ROBUST ACADEMIC POSTER RECOGNITION

Thanh Vu and Amir Sadovnik

Department of Computer Science, Lafayette College

Introduction

- Image recognition has been one of the most well researched tasks in computer vision.
- Traditionally this is done by comparing features extracted from query images to those extracted from an image database.
- Despite producing state of the art results for many types of objects, this approach does not generalize to all domains.
- In this paper we address one such domain: academic poster recognition.

Fig. 1. We attempt to develop a robust and fast algorithm to recognize academic posters in conferences by detecting and recognizing their titles

In this study, we present:

- The reason why the traditional approach of feature matching fails
- A new approach based on object detection using convolutional neural network, which outperforms the traditional methods

Dataset

Our dataset contains 100 academic posters from different departments at Lafayette College:

- For each poster, we take one single front facing photo of the entire poster for training, and five additional ones for testing.
- The test images are taken at different distances, angles, cropping and occlusions, all using a mobile device .
- We use one image (per poster) to generate a training set of 2000 images for the CNN by applying different image transformations.

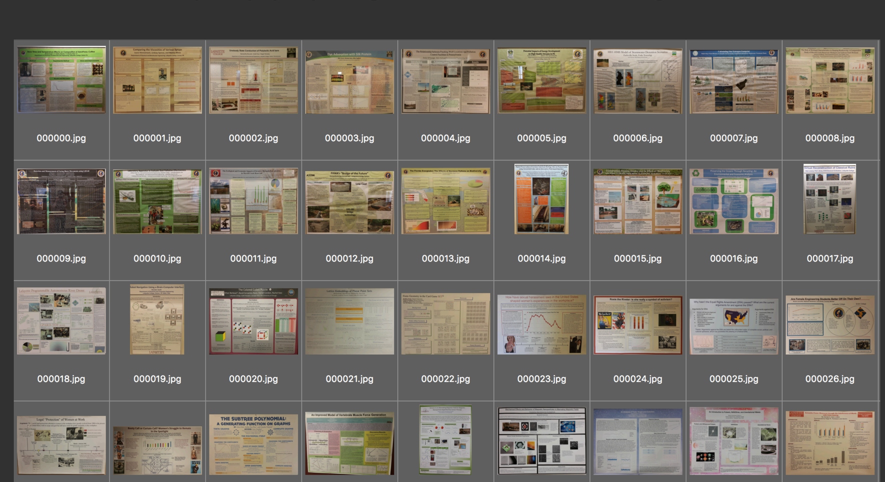


Fig. 2. Examples of posters in our database

Methods

- First, we generate synthetic training data using linear transformations.
- Then, we train a CNN based on YOLOv2, treating poster titles as objects.
- Lastly, we simply pass a new image once through the trained CNN to detect which poster is in the image

