BALF

Simple and Efficient Blur Aware Local Feature Detector

WACV 2024

Zhenjun Zhao

香港中文大學
The Chinese University of Hong Kong

Local Feature Detector

Sharp vs Sharp

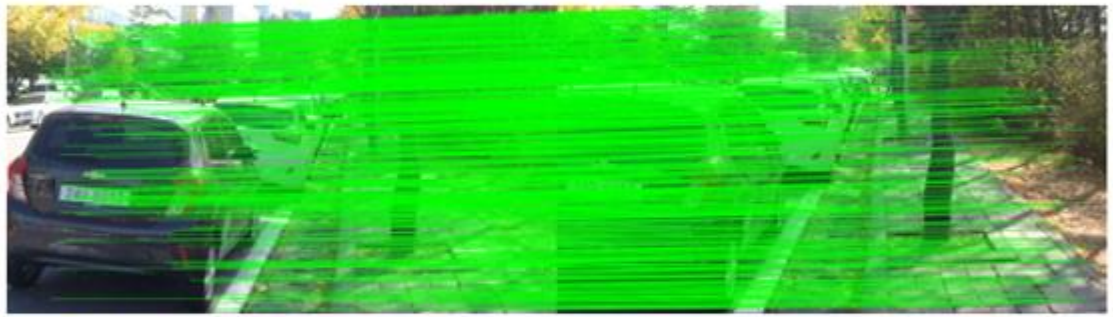
Blur vs Sharp

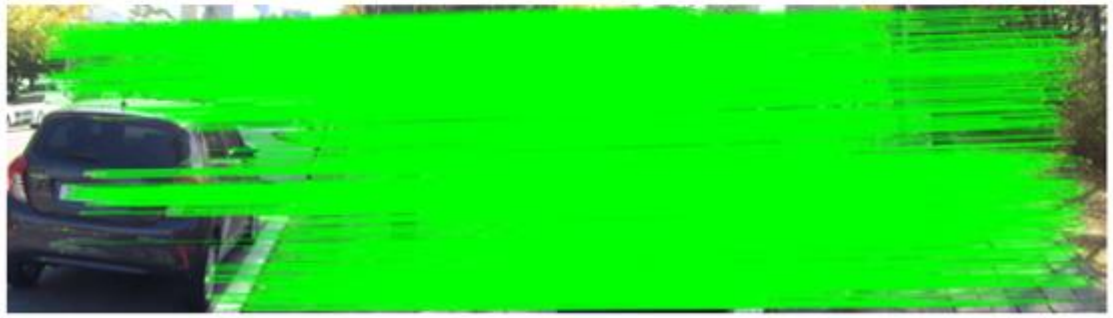
SIFT (IJCV'04)

SuperPoint (CVPRW'18)

R2D2 (NIPS'19)








BALF

First pure MLP-based network for local feature detection

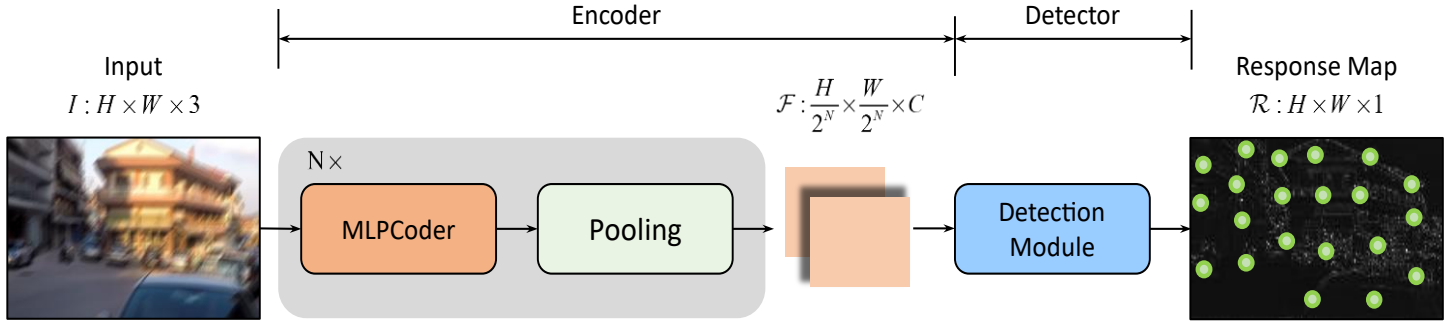
Not only for blurred images, but also for sharp images

