ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Object Detection and Tracking of Pedestrian for ADAS

Miss.Kaiche Pranali R
PG Student,
E & TC Department,
Matoshri College of Engineering,
Nashik, Maharashtra, India.

Prof.Khule R.S
Professor,
E&TC Department,
Matoshri College of Engineering,
Nashik, Maharashtra, India.

Abstract:

Pedestrian protection system in advanced driver assistance system is become active research area in many developed and also developing countries for improving traffic safety. The main challenge of pedestrian detection is development of reliable and affordable on board pedestrian detection system. Due to non-rigid appearance of pedestrians like different clothes, aspect ratio, changing size and dynamic shape etc. and varying environmental conditions, it is very hard to cope with the high robustness of this system. In this survey we explicitly addressed the detail overview of current advances in the field. Image segmentation, feature extraction and subsequent classification process are discussed separately to focus on novelty of recent research. Moreover, we discussed limitation of state of art and promising direction for further work.

Keywords:

Classifiers, computer vision, feature extraction, fusion algorithm, pedestrian detection.

I. Introduction:

Detecting pedestrians in still images and real time videos is one of the fundamental tasks of pattern recognition and computer vision. Recently, pedestrian detection becomes one of the most interesting and potentially useful challenges for researchers. Two problems facing researchers because of which many proposed systems or algorithms cannot be get implemented are lack of public benchmarks and difficulty in reproducing many of proposed methods, which makes it difficult to compare approaches.

In this paper we present steps for detecting pedestrians, methods discussed which implemented by earlier researchers. The objective of a pedestrian detection system is to detect the presence of both stationary and moving people in a specific area of interest around the moving host vehicle in order to warn the driver, perform braking actions, and deploy external airbags if a collision is unavoidable. Fig 1 shows workflow of pedestrian detection system.

A. Challenges and characteristics of domain

- The appearance of pedestrians shows very high variability as they can change pose, wear different clothes, carry different conditions that vary the quality of the sensed image. In addition, pedestrians may be partially occluded by common urban elements, like parked vehicles.
- Pedestrians must be identified in highly dynamic scenes since both the pedestrian and camera are in motion, which make difficulty in tracking and movement analysis. Also, pedestrians appear at different viewing angles and the system must work over a large range of distances.
- The required performance is must be in terms of system reaction time and robustness of objects, and have a considerable range of sizes (specially interis of height).Pedestrians must be identified in outdoor urban scenarios, i.e., they must be detected in the context of a cluttered background under a wide range of light and weather.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The methods of pedestrian detection is based on feature extraction and machine learning. Pedestrian detection is very difficult task from a machine vision perspective. The lack of direct models leads to the use of machine learning techniques, where an indirect representation is learned from examples. As such, it is an instantiation of the multiclass object categorization problem (e.g., [1]).The pedestrian detection task has some of its own characteristics, which can effect the methods of choice. The mostly used machine learning algorithm include Adaboost [2], SVM [3], neural networks [4–5]. Extracting effective features and developing powerful learning algorithms have always been the research focus for pedestrian detection problem. In the case of feature representation, various features are proposed.

The FDF (Four Direction Feature) is a comparatively simple feature [6–7], of which the computational complexity is between the Haar feature [8] and HOG (Histograms of Oriented Gradients) [9]. The FDF can extract gradient feature of the pedestrian which is suitable for pedestrian detections. Multiple features are used for pattern recognition applications, for example face recognition [10–11] and human detection [12]. The detection of pedestrian is typically part of a system situated in a physical environment, which means that prior scene knowledge is often available to improve performance. Also very much efforts have been spent to collect extensive databases. Many techniques have been proposed for pedestrian detection in terms of features, models, and general architectures.

The picture is blurred on the experimental side. Reported performances get differ by several orders of magnitude (e.g., within the same study [13]or [14] versus [13]). This comes from the different types of image data used (degree of background change), the limited size of the test data sets, and the different (often, not fully specified) evaluation criteria such as localization tolerance, coverage area, etc.

