From ImageNet to Serengeti: Recognizing Animals in Wild Scenes

Guangnan Ye, Maja Rudolph

05/08/2014

Motivation

For the problem of animal classification, most existing datasets are collected from the web. There remains a problem whether our models trained with the web images would work for the real world application.

Serengeti Dataset

Animals are captured from several fixed cameras.

Overall 48 animal categories with 901291 positive images in wild scene.

Proposed Idea

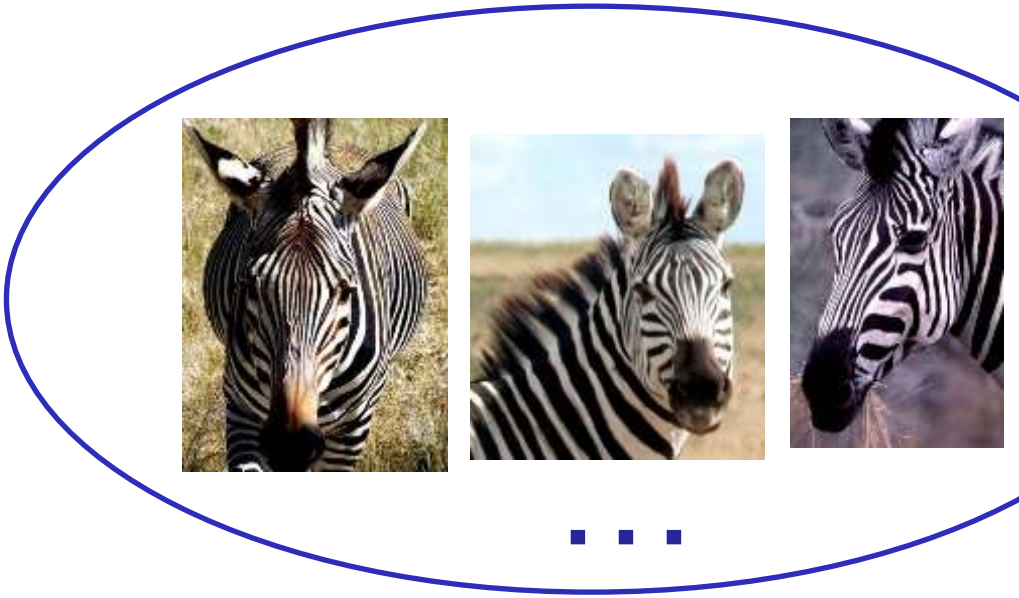
Test Image

Robust Foreground Extraction

Detection and Classification with Domain Adaptation

Background Modeling with Background Images

IMAGENET



Related Work

- Butzer, Hoferlin, Heidemann, Evaluation of background subtraction techniques for video surveillance (CVPR 2011)

  (*Unfortunately can't use GMM/background modeling as implemented in openCV it only works well for videos)*

- Stephen Gould, Jim Rodgers, David Cohen, Gal Elidan, Daphne Koller: Multi-class Segmentation with Relative Location Prior (IJCV 2007)
- Lei Zhang and Qiang Ji: Image Segmentation with a Unified Graphical Model (PAMI 2010)
- Oliver, Nuria M., Barbara Rosario, and Alex P. Pentland. "A Bayesian computer vision system for modeling human interactions." PAMI 2000
- P. Felzenszwalb, D. McAllester, D. Ramanan. A Discriminatively Trained, Multiscale, Deformable Part Model. CVPR 2008

Background Modeling

Preprocessing (tesseract-ocr)

09-11-2010 08:01:10

10-12-2010 16:43:49

09-11-2010 18:50:53

Background Modeling with Different Time Stamps

Background Modeling

PCA

Training vectors

$$\{x_t\}_{t=1:N} \qquad x_t \in \mathbb{R}^{h \cdot w}$$

Mean and covariance:

$$\mu = \frac{1}{N} \sum_{t=1}^{N} x_t$$

$$\Sigma = \frac{1}{N} \sum_{t=1}^{N} (x_t - \mu)(x_t - \mu)^T$$

The goal is to find a unitary matrix Q s.t.

