Statistical Methods in Artificial Intelligence

Human Detection in Images

Project Guide - Prof. Bapi Raju Surampudi

Members

Vinod Adwani	201506633
Madhudeep Petwal	201506585
Abhishek Singhal	201506572

Abstract

In surveillance systems, it is a basic task to detect and track the presence of human figure. Several applications like traffic monitoring, event detection etc. rely on standard Human Detection Algorithm.

Challenges include change in lighting conditions, occlusions, viewpoint variance poor video quality

Variations arise not only from changes in illumination and viewpoint, but also due to non-rigid deformations, and intraclass variability in shape and other visual properties. For e.g. people wear different clothes and take variety of poses.

We have used HOG (Histograms of Oriented Gradients) descriptors, a paper by **Navneet Dalal and Bill Triggs** and **Linear SVM Classifier** for Human detection.

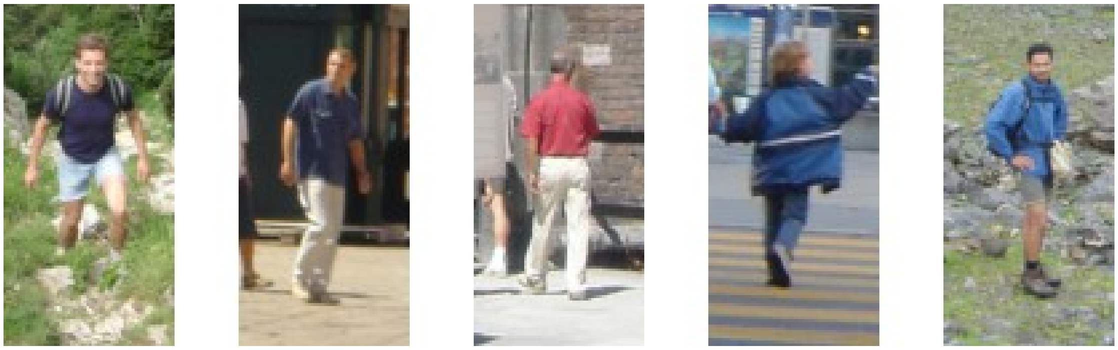
HOG uses a “global” feature to describe a person rather than a collection of “local” features. Put simply, this means that the entire person is represented by a single feature vector, as opposed to many feature vectors representing smaller parts of the person.

The HOG person detector uses a sliding detection window which is moved around the image. At each position of the detector window, a HOG descriptor is computed for the detection window. This descriptor is then shown to the trained SVM, which classifies it as either “person” or “not a person”.

To recognize persons at different scales, the image is subsampled to multiple sizes. Each of these subsampled images is searched.

Gradient Histograms

The HOG person detector uses a detection window that is 64 pixels wide by 128 pixels tall.
Below are some of the original images used to train the detector, cropped in to the 64×128 window



To compute the HOG descriptor, we operate on 8×8 pixel cells within the detection window. These cells will be organized into overlapping blocks, but we'll come back to that.
Here's a zoomed-in version of one of the images, with an 8×8 cell drawn in red, for an idea of the cell size and image resolution.

Feature Extraction

Within a cell, we compute the gradient vector at each pixel. We take the 64 gradient vectors (in our 8×8 pixel cell) and put them into a 9-bin histogram. The Histogram ranges from 0 to 180 degrees, so there are 20 degrees per bin.