Be able to run in real-time (~34.46 FPS) with a VGA resolution image

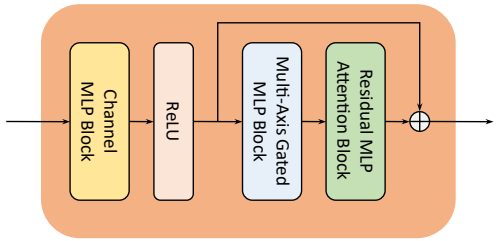
Method

BALF Framework:

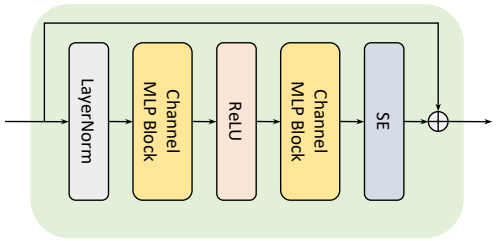
- **Pure MLP-based** network
- **MLP-based encoder** and **MLP-based detection module**



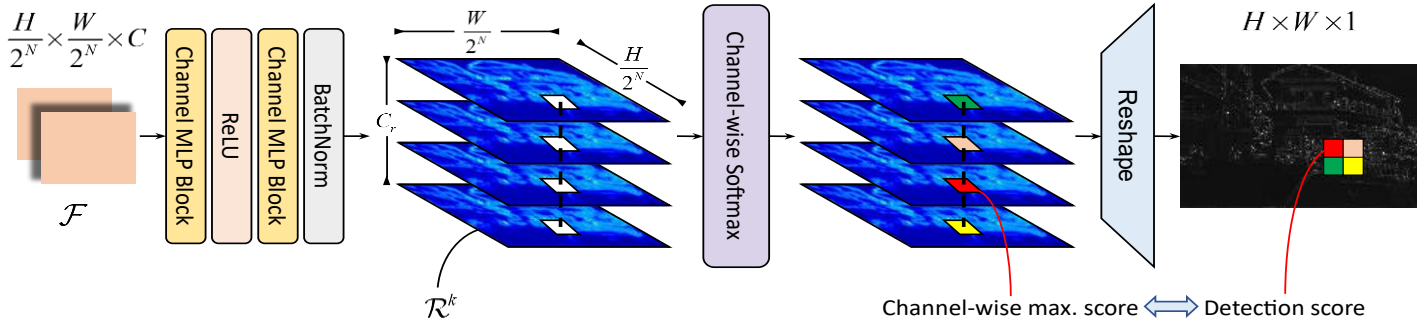
MLPCoder Block:



Residual MLP Attention Block (RMAB):



MLP-Based Detection Module:



Results

Evaluation on Sharp/Blur-HPatches datasets

Method	Reference: Sharp Target: Sharp		
	Viewpoint ↑	Illumination ↑	Total ↑
SIFT [34]	60.29	60.44	60.36
SURF [6]	62.67	64.01	63.33
Harris-Laplace [38]	63.89	62.91	63.41
Shi-Tomasi [56]	69.28	64.13	66.74
MSER [36]	52.45	50.58	51.53
KAZE [2]	67.30	65.67	66.50
AKAZE [3]	66.08	69.07	67.55
FAST [47]	66.08	63.65	64.88
LIFT [74]	56.97	60.73	58.82
Key.Net [25]	68.99	67.47	68.24
SuperPoint [12]	**69.53**	68.92	69.23
LF-Net [41]	68.41	**73.61**	**70.96**
D2-Net [14]	53.99	62.80	58.32
R2D2 [44]	61.68	61.93	61.80
BALF (ours)	67.21	73.51	70.28

