Using Gaussian Distributions for Training and Evaluating Oriented Object Detectors

Claudio Rosito Jung

Institute of Informatics - Federal University of Rio Grande do Sul

In collaboration with Jeffri Erwin Murrugarra Llerena

LATAM Research Program Workshop

March 16, 2024

INSTITUTO DE INFORMÁTICA UFRGS

UFRGS

Introduction

- Deep Learning (DL) greatly improved object detectors
  - Used initially as a classifier module only, coupled with sliding windows
  - Moved to an evaluation at sparse windows only (proposals)
  - Current methods predict object labels and regress the object representation

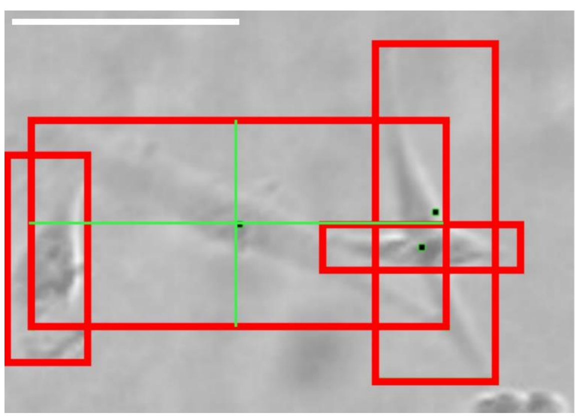
Detection result using EfficientDet (Tan et al., CVPR 2020)

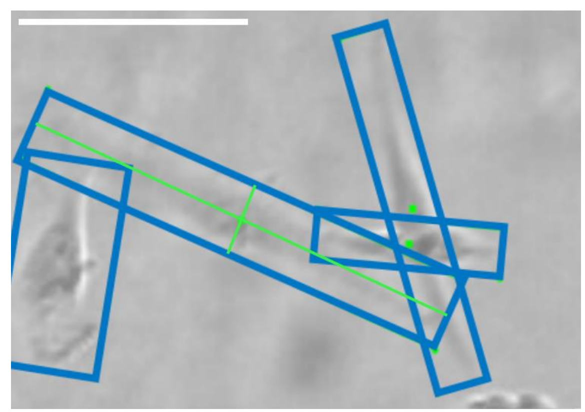
Introduction

- Several aspects influence the quality (and speed) of a DL-based object detector
  - Backbone (feature extraction)
  - Classification and Localization Heads
  - Training dataset (and data augmentation primitives)
  - **Representation choice for the object**
  - Classification and **Regression (Localization) loss**
- There is some coupling among these modules, but
  - In general, we can change one of them and plug it into an existing detector

Object Representation

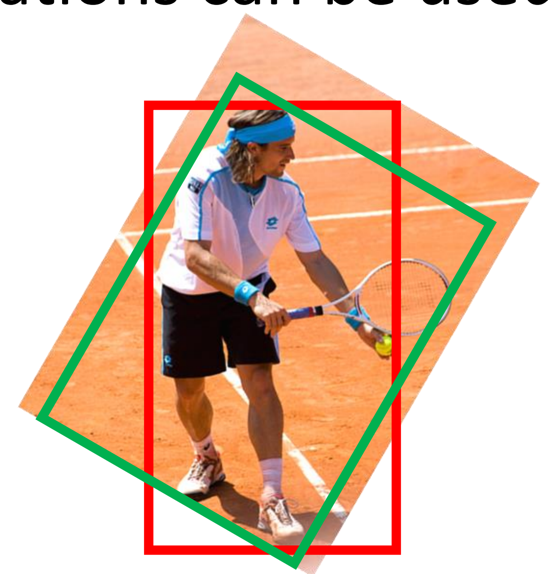
- *Horizontal Bounding Boxes* (HBBs) are by far the most popular choice
  - Very easy to annotate, just four parameters to regress
  - Not suitable for oriented elongated objects
- *Oriented Bounding Boxes* (OBBs) are an alternative
  - Better representation power, annotation not that harder, additional parameter (angle) must be regressed
  - Require a thorough parametrization to avoid ambiguities