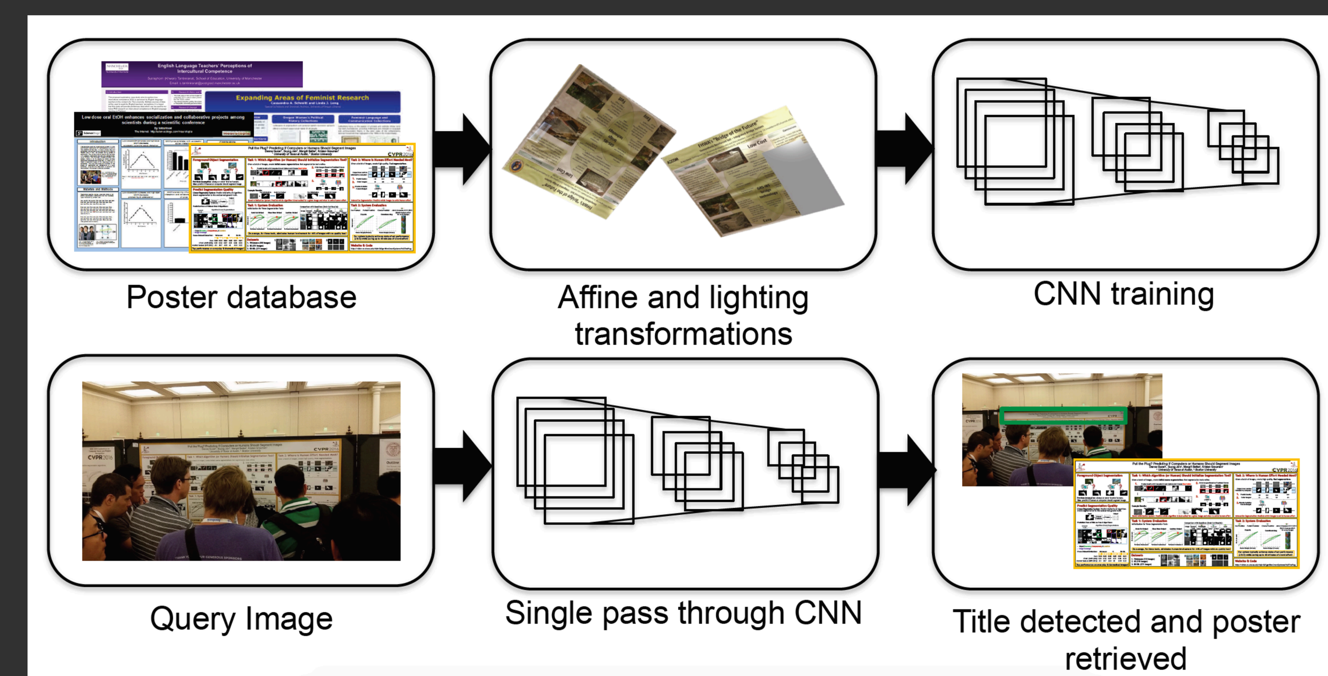


Fig. 3. The general workflow of our method

1. Occlusions

- Focus on only the title.
- Simulate the audience (occlusion) as gray rectangles randomly placed within the images.

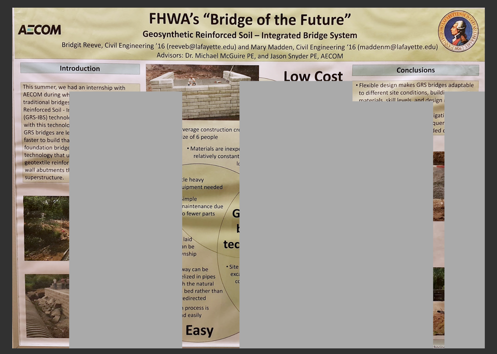


Fig. 4. Occlusion example

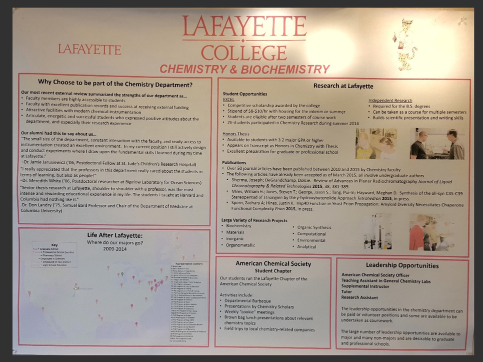


Fig. 5. Lighting example

2. Lighting Transformation

- Randomly generate light blobs using a Gaussian distribution with arbitrary mean and a standard deviation of *R/3*, with *h/2 < R < h*2*.

3. Blurring

- To simulate out-of-focus images, apply a standard average blurring method, with kernel K:

$$K_{n,n} = \frac{1}{n^2} \times \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

