DATASET COMPARISON

DATASETS:

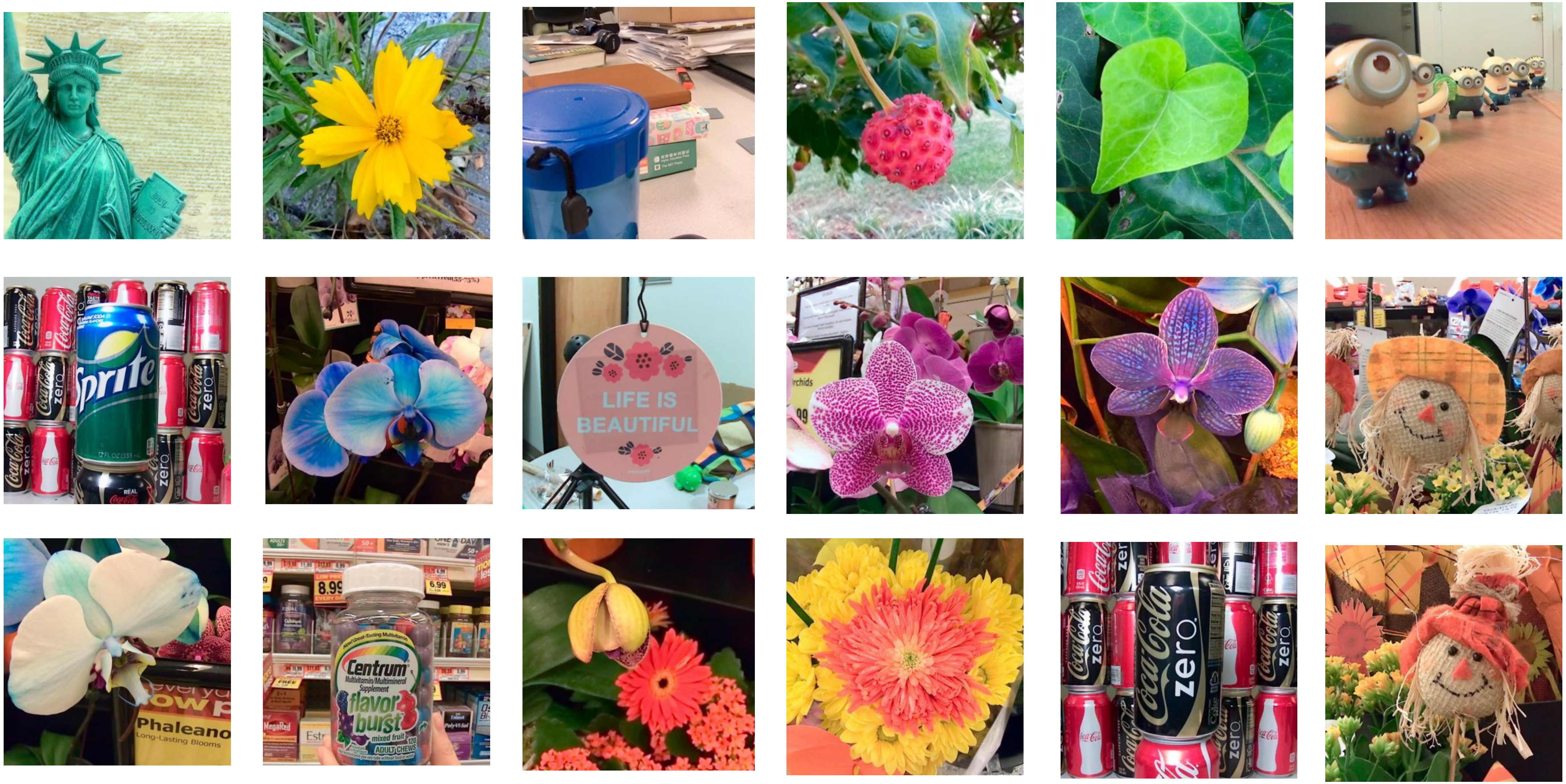
- **LFSD** (2014)
  available
- **HFUT-Lytro** (2017)
  NOT available
- **DUTLF** (DUTLF-FS: 2019, DUTLF-V2: 2022)
  available
- **Lytro-Illum** (2020)
  available
- **PKU-LF** (2023)
  NOT available

TABLE I
SUMMARY OF FOUR EXISTING DATASETS AND THE PROPOSED PKU-LF DATASET

Dataset	Year	#Scale	Resulution	Light Field								Annotation					Device
				Cal.	Raw.	Foc.	Dep.	Sub.	Mic.	Rel.	Spl.	Scr.	Bou.	Obj.	Ins.	Edg.	
LFSD [18]	2014	100	360 × 360		✓	✓	✓							✓			Lytro
HFUT [25]	2017	255	328 × 328			✓	✓	✓						✓			Lytro
DUT-LF [20]	2019	1,462	600 × 400			✓	✓				✓			✓			Lytro Illum
Lytro Illum [22]	2019	640	400 × 590		✓				✓					✓			Lytro Illum
PKU-LF	2021	5,000	2022 × 1404	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Lytro Illum

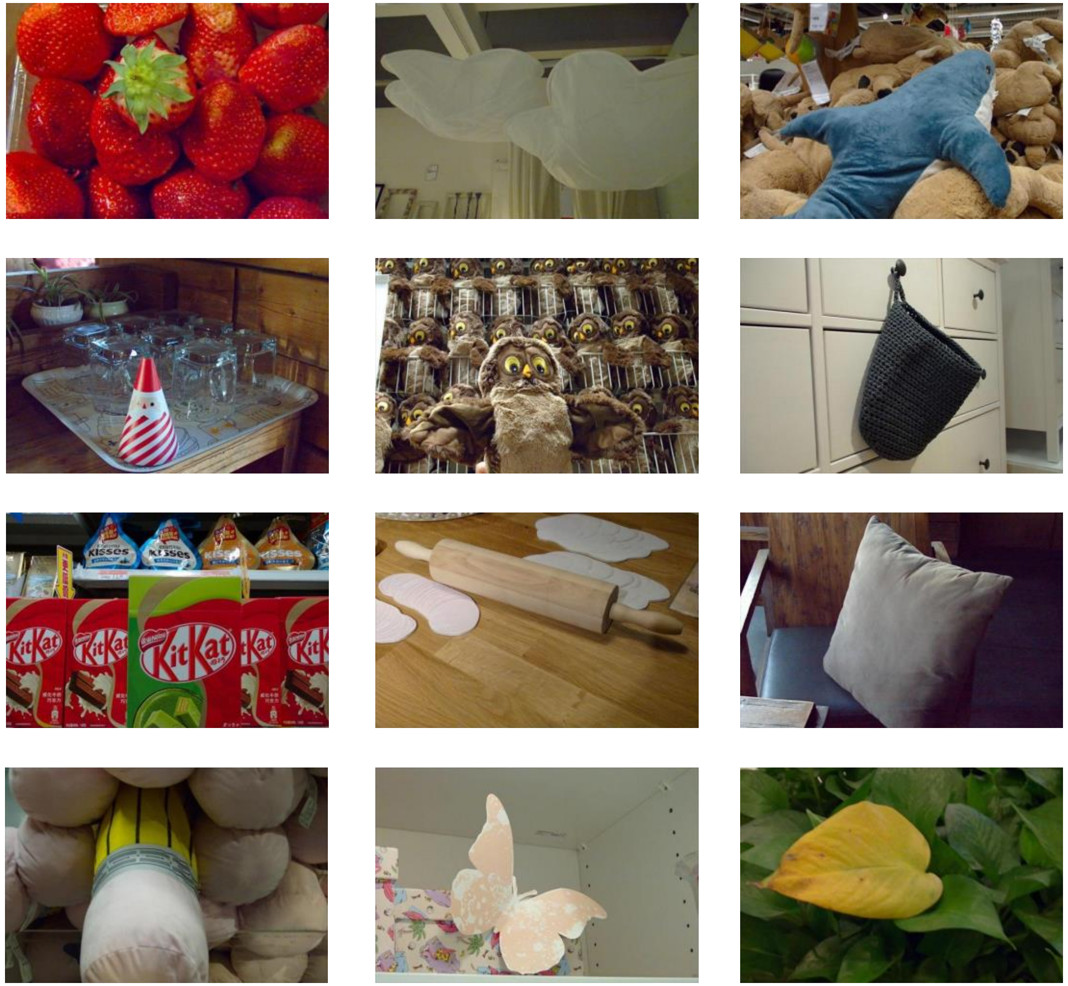
Cal.: Calibration data of the camera. Raw.: Raw light fields. Foc.: Focal stacks. Dep.: Depth maps. Sub.: Sub-aperture images. Mic.: Micro-lens images. Rel.: Relative-depth-of-field coordinates. Spl.: Official training/testing set split. Scr.: Scribble annotations. Bou.: Bounding boxes. Obj.: Object-level annot tions. Ins.: Instance-level annotations. Edg.: Edge annotations.

