

HOG Implementation approach

Preprocessing of Input Image

Image is converted from RGB to Gray scale and then resized to fixed dimension. Here, Images are resized to 128*64 pixels as suggested by the authors of INRIA datasets with main aim to obtain better results on detection of pedestrian.

Gradients calculation

- Images are converted to numpy array.
- Luminescence Gradients of image in x and y direction are calculated at each pixel using finite difference convolution mask i.e. [-1,0,1] in x direction and [1,0,-1] from top to bottom in y direction.
- Gradient magnitude is calculated as follows:

$$Magnitude(\mu) = \sqrt{G_x^2 + G_y^2}$$

-
- Gradient Orientation is calculated as follows:

$$Angle(\theta) = |\tan^{-1}(G_y/G_x)|$$

-
- where G(x), G(y) are gradients in x and y direction
- theta is orientation of gradient

HOG for each cell

9 – point histogram bins are created for each cell of 8*8 pixels using pixels magnitude and orientation. Each bin range/step size is of 20 degree. Bins from 0 to 8. J th bin having range J*20 degree to (J+1)*20degree.

Calculation of Bin height/frequency: If a pixel have angle theta between J*20 and (J+1)*20. Magnitude of that pixel is distributed between J and J+1 bins as follows:

For J : height increased = $Magnitude * ((J+1)*20 - \theta) / \text{step size (here 20 degrees)}$

And For J+1: height increased = $Magnitude * (\theta - J*20) / \text{step size (here 20 degrees)}$

Then 9 bins for each cell are prepared

HOG Normalization

Block containing 2*2 cells each are used for normalization of bins. This Clubbing of cells is done in overlapping manner with stride of 1 cell.

Normalization is performed using following equation:

$$v(n) = \frac{v(n)}{\sqrt{\left(\sum_{k=1}^{3 \times 3 \times 9} v(k)^2\right) + 1}}$$

where v(k) in denominator denotes magnitude of bins in a block ,v(n) in numerator denotes magnitude of bin under consideration ,V(n) in left denotes normalized bin magnitude

Feature Descriptor

In each block there are 4 cells each containing 9 normalized histogram bins i.e. total 36 features in each block. Image size is 128*64 pixels. Stride of block is 1 cell = 8 pixels. Means 16*8 cell image and 2*2 cell block implies movements = 15 in horizontal and 7 in vertical

Feature Descriptor Size = 15*7*36 = 3780.

Image is also analyzed for

- change in scale, by increasing size from 128*64 depending on scale factor provided to study effect of change of scale on HOG feature descriptor.
- Change in illumination by multiplying pixel gray level with illumination factor.
- and change in rotation by changing the rotation angle.

RMSE of feature vectors after and before change are calculated to quantify the effect of these changes. We also included many visualizations for more clarity.

Classification

Used Support vector machine classifier with linear kernel to classify images (from Yoga position dataset) based on the extracted features.

References: <https://medium.com/analytics-vidhya/a-gentle-introduction-into-the-histogram-of-oriented-gradients-fdee9ed8f2aa>

<http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>

<https://www.youtube.com/watch?v=QmYJCxJWdEs>