Object Representation

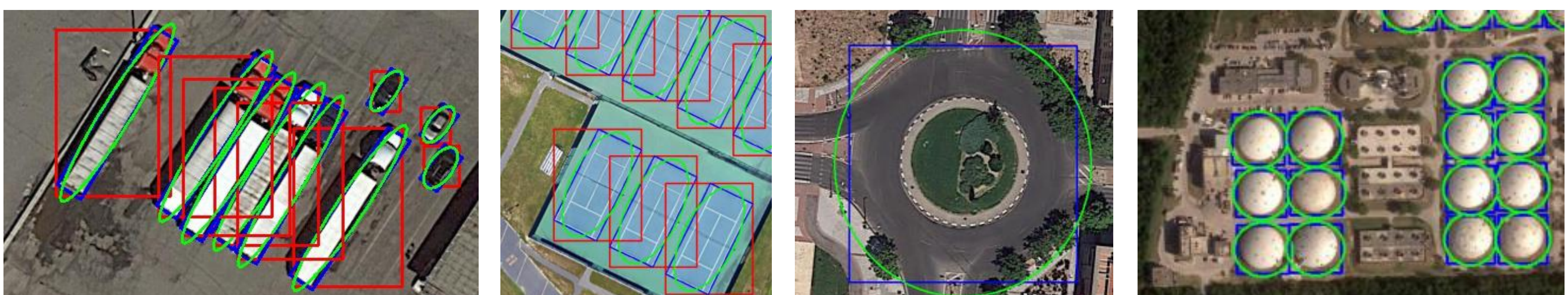
- Data augmentation is a popular choice for enlarging training data
- Geometrical transformations limited for HBB representations
  - The HBB of a rotated object is not the direct rotation of the HBB
- Arbitrary rotations can be used of OBBs



We cannot predict the “rotated” HBB (right) from the original annotation (left)

Object Representation

- OBBs are an improvement, <u>but</u>
  - Roughly circular or irregular objects generate a roughly square bounding box
  - Circular regions generate OBBs with angular ambiguity
    - Any square with arbitrary rotation can be chosen
- Can we use elliptical regions (EBBs)?
  - They can also represent oriented objects
  - Natural fit for circular regions



Regression Loss

- Most approaches use a corner-based or parameter-based cost (ℓ_1, ℓ_2, Huber, etc.)
  - Such strategies ***do not correlate directly*** with the IoU
- The Generalized IoU (GIoU) and variants became popular for regressing HBBs
- Extensions of IoU and GIoU exist for OBBs as well, but
  - The formulation is more complex
  - Customized implementations are required for backpropagation

Main ideas and contributions

- Use a 2D Gaussian function as a fuzzy representation for the objects (GBB)
- Use Elliptical Regions (EBBs) for representing oriented objects
  - More coherent with GBBs than OBBs
- A similarity metric (ProbIoU) based on GBBs
  - Differentiable closed-form expression
    - Can be used as a regression loss function to train detectors
  - Strongly correlates with the IoU between the respective EBBs
    - Can be used as an evaluation metric

2D Gaussian Distributions

- Can use a 2D Gaussian function as a fuzzy representation for oriented objects (Gaussian Bounding Box, or GBB)
  - Used as intermediate representations in (Yang et al, ICML 2021, NeurIPS 2021)
- Can compute closed-form "distance" terms between two 2D Gaussians to train object detectors
  - Gauss-Wesserstein Distance (GWD) used in (Yang et al, ICML 2021), Kullback-Leibler Divergence (KLD) used in (Yang et al, NeurIPS 2021)
    - GWD and KLD present poor results when used as regression terms (empirical non-linear functions used)
    - No theoretical or empirical relationship with IoU
- Mapping from OBB to 2D Gaussians is <u>not bijective</u>
  - Any rotated square maps to a single isotropic Gaussian