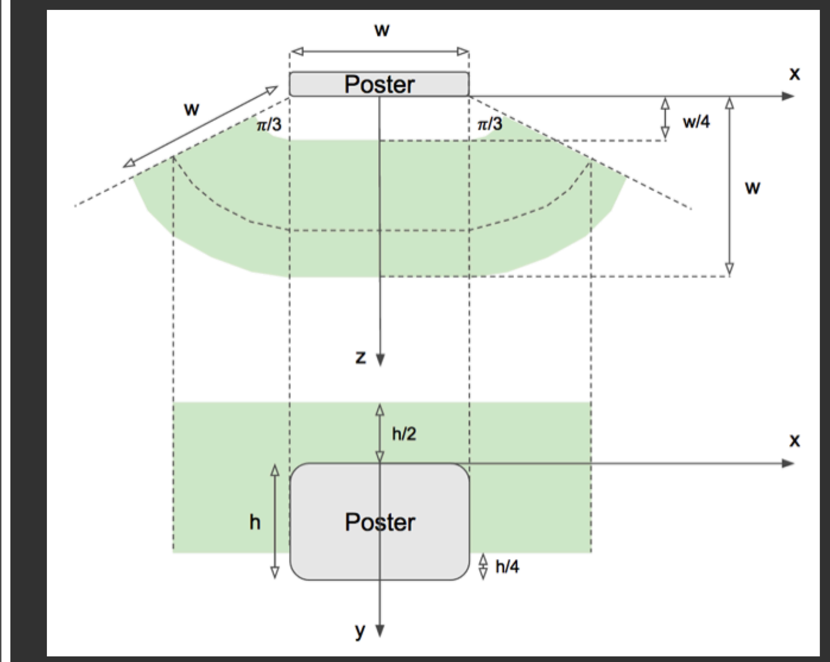


Fig. 6. 3D point selection range (green area) for the perspective transformation

4. Scale, Translation, and Rotation

- Scale and translate the poster, relative to the training image
- Rotate by an angle between -π/4 & π/4

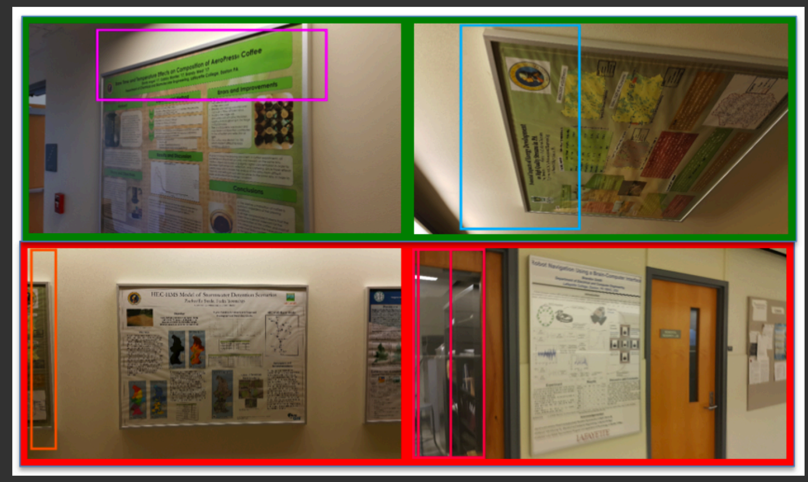
5. Perspective Transformation

- Randomly choose a viewpoint
- Create a 3D perspective transformation
- Convert the 3D transformation to 2D