LFSD **ALL-IN-FOCUS** (+ GT, focal stack, results)

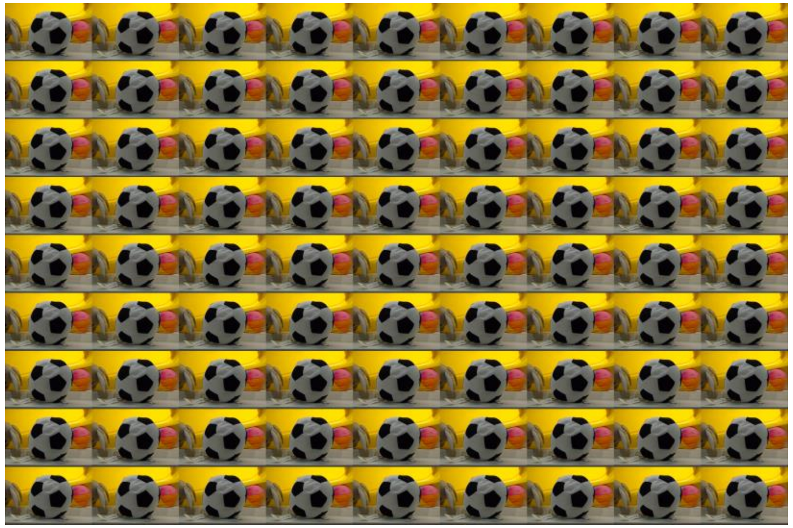


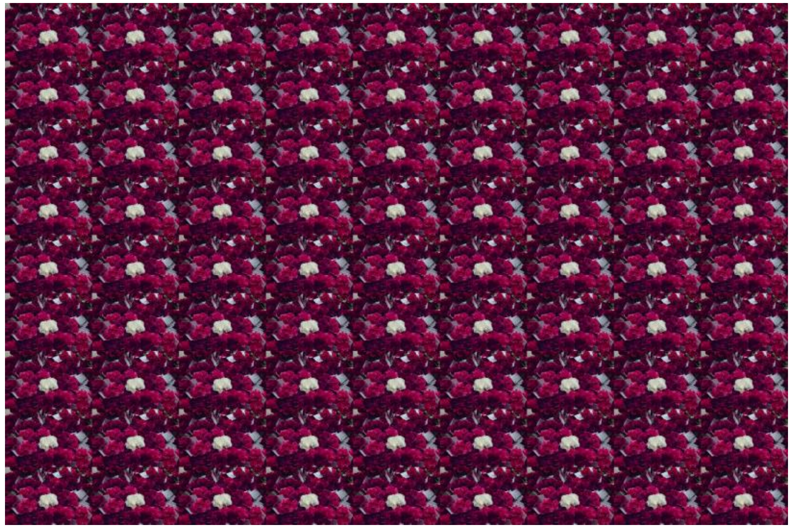
DUTLF-Fs **ALL-IN-FOCUS**

(+ GT, focal stack, results)



DUTLF-V2 only **FOCAL STACK**





2014 LFS

- **PRC**
- → Precision (% saliency pixels correctly assigned)
- → Recall (detected salient region wrt GT)

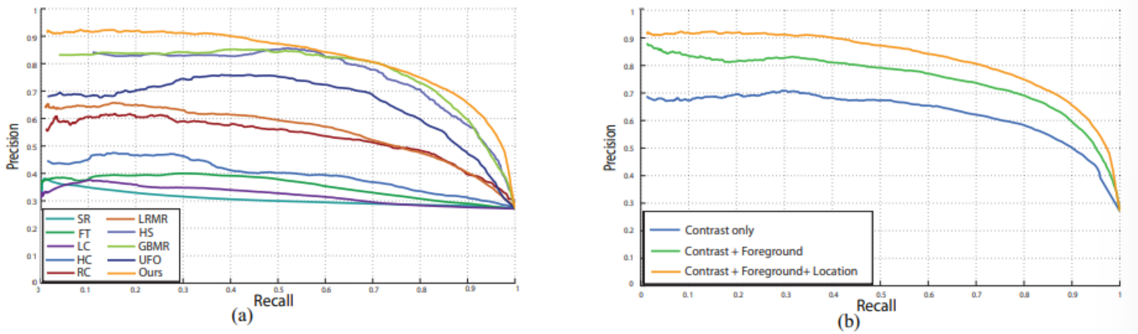


Figure 6. (a) PRC comparisons on our light field dataset; (b) PRC comparisons using different cues in our approach.

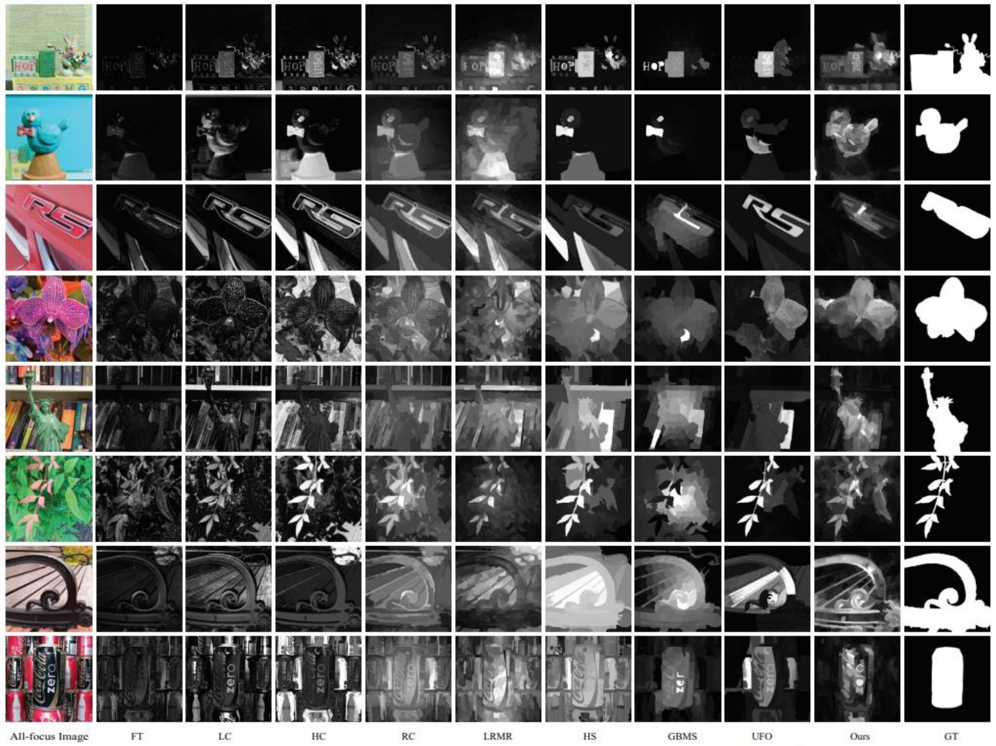
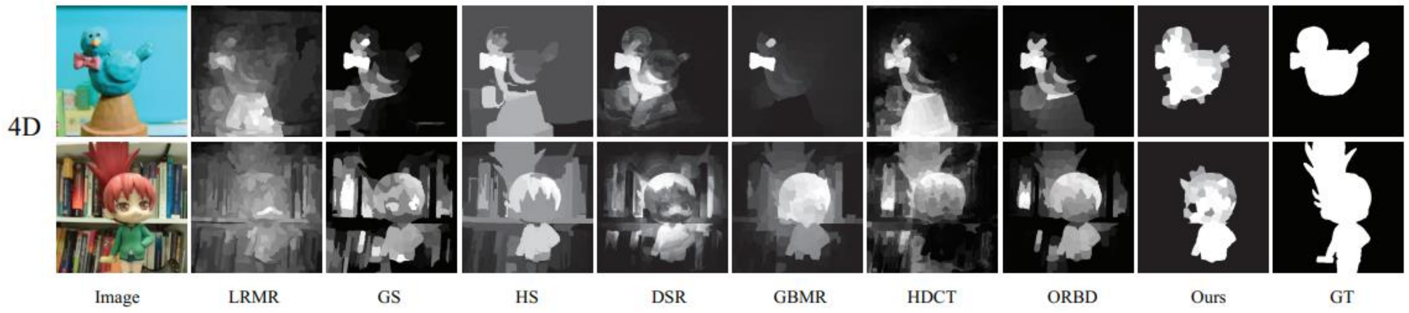
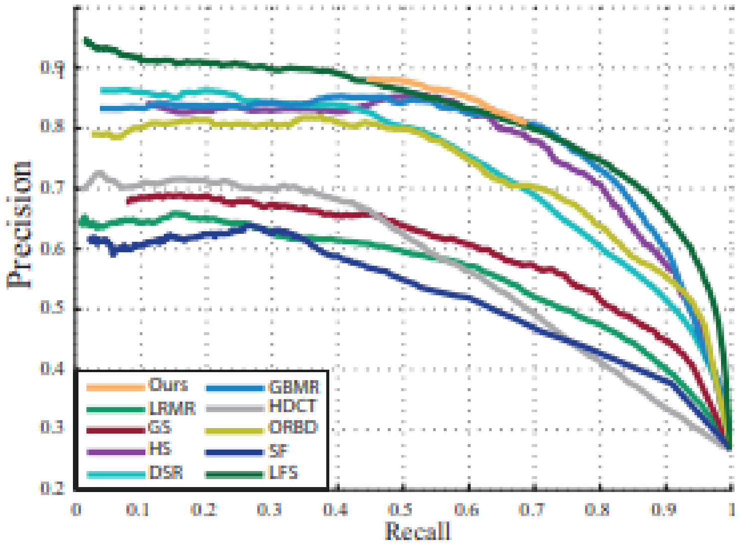
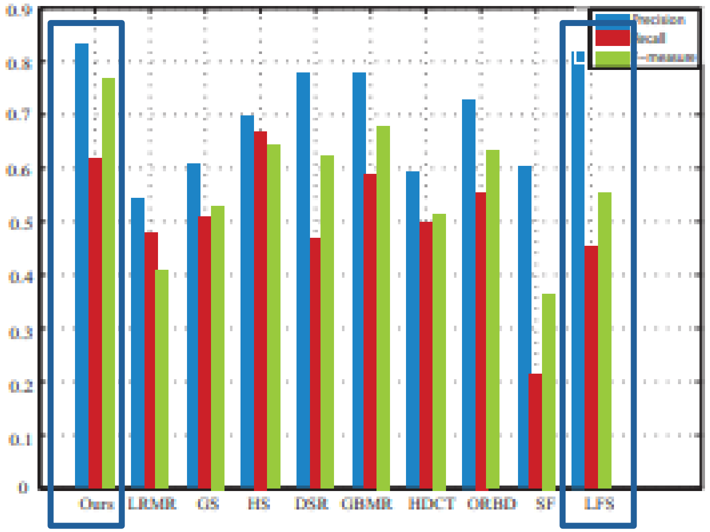