II. WORKFLOW OF SYSTEM

Working flow of system is depicted as below

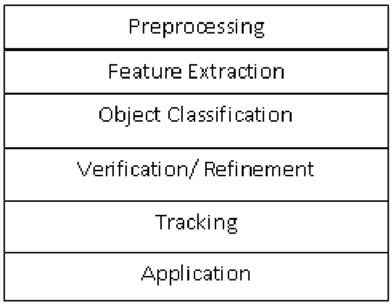

Fig.1. work flow of pedestrian detection system

A. PREPROCESSING

The preprocessing consists of tasks for instance exposure time, gain adjustments, and camera calibration etc. Low-level adjustments are normally not illustrate in ADAS literature, some researchers have targeted image enhancement through these systems. There are two approaches monocular [15] and stereo vision based [16].

B. FEATURE EXTRACTION

After pre-processing feature extraction is the step in most object detection and pattern recognization algorithm. Feature extraction and representation is a important step for multimedia processing. Extraction of ideal features that can reflect the intrinsic content of the images is still a challenging problem in computer vision. The common visual features consist of color, texture and shape, etc. [1-9], and most image annotation and retrieval systems have been constructed based on these features. However, their performance is heavily dependent on the use of image features.

In general, there are three basic feature representation methods, which include global, block-based, and region-based features. In feature extraction process, dominant features are extracted from a enormous number of training samples. These obtained features are then used to train a classifier. This general method has shown to work very well in detection of many different objects detection schemes.
Different features including local receptive fields (LRF) feature [4], and Haar wavelets [5] are used to train neural networks, support vector machines (SVM)

[6], [7] and k-NN classifiers. A widely used features for pedestrian detections are HOG which encode high frequency high frequency gradients information, Haar wavelet it encode lower frequency changes in color channels and oriented histograms of flow features which exploit optical flow and thus a complementary cue.

We classify the different features as:

- General features: Application independent features such as color, texture, and shape. According to the abstraction level, they can be divided into: Pixel-level features: Features calculated at each pixel, for ex. color, location.
- Local features: features calculated over the results of subdivision of the image band on image segmentation or edge detection. Global features: Features calculated over the entire image or just regular sub-area of an image.
- Domain-specific features: Application dependent features as fingerprints, human faces and also conceptual features. These features are frequently a synthesis of low-level features for a specific domain.

All features can be classified into low-level features and high-level features. Low-level features are extracted directed from the original images and high-level feature extraction is based on low level features [8].

C. CLASSIFICATION

The first step in classification is feature extraction, in which each instance in the training set is expressed as a vector of measurements. Images are classified here, this vector might be the made from the pixel intensities, but a feature reduction step is used to reduce this to a more effortlessly handled length. The measurements are referred to as features. And these may be integer, real or categorical. The space spanned by all possible combinations of features is called as the feature space.

Because of some reasons not all of the measurements are available, and this is called as missing data. Once features have been extracted there are now two possible cases, which are in general handled quite differently. In the first case, the actual class of each instance in the training set is made available to the classifier, is called supervised classification. In the second case, the information is not made available and this is called unsupervised classification or clustering. Supervised classification is where the actual classes of each of the instances in the training set are known. Sometimes an instance has been given an incorrect class, and this is called a labeling error.

The most frequently used supervised classifiers are binary classifiers. These classifiers differentiate between two types of objects or events. The output of the classifier will be a single output at each point in feature space specifying how strongly the classifier believes that point to have been generated by a specific class. This function is known as a discriminant (for example in Fisher's linear discriminant or quadratic discriminant analysis).

Multiple class classifiers are mostly generated from a number of binary classifiers combined pairwise or else using one against the rest. Types of supervised classifiers are Fisher's linear discriminant, Quadratic discriminant analysis, Nearest neighbours, Parzen window methods, Support Vector Machines, Ensemble methods. Types of unsupervised classifiers is Gaussian mixture models.

