

FlickrMFC dataset

Gunhee Kim Eric P. Xing
School of Computer Science, Carnegie Mellon University
{gunhee, epxing}@cs.cmu.edu

1. The FlickrMFC Dataset

1.1. Overview

We introduce a new dataset called FlickrMFC for the benchmark purpose of multiple foreground cosegmentation. They are sampled images from Flickr photostreams, each of which is taken by a single user for a specific event in a single day. Hence, a fixed number of subjects (or foregrounds) frequently occur across the photostream, but an unknown subset of them appears in every single image. We also provide hand-labeled pixel-level ground truths.

The FlickrMFC consists of 14 groups, each of which contains 12~20 images. Table 1 summarizes some key information about the collected groups including the number of images and the description of foregrounds. The group names are identical to the search keywords that are used for image downloading from Flickr. Fig.1 and Fig.2 show all images of the whole 14 groups.

1.2. Usage Detail

The images and their ground truths are stored in `img` and `gt` folders. All image formats are JPG, and all ground truth formats are PNG. The image and its corresponding ground truth share the same file name only except the extension type.

Our ground truth files are made by following the conventions of PASCAL VOC 2012 dataset[1].

- Each pixel is labeled with a unique foreground number or background (index 0). The foreground numbers are stored in `GroupInfo.mat`.
- Bordering regions with a width of five pixels are marked as a *void* label (index 255). All void pixels can be ignored when computing segmentation accuracies.

The `GroupInfo.mat` includes the `GroupInf`, a 14-length structure array, each of which stores foreground information about each group. It consists of three fields:

- `name`: Group name
- `fgname`: Foreground names (See Table 1).
- `fgid`: The unique foreground indices of the group.

- `labels`: The foreground incides that each image contains.

2. Notes

- The FlickrMFC dataset is available at our webpage <http://www.cs.cmu.edu/~gunhee/r.mfc.html>.
- Please cite [2] when you use this dataset.
- This document is last updated at July 13, 2012.

References

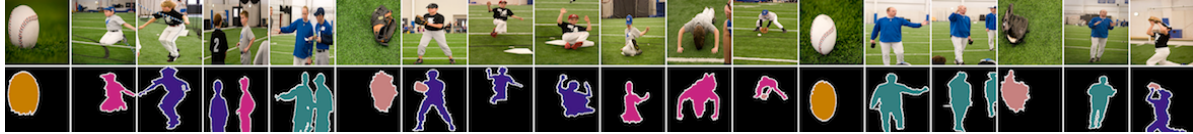
- [1] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. <http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html>. 1
- [2] G. Kim and E. P. Xing. On Multiple Foreground Cosegmentation. In *25th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012)*, 2012. 1

Group	M	K	Foreground names
<i>apple+picking</i>	20	6	apple+bucket, baby, boy+blue, girl+blue, girl+red, pumpkin
<i>baseball+kids</i>	18	5	ball, boy+black, boy+gray, coach, glove
<i>butterfly+blossom</i>	18	8	beetle, butterfly+orange, butterfly+tiger, butterfly+yellow, flower+pink, flower+red, ladybug, leaf
<i>cheetah+ safari</i>	20	5	cheetah, eagle, elephant, lion, monkey
<i>cow+pasture</i>	20	5	cow+black, cow+brown, man+blue, man+red+cap, truck
<i>dog+park</i>	20	4	dog+black, dog+brown, dog+white, woman
<i>dolphin+aquarium</i>	18	3	killer+whale, dolphin+gray, seal
<i>fishing+alaska</i>	18	5	flower, man+gray, man+white, salmon, woman+gray
<i>gorilla+zoo</i>	18	4	boy, girl, gorilla+black, orangutan+brown
<i>liberty+statue</i>	18	4	boat+blue, boat+red, empire+state+building, liberty+statue
<i>parrot+zoo</i>	18	5	hand, parrot+green, parrot+red, parrot+white, parrot+yellow
<i>stonehenge</i>	20	5	bird, cow+black, cow+white, person, stonehenge
<i>swan+zoo</i>	20	3	flower+yellow, swan+black, swan+gray
<i>thinker+Rodin</i>	17	4	sculpture+thinker, sculpture+venus, van+gogh, woman

Table 1. Summary of 14 groups of *FlickrMFC* dataset. M and K denote the number of images and foregrounds, respectively.



(a) *apple+picking*



(b) *baseball+kids*



(c) *butterfly+blossom*



(d) *cheetah+safari*



(e) *cow+pasture*

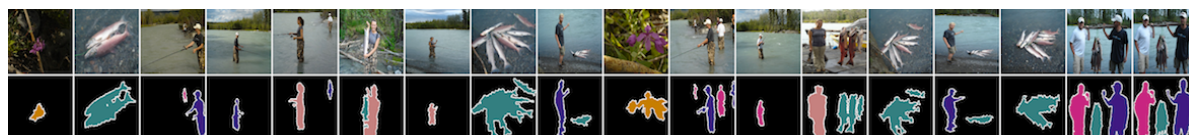
Figure 1. The *FlickrMFC* dataset. In each group, the first row shows input images, and the second row illustrates hand-labeled pixel-level ground truths. Every image is resized as a square for the sake of better visualization even though it is actually not.



(f) *dog+park*



(g) *dolphin+aquarium*



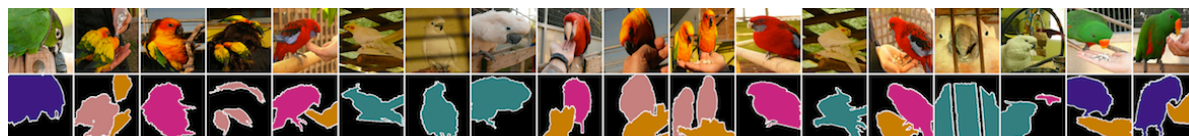
(h) *fishing+alaska*



(i) *gorilla+zoo*



(j) *liberty+statue*



(k) *parrot+zoo*



(l) *stonehenge*



(m) *swan+zoo*



(n) *thinker+Rodin*

Figure 2. The FlickrMFC dataset. In each group, the first row shows input images, and the second row illustrates hand-labeled pixel-level ground truths. Every image is resized as a square for the sake of better visualization even though it is actually not.