Figure 5. Visual Comparisons of different saliency detection algorithms vs. ours on our light field dataset.

2015 WSC







(d) LFSD

2015 DILF

- **PRC**
- **ROC** (based on true & false positives obtained after calculation of PR curve)
- **F-measure**
- **AUC** (Area under curve)
- **MAE** (measure the average per-pixel difference bt binary GT & saliency map)

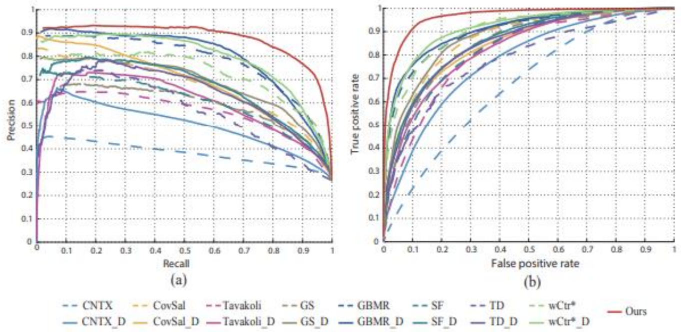


Figure 5: Quantitative results of our approach, state-of-the-art 2D approaches and their depth-extended versions. (a) PR curves; (b) ROC curves.

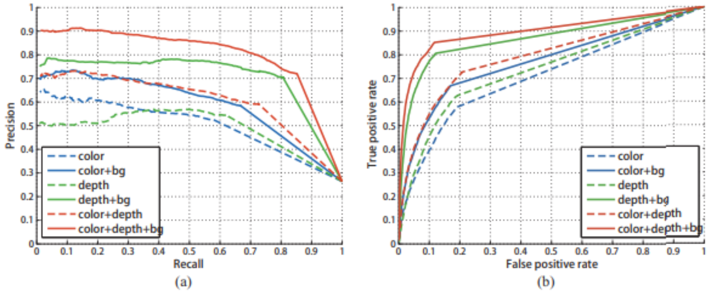


Figure 3: Quantitative measurements of various light field properties on LFSD datasets. (a) PR curves; (b) ROC curves.

Model	F-measure	AUC	MAE
Color	0.5923	0.7089	0.2367
Color+Bg	0.6390	0.7708	0.2157
Depth	0.5587	0.7354	0.2421
Depth+Bg	<u>0.7297</u>	<u>0.8676</u>	<u>0.1708</u>
Color+Depth	0.6422	0.7904	0.2255
Color+Depth+Bg	**0.7749**	**0.8982**	**0.1605**

Table 1: Comparisons of F-measure, ROC and MAE from different light field properties (bold: best; underline: second best).

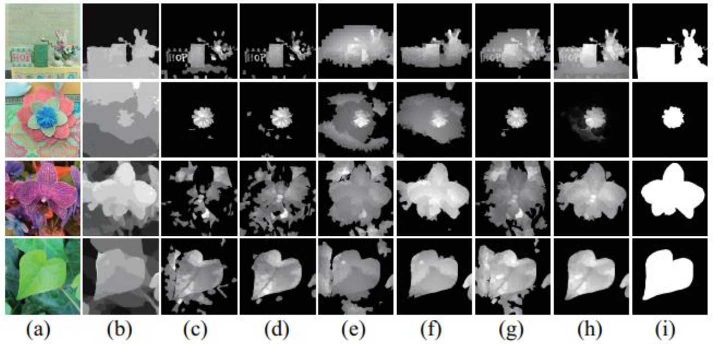


Figure 4: Visual comparisons of saliency estimation from different light field properties. (a) all-focus image; (b) depth map; (c) color; (d) color+bg; (e) depth; (f) depth+bg; (g) color+depth; (h) color+depth+background; (i) GT.

Model	F-measure	AUC	MAE
CNTX	0.3643	0.6700	0.3574
CNTX_D	0.4123	0.7718	0.3514
CovSal	0.6335	0.8599	0.2417
CovSal_D	0.6373	0.8466	0.2850
Tavakoli	0.5498	0.8078	0.2551
Tavakoli_D	0.5711	0.8276	0.2903
GS	0.5944	0.8443	0.2395
GS_D	0.6217	0.8792	0.2843
GBMR	0.7461	0.8965	<u>0.1822</u>
GBMR_D	0.7536	0.9072	0.2415
SF	0.4678	0.8301	0.2468
SF_D	0.4704	0.8552	0.2903
TD	0.5766	0.7775	0.2623
TD_D	0.5999	0.8490	0.2951
wCtr*	0.6996	0.8991	0.1878
wCtr*_D	0.7382	0.9156	0.2475
DVS	0.2723	0.5354	0.3274
DVS_Bg	0.2851	0.5509	0.2846
ACSD	0.7905	<u>0.9467</u>	0.1830
ACSD_Bg	<u>0.8025</u>	0.8361	0.1668
LFS	0.7500	0.9272	0.2077
Ours	**0.8186**	**0.9641**	**0.1363**

ble 2: Comparisons of F-measure, AUC, and MAE from our proach, state-of-the-art 2D/3D approaches and their light field-tended methods (bold: best; underline: second best).

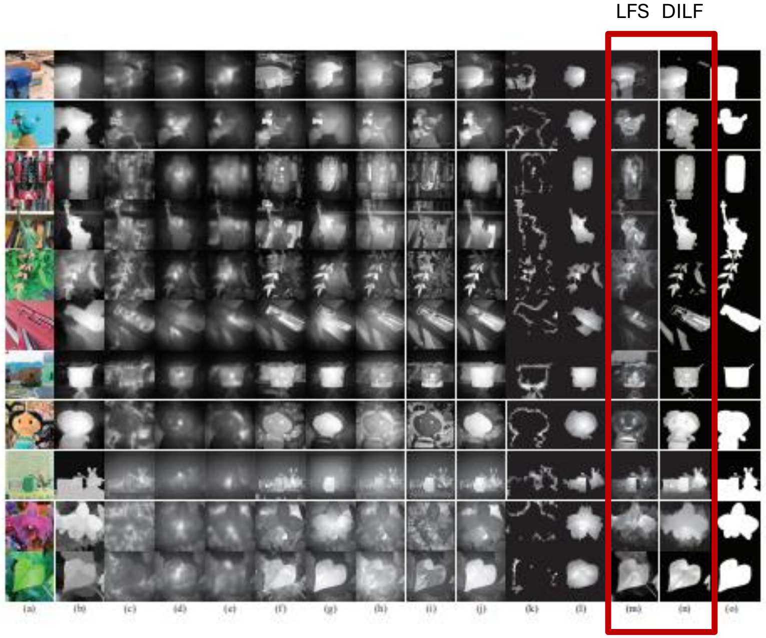


Figure 7: Visual comparisons of our approach and 2D/3D extended methods. (a) all-focus image; (b) depth map; (c) CNTX_D; (d) CovSal_D; (e) Tavakoli_D; (f) GS_D; (g) GBMR_D; (h)SF_D; (i) TD_D; (j) wCtr*_D; (k) DVS_Bg; (l) ACSD_Bg; (m) LFS; (n) Ours; (o) GT.

