

Object Detection and Classification with Limited Training Data

Patrick Wong ^{1,*}  and Yifan Zhao ² 

¹ School of Computing and Communications, Faculty of Science, Technology, Engineering and Mathematics, Open University, Walton Hall, Milton Keynes MK7 6AA, UK

² Faculty of Engineering and Applied Sciences, Cranfield University, College Road, Cranfield, Bedfordshire MK43 0AL, UK; yifan.zhao@cranfield.ac.uk

* Correspondence: patrick.wong@open.ac.uk

1. Introduction

Since the rise of deep learning around a decade ago, the field of object detection and classification using a convolutional neural network (CNN) and its variants has grown exponentially. This technology has been applied in domains such as the manufacturing, construction, surveillance and monitoring, sports, transports, and medical sectors. An attractive quality of CNNs is their ability to take images in their raw form without the traditional feature extraction step of reducing input dimensionality. However, a significantly larger amount of training samples is required for CNNs to extract features and automatically classify objects. Labelling a large amount of data samples is costly and time-consuming, and as a result, the availability of training data is limited, particularly in domains outside of the science sectors. The labelling quality can also impact classification performance and reliability as wrong or erratic labelling can lead to poor or biased classification performance. The availability of quality data has become a bottleneck, and it hinders CNNs' use in wider applications.

2. Learning with Limited Training Data

To address this bottleneck, the research community has developed various strategies to reduce the reliance on large amounts of training data. Data augmentation and data generation is one approach to produce more data from existing data. While data augmentation produces more variations of data by transforming the existing data in different ways, data generation achieves this by using generative models such as Generative Adversarial Networks [1]. However, as the produced data are derived from the original data, they do not often contain the extra features found in the original data samples.

Transfer learning [2] is a different approach to address the data scarcity bottleneck. It re-trains some parts, often the outer layers, of an existing model that was designed for classifying different but related objects. As the inner layers of the existing model have already learnt to recognise certain common features between old and new objects, less data are required to train the model to classify new objects.

Based on a similar idea, n-shot learning [3] aims to learn how to recognize new objects with just one or a few samples. It typically involves a meta learning process which aims to help a model quickly learn new tasks by training it on a variety of tasks with a few data samples. Training a model to predict the similarity between data points helps it generalise well with new tasks. Models also commonly use an embedding method in which the model learns to map inputs into a space where similar items are situated nearby. The model is often fine-tuned for a specific task with a small dataset through transfer learning.



Received: 10 June 2025
Accepted: 16 June 2025
Published: 18 June 2025

Citation: Wong, P.; Zhao, Y. Object Detection and Classification with Limited Training Data. *Appl. Sci.* **2025**, *15*, 6842. <https://doi.org/10.3390/app15126842>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

3. Conclusions

With advances in data generation and transfer and meta learning, it is becoming feasible to train a model with limited labelled data. This will enable object detection to be applied to domains in which this was previously impossible due to a lack of labelled data. However, regarding computer vision, current object detection applications are still merely focused on detecting specific objects in images or videos rather than attempting to understand the holistic view of the image that is being represented. The first step of understanding an image is knowing what objects are in the image. However, understanding the relationships and interactions between these objects and predicting their future actions and behaviours, for example, are also necessary to reveal the deeper context of an image. Many more studies are needed to bring computer vision closer to human vision.

Author Contributions: Writing—original draft preparation, P.W.; writing—review and editing, Y.Z. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Goyal, M.; Mahmoud, Q.H. A Systematic Review of Synthetic Data Generation Techniques Using Generative AI. *Electronics* **2024**, *13*, 3509. [[CrossRef](#)]
2. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. *A Comprehensive Survey on Transfer Learning*; IEEE: Piscataway, NJ, USA, 2019. [[CrossRef](#)]
3. Parnami, A.; Lee, M. Learning from Few Examples: A Summary of Approaches to Few-Shot Learning. *arXiv* **2022**, arXiv:2203.04291.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



DSPACE

<https://dspace.org/>

Object detection and classification with limited training data

Wong, Patrick; Zhao, Yifan

2025-06-18

MDPI

<https://doi.org/10.3390/app15126842>

Wong P, Zhao Y. (2025) Object detection and classification with limited training data. Applied Sciences, Volume 15, Issue 12, June 2025, Article number 6842