1802

ICIP 2021 IEEE International Conference on Image Processing

IEEE Signal Processing Society

DATA MINING LAB

Georgia State University

Multiscale IoU

A Metric for Evaluation of Salient Object Detection with Fine Structures

Azim Ahmadzadeh, Dustin J. Kempton, Yang Chen, and Rafal A. Angryk
Department of Computer Science, Georgia State University, USA

Introduction

◉ Computer Vision has shown unprecedented success in the object detection and segmentation task, thanks to several exuberant and popular competitions (e.g., PASCAL VOC, ILSVRC, COCO, RVC, etc.).

◉ What they have in common: (1) targets are "everyday" objects, (2) objects are annotated coarsely, (3) area-based metrics are used for evaluation.

The by product of such settings is that, ...

… often a pixel-level precise detection is not of high priority!

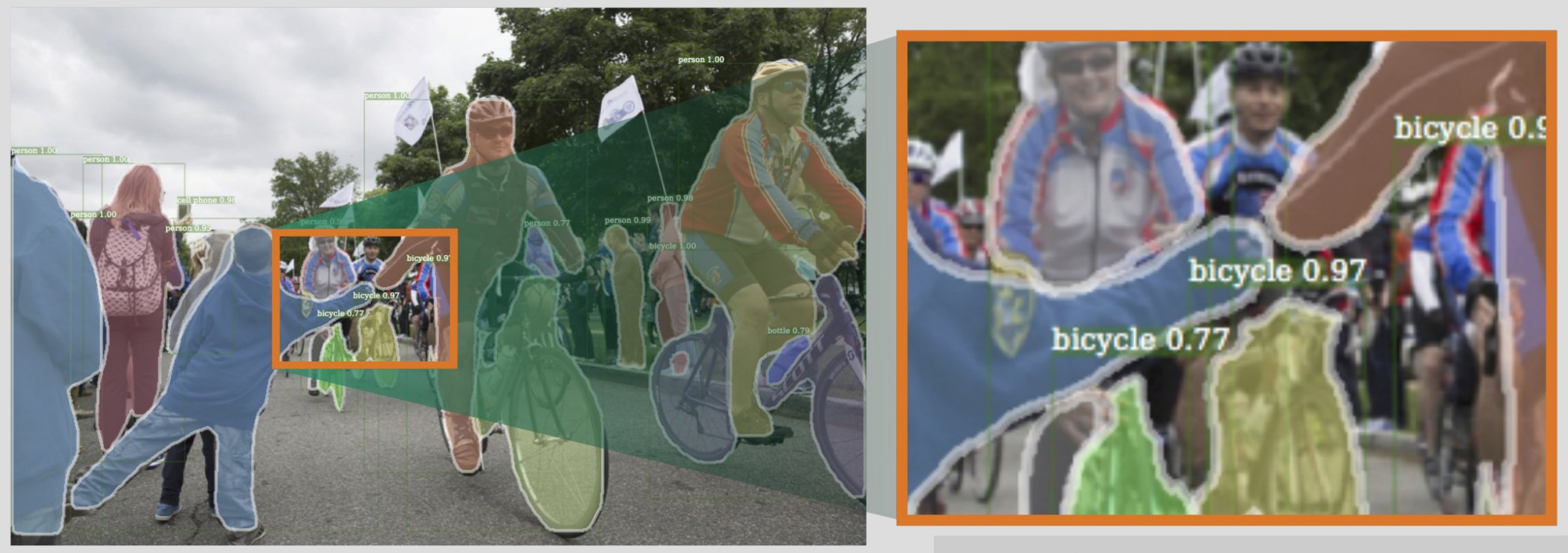
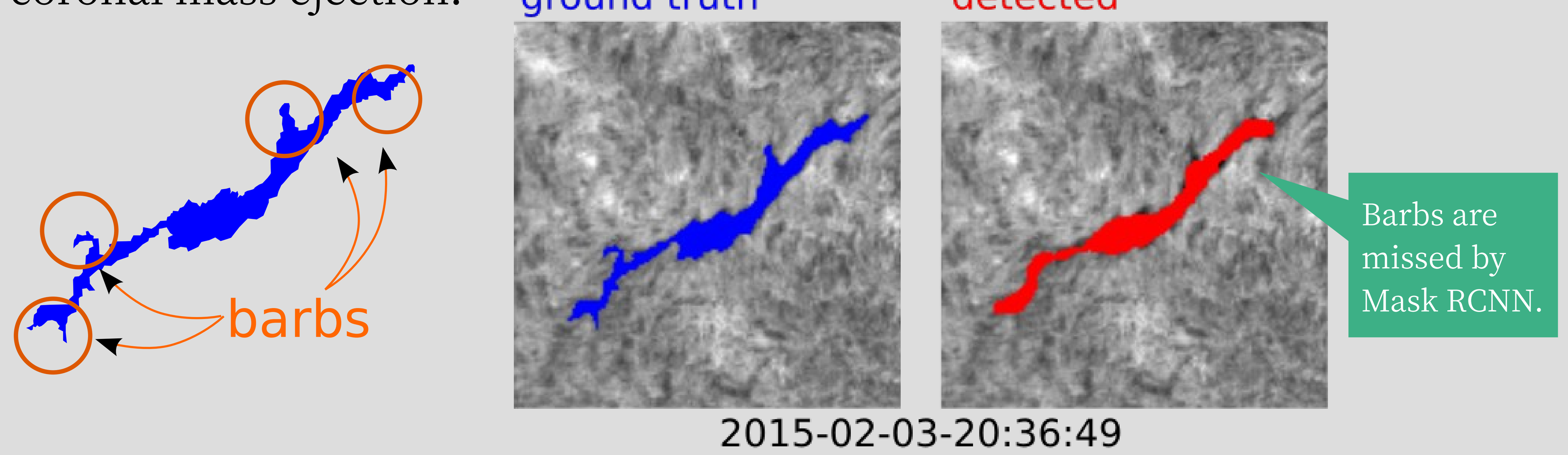