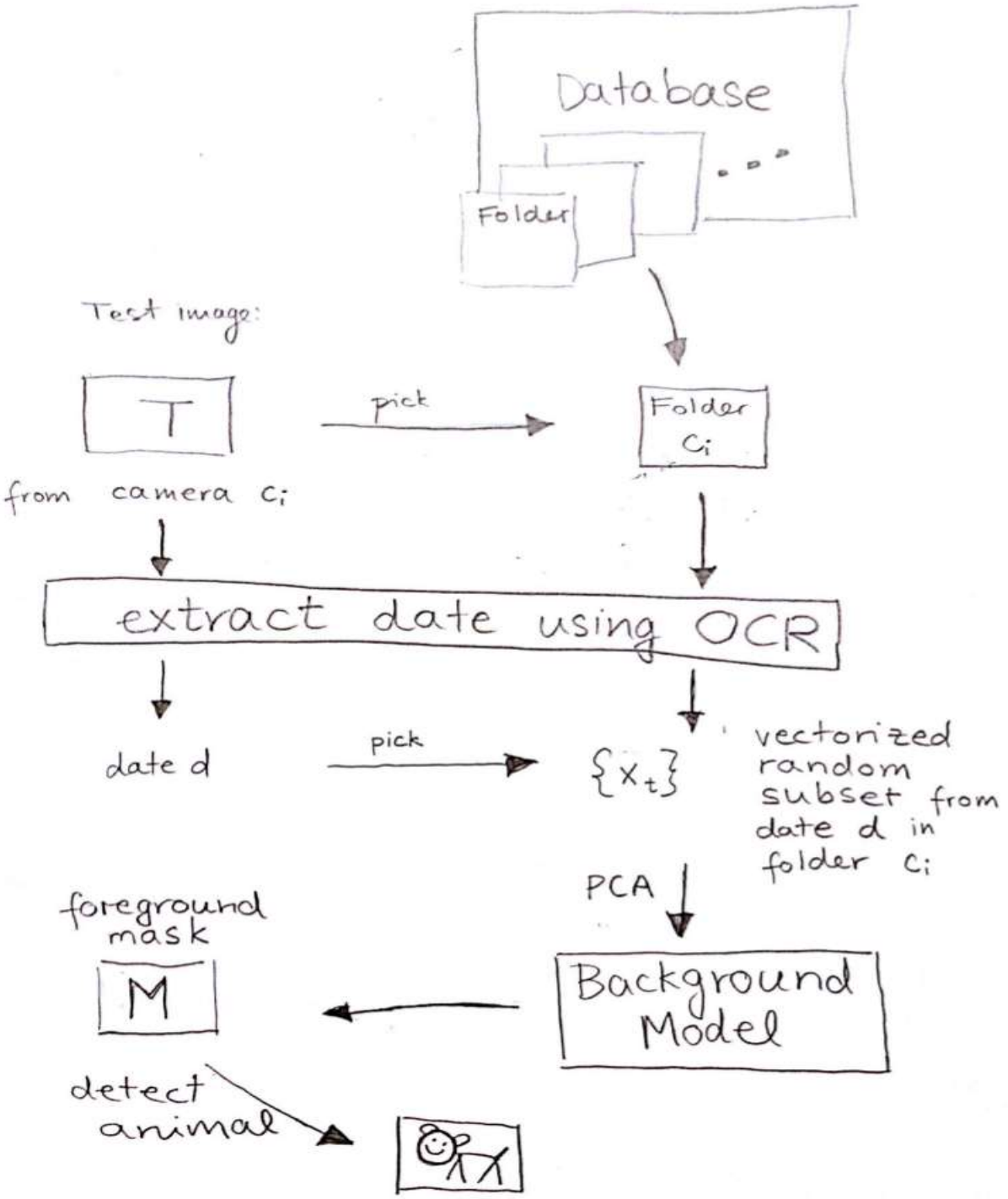
$$\Lambda = Q^T \Sigma Q$$

is diagonal with diagonal entries sorted by magnitude.

