



Insights of Object Proposal Evaluation

Yuantian Wang, Lei Huang, Tongwei Ren, Han Gu

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

Introduction

The existing **object proposal evaluation criteria** based on **recall** cannot evaluate the real **objectness measurement ability** of different object proposal methods

- Ignore the **position** and **size** of objects which may influence recall value
- Cannot judge whether a method is **worse than random sampling** if it obtains acceptable recall

Method

Basic Idea

Define objectness measurement ability (OMA) based on the probability to hit an object by non-repetitive random sampling (HPRS), and extend the commonly used object proposal evaluation criteria by replacing recall with OMA

Contribution

- We analyze and calculate **HPRS** for OMA definition
- We propose new **OMA based criteria** in object proposal evaluation
- We validate our proposed criteria on PASCAL VOC 2007, which is **superior** to current criteria in evaluating different object proposal methods

HPRS Calculation

$$HPRS(o, k) = 1 - \frac{C_{N_{tol} - N_{hit}}^k}{C_{N_{tol}}^k}$$

$HPRS(o, k)$: hit probability of object o with k randomly sampled candidates

N_{tol} : number of possible candidates in the image

N_{hit} : number of hit candidates to object o

OMA Calculation

$$OMA = \frac{1}{N_{img}} \sum_{i=1}^{N_{img}} \frac{1}{|O_i|} \left(|H_i| - \sum_{j=1}^{|O_i|} HPRS(o_j^i, k) \right)$$

N_{img} : number of images in dataset

H_i : the set of hit objects on the i th image

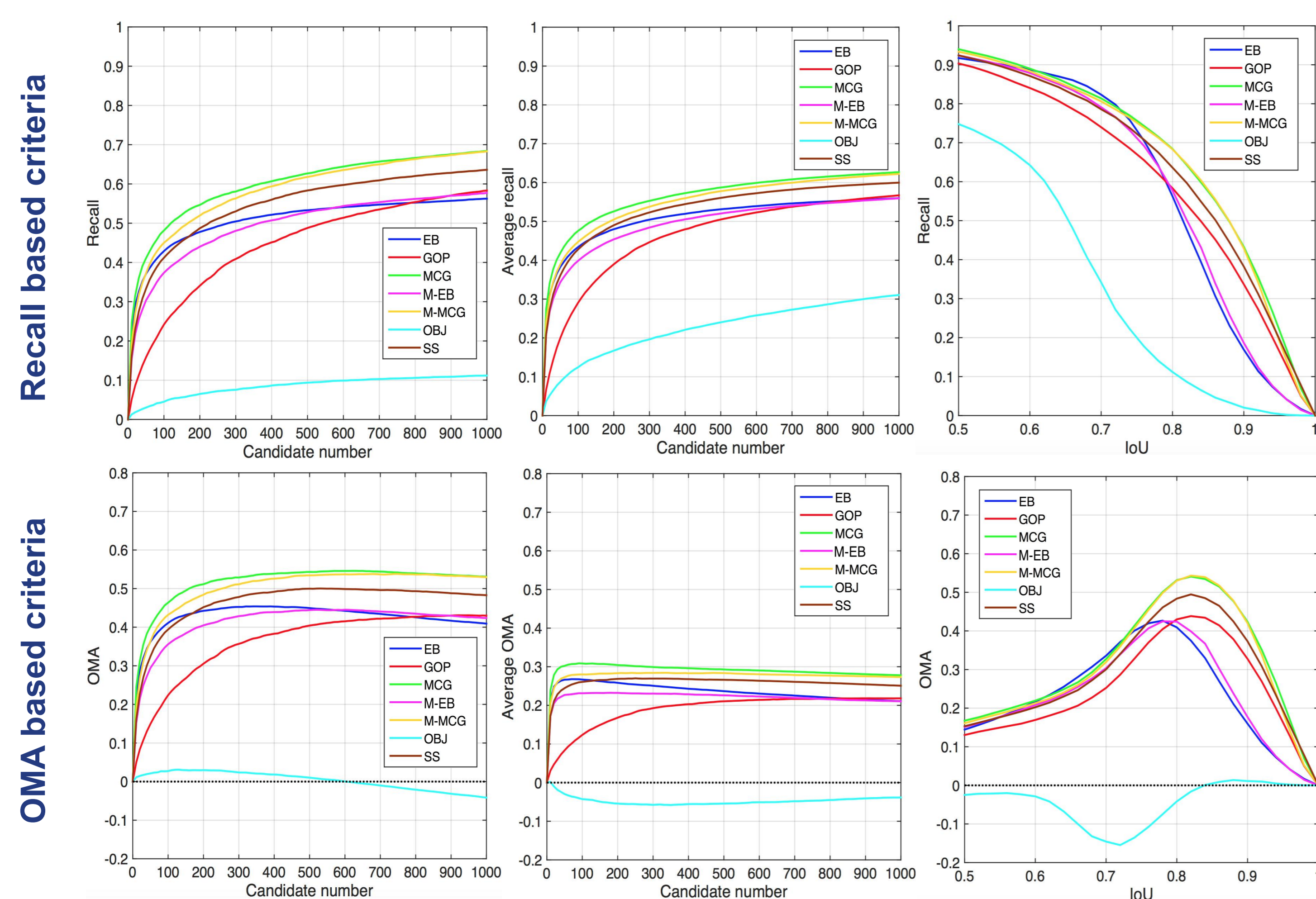
O_i : the set of all objects on the i th image

