Self-supervised Visual Priors for Softmax Classification

Ishan Nigam
ishann@cs.utexas.edu

1. Proposal

Softmax classifiers are extremely effective at mapping visual structure to arbitrary labels [17]. However, softmax classifiers assume mutual exclusion in label space when this is clearly not the case as we move towards real-world long-tailed benchmarks [5].

Past attempts such as training independent binary cross-entropy classifiers, have not received much traction because they are (counter-intuitively) less effective even for multi-label classification problems [10].

We propose to reconcile top-down knowledge with bottom-up visual structure for interpretable visual recognition. We hope to supplement the softmax classifier with a self-supervised visual prior. Visual recognition should be treated as a multi-label classification problem. There is some indirect evidence to support this [1]. Apart from 1-hot encoded softmax objectives, we additionally view recognition as a structured output prediction problem in the label space based on the visual structure of the object. For example, a toy deer could be 0.5 toy and 0.5 deer, and a white horse could be 0.8 horse and 0.2 zebra.

2. Related Work

- Label aware margin distribution loss: [2].
- Knowledge distillation, label smoothening, subclass distillation: [7], [12], [11].
- Decoupling representation and classification: [8].
- The devil is in the tails: [14].
- Effective number of samples: [4].
- Label Refinery [1].

3. Technical Details

Notion of inter-class similarity Perform instance-aware self-supervised embedding learning. Cast embeddings onto unit hypersphere. Cosine distance of centroid embeddings is the notion of inter-class similarity.

Self-supervised Learning Self-supervised instance aware embedding learning has received quite a lot of attention in recent times [15, 18, 6, 3]. We intend to work with [18].

Benchmarks Imbalanced CIFAR-100 [9], iNaturalist [13].
Nearest Neighbors [16, 2, 8]

4. Hopes and Dreams

We do not expect our proposed method to outperform the classical softmax classifier on traditional balanced benchmarks such as CIFAR-10 or ImageNet-1k. Our initial experiments indicate that a several hundred examples per class are usually sufficient to learn an effective representation. However, we do expect to see improvements on balanced benchmarks in terms of computational speedups, since we expect our proposed algorithm to converge faster.

However, we expect that our proposed method should be able to outperform the softmax classifier with imbalanced long-tailed distributions. We expect that the tail classes would be able to borrow information from the head classes.

The evaluation metric precisely measures softmax classification. Hence, we do not expect our method to replace softmax classification. Instead, we expect it to supplement softmax classification. Our proposed method will be more general at train time, and more interpretable at test time.

5. (Expected) Timeline

- Mar 15: Literature Survey
- Apr 15: Prototyping on Imbalanced CIFAR-100
- May 15: Scale up results on iNaturalist
References