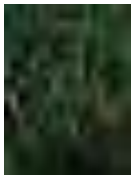
Given a testing image x we can model the background B and extract the foreground F as follows:

$$B = Q_M^T Q_M (x - \mu) + \mu$$

$$F = x - B$$

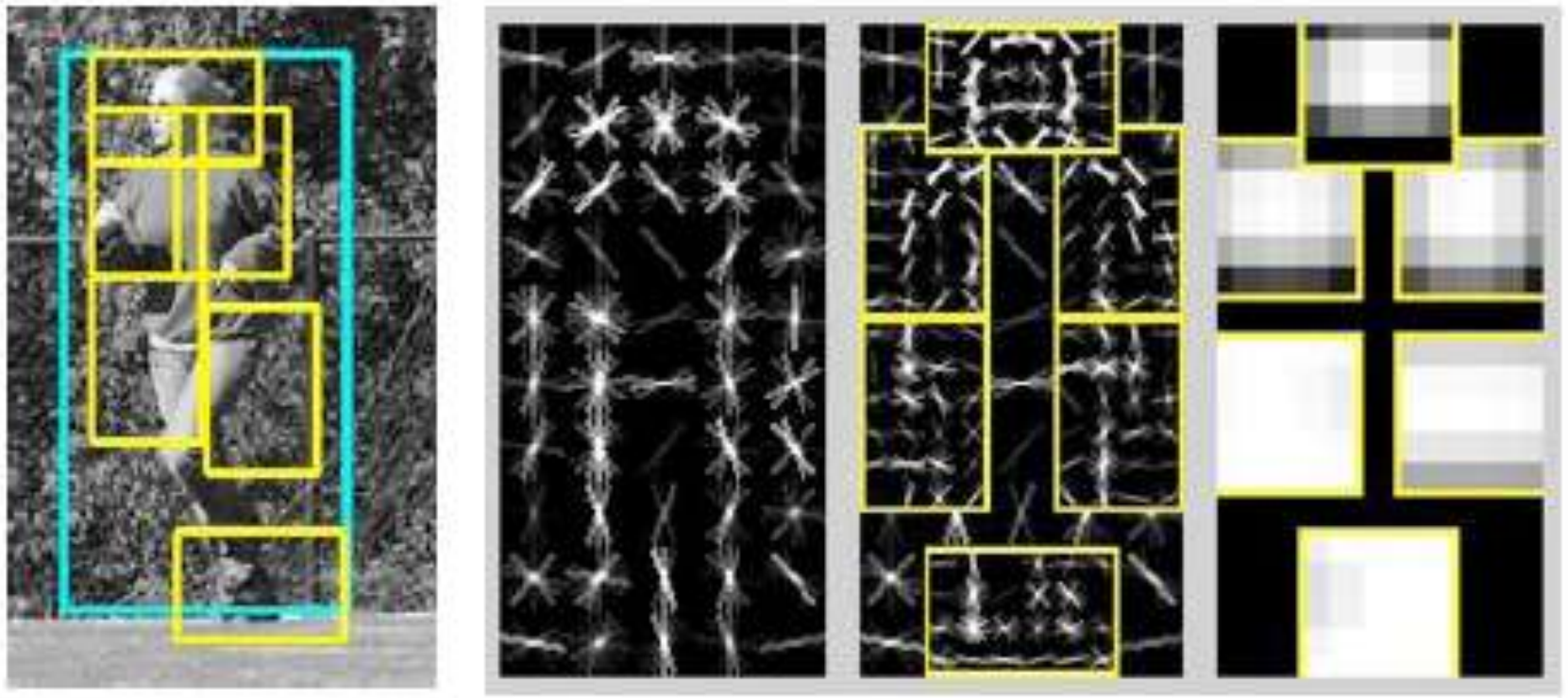

Database
Folder
Test image:
T
pick
Folder c_i
from camera c_i
extract date using OCR
date d
pick
$\{x_t\}$
vectorized random subset from date d in folder c_i
PCA
foreground mask
M
Background Model
detect animal

