Dense Label Encoding for Boundary Discontinuity Free Rotation Detection

Xue Yang[1,2], Liping Hou[4], Yue Zhou[3], Wentao Wang[1], Junchi Yan[1,2]
[1]Department of CSE, SJTU [2]MoE Key Lab of Artificial Intelligence, AI Institute, SJTU
[3]Department of EE, SJTU [4]University of Chinese Academy of Sciences

CVPR VIRTUAL JUNE 19-25

Introduction:

- Task: Design a novel boundary discontinuity free rotation detector based on angle classification.
- Challenges:
  - Thick prediction layer: From the perspective of Gflops, and Param, detectors based on CSL have increased by about 82.96% and 45.63%.
  - Unfriendliness to small aspect ratio objects: long-side definition method is not suitable for square-like box and will suffer a special problem.
- Our main contributions:
  - To improve the robustness especially for objects with small aspect ratio, we propose Angle Distance and Aspect Ratio Sensitive Weighting (ADARSW), which further improves accuracy by making our proposed DCL-based detector sensitive to angular distance and object's aspect ratio. In contrast, the existing CSL-based detector suffers from its long-side definition for detecting square-like objects.
  - We compare the impact of two classic Densely Coded Labels (DCL) by introducing them to the angle classification task for potential speedup, namely Binary Coded Label (BCL) and Gray Coded Label (GCL), which are more compact than existing CSL. We empirically show that DCL, especially BCL can lead to notable training speed boost (about three times) as well as detection accuracy.
  - Extensive experiments and visual analysis on different datasets and detectors prove the efficacy of our techniques.
- Codes:
  - https://github.com/yangxue0827/RotationDetection
  - https://github.com/Thinklab-SJTU/DCL_RetinaNet_Tensorflow

Proposed Approach

- Densely Coded Label: Binary Coded Label and Gray Coded Label are two Densely Coded Label methods commonly used in the field of electronic communication. Their advantage is that they can represent a larger range of values with less coding length. Thus, they can effectively solve the problem of excessively long coding length in CSL and One-Hot based methods, as shown in Figure 1.
- ADARSW: We add a periodic trigonometric function to make the model sensitive to the distance of the angle and aspect ratio.

$$W^*(\Delta\theta) = |\sin(\alpha(\Delta\theta))| = |\sin(\alpha(\theta_{gt} - \theta_{pred}))|$$

$$\alpha = \begin{cases} 1, & (h_{gt}/w_{gt}) > r \\ 2, & otherwise \end{cases}$$

Experiments:

- Comparison of GFlops and Param over rotation detectors, under the same setting and hyperparameters.

