The Eleventh International Conference on
Engineering Computational Technology
23–25 August, 2022 | Montpellier, France

Proceedings of the Eleventh International Conference on
Engineering Computational Technology
Edited by B.H.V. Topping and P. Iványi
Civil-Comp Conferences, Volume 2, Paper 7.4
Civil-Comp Press, Edinburgh, United Kingdom, 2022, doi: 10.4203/ccc.2.7.4

Analysis of Object Detection Datasets for Machine Learning with Small and Tiny Objects

L. Fichtel, D. Erbacher, L. Heller, A. Frühwald, L. Hösch and C. Bachmeir

University of Applied Sciences Würzburg-Schweinfurt, Germany

Abstract

Deep Learning models are trained to detect humans, cars, and other large objects which are centered in the images. The same models struggle with detecting small and tiny objects because of architecture design decisions that reduce the entropy of small and tiny objects during training. These small and tiny objects are essential for damage identification and maintenance including inspection and documentation of aeroplanes, constructions, offshore structures, and forests. Our work defines the terms tiny and small in context of deep learning models to evaluate possible approaches to resolve the issue of low accuracy in detecting these objects. We analyse the currently applied common datasets Common Objects in Context, ImageNet and Tiny Object Detection Challenge dataset. In addition we compare these datasets and present the differences in terms of object instance size. The COCO dataset, ImageNet dataset and TinyObjects dataset are analysed regarding size categorization and relative object size. The results show the large differences between the size ratios of the three chosen datasets, with ImageNet having by far the largest object instances, COCO being in the middle and TinyObjects having the smallest objects as its name would indicate. Since the objects themselves are larger in terms of total pixel width and height, they therefore make up a bigger percentage on the superordinate picture. Looking at the size categories of the COCO dataset and our extension of the tiny and very small category, the results confirm the size hierarchy of the datasets. With ImageNet having most of its objects in the large category, COCO respectively in the medium category and TinyObjects in the very small category. By taking these results into account, the reader is able to choose a fitting dataset for their tasks. We expect our analysis to help and improve future research in the area of small and tiny object detection.

Keywords: machine learning , damage identification, object detection, inspection, tiny objects, small objects, maintenance , dataset

1 Introduction

Since the general goal of object detection in computer vision is the determination of object instances including their location, size, and object category, it is very important to be able to detect said objects in an image [1, 2]. In the last few years, researchers have begun building general purpose object detection systems with the focus of rivalling the human object detection ability, instead of focusing on the detection of a single or a few specific object instance categories. At the same time, there is another challenge that has been less studied: detecting small and tiny objects in images [3]. It is very difficult to detect tiny objects because of different factors such as a lack of datasets as well as the fact that tiny objects often get filtered out while going through the different neural network layers. The reason for this loss of information is the fact that tiny objects are represented by only a few pixels, which makes it consequently more difficult to distinguish them from background noise. Furthermore, an accurate localization is more difficult and requires object detection methods to have higher precision requirements. The experience and knowledge of small object detection is very limited, as most efforts to date have focused on large object detection [4]. However, the detection of these tiny objects is a very important challenge with economic use cases such damage identification and maintenance including inspection and documentation of aeroplanes, constructions, offshore structures, and forests [3,5,6].

For an AI model to generalize well and work safely when used, it needs a data set that correctly represents its task in as many situations as possible. Collecting sufficient training data is often expensive, time-consuming, or even unrealistic in many scenarios. Following the amount of relevant data is still a problem for the use of AI. A solution for this problem can be transfer learning. Transfer learning has the advantage of higher accuracy even on small data sets and faster training due to a better baseline.

For this reason, this work concentrates on the comparison between common datasets in terms of their respective object sizes.

2 Methods

There are multiple research papers that thematize datasets for common object detection and therefore serve as base for this scientific work. It's important to understand the size tendencies of these popular datasets.

Tuggener et al. argue that the common datasets focus on images with "a few large objects".

Wang et al. take a look at tiny object detection in areal images and examine the occurring problems like a lower pixel count for example [7]. The result of very small objects with a low pixel density is that some of these tiny objects will be filtered out in the down sampling progress (i.e. Faster R-CNN) and therefore are missing in the final feature map.

In contrast to scientific work that often compares datasets under specific conditions, this work focuses on comparing the different datasets based on their object instances' total and relative size as well as their position on the respective image, regardless of the object instance category.

We selected three datasets for our analysis.

