# DIP - PROJECT PROPOSAL

## One Shot Detection with Laplacian Object and Fast Matrix Cosine Similarity

SIDDHANT KATYAN

2018801005

## Overview

**Recent research in visual recognition has attracted interest in the study of the following two big questions — i) how to make generalizations for solving large scale visual recognition problem ii) How to encode image attributes robustly to represent fine grained distinction/similarity. In this paper, we focus on a particular variety of the fine-grained visual recognition where the objective is to detect visual similarities across images without the involvement of extensive (or any) training.**

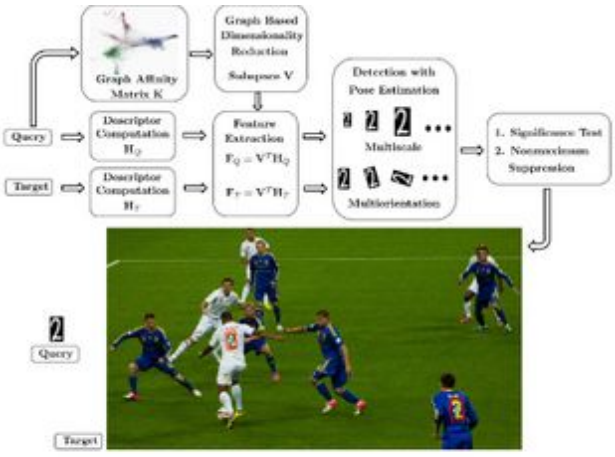
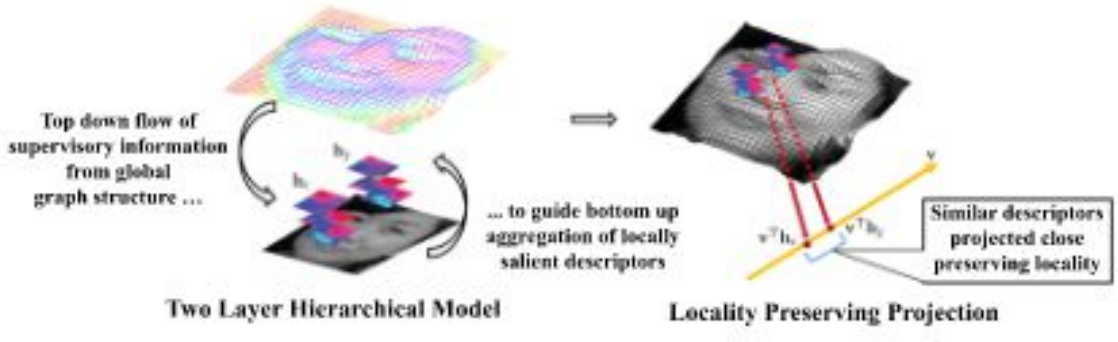



Fig. 1. Overview of our one shot detection scheme: we aim to detect a given query [15] (e.g., symbol, face, human pose, car, flower) appearing in a visually similar manner in a bigger target image [1]

# Problem

**One shot, generic object detection involves searching for a single query object in a larger target image. Relevant approaches have benefited from features that typically model the local similarity patterns. In this project we aim to combine local**
**similarity (encoded by local descriptors) with a global context (i.e., a graph structure) of pairwise affinities among the local descriptors, embedding the query descriptors into a low dimensional but discriminatory subspace.**



# Goals

**In this project , we focus on a particular variety of the fine-grained visual recognition where the objective is to detect visual similarities (Fig. 1) across images without the involvement of extensive (or any) training. In general, the one shot, generic object detection approaches take a single query image as input, and the dominant object present in the query image is detected in a bigger target image**

# Approach

**1) We combine local similarity (encoded by local descriptors) with a global context**
**(i.e., a graph structure) of pairwise affinities among the local descriptors, embedding the query descriptors into a low dimensional but discriminatory subspace. Unlike principal components that preserve global structure of feature space, we actually seek a linear approximation to the Laplacian eigenmap that permits us a locality preserving embedding of high dimensional region Descriptors.**

**2) Our second contribution is an accelerated but exact computation of matrix cosine similarity as the decision rule for detection, obviating the computationally expensive sliding window search. We leverage the power of Fourier transform combined with integral image to achieve superior runtime efficiency that allows us to test multiple hypotheses (for pose estimation) within a reasonably short time.**

## References


1. He, Xiaofei, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang, Face recognition using Laplacianfaces, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005
2. Eli Shechtman and Michal Irani, Matching local self-similarities across images and videos, IEEE Conference on Computer Vision and Pattern Recognition, 2007
3. Hae Jong Seo and Peyman Milanfar, Training-free, generic object detection using locally adaptive regression kernels, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010
4. Dubout, Charles, and François Fleuret, Exact acceleration of linear object detectors, European Conference on Computer Vision, 2012


## Github Link to the project: One Shot Detection with Laplacian Object and Fast Matrix Cosine Similarity