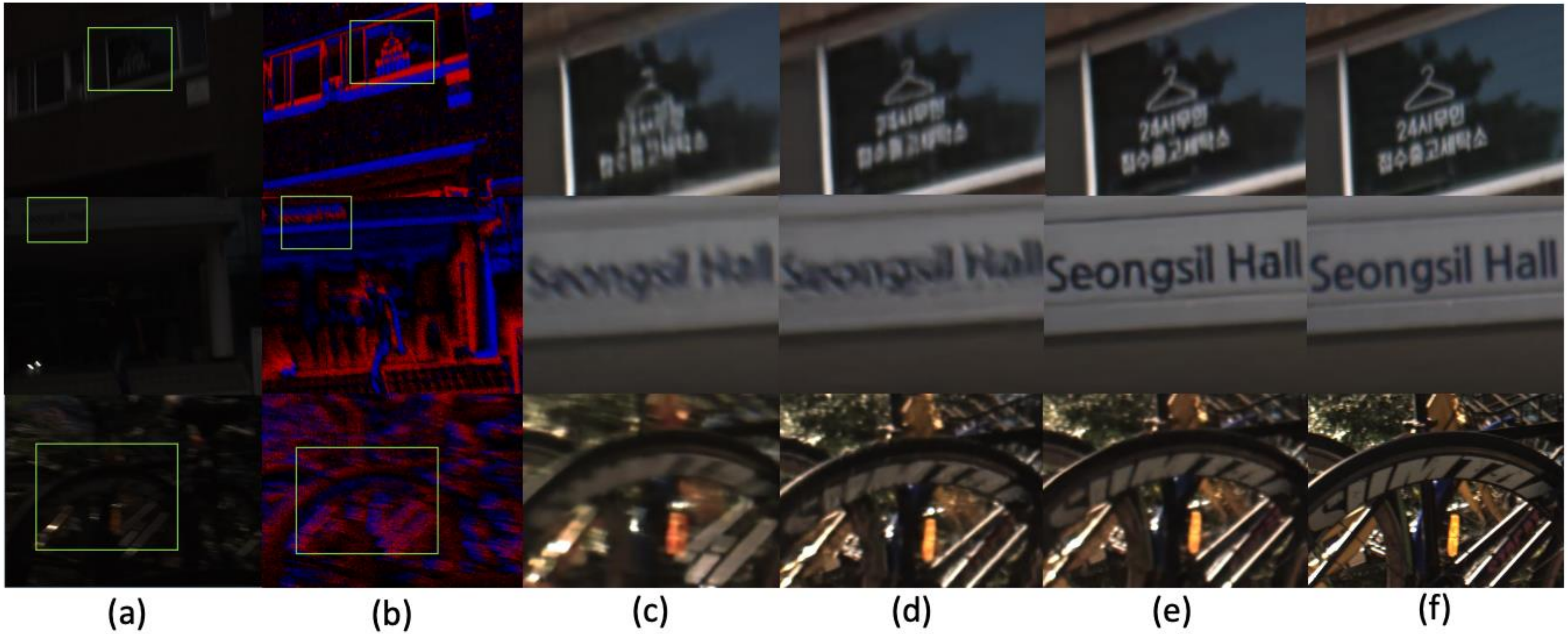


Fig. 6: Visual comparisons on the RELED datasets. In the figures, (a) to (f) depict the following: (a) input low-light blurry images, (b) low-light events, (c) LLFormer [40], (d) EFNet [32], (e) Ours, and (f) GT normal-light sharp image.

Results

Category	Methods	PSNR	SSIM	Params(MB)
Image-based LLE	SNRNet [47]	26.47	0.851	40.08
	SDSDNet [38]	28.47	0.887	4.43
	LLFormer [40]	26.62	0.862	13.15
	RetinexFormer [1]	26.66	0.865	1.61
Image-based beblur	MPRNet [51]	26.89	0.867	20.13
	MIMOUNet+ [7]	26.52	0.866	16.11
	NAFNet [3]	26.77	0.862	67.91
	RNN-MBP [2]	29.52	0.902	14.16
	DSTNet [25]	29.59	0.903	7.53
Video-based LLE+blur	LEDNet [58]	26.47	0.856	7.41
Event-guided Image-based beblur	e-SLNet† [37]	19.45	0.663	0.17
	REDNet† [46]	29.19	0.903	9.7
	EFNet† [32]	29.85	0.905	8.47
Event-guided Video-based beblur	UEVD† [13]	29.93	0.905	27.88
	GEM† [53]	26.04	0.810	2.36
	REFID† [33]	30.10	0.913	15.9
	Ours-s†	<u>30.98</u>	<u>0.919</u>	5.3
	Ours†	**31.30**	**0.925**	12.8

Datasets	Image resolution	color	Real event(resolution)	Low-light type	Blur type
LoL-Blur [58]	1120×640	Yes	Not provide	Synthetic low-light	Synthetic blur
RELED(Ours)	**1024×768**	**Yes**	**Yes(1024×768)**	**Real low-light**	**Real blur**

Outline

- Introduction
- Framework
- Method
- Experiment
- Conclusion

Conclusion

- Addresses the novel research problem of event-guided low-light video enhancement and deblurring
- Designed a hybrid camera sys- tem using beam splitters and constructed the RELED dataset containing low- light blurry images, normal sharp images, and event streams
- Developed a tailored framework for the task and validated its effectiveness and finally achieved significant performance improvement on the proposed dataset, surpassing both event-guided and frame-based methods.