<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>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>evolution of ChnFtrs [source code]</td>
</tr>
<tr>
<td>ACF++</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>evolution of ChnFtrs [source code]</td>
</tr>
<tr>
<td>ACF-Caltech</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>uses deeper trees and denser sampling</td>
</tr>
<tr>
<td>ACF-Caltech+</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>SDt = Stabilized Dt (motion features)</td>
</tr>
<tr>
<td>ACF+SDt</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>ImageNet+CityPersons data</td>
</tr>
<tr>
<td>AdaptFasterRCNN</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet pre-training</td>
</tr>
<tr>
<td>ADM</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+</td>
<td>accelerated version ofFeatSynth</td>
</tr>
<tr>
<td>AFS</td>
<td>multiple</td>
<td>linear SVM</td>
<td>INRIA</td>
<td>variant of AFS with geometry constraints</td>
</tr>
<tr>
<td>AFS+Geo</td>
<td>multiple</td>
<td>linear SVM</td>
<td>INRIA</td>
<td>ImageNet pre-training [source code]</td>
</tr>
<tr>
<td>AR+PED</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+</td>
<td>Checkerboards + flow-based features from [47]</td>
</tr>
<tr>
<td>CCF</td>
<td>deep</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>updated (see addendum on author website)</td>
</tr>
<tr>
<td>CCF+CF</td>
<td>deep+channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>uses Caltech+ETH+Daimler for training</td>
</tr>
<tr>
<td>Checkerboards</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>100 fps on a CPU</td>
</tr>
<tr>
<td>Checkerboards+</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
<td>ImageNet data</td>
</tr>
<tr>
<td>ChnFtrs</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td>ImageNet+ETH+TudBrussels data</td>
</tr>
<tr>
<td>CompACT-Deep</td>
<td>multiple</td>
<td>boosting</td>
<td>Caltech</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>ConvNet</td>
<td>pixels</td>
<td>DeepNet</td>
<td>INRIA</td>
<td>ImageNet+ETH+TudBrussels data</td>
</tr>
<tr>
<td>Crosstalk</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td>ImageNet data</td>
</tr>
<tr>
<td>DBN-Isol</td>
<td>HOG</td>
<td>DeepNet</td>
<td>INRIA</td>
<td>ImageNet+ETH+TudBrussels data</td>
</tr>
<tr>
<td>DBN-Mut</td>
<td>HOG</td>
<td>DeepNet</td>
<td>INRIA/Caltech</td>
<td>ImageNet+ETH+TudBrussels data</td>
</tr>
<tr>
<td>DeepCascade</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>DeepCascade+</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>DeepParts</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>FastCF</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
<td>accelerated variant of ChnFtrs</td>
</tr>
<tr>
<td>FasterRCNN+ATT</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet data</td>
</tr>
<tr>
<td>F-DNN</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet+ETH+TudBrussels data</td>
</tr>
<tr>
<td>F-DNN+SS</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>F-DNN2+SS</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>ImageNet+Cityscapes+ETH+TudBrussels data</td>
</tr>
<tr>
<td>FtrMine</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td>multiple occlusion specific models</td>
</tr>
<tr>
<td>Franken</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
<td>ImageNet data</td>
</tr>
<tr>
<td>GDFL</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
<td>boundary effect fixed since publication</td>
</tr>
<tr>
<td>HikSvm</td>
<td>HOG</td>
<td>HIK SVM</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>HOG</td>
<td>HOG</td>
<td>linear SVM</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>HOG-LBP</td>
<td>HOGLBP</td>
<td>linear SVM</td>
<td>INRIA</td>
<td></td>
</tr>
<tr>
<td>InformedHaar</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
<td></td>
</tr>
</tbody>
</table>