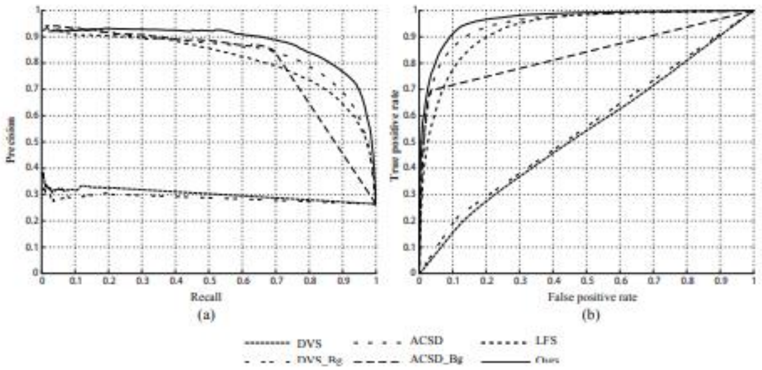


Figure 8: Quantitative results of our approach, LFS, state-of-the-art 3D approaches and their focusness-extended versions. (a) PR curves; (b) ROC curves.

2017 MULTI-CUE

Developed HFUT-Lytro dataset (not available)

- **PRC**
- **AP**
- **F-measure**
- **MAE**

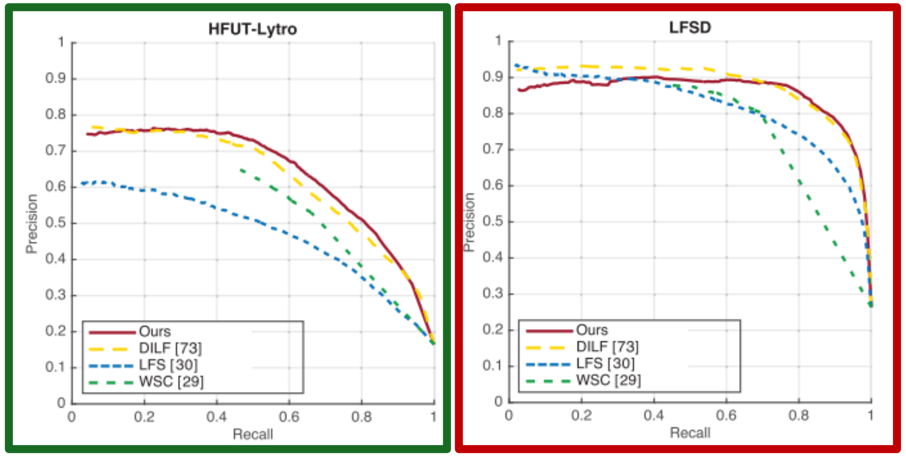


Fig. 5. The PR curves obtained by different methods on the HFUT-Lytro (left) and LFSD (right) datasets.

Table II. The Precision, Recall, F-measure, AP, and MAE Obtained by Different Methods on the HFUT-Lytro and LFSD Datasets (Bold: Best; Underline: Second Best)

Dataset	Metric	Ours	DILF[73]	WSC[29]	LFS[30]
HFUT-Lytro	Precision	**0.5928**	0.5186	0.5254	0.4753
	Recall	0.6726	**0.7147**	0.6673	0.5354
	F-measure	**0.6095**	0.5537	0.5525	0.4880
	AP	**0.6354**	0.6221	0.4743	0.4718
	MAE	**0.1388**	0.1578	0.1454	0.2214
LFSD	Precision	**0.8542**	0.8271	0.8076	0.8115
	Recall	0.7397	**0.7916**	0.6783	0.6083
	F-measure	**0.8247**	0.8186	0.7735	0.7534
	AP	0.8625	**0.8787**	0.6832	0.8161
	MAE	0.1503	**0.1363**	0.1453	0.2072

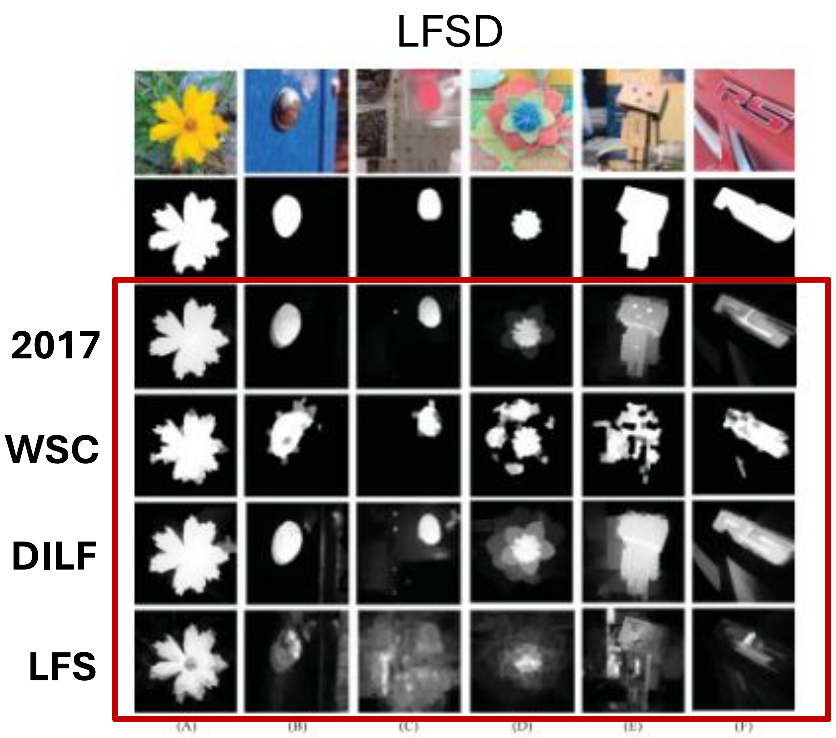


Fig. 4. Saliency detection results of different methods on the LFSD dataset. From top to bottom: all-in-focus images, ground-truth maps, and saliency maps obtained by our approach, WSC [29], DILF [73], and LFS [30].

w/ and w/o SLIC (for segmentation)

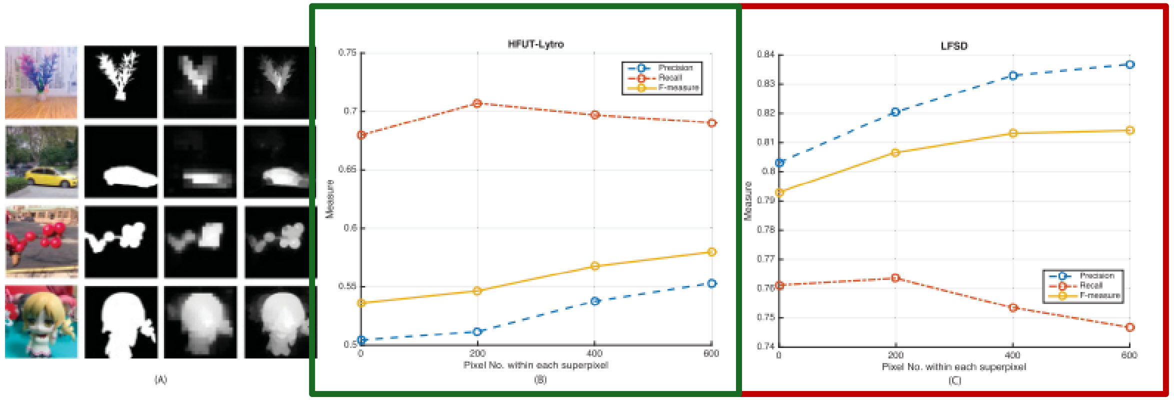


Fig. 12. Precision, recall, F-measure, and qualitative comparisons of saliency maps with and without SLIC. (A) Qualitative comparison of saliency maps. From left to right: all-focus images, ground truth, w/o SLIC, and w/ SLIC ($N = 400$); ((B) and (C)) Quantitative results on the HFUT-Lytro and LFSD datasets w.r.t. different numbers of pixels within each superpixel. 0 indicates regular grid sampling.

Table III. Computational Time of Our Approach and the State-of-the-art Methods for Processing One Image

Methods	Our approach	DILF [73]	WSC [29]	LFS [30]
Runtime (seconds)	4.2	0.9	8.5	8.1

w/ and w/o refinement (to optimize the saliency graph)

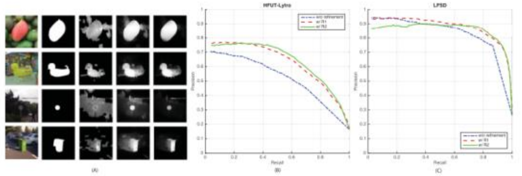


Fig. 10. PR-curve and qualitative comparisons of saliency maps with and without refinement. (A) Qualitative comparison of saliency maps. From left to right: all-focus images, ground truth, w/o refinement, w/ R1, and w/ R2; (B) HFUT-Lytro dataset; (C) LFSD dataset.

