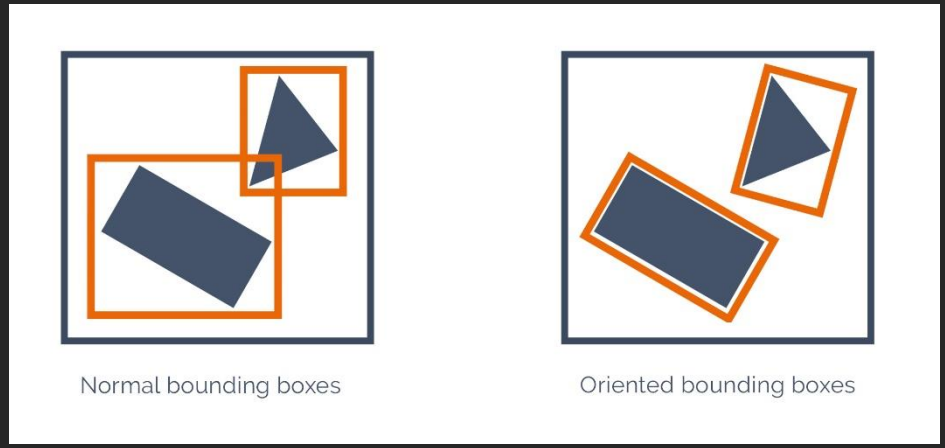
Can we trust bounding box annotations for object detection? (2022)

Problem statement

Both HBB and OBB annotations are not objective because:

- Quality decreases when the annotators are tired.
- BBs only provide the view of one human annotator.
- Occluded object annotation is not consistent.

Discrepancy in annotating might significantly impact the IOU and mAP, hence the detector's performance reliability.



(Left) Horizontal bounding box (HBB) and
(Right) Oriented bounding box (OBB)

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Method

FYI

Bounding box (BB) is the smallest rectangle that contains the object.

According to Nguyen et al., available annotations (including from MS COCO, VOC2012, …) are only approximations of the actual GT.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Nguyen, T. T. D., Rezatofighi, H., Vo, B. N., Vo, B. T., Savarese, S., & Reid, I. (2022). *How trustworthy are performance evaluations for basic vision tasks?*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(7), 8538-8552.

Method

To show that discrepancy in annotating significantly impacts the IOU (hence mAP), the authors emulate the human error by choosing datasets with both HBB (or OBB) annotations and segmentation masks.

- HBB human Annotations → AHBB
- HBB generated from Seg. masks → SHBB

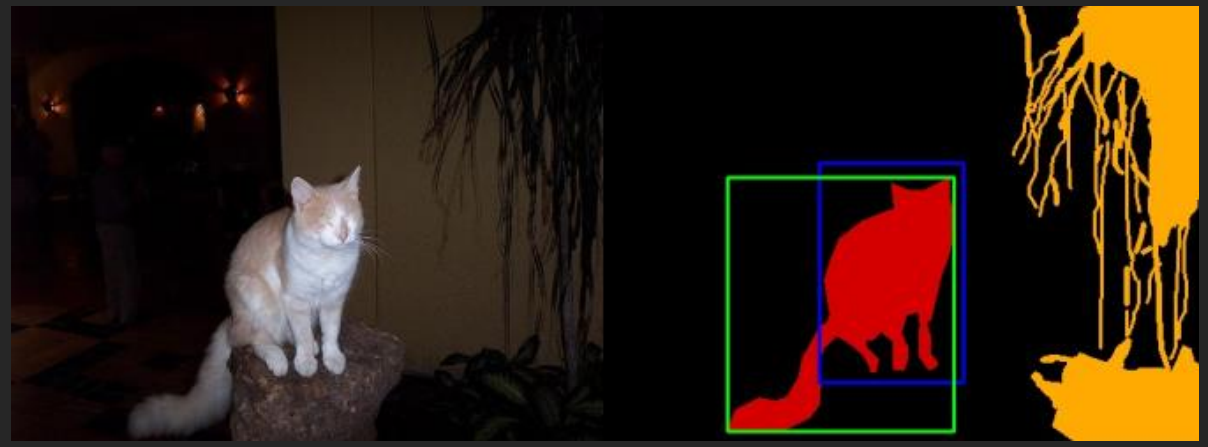
In simpler words

The authors create new human-like HBB annotations as if the dataset had 2 human annotators to evaluate the discrepancy.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Method