Our contributions

- Jointly tackle the problems of *training*, *representing*, and *evaluating* oriented object detectors
- Introduce a similarity measure called Probabilistic IoU (ProbIoU) that induces a distance function $\mathcal{L}_{ProbIoU}$ = 1 – ProbIoU
- $\mathcal{L}_{ProbIoU}$ can be directly used as a regression loss term to train object detectors
- Propose a bijective mapping from GBBs to EBBs
- Show that show that the IoU between two EBBs strongly correlates with the ProbIoU between the corresponding GBBs.
  - Potential use for evaluating oriented object detectors

.inf INSTITUTO DE INFORMÁTICA UFRGS

UFRGS

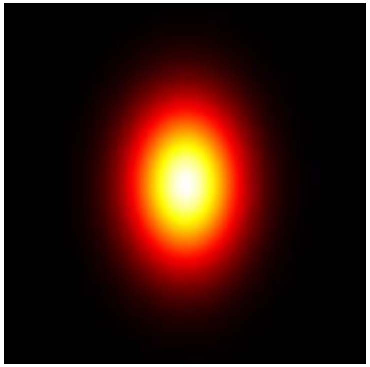
Gaussian Bounding Boxes

- A 2D Gaussian is parametrized by a mean vector $\boldsymbol{\mu}$ and a covariance matrix Σ

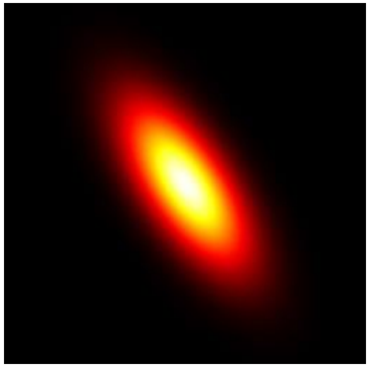
$$\boldsymbol{\mu} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$$\Sigma = \begin{bmatrix} a & c \\ c & b \end{bmatrix} = \begin{bmatrix} a' \cos^2\theta + b' \cos^2\theta & \frac{1}{2}(a' - b') \sin 2\theta \\ \frac{1}{2}(a' - b') \sin 2\theta & b' \cos^2\theta + a' \cos^2\theta \end{bmatrix}$$

- a' and b' are the uncorrelated variances and θ the orientation



Uncorrelated



Rotated

Gaussian Bounding Boxes

- GBBs from a segmentation mask Ω:

$$\boldsymbol{\mu} = \frac{1}{\#\Omega} \int_{x \in \Omega} \boldsymbol{x} d\boldsymbol{x},$$

$$\Sigma = \frac{1}{\#\Omega} \int_{x \in \Omega} (\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^T d\boldsymbol{x},$$

- Closed-form expression if Ω is an HBB with dimensions $W \times H$:

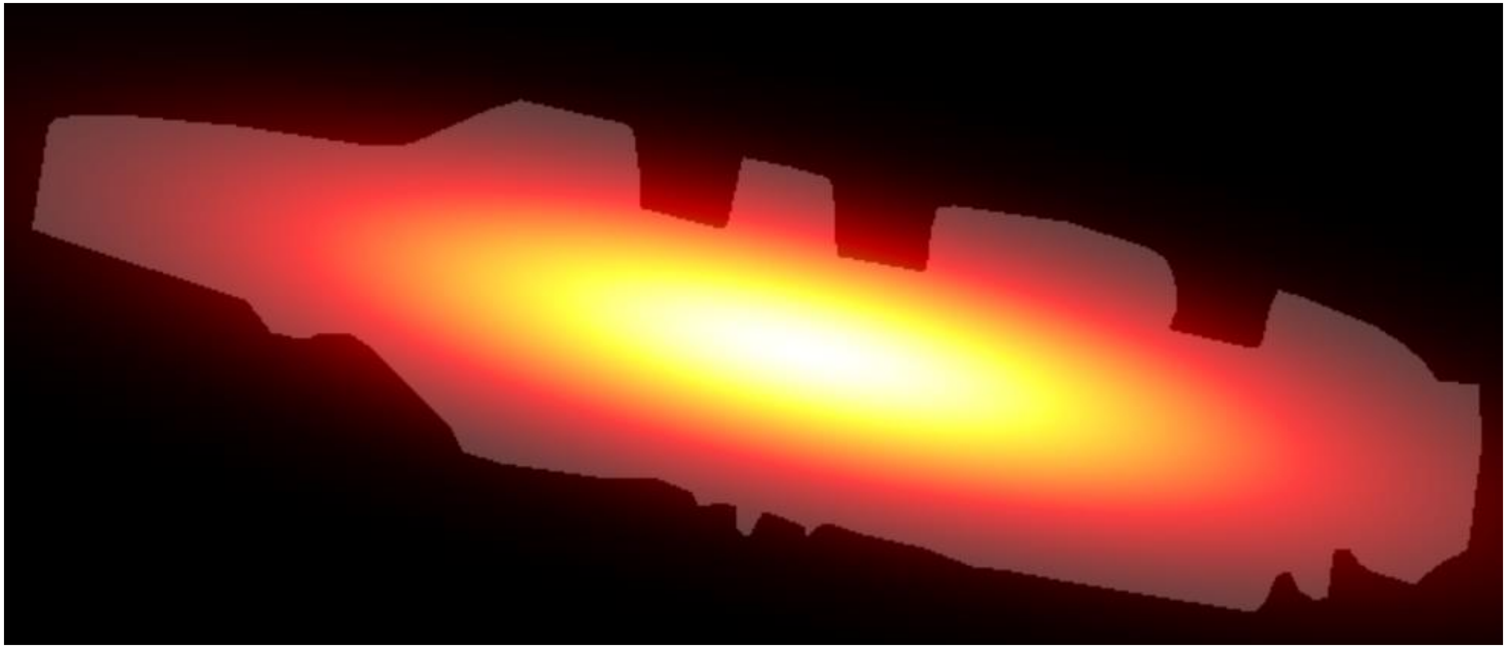
$\boldsymbol{\mu}$ is the centroid

$$\Sigma = \frac{1}{12} \begin{bmatrix} W^2 & 0 \\ 0 & H^2 \end{bmatrix}$$

- Similar formulation with rotation if Ω is an OBB

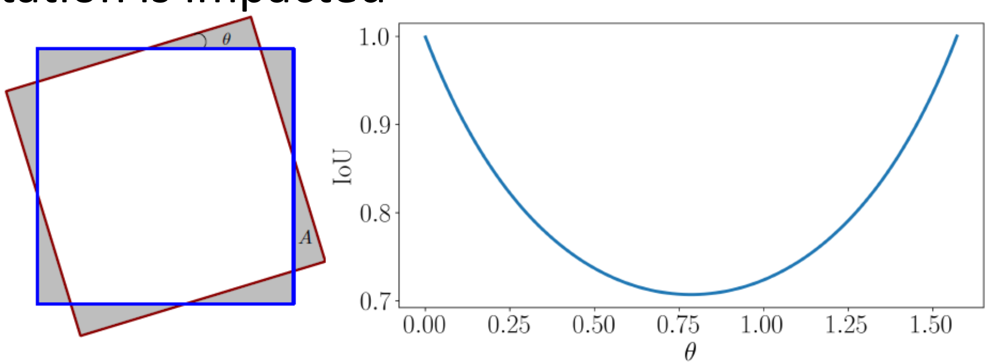
Gaussian Bounding Boxes

- GBBs from a segmentation mask Ω:



Binary representation from GBBs

- Can map a GBB to an OBB
- Eigen-decomposition of Σ
  - Eigenvalues relate to the OBB dimensions
  - Eigenvectors relate to the rotation
- Mapping from GBB to OBB <u>is not unique</u>
  - Any rotated square map to the same GBB
    - OBB orientation cannot be retrieved
    - IoU computation is impacted