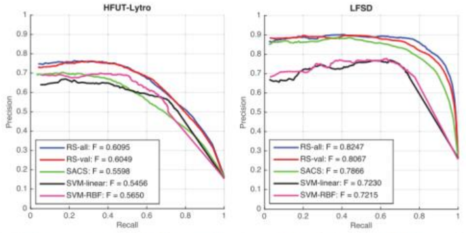


Fig. 11. Performance comparison of different multi-cue integration methods on the HFUT-Lytro and LFSD datasets. Here, we illustrate the PR curves and F-measures.

Contribution of individual cues

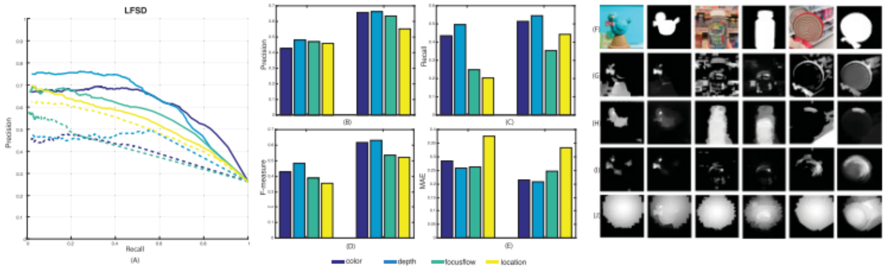


Fig. 7. Quantitative and qualitative comparisons of saliency maps from our framework with individual light-field cues and their refinements on the LFSD dataset. (A) PR curves of individual saliency maps (dashed line) and their refinements (solid line); (B)–(E) quantitative results of individual saliency maps (left) and their refinements (right) for each metric; (F) all-in-focus images and the corresponding ground-truth saliency maps; (G)–(J) color-, depth-, focusing flow-, and position-driven saliency maps and their refined versions with the structure cue.

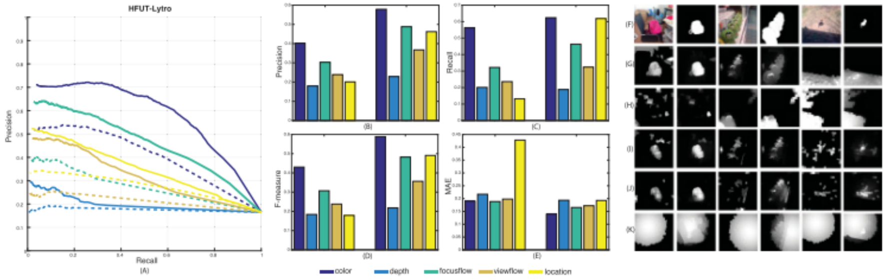
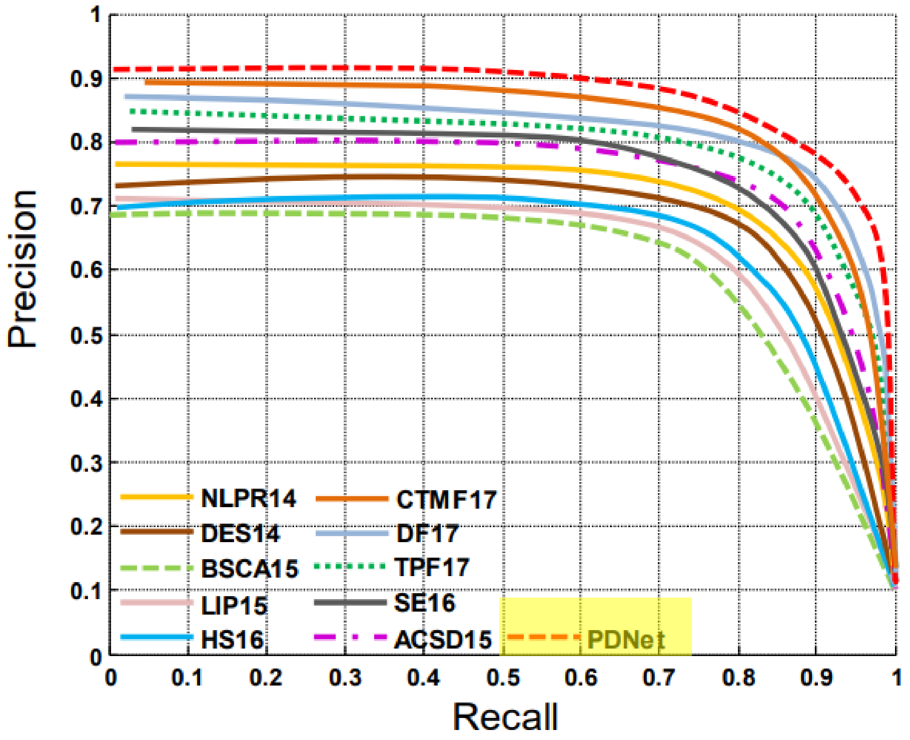


Fig. 8. Quantitative and qualitative comparisons of saliency maps from our framework with individual light-field cues and their refinements on the HFUT-Lytro dataset. (A) PR curves of individual saliency maps (dashed line) and their refinements (solid line); (B)–(E) quantitative results of individual saliency maps (left) and their refinements (right) for each metric; (F) all-in-focus images and the corresponding ground-truth saliency maps; (G)–(K) color-, depth-, focusing flow-, viewing flow-, and position-driven saliency maps and their refined versions with the structure cue.

2018 – PDNet

Method	NJU2000-TE [15]		NLPR-TE [14]		LFSD [21]		RGBD135 [13]		SSD100 [22]	
	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE
PDNet	**0.8503**	**0.0689**	**0.8478**	**0.0491**	**0.8219**	**0.0752**	**0.8805**	**0.0384**	**0.8152**	**0.0812**
CTMF17 [5]	-	-	-	-	0.8025	0.0912	0.8102	0.0653	0.7925	0.0912
DF17 [17]	-	-	-	-	0.8109	0.0815	0.8059	0.6871	0.7858	0.0846
TPF17 [22]	0.7213	0.1488	0.7190	0.0852	0.7925	0.1058	0.7395	0.0891	0.7541	0.1217
SE16 [16]	0.6946	0.1687	0.7101	0.0904	0.7568	0.1156	0.5807	0.1253	0.6666	0.1648
ACSD15 [15]	0.6747	0.1939	0.6019	0.1624	0.7865	0.1425	0.6851	0.1518	0.6382	0.2010
NLPR14[14]	0.6165	0.1669	0.5957	0.1087	0.7356	0.1547	0.4912	0.1165	0.6415	0.1784
DES14 [13]	0.6202	0.4465	0.5915	0.3207	0.7254	0.2168	0.5410	0.3079	0.5797	0.3132
BSCA15 [8]	0.6290	0.2148	0.5925	0.1754	0.7126	0.1894	0.5826	0.1851	0.5755	0.2386
LIP15 [10]	0.5692	0.2059	0.5890	0.1252	0.7176	0.1880	0.5452	0.1406	0.5935	0.1960
HS16 [9]	0.6090	0.2516	0.6003	0.1909	0.7248	0.1751	0.5361	0.1849	0.5716	0.2582



(c) PR curve on LFSD

2019 –single view

- **S-measure**
- **E-measure** (pixel-level matching & img-level statistics)
- **F-measure**
- **MAE**

Developed DUTLF-FS dataset

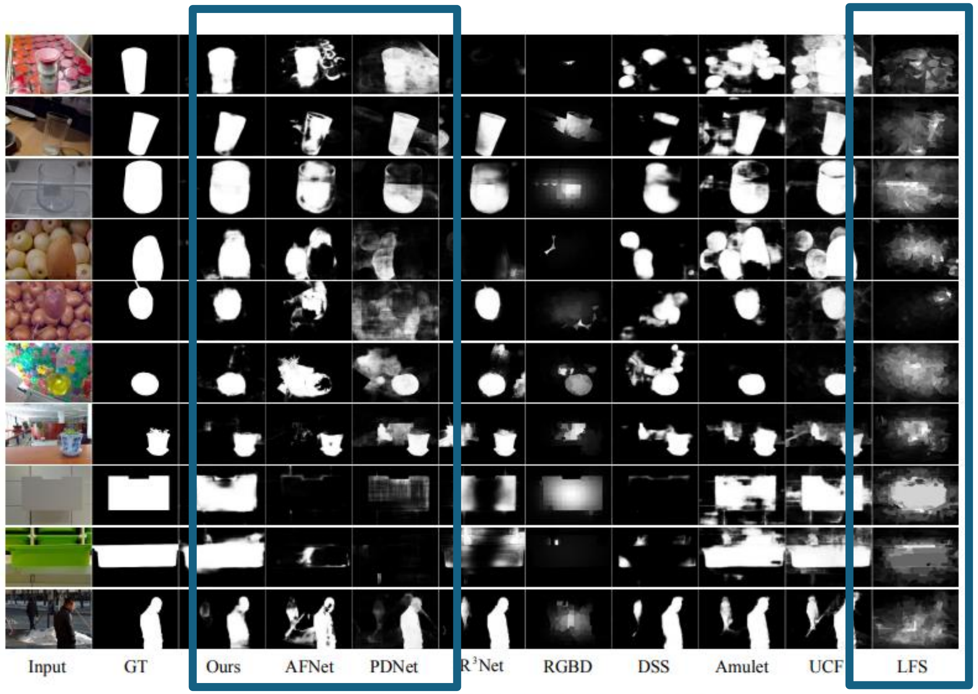


Figure 6: Visual comparsion of saliency maps on the proposed dataset.