Caltech Pedestrian Dataset: Evaluated Algorithms
<table>
<thead>
<tr>
<th>features</th>
<th>classifier</th>
<th>training</th>
<th>notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>JointDeep</td>
<td>color+gradient</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>Katamari</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>LatSvm-V1</td>
<td>HOG</td>
<td>latent SVM</td>
<td>PASCAL</td>
</tr>
<tr>
<td>LatSvm-V2</td>
<td>HOG</td>
<td>latent SVM</td>
<td>INRIA</td>
</tr>
<tr>
<td>LDCF</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>LDCF++</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>LFOV</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech</td>
</tr>
<tr>
<td>MLS</td>
<td>HOG</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>MOCO</td>
<td>HOG+LBP</td>
<td>latent SVM</td>
<td>Caltech</td>
</tr>
<tr>
<td>MRFC+Semantic</td>
<td>channels</td>
<td>boosting</td>
<td>Caltech+</td>
</tr>
<tr>
<td>MS-CNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>MT-DPM</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech</td>
</tr>
<tr>
<td>MT-DPM+Context</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech+</td>
</tr>
<tr>
<td>MultiFtr</td>
<td>multiple</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>MultiFtr+CSS</td>
<td>multiple</td>
<td>linear SVM</td>
<td>TUD-Motion</td>
</tr>
<tr>
<td>MultiFtr+Motion</td>
<td>multiple</td>
<td>linear SVM</td>
<td>TUD-Motion</td>
</tr>
<tr>
<td>MultiResC</td>
<td>HOG</td>
<td>latent SVM</td>
<td>Caltech</td>
</tr>
<tr>
<td>MultiSDP</td>
<td>HOG+CSS</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>NAMC</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>pAUCBoost</td>
<td>HOG+COV</td>
<td>pAUCBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>PCN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>Pls</td>
<td>multiple</td>
<td>PLS+QDA</td>
<td>INRIA</td>
</tr>
<tr>
<td>PoseInv</td>
<td>HOG</td>
<td>AdaBoost</td>
<td>INRIA+</td>
</tr>
<tr>
<td>posesvm</td>
<td>HOG</td>
<td>kernel SVM</td>
<td>INRIA+</td>
</tr>
<tr>
<td>RandForest</td>
<td>HOG+LBP</td>
<td>random forest</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>Roerei</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>RPN+BF</td>
<td>pixels</td>
<td>DeepNet+AdaBoost</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>SA-FastRCNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>SCCF</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>SCF+AlexNet</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>SDN</td>
<td>pixels</td>
<td>deep net</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>SDS-RCNN</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>Shapelet</td>
<td>gradients</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>Shapelet-orig</td>
<td>gradients</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>ShearFtrs</td>
<td>channels</td>
<td>AdaBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>SketchTokens</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA+</td>
</tr>
<tr>
<td>SpatialPooling</td>
<td>multiple</td>
<td>pAUCBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>SpatialPooling+</td>
<td>multiple</td>
<td>pAUCBoost</td>
<td>Caltech</td>
</tr>
<tr>
<td>TLL-TFA</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+</td>
</tr>
<tr>
<td>TA-CNN</td>
<td>pixels</td>
<td>DeepNet</td>
<td>Caltech+</td>
</tr>
<tr>
<td>UDNN+</td>
<td>pixels</td>
<td>deep net</td>
<td>Caltech+ImageNet</td>
</tr>
<tr>
<td>VeryFast</td>
<td>channels</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>VJ</td>
<td>Haar</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>VJ-OpenCV</td>
<td>Haar</td>
<td>AdaBoost</td>
<td>INRIA</td>
</tr>
<tr>
<td>WordChannels</td>
<td>WordChannels</td>
<td>AdaBoost</td>
<td>INRIA/Caltech</td>
</tr>
<tr>
<td>*+2Ped</td>
<td>HOG</td>
<td>latent SVM</td>
<td>INRIA+</td>
</tr>
</tbody>
</table>

CamVid+SiftFlow+KITTI data for segm. ImageNet pre-training context obtained from a vehicle detector trained with annotated silhouettes trained with annotated silhouettes Caltech results include context (CGP) ImageNet pre-training ImageNet pre-training ImageNet pre-training ImageNet pre-training with boundary effects fixed [61] original implementation features based on shearlet transform Sketch Tokens were trained on BSDS improved version of [43, 44] + flow features augmented with external data ImageNet pre-training implementation from [61] implementation from OpenCV

References

Pedestrian Detection with a Large-Field-Of-View Deep Network
ICRA 2015, Seattle, WA. 2

2
   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. 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. 2

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

[8] G. Brazil and X. Liu
   Pedestrian Detection with Autoregressive Network Phases
   CVPR 2019, Long Beach, CA. 1

   Learning Complexity-Aware Cascades for Deep Pedestrian Detection
   ICCV 2015, Santiago, Chile. 1

[10] Z. Cai, Q. Fan, R. Feris, and N. Vasconcelos
    A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection
    ECCV 2016, Amsterdam, The Netherlands. 2

    Detection Evolution with Multi-order Contextual Co-occurrence.
    CVPR 2013, Portland, OR. 2

[12] A. D. Costea and S. Nedevschi
    Word Channel Based Multiscale Pedestrian Detection
    Without Image Resizing and Using Only One Classifier
    CVPR 2014, Columbus, Ohio. 2