Different learning algorithms used in classification process are

1) SVM[17] : It finds decision boundary by maximizing the margin between the different classes. In case of SVM decision boundary can be linear. Also, data can be of any type, i.e. scalar or vector features, intensity images etc. It uses features intensity image[18,19], Haar wavelet [20,21,22], HOG [23, 24], edgelet etc.

2) AdaBoost[25] : It constructs a strong classifier by attaching weak classifiers. And each new classifier focuses on misclassified instances. Speed get optimized if we use AdaBoost in cascades. It can be combined with any classifer to find weak classifier.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

3) Neural Networks[26] : In NN different layers of neurons provide a nonlinear decision. In NN many configurations and parameters are needed to choose. Also, raw data is often get used, so no explicit feature extraction process is needed.

D. VERIFICATION/ REFINEMENT

This module verifies and refines the ROIs classified as pedestrians. The verification step filters false positives, using criteria that do not overlap with the classifier, while the refinement step carry out a fine segmentation of the pedestrian to give an exact distance evaluation or to support the successive tracking module.

E. TRACKING

The most developed systems use a tracking module to track detected pedestrians over time. This step has several reasons like avoiding false detections over time, predicting future pedestrians positions, thus feeding the foreground segmentation algorithm with pre-candidates, and, at a higher level, making useful inferences about pedestrian behavior (e.g., walking direction).

F. APPLICATION

The last step of a pedestrian detection system takes high-level decisions based on the information from previous modules.

III. FUSION ALGORITHM

Based on studying various classification algorithm we can state that fusion of classifiers is generally better than a single classifier. The two mostly used fusion strategies are boosting [27] and bagging [28]. Numerous theoretical studies explain the success of Boosting by proving bounds and margins on the error. The boosting strategy is to yield a series of classifier trained using subsequent training datasets so as to produce new classifiers that are better to predict examples for which the current method's performance is poor. The bagging strategy is to compose an fusion of classifiers where each one is trained with a subset based on a random redistribution of the training dataset. In fact, each individual classifier in the fusion is generated with a different random sampling of the training set. Fusion algorithm adopts fusion of different classifiers, all of them are trained with the same training dataset. The motivation is to compose classifier ensembles with diversity of behavior, in order to increase the accuracy.

This kind of classifier fusion has been applied by many research group on pedestrian detection with, often, better result than single classifiers. Therefore, this work proposes a novel fusion scheme. Actually the fusion algorithm is also a classifier that takes the likelihoods from the others single-classifiers and decides the class. Both the single classifiers and the fusion algorithm are trained with the same training dataset.

However, the single-classifiers are trained before the fusion algorithm in order to create a likelihood training dataset which is used in combination with the training labels in the fusion algorithm training process. In feature vector fusion approaches, some preprocessing is done for each sensor to build a set of features for each one.

IV. DATASET

Datasets are a fundamental tool for comparing detection algorithms, fostering advances in the state of the art. Datasets used for pedestrian detection includes INRIA person database, TUD pedestrian dataset, Caltech pedestrian dataset, MIT LABELME dataset, DC PEDESTRIAN CLASSIFICATION dataset, Penn-Fudan pedestrian dataset, Daimler Pedestrian dataset, CVC-07 DPM Virtual World Pedestrian Dataset etc.

A. INRIA person database : It is very well known in pedestrian detection both for training detectors and

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

reporting results. This dataset from Navneet Dalal and Bill Triggs consists of training and testing data. The training contains 1805 images and X people normalized to 64x128 pixels. The TUD pedestrians training dataset from Micha Andriluka, Stefan Roth and Bernt Schiele consists of 210 and 400 training images with X pedestrians with significant variation in clothing and articulation.

B. **CVC-07 DPM Virtual World Pedestrian Dataset**: It contains 2534 pedestrian images and background images. Fig.3 shows some examples of this dataset. The pedestrian images have frontal view and left view, Which are annotated as 'M' and 'L'. You may flip the pedestrians to get right view examples. Part annotations are also provided[29-31].

C. **The Caltech Pedestrian Dataset**: It 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. Fig. 5 shows some examples of this dataset.

