Supplementary Material for Anchor Loss: Modulating Loss Scale based on Prediction Difficulty

Serim Ryou
California Institute of Technology

Seong-Gyun Jeong
CODE42.ai

Pietro Perona
California Institute of Technology

1. Anchor Design

In the paper, we set the anchor probability to the target class prediction score and modulate loss of the background class. Here we further study how to design anchor probability that affects behavior of the loss. We first define the basic formulation of anchor loss (AL) with sigmoid-binary cross entropy:

$$\ell(p, q; \gamma) = -\underbrace{(1 - q + q_{pos})^{\gamma_t} p \log(q)}_{\text{target class}} - \underbrace{(1 + q - q_{neg})^{\gamma_b} (1-p) \log(1-q)}_{\text{background class}}. \tag{1}$$

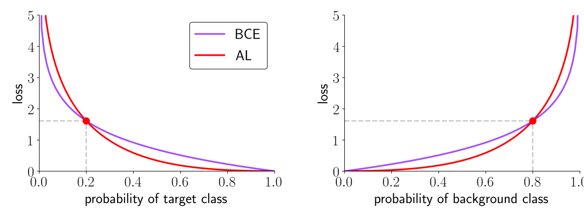
Anchor probability is a reference value for determining the prediction difficulty, which is defined as a confidence score gap between the target and background classes. The prediction difficulty is used to modulate loss values either by (i) pushing the loss of target class high, (ii) suppressing the loss of background classes, or (iii) using both ways around. The details of parameter setting for each case are as follows:

(i) **Modulate loss for target class**: We set the anchor probability to the maximum prediction score among background classes. Hence, target class loss gets more penalty when its score is lower than the anchor probability.

$$q_{pos} = \max_{i, \forall p_i = 0} q_i, \quad \gamma_t = \gamma \text{ and } \gamma_b = 0. \tag{2}$$

(ii) **Modulate loss for background classes**: We set the anchor probability to prediction score of the target class. Anchor loss is penalized more when output scores of the background classes are higher than the target.

$$q_{neg} = q_j, \text{ for } j, p_j = 1, \quad \gamma_t = 0 \text{ and } \gamma_b = \gamma. \tag{3}$$



(i) Modulate target loss ($q_{pos} = 0.2$) (ii) Modulate background loss ($q_{neg} = 0.8$)

Figure 1. How an anchor probability modulates loss values. When the prediction score of target class is lower than $q_{pos} = 0.2$, anchor loss penalizes more than binary cross entropy (i). On the contrary, when the prediction score of background class is higher than $q_{neg} = 0.8$, the loss value becomes higher than the binary cross entropy (ii).

(iii) **Modulate loss for both target and background classes**: We modulate loss on both directions by combining the above cases.

$$q_{pos} = \max_{i, \forall p_i = 0} q_i, \quad q_{neg} = q_j, \text{ for } j, p_j = 1, \quad \gamma_t = \gamma_b = \gamma. \tag{4}$$

We report image classification performance on CIFAR-100 by varying the way of designing anchor probability in Table 1. We achieve the best performance by modulating the loss for background classes (ii).

Table 1. Classification accuracies on CIFAR-100 with different anchor probabilities

loss fn.	Top-1	Top-5
BCE	73.88 $\pm$ 0.22	92.03 $\pm$ 0.42
(i)	74.06 $\pm$ 0.53	92.32 $\pm$ 0.24
(ii)	**74.25 $\pm$ 0.34**	**92.62 $\pm$ 0.50**
(iii)	73.90 $\pm$ 0.40	92.24 $\pm$ 0.06

Figure 2. Qualitative results for human pose estimation. Top row shows the output images with baseline (MSE) and bottom row represents the outcomes with anchor loss.

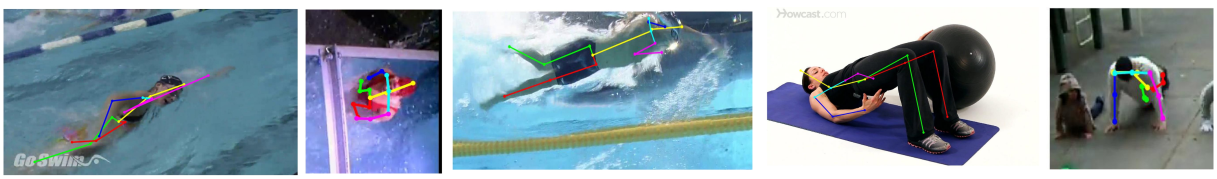


Figure 3. Failure cases on human pose estimation. Network trained with anchor loss still fails to detect correct body part locations when the body part is blurred or self-occluded.

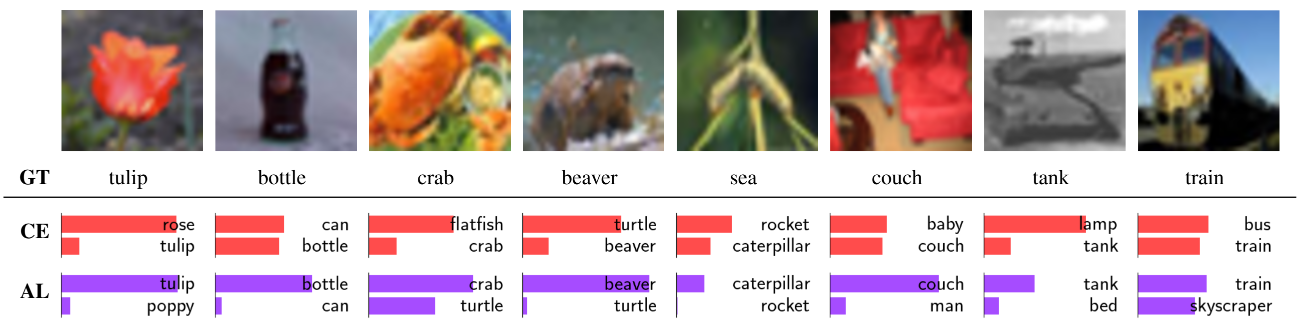


Figure 4. Image classification results on CIFAR-100. We compare the top-2 prediction scores of ResNet-110 with cross entropy (CE) and anchor loss (AL). Network trained with anchor loss successfully classifies difficult examples even though the model trained with cross entropy fails.

2. Qualitative figures

We visualize qualitative results for human pose estimation (Fig. 2, 3) and image classification (Fig. 4). Network trained with anchor loss has shown improvement over the baseline losses for both tasks. Specifically, anchor loss shows its potential use for multi-person pose estimation by finding correct body parts when the target person is occluded or overlapped by other person (last two columns of Fig. 2).