Method	Reference: Sharp Target: Blur			Reference: Blur Target: Blur		
	EASY ↑	HARD ↑	TOUGH ↑	EASY ↑	HARD ↑	TOUGH ↑
SIFT [34]	55.92	56.80	53.49	56.99	53.49	45.94
SURF [6]	58.88	56.23	56.24	61.08	58.04	53.60
Harris-Laplace [38]	36.70	37.97	34.98	35.76	31.95	27.47
Shi-Tomasi [56]	57.33	55.11	49.11	56.29	53.75	51.37
MSER [36]	44.19	41.97	37.05	41.81	38.24	34.59
KAZE [2]	49.90	46.84	39.98	63.29	58.71	46.90
AKAZE [3]	54.15	50.51	45.49	65.16	62.20	51.54
FAST [47]	61.98	61.77	51.37	57.84	53.35	51.17
LIFT [74]	50.69	50.17	46.99	48.34	46.57	46.53
Key.Net [25]	60.34	54.71	44.69	62.77	58.17	49.25
SuperPoint [12]	65.64	62.22	52.84	58.60	50.03	43.28
LF-Net [41]	63.54	61.19	56.78	60.45	59.07	57.71
D2-Net [14]	49.71	47.30	44.32	51.80	51.05	50.53
R2D2 [44]	57.99	51.73	40.57	57.49	55.31	46.86
BALF (ours)	**74.12**	**74.45**	**71.84**	**70.48**	**68.43**	**67.71**

Our method achieves superior detection performance over prior works on motion blurred images, while keeping comparable performance for sharp images

Evaluation on Deblurred Images Using Comparing Methods

Method	Reference: Sharp, Target: Deblur						Reference: Deblur, Target: Deblur					
	SRN-DeblurNet [60]			DeblurGAN-v2 [24]			SRN-DeblurNet [60]			DeblurGAN-v2 [24]		
	EASY ↑	HARD ↑	TOUGH ↑	EASY ↑	HARD ↑	TOUGH ↑	EASY ↑	HARD ↑	TOUGH ↑	EASY ↑	HARD ↑	TOUGH ↑
SIFT [34]	56.62	55.36	53.83	57.63	56.52	56.50	59.75	58.13	50.63	59.44	57.98	51.21
SURF [6]	61.89	59.13	54.88	61.97	59.57	56.34	62.44	61.26	55.27	62.07	60.81	55.09
Harris-Laplace [38]	17.15	16.87	20.54	16.60	16.90	20.24	36.98	35.73	32.23	37.09	35.97	31.54
Shi-Tomasi [56]	60.56	56.87	48.78	61.75	59.10	51.56	63.18	61.03	50.88	63.58	61.89	53.76
MSER [36]	46.65	43.23	37.90	47.62	45.14	40.70	47.70	45.40	37.49	47.84	45.56	38.01
KAZE [2]	65.14	63.10	60.16	65.23	63.18	61.41	64.20	62.45	53.41	64.13	61.87	54.19
AKAZE [3]	66.03	64.02	60.64	66.29	64.50	62.72	65.71	64.08	56.10	65.75	63.75	57.35
FAST [47]	61.77	59.67	61.60	62.00	60.44	58.74	62.72	61.14	50.61	63.40	61.70	55.43
LIFT [74]	54.98	52.64	46.75	56.59	53.54	49.09	55.88	53.64	45.31	56.68	55.31	50.44
Key.Net [25]	63.28	58.01	47.10	63.99	59.16	49.35	62.86	60.44	50.74	62.73	60.58	52.96
SuperPoint [12]	67.72	64.05	55.26	67.95	65.86	58.22	66.38	63.16	49.52	66.50	63.71	52.09
LF-Net [41]	62.22	59.90	54.73	62.59	60.24	54.81	63.06	62.03	57.28	63.00	61.79	57.85
D2-Net [14]	51.81	49.49	45.94	52.64	50.21	45.88	53.60	53.00	50.93	53.93	53.29	50.74
R2D2 [44]	60.31	55.43	43.26	60.46	55.68	45.38	58.11	54.80	45.77	57.95	55.03	47.86
BALF (ours)	**74.12 / 74.45 / 71.84** (EASY / HARD / TOUGH)						**70.48 / 68.43 / 67.71** (EASY / HARD / TOUGH)					