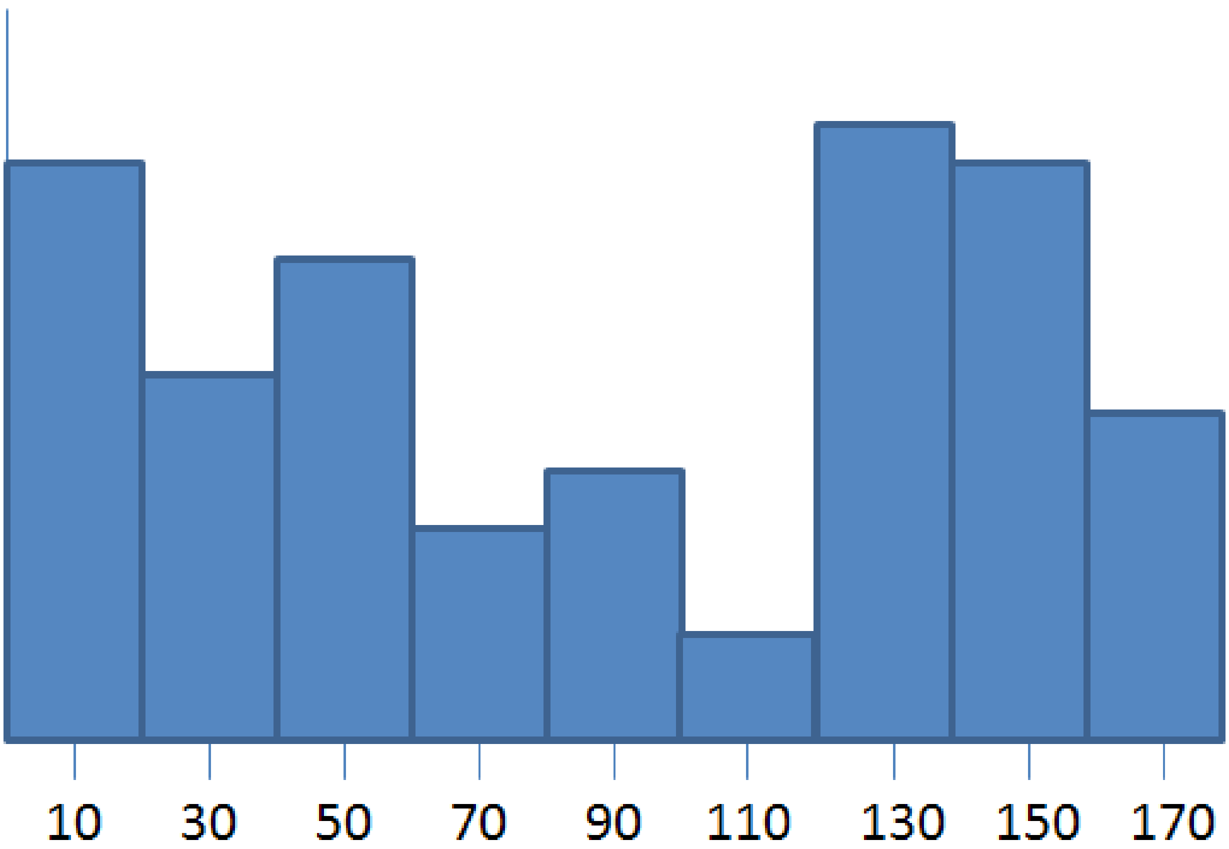


Fig. Feature extracted in form of bins

For each gradient vector, its contribution to the histogram is given by the magnitude of the vector (so stronger gradients have a bigger impact on the histogram). We split the contribution between the two closest bins. So, for example, if a gradient vector has an angle of 85 degrees, then we add 1/4th of its magnitude to the bin centered at 70 degrees, and 3/4ths of its magnitude to the bin.

Datasets

We have used 7000 positive dataset and 8000 images as negative dataset.