IOU=0.38



IOU=0.38

Examples of AHBB (Blue) and SHBB (Green) from VOC2012.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

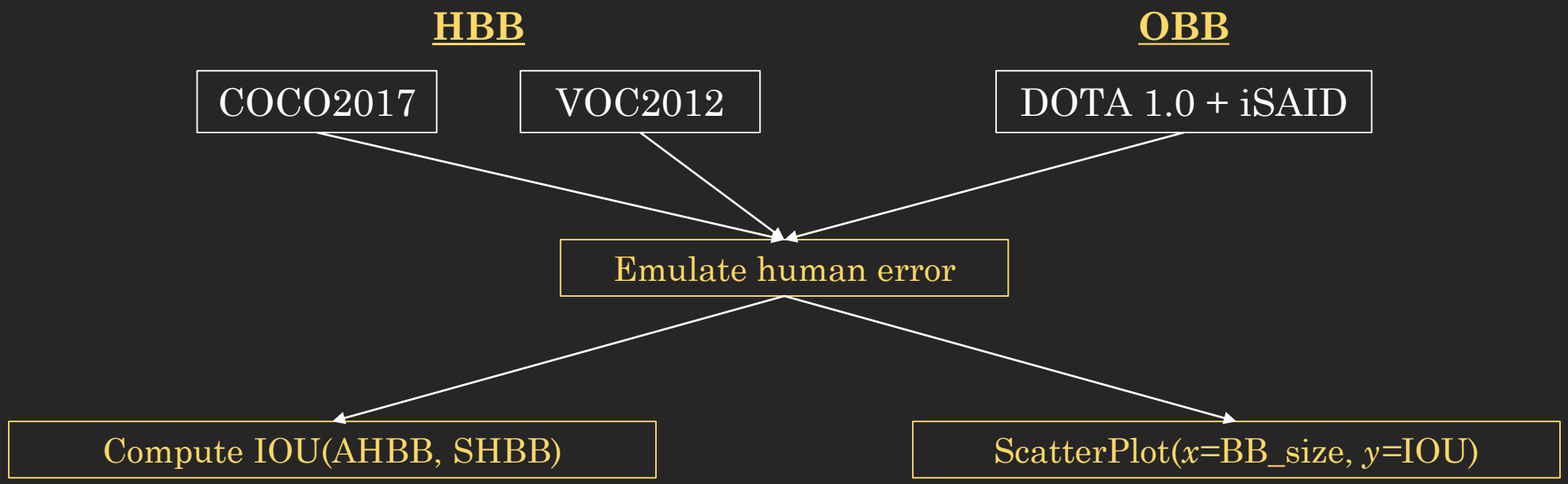
Method

IOU=0.57 IOU=0.61

Examples of AOBB (Red) and SOBB (Blue) from DOTA 1.0.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Method



Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – COCO2017

The authors compare two bounding box representations:

1) Available human annotations (AHBB)

2) Human-emulated annotations (SHBB)

There is only a sub-pixel difference in the two representations.