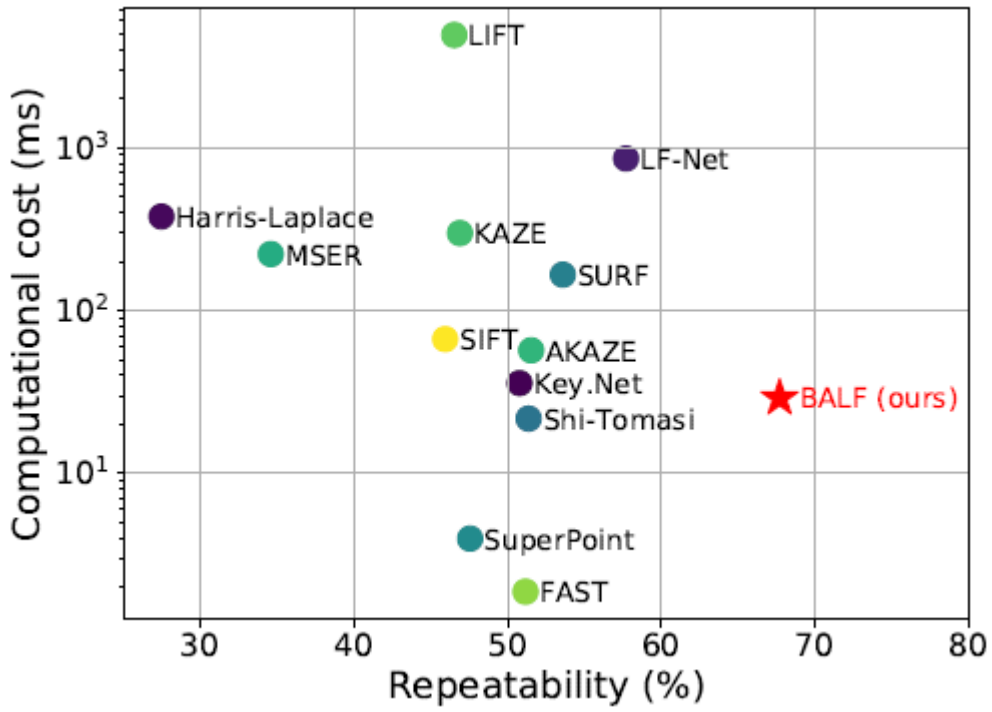
To design keypoint detector from blurred image directly would be a better option compared to that of detecting keypoints from the intermediate deblurred image.

Efficiency Results

Computational cost (ms)