Dikdik Detector: Before and After



Object Detector

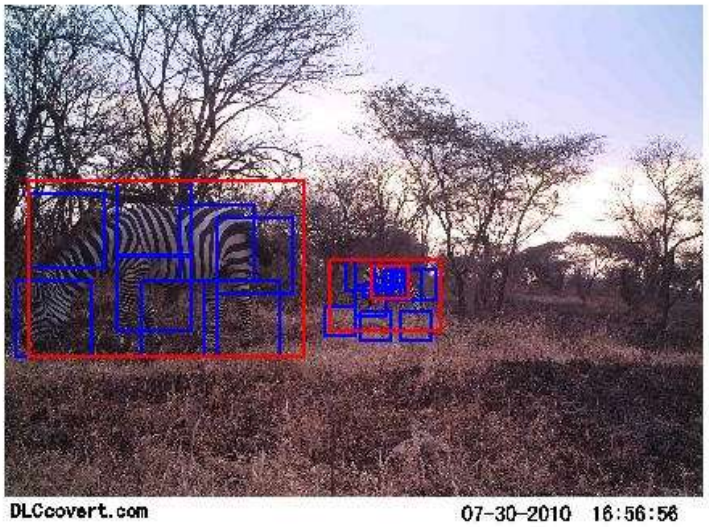
Part-based model Detector

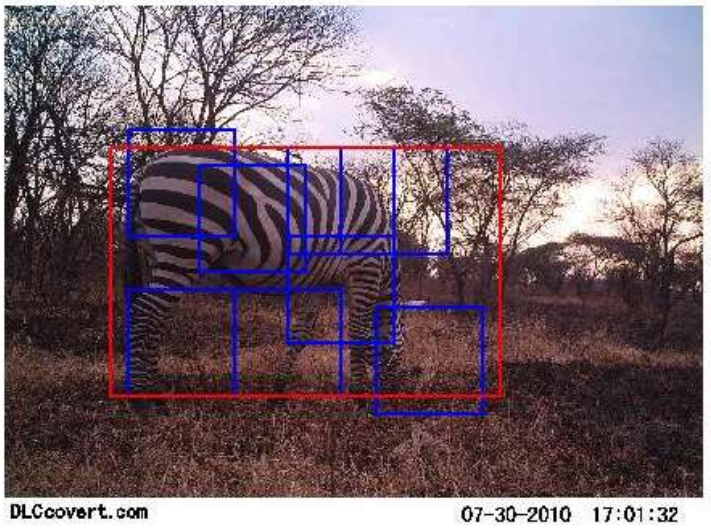


P. Felzenszwalb, D. McAllester, D. Ramanan. **A Discriminatively Trained, Multiscale, Deformable Part Model. CVPR 2008**

Detector Results on Serengeti

Training with 457 ImageNet “zebra” positive Images and 100 Serengeti background images.

