Supplementary Material

1. The INRIA Dataset

INRIA dataset was collected as part of research work on detection of upright people in images and video. The research is described in detail in CVPR 2005 paper [Histograms of Oriented Gradients for Human Detection](http://lear.inrialpes.fr/pubs/2005/DT05/) and their [PhD thesis](http://lear.inrialpes.fr/people/dalal/NavneetDalalThesis.pdf). The dataset is divided in two formats: (a) original images with corresponding annotation files, and (b) positive images in normalized 64x128 pixel format (as used in the CVPR paper) with original negative images.

The data set contains images from several different sources: Images from [GRAZ 01](http://www.emt.tugraz.at/~pinz/data/GRAZ_01/" \t "_blank) dataset, though annotation files are completely new. Images from personal digital image collections taken over a long time period. Usually the original positive images were of very high resolution (approx. 2592x1944 pixels), so we have cropped these images to highlight persons. Many people are bystanders taken from the backgrounds of these input photos, so ideally there is no particular bias in their pose. Few of images are taken from the web using google images. Only upright persons (with person height > 100) are marked in each image. Annotations may not be right; in particular at times portions of annotated bounding boxes may be outside or inside the object.

**Original Images.** Folders 'Train' and 'Test' correspond, respectively, to original training and test images. Both folders have three sub folders: (a) 'pos' (positive training or test images), (b) 'neg' (negative training or test images), and (c) 'annotations' (annotation files for positive images in [Pascal Challenge](http://www.pascal-network.org/challenges/VOC/) format).

**Normalized Images.** Folders 'train\_64x128\_H96' and 'test\_64x128\_H96' correspond to normalized dataset as used in above referenced paper. Both folders have two sub folders: (a) 'pos' (normalized positive training or test images centered on the person with their left-right reflections), (b) 'neg' (containing original negative training or test images). Note images in folder 'train/pos' are of 96x160 pixels (a margin of 16 pixels around each side), and images in folder 'test/pos' are of 70x134 pixels (a margin of 3 pixels around each side). This has been done to avoid boundary conditions (thus to avoid any particular bias in the classifier). In both folders, use the centered 64x128 pixels window for original detection task.

You may download the whole data set from [here](ftp://ftp.inrialpes.fr/pub/lear/douze/data/INRIAPerson.tar) (970MB). To avoid duplicating images, 'neg' image folder in 'train\_64x128\_H96' and 'test\_64x128\_H96' are referenced using symbolic links.

1. The Caltech Pedestrian Dataset

The Caltech Pedestrian Dataset consists of approximately 10 hours of 640x480 30Hz video taken from a vehicle driving through regular traffic in an urban environment. About 250,000 frames (in 137 approximately minute long segments) with a total of 350,000 bounding boxes and 2300 unique pedestrians were annotated. The annotation includes temporal correspondence between bounding boxes and detailed occlusion labels. More information can be found in [PAMI 2012](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/files/PAMI12pedestrians.pdf) and [CVPR 2009](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/files/CVPR09pedestrians.pdf) benchmarking papers.

[**Caltech Pedestrian Dataset**](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/USA/)**.** The training data (set00-set05) consists of six training sets (~1GB each), each with 6-13 one-minute long seq files, along with all annotation information (see the paper for details). The testing data (set06-set10) consists of five sets, again ~1GB each. New: annotations for the entire dataset are now also provided. Output files containing detection results for all evaluated algorithms are also available.

[**Seq video format**](https://pdollar.github.io/toolbox/videos/seqIo.html)**.** An seq file is a series of concatenated image frames with a fixed size header. Matlab routines for reading/writing/manipulating seq files can be found in [Piotr's Matlab Toolbox](https://pdollar.github.io/toolbox/index.html)(version 3.20 or later required). These routines can also be used to extract an seq file to a directory of images.

[**Matlab evaluation/labeling code (3.2.1)**](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/code/code3.2.1.zip)**.** The annotations use a custom "video bounding box" (vbb) file format. The code also contains utilities to view seq files with annotations overlaid, evaluation routines used to generate all the ROC plots in the paper, and also the vbb labeling tool used to create the dataset (see also this somewhat outdated [video tutorial](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/files/vbbLabelerTutorial-divx.avi)).

[**Additional datasets in standardized format**](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/)**.** For convenience we are posting full images/annotations in seq/vbb format as well as detection results for all evaluated algorithms on a number of additional datasets. This facilitates training/testing on these additional datasets and exact reproduction of all ROC curves. Full copyright remains with the original authors, please see the respective website for additional information including how to cite evaluation results on these datasets. [INRIA pedestrian dataset](http://pascal.inrialpes.fr/data/human/).[[converted](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/INRIA/)], [ETH pedestrian dataset](http://www.vision.ee.ethz.ch/~aess/dataset/) [[converted](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/ETH/)], [TUD-Brussels pedestrian dataset](http://www.d2.mpi-inf.mpg.de/tud-brussels) [[converted](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/TudBrussels/)], [Daimler pedestrian dataset](http://www.gavrila.net/Research/Pedestrian_Detection/Daimler_Pedestrian_Benchmark_D/Daimler_Mono_Ped__Detection_Be/daimler_mono_ped__detection_be.html) [[converted](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/datasets/Daimler/)].