



Nice, France

October 21-25 2019

P4C-06

Instance of Interest Detection

Fan Yu^{1,3}, Haonan Wang¹, Tongwei Ren^{1,3}, Jinhui Tang², Gangshan Wu¹

¹ State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

² School of Computer Science, Nanjing University of Science and Technology, Nanjing, China

³ Shenzhen Research Institute of Nanjing University, Shenzhen, China



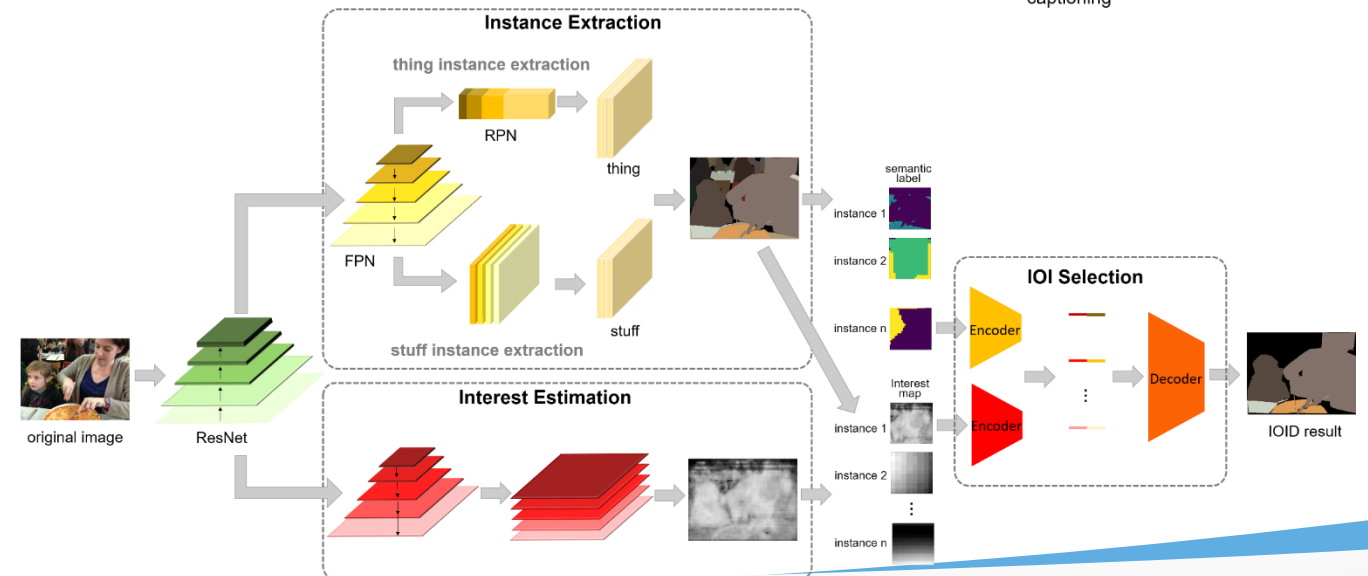
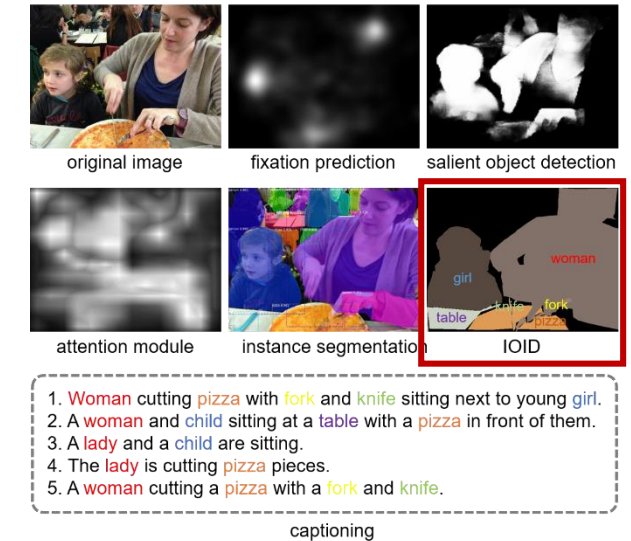
MAGUS

Media recoGnition
and UnderStanding

Motivation and Solution



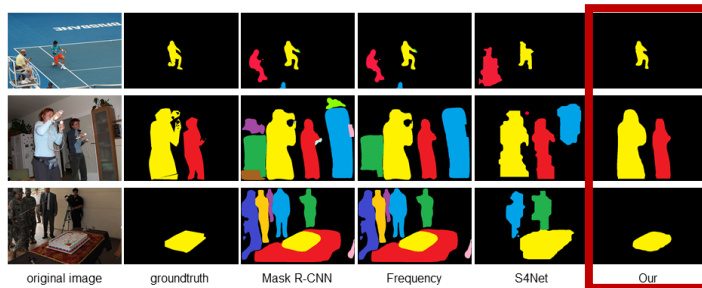
- **Instance of interest detection (I OID)** aims to provide instance-level user interest model for image semantic description
 - **Instance of Interest (IOI)**: the instances which are beneficial to represent image content
- Our solution: **Cross-Influential Network (CIN)**
 - **Instance Extraction**: containing a thing extraction branch and a stuff extraction branch
 - **Interest Estimation**: estimate pixel-interest according to feature maps
 - **IOI Selection**: select IOIs with a Cross-influential Encoder-decoder Network



Experiments



- Datasets
 - Construct a dataset based on MSCOCO dataset
 - Training set contains 36,000 images with 165,094 IOIs, and test set contains 9,000 images with 40,617 IOIs
- Evaluation criteria: precision, recall, F , recall*, F^*
- Comparison



Qualitative examples

Method	precision	recall	F	recall*	F^*
Thing [14]	87.06	9.66	30.56	26.00	56.47
Stuff [4]	19.91	2.59	7.82	15.04	18.52
Our	68.47	30.15	52.95	49.80	63.02

Different instance extraction

Method	precision	recall	F	recall*	F^*
Binary	40.93	35.71	39.59	58.98	44.04
RNN	46.57	49.10	47.13	81.12	51.64
Our	68.47	30.15	52.95	49.80	63.02

Different interest estimation

Method	precision	recall	F	recall*	F^*
DSS [32]	68.78	15.24	37.99	25.18	49.14
MSRNet [11]	63.87	29.92	50.62	49.42	59.83
NLDF [31]	67.33	23.18	46.77	38.28	57.30
PiCANet [30]	67.63	24.36	47.97	40.24	58.45
SalGAN [19]	60.31	23.66	44.43	39.09	53.59
SAT [37]	52.09	30.73	44.89	50.76	51.78
Our	68.47	30.15	52.95	49.80	63.02

Different IOI selection



Dataset construction tool

Method	precision	recall	F	recall*	F^*
Mask R-CNN [14]	41.48	37.14	40.39	100.00	47.95
Frequency	50.36	32.76	44.81	88.19	55.90
S4Net [9]	40.70	18.63	31.96	100.00	47.16
Our	68.47	30.15	52.95	49.80	63.02

Overall