Caltech Pedestrian Dataset:
Evaluated Algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Features</th>
<th>Classifier</th>
<th>Training</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>[19]</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td>evolution of ChnFtrs [source code]</td>
</tr>
<tr>
<td>ACF++</td>
<td>[35]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>ACF-Caltech</td>
<td>[19]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>ACF-Caltech+</td>
<td>[34]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>ACF+SDt</td>
<td>[45]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>AFS</td>
<td>[25]</td>
<td>linear SVM</td>
<td>INRIA</td>
<td>accelerated version ofFeatSynth</td>
</tr>
<tr>
<td>AFS+Geo</td>
<td>[25]</td>
<td>linear SVM</td>
<td>INRIA</td>
<td>variant of AFS with geometry constraints</td>
</tr>
<tr>
<td>CCF</td>
<td>[59]</td>
<td>deep</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>CCF+CF</td>
<td>[59]</td>
<td>deep+channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>Checkerboards</td>
<td>[64]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>Checkerboards+</td>
<td>[64]</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>ChnFtrs</td>
<td>[18]</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>CompACT-Deep</td>
<td>[8]</td>
<td>multiple</td>
<td>boosting</td>
<td>Caltech</td>
</tr>
<tr>
<td>ConvNet</td>
<td>[48]</td>
<td>pixels</td>
<td>DeepNet</td>
<td></td>
</tr>
<tr>
<td>Crosstalk</td>
<td>[15]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>DBN-Isol</td>
<td>[36]</td>
<td>HOG</td>
<td>DeepNet</td>
<td>INRIA</td>
</tr>
<tr>
<td>DeepParts</td>
<td>[51]</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech</td>
</tr>
<tr>
<td>F-DNN</td>
<td>[20]</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
</tr>
<tr>
<td>F-DNN+SS</td>
<td>[20]</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
</tr>
<tr>
<td>F-DNN2+SS</td>
<td>[21]</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
</tr>
<tr>
<td>FeatSynth</td>
<td>[3]</td>
<td>multiple</td>
<td>linear SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>FisherBoost</td>
<td>[49]</td>
<td>HOG+COV</td>
<td>FisherBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>FPDW</td>
<td>[16]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>FtrMine</td>
<td>[17]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>Franken</td>
<td>[32]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>HikSvm</td>
<td>[38]</td>
<td>HOG</td>
<td>HIK SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>HOG</td>
<td>[14]</td>
<td>HOG</td>
<td>linear SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>HOG-LBP</td>
<td>[56]</td>
<td>HOG+LBP</td>
<td>linear SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>InformedHaar</td>
<td>[63]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>JointDeep</td>
<td>[37]</td>
<td>color+gradient</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>Katamari</td>
<td>[6]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>LatSvm-V1</td>
<td>[22]</td>
<td>HOG</td>
<td>latent SVM</td>
<td>PASCAL</td>
</tr>
<tr>
<td>LatSvm-V2</td>
<td>[23]</td>
<td>HOG</td>
<td>latent SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>LDCF</td>
<td>[34]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>LDCF++</td>
<td>[35]</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>LFOV</td>
<td>[1]</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech</td>
</tr>
<tr>
<td>MLS</td>
<td>[33]</td>
<td>HOG</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>MOCO</td>
<td>[10]</td>
<td>HOG+LBP</td>
<td>latent SVM</td>
<td>Caltech</td>
</tr>
<tr>
<td>features</td>
<td>classifier</td>
<td>training</td>
<td>notes</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>MRFC+Semantic</td>
<td>channels</td>
<td>boosting</td>
<td>Caltech+CamVid+SiftFlow+KITTI data for segm.</td>
<td></td>
</tr>
<tr>
<td>MS-CNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet pre-training</td>
<td></td>
</tr>
<tr>
<td>MT-DPM</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>MT-DPM+Context</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech+context obtained from a vehicle detector</td>
<td></td>
</tr>
<tr>
<td>MultiFtr</td>
<td>multiple</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>MultiFtr+CSS</td>
<td>multiple</td>
<td>linear SVM</td>
<td>TUD-Motion</td>
<td></td>
</tr>
<tr>
<td>MultiFtr+Motion</td>
<td>multiple</td>
<td>linear SVM</td>
<td>TUD-Motion</td>
<td></td>
</tr>
<tr>
<td>MultiResC</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>MultiSDP</td>
<td>HOG+CSS</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>pAUCBoost</td>
<td>HOG+COV</td>
<td>pAUCBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>PCN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>Pia</td>
<td>multiple</td>
<td>PLS+QDA</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>PoseInv</td>
<td>HOG</td>
<td>AdaBoost</td>
<td>INRIA+</td>
<td></td>
</tr>
<tr>
<td>PoseInvSvm</td>
<td>HOG</td>
<td>kernel SVM</td>
<td>INRIA+</td>
<td></td>
</tr>
<tr>
<td>RandForest</td>
<td>HOG+LBP</td>
<td>random forest</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>Roerei</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>RPN+BF</td>
<td>pixels</td>
<td>DeepNet+AdaBoost</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>SA-FastRCNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>SCCPrriors</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>SCF+AlexNet</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>SDN</td>
<td>pixels</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>SDS-RCNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>Shapelet</td>
<td>gradients</td>
<td>AdaBoost</td>
<td>INRIA with boundary effects fixed [57]</td>
<td></td>
</tr>
<tr>
<td>Shapelet-orig</td>
<td>gradients</td>
<td>AdaBoost</td>
<td>INRIA original implementation</td>
<td></td>
</tr>
<tr>
<td>SketchTokens</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA+</td>
<td></td>
</tr>
<tr>
<td>SpatialPooling+</td>
<td>multiple</td>
<td>pAUCBoost</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>SpatialPooling</td>
<td>multiple</td>
<td>pAUCBoost</td>
<td>Caltech</td>
<td></td>
</tr>
<tr>
<td>TA-CNN</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech++</td>
<td></td>
</tr>
<tr>
<td>UDN+</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
<td></td>
</tr>
<tr>
<td>VeryFast</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>VJ</td>
<td>Haar</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>VJ-OpenCV</td>
<td>Haar</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>WordChannels</td>
<td>WordChannels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
<tr>
<td>*+2Ped</td>
<td>HOG</td>
<td>latent SVM</td>
<td>INRIA+</td>
<td></td>
</tr>
</tbody>
</table>

