Hierarchical Novelty Detection for Visual Object Recognition

Cheng Guan

August 5, 2018

Abstract

Deep neural networks have achieved impressive success in large-scale visual object recognition tasks with a predefined set of classes. However, recognizing objects of novel classes unseen during training still remains challenging. The problem of detecting such novel classes has been addressed in the literature, but most prior works have focused on providing simple binary or regressive decisions, e.g., the output would be "known," "novel," or corresponding confidence intervals. In this paper, the authors study more informative novelty detection schemes based on a hierarchical classification framework. For an object of a novel class, they aim for finding its closest super class in the hierarchical taxonomy of known classes. To this end, they propose two different approaches termed top-down and flatten methods, and their combination as well.

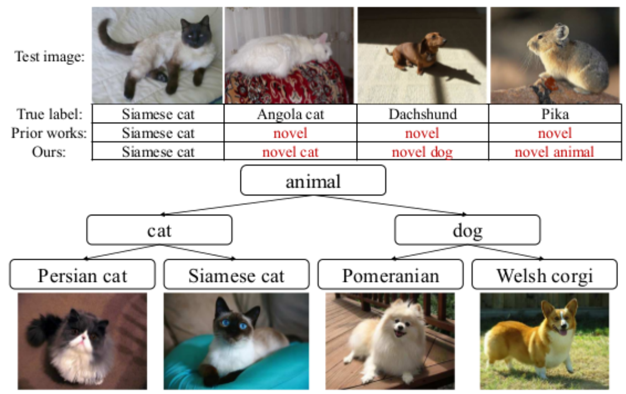

Figure 1. An illustration of their proposed hierarchical novelty detection task. In contrast to prior novelty detection works, the authors aim to find the most specific class label of a novel data on the taxonomy built with known classes.

1. Introduction

Object recognition in large-scale image datasets has achieved impressive performance with deep convolutional neural networks (CNNs) [4, 5, 7, 8]. The standard CNN architectures are learned to recognize a predefined set of classes seen during training. However, in practice, a new type of objects could emerge (e.g., a new kind of consumer product). Hence, it is desirable to extend the CNN architectures for detecting the novelty of an object (i.e., deciding if the object does not match any previously trained object classes). There have been recent efforts toward developing efficient novelty detection methods [1], but most of the existing methods measure only the model uncertainty, i.e., confidence score, which is often too ambiguous for practical use. For example, suppose one trains a classifier on an animal image dataset as in Fig. 1. A standard novelty detection method can be applied to a cat-like image to evaluate its novelty, but such a method would not tell whether the novel object is a new species of cat unseen in the training set or a new animal species.

To address this issue, the authors design a new classification framework for more informative novelty detection by utilizing a hierarchical taxonomy, where the taxonomy can be extracted from the natural language information, e.g., WordNet hierarchy [6]. Their approach is also motivated by a strong empirical correlation between hierarchical semantic relationships and the visual appearance of objects. For example, as illustrated in Fig. 1, Their goal is to distinguish "new cat," "new dog," and "new animal," which cannot be achieved in the standard novelty detection tasks. they call this problem *hierarchical novelty detection* task.

In contrast to standard object recognition tasks with a closed set of classes, Their proposed framework can be useful for extending the domain of classes to an open set with taxonomy information (i.e., dealing with any objects unseen in training). In practical application scenarios, Their framework can be potentially useful for automatically or interactively organizing a customized taxonomy (e.g., companys product catalog, wildlife monitoring, personal photo library) by suggesting closest categories for an image from novel categories (e.g., new consumer products, unregistered animal species, untagged scenes or places).

2. Related Work

Novelty detection. For robust prediction, it is desirable to detect a test sample if it looks unusual or significantly differs from the representative training data. Novelty detection is a task recognizing such abnormality of data A confidence score about novelty can be measured by taking the maximum predicted probability , ensembling such outputs from multiple models , or synthesizing a score based on the predicted categorical distribution.

Object recognition with taxonomy. Incorporating the hierarchical taxonomy for object classification has been investigated in the literature, either to improve classification performance, or to extend the classification tasks to obtain more informative results [3].

Generalized zero-shot learning (GZSL). they remark that GZSL [2] can be thought as addressing a similar task as theirs. While the standard ZSL tasks test classes unseen during training only, GZSL tasks test both seen and unseen classes such that the novelty is automatically detected if the predicted label is not a seen class.

References

[1] A. Bendale and T. E. Boult. Towards open set deep networks. In *CVPR*, 2016. 1

[2] W.-L. Chao, S. Changpinyo, B. Gong, and F. Sha. An empirical study and analysis of generalized zero-shot learning for object recognition in the wild. In *ECCV*, 2016. 2

[3] J. Deng, J. Krause, A. C. Berg, and L. Fei-Fei. Hedging your bets: Optimizing accuracy-specificity trade-offs in large scale visual recognition. In *CVPR*, 2012. 2

[4] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *ICCV*, 2015. 1

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 1

[6] G. A. Miller. WordNet: A lexical database for english. *Communications of the ACM*, 1995. 1

[7] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1

[8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *CVPR*, 2015. 1