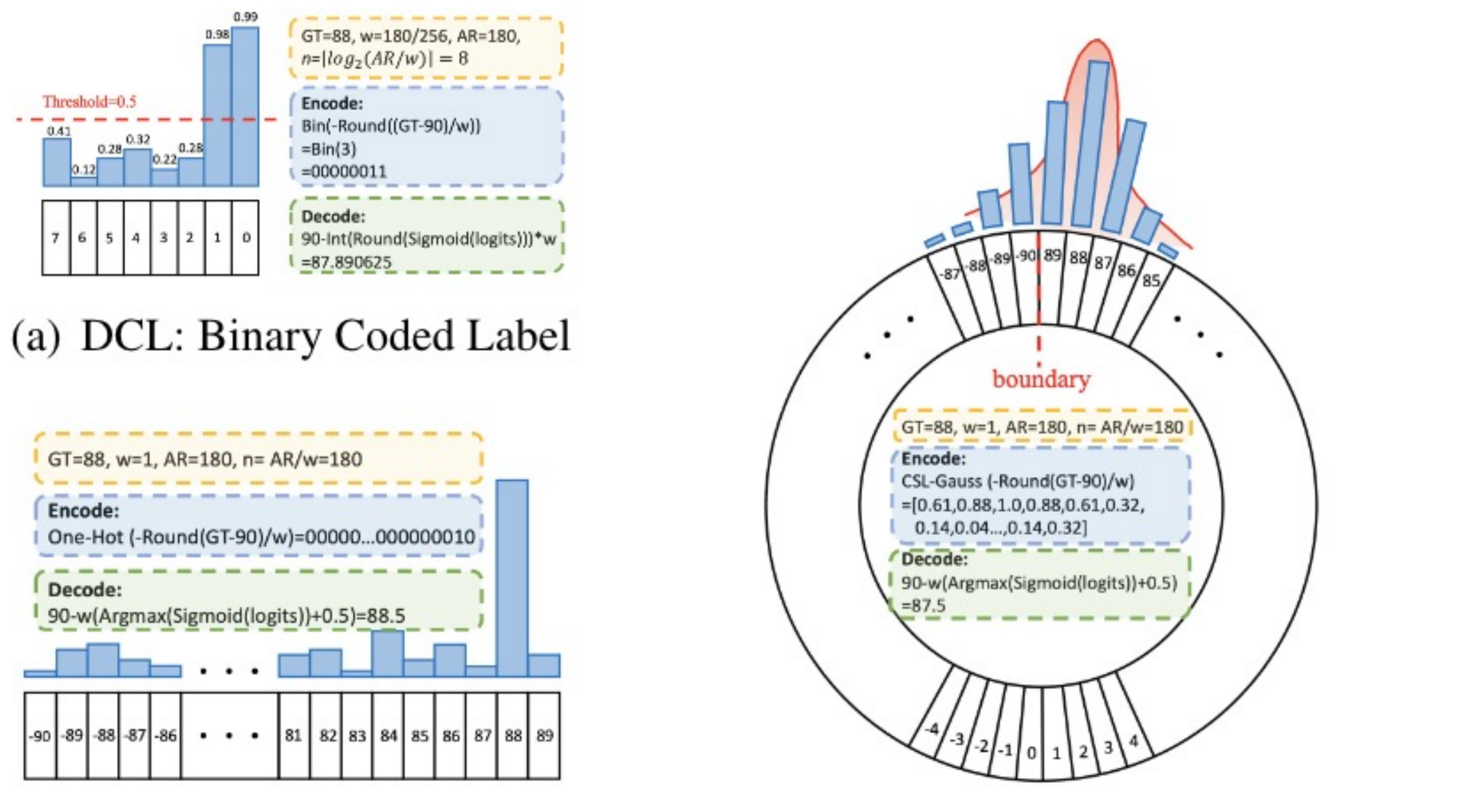
Base Model	ω	GFlops	ΔGFlops	Params (M)	ΔParams	Training Time
RetinaNet-Reg	-	139.35	-	36.97	-	-
RetinaNet-CSL	1	254.96	+82.96%	45.63	+23.42%	~3x
RetinaNet-BCL	1	143.87	**+3.24%**	37.31	**+0.92%**	**~1x**
RetinaNet-GCL	1	143.87	**+3.24%**	37.31	**+0.92%**	**~1x**

- Ablation study of four orientation detectors on DOTA test dataset. 5-mAP means the performance of the five categories listed.

Method	BR	SV	LV	SH	HA	5-mAP$_{50}$	mAP$_{50}$
RetinaNet-Reg	38.31	60.48	49.77	68.29	51.28	53.63	64.17
RetinaNet-CSL	40.55	66.77	51.50	73.60	53.76	57.24 (+3.61)	65.69 (+1.52)
RetinaNet-BCL	41.58	67.98	**57.34**	**74.66**	54.28	**59.17 (+5.54)**	**66.53 (+2.36)**
RetinaNet-GCL	**42.55**	**68.38**	56.40	73.53	**54.36**	59.04 (+5.41)	66.27 (2.10)

- ADARSW on small aspect ratio objects in DOTA.

Method	ADARSW	PL	BD	GTF	TC	BC	ST	SBF	RA	SP	HC	10-mAP$_{50}$	mAP$_{50}$
BCL		88.63	71.62	65.18	90.70	76.32	78.47	52.26	60.25	66.61	49.15	69.92	66.53
	✓	88.92	72.11	66.32	90.79	79.86	79.03	54.11	63.18	67.86	60.04	**72.22**	**67.39**
GCL		88.52	73.58	64.38	90.80	77.66	76.38	50.84	59.46	65.83	48.42	69.59	66.27
	✓	88.96	75.20	65.24	90.78	79.13	77.95	55.60	61.90	66.18	56.27	**71.72**	**67.02**



(a) DCL: Binary Coded Label

(b) SCL: One-Hot Label

(c) SCL: Circular Smooth Label

Figure 1: Examples of encoding and decoding process of One-Hot, CSL-Gaussian and BCL for angle prediction.

- Comparison of detection results under different angle discretization granularities denoted by ω.