Image Credit: Detectron (https://github.com/facebookresearch/detectron)

So, the competing algorithms ignore this important objective.

Fine Structures Matter!

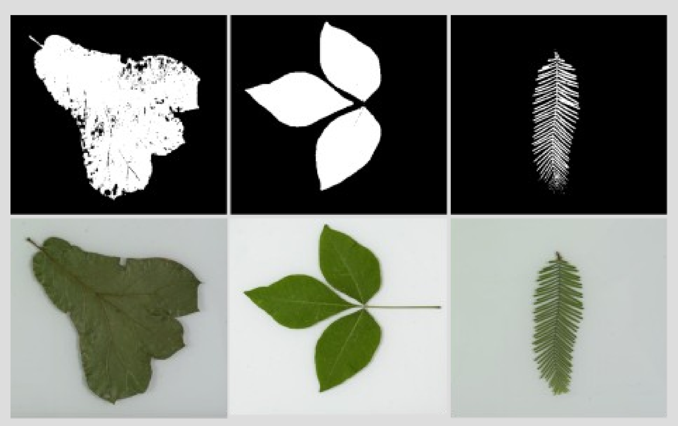
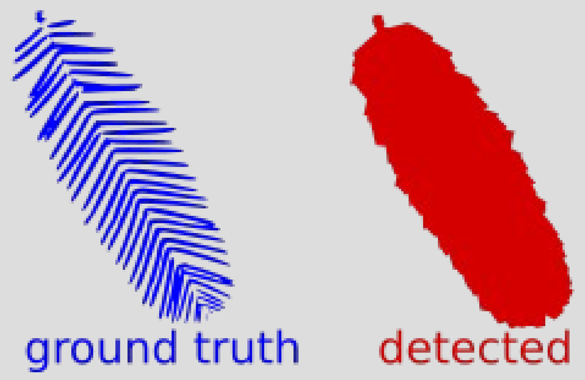
◉ In Heliophysics, the spatial information of solar filaments can be used to determine the magnetic field orientation in a potentially associated coronal mass ejection.



