



GAUCHO: GAUSSIAN DISTRIBUTIONS WITH CHOLESKY DECOMPOSITION FOR ORIENTED OBJECT DETECTION

José H. L. Marques¹†, Jeffri Murrugarra-Llerena²†, Claudio R. Jung¹. Equal Contribution (†)

¹Institute of Informatics, Federal University of Rio Grande do Sul, Brazil, ²Computer Science Department, Stony Brook University

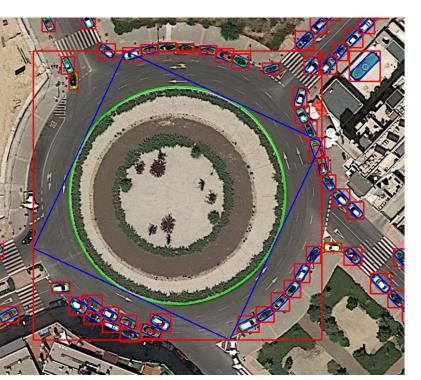


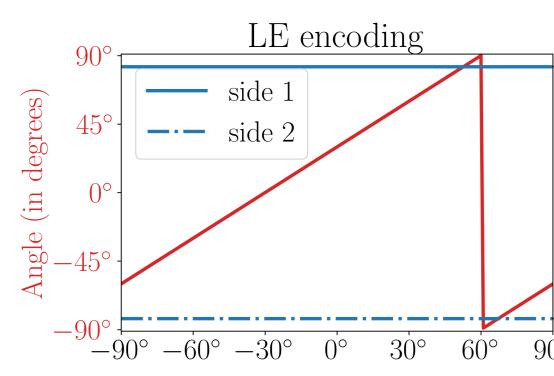


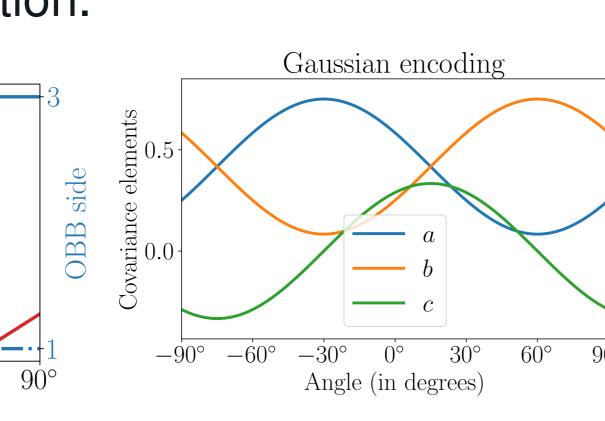
Motivation & Goals

- Oriented object detection (OOD) is an essential application of computer vision that extends horizontal (HBB) object detection by considering the orientation of objects in images.
- Oriented Bounding Boxes (OBBs) are typically represented as (x, y, w, h, θ) , but the parametrization is not unique: Long Edge (LE), OpenCV (OC) present the **boundary discontinuity problem**
- Gaussian representations are unique, but the mapping from OBB to Gaussian still generates training ambiguities.
- In this work, we propose a new paradigm for OOD by regressing the parameters of a **Gaussian distribution directly from the network**, avoiding the intermediate use of OBBs that theoretically **mitigates the boundary discontinuity problem**
- To avoid a constrained optimization imposed by positive-definiteness of the covariance matrix, we explore the **Cholesky Decomposition** to develop the **GauCho** head.
- We show a one-to-one mapping between GauCho and Oriented Ellipses (OEs), and advocate their use as an alternative representation for oriented object detection.









Mathematical Background

Gaussian Bounding Boxes

• Represent an OBB with center (x,y), dimensions (w,h) and orientation $\theta \in [-90^\circ, 90^\circ)$ as a Gaussian distribution with mean μ and covariance matrix C.

$$\mu = (x, y)^T, \quad C = \begin{bmatrix} a & c \\ c & b \end{bmatrix} = R\Lambda R^T, R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_w & 0 \\ 0 & \lambda_h \end{bmatrix} = s \begin{bmatrix} w^2 & 0 \\ 0 & h^2 \end{bmatrix}, \quad (1)$$

where R is the rotation and λ_* the eigenvalue matrices, and s is a scaling factor.

$$C = \begin{bmatrix} \lambda_w \cos^2 \theta + \lambda_h \sin^2 \theta & \frac{1}{2} (\lambda_w - \lambda_h) \sin(2\theta) \\ \frac{1}{2} (\lambda_w - \lambda_h) \sin(2\theta) & \lambda_w \sin^2 \theta + \lambda_h \cos^2 \theta \end{bmatrix}.$$
 (2)