(a) Rotated squares (b) θ vs. IoU

Binary representation from GBBs

- Ellipses are a natural choice (EBBs)
  - Can be extracted from the level-sets of the Mahalanobis distance:

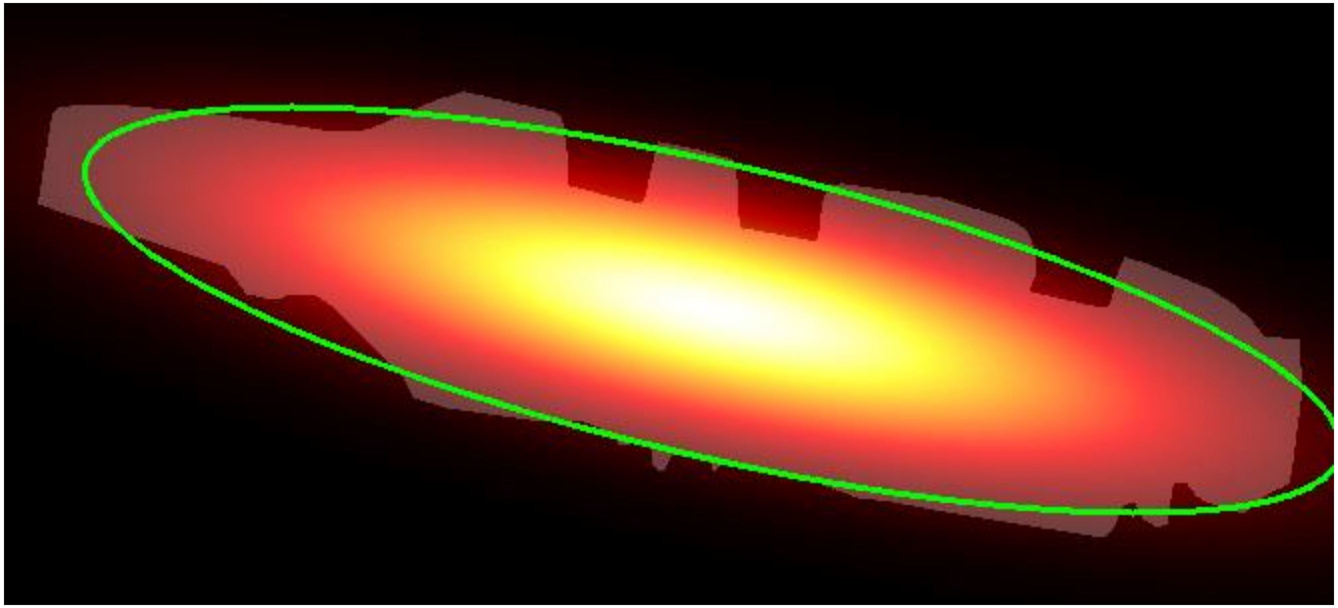
$$(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu}) = r^2$$

  - Choose r such that generated ellipse matches area of OBB

- GBB to EBB mapping <u>is a bijection</u>
  - Natural solution for circular objects
- However, computing the IoU between EBBs is <u>hard</u>!