Filament observations from BBSO observatory (https://www.bbso.njit.edu/)

The barbs are entirely missed in the detected (red) region. Such detections won't be useful for classification of filaments' chirality which is determined by the angle of these barbs.

◉ In Botany, to build a content-based image retrieval system, one of the key features needed is the leaves' shape.

Data: Leafsnap (http://leafsnap.com/dataset/)

See how the fine structure of the leaf sample (blue) is largely disregarded in the detected region (red).

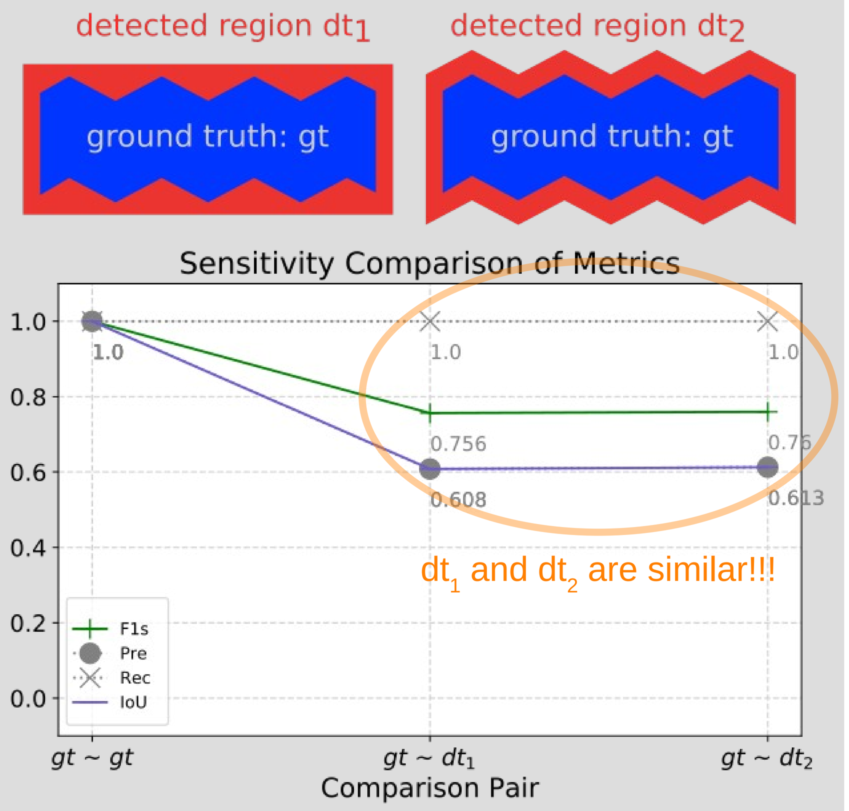
Problem Description

◉ Given a ground-truth region (gt) and multiple detected regions (dt_i) area-based metrics fail to capture the fine details, as long as dt_is' overlaps with gt are roughly the same.

$$\forall i \quad gt \cap dt_i \approx c \quad \text{where } c \in \Re$$

The simplest form of this issue is illustrated on the right.

The scores corresponding to $gt\sim dt_1$ and $gt\sim dt_2$ are similar, despite the significant difference between dt_1 and dt_2.



Proposed Solution (MIoU)

◉ Multiscale IoU is the marriage of two concepts:

MIoU : IoU + fractal dimension

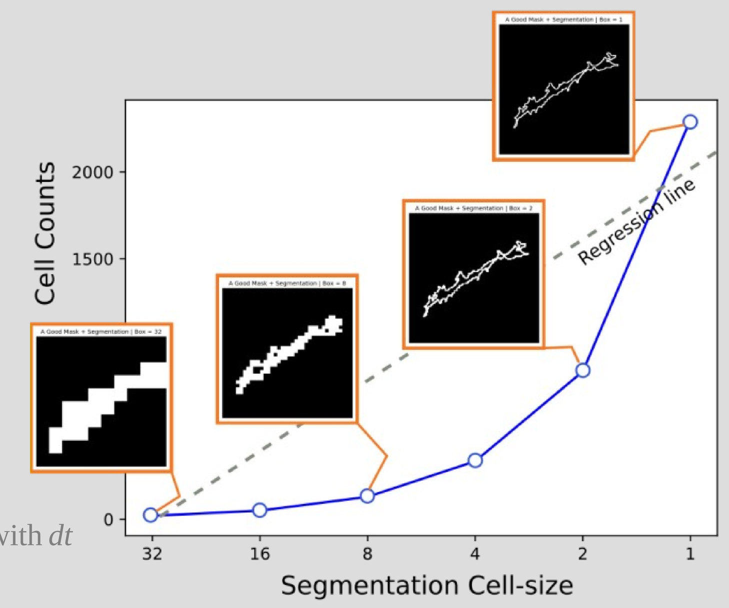
◉ **Intersection over Union** (IoU) is a very popular area-based metric that quantifies the degree of which gt and dt intersect relative to the area occupied by their union.

$$IoU = \frac{gt \cap dt}{gt \cup dt}$$

◉ **Fractal Dimension** is a concept from Fractal Geometry that is defined to quantify the complexity of fractals. **Box counting** is a popular method for measuring the fractal dimension of shapes.

$$D_{box} = \lim_{\delta \to 0} \frac{\log(n(dt, \delta))}{\log(1/\delta)}$$

δ: cell size of a grid
n: number of cells of size δ, which overlap with dt

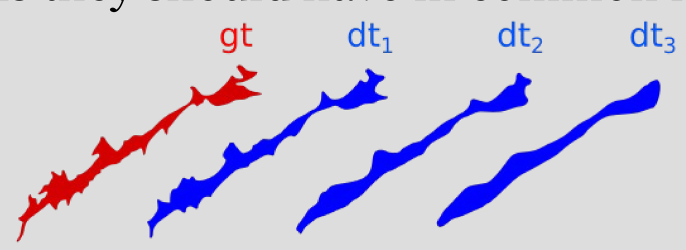


◉ MIoU fuses the multi-resolution concept of fractal dimension with IoU.

1. **Intersection Ratio (r)** is the ratio of the number of cells gt and dt have in common, over the number of cells they should have in common if they perfectly align.

$$r(gt, dt, \delta_i) = \frac{n(s(gt,\delta)) \cap n(s(dt,\delta))}{n(s(gt,\delta))}$$

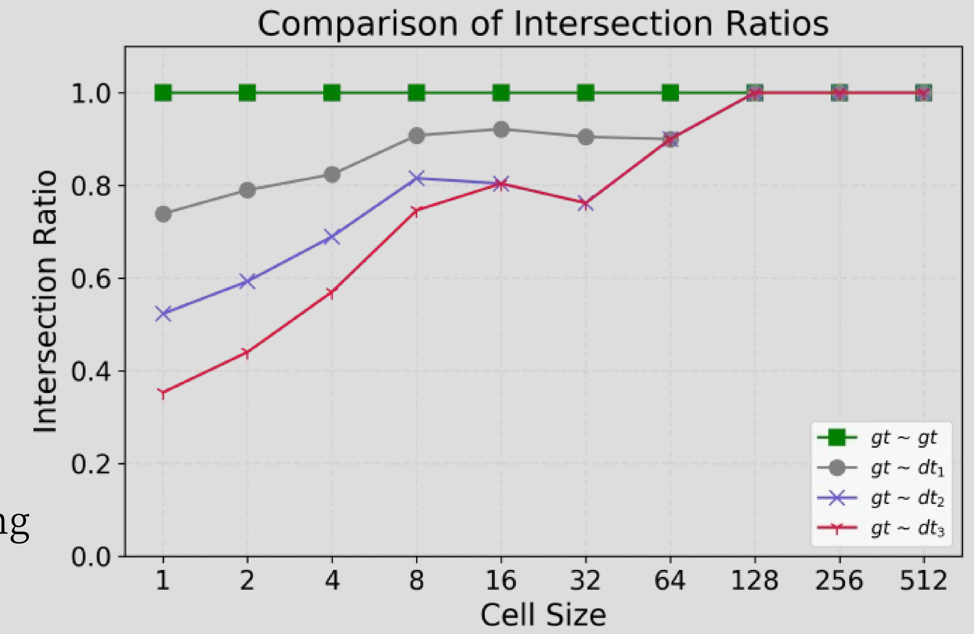
s: segmentation function
$\delta_i \in \Delta$ (set of all cell sizes)



