Introduction

An Image Classifier
Brendan Miller

- An image classifier based on the generative/discriminative framework presented in Yi Li's paper.
- A "bag of words" classifier.
- Uses SIFT key-point descriptors and color.

Basic Idea

A quick high level summary of Li's classification framework:

- Build a mixture model of features in an images containing a class of objects.
- Determine how well a training set of images match each component of the model.
- Build a vector of these match values (joint probabilities).
- Use these values as the inputs to train a classifier (a neural network).

Generative/discriminative model part 1

Generative model.

- Gaussian mixture model of distribution of features constructed using EM algorithm.

$$P(X^a|o) = \sum_{m=1}^{M^a} w_m^a N(X^a; \mu_m^a, \Sigma_m^a)$$

- Joint probability of image features with each model component is calculated.

$$P(X_{i,r}^a, m^a) = w_m^a N(X_{i,r}^a, \mu_m^a, \Sigma_m^a)$$

- For each image and gaussian component, the maximum joint probability is found.

$$P(I_i, m^a) = max(\{P(X_{i,r}^a, m^a) | r \in \text{features of type } a \text{ in the } i\text{th image}$$

Generative/discriminative model part 2

Discriminative classifier.

- The maximum joint probabilities are used as inputs to train a neural network.
- The neural network acts as a classifier.
- The classifier understands not just how well an example image fit the (mixture of gaussians) model, but how well it fit *each component* of the model.
- Some components of the model found during the generative phase may also be present in images which do not contain the object we are looking for. Some components may be necessary but not sufficient to establish the presence of an object. The discriminative phase deals with these problems.

Features

- Original paper used regional features e.g. color, texture.
- My project also uses key-point descripters i.e. SIFT.
- I also keep color. Mainly useful for improving accuracy on face detection where color can detect the distinctive skin tone.
- Color is not that useful for classifying cars, planes, motorcycles which may be painted any color.

SIFT

- SIFT alone produces good results.
- Increasing the number of Gaussian components (50+) helps accuracy.
- Objects are made of many parts. Ideally each Gaussian component matches a part.

Correct classification

False Negative

Face detection more reliant on color than others.

Accuracy

Image Type	Accuracy	False Positives	False Negatives
Cars	0.94	0.035	0.025
Planes	0.78	0.11	0.11
Faces	0.9	0.055	0.045
Bikes	0.805	0.08	0.115

Future work

- One idea: How can we find which parts of an image contain the classified object?
- Given a classified image, it would be nice to be able to work backwards and find the features which contributed the most to the classification.
- Given the most important features, a bounding box could be created.
- A 3 layer neural network might not be ideal for this. It would be easier to work backwards through a simpler classifier.