Visual Object Detection in Live Video Streams

Apical LTD, UK and University of Leicester, UK

Introduction

Computer vision systems are at a key point in their evolution. Their importance stretches across robotics and autonomous vehicles to automated approaches in health care. Research and development in this field has progressed over the last 60 years. Key themes in the literature include perceptrons, statistical theory of pattern classification, support vector machines, and neural networks.

Recently, extremely capable new technologies arrived: Convolutional Neural Networks, Deep Learning, Deep Belief, Random Field, and residual Neural Networks. They show remarkable performance in benchmark tests and applications.

Despite this success several obstacles impede further progress and limit the impact of these discoveries on society. These are:

1. Computational costs.
2. Autonomy and energy consumption.
3. Training of these advanced systems requires large data set of samples that is not necessarily available for all objects of interest and applications.

Solving all or a part of these limitations will enable immediate and broad access of the technology to society.

Specific problem

Apical LTD, a UK-based company, has developed a prototype of a low-cost and a low-energy system, featuring image representation with Histograms of Oriented Gradients at multiple spatial scales and linear Support Vector Machine-type classifiers for object detection. The system has been fully implemented in the hardware; this allows processing of high-resolution images and video streams in near real-time speed of approximately 30 frames per second. The system, however, has been trained to detect a restricted set of

objects, including human shapes and hands. The task is to train the system to detect a given new class of objects that is not in the list of objects it already knows.

Approach and work plan

On the first day of the school participating students will be provided with a tutorial on the particular system at hands and on the basics of Machine Learning Theory that is relevant for the project. Remaining days will be devoted to independent yet supervised take on the problem. In the process of solving this task the students will obtain first-hands experience in training a real functional object detection system. They will also acquire practical data mining and data analysis skills and will experiment with measure concentration effects in high dimensions. The class of the object detectors will initially be limited to Support Vector Machines. It can, however, be broadened at a later stage depending on the degree of success and level of engagement.

Recommended Literature

[1] R. O. Duda, G. Stork, P. E. Hart. Pattern classification and scene analysis, Wiley, 2000.
[2] V. Vapnik. The Nature of Statistical Learning Theory, Springer, 2000.
[3] N. Dalal, B. Triggs. Histograms of Oriented Gradients for Human Detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 1063-6919.