Few-shot Object Detection with Camouflage Animals

Yi Lei*, Lei Lei

1Network and Data Security Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China
2Primary school in Gao sheng, Zi yang, China
*Corresponding author’s e-mail: yilei@std.uestc.edu.cn

Abstract. Object detection is a hot issue in the field of computer vision, which is widely used in intelligent video surveillance, medical image analysis, and practice in the fields of military strategy. Previous object detection is based on a lot of training data, and for some special target, it is difficult to find enough annotation data to training model, so few-shot object detection was born. In this paper, a data set of animals with protective colors is constructed(CAD), which contains ten categories of different objects with high quality annotations. We combine the existing attention RPN and Feature Reweighting module to train our net, the experimental results prove that the Camouflage animals in nature can be well identified, and our results obtain 1.9% AP50, 1.2% AP75 improvement compared with before work.

1. Introduction
Object detection can be used for many tasks, such as intelligent video surveillance, medical image analysis, few shot object detection[1,2,3,4] is a hot direction in recent years. Over the past few years, a number of deep learning approaches have achieved remarkable performance in few shot object detection, such as: LSTD[7], RepMet[8], Feature Reweighting[9]. This paper studies the target detection of animals with protective colors, aiming to detect animals with protective colors in nature, and then it can be used in the military, hunters can use this method to find animals hidden in the grass. Therefore, it is of practical significance to study this topic.

2. Related work

2.1. LSTD
Chen et al.(2018) proposed a new few shot transfer detector, where they leveraged rich source-domain knowledge to create an useful target-domain detector, only use a few training examples. They designed a flexible deep architecture to alleviate transfer difficulties in few shot object detection, which integrated the advantages of both SSD and Faster RCNN. They also created a new regularized transfer learning framework for few shot object detection, where the background depression regularizations and transfer knowledge are proposed to leverage object knowledge respectively from source and target domains, which is helpful to enhance fine-tuning with a few object pictures[7].

2.2. RepMet
Leonid et al.(2019) proposed a new method for distance metric learning, it can learns the multi-modal distribution of each of the training categories in that space, the embedding space, and the backbone...
network parameters at the same time. They also offered a novel episodic benchmark based on the ImageNet dataset for few shot object detection task[8].

2.3. Feature Reweighting
Kang et al.(2019) developed a new object detector about few-shot detection, which can learn to detect new objects, and no fine-tuning needed, they created model by high quality labeled base classes, which can detected to novel classes quickly[12].

Table 1. This dataset is of great diversity with large variance.

|                     | Train       | Test |
|---------------------|-------------|------|
| No.Image            | 66520       | 10   |
| No.Box              | 182591      | 10   |
| No.Class            | 1000        | 10   |
| Avg No.Box / Img    | 2.74        | 1    |
| min N.img/C         | 22          | 1    |
| max N.img/C         | 208         | 1    |
| avg N.img/C         | 74.98       | 1    |

3. Materials and methods

3.1. Materials
Our model is finished on 2 NVIDIA 1080ti gpus. We found that pre-training model on both ImageNet[10] and MS COCO[11] can get great results. We used AP, AP50 and AP75 for evaluation.

3.2. Camouflage animals Dataset
In this paper, We created a new dataset about Camouflage animals, which contains ten categories of different objects with high quality annotations[5]. Detailed statistics are shown in Table 1. We are among the first to study the problem about Camouflage animals of few shot object detection, which is of great practical values but a less explored task[6].

3.3. Attention RPN
In few shot object detection, RPN is of great importance to detection. We made the support feature as $X \in t^{S \times S \times C}$ and the map of the query as $Y \in t^{H \times W \times C}$, which defined as

$$G_{h,w,c} = \sum_{m,n} X_{m,n,c} \cdot Y_{h+m-1,w+n-1,c}, \quad m, n \in \{1, ..., S\}$$

the resultant attention feature map is G, and X is used to slide on the query feature map in a depth-wise cross correlation manner. Figure 1 shows Attention RPN structure[12].
4. Results
Few-shot object detection with camouflage animals, the difficulty lies in the target object and the background has a high similarity, it is difficult to be detected. This paper is based on attention RPN and multiple relations detector, using fsod as training dataset and our camouflage animal dataset as testing dataset, experiments show that animals with camouflage in nature can be accurately identified. Refer to Figure 2 and Table 2. We trained our model by different dataset, and our results obtain 1.9% AP50, 1.2% AP75 improvement compared with before work.
Table 2. Experimental results on COCO and FSOD dataset for Camouflage Animals.

| Method     | Dataset        | AP50 | AP75 |
|------------|----------------|------|------|
| LSTD[7]    | COCO           | 37.4 | -    |
| RepMet[8]  | COCO           | 39.6 | -    |
| Ours       | COCO           | 42.1 | 20.8 |
| Ours       | FSOD           | 44.3 | 31.2 |
| Ours       | Camouflage Animals | 46.2 | 33.4 |

5. Conclusion

We created a new Camouflage Animals dataset, which contains ten kinds of different objects with great annotations. This model trained on FSOD and tested on CAD, which can detect objects of camouflage animals well. In these years, Few shot object detection is a useful research topic, but about the image detection that has a great similarity with the background image, there is little work at present, so we need to continue to work hard.

References

[1] Fan, Q., Zhuo, W., Tang, C. K., Tai, Y. W. (2020). Few-Shot Object Detection With Attention-RPN and Multi-Relation Detector. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 4012-4021.
[2] Dong, X.Y., Zheng, L., Ma, F., Yang, Y., Meng, D.Y. (2018). Few-example object detection with model communication. IEEE transactions on pattern analysis and machine intelligence, 41(7):1641-1654.
[3] Pedro F., Ross B G., David M. A., Deva R. (2010). Object detection with discriminatively trained part-based models. IEEE transactions on pattern analysis and machine intelligence, 32(9):1627-1645.
[4] Ranjay K., Yuke Z., Oliver G., Justin J., Kenji H., Joshua K., Stephanie C., Yannis K., Li, L. J., David A. S., etal. (2017) Visual genome: Connecting language and vision using crowdsourced dense image annotations. International Journal of Computer Vision, 123(1):32 – 73.
[5] George A. M. (1995). Wordnet: a lexical database for english. Communications of the ACM, 38(11):39 - 41.
[6] Chelsea, F., Pieter, A., Sergey, L. (2017). Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, PMLR 70.
[7] Chen, H., Wang Y. L., Wang G. Y., Qiao, Y. (2018). LSTD: A Low-Shot Transfer Detector for Object Detection. AAAI. 2836-2843.
[8] Karlinsky, L., Shtok, J., Harary, S., Schwartz, E., & Bronstein, A. M. (2019). RepMet: Representative-Based Metric Learning for Classification and Few-Shot Object Detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 5197-5206.
[9] Kang, B., Liu, Z., Wang, X., Yu, F., & Darrell, T. (2018). Few-shot object detection via feature reweighting.CVPR. 8420-8429.
[10] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: a large-scale hierarchical image database. Proc of IEEE Computer Vision & Pattern Recognition, 248-255.
[11] Lin, T. Y., Maire, M., Belongie, S., Hays, J., & Zitnick, C. L. (2014). Microsoft coco: common objects in context. ECCV.740-755.
[12] Wang, Y., Tang, Y., Ma, J., & Zhen, Q. (2015). Gender Prediction Based on Data Streams of Smartphone Applications. International Conference on Big Data Computing and Communications. Springer, Cham.115 – 125.