Fig. 2 Some examples of CVC-07 DPM Virtual World Pedestrian Dataset .

Fig 3 CVC-06 Partially Occluded Virtual-World Pedestrian Dataset

Fig. 4 The Caltech Pedestrian Dataset

V. STATE OF ART

A. Pedestrian detection using HOG features and SVM classifiers

Works presented by Zhu et al. (2006) use a rejection cascade with HOG features. Author make use of blocks of various sizes (v- HOG), as well as to train the classifier for each stage, a small subset of blocks is selected randomly. In the work presented by Premebida et al. (2009), pedestrians are detected with a laser scanner using multidimensional features to illustrate the geometrical properties of the detections, and features of Histograms of Oriented Gradients (HOG) through SVM classification. Also, Bayesian modeling method is used for fusion. Perez Grassi et al.(2010) make a use of Invariant Features, with SVM, for pedestrian detection. Ludwig et al. (2011) make a use of HOG features along with SVM classifier for pedestrian detection. Zhao et al. (2014) use multiple feature fusion based on different feature extraction method for vision approaches like Naive Bayesian, Support Vector Machine (SVM), and Minimum Distance Classifier. Decentralized fusion architecture methods carry out detections and classifications used for every sensor separately and a final stage combines the detections as per the detections and sensors trustability.

B. Pedestrian detection using Adaboost classifiers

Spinello et al. (2008) make a use of Adaboost algorithm for pedestrian detection. Gaussian Mixture Model classifier (GMM) used for laser scanner based pedestrian detection, as well as a Bayesian decisor is used to merge detections of both subsystems. Guo et al.(2012) employ Adaboost as well as SVM classifier based system for pedestrian detection in computer vision. Author Hulin Kuang et al. (2013) presents multi cascade approach for pedestrian detection. It consists of Gentle Adaboost (GAB) cascade and the Four Direction Feature(FDF). This is most effective and also efficient feature detection method.

C. Pedestrian detection using Kalman Filter (KF)

In the work of Li et al.(2013) presents background foreground identifications increases the detection and combination with Camshift tracker and a Kalman Filter (KF) which gives trustable pedestrian detection with tracking. In Fan et al. (2013) uses deformable part models and KF for visual based pedestrian detection along with tracking through JPDA association technique. In the work of Schneider et al (2013), comparison between Extended KF with Interacting Multiple Models (IMM) tracking provided is presented for stereovision based pedestrian detection.

D. Pedestrian detection using fuzzy logic

Castro et al. (2011) presents an intelligent system based on fuzzy logic which is specially designed to avoid pedestrian accidents.

E. Pedestrian detection using other classifiers

Broggi et al. (2008) make a use of information from laser scanner to locate regions of the environment where pedestrians could be located and visibility is reduced (for example space between two vehicles) and does detections using a vision based approach approach. Ludwig et al.(2010) and also Premebid et al.(2009) uses method of extracting features for each sensor separately, accordingly a fresh data set is build; authors represent different methods, if combining or not the different features of the different sensors and comparing results. The final classification subsequent to fusion is compared with other methods like Naive Bayes, Gaussian Mixture Models, Neural Networks. Sanchez et al. (2009) and Gomez-Romero et al.(2011) present tracking procedures which is used for surveillance applications by taking benefit of the context information in complex situations. In the work presented by Garcia et al. (2012) uses radar along with computer vision. Featured based optical flows are used for vehicle overtaking detection. In the range of data fusion, target tracking is the main characteristics. Premebida et al. (2013) uses lidar ROI detection to provide high level fusion based on a Bayesian approach. Bohmlander et al. (2013) uses mono camera as well as a capacitive sensor are used for pedestrian detection and expert systems researching fields.

By combining strong classification algorithms and trustable tracking procedures, pedestrian detection can be performed. Conesa et al. (2013) make a use of vehicles driving in opposite direction are detected by use of agent based architecture. Abellan et al (2013) uses decision tree based algorithm to analyze and recognize the severity of the accidents. Jo et al. (2014) identifies driver drowsiness by means of the use of driver specific biological measurements and computer vision based algorithms for eye state and blinking detection.

