2413, Machine Learning, Assignment 2
Universität Bern
Due date: 23/11/2016

1 Introduction

In the previous homework assignment we trained a classifier that distinguishes between face and non-face images. However this system lacks a few things that makes it useful in the real world. When we recognise a face we want to find it on arbitrary images, and not only decide whether 24×24 pixel image contains a face. This can be done using the sliding window approach. 24×24 sized patches can be extracted from the image. We can 'slide' through the image and use the classifier on the patches extracted from each location. When the classifier says the patch contains a face, we consider that as a detection, and the bounding box of the image patch gives us the location of the face. Multiple scales can be handled by resizing the image, and use the sliding window technique on the resized images.

Previously we used the raw pixel of the images to classify faces. This is the simplest feature representation we can use. There are features, which were crafted to achieve high performance on the detection task. In this homework assignment you will explore the effect of the feature representation on the performance. You will also have to evaluate the performance of the different learning algorithm we studied so far. The trained detector (hopefully) will be able to detect faces on real world images.

There are presentation slides in the references folder that explain and visualise the sliding window approach. They also contain the main ideas about the Histograms of Oriented Gradients (HOG), a very useful feature representation. You can also find two research papers in the references folder. They are not compulsory reading material for the course, but feel free to study it if you are interested. The first [1] explains the HOG descriptors in detail and the second [2] shows an effective SVM solver called Pegasos. It solves the SVM problem with stochastic gradient descent with a clever way of setting the learning rates.

2 Exercises

1. The script 'exercise1.m' learns a linear classifier based on Support Vector Machines on the face/pedestrian dataset (the same as in the previous assignment). You do not need to implement the SVM classifier, a solver Pegasos is provided. You need to run and evaluate the training on two datasets, and answer the questions below. Use the raw pixel values as a feature representation.

   - Train the SVM classifier with different choices of the regularisation parameter λ. Plot a graph where you show the classification performance versus λ. What is the optimal regularisation parameter?
   - Compare SVM to the previously implemented learning algorithms (logistic regression and GDA) in terms of classification accuracy.

2. In this exercise you need to use the HOG features for representing the samples. Try training the classifier with the three different learning algorithms (SVM, logistic regression and GDA) on two datasets.

   - Compare the HOG features to the raw pixels. Which feature representation is better for classification?
   - Compare the training algorithms (SVM, logistic regression and GDA) in terms of classification accuracy.
   - According to the experiments what is more important? The learning algorithm, or the feature representation?
   - On which dataset do HOG features help more? Justify your answer.

3. It is time to try machine learning in the real world. The script 'exercise3.m' implements the sliding window approach for face detection. The positive examples are the same as before, but the negatives are collected from the negative images by the sliding window. You can specify which features you want to use (raw pixels or HOG) in the first few lines of the code.

   - Before you run the experiment, what kind of performance do you expect from the raw pixels and the HOG features (in terms of average precision)?
   - What is the difference between the performances on the dataset used previously and the dataset used in this experiment? Explain why you get these results.

4. After training the sliding window classifier, run the detector (in 'demo.m') on your own images, or images downloaded from the internet. Show detections for images that contain faces and also on images that do not. Use the same threshold for all the images (for each feature representation). Comment on the usefulness of this face detector on real images.

3 References

1. Navneet Dalal and Bill Triggs: Histograms of Oriented Gradients for Human Detection, CVPR 2005

2. Shai Shalev-Shwartz, Yoram Singer, Nathan Srebro, Andrew Cotter: Pegasos: primal estimated sub-gradient solver for SVM, Mathematical Programming 2011