Method	ω	BR	SV	LV	SH	HA	5-mAP$_{50}$	mAP$_{50}$	mAP$_{75}$	mAP$_{50:95}$
Reg	-	34.52	51.42	50.32	73.37	55.93	53.12	62.21	26.07	31.49
CSL	180/180	35.94	53.42	61.06	81.81	62.14	58.87	64.40	32.58	35.04
BCL	180/4	30.74	40.54	50.98	72.07	59.54	50.77	62.38	24.88	31.01
	180/8	36.65	52.58	60.46	82.24	61.60	58.71	**66.17**	33.14	35.77
	180/32	**39.83**	54.41	60.62	80.81	60.32	**59.20**	65.93	**35.66**	**36.71**
	180/64	38.22	**54.70**	60.16	80.75	60.11	58.79	65.00	34.31	36.00
	180/128	36.76	53.73	**61.35**	**82.52**	58.42	58.56	65.14	34.28	35.69
	180/180	37.42	53.72	58.70	80.73	**63.31**	58.78	65.83	33.94	36.35
	180/256	37.66	53.83	60.66	80.43	60.74	58.66	64.97	33.52	35.21
	180/512	37.93	53.85	58.52	80.04	60.87	58.24	64.88	33.09	34.99
GCL	180/4	30.90	41.20	48.30	72.93	60.16	50.70	62.98	23.83	30.81
	180/8	36.88	51.10	59.81	82.40	61.57	58.35	65.23	33.92	35.29
	180/32	38.04	**54.77**	60.88	**82.75**	61.24	**59.54**	65.11	**34.67**	36.15
	180/64	**38.05**	54.36	60.59	81.84	60.39	59.05	64.78	33.23	35.67
	180/128	37.74	54.36	59.43	81.15	60.51	58.64	**66.13**	33.65	**36.34**
	180/256	35.81	53.78	58.35	81.45	59.84	57.85	64.87	33.77	35.97
	180/512	37.99	54.23	**61.61**	80.84	**62.13**	59.36	64.34	34.08	35.92

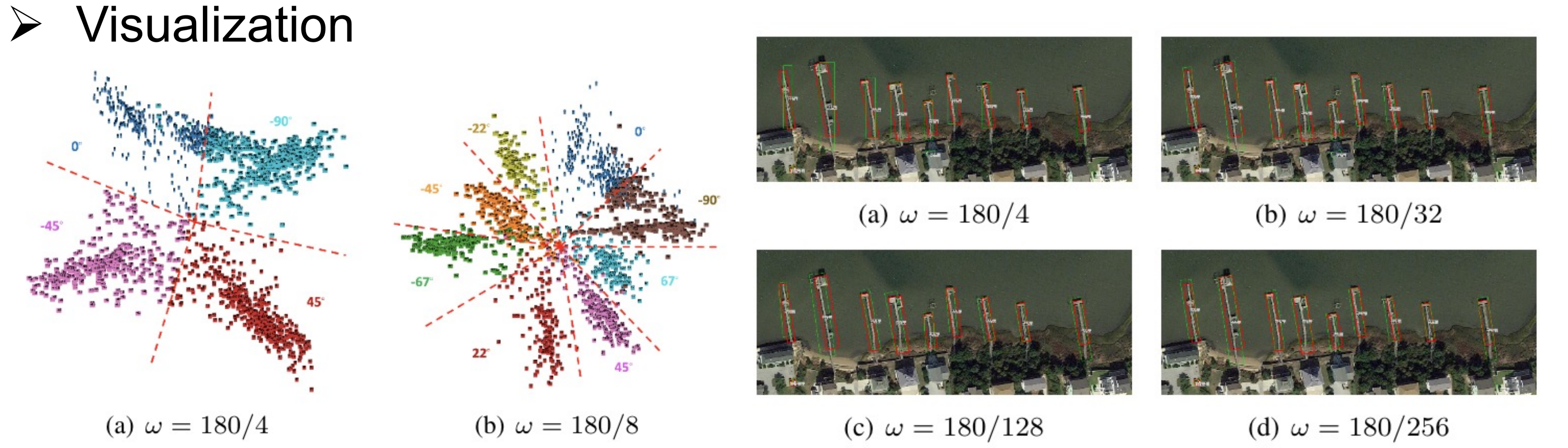
- More results of classification and regression-based methods.

Method	ICDAR2015	UCAS-AOD car(07/12)	UCAS-AOD plane(07/12)	UCAS-AOD mAP$_{50}$ (07/12)	MLT
Reg.	82.38	87.28/90.79	90.42/97.52	88.85/94.16	64.01
CSL	**83.81**	88.09/**92.93**	90.38/97.22	89.23 /95.07	65.08
BCL	83.17	**88.15**/92.35	**90.57/97.86**	**89.36/95.10**	**65.26**

- Detection accuracy on DOTA.