o_j^i : the j th object in O_i

k : the number of candidates on each image

Experiments

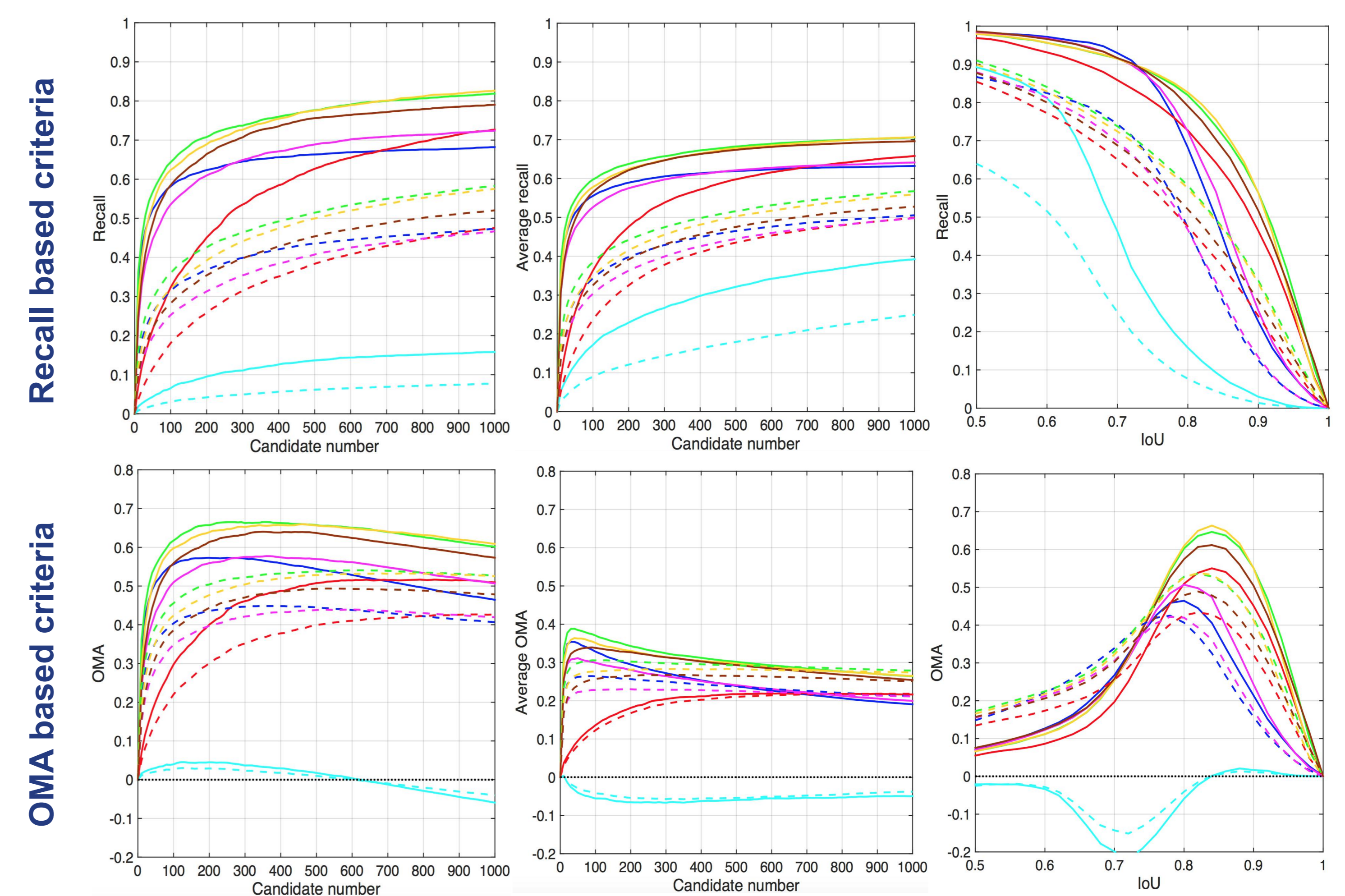
Experiment A: we evaluate seven object proposal methods using current recall based criteria and our OMA based criteria on VOC 2007



Result

- Most methods perform **more stable** in OMA
- OBJ slightly **underperforms random sampling**
- Most methods obtain the best performance around **IoU = 0.8**

Experiment B: we decompose VOC 2007 into VOC 2007-few and VOC 2007-multiple according to the object number in each image



Result

- The **distance** between the evaluation results of the same method on two datasets are **smaller** under the OMA based criteria.