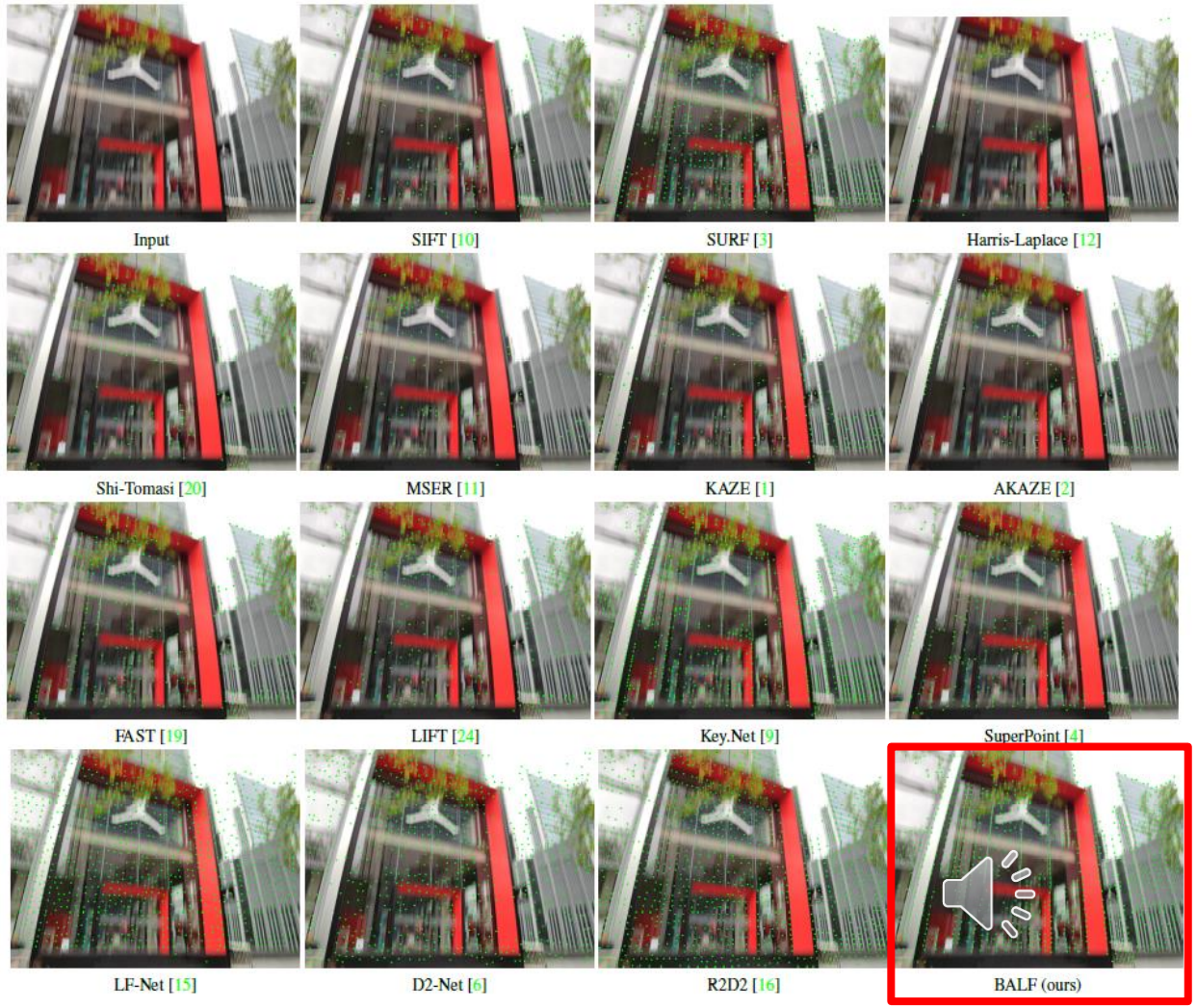
Method	240×320 pixels ↓	480×640 pixels ↓
SIFT [34]	21.80	66.70
SURF [6]	148.46	165.78
Harris-LapLace [38]	110.41	377.13
Shi-Tomasi [56]	5.20	21.69
MSER [36]	64.19	221.79
KAZE [2]	105.85	298.43
AKZE [3]	18.02	56.93
FAST [47]	0.89	1.88
LIFT [74]	2209.03	4901.38
Key.Net [25]	15.64	35.82
SuperPoint [12]	2.41	3.98
LF-Net [41]	282.77	855.77
BALF (ours)	8.15	29.02

Efficiency-Effectiveness



Our motion blur aware detector is able to run in real-time (~34.46 FPS) with a VGA resolution image (480×640 pixels).

Qualitative Detection Results



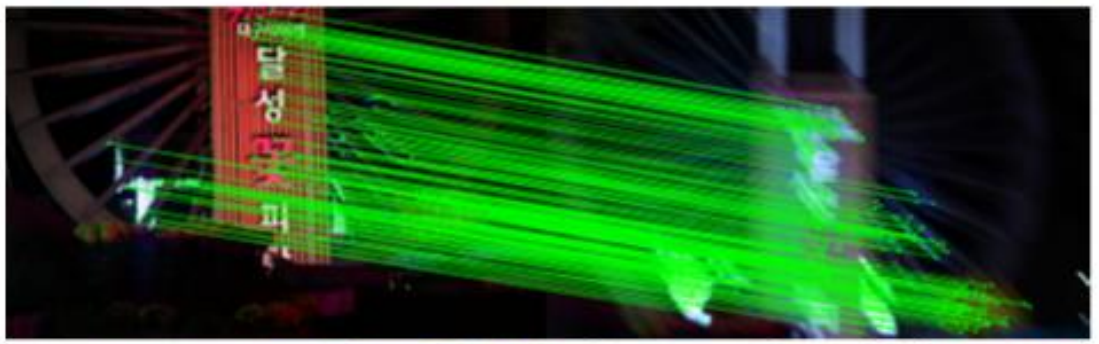
Qualitative Matching Results

SuperPoint (CVPRW'18)

BALF (ours)

Indoor

Outdoor



BALF

Code, data, and more results at
ericzzj1989.github.io/balf