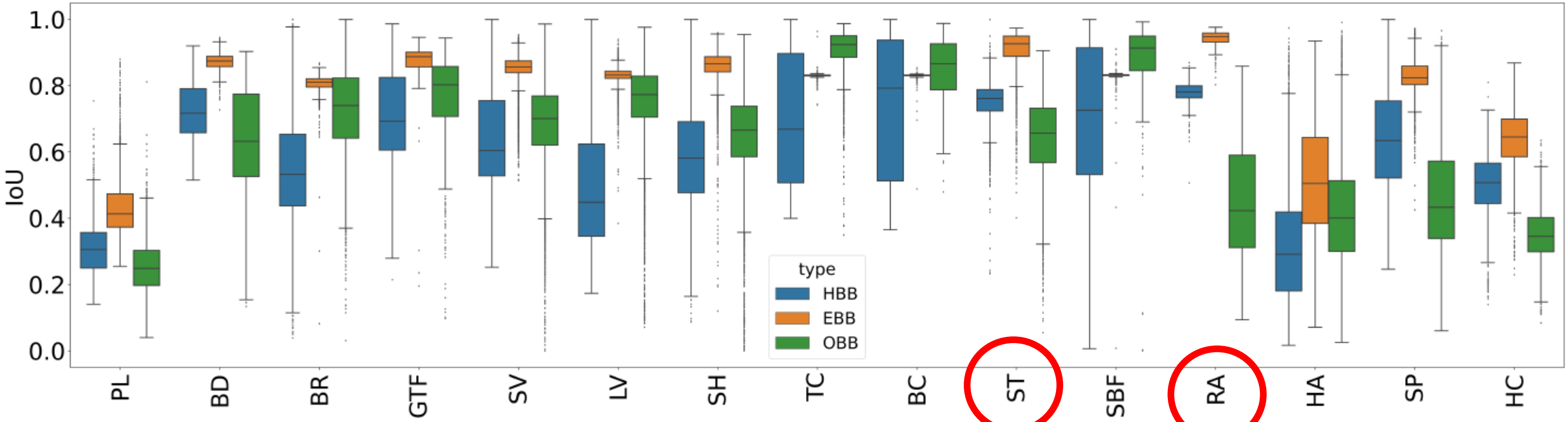
Binary representation from GBBs

- Example of EBBs from segmentation masks



HBB, OBB, EBB vs. Segmentation Masks

- Comparing HBB, OBB, and EBBs with segmentation masks for DOTA/iSaid datasets (aerial images)
  - 15 categories
- Mean IoU values over all categories
  - HBBs: 0.61
  - OBBs: 0.63
  - EBBs: 0.79



Probabilistic IoU

- The Bhattacharyya Coefficient measures the ***similarity*** between two distributions p and q

$$B_C(p,q) = \int \sqrt{p(\boldsymbol{x})q(\boldsymbol{x})}d\boldsymbol{x}$$

- The Bhattacharyya Distance measures their ***dissimilarity***:

$$B_D(p,q) = -\ln B_C(p,q)$$

  - it is not an actual distance (does not satisfy the triangle inequality), but the Hellinger distance is:

$$H_D(p,q) = \sqrt{1 - B_C(p,q)}$$

- $\mathrm{ProbIoU}(p,q) = 1 - H_D(p,q)$ is a similarity metric

Probabilistic IoU

- ProbIoU is easy to compute for Gaussian distributions:

$$B_D = \frac{1}{8}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \frac{1}{2} \ln\left(\frac{\det\Sigma}{\sqrt{\det\Sigma_1 \det\Sigma_2}}\right),$$

$$\Sigma = \frac{1}{2}(\Sigma_1 + \Sigma_2)$$

  - Note that B_C and hence H_D present a ***differentiable, closed-form expression.***
  - Important characteristics for a ***regression loss!***

