De-Correlating CNN Features for Generative Classification

- Chaitanya Desai, Jayan Eledath, Harpreet Sawhney, Mayank Bansal
SRI International

IEEE Winter Conference on Applications of Computer Vision

Problem Definition:

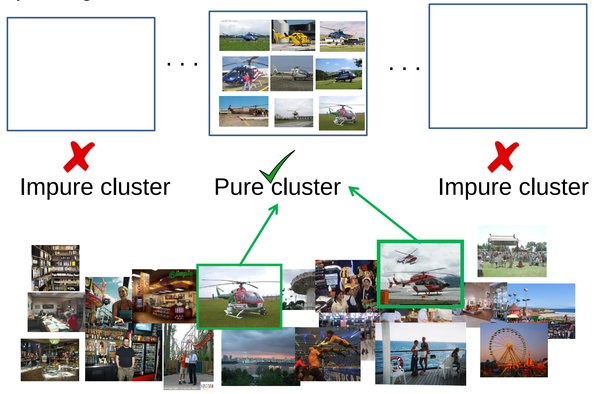
- Given a handful of images as **positive** examples for the class of interest, how do I build a classifier **without** any **labeled negative** training data?

Motivating Applications:

1. Content based Searching and Tagging:
   - Suppose you go *Rafting* with friends and take pictures:

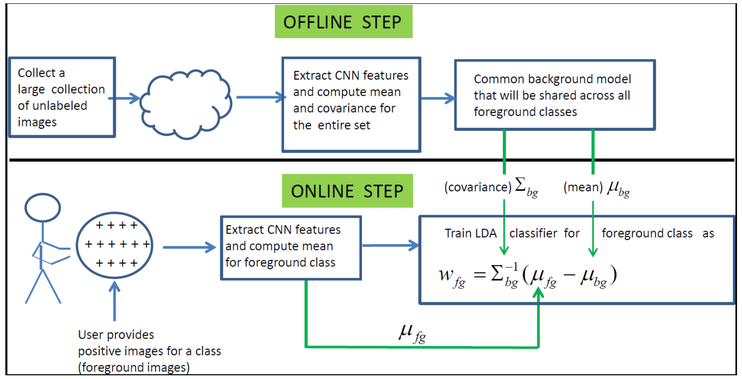
   - How do you search for similar *Rafting* pictures in your photo collection using only the images taken on that day as source of class specific *training* data?

2. Improving *Recall* for a *Pure* Cluster:



   - Given the output of a clustering algorithm (a mix of pure and impure clusters), how can I use the images of a “pure” cluster to bootstrap a model that can help improve its recall?

Pipeline of the proposed LDA based approach:



Closest Related Work (*Malik et. at*. ECCV 2012)

$M^1_{bicycle} = \Sigma_{bg}^{-1}(\mu^1_{bicycle} - \mu_{bg})$ $M^2_{bicycle} = \Sigma_{bg}^{-1}(\mu^2_{bicycle} - \mu_{bg})$ $M^3_{bicycle} = \Sigma_{bg}^{-1}(\mu^3_{bicycle} - \mu_{bg})$

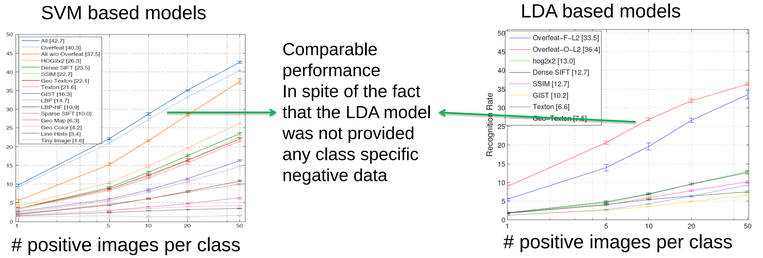
- Different viewpoints of an object have different appearances in HOG space, and are modeled using different templates.
- Each template is built using viewpoint specific foreground images and a common background model

$M^1_{bicycle} \bullet \Phi(I) \rightarrow s_1$
$M^2_{bicycle} \bullet \Phi(I) \rightarrow s_2$
$M^K_{bicycle} \bullet \Phi(I) \rightarrow s_K$
$M_{bicycle} = (M^1_{bicycle}, M^2_{bicycle}, \ldots M^K_{bicycle})$

Outputs of different mixtures need to be calibrated using class specific negatives
Whereas our model implicitly models mixtures and needs no calibration

Experiments:

- Quantitative Scene Classification results on SUN-Scene database



- Qualitative results of using our LDA based classifiers:

- Application to Query Expansion (QE):