    Semantic Channels for Fast Pedestrian Detection
    CVPR 2016, Las Vegas, Nevada. 2

[14] A. D. Costea, A. Vesa, and S. Nedevschi
    Fast Pedestrian Detection for Mobile Devices
    ITSC 2015, Canary Islands. 1

    Histogram of Oriented Gradient for Human Detection
    CVPR 2005, San Diego, California. 1
[16] P. Dollár, R. Appel and W. Kienzle
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, 2

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

[23] P. Felzenszwalb, D. McAllester, D. Ramanan
A Discriminatively Trained, Multiscale, Deformable Part Model
CVPR 2008, Anchorage, Alaska. 2

Object Detection with Discriminatively Trained Part Based Models
PAMI 2010. 2

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

[27] J. Li, X. Liang, S. Shen, T. Xu, and S. Yan
Scale-aware Fast R-CNN for Pedestrian Detection
arXiv, 2016. 2

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

[29] C. Lin, L. Jiwen, G. Wang, and J. Zhou
Graininess-Aware Deep Feature Learning for Pedestrian Detection
ECCV 2018, Munich, Germany. 1
A Pose-Invariant Descriptor for Human Detection and Segmentation
ECCV 2008, Marseille, France. 2

[31] 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

[34] M. Mathias, R. Benenson, R. Timofte, L. Van Gool
Handling Occlusions with Franken-classifiers
ICCV 2013, Sydney, Australia. 1

[35] W. Nam, B. Han, and J. H. Han
Improving Object Localization Using Macrofeature Layout Selection
ICCV Workshop on Visual Surveillance 2011, Barcelona, Spain. 2

[36] W. Nam, P. Dollár, and J. H. Han
Local Decorrelation For Improved Pedestrian Detection
NIPS 2014, Montreal, Quebec. 1, 2

[37] E. Ohn-Bar and M. Trivedi
To Boost or Not to Boost? On the Limits of Boosted Trees for Object Detection
ICPR 2016, Cancun, Mexico. 1, 2

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

[39] W. Ouyang and X. Wang
Joint Deep Learning for Pedestrian Detection
ICCV 2013, Sydney, Australia. 2

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

[41] W. Ouyang, X. Zeng and X. Wang
Modeling Mutual Visibility Relationship with a Deep Model in Pedestrian Detection
CVPR 2013, Portland, OR. 1

Jointly learning deep features, deformable parts, occlusion and classification for pedestrian detection
PAMI, 2017. 2

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

[46] 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, 2

[48] L. Pfeifer  
Shearlet Features for Pedestrian Detection  
Journal of Mathematical Imaging and Vision, 2019. 2

[49] 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

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

[52] C. Shen, P. Wang, S. Paisitkriangkrai, A. van den Hengel  
Training Effective Node Classifiers for Cascade Classification  
IJCV 2013. 1

[53] T. Song, L. Sun, D. Xie, H. Sun, S. Pu  
Small-scale Pedestrian Detection Based on Somatic Topology Localization and Temporal Feature Aggregation  
ECCV 2018, Munich, Germany. 2

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

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

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

[57] P. Viola and M. Jones  
Robust Real-Time Face Detection  
IJCV 2004. 2
[58] 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

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

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

Robust Multi-Resolution Pedestrian Detection in Traffic Scenes
CVPR 2013, Portland, OR. 2

[63] B. Yang, J. Yan, Z. Lei, and S. Z. Li
Convolutional Channel Features
ICCV 2015, Santiago, Chile. 1

[64] Y. Yang, Z. Wang, and F. Wu
Exploring Prior Knowledge for Pedestrian Detection
BMVC 2015, Swansea, UK. 2

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

[66] L. Zhang, L. Lin, X. Liang, and K. He
Is Faster R-CNN Doing Well for Pedestrian Detection?
ECCV 2016, Amsterdam, The Netherlands. 2

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

Filtered channel features for pedestrian detection
CVPR 2015, Boston, Massachusetts. 1

[69] S. Zhang, R. Benenson, and B. Schiele
CityPersons: A Diverse Dataset for Pedestrian Detection
CVPR 2017, Honolulu, Hawaii. 1

[70] S. Zhang, J. Yang, and B. Schiele
Occluded Pedestrian Detection Through Guided Attention in CNNs
CVPR 2018, Salt Lake City, Utah. 1

[71] X. Zhang, L. Cheng, B. Li, and H. Hu
Too Far to See? Not Really!—Pedestrian Detection with Scale-aware Localization Policy
TIP, 2018. 1