References

Pedestrian Detection with a Large-Field-Of-View Deep Network
ICRA 2015, Seattle, WA. 1

Real-Time Pedestrian Detection With Deep Network Cascades
BMVC 2015, Swansea, UK. 1

Part-based Feature Synthesis for Human Detection
ECCV 2010, Crete, Greece. 1
Pedestrian detection at 100 Frames Per Second  
CVPR 2012, Providence, Rhode Island. 1, 2

Seeking the strongest rigid detector  
CVPR 2013, Portland, OR. 2

Ten years of pedestrian detection, what have we learned?  
ECCV-CVRSUAD 2014, Zurich, Switzerland. 1

[7] G. Brazil, X. Yin, and X. Liu  
Illuminating Pedestrians via Simultaneous Detection & Segmentation  
ICCV 2017, Venice, Italy. 2

[8] Z. Cai, M. Saberian, and N. Vasconcelos  
Learning Complexity-Aware Cascades for Deep Pedestrian Detection  
ICCV 2015, Santiago, Chile. 1

A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection  
ECCV 2016, Amsterdam, The Netherlands. 2

[10] G. Chen, Y. Ding, J. Xiao, and T. Han  
Detection Evolution with Multi-order Contextual Co-occurrence.  
CVPR 2013, Portland, OR. 1

Word Channel Based Multiscale Pedestrian Detection  
Without Image Resizing and Using Only One Classifier  
CVPR 2014, Columbus, Ohio. 2

[12] A. D. Costea and S. Nedevschi  
Semantic Channels for Fast Pedestrian Detection  
CVPR 2016, Las Vegas, Nevada. 2

Fast Pedestrian Detection for Mobile Devices  
ITSC 2015, Canary Islands. 1

Histogram of Oriented Gradient for Human Detection  
CVPR 2005, San Diego, California. 1

Crosstalk Cascades for Frame-Rate Pedestrian Detection  
ECCV 2012, Florence Italy. 1

The Fastest Pedestrian Detector in the West  
BMVC 2010, Aberystwyth, UK. 1

Feature Mining for Image Classification  
CVPR 2007, Minneapolis, Minnesota. 1
Integral Channel Features

Fast Feature Pyramids for Object Detection
PAMI, 2014. 1

Fused DNN: A deep neural network fusion approach to fast and robust pedestrian detection
arXiv, 2016. 1

Fused Deep Neural Networks for Efficient Pedestrian Detection
arXiv, 2018. 1

[22] P. Felzenszwalb, D. McAllester, D. Ramanan
A Discriminatively Trained, Multiscale, Deformable Part Model
CVPR 2008, Anchorage, Alaska. 1

[23] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan
Object Detection with Discriminatively Trained Part Based Models
PAMI 2010. 1

Taking a Deeper Look at Pedestrians
CVPR 2015, Boston, Massachusetts. 2

Fast multiple-part based object detection using KD-Ferns
CVPR 2013, Portland, OR. 1

Scale-aware Fast R-CNN for Pedestrian Detection
arXiv, 2016. 2

[27] J. Lim, C. Lawrence Zitnick, P. Dollár
Sketch Tokens: A Learned Mid-level Representation for Contour and Object Detection
CVPR 2013, Portland, OR. 2

[28] Z. Lin and L. Davis
A Pose-Invariant Descriptor for Human Detection and Segmentation
ECCV 2008, Marseille, France. 2

[29] P. Luo, Y. Tian, X. Wang, and X. Tang
Switchable Deep Network for Pedestrian Detection
CVPR 2014, Columbus, Ohio. 2

Classification Using Intersection Kernel Support Vector Machines is efficient
CVPR 2008, Anchorage, Alaska. 1