Cholesky Decomposition

ullet The Cholesky decomposition for a positive-definite matrix C involves a lower-triangular matrix

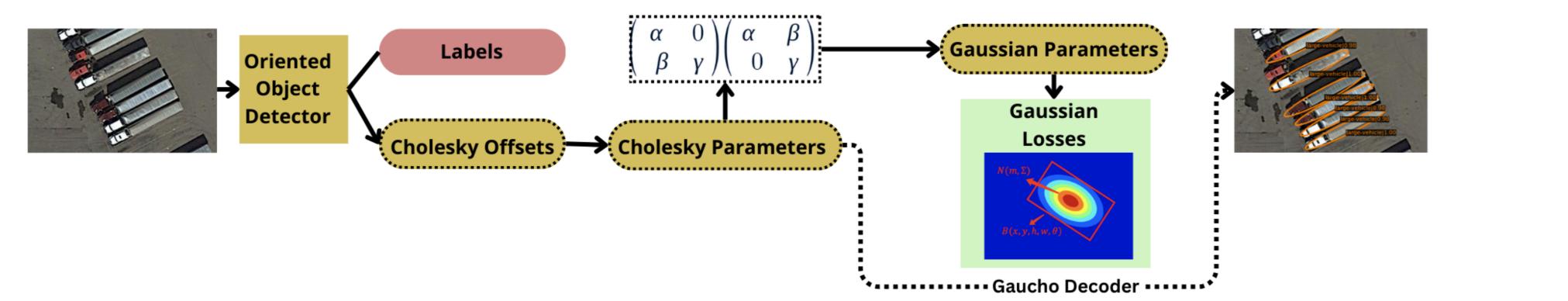
$$L = \begin{bmatrix} \alpha & 0 \\ \gamma & \beta \end{bmatrix}, C = LL^T = \begin{bmatrix} \alpha^2 & \alpha\gamma \\ \alpha\gamma & \beta^2 + \gamma^2 \end{bmatrix} = \begin{bmatrix} a & c \\ c & b \end{bmatrix}.$$
 (3)

- The parameters (α, β, γ) provide a **unique** mapping to a Gaussian
- The GauCho head aims to directly regress $\alpha, \beta > 0, \gamma \in \mathbb{R}$ as an alternative to an OBB head.
- Theoretical relationships between (α, β, γ) and OBB dimensions (w, h) (see paper).

GauCho Object Detectors

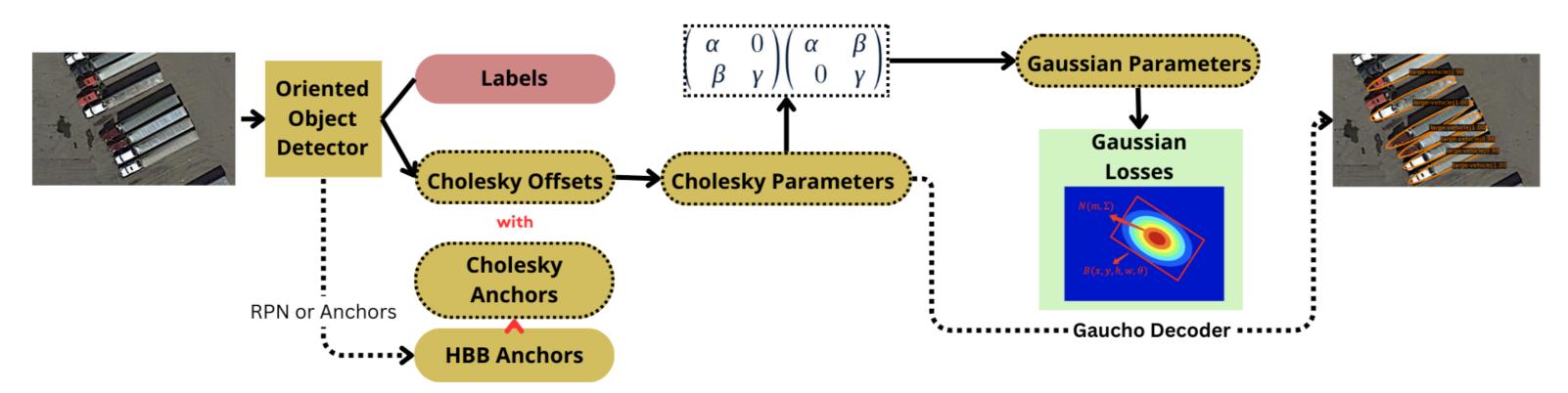
• Anchor-Free Detectors Given (p_x, p_y) a central point, t the stride. We regress the offset d_x, d_y and $d_\alpha, d_\beta, d_\gamma$ with linear activations. Then:

$$x = p_x + td_x, \ y = p_y + td_y, \alpha = te^{d_\alpha}, \ \beta = te^{d_\beta}, \ \gamma = td_\gamma, \tag{4}$$