2. **MIoU** is the area under the curve of $r(gt, dt, \delta_i)$ for all $\delta_i \in \Delta$.

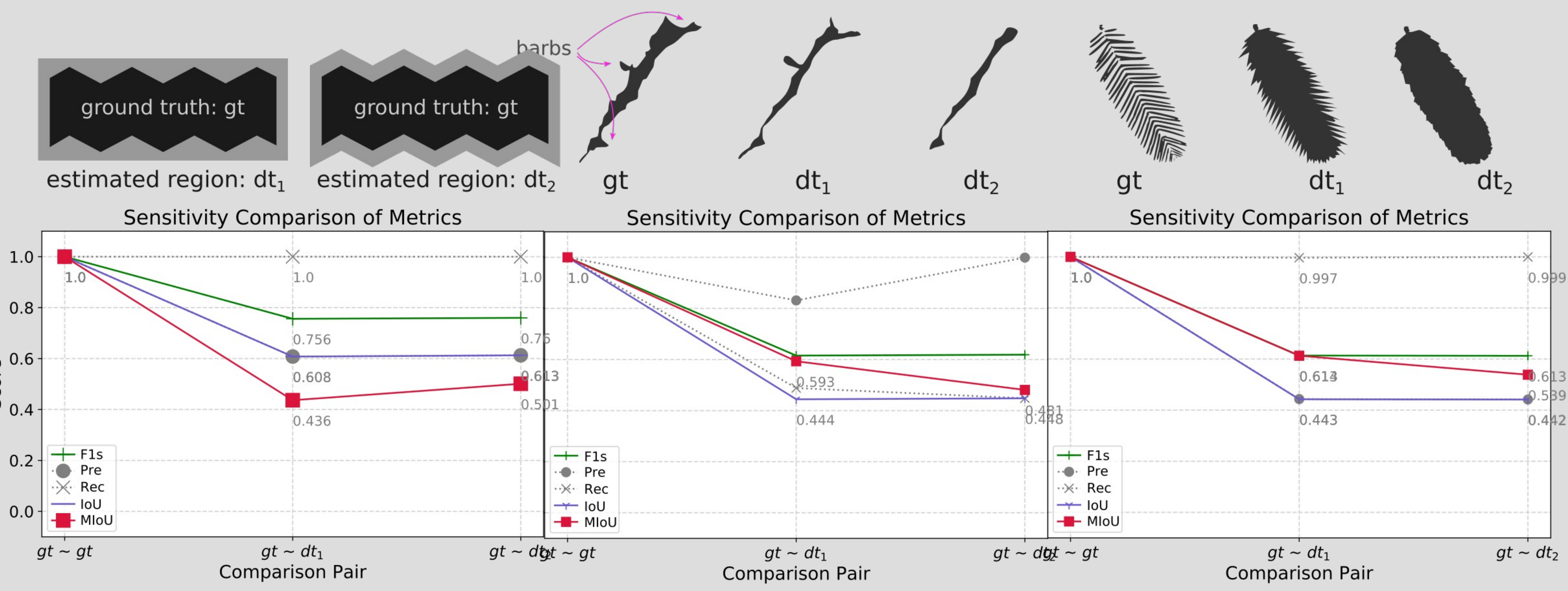
$$MIoU(gt, dt) = \int_0^1 r(gt, dt, \delta)\, d\delta$$

By choosing $d\delta = \frac{1}{|\Delta| - 1}$ and transforming δ to the range [0,1], then MIoU∈[0,1].



