Discovering Localized Attributes for Fine-grained Recognition

Kun Duan¹, Devi Parikh², David Crandall¹, and Kristen Grauman³ ¹Indiana University, ²Toyota Technological Institute at Chicago, ³University of Texas at Austin

1. Overview

 Objective: Given images with fine-grained category-labels, discover a vocabulary of localized attributes that are both semantically meaningful and discriminative

red stripes on wings white helly vellow belly

We employ an iterative and interactive approach:

- A. Find discriminative attribute candidates for 2 similar classes with a latent CRF
- B. Use a recommender system to identify candidates likely to be meaningful
- C. Present them to a human for naming and verification. Repeat.

2. Discovering Localized Attributes with Latent CRFs

- Goal: Find discriminative local attribute candidates given images from two categories
- Define a latent CRF to find regions in positive images that are similar to one another but dissimilar from negative image regions

Nodes: images, Labels: segments



When finding multiple candidates, also encourage diversity;

$$E(\mathcal{L}|\mathcal{I}) = \sum_{k=1}^{K} E(L_k|\mathcal{I}) + \sum_{i=1}^{M} \sum_{k,k'} \delta(l_i^k, l_i^{k'}|\mathcal{I}_i)$$
 Spatial overlap between regions

- · Attribute detection: add test image to CRF and run inference
- Implementation details:
- Generating regions: hierarchical segmentation
- Region features: color, qPb contour, size, shape, and spatial location
- Distances: L2 for spatial location, chi-squared for other features



forewir

4. Image-to-text Generation Results

We can annotate unseen images with region labels using our discovered local attributes:



5. Image Classification Results

- We run attribute detection to produce a binary feature vector for each image, then use these vectors for fine-grained recognition using Nearest Neighbor and SVM classifiers.
- We compare four approaches to generate localized attributes:
- Proposed: our proposed method that focuses on discriminative power and semantics
- Hand-listed (focus on semantics): expert-generated attributes
- Discriminative only (focus on discrimination): non-semantic candidates removed in post-process
- Upper bound (price paid for semantics): all discriminative candidates including non-semantic ones

Our proposed method performs significantly better than existing approaches



Recommender aathers more attributes and achieves higher accuracy for same amount of user effort



6. Summary and Conclusions

- Finds local attributes that are both discriminative and human understandable.
- Recommender system prioritizes candidates likely to be meaningful, saving user time.
- Compares favorably to existing approaches on two fine-grained recognition datasets.
 - More information at http://vision.soic.indiana.edu/attributediscovery



INDIANA UNIVERSITY