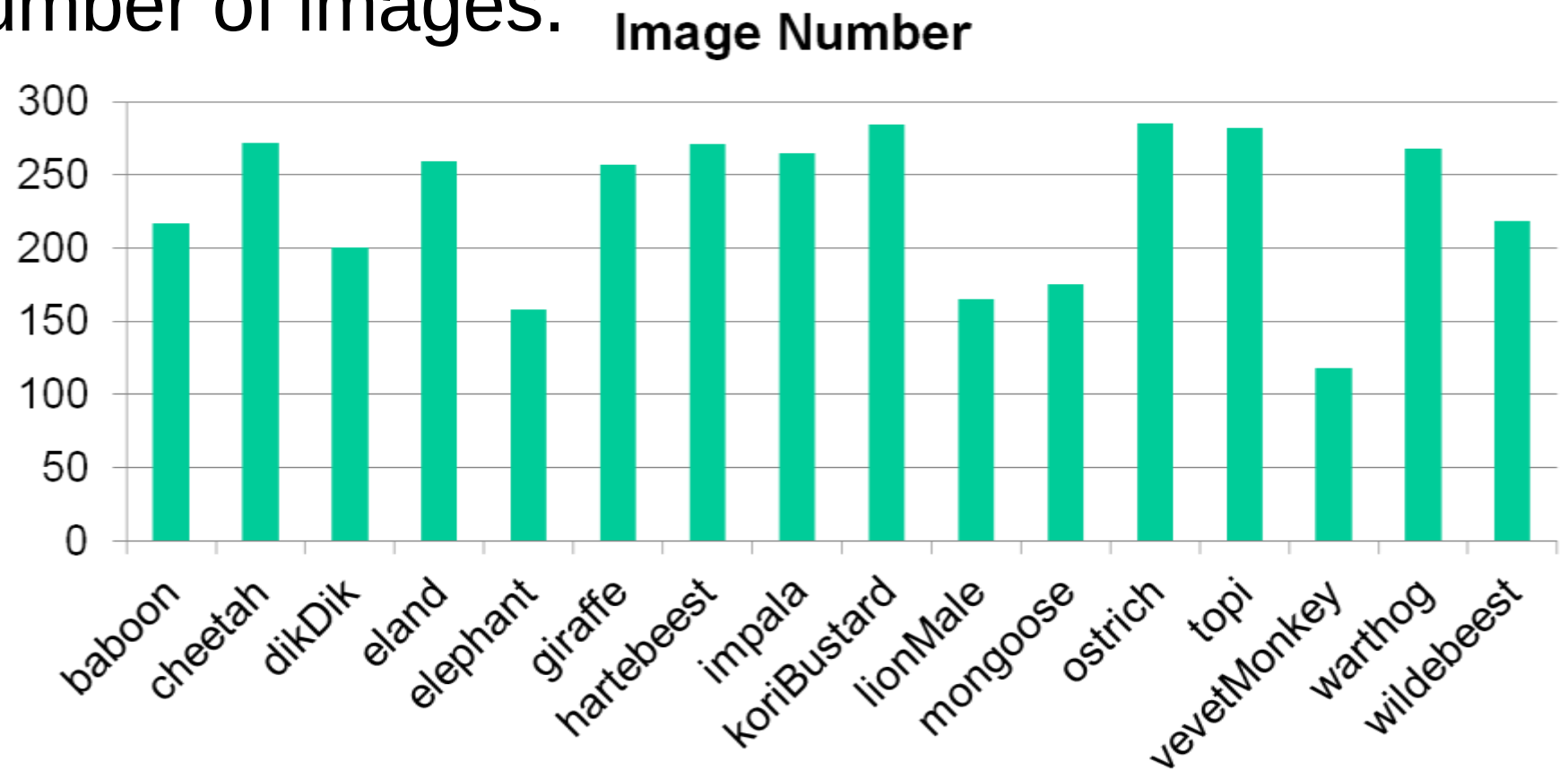




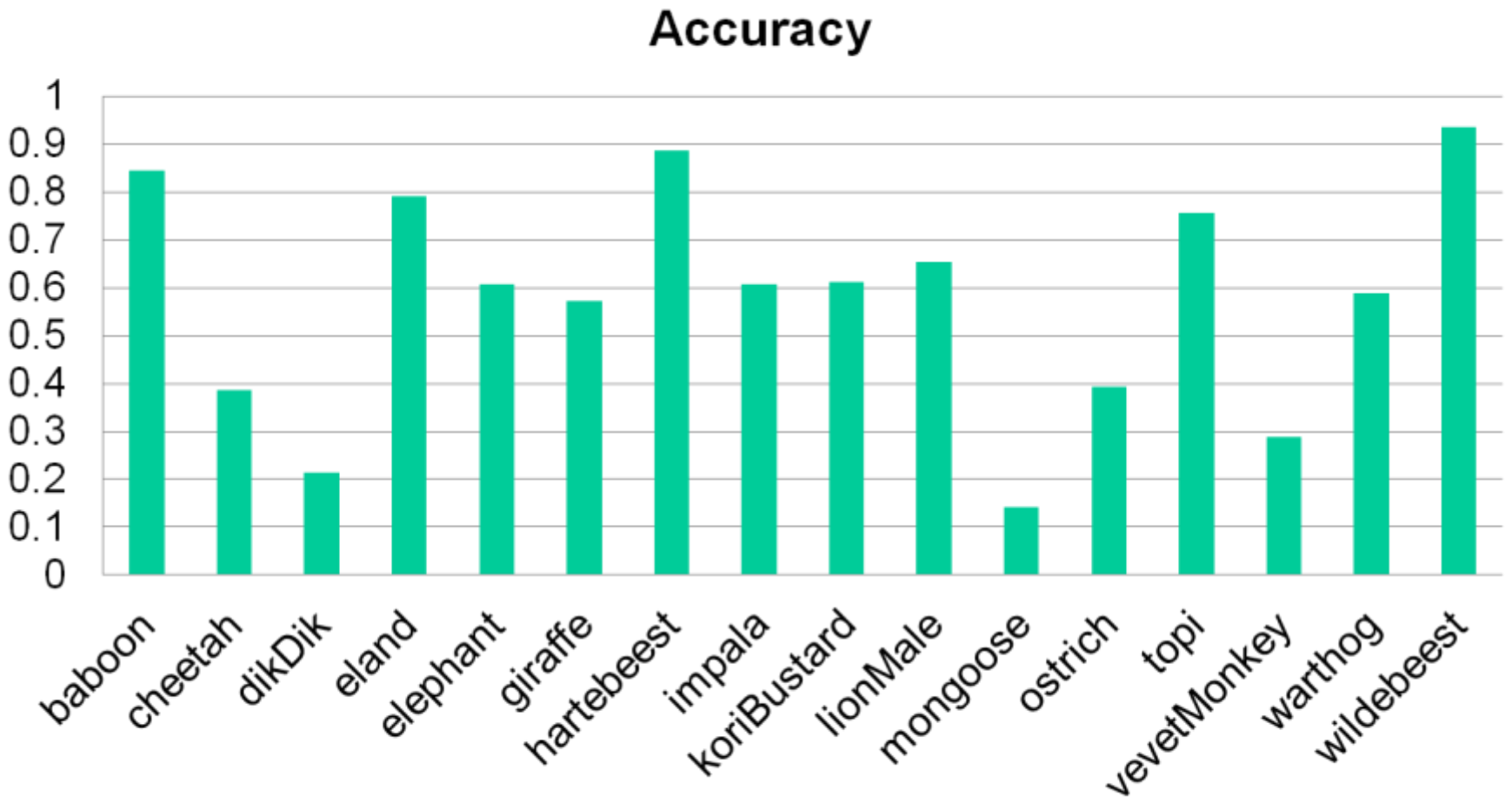
Dataset Setup

OCR is applied on the time stamp so that pictures captured between 7:00am-5:00pm are selected.

16 popular categories are selected based on number of images.