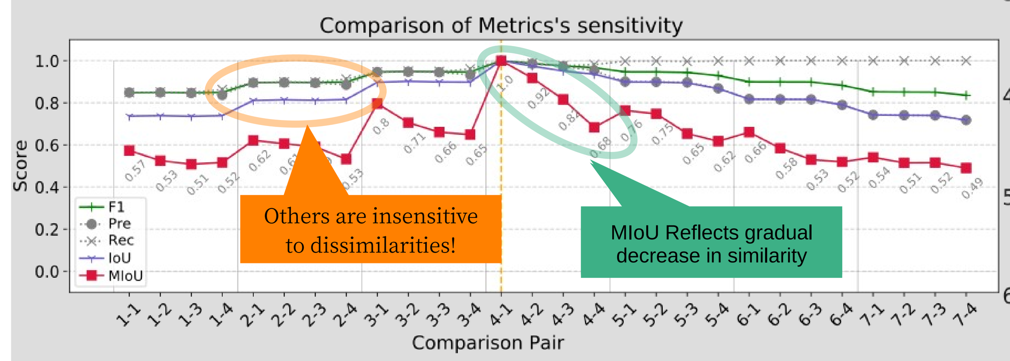
Experiments

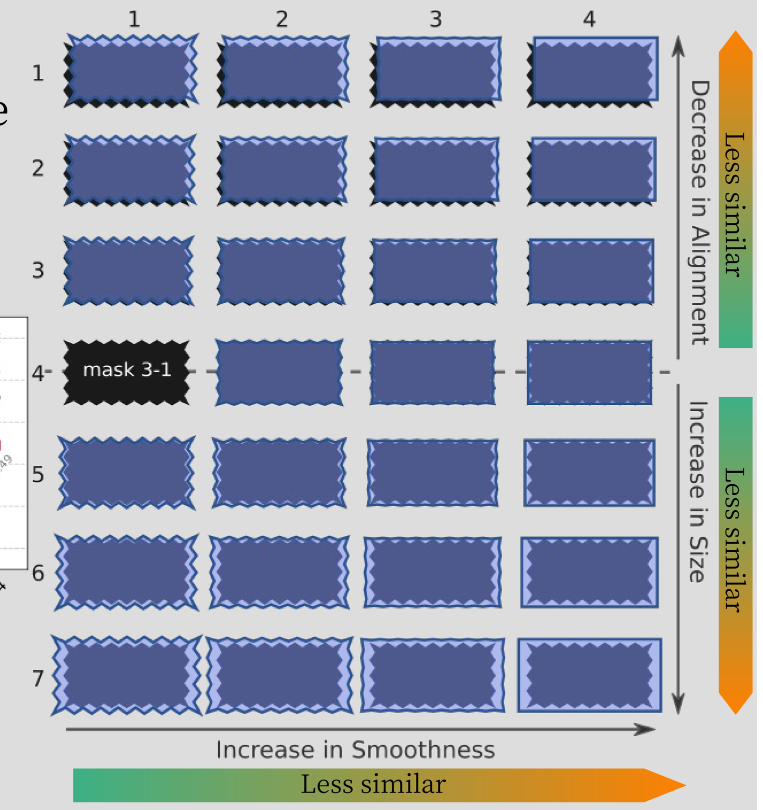
A On Evidential Cases: MIoU sees what others may miss!



B On Synthetic Regions:

To control for the confounding factors, we generated 28 proposed regions for one ground-truth region (mask 3-1).

