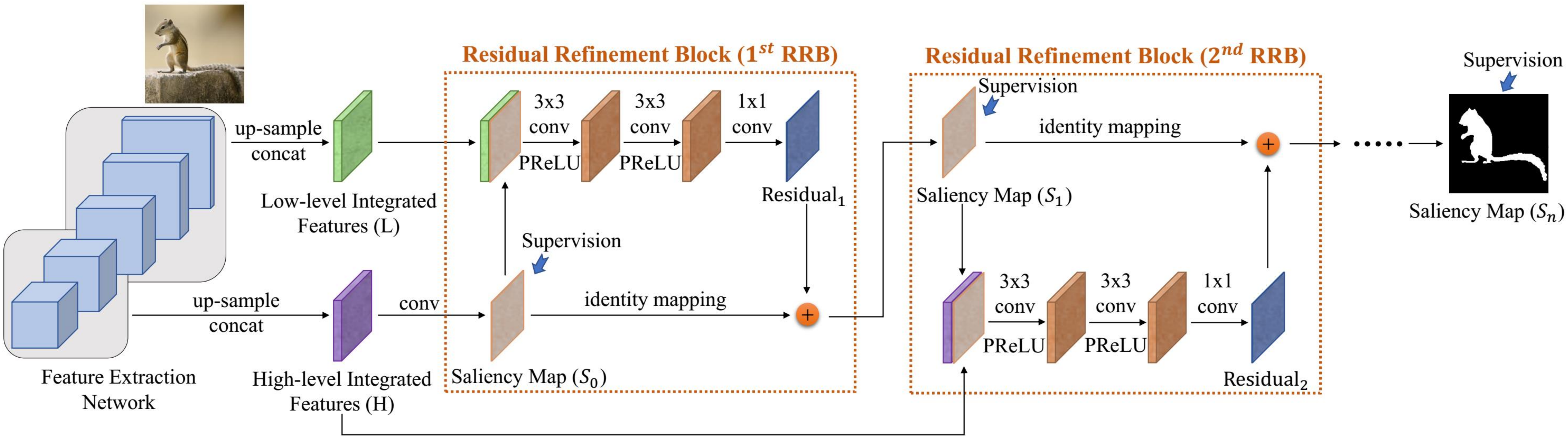
R^3Net: Recurrent Residual Refinement Network for Saliency Detection

Zijun Deng[1,*], Xiaowei Hu[2,*], Lei Zhu[3,2], Xuemiao Xu[1], Jing Qin[3], Guoqiang Han[1], and Pheng-Ann Heng[2,4]

[1]South China University of Technology
[2]The Chinese University of Hong Kong
[3]The Hong Kong Polytechnic University
[4]Shenzhen Institutes of Advanced Technology

Architecture



Introduction

- **Problem Definition**

Saliency detection aims to highlight the most visually distinctive objects in an image.

- **Contributions**

➢ Design a novel residual refinement block (RRB) to learn the residual between the ground truth and the saliency map at each recurrent step. This learning strategy can make the network easy to train and help to learn the complementary information of previous prediction for the refinement.