We applied linear SVM classifier and following are some results

Results

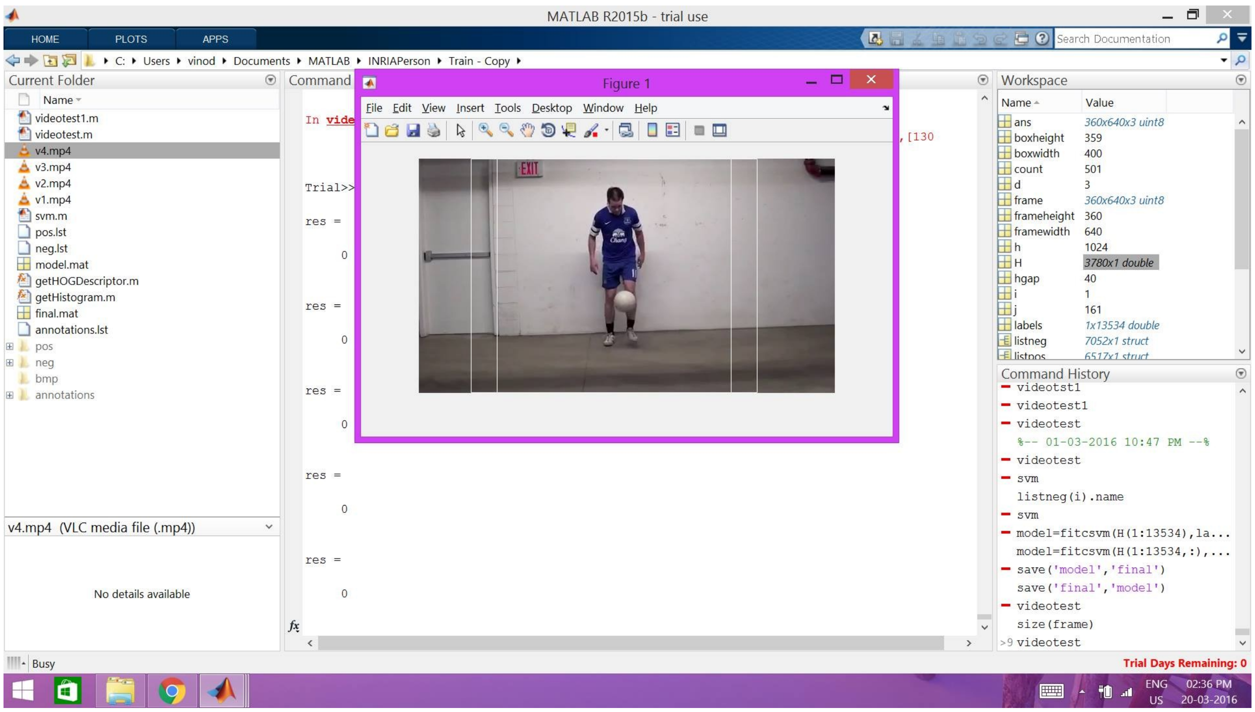


Fig. Detection of Human in videos

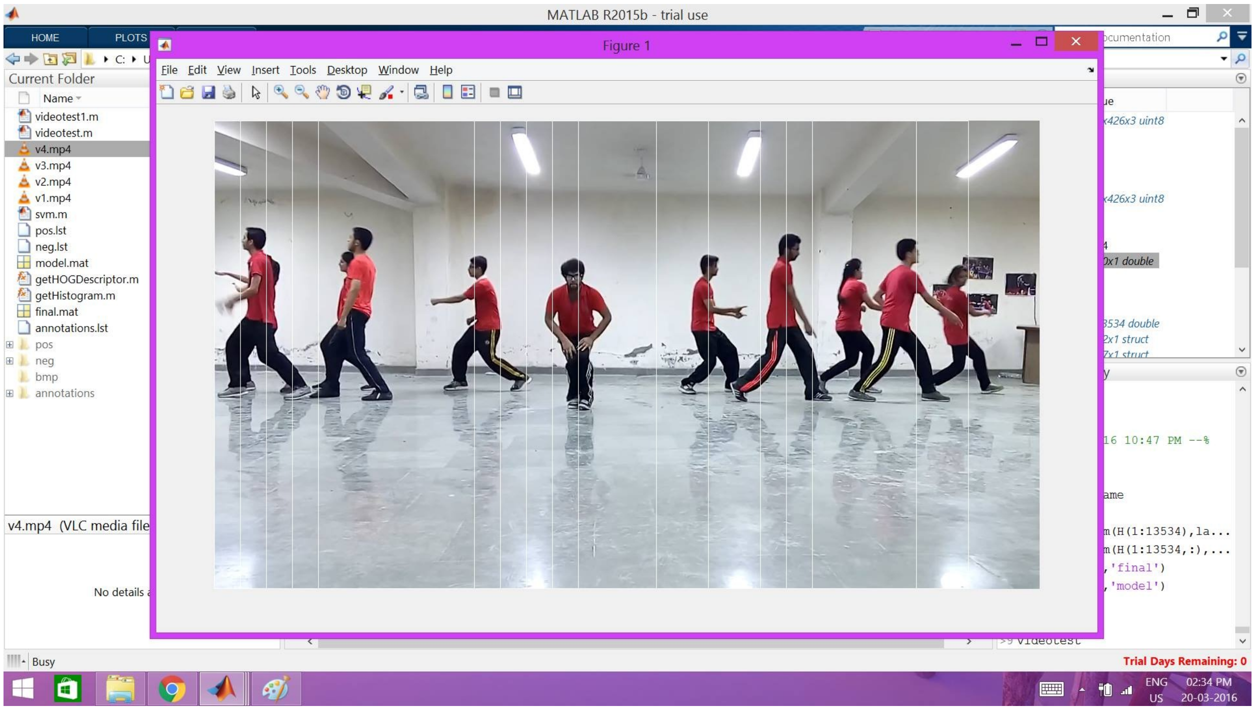


Fig. Detection of Human in videos

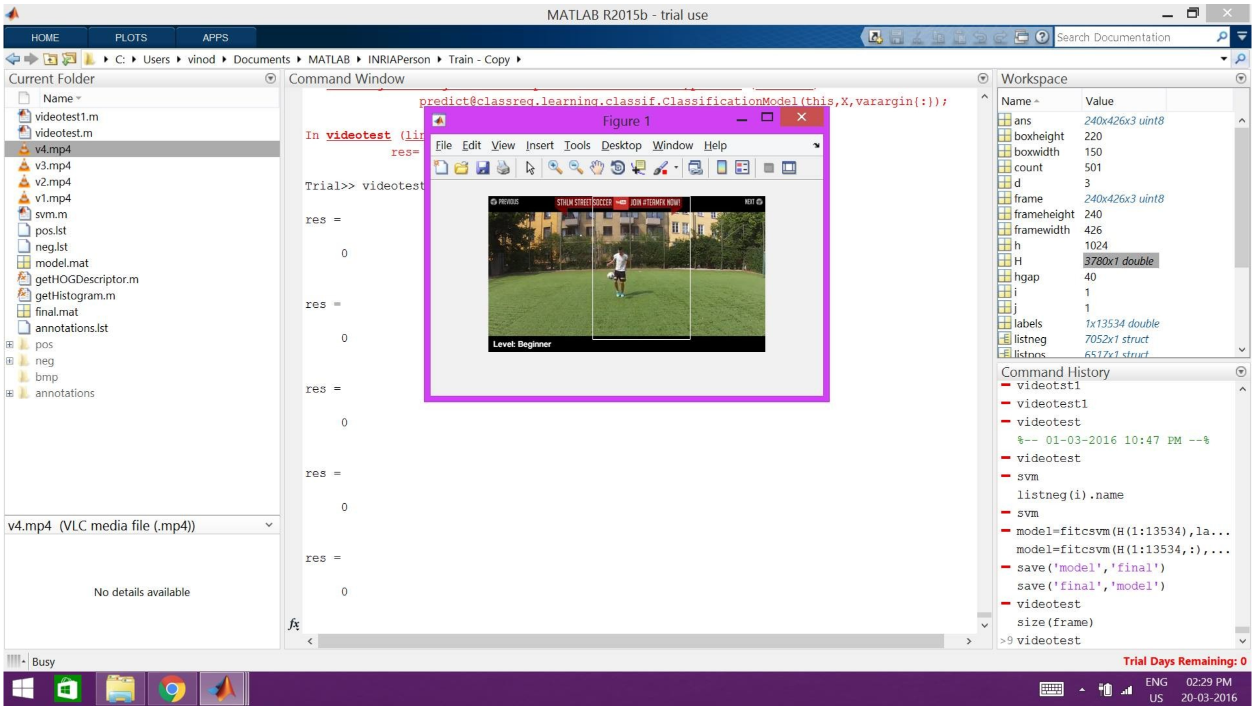