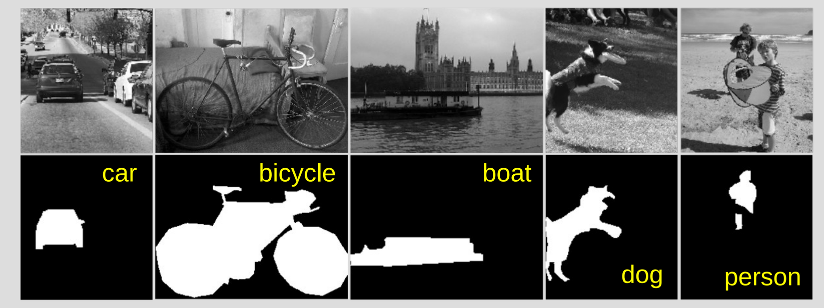


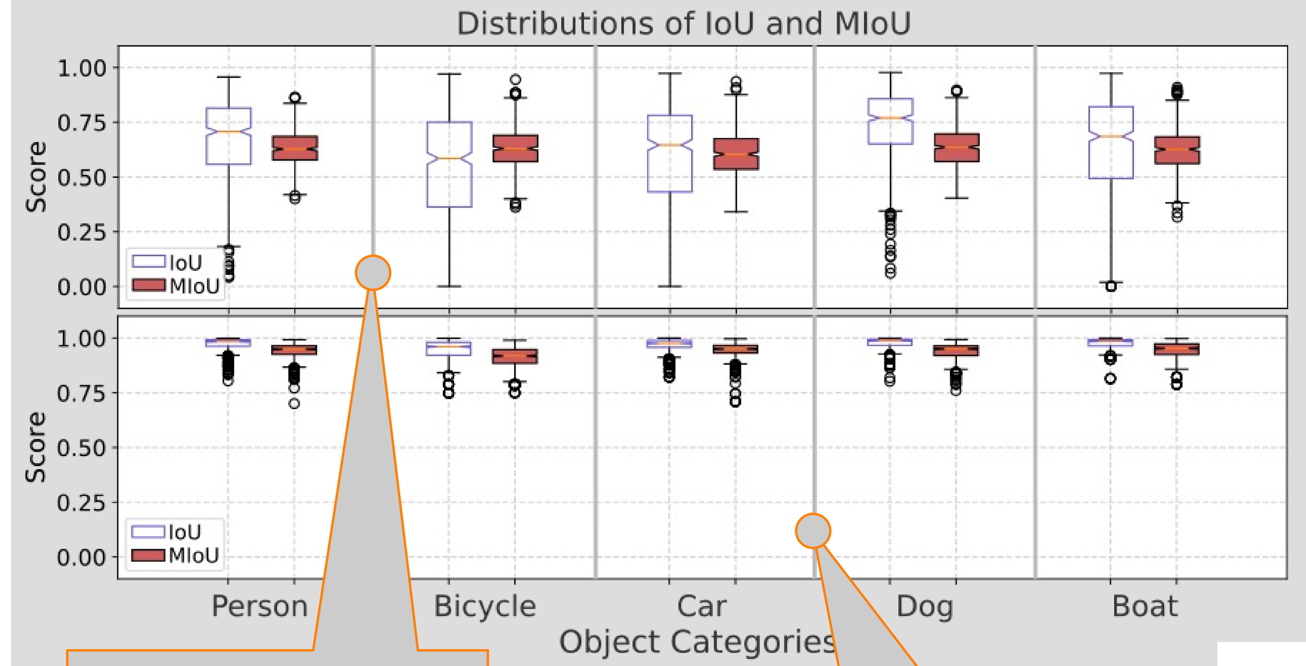


C On Real Regions:

To test the reliability of MIoU as an alternative to IoU, we compared the two metrics on COCO data set.

$\Delta = [2^n, n < 10]$





With random rotation (±10 deg) and translations (±10 px).

With Gaussian smoothing followed by thresholding.

MIoU: is similar to IoU, but with smaller variance; it captures the structural Similarities despite small rotation and translation.

MIoU: is similar to IoU, but with lower median; it captures the loss of fine structures after smoothing.

MIoU seems to be a good alternative to IoU, specially for segmentation of scientific images! Perhaps more challenging test cases shed more light on this proposal.

Link to code