Probabilistic IoU for Training

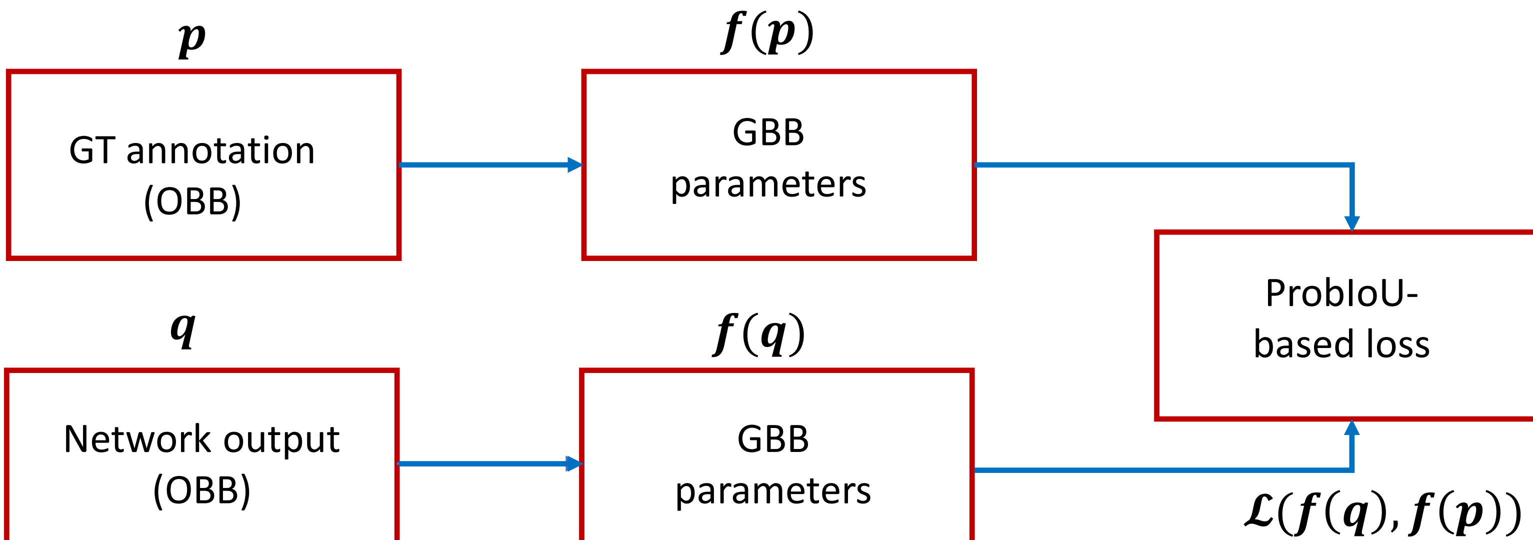
- Can treat p as the GT annotation and q as the (Gaussian) distribution to be regressed

$$\mathcal{L}_{ProbIoU}\,(p,q) = H_D(p,q) = 1 - \mathrm{ProbIoU}(p,q) \in [0,1)$$

- $\mathcal{L}_{ProbIoU}$ *does not produce vanishing gradients (unlike the IoU)!*
  - Recall the infinite support of a Gaussian
- Can also use $\mathrm{ProbIoU}(p,q)$ to *validate* detection results as an alternative to the IoU.

Probabilistic IoU for Training

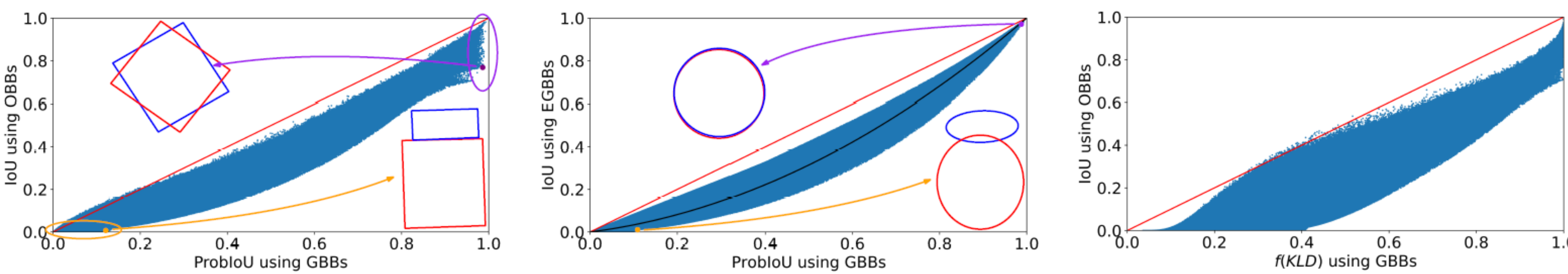
- Can be used directly if the GT annotations and objects are parametrized as GBBs
- Can be seamlessly integrated into OBB-based object detectors:



  - All connections in blue are ***differentiable*** w.r.t. to OBB parameters!