	Method	Backbone	MS	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP$_{50}$
Two-stage methods	ICN [2]	ResNet101	✓	81.40	74.30	47.70	70.30	64.90	67.80	70.00	90.80	79.10	78.20	53.60	62.90	67.00	64.20	50.20	68.20
	RoI-Transformer [6]	ResNet101	✓	88.64	78.52	43.44	75.92	68.81	73.68	83.59	90.74	77.27	81.46	58.39	53.54	62.83	58.93	47.67	69.56
	CAD-Net [54]	ResNet101		87.8	82.4	49.4	73.5	71.1	63.5	76.7	90.9	79.2	73.3	48.4	60.9	62.0	67.0	62.2	69.9
	SCRDet [51]	ResNet101	✓	89.98	80.65	52.09	68.36	68.36	60.32	72.41	90.85	**87.94**	86.86	65.02	66.68	66.25	68.24	65.21	72.61
	Gliding Vertex [44]	ResNet101		89.64	85.00	52.26	**77.34**	73.01	73.14	86.82	90.74	79.02	86.81	59.55	**70.91**	72.94	70.86	57.32	75.02
	Mask OBB [38]	ResNeXt101 [43]	✓	89.56	**85.95**	54.21	72.90	76.52	74.16	85.63	89.85	83.81	86.48	54.89	69.64	73.94	69.06	63.32	75.33
	FFA [9]	ResNet101	✓	90.1	82.7	54.2	75.2	71.0	79.9	83.5	90.7	83.9	84.6	61.2	68.0	70.7	76.0	63.7	75.7
	APE [59]	ResNeXt101		89.96	83.62	53.42	76.03	74.01	77.16	79.45	90.83	87.15	84.51	67.72	60.33	**74.61**	71.84	65.55	75.75
	CenterMap OBB [39]	ResNet101	✓	89.83	84.41	54.60	70.25	77.66	78.32	87.19	90.66	84.89	85.27	56.46	69.23	74.13	71.56	66.06	76.03
	FPN-CSL [48]	ResNet152	✓	**90.25**	85.53	**54.64**	75.31	70.44	73.51	77.62	90.84	86.15	86.69	**69.60**	68.04	73.83	71.10	68.93	76.17
Single-stage methods	PIoU [4]	DLA-34 [53]		80.9	69.7	24.1	60.2	38.3	64.4	64.8	90.9	77.2	70.4	46.5	37.1	57.1	61.9	64.0	60.5
	O^2-DNet [40]	Hourglass104 [28]	✓	89.31	82.14	47.33	61.21	71.32	74.03	78.62	90.76	82.23	81.36	60.93	60.17	58.21	66.98	61.03	71.04
	BBAVectors [52]	ResNet101	✓	88.35	79.96	50.69	62.18	78.43	78.98	**87.94**	90.85	83.58	84.35	54.13	60.24	65.22	64.28	55.70	72.32
	DRN [29]	Hourglass104	✓	89.71	82.34	47.22	64.10	76.22	74.43	85.84	90.57	86.18	84.89	57.65	61.93	69.30	69.63	58.48	73.23
	R^3Det [45]	ResNet152		89.49	81.17	50.53	66.10	70.92	78.66	78.21	90.81	85.26	84.23	61.81	63.77	68.16	69.83	67.17	73.74
	RSDet [30]	ResNet152		90.1	82.0	53.8	68.5	70.2	78.7	73.6	**91.2**	87.1	84.7	64.3	68.2	66.1	69.3	63.7	74.1
	RetinaNet-DCL (Ours)	ResNet152	✓	89.10	84.13	50.15	73.57	71.48	58.13	78.00	90.89	86.64	86.78	67.97	67.25	65.63	**74.06**	67.05	74.06
	R^3Det-DCL (Ours)	ResNet152		89.78	83.95	52.63	69.70	76.84	81.26	87.30	90.81	84.67	85.27	63.50	64.16	68.96	68.79	65.45	75.54
	R^3Det-DCL (Ours)	ResNet101	✓	89.14	83.93	53.05	72.55	78.13	81.97	86.94	90.36	85.98	86.94	66.19	65.66	73.72	71.53	68.69	76.97
	R^3Det-DCL (Ours)	ResNet152	✓	89.26	83.60	53.54	72.76	**79.04**	**82.56**	87.31	90.67	86.59	**86.98**	67.49	66.88	73.29	70.56	**69.99**	**77.37**

- Visualization



(a) ω = 180/4 (b) ω = 180/8

(a) ω = 180/4 (b) ω = 180/32 (c) ω = 180/128 (d) ω = 180/256