Random Forests of Local Experts for Pedestrian Detection
ICCV 2013, Sydney, Australia. 2
Handling Occlusions with Franken-classifiers
ICCV 2013, Sydney, Australia. 1

[33] W. Nam, B. Han, and J. H. Han
Improving Object Localization Using Macrofeature Layout Selection
ICCV Workshop on Visual Surveillance 2011, Barcelona, Spain. 1

[34] W. Nam, P. Dollár, and J. H. Han
Local Decorrelation For Improved Pedestrian Detection
NIPS 2014, Montreal, Quebec. 1

To Boost or Not to Boost? On the Limits of Boosted Trees for Object Detection
ICPR 2016, Cancun, Mexico. 1

[36] W. Ouyang and X. Wang
A Discriminative Deep Model for Pedestrian Detection with Occlusion Handling
CVPR 2012, Providence, RI. 1

[37] W. Ouyang and X. Wang
Joint Deep Learning for Pedestrian Detection
ICCV 2013, Sydney, Australia. 1

[38] W. Ouyang and X. Wang
Single-pedestrian detection aided by multi-pedestrian detection.
CVPR 2013, Portland, OR. 1, 2

Modeling Mutual Visibility Relationship with a Deep Model in Pedestrian Detection
CVPR 2013, Portland, OR. 1

[40] W. Ouyang, H. Zhou, H. Li, Q. Li, J. Yan and X. Wang
Jointly learning deep features, deformable parts, occlusion and classification for pedestrian detection
PAMI, 2017. 2

[41] S. Paisitkriangkrai, C. Shen, A. van den Hengel
Efficient pedestrian detection by directly optimize the partial area under the ROC curve
ICCV 2013, Sydney, Australia. 2

Strengthening the Effectiveness of Pedestrian Detection
ECCV 2014, Zurich, Switzerland. 2

Pedestrian Detection with Spatially Pooled Features and Structured Ensemble Learning
arXiv, 2014. 2

[44] D. Park, D. Ramanan, C. Fowlkes
Multiresolution models for object detection
ECCV 2010, Crete, Greece. 2

Exploring Weak Stabilization for Motion Feature Extraction
CVPR 2013, Portland, OR. 1
[46] P. Sabzmeydani and G. Mori
Detecting pedestrians by learning shapelet features
CVPR 2007, Minneapolis, Minnesota. 2

Human Detection Using Partial Least Squares Analysis
ICCV 2009, Kyoto, Japan. 2

[48] P. Sermanet, K. Kavukcuoglu, S. Chintala, Y. LeCun
Pedestrian Detection with Unsupervised Multi-Stage Feature Learning
CVPR 2013, Portland, OR. 1

Training Effective Node Classifiers for Cascade Classification
IJCV 2013. 1

[50] Y. Tian, P. Luo, X. Wang, and X. Tang
Pedestrian Detection aided by Deep Learning Semantic Tasks
CVPR 2015, Boston, Massachusetts. 2

[51] Y. Tian, P. Luo, X. Wang, and X. Tang
Deep Learning Strong Parts for Pedestrian Detection
ICCV 2015, Santiago, Chile. 1

[52] C. Toca, M. Ciuc, and C. Patrascu
Normalized Autobinomial Markov Channels For Pedestrian Detection
BMVC 2015, Swansea, UK. 2

[53] P. Viola and M. Jones
Robust Real-Time Face Detection
IJCV 2004. 2

[54] S. Walk, N. Majer, K. Schindler, B. Schiele
New Features and Insights for Pedestrian Detection
CVPR 2010, San Francisco, California. 2

PCN: Part and context information for pedestrian detection with CNNs
BMVC 2017, London, UK. 2

[56] X. Wang, T. X. Han, and S. Yan
An HOG-LBP Human Detector with Partial Occlusion Handling
ICCV 2009, Kyoto, Japan. 1

[57] C. Wojek and B. Schiele
A Performance Evaluation of Single and Multi-Feature People Detection
DAGM 2008, Munich, Germany. 2

[58] J. Yan, X. Zhang, Z. Lei, S. Liao, S. Z. Li
Robust Multi-Resolution Pedestrian Detection in Traffic Scenes
CVPR 2013, Portland, OR. 2

[59] B. Yang, J. Yan, Z. Lei, and S. Z. Li
Convolutional Channel Features
ICCV 2015, Santiago, Chile. 1
[60] Y. Yang, Z. Wang, and F. Wu
Exploring Prior Knowledge for Pedestrian Detection
BMVC 2015, Swansea, UK. 2

[61] X. Zeng, W. Ouyang, X. Wang
Multi-Stage Contextual Deep Learning for Pedestrian Detection
ICCV 2013, Sydney, Australia. 2

Is Faster R-CNN Doing Well for Pedestrian Detection?
ECCV 2016, Amsterdam, The Netherlands. 2

[63] S. Zhang, C. Bauckhage, and A. B. Cremers
Informed Haar-like Features Improve Pedestrian Detection
CVPR 2014, Columbus, Ohio. 1

[64] S. Zhang, R. Benenson, and B. Schiele
Filtered channel features for pedestrian detection
CVPR 2015, Boston, Massachusetts. 1