CUB (Caltech-UCSD Birds-200-2011) dataset description

General information

CUB-200-2011 is an extended version of the CUB-200 dataset (CUB-200, as example, contains half as many images). The dataset has an unique properties: subordinate category recognition, multi-class object detection, attribute-based methods, crowdsourcing and user studies.

Attention

Authors warn that images in this dataset overlap with images in ImageNet. Exercise caution when using networks pretrained with ImageNet (or any network pretrained with images from Flickr) as the test set of CUB may overlap with the training set of the original network.

The dataset page

http://www.vision.caltech.edu/visipedia/CUB-200-2011.html

View the dataset

http://www.vision.caltech.edu/visipedia-data/CUB-200-2011/browse/index.html

Details

The dataset contains about 200 categories, 11,788 images, 15 part locations, 312 binary attributes, 1 bounding box.

Structure of the dataset & folders

Segmentation

Has 200 subfolders (respectively for each bird) with segmented photos.

All Images and Annotations

Has images and class labels, bounding boxes, part locations, attribute labels. For example, images and class labels contains:

1. The list of image file names is contained in the file images.txt, with each line corresponding to one image: <image_id> <image_name>.

2. The suggested train/test split is contained in the file train_test_split.txt, with each line corresponding to one image: <image_id> <is_training_image>.

3. The list of class names (bird species) is contained in the file classes.txt, with each line corresponding to one class: <class_id> <class_name>.

4. The ground truth class labels (bird species labels) for each image are contained in the file image_class_labels.txt, with each line corresponding to one image: <image_id> <class_id>.

Benchmarks

The authors (Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, Serge Belongie) introduce a set of benchmarks and baseline experiments for studying bird species categorization, detection, and part localization:

Localized Species Categorization

Given the ground truth part locations, assign each image to one of 200 bird classes.

Part Localization

Given the full, uncropped bird images, predict the location and visibility of each bird part.

Species Categorization/Detection

Using only the full, uncropped bird images, assign each image to one of 200 bird classes.

Example

For example, consider the one of Scarlet Tanager's photos from this dataset:

Image marking

Scarlet_Tanager_0086_138272.jpg

Some of attribute labels

Has_Breast_Pattern: Solid (definitely, 2.4410sec)	Has_Back_Color: Red (definitely, 3.1130sec)
Has_Tail_Shape: (not visible, 2.8420sec)	Has_Upper_Tail: Color_Red (probably, 6.7530sec)
Has_Head_Pattern: Plain Eyeline (definitely, 24.4490sec)	Has_Breast_Color: Red (definitely, 2.3130sec)

How to download

To download a dataset, you need to go to the dataset website (See the link under the heading «The dataset page») and download the segmented data (37 MB) and, accordingly, the source data with the attributes (1,1 GB) via links.