The Common Objects in Context dataset, in short COCO, was created for the purpose of scene understanding and therefore consists of images that depict the natural context of day-to-day scenes [8]. The dataset provides data containing classification and object localization in the form of bounding boxes and polygons for segmentation. For the analysis in this paper, we analysed 118 000 pictures and 850 000 object instances, which results in approximately 7.2 objects per image.

The ImageNet dataset is an image collection organized according to the WordNet hierarchy. We analysed 1 073 728 pictures with a total of 1 564 348 object instances resulting in an average of approximately 1.46 object instances per image [9].

The TinyObjects dataset is an image collection consisting of real-world images recorded from UAVs. Because of the high distance between air vehicle and the ground, the recognized objects are rather small (as the name of the dataset indicates) [10]. The differences between angles and distances towards the recorded people results in a complex diversity. We analysed 794 pictures with a total of 37 288 object instances resulting in an average of approximately 47 object instances per image. Similarly, to the COCO dataset, the TinyObjects dataset gives detailed segmentation data in form of polygons for each object. However, the ImageNet dataset does only provide annotation via bounding boxes, so the analysis of this dataset focusses only on the bounding box annotation.

3 Results

The results of the analysis are shown in Figure 1. 34.23% of COCO objects belong to the category "Medium". The small and large categories are represented with 22.27% and 23.96% respectively. 14.05% of the objects belong to the category very small. The remaining 5.49% of the objects are tiny.

TinyObjects mostly consists of ”very small” objects (38.45%). The tiny and small categories are represented with respectively 29.42% and 20.08%, while 11.31% of the objects can be categorized as ”medium” and the remaining 0.74% of the objects are in the ”large” category.

Imagenet’s tiny and small categories only are respectively 0.10% and 1.17% and only 4.08% of the object instances can be categorized as ”small”. 17.21% of the objects can be described as ”medium” according to our categorization. By a large margin, 77.44% of all ImageNet objects are larger than 96x96 pixels and can therefore be called ”large”.

20% of the Coco objects are smaller than 0.1% of the image on which they can be found. Nearly 38% of the objects are between 0.1% and 1% of total image size and 16.3% lie between the 1% and 3% size window. More than 74% of the objects are smaller than 3% in relation. The percentage of objects between 3% and 10% of the original image is 14.07%. Objects that are larger than 10% of the of superordinate image account for only 11.51%.

63% of TinyObjects objects are smaller than 0.01% of the images on which they can be found, while there are 16% of objects that are between 0.01% and 0.02% and 6.72% that lie between the 0.02% and 0.03% size window. The percentage of objects that are between 0.03% and 0.05% of the original image is 6.06% with approximately 8.33% of objects that are larger than 0.05%.

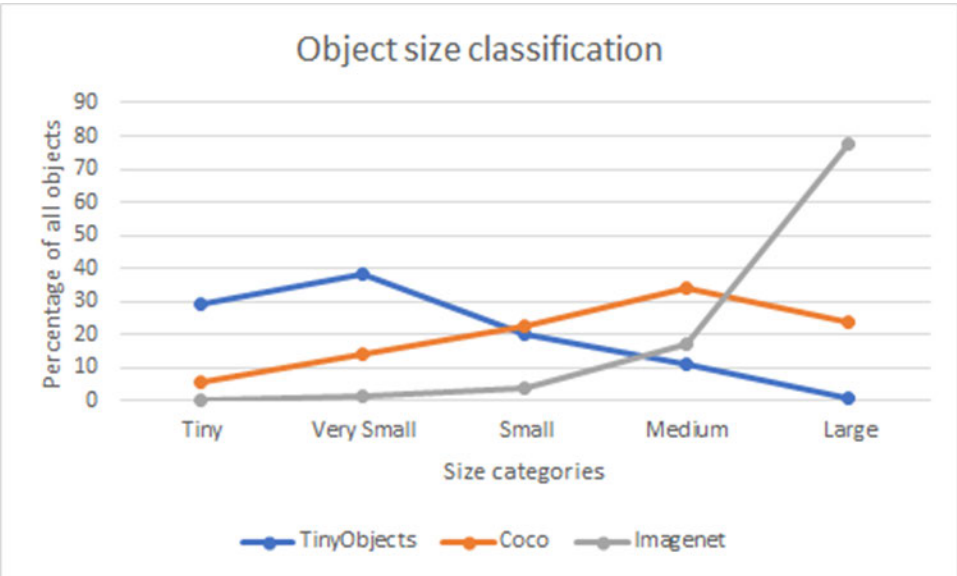

Figure 1: Percentual Representation of Tiny, Very Small, Small, Medium and Large Objects for the
TinyObjects (blue), Coco (orange) and Imagenet (grey) dataset.