RSF- = without rich saliency feature extraction technique
MAVM- = average the warped multi-view saliency maps

Model	S-measure	F-measure	E-measure	MAE
RSF^-	0.801	0.731	0.849	0.097
$MVAM^-$	0.803	0.730	0.845	0.094
Ours	**0.806**	**0.749**	**0.861**	**0.088**

Table 2: Ablation analysis on the proposed dataset.

Model	Input	S-measure	F-measure	E-measure	MAE
DSR	2D	0.640	0.574	0.767	0.173
BSCA	2D	0.673	0.605	0.777	0.198
MST	2D	0.637	0.548	0.738	0.179
ACSD	3D	0.675	0.637	0.792	0.188
RGBD	3D	0.535	0.567	0.732	0.179
LFS	4D	0.538	0.423	0.717	0.242
UCF	2D	0.788	0.709	0.814	0.136
UCF+	2D	0.792	0.715	0.821	0.131
Amulet	2D	0.801	0.734	0.839	0.104
DSS	2D	0.740	0.709	0.228	0.112
DSS+	2D	0.731	0.674	0.795	0.141
R^3Net	2D	0.793	**0.749**	0.851	0.089
PDNet	3D	0.761	0.692	0.827	0.126
AFNet	3D	0.731	0.687	0.822	0.109
Ours	2D	**0.806**	**0.749**	**0.861**	**0.088**

Table 1: Quantitative comparison of S-measure, F-measure, E-measure, and MAE scores. The retrained models are denoted as "XX+". (bold: best; underline: second best).

2019 – dl4lfsd

Datasets		MST	BSCA	DCL	DHS	DSS	Amulet	UCF	PAGRN	PiCANet	R^3Net	DFRGBD	RGBD	ACSD	DILF	LFS	WSC	Ours
LFSD	maxF	0.704	0.795	0.780	0.856	0.768	0.863	0.865	0.840	0.867	0.876	0.841	0.841	0.780	0.849	0.779	0.786	0.863
	MAE	0.209	0.205	0.161	0.115	0.178	0.093	0.143	0.132	0.111	0.098	0.180	0.197	0.218	0.153	0.239	0.168	0.093
	S-m	0.646	0.725	0.742	0.803	0.678	0.801	0.808	0.766	0.822	0.811	0.732	0.650	0.681	0.801	0.655	0.700	0.826
	E-m	0.720	0.766	0.784	0.844	0.865	0.847	0.844	0.791	0.847	0.852	0.737	0.650	0.675	0.845	0.625	0.770	0.877
Ours	maxF	0.545	0.642	0.716	0.816	0.735	0.782	0.789	0.849	0.851	0.761	0.722	0.570	0.262	-	0.439	-	0.868
	MAE	0.210	0.215	0.156	0.095	0.132	0.070	0.153	0.084	0.089	0.114	0.163	0.202	0.337	-	0.259	-	0.070
	S-m	0.594	0.66	0.710	0.803	0.714	0.777	0.770	0.810	0.838	0.733	0.687	0.5	0.357	-	0.517	-	0.852
	E-m	0.717	0.742	0.781	0.865	0.784	0.843	0.828	0.841	0.872	0.808	0.684	0.432	0.545	-	0.545	-	0.905

Table 2. Quantitative comparison of maximum F-measure, MAE, S-measure, E-measure scores on two datasets. The color in red and blue represent the best and second scores.

2020

2020 2017

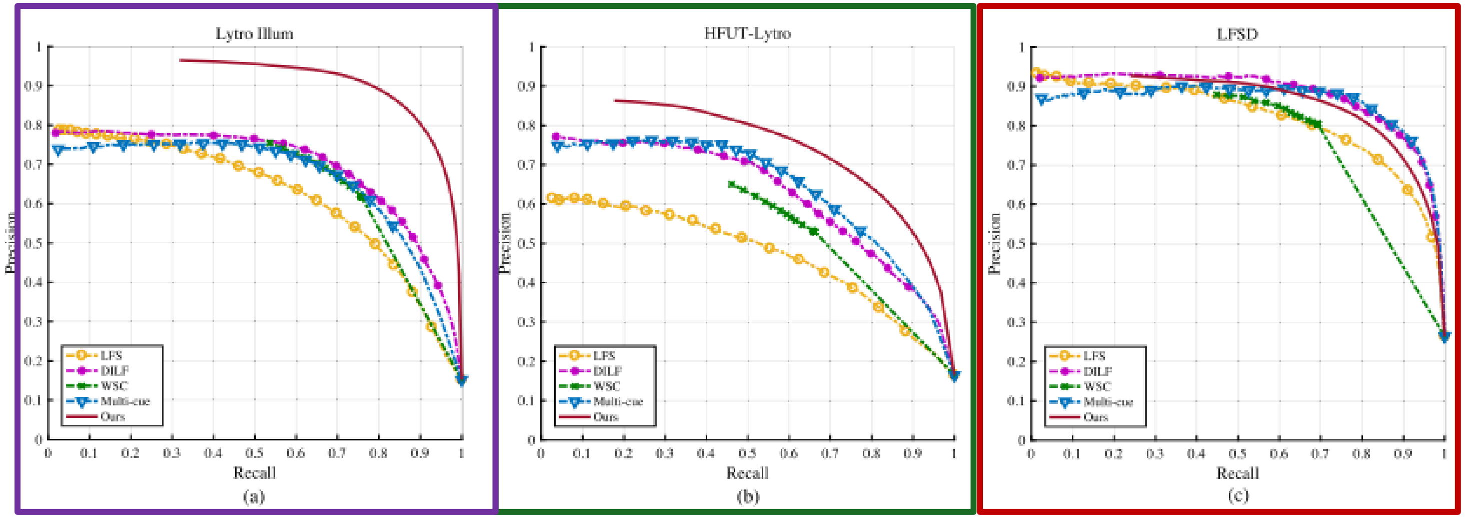


Fig. 13. Comparison on three datasets in terms of PR curve. (a) The proposed Lytro Illum dataset. (b) The HFUT-Lytro dataset. (c) The LFSD dataset.

Developed LYTRO ILLUM dataset

2020 LFS DILF WSC Multi-cue

Lytro-Illum

HFUT-Lytro

LFSD

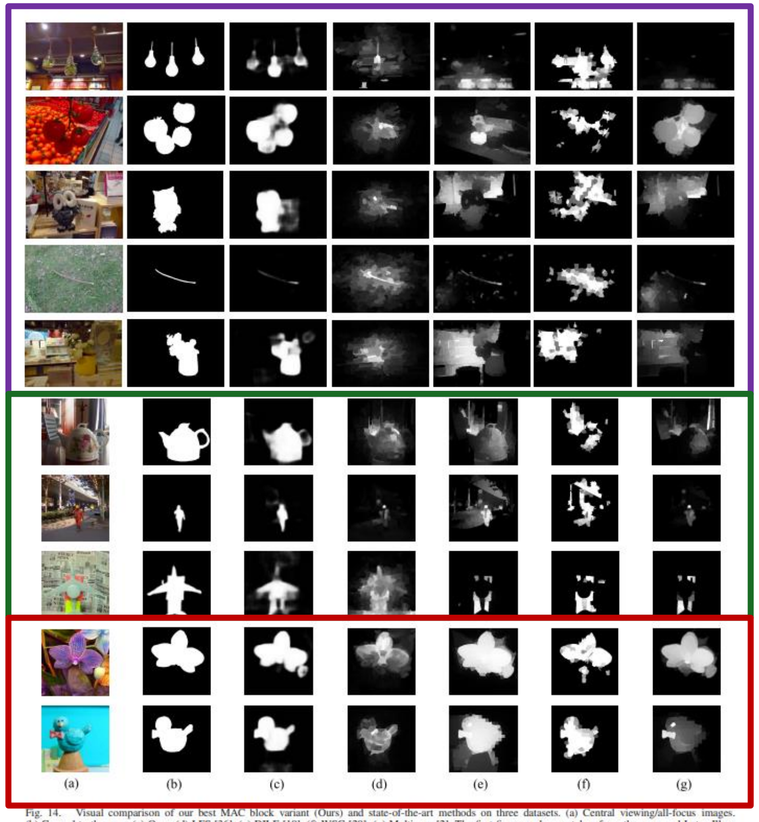


Fig. 14. Visual comparison of our best MAC block variant (Ours) and state-of-the-art methods on three datasets. (a) Central viewing/all-focus images. (b) Ground truth maps. (c) Ours. (d) LFS [36]. (e) DILF [19]. (f) WSC [20]. (g) Multi-cue [2]. The first five samples are taken from the proposed Lytro Illum ... the last two samples are taken from the LFSD dataset.

