1-7

Chirality Created by Semi-Supervised Object Detection with Consistency Loss

Yuji YAMAUCHI Yoko UWATE and Yoshifumi NISHIO
(Tokushima University)

Introduction

Large datasets are essential to successful object detection. In addition, labeling objects can be very costly, time consuming, and labor intensive. For example, Caltech's pedestrian detection benchmark took about 400 hours to annotate 250,000 images. Semi-supervised learning uses unlabeled data in combination with box-level labeled data to reduce the cost of labeling. Among them, we are focusing on the method using consistency loss. In the model using consistency loss, a pair of unlabeled data and data that is flipped horizontally is used as input, and the difference between outputs is defined as consistency loss. Furthermore, it was found that visual cirality also occurs in the data expansion that has been used as a matter of course in order to prepare the large-scale data set required for object detection.

Therefore, we considered that the image with perturbation added in semi-supervised learning would have chirality, which would affect the accuracy of detection. We aim to create a semi-supervised object detection model that shows the latest accuracy, and in this study, as a first step, we will investigate the nature of semi-supervised learning using consistency loss.

Effect of perturbation on detection accuracy

Mathod. We compared and evaluated the detection accuracy due to the difference in perturbations applied to the input image during the training process in the semi-supervised learning using consistency loss. Supervised learning with labeled data (VOC07) only, semi-supervised learning with VOC07 and horizontally inverted unlabeled data (VOC12), semi-supervised learning with VOC07 and upside down VOC12, each learning The models that have been made are compared.

Results. Table 1 shows the detection accuracy by the model. From this result, it was found that the detection accuracy is affected by the difference in perturbation applied to the input image. We considered that one of the reasons for this result was that chirality influences detection accuracy.

Table 1: Detection accuracy.

Labeled data	Unlabeled data	mAP
VOC07	-	70.3
VOC07	VOC12(Horizontal)	72.6
VOC07	VOC12(Vertical)	72.1

Chirality created in the process of the model

Mathod. We investigate whether chirality can be seen in the inverted image generated in the process of the semi-supervised learning model using the above-mentioned consistency loss. Detection models trained with VOC07 or VOC07 (no-text) which leaves only the image without text, detects the inverted image in the test data.

Results. Table 2 shows the result of the detection model detecting an image in which an object with chirality is reflected. This result indicates that the inverted image generated in the process of the semi-supervised learning model that takes advantage of the inconsistency may have chirality. Figure 1 shows what part the detection model is looking at when it detects the inverted image. The red part shows the chirality more strongly. In Fig.1, you can see that the letters show strong chirality.

Table 2: Inverted image detection accuracy.

Training data	Test accuracy (%)
VOC07	81.2
VOC07(no text)	73.5

(a)Original (b)Flipped

Figure 1: Image with chirality-revealing regions highlighted.

Conclusion

In this research, the visual chirality that occurs in image inversion, which has been used in data expansion as a matter of course, is also applied to the perturbations added to images in semi-supervised learning, which may affect the detection accuracy. We wondered if this was the case and investigated the nature of semi-supervised learning using coherence loss.

The results showed that the difference in perturbations applied to the input image in the model affects the detection accuracy and that the inverted image generated in the process shows chirality.