Fig. Detection of Human in videos

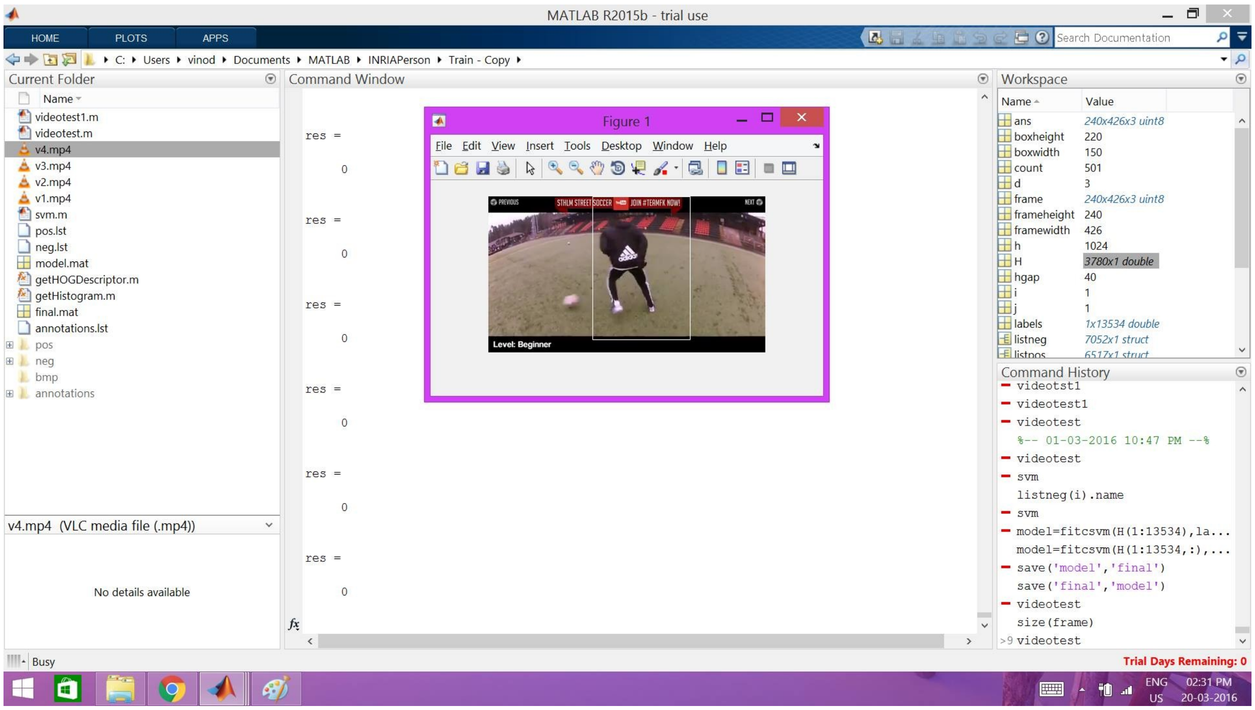


Fig. Detection of Human in videos

Bibliography

1. Navneet Dalal and Bill Triggs paper on Histograms of Oriented Gradients for Human Detection.
   https://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf

2. Tutorial on HOG descriptor
   https://chrisjmccormick.wordpress.com/2013/05/09/hog-person-detector-tutorial/

3. Video on HOG descriptor
   https://www.youtube.com/watch?v=0Zib1YEE4LU

4. Matlab official page for various commands
   http://in.mathworks.com/products/matlab/