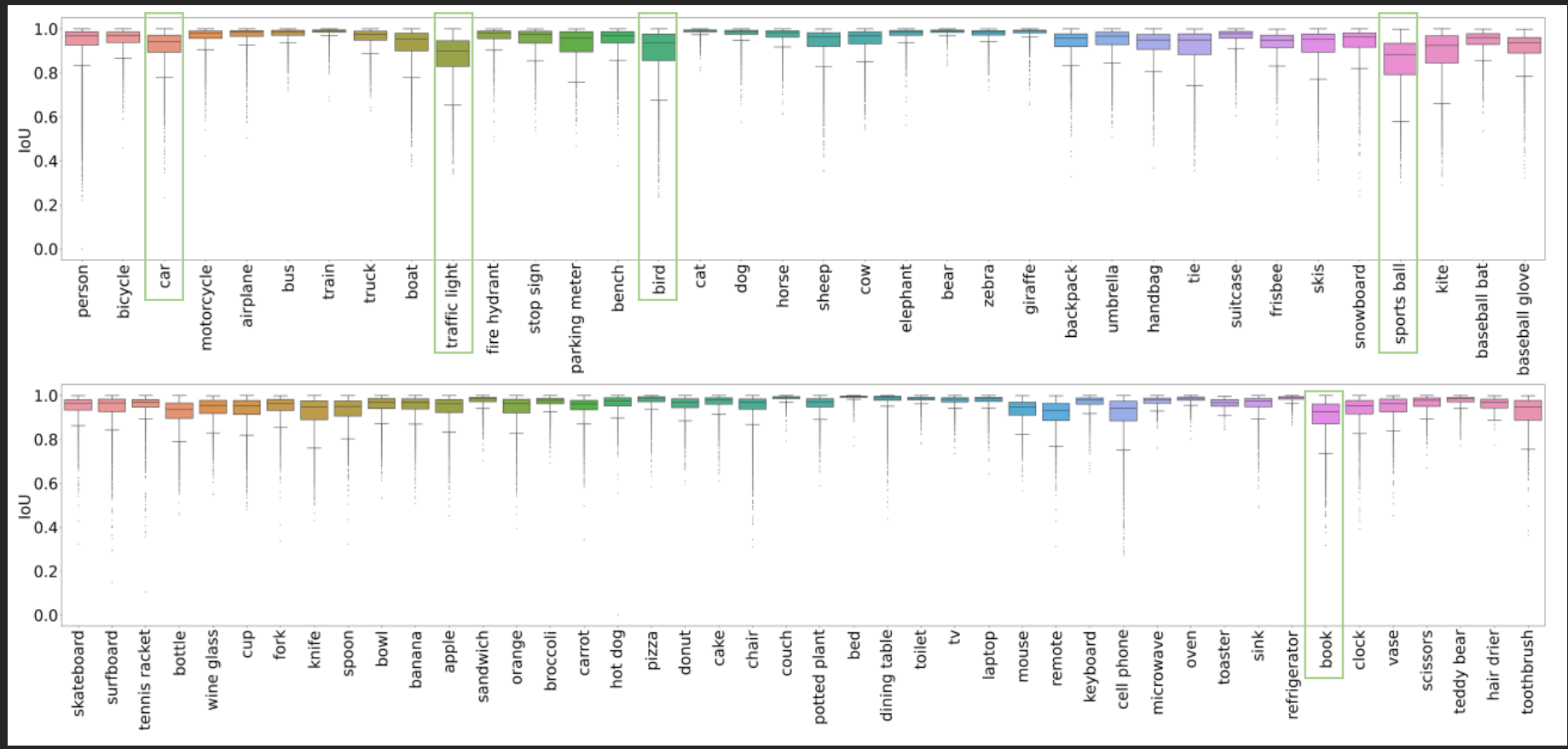
Example of AHBB (Blue) and "emulated" SHBB (Green).

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – COCO

Marked categories suffer from significant degradation even with sub-pixel discrepancy.

Even if the detector outputs perfect result (SHBB), the AP can still be degraded when compared to AHBB, let alone the unknown GT.