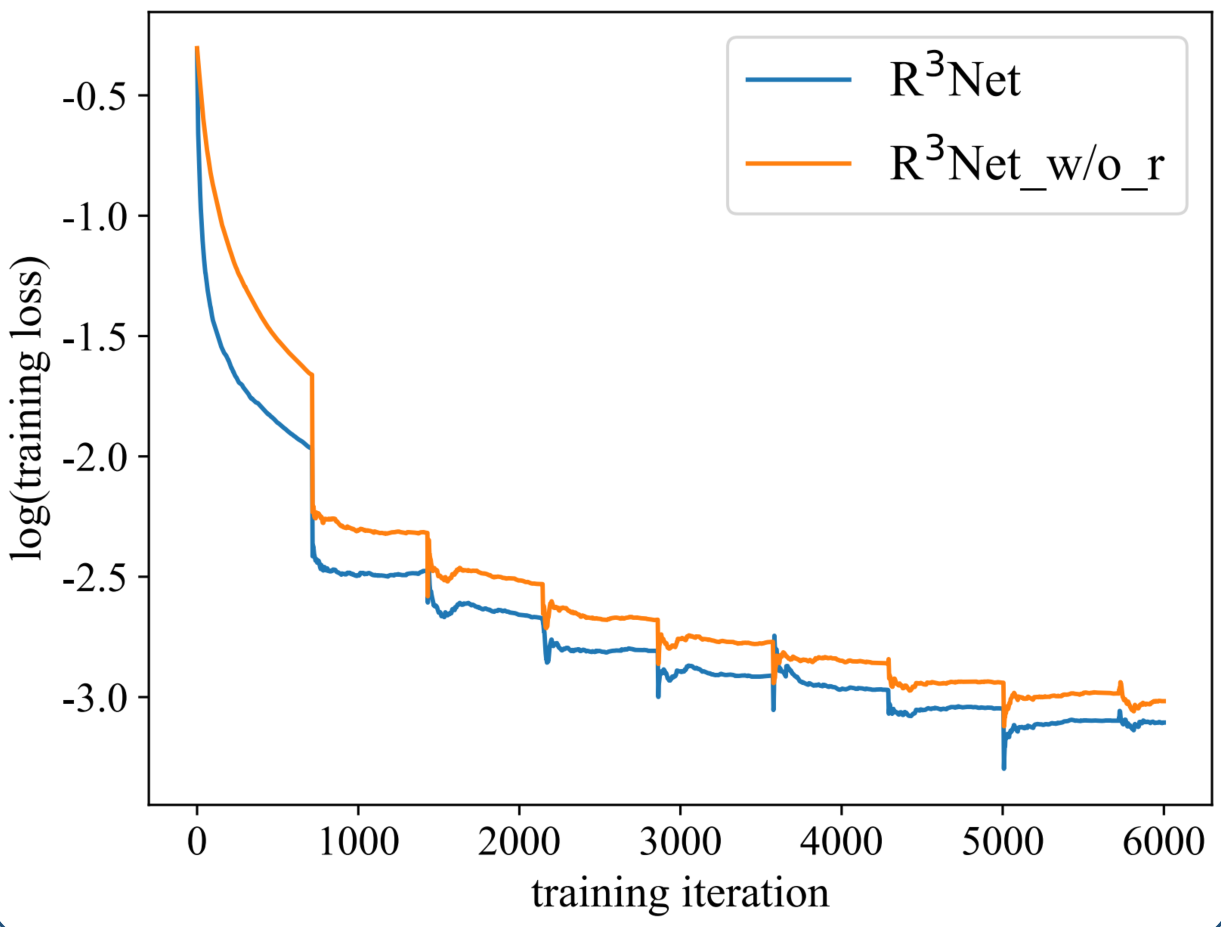
➢ Develop a recurrent residual refinement network (R^3Net) to progressively refine the saliency maps by building a sequence of RRBs, which alternatively use the low-level features and high-level features.

➢ Achieve the best performance on all the five famous benchmarks when comparing to 16 state-of-the-art saliency detectors.

