Features for Nearly-Infallible Classification

Jonathan Krause
Computer Science Department, Stanford University
Stanford, CA
jkrause@cs.stanford.edu

Future Distribution Permission

The author(s) of this report give permission for this document to be distributed to Stanford-affiliated students taking future courses.

1. Introduction

Within the field of object classification, the goal of nearly-infallible classification is to maintain an accuracy level of $1 - \epsilon$. For relatively small ϵ on any dataset of reasonable size, this is not feasible with current state-of-the-art classifiers. However, if a semantic hierarchy exists, then one can achieve $1 - \epsilon$ accuracy by predicting that certain examples are members of more general classes, such as "Dog" instead of "German Shepherd" and "Object" instead of "Animal". The goal of this work is both to examine the efficacy of such a prediction scheme under varying choices of features and to incorporate today's state-of-the-art classification algorithms into an infallible system.

1.1. Reward Metric

Formally, the objective for (strong) infallibility is

$$\begin{aligned} \underset{f(x)}{\text{maximize}} \quad & \mathop{\mathbb{E}}_{x,y} r_{f(x)}[f(x) \in \pi(y)] \\ \text{subject to} \quad & \forall x \mathop{\mathbb{E}}_{x,y}[f(x) \in \pi(y)] \geq 1 - \epsilon \end{aligned}$$

where x is an input image, y is its correct class as a leaf node in the semantic hierarchy, $r_{f(x)}$ is a reward associated with class $f(x)$, and $\pi(y)$ is the path from y to the root node in the hierarchy.

1.2. Data

The data that will be used, at least initially, is a subset of 57 leaf nodes in the ImageNet [1] dataset. The ultimate goal would be to run the final algorithm on all of the ImageNet dataset. As the ImageNet dataset is semantically very rich, this could conceivably become a system capable of classifying any object with high accuracy.

1.3. Algorithms

The current approach to nearly-infallible classification uses Locality-constrained Linear Coding [3] for features. The first step would then be to reproduce those results independently, after which other state-of-the-art methods can be examined. Exactly which algorithms are chosen depends on further reading. One such candidate is Object Bank [2], although certainly others are feasible as well. It might also be interesting to examine how well some weaker classifiers perform, if only to prove that infallibility can be achieved with any set of features, albeit with a smaller reward.

References

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[2] L. Li, H. Su, E. Xing, and L. Fei-Fei. Object bank: A high-level image representation for scene classification & semantic feature sparsification. In *Neural Information Processing Systems (NIPS)*.

[3] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Locality-constrained linear coding for image classification. 2010.