• Anchor-Based Detectors Given axis-aligned anchors characterized by (a_x, a_y, a_w, a_h) , we regress center (d_x, d_y) and the multiplicative $(d_\alpha, d_\beta, d_\gamma)$ offsets with linear activations. Then:

$$x = x_a + a_w d_x$$
, $y = y_a + a_h d_y$, $\alpha = \sqrt{s} a_w e^{d_\alpha}$, $\beta = \sqrt{s} a_h e^{d_\beta}$, $\gamma = \sqrt{s} \max\{\delta, |a_w - a_h|\} d_\gamma$, (5)



Experimental Results

We used detectors (FCOS [22], RetinaNet [14], R³Det [31], Rol-Transformer [2]) and compared GauCho with OBB head with Gaussian-based loss (GWD [33], KLD [35], and ProbloU [20]).

Detector	Head-Loss	AP_{50}	AP_{75}	AP	AP_{50}	AP_{75}	AP	AP_{50}	AP_{75}	AP
		HRSC (OBB)			UCAS-AOD (OBB/OE)			DOTA v1.0 (OBB)		
FCOS	OBB-GWD	88.93	76.67	84.93	90.22/90.26	55.75/65.42	53.73/ 59.52	69.76	34.68	37.89
	GauCho-GWD	89.76	76.30	85.26	90.17/90.17	53.84/64.84	52.33/58.55	71.22	35.85	38.63
	OBB-KLD	88.38	66.42	82.24	90.22/90.26	50.03/64.96	52.48/59.04	71.74	28.30	36.18
	GauCho-KLD	89.94	78.99	87.86	90.04/90.07	55.01/65.06	52.72/59.37	72.16	33.27	38.46
	OBB-ProbloU	90.08	76.84	87.27	90.17 /90.16	46.73/64.83	52.27/ 59.27	71.31	37.34	39.80
	GauCho-ProbloU	89.86	78.21	87.58	90.14/ 90.18	55.35/65.27	53.03 /59.08	72.86	37.69	40.65
RetinaNet-ATSS	OBB-GWD	89.47	75.65	83.83	89.72/89.83	34.37/60.16	46.28/56.08	71.51	36.34	39.59
	GauCho-GWD	90.32	78.34	86.39	89.79/89.83	50.40/62.69	51.55/57.92	71.36	38.00	40.29
	OBB-KLD	90.17	77.62	86.00	89.64/89.65	49.33/62.98	50.73/ 57.10	72.05	37.72	40.47
	GauCho-KLD	90.40	80.45	88.56	89.71/89.71	50.18/63.01	50.84 /57.08	72.71	38.47	40.57
	OBB-ProbloU	90.20	77.67	87.37	89.87/89.87	48.93/ 63.16	51.03 /57.09	72.14	39.77	40.97
	GauCho-ProbloU	90.48	80.35	88.56	89.78/89.74	50.61 /63.04	51.34/ 57.43	73.21	37.63	40.91
R ³ Det-ATSS	OBB-GWD	89.66	65.68	81.90	90.02/90.07	38.60/61.40	47.54/56.68	67.98	34.89	37.11
	GauCho-GWD	89.52	65.83	81.77	89.94/89.95	49.87/62.15	51.41/56.72	70.53	35.74	39.07
	OBB-KLD	89.92	53.46	79.32	89.96/90.00	52.05/63.87	52.07/57.35	70.77	36.98	38.90
	GauCho-KLD	89.65	62.66	82.97	89.90/89.93	49.79/63.65	51.48/57.11	70.83	33.48	37.65
	OBB-ProbloU	89.19	51.37	78.40	89.98/90.19	44.85/ 64.28	50.23/ 57.67	70.85	36.66	38.91
	GauCho-ProbloU	90.02	76.43	85.76	89.95/89.96	51.72 /63.95	52.01 /57.41	71.23	33.64	37.89
Rol Transformer	OBB-GWD	90.35	88.51	80.40	90.31/90.32	58.37/69.07	55.20 /59.54	75.38	42.53	42.87
	GauCho-GWD	90.35	59.28	79.72	90.28/90.31	58.53/69.47	54.84/ 59.54	75.66	41.05	42.38
	OBB-KLD	90.52	89.36	90.25	90.35/90.35	64.15/73.71	57.42 /61.32	76.55	47.54	45.96
	GauCho-KLD	90.50	88.80	90.12	90.32/90.34	56.90/70.34	54.60/ 61.40	76.35	43.79	44.32
	OBB-ProbloU	90.54	89.12	90.16	90.35/90.37	63.05/73.40	56.76 /60.81	75.49	46.31	45.18
	GauCho-ProbloU	90.58	89.13	90.20	90.32/90.33	61.41/70.59	55.57/ 60.91	76.09	42.60	43.90