Residual Refinement Block

$$residual_j = \varphi_j(Cat(S_{j-1}, F))$$

$$S_j = S_{j-1} + residual_j$$

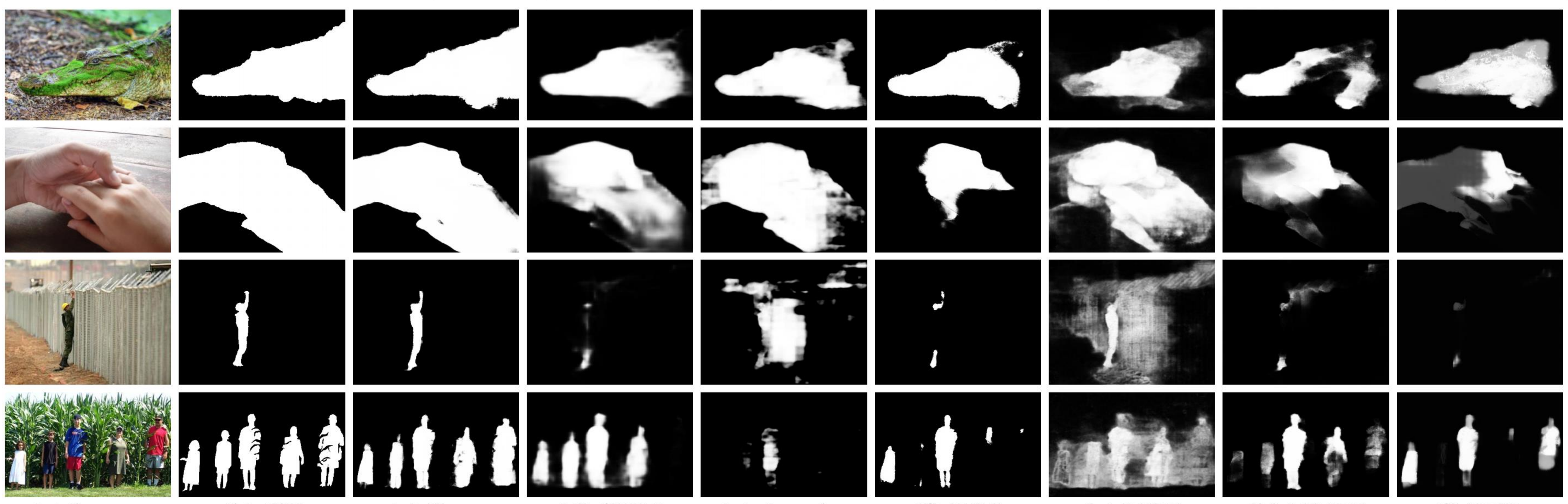


https://github.com/zijundeng/R3Net

Experimental Results

- **Comparison with the State-of-the-arts**

Method	ECSSD		HKU-IS		PASCAL-S		SOD		DUT-OMRON	
	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE
MR [Yang *et al.*, 2013]	0.736	0.189	0.715	0.174	0.666	0.223	0.619	0.273	0.610	0.187
wCtr* [Zhu *et al.*, 2014]	0.716	0.171	0.726	0.141	0.659	0.201	0.632	0.245	0.630	0.144
BSCA [Qin *et al.*, 2015]	0.758	0.183	0.723	0.174	0.666	0.224	0.634	0.266	0.616	0.191
MDF [Li and Yu, 2015]	0.831	0.108	0.860	0.129	0.759	0.142	0.785	0.155	0.694	0.092
RFCN [Wang *et al.*, 2016]	0.898	0.097	0.895	0.079	0.827	0.118	0.805	0.161	0.747	0.095
DCL [Li and Yu, 2016]	0.898	0.071	0.904	0.049	0.822	0.108	0.832	0.126	0.757	0.080
DHSNet [Liu and Han, 2016]	0.907	0.059	0.892	0.052	0.827	0.096	0.823	0.127	-	-
NLDF [Luo *et al.*, 2017]	0.905	0.063	0.902	0.048	0.831	0.099	0.810	0.143	0.753	0.080
UCF [Zhang *et al.*, 2017b]	0.910	0.078	0.886	0.073	0.821	0.120	0.800	0.164	0.735	0.131
DSS [Hou *et al.*, 2017]	0.916	0.053	0.911	0.040	0.829	0.102	0.842	0.118	0.771	0.066
Almulet [Zhang *et al.*, 2017a]	0.913	0.059	0.887	0.053	0.828	0.095	0.801	0.146	0.737	0.083
SRM [Wang *et al.*, 2017]	0.917	0.056	0.906	0.046	0.844	0.087	0.843	0.126	0.769	0.069
R^3Net (ours)	0.935	0.040	0.916	0.036	0.845	0.100	0.847	0.124	0.805	0.063



(a) Inputs (b) GT (c) Ours (d) SRM (e) Amulet (f) DSS (g) UCF (h) NLDF (i) DCL

- **Ablation Analysis**

Method	ECSSD		HKU-IS		PASCAL-S		SOD		DUT-OMRON	
	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE	F_β	MAE
R^3Net-0	0.918	0.049	0.900	0.044	0.831	0.101	0.816	0.128	0.769	0.079
R^3Net-1	0.926	0.044	0.910	0.038	0.841	0.100	0.833	0.125	0.783	0.073
R^3Net-2	0.931	0.043	0.911	0.038	0.844	0.104	0.836	0.127	0.787	0.073
R^3Net-3	0.934	0.041	0.915	**0.036**	0.847	0.100	0.841	0.123	0.794	0.066
R^3Net-4	0.932	0.042	0.912	0.038	0.843	0.102	0.841	0.125	0.782	0.073
R^3Net-5	0.933	0.042	0.913	0.037	0.845	0.100	0.841	0.122	0.791	0.069
R^3Net-6	**0.935**	**0.040**	**0.916**	**0.036**	0.845	0.100	**0.847**	0.124	**0.805**	**0.063**
R^3Net-7	0.934	**0.040**	0.914	**0.036**	**0.848**	**0.096**	0.842	**0.121**	0.804	**0.063**
R^3Net_w/o_r	0.931	0.042	0.910	0.039	0.839	0.103	0.839	**0.121**	0.782	0.077
R^3Net_w_s	0.933	0.041	0.914	0.037	0.841	0.102	0.842	0.122	0.794	0.070
R^3Net_LL	0.932	0.041	0.910	0.038	0.844	0.100	0.839	0.125	0.778	0.080
R^3Net_HH	0.926	0.046	0.902	0.042	0.836	0.101	0.819	0.128	0.786	0.071
R^3Net-D	0.928	0.046	0.907	0.042	0.829	0.112	**0.847**	0.127	0.793	0.067
R^3Net-V	0.913	0.049	0.891	0.047	0.814	0.105	0.818	**0.121**	0.746	0.089

e-mail: zijdeng@gmail.com