VI. DISCUSSION

A perfect on board pedestrian detection systems with affordable cost, must detect presence of people in the way of vehicle and react according to risk like warn the driver, brake the vehicle (if it is in autopilot mode), deploy external airbags etc. but not disturbing the driver if there is no risk at all. Also, this system should work excellent independent of environmental conditions, time, road etc. Moreover, cost of this pedestrian detection module should be relatively small compared to cost of vehicle. It is clear, in reviewed literature that, enormous research has been done in this field with different methods of feature extraction and classifiers. Also, fusion of two or more classifiers and feature extraction techniques used for getting better results as there is tradeoff between accuracy and time being needed for computation.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

VII. CONCLUSIONS

Intelligent vehicles represent a key technology for minimizing number of accidents between pedestrians and vehicles. Problems occurring during designing such a system must be overcome i.e.real time detection of changing targets in outdoor scenarios which is uncontrolled. Pedestrian detection is not easy task , consequently a plethora of researcher working for this system. We have summarized work done by earlier researchers and found that there is great need for further work in this domain to overcome reported shortcomings.

REFERENCES

[1] H. Zhang, A. Berg, M. Maire, and J. Malik, “SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition,” Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition, 2006.

[2] P. Viola, M. Jones, “Robust real-time object detection” ,Int.J.Comput.Vis.57(2) (2004)137–154.

[3] Q. Zhu, S. Avidan, M. C. Yeh, K. T. Cheng, “Fast human detection using a cascade of histograms of oriented gradient”s, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,vol.2,pp.1491–1498, 2006.

[4] D.S Huang,” Radial basis probabilistic neural networks: Model and application”, Int. J. Pattern Recognit.Artif.Intell.13(7)(1999)1083–1101.

[5] D.S. Huang, Ji-Xiang Du, “A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks”, IEEE Trans. Neural Netw. 19(12)(2008)2099–2115.

[6] CHONG Yan-Wen, KUANG Hu-Lin, LI Qing-Quan, “Two- stage pedestrian detection based on multiple features and machine learning”, Acta Autom.Sin.38(3) (2012)375–581.

[7] Mineki Soga, Shigeyoshi Hiratsuka. “Pedestrian detection for a near infrared imaging system”, in: Proceedings of the Eleventh International IEEE Conference on Intelligent Transportation Systems,pp.12–15,2008.

[8] R. Rios-Cabrera, T. Tuytelaars, L. Van Gool ,” Efficient multi-camera detection, tracking, and identification using a shared set of Haar-features”, in: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp.65–71,2011.

[9] Shen Jifeng, Sun Changyin, Yang Wankou, Sun Zhongxi, “Fast human detection based on enhanced variable size HOG features”, in: Proceedings of the Eighth International Symposium on Neural Networks, ISNN2011,2011,pp.342–349.

[10] Z.-Q. Zhao, D. S. Huang, B.- Y. Sun, “Human face recognition based on multiple features using neural networks committee”, Pattern Recognit.Lett.25(12) (2004)1351–1358.

[11] Z.-Q. Zhao, D.S. Huang, B.- Y. Sun, “Human facial recognition based on multiple feature domains”, The Fifth World Congress on Intelligent Control and Automation (WCICA’04), Hangzhou, China, pp.4150–4155,2004.

[12] Xiaoyu Wang Tony X. Han Shuicheng Yan. “An HOG- LB human detector with partial occlusion handling”, in :Twelfth International Conference on Computer Vision, pp.32–39, 2009.

[13] P. Viola, M. Jones, and D. Snow, “Detecting Pedestrians Using Patterns of Motion and Appearance,” Int’l J. Computer Vision, vol. 63, no. 2, pp. 153-161, 2005.

[14] B. Leibe, N. Cornelis, K. Cornelis, and L.V. Gool, “Dynamic 3D Scene Analysis from a Moving Vehicle,” Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition, 2007.