TABLE VII

QUANTITATIVE RESULTS ON THE LFSD DATASET. BOLD: BEST, UNDERLINED: SECOND BEST

	Method	F-measure	WF-measure	MAE	AP
2014	LFS [1]	0.7525	0.5319	0.2072	0.8161
2015	WSC [20]	0.7729	<u>0.7371</u>	0.1453	0.6832
2015	DILF [19]	<u>0.8173</u>	0.6695	<u>0.1363</u>	**0.8787**
2017	Multi-cue [2]	**0.8249**	0.7155	0.1503	<u>0.8625</u>
	Ours	0.8105	**0.7378**	**0.1164**	0.8561

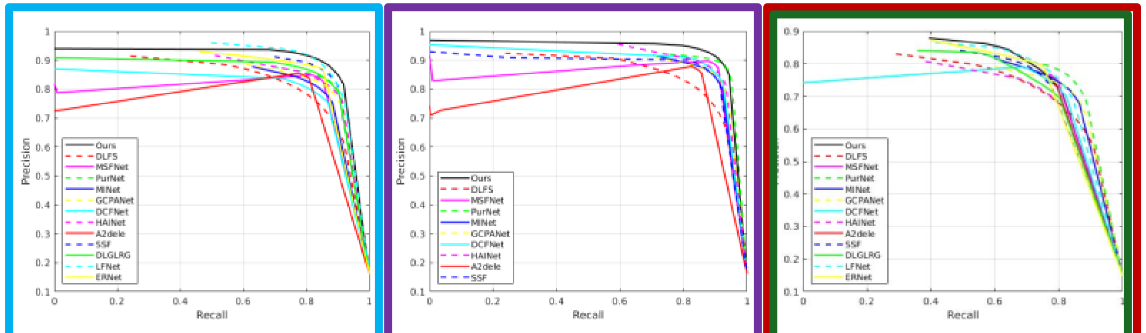
QUANTITATIVE RESULTS ON THE PROPOSED LYTRO ILLUM DATASET. BOLD: BEST, UNDERLINED: SECOND BEST

Method	F-measure	WF-measure	MAE	AP
LFS [1]	0.6107	0.3596	0.1697	0.6193
WSC [20]	0.6451	<u>0.5945</u>	<u>0.1093</u>	0.5958
DILF [19]	0.6395	0.4844	0.1389	<u>0.6921</u>
Multi-cue [2]	<u>0.6648</u>	0.5420	0.1197	0.6593
Ours	**0.8116**	**0.7540**	**0.0551**	**0.9124**

QUANTITATIVE RESULTS ON THE HFUT-LYTRO DATASET. BOLD: BEST, UNDERLINED: SECOND BEST

Method	F-measure	WF-measure	MAE	AP
LFS [1]	0.4868	0.3023	0.2215	0.4718
WSC [20]	0.5552	0.5080	0.1454	0.4743
DILF [19]	0.5543	0.4468	0.1579	0.6221
Multi-cue [2]	<u>0.6135</u>	<u>0.5146</u>	<u>0.1388</u>	<u>0.6354</u>
Ours	**0.6721**	**0.6087**	**0.1029**	**0.7390**

2022



(a) DUTLF-V2 (b) LYTRO ILLUM (c) HFUT-LFSD

Fig. 6. PR curves of the proposed method and other representative state-of-the-art methods on three light field saliency datasets.

TABLE II

QUANTITATIVE COMPARISONS OF CROSS ENTROPY ON THREE LIGHT FIELD DATASETS. BOLD: BEST

Type	Methods	DUTLF-V2	LYTRO ILLUM	HFUT-LFSD
4D	DLFS	.082	.132	.142
4D	Ours	**.062**	**.106**	.104
4D	DLGLRG	.079	-	.154
	LFNet	.066	-	**.098**
	ERNet	.079	-	.136
	MoLF	.083	-	.112
3D	DCFNet	.098	.110	.235
	HAINet	.066	.152	.167
	S2MA	.077	.111	.108
	A2dele	.086	.136	.215
	SSF	.065	.124	.143
	BBSNet	.074	.111	.183
	DMRA	.067	.124	.105
2D	MSFNet	.116	.144	.280
	PurNet	.072	.119	.143
	MINet	.076	.117	.130
	GCPANet	.110	.109	.238
	F^3Net	.101	.109	.169
	EGNet	.076	.116	.117
	CPD	.072	.112	.183
	PoolNet	.077	.182	.114
	R^3Net	.179	.185	.466

TABLE I

QUANTITATIVE COMPARISONS OF E-MEASURE, S-MEASURE, WEIGHTED F-MEASURE, F-MEASURE AND MAE SCORES ON THREE LIGHT FIELD DATASETS. BOLD: BEST, UNDERLINE: SECOND BEST (* REPRESENTS CONVENTIONAL METHODS, - MEANS NO AVAILABLE RESULTS)

Type	Methods	Years	DUTLF-V2					LYTRO ILLUM					HFUT-LFSD				
			$E_s\uparrow$	$S_\alpha\uparrow$	$F_\beta^w\uparrow$	$F_\beta\uparrow$	MAE$\downarrow$	$E_s\uparrow$	$S_\alpha\uparrow$	$F_\beta^w\uparrow$	$F_\beta\uparrow$	MAE$\downarrow$	$E_s\uparrow$	$S_\alpha\uparrow$	$F_\beta^w\uparrow$	$F_\beta\uparrow$	MAE$\downarrow$
4D	DLFS	-	.868	.810	.676	.739	.080	.876	.830	.712	.774	.072	.799	.759	.611	.650	.097
4D	Ours	-	**.931**	**.882**	**.818**	**.852**	**.041**	**.936**	**.906**	**.858**	**.884**	**.038**	.848	.789	.672	.724	**.061**
	△gains		↑.063	↑.072	↑.142	↑.113	↑.039	↑.060	↑.076	↑.146	↑.110	↑.034	↑.049	↑.030	↑.061	↑.074	↑.036
4D	DLGLRG	ICCV'21	.908	.861	.780	.816	.046	-	-	-	-	-	.843	.765	.634	.709	.071
	LFNet	TIP'20	.915	.873	.799	.819	.047	-	-	-	-	-	.852	**.807**	.693	.718	.062
	ERNet	AAAI'20	.924	.852	.792	**.852**	.050	-	-	-	-	-	**.858**	.778	.687	**.753**	.069
	MoLF	NIPS'19	.915	.877	.803	.821	.047	-	-	-	-	-	.851	.795	.684	.722	.068
	DILF*	IJCAI'15	.733	.648	.388	.504	.187	.785	.731	.500	.608	.149	.736	.695	.458	.555	.131
	LFS*	CVPR'14	-	-	-	-	-	-	-	-	-	-	.686	.579	.264	.430	.205
3D	DCFNet	CVPR'21	.878	.821	.732	.772	.065	.908	.879	.823	.845	.046	.822	.779	.670	.703	.077
	HAINet	TIP'21	.886	.845	.760	.794	.060	.911	.882	.818	.848	.049	.792	.755	.628	.672	.097
	S2MA	CVPR'20	.844	.803	.679	.729	.087	.883	.868	.776	.802	.061	.770	.729	.573	.616	.112
	A2dele	CVPR'20	.888	.836	.771	.807	.048	.898	.854	.800	.834	.050	.833	.782	.688	.715	.069
	SSF	CVPR'20	.917	.869	.804	.831	.043	.903	.874	.810	.836	.049	.835	.781	.673	.714	.068
	BBSNet	ECCV'20	.893	.851	.754	.794	.059	.902	.874	.802	.834	.050	.806	.759	.612	.675	.086
	DMRA	ICCV'19	.897	.822	.740	.800	.060	.903	.845	.781	.837	.059	.844	.765	.646	.705	.073
2D	MSFNet	ACMMM'21	.888	.848	.784	.815	.050	.906	.882	.835	.857	.049	.809	.776	.684	.712	.083
	PurNet	TIP'21	.886	.849	.771	.802	.059	.918	.897	.846	.862	.046	.825	.798	**.703**	.721	.084
	MINet	CVPR'20	.870	.828	.736	.781	.065	.890	.864	.795	.825	.057	.800	.777	.670	.704	.090
	GCPANet	AAAI'20	.869	.838	.743	.782	.071	.888	.874	.800	.826	.057	.788	.772	.663	.683	.108
	F^3Net	AAAI'20	.878	.841	.756	.803	.063	.901	.876	.812	.843	.052	.810	.776	.673	.707	.094
	EGNet	ICCV'19	.855	.821	.710	.746	.078	.902	.884	.820	.843	.052	.794	.772	.634	.672	.094
	CPD	CVPR'19	.886	.836	.753	.794	.062	.895	.874	.802	.831	.052	.811	.766	.652	.691	.097
	PoolNet	CVPR'19	.876	.832	.732	.774	.069	.889	.867	.785	.820	.059	.803	.777	.652	.685	.092
	R^3Net	IJCAI'18	.842	.767	.665	.712	.083	.901	.857	.807	.837	.051	.741	.726	.622	.663	.136

