

A Study of Object Detection Based on Weakly Supervised Learning

Guangyao Wang

School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, 454000, China.

ABSTRACT

Object detection is one of the important research contents in the field of computer vision. At present, the classical object detection methods can be divided into two categories: fully supervised-based target detection and weakly supervised-based target detection. Since the fully supervised object detection model requires a large number of training data with category labels and target bounding boxes, and such labeled data is difficult to obtain, it is of great significance to explore the weakly supervised object detection method that only needs category label data.

KEYWORDS

Weakly Supervised Learning; Object Detection; Detect the Model

1. INTRODUCTION

Since the beginning of the 20th century, object detection has attracted the interest of a large number of researchers, and has been focused on and studied, and the main purpose of object detection is to find out as many targets as possible and determine the location and size of the target on the basis of judging and identifying whether there is an object in the image [1]. Due to the different degrees of differences in shape, position and clarity of various targets, and the influence of daylight or other different factors when forming target images, the research of object detection tasks has always been very challenging.

2. OBJECT DETECTION METHODS FOR WEAKLY SUPERVISED LEARNING

In the common methods of object detection, the bounding box annotation of the target and the corresponding category annotation are required to optimize and train the parameters of the neural network, and it has gradually developed into the main model architecture of object detection, and the accuracy of its detection results has also been greatly improved. With the rapid development of Internet-related technologies and the widespread application of relevant principles in the field of computer vision, object detection, as a highly discussed research task, has a wide range of application value in various fields such as intelligent video surveillance systems, high-tech military target detection, traffic management detection systems, face detection, and medical treatment [2]. For example, the object detection model is used to detect faces in video surveillance or digital cameras; Help doctors to complete the CT lesion detection in medical imaging; Object detection is combined with satellite imagery or on-board camera imagery to detect road conditions, various obstacles on the road, and much more. Object detection is mainly divided into three forms: strongly supervised,

unsupervised[3] and weakly supervised, and most of the current object detection is based on strongly supervised learning, and the accuracy of the detection results based on strongly supervised learning object detection will be affected by the accuracy of the bounding box annotation of the detected target to a high extent, and the results of the target bounding box annotation are easily affected by subjective judgment, and it is extremely time-consuming and expensive to obtain a large number of accurate bounding box annotations in strongly supervised learning. It takes a lot of time and energy of professionals to meet the needs of strongly supervised learning for datasets through repeated operations, sample data collection, data collation, and subsequent data labeling. In addition, there will be various changes in the data in some tasks through the evolution of the task, resulting in the need for the data to be constantly re-labeled, so the applicability of strong supervised learning is also limited. Therefore, in order to solve the problem of limited applicability of strongly supervised learning, a series of weakly supervised learning object detection methods that only need to provide image-level labels of the target are proposed in the research process of object detection.

In order to solve the problems in strongly supervised learning, researchers have proposed a series of weakly supervised object detection algorithms that only need to provide only image-level labeled datasets to detect targets. Weakly supervised learning only needs to indicate whether a certain type of object is included in the dataset, which is easier to obtain labels than strongly supervised learning, which solves the problem of difficult labeling and expensive cost of strongly supervised learning object detection. In 2016, a weakly supervised depth detection network architecture, WSDDN, was proposed in Ref. [4], which is stronger than other weakly supervised detection models on the PASCALVOC dataset. In 2019, Ref. [5] proposed a sequential label propagation and augmentation network model, Label-PEnet, which realizes object detection and instance segmentation in images by gradually converting image-level labels into pixel-level labels. In the same year, in order to solve the problem that the model only focuses on the most discriminating part of the target in the image rather than the whole object, an end-to-end trainable network model was proposed in Ref. [6], and an object detection module and a perceptual triplet state loss were added to enhance the ability of weakly supervised learning to detect foreground objects. In 2020, Ref. [7] proposed a method to generate proposal clusters, and then learned from the refined instance classifier through an iterative process, which realized the weakly supervised object detection of datasets with only image-level annotations. In 2018, Ref. [8] mainly used domain migration technology to convert the source domain data in the source domain into images with example labels, or to generate corresponding pseudo-label labels for the target domain data through pseudo-labeling. In 2019, a continuous multi-example learning model was proposed in Ref. [9], which continuously optimizes the multi-example function through the smoothing loss function to alleviate the non-convexity problem of MIL. However, most of the labeling of weakly supervised learning is incomplete, inaccurate or inaccurate, so there is still a lot of room for improvement in the accuracy of object detection algorithms based on weakly supervised learning.

3. CONCLUDING REMARKS

Although the object detection model based on weakly supervised learning has achieved very significant research results, due to the problems of incomplete, inaccurate or inaccurate annotation in the dataset data required by weakly supervised learning, the object detection model based on weakly supervised learning is still imperfect in terms of detection performance, so it is necessary to make further research on the object detection model based on weakly supervised learning to improve the accuracy of weakly supervised object detection. The conduct of this research topic is not only of great significance in academic research, but also has a profound positive effect in the application of various fields in practical life.

REFERENCES

- [1] X. L. Zhou, X. J. Chen, S. Y. Chen and B. J. Lei, Weakly Supervised Learning-based Object Detection: A Survey. *Computer Science*,2019,46(11):49-57.
- [2] S. F. Zhang, Infrared small target detection and tracking based on background suppression. Huazhong University of Science and Technology,2017.
- [3] B. K. Tripathy, S. Anveshritha and S. Ghela. *Unsupervised Learning Approaches for Dimensionality Reduction and Data Visualization*. CRC Press:2021-04-15.
- [4] X. W. Xie, J. F. Shi, Research on multi-objective image detection based on weakly supervised convolutional neural network. *Journal of Electronic Measurement and Instrumentation*,2019,33(06):31-37.
- [5] H. Bilen, A. Vedaldi, Weakly Supervised Deep Detection Networks// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2016:2846-2854.
- [6] W. Ge, W. Huang, S. Guo and M. Scott, "Label-PEnet: Sequential Label Propagation and Enhancement Networks for Weakly Supervised Instance Segmentation," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 3344-3353, doi: 10.1109/ICCV.2019.00344.
- [7] Y. Chen, W. H. Hsu, "Saliency Aware: Weakly Supervised Object Localization," ICASSP 2019 - 2019 IEEE International Conference on Acous.
- [8] P. Tang et al., "PCL: Proposal Cluster Learning for Weakly Supervised Object Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 1, pp. 176-191, 1 Jan. 2020, doi: 10.1109/TPAMI.2018.2876304.
- [9] N. Inoue, R. Furut and T. Yamasaki, K. Aizawa, "Cross-Domain Weakly-Supervised Object Detection Through Progressive Domain Adaptation," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 5001-5009.