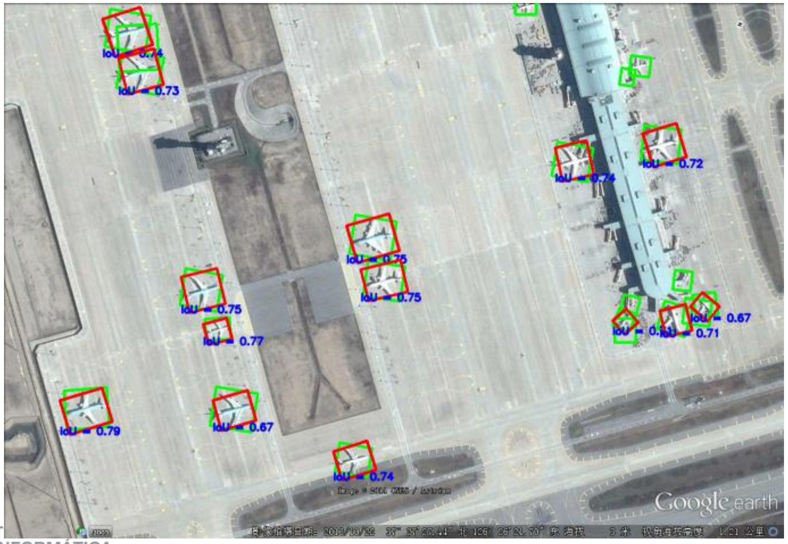
Probabilistic IoU for Evaluation

- Relationship between ProbIoU and IoU for OBBs and EBBs
  - Empirical evaluation using millions of OBB/EBB pairs
  - ProbIoU using the corresponding GBBs
    - Both present strong correlation with the IoU
  - Results for KLD provide a Loose relationship!



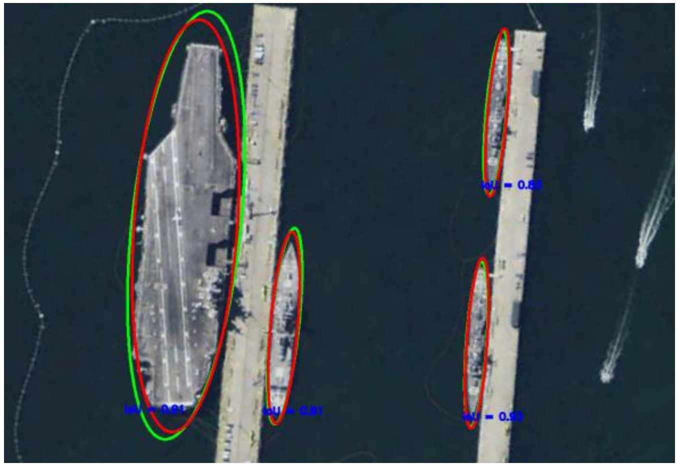
Results with ProbIoU-loss

- Results using OBBs as final representation



Results with ProbIoU-loss

- Results using EBBs as final representation



Conclusions

- GBBs seem a promising direction for oriented object detection]
- Existing approaches based on GBBs face limitations
  - Require non-linear mapping for regression loss
  - No relationship with the IoU
  - Mapping from GBB to OBB is not bijective
- We show that EBBs are viable representations
- The proposed ProbIoU strongly correlates with EBB-IoU
- The ProbIoU-loss can be directly used as a regression term

Paper:

Probabilistic Intersection-Over-Union for Training and Evaluation of Oriented Object Detectors, IEEE Transaction on Image Processing 2024

Github:

https://github.com/ProbIOU/

Thank you!

crjung@inf.ufrgs.br

INSTITUTO DE INFORMÁTICA UFRGS

UFRGS