2023

TABLE II

BENCHMARKING RESULTS OF REPRESENTATIVE 2D, 3D, AND 4D MODELS ON FOUR EXISTING AND OUR PROPOSED DATASETS

	Metric	Traditional LFS‡ [18]	WSC‡ [55]	DILF‡ [56]	Deep Learning-based MoLF‡ [21]	ERNet‡ [23]	BBS† [17]	JLDCF† [5]	SSF† [71]	UCNet† [1]	D3Net† [15]	S2MA† [72]	cmMS† [73]	HDF† [74]	ATSA† [16]	MINet* [13]	GCPA* [14]	$\text{Ours}_1^\ddagger$	$\text{Ours}_2^\ddagger$	$\text{Ours}_3^\ddagger$
LFSD [18]	S_α ↑	.681	.702	.811	.831	.835	.864	.862	.859	.858	.825	.837	.850	.846	.858	.815	.830	.859	.864	.871
	F_β^{max} ↑	.744	.743	.811	.834	.850	.858	.867	.868	.859	.812	.835	.858	.837	.866	.790	.811	.868	.871	.877
	F_β^{mean} ↑	.513	.722	.719	.809	.836	.842	.848	.862	.848	.797	.806	.850	.818	.856	.781	.800	.853	.864	.868
	F_β^{adp} ↑	.735	.743	.795	.819	.839	.840	.827	.862	.838	.788	.803	.857	.818	.852	.810	.816	.860	.866	.872
	E_ϕ^{max} ↑	.809	.789	.861	.888	.888	.900	.902	.901	.898	.863	.873	.896	.880	.902	.840	.850	.905	.907	.909
	E_ϕ^{mean} ↑	.567	.753	.764	.872	.883	.883	.894	.890	.893	.850	.855	.881	.869	.899	.834	.839	.902	.902	.905
	E_ϕ^{adp} ↑	.773	.788	.846	.886	.887	.889	.882	.896	.890	.853	.863	.890	.872	.897	.864	.869	.906	.905	.910
	$\mathcal{M}$ ↓	.205	.150	.136	.088	.082	.072	.070	.067	.072	.095	.094	.073	.086	.068	.096	.093	.067	.065	.062
HFUT [25]	S_α ↑	.565	.613	.672	.742	.778	.751	.789	.725	.748	.749	.729	.723	.763	.772	.769	.775	.810	.789	.834
	F_β^{max} ↑	.427	.508	.601	.662	.722	.676	.727	.647	.677	.671	.650	.626	.690	.729	.692	.701	.784	.779	.810
	F_β^{mean} ↑	.323	.493	.513	.639	.709	.654	.707	.639	.672	.651	.623	.617	.669	.706	.683	.682	.771	.770	.805
	F_β^{adp} ↑	.427	.485	.530	.627	.706	.654	.677	.636	.675	.647	.588	.636	.653	.689	.691	.687	.769	.767	.804
	E_ϕ^{max} ↑	.637	.695	.748	.812	.841	.801	.844	.778	.804	.797	.777	.784	.801	.833	.804	.812	.868	.841	.884
	E_ϕ^{mean} ↑	.524	.684	.657	.790	.832	.765	.825	.763	.793	.773	.756	.746	.788	.819	.787	.794	.864	.836	.879
	E_ϕ^{adp} ↑	.666	.680	.693	.785	.831	.804	.811	.781	.810	.789	.744	.779	.789	.810	.816	.822	.865	.838	.880
	$\mathcal{M}$ ↓	.221	.154	.150	.094	.082	.089	.075	.100	.090	.091	.112	.097	.095	.084	.088	.094	.067	.072	.057
DUT-LF [20]	S_α ↑	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	F_β^{max} ↑	.533	.621	.585	.903	.908	.852	.878	.887	.816	.797	.754	.803	.801	.915	.855	.866	.928	.929	.941
	F_β^{mean} ↑	.358	.610	.492	.855	.891	.834	.846	.879	.806	.776	.733	.774	.776	.900	.845	.849	.906	.911	.928
	F_β^{adapt} ↑	.525	.619	.597	.843	.885	.848	.835	.885	.803	.784	.735	.819	.778	.898	.862	.865	.906	.912	.929
	E_ϕ^{max} ↑	.711	.789	.757	.939	.949	.900	.925	.923	.876	.860	.839	.879	.864	.941	.895	.904	.959	.962	.965
	E_ϕ^{mean} ↑	.511	.762	.635	.921	.943	.879	.911	.907	.870	.841	.817	.817	.848	.937	.886	.894	.954	.958	.961
	E_ϕ^{adapt} ↑	.742	.789	.784	.923	.943	.908	.910	.918	.878	.869	.842	.870	.867	.938	.915	.922	.956	.961	.964
	$\mathcal{M}$ ↓	.227	.149	.165	.051	.039	.066	.058	.050	.081	.083	.102	.079	.091	.041	.060	.061	.033	.030	.027
Lytro Illum [22]	S_α ↑	.619	.708	.751	.834	.843	.876	.881	.822	.852	.860	.856	.872	.856	.882	.862	.874	.878	.884	.890
	F_β^{max} ↑	.545	.663	.688	.820	.827	.848	.868	.787	.827	.836	.832	.849	.835	.875	.829	.845	.868	.873	.888
	F_β^{mean} ↑	.385	.646	.599	.766	.800	.830	.840	.776	.817	.809	.795	.832	.806	.848	.812	.823	.850	.856	.877
	F_β^{adapt} ↑	.547	.640	.666	.747	.796	.830	.826	.780	.821	.801	.788	.833	.809	.842	.818	.825	.846	.855	.878
	E_ϕ^{max} ↑	.721	.804	.827	.908	.911	.909	.926	.877	.899	.905	.903	.907	.901	.929	.888	.898	.928	.928	.936
	E_ϕ^{mean} ↑	.546	.792	.721	.882	.900	.896	.914	.865	.893	.889	.882	.897	.887	.919	.881	.887	.916	.922	.932
	E_ϕ^{adapt} ↑	.771	.798	.817	.876	.900	.911	.914	.885	.905	.901	.886	.912	.894	.917	.900	.908	.917	.925	.936
	$\mathcal{M}$ ↓	.197	.115	.127	.065	.056	.047	.044	.066	.053	.055	.060	.045	.056	.042	.051	.050	.045	.043	.037
PKU-LF	S_α ↑	.579	.641	.618	.809	.826	.847	.854	.841	.792	.802	.765	.846	.822	.860	.841	.852	.862	.876	.887
	F_β^{max} ↑	.424	.540	.529	.776	.781	.802	.811	.798	.736	.742	.683	.797	.775	.839	.787	.804	.843	.863	.878
	F_β^{mean} ↑	.325	.519	.507	.716	.761	.775	.785	.786	.730	.717	.660	.782	.753	.814	.776	.785	.820	.854	.870
	F_β^{adapt} ↑	.435	.505	.498	.683	.754	.761	.778	.787	.737	.715	.651	.777	.744	.800	.784	.784	.809	.852	.870
	E_ϕ^{max} ↑	.685	.750	.745	.883	.883	.882	.889	.884	.852	.858	.811	.880	.873	.909	.870	.876	.918	.919	.930
	E_ϕ^{mean} ↑	.546	.730	.731	.850	.870	.864	.878	.876	.842	.828	.791	.869	.858	.898	.858	.862	.909	.914	.927
	E_ϕ^{adapt} ↑	.693	.722	.727	.832	.867	.876	.880	.889	.864	.855	.808	.881	.861	.891	.885	.888	.909	.917	.930
	$\mathcal{M}$ ↓	.214	.132	.143	.066	.059	.056	.049	.052	.070	.067	.100	.052	.065	.045	.050	.055	.047	.042	.035

Note that the 2D, 3D, and 4D SOD models are marked with "*", "†", and "‡", respectively. "↑"/"↓" indicates that larger/smaller is better. The top-performance