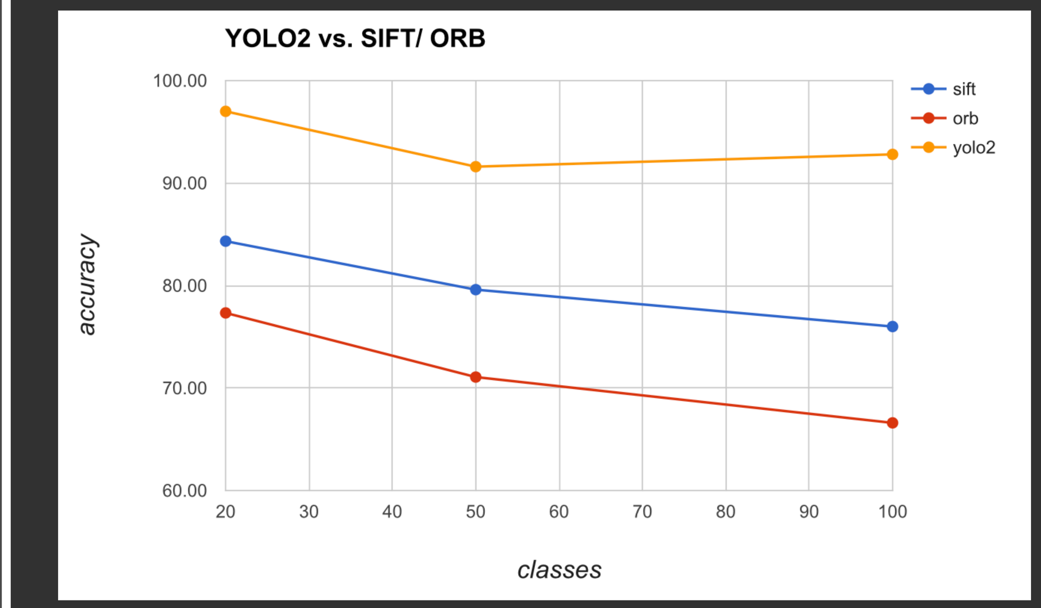
Results

Here are some examples of our results. The rest can be found at:

http://www.cs.lafayette.edu/~vut/poster-recognition/yolo2-100c/results.html



Performance compared to SIFT/ORB

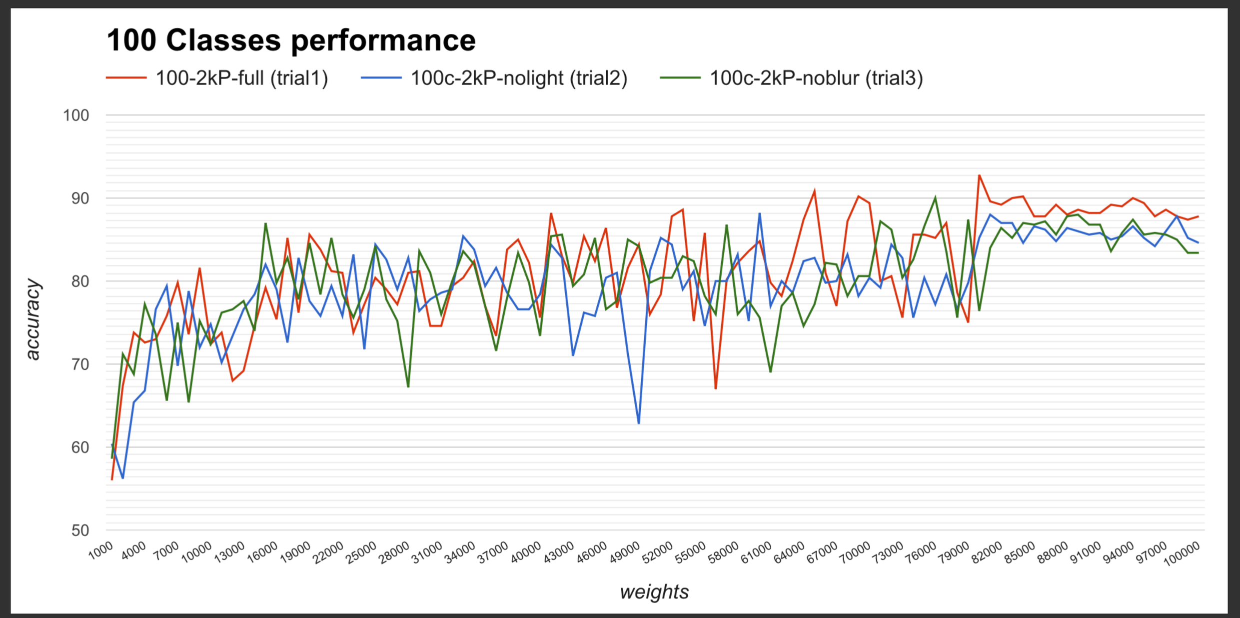


Training time

20 Posters		
Iteration	Time(hr)	Acc.
0.9k	2	90.0%
10k	24	97.0%
34k	36	100.0%

100 Posters		
Iteration	Time(hr)	Acc.
19k	36	85.6%
65k	132	90.8%
80k	168	92.8%

Effect of blurring and lighting



Conclusion

- We develop a fast and robust academic poster recognition algorithm, which uses deep convolutional neural networks for object detection
- Our proposed algorithm are more optimal for these types of images and has the potential for better prediction, especially in real life setting where posters tend to be partially occluded.
- We take advantage of the fact that academic posters are planar and propose a set of transformations in order to generate a large training set for the convolutional neural network. We then show through experiments how our method outperforms more traditional baselines.