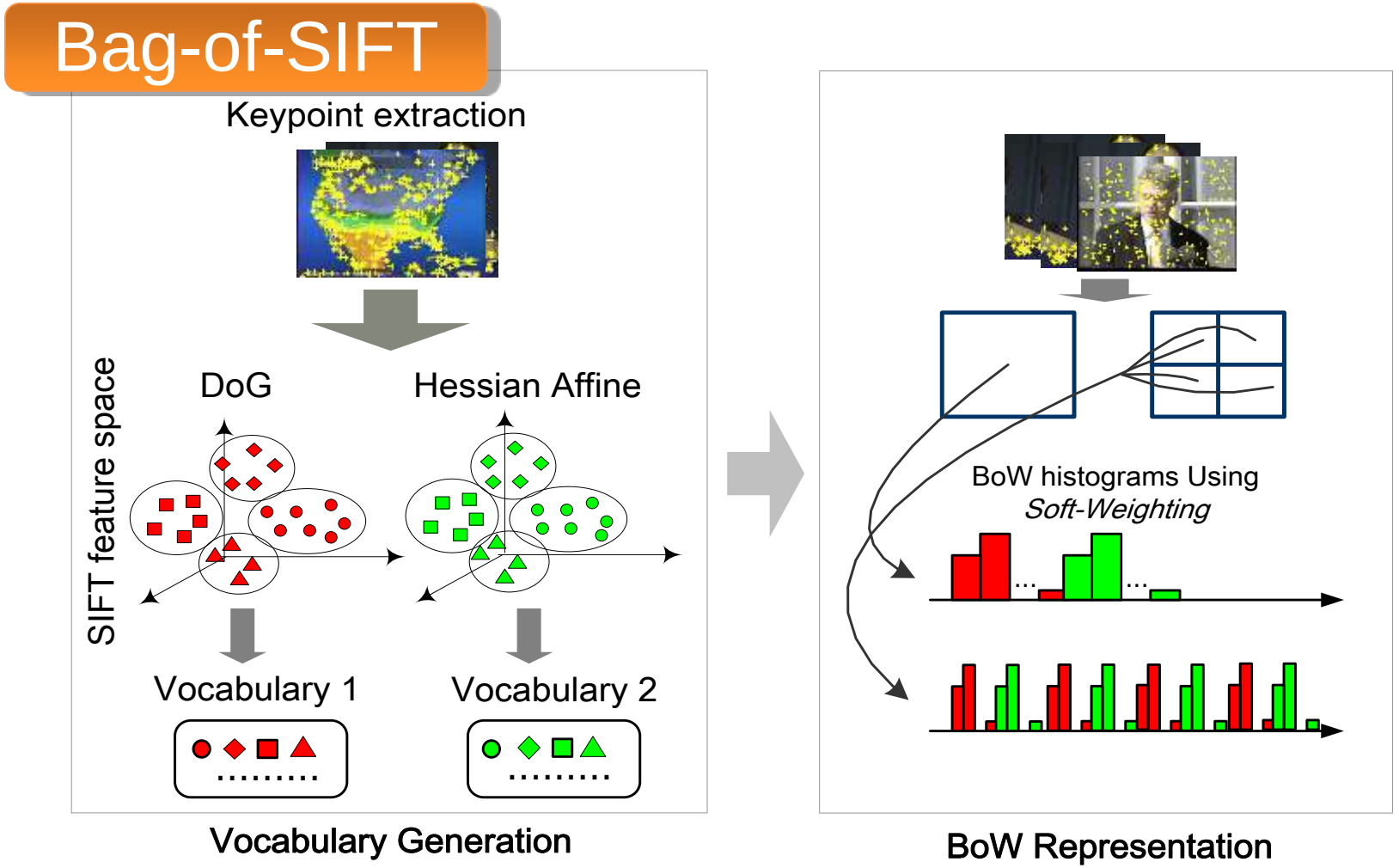


Detector Performance

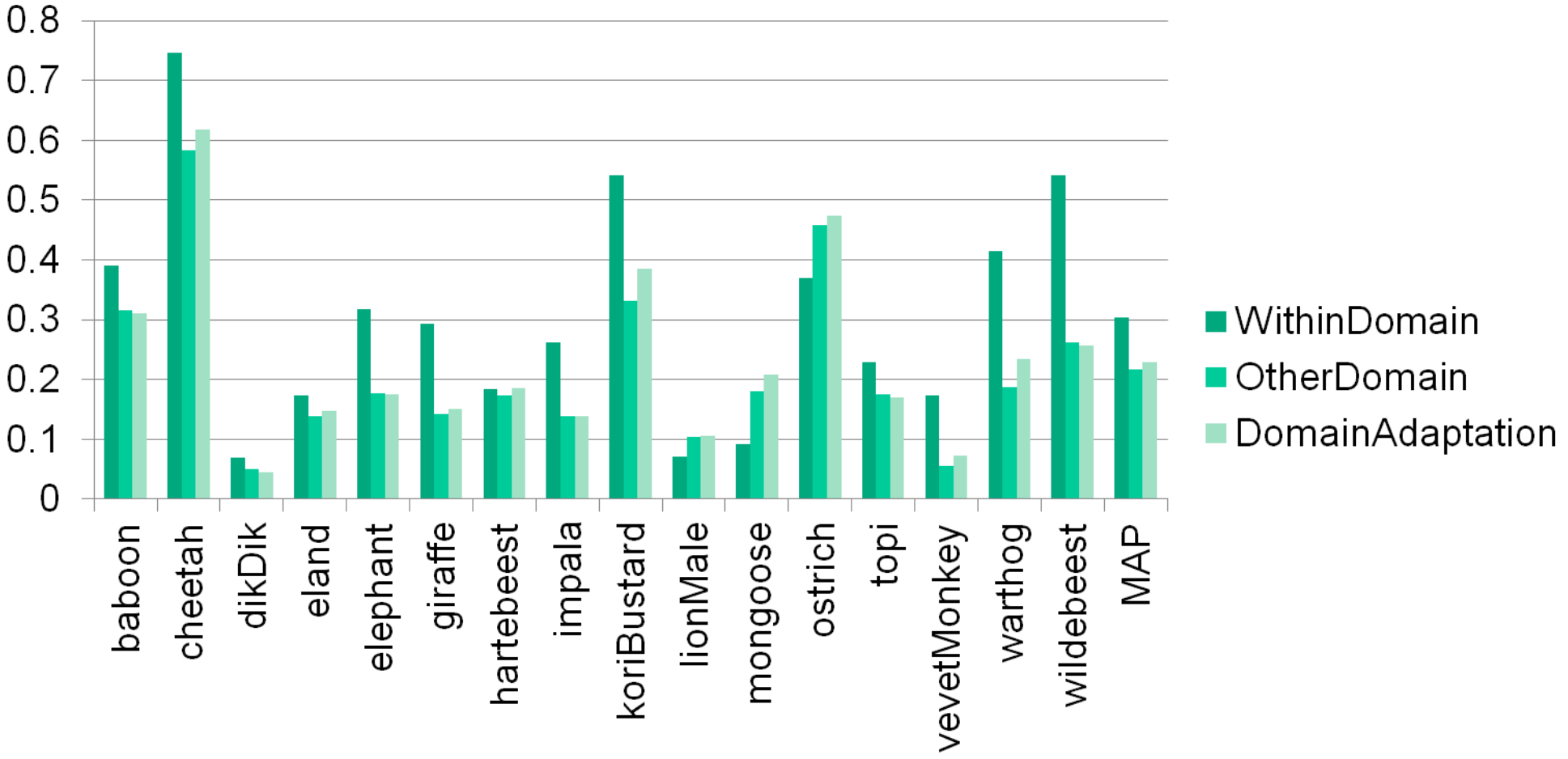


Bag-of-SIFT Representation

- **SIFT**
- **Soft weighting** (Jiang, Ngo and Yang, ACM CIVR 2007)



Performance Comparison



Domain Adaptation Method: Geodesic Flow Kernel for Unsupervised Domain Adaptation. B. Gong, Y. Shi, F. Sha, and K. Grauman.CVPR 2012.

Conclusion

Proposed Idea:

Background model, detection, classification have been tested on Serengeti Dataset.

Future Work:

Joint model among segmentation, detection, and classification