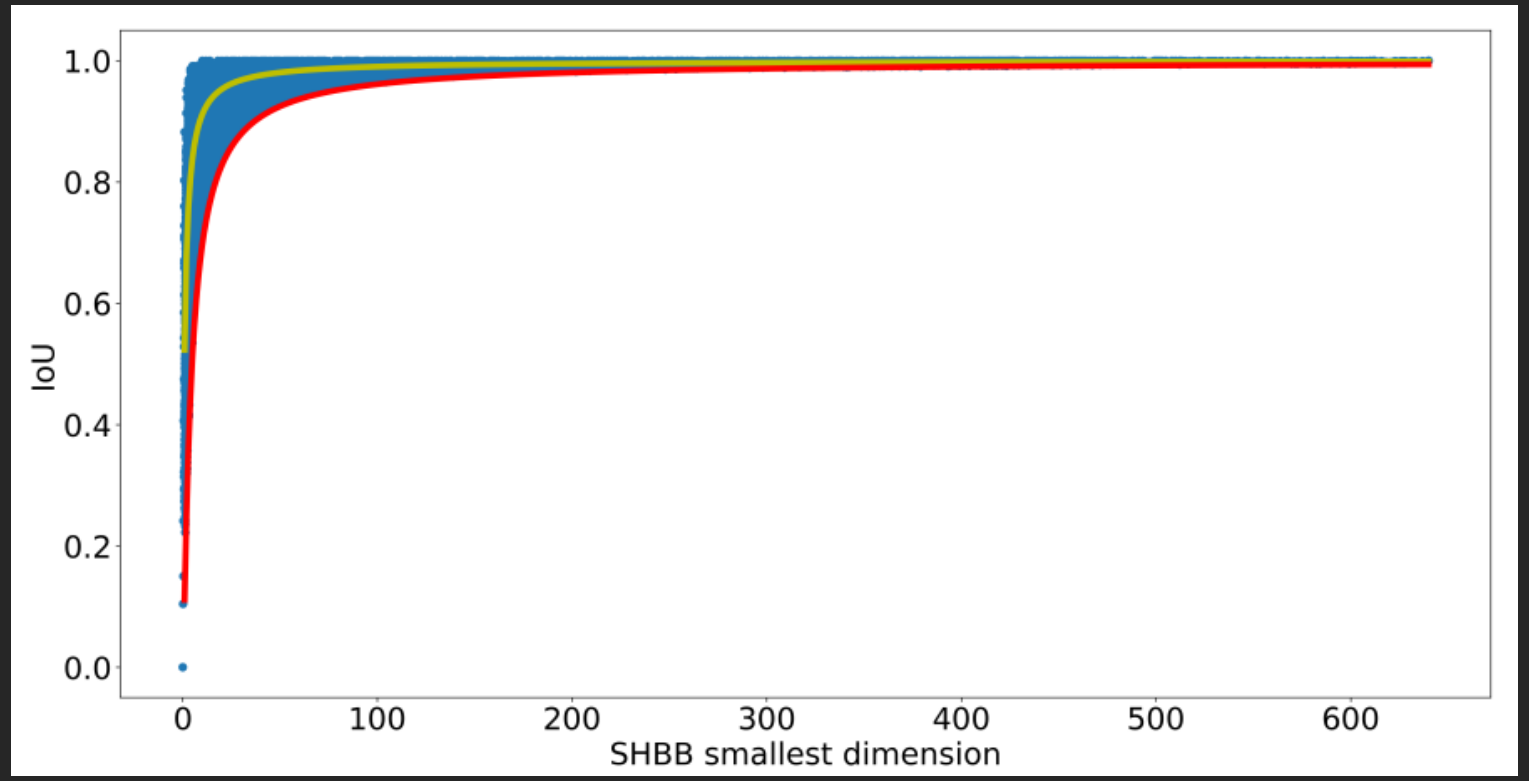


Per-category Iou between AHBBs and HBBs generated from segmentation masks (SHBBs) in COCO.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – COCO

- The effect on IOU is dependent on BB size (especially for small BB).
- Small and medium objects are more susceptible to sub-pixel errors than large objects.



Scatter plot of the smallest dimension - min(W, H) - of SHBB and IOU(AHBB, SHBB).

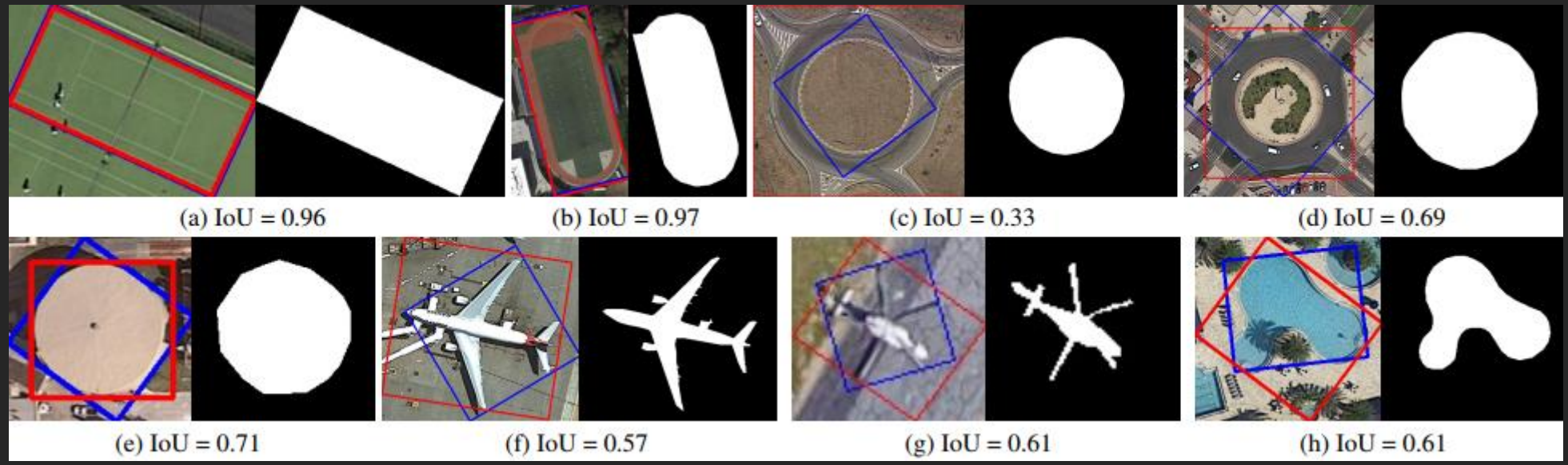
Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – VOC

VOC experiments yields similar results.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – DOTA 1.0



Examples of AOBB (Red) and SOBB (Blue) from DOTA 1.0.