7.78% of the Imagenet objects are smaller than 1% of the images they are located on. There are 12.19% objects that are between 1% and 5% while 8.12% belong to the 5% to 10% window. Approximately 28% of the objects are smaller than 10% in relation to their parent image. 16.80% of all objects make up more than 10% and less than 25% of the respective image size while 21.58% of all instances lies between 25% and 50%. Another 24,36% of objects make up over 50% and 90% of their parent image. While there are 9.17% of object instances that make up more than 90% of the superordinate image.

4 Conclusions and Contributions

Our data analysis presents the representation of small and tiny objects in common object detection datasets.

The COCO dataset, ImageNet dataset and TinyObjects dataset are analysed regarding size categorization and relative object size.

The results show the large differences between the size ratios of the three chosen datasets, with ImageNet having by far the largest object instances, COCO being in the middle and TinyObjects having the smallest objects as its name would indicate. Since the objects themselves are larger in terms of total pixel width and height, they therefore make up a bigger percentage on the superordinate picture.

Looking at the size categories of the COCO dataset and our extension of the tiny and very small category, the results confirm the size hierarchy of the datasets. With ImageNet having most of its objects in the large category, COCO respectively in the medium category and TinyObjects in the very small category.

By taking these results into account, the reader is able to choose a fitting dataset for their tasks.

We expect our analysis to help and improve future research in the area of small and tiny object detection.

Acknowledgements

The project ‘AI-Inspection Drone’ (Project Number 20X1707) is funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) of Germany. We are thankful for this support.

References

[1] M. Everingham, L. Gool, C. K. Williams, J. Winn, A. Zisserman, “The pascal visual object classes(voc) challenge, ”Int. J. Comput. Vision, vol. 88,no. 2, p. 303–338, Jun. 2010. doi:10.1007/s11263-009-0275-4

[2] O. Russakovsky, J. Deng, H. Su, J. Krause,S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla,M. Bernstein, A. Berg, L. Fei-Fei, “Imagenet large scale visual recognition challenge,”Interna-tional Journal of Computer Vision, vol. 115, 2014.

[3] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen,X. Liu, M. Pietikainen, “Deep learning for generic object detection: A survey, ”International journal of computer vision, vol. 128, no. 2, 261–318, 2020.

[4] K. Tong, Y. Wu, “Recent advances in small object detection based on deep learning: A review, ”Image and Vision Computing, vol. 97, 2020.

[5] L. Fichtel, A. M. Frühwald, L. Hösch, V. Schreibmann, C. Bachmeir, F. Bohlander. Tree Localization and Monitoring on Autonomous Drones employing Deep Learning. Conference of Open Innovation Association, FRUCT, pp. 132–140. 2021. doi:10.23919/FRUCT52173.2021.9435549

[6] R. Ashour, T. Taha, F. Mohamed, E. Hableel, Y. Kheil, M. Elsalamouny, M. Kadadha, K. Rangan, J. Dias, L. Seneviratne, G. Cai, Site inspection drone: A solution for inspecting and regulating construction sites. 1-4., 2016. doi:10.1109/MWSCAS.2016.7870116

[7] J. Wang, W. Yang, H. Guo, R. Zhang, G.-S.Xia, “Tiny object detection in aerial images,” in 2020 25th International Conference on Pattern Recognition(ICPR), pp. 3791–3798, 2021.

[8] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B.Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollar, C. L. Zitnick, “Microsoft COCO: commonobjects in context,”CoRR, vol. abs/1405.0312, 2014.

[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, “Imagenet: A large-scale hierarchical im-age database,” in 2009 IEEE Conference on ComputerVision and Pattern Recognition, 248–255, 2009.

[10] X. Yu, Z. Han, Y. Gong, N. Jiang, J. Zhao, Q. Ye, J. Chen, Y. Feng, B. Zhang, X. Wang, Y. Xin, J. Liu, M. Mao, S. Xu, B. Zhang, S. Han, C. Gao, W. Tang, L. Jin, M. Hong, Y. Yang, S. Li, H. Luo, Q. Zhao, H. Shi “1st tiny object detection challenge: Methods and results,”CoRR, vol. abs/2009.07506, 2020.