Learning Compositional Representations for Few-Shot Recognition

Pavel Tokmakov
Yu-Xiong Wang
Martial Hebert

Category distribution in the wild

- Objects in the world distribute according to Zipf’s law
- Most of the categories are extremely rare
- Traditional learning-based approaches do not scale to this scenario
- Tackled in the few-shot learning setting

Analysis of existing approaches

- Observations differ as network depth increases
- Simple cosine classifier works surprisingly well on modern architectures
- Should we focus on learning better representations?

Compositional representations

- Given a set of frequent categories, what representation learned on it will transfer better to the rare ones?
- Compositionality is the key property for generalization in humans [Fodor, Harvard Univ. Press ’75]
- Can we enforce compositionality in deep representations using part annotations?

Compositionality regularization

- On base categories, embed images and attributes into a feature space
- Force image embedding to be equal to the sum of attribute embeddings
- Train jointly with the image classification loss to constrain model selection
- Relax the hard constraint to allow for non-exhaustive attribute annotations

Databases

- Support a wide spectrum of attributes, from concrete to highly abstract
- Attributes are annotated on the category level
- Collect 150 attributes for a subset of ImageNet categories

Ablation analysis

- On CUB all variants result in improvements
- Overall up to 8% improvement

Comparison to the state of the art

- Marginal improvements on base
- Indeed select the model with better generalization properties

Analysis of representations

- Cosine classifier with compositionality constraint outperforms complex methods across the board
- Our ImageNet annotations are effective
- Data augmentation is important

Baseline on SUN397: 331 interpretable units, 119 concepts [Bau et al., CVPR’17]

CUB-200-2011
SUN397
ImageNet

Ours on SUN397: 331 interpretable units, 119 concepts [Bau et al., CVPR’17]

CUB-200-2011
SUN397
ImageNet

Baseline on SUN397: 169 interpretable units, 92 concepts