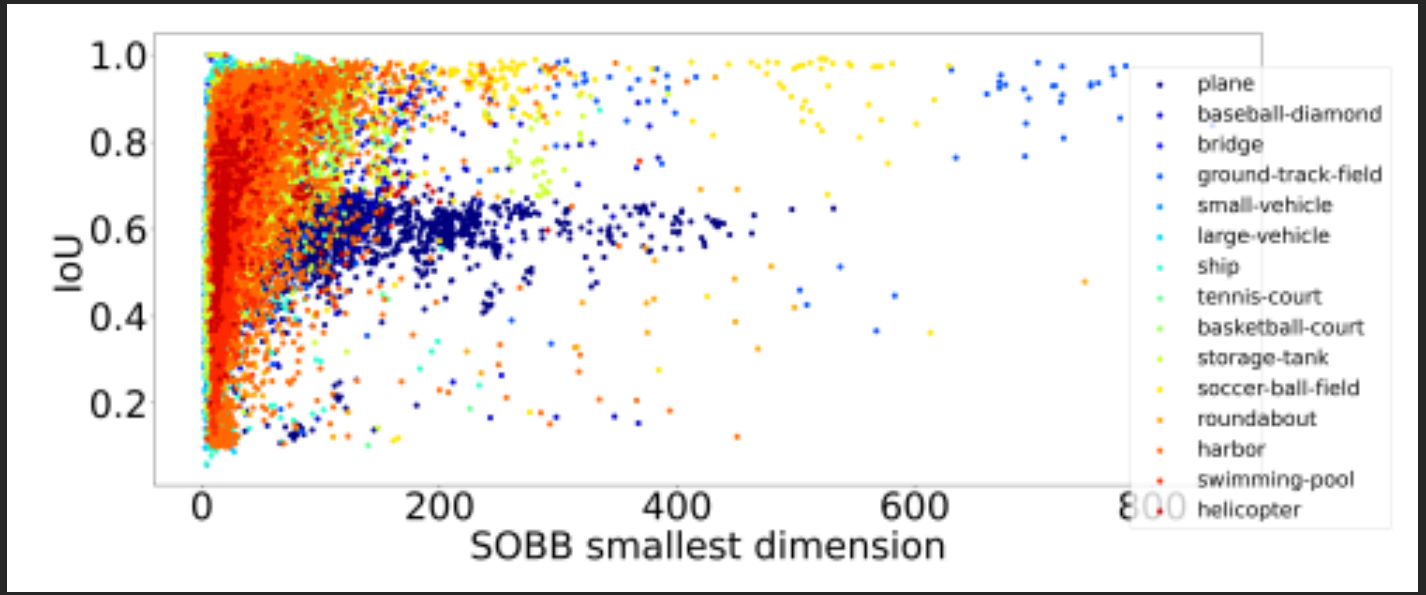
- Rectangular shaped objects yield high IOU.
- Circular shape leads to ambiguous BB orientation (any orientation yields the same area).
- Airplanes and irregularly-shaped objects have arbitrary orientations.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Findings – DOTA 1.0

- Not a clear monotonic relationship like previously.
- Discrepancy between AOBB and SOBB is caused by other factors:
  - human-center bias
  - orientation ambiguity

Discrepancy between OBBs is deeper than HBBs, hence its more likelihood to affect IOU and mAP.



Scatter plot of the smallest dimension of SOBB and IOU(AOBB, SOBB).

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).

Summary

- Blind use of IOU for BBs is dangerous.
- High IOU with the annotations doesn't equate high IOU with the GT.
- The effect on IOU is dependent on BB size (especially for small BB).
- Small and medium objects are more susceptible to sub-pixel errors than large objects.

Murrugarra-Llerena, J., Kirsten, L. N., & Jung, C. R. (2022). *Can we trust bounding box annotations for object detection?*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4813-4822).