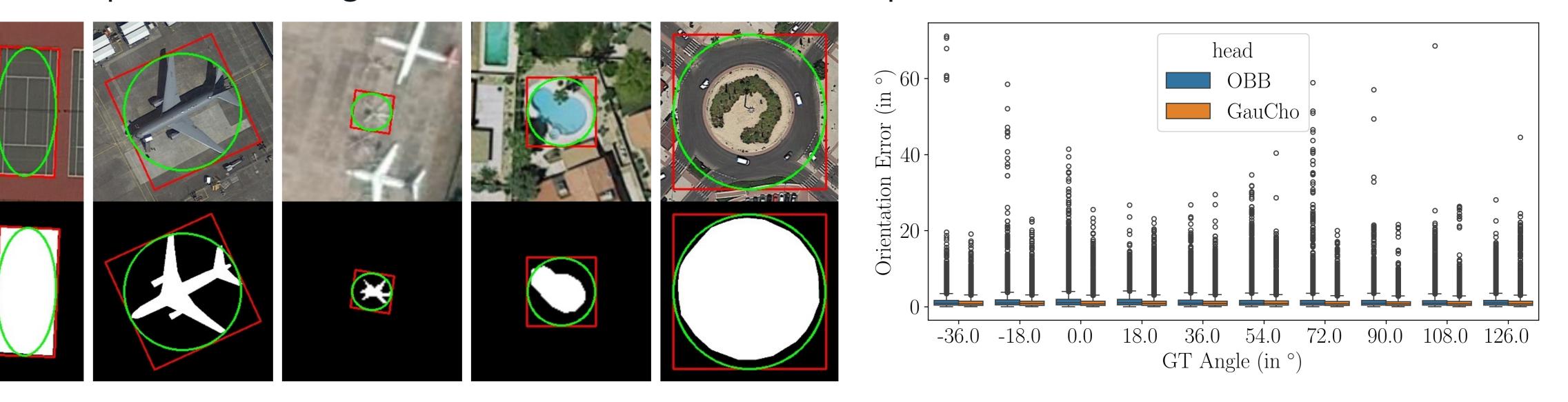
Discussion

A Critical Analysis of OBBs vs. OEs

- Circular or irregular objects might lead to **square-like OBBs**, for which the **orientation is ambiguous** and generates inconsistencies for augmentation based on rotations.
- Oriented Ellipses mitigate this problem, and are still suitable for elongated objects. However, they do not capture the orientation of oriented square objects (as for some planes, as shown below).
- We can compute the AP metrics using OEs instead of OBBs (larger differences for datasets with square objects see UCAS-AOD with restrictive IoU thresholds)

Orientation Consistency

- Our method is fully compatible with *Rotation Equivariance* techniques and is not affected by boundary discontinuity and orientation inconsistencies.
- Compared to the standard parametrization, our method shows better angular consistency
- We computed the histogram of orientation errors for multiple rotations of the HRSC dataset



Conclusions

- We proposed a novel representation for OOD based on Gaussian-Cholesky decomposition (GauCho)
- GauCho can be **integrated with existing object detection paradigms** being a viable alternative to the traditional OBB head.
- We showed that GauCho mitigates the angular discontinuity problem present in OBB representations and provided theoretical bounds that relate GauCho parameters with actual OBB dimensions.
- GauCho produces similar AP metrics compared to OBB heads for different detectors (including SOTA techniques), Gaussian-based loss functions, and datasets.

References

Reference numbers are according to the manuscript.

This study was partially funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We also want to acknowledge to LXAI supercomputer program, Gobierno de Jalisco, and Nvidia Corp for providing access to a DGX A-100 640 server.