**P4C-06** 



# **Instance of Interest Detection**

Fan Yu, Haonan Wang, Tongwei Ren, Jinhui Tang, Gangshan Wu

# Introduction

Instance of Interest (IOI) are the instances which are beneficial to represent image content, including things and stuff.

**Instance of interest detection (IOID)** aims to provide instance-level user interest model for image semantic description.

**Contribution:** we **firstly** define the IOID task, propose a novel IOID method, and construct an IOID dataset.











attention module instance segmentation

## **Method**

We propose a Cross-Influential Network (CIN) to handle the subtasks in IOID, including:

- Instance Extraction: containing a thing extraction branch and a stuff extraction branch.
- **Interest Estimation:** estimating pixelinterest according to feature maps.
- IOI Selection: select IOIs with a Cross-influential Encoder-decoder Network.

ResNet101 acts as the backbone and the feature maps are shared in the three key modules.



### **Experiments**

#### Dataset

We construct an **IOID dataset** based on MSCOCO dataset, in which the training set and the test set contain 36,000 images with 165,094 IOIs and 9,000 images with 40,617 IOIs, respectively.

#### **Evaluation criteria**

Precision, recall, F, recall\*, F\*

FN\* for recall\* and F\*: missed IOIs in projected groundtruth and mismatched IOIs in groundtruth.

#### **Parameter setting**

| threshold               | 0.30         | 0.35  | 0.40  | 0.45  | 0.50  | 0.55  |
|-------------------------|--------------|-------|-------|-------|-------|-------|
| precision               | 41.76        | 53.99 | 66.02 | 75.66 | 82.17 | 86.55 |
| $\operatorname{recall}$ | <b>49.98</b> | 41.67 | 33.25 | 24.65 | 16.78 | 10.07 |
| F                       | 43.41        | 50.54 | 53.78 | 51.21 | 43.26 | 31.45 |
|                         |              |       |       |       |       |       |

#### **Qualitative results**



#### **Component analysis**

| Method       | precision | recall       | F     | $\operatorname{recall}^*$ | $F^*$ |
|--------------|-----------|--------------|-------|---------------------------|-------|
| 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 | <b>49.80</b>              | 63.02 |
| Method       | precision | recall       | F     | $\operatorname{recall}^*$ | $F^*$ |
| 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 |
| Method       | precision | recall       | F     | $\operatorname{recall}^*$ | $F^*$ |
| Binary       | 40.93     | 35.71        | 39.59 | 58.98                     | 44.04 |
| RNN          | 46.57     | <b>49.10</b> | 47.13 | 81.12                     | 51.64 |
| Our          | 68.47     | 30.15        | 52.95 | 49.80                     | 63.02 |

#### **Comparison with the state-of-the-arts**

| Method          | precision | recall | F     | $\operatorname{recall}^*$ | $F^*$ |
|-----------------|-----------|--------|-------|---------------------------|-------|
| 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 |

#### Fan Yu: yf@smail.nju.edu.cn

#### Tongwei Ren: rentw@nju.edu.cn