Dense Label Encoding for Boundary Discontinuity Free Rotation Detection

Xue Yang - Shanghai Jiao Tong University

X. Yang, et al. “Dense Label Encoding for Boundary Discontinuity Free Rotation Detection.” In CVPR21.

Virtual. 2021
Limitations of CSL

- **Issue 1:** Thick prediction layer
  \[ Th_{reg.} = A \]
  \[ Th_{onehot} = Th_{csl} = A \times AR/\omega \]

- **Issue 2:** Unfriendliness to small aspect ratio objects

---

**Case 3**

**Square-Like Problem**

The smaller the aspect ratio, the less sensitive the loss function is to angle.

**Long Edge Definition**

- Anchor/Proposal: (0,0,45,44,0°)
- Ground-Truth: (0,0,45,43,−60°)
- Predict box: (0,0,45,44,−60°)

\[ w \approx w \approx h \approx h, |\theta - 0| = 0° \]
\[ IoU < G, P > \approx 1 \]
\[ Smooth-L1 \text{ Loss} < G, P \approx 0 \]

- Anchor/Proposal: (0,0,45,44,0°)
- Ground-Truth: (0,0,45,43,−60°)
- Predict box: (0,0,45,44,30°)

\[ w \approx w \approx h \approx h, |\theta - 0| = 90° \]
\[ IoU < G, P > \approx 1 \]
\[ Smooth-L1 \text{ Loss} < G, P > \approx 0 \]
Densely Coded Label (DCL)

- Use Densely Coded Label (DCL) instead of Sparsely Coded Label (SCL) (for Issus 1)

\[
\begin{align*}
\text{Th}_{bcl} &= \text{Th}_{gcl} = A \times \left\lceil \log_2 \left( \frac{AR}{W} \right) \right\rceil \\
\text{Th}_{dcl} &= 
\end{align*}
\]

where A indicates the number of anchors. AR represents angle range. W indicates the angle discretization granularity.

- Example: A=21, AR=180, w=1

\[
\begin{align*}
\text{Th}_{\text{reg}} &= 21, \\
\text{Th}_{\text{onehot}} &= \text{Th}_{\text{csl}} = 3780, \\
\text{Th}_{\text{dcl}} &= 168
\end{align*}
\]

| Base Model       | ω | GFlops  | ∆GFlops | Params (M) | ∆Params  | 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           |
Densely Coded Label (DCL)

- Use Densely Coded Label (DCL) instead of Sparsely Coded Label (SCL) (for Issue 1)
Densely Coded Label (DCL)

- Use Densely Coded Label (DCL) instead of Sparsely Coded Label (SCL) (for Issus 1)
Angle Distance and Aspect Ratio Sensitive Weighting (for Issus 2)

\[ W_{ADARSW}(\Delta \theta) = | \sin(\alpha(\Delta \theta)) | = | \sin(\alpha(\theta_{gt} - \theta_{pred})) | \]

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

(a) Ground Truth  
(b) Prediction after using ADARSW
Ablation Experiments

- When angle discretization granularity $w$ is too small, too many angle categories, then classification affects performance.
- When angle discretization granularity $w$ is too large, the theoretical error is too large, thus the upper limit of performance is low.

| Method | $\omega$     | BR  | SV  | LV  | SH  | HA  | 5-mAP | mAP50 | mAP60 | mAP75 | mAP70:95 |
|--------|--------------|-----|-----|-----|-----|-----|-------|-------|-------|-------|----------|
| 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 |
|        | 180/64       | 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 | 58.91 | 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 |
| BCL    | 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 |
Ablation Experiments

(a) $\omega = 180/4$

(b) $\omega = 180/32$

(c) $\omega = 180/128$

(d) $\omega = 180/256$
Ablation Experiments

- Angle Distance and Aspect Ratio Sensitive Weighting

| 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      |
|        | ✓      |      |      |      |      |      |      |      |      |      |      | **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      |
|        | ✓      |      |      |      |      |      |      |      |      |      |      | **71.72**      | **67.02**  |

- Verification on different datasets

| Method       | ICDAR2015 | UCAS-AOD | MLT   |
|--------------|-----------|----------|-------|
|              | Recall    | Precision | Hmean | Recall | Precision | Hmean |
| RetinaNet-Reg | 81.49     | 83.29     | 82.38 |        |           |       |
|              | car(07/12) | plane(07/12) | mAP$_{50}$ (07) | mAP$_{50}$ (12) | Recall | Precision | Hmean |
| RetinaNet-CSL | 80.50     | **87.40** | 83.81 (+1.43) |        |           |       |
| RetinaNet-BCL | **81.61** | 84.79     | 83.17 (+0.79) |        |           |       |
|              | 88.15/92.35 | 90.57/97.86 | **89.36 (+0.51)** | 95.10 (+0.94) |        |           |       |
Ablation Experiments

- Visualization

(a) $\omega = 180/4$

(b) $\omega = 180/8$
Thank You!

- Paper: [https://arxiv.org/abs/2011.09670](https://arxiv.org/abs/2011.09670)
- Code: [https://github.com/yangxue0827/RotationDetection](https://github.com/yangxue0827/RotationDetection)
- Contact:
  - Xue Yang: [yangxue-2019-sjtu@sjtu.edu.cn](mailto:yangxue-2019-sjtu@sjtu.edu.cn)
  - Junchi Yan: [yanjunchi@sjtu.edu.cn](mailto:yanjunchi@sjtu.edu.cn)
- Homepage of our Lab:
  - [http://thinklab.sjtu.edu.cn/](http://thinklab.sjtu.edu.cn/)