[15] Markus Enzweiler, Student Member, IEEE, and Dariu M. Gavrila, “Monocular Pedestrian Detection:Survey and Experiments” IEEE Transaction on pattern analysis and machine intelligence, vol. 31, no. 12, December 2009.

[16] A. Broggi, A. Fascioli, I. Fedriga, A. Tibaldi, and M.D. Rose, “Stereo-Based Preprocessing for Human Shape Localization in Unstructured Environments,” Proc. IEEE Intelligent Vehicles Symp., pp. 410-415, 2003.

[17] V. Vapnik, The Nature of Statistical Learning Theory. Springer, 1995.

[18] Q. Tian, H. Sun, Y. Luo, and D. Hu, "Nighttime Pedestrian Detection with a Normal Camera Using SVM Classifier," Proc. Int'l Symp. Neural Networks, pp. 189-194, 2005.

[19] F. Xu, X. Liu, and K. Fujimura, "Pedestrian Detection and Tracking with Night Vision," IEEE Trans. Intelligent Transportation Systems, vol. 6, no. 1, pp. 63-71, Mar. 2005.

[20] C. Papageorgiou and T. Poggio, "A Trainable System for Object Detection," Int'l J. Computer Vision, vol. 38, no. 1, pp. 15-33, 2000.

[21] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-Based Object Detection in Images by Components," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 4, pp. 349-361, Apr. 2001.

[22] L. Andreone, F. Bellotti, A.D. Gloria, and R. Lauletta, "SVM-Based Pedestrian Recognition on Near-Infrared Images," Proc. Fourth Int'l Symp. Image and Signal Processing and Analysis, pp. 274-278, 2005.

[23] L. Zhang, B. Wu, and R. Nevatia, "Pedestrian Detection in Infrared Images Based on Local Shape Features," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[24] C. Wojek and B. Schiele, "A Performance Evaluation of Single and Multi-Feature People Detection," Proc. DAGM Symp., pp. 82-91, 2008.

[25]Y. Freund and R. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," J. Computer and System Sciences, vol. 55, no. 1, pp. 119-139, 1997.

[26] C. Bishop, Neural Networks for Pattern Recognition. Oxford Univ. Press, 1995.

[27] Freund, Y. Schapire, R., "Experiments with a new boosting algorithm" Proceedings of the Thirteenth International Conference on Machine Learning, 148-156 Bari, 1996.

[28] Breiman, L., "Bagging predictors", Machine Learning, 24(2), 123-140, 1996.

[29] David Geronimo, Antonio M. Lopez, Angel D. Sappa, Member, IEEE, and Thorsten Graf , "Survey of Pedestrian Detection for Advanced Driver Assistance Systems", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 32, NO. 7, JULY 2010.

[30] Oswaldo Ludwig Junior, David Delgado, Valter Gonc¸alves, Urbano Nunes, "Trainable Classifier-Fusion Schemes: an Application to Pedestrian Detection".

[31]] J. Xu, D. Vazquez, A. M. Lopez, J. Marin, D. Ponsa, "Learning a Part-based Pedestrian Detector in Virtual World". In IEEE Transactions on Intelligent Transportation Systems (**T-ITS**), 2014.

[32]J. Xu, S. Ramos, D. Vazquez, & A. M. Lopez, Domain "Adaptation of Deformable Part-Based Models". In IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**), 2014.

[33]D. Vázquez, J. Marín, A.M. López, D. Ponsa and D. Gerónimo, "Virtual and Real World Adaptation for Pedestrian Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**), 2014.

[34] J. Marín, D. Vázquez, A.M. López, J. Amores and L.I. Kuncheva, "Occlusion handling via random subspace classifiers for human detection", IEEE Transactions on Systems, Man, and Cybernetics (Part B) (**ITSMC-B**), 2013.

[35]Q. Zhu, S. Avidan ,M .C .Yeh, K. T. Cheng, "Fast human detection using a cascade of histograms of oriented gradients" ,in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,vol.2,pp.1491–1498,2006.