A Review on Pedestrian Detection in the Vehicle Dashpot and Surveillance Camera Video Footage

B. Thiyaneswaran	K. Anguraj	P. Keerthana	D. B. Ramya
Department of ECE	Department of ECE	Department of ECE	Department of ECE
Sona College of	Sona College of	Sona College of	Sona College of
Technology, Salem	Technology, Salem	Technology, Salem	Technology, Salem

ABSTRACT
Automation plays an important role in the universe. The automation is having more impact in vehicle. A pedestrian detection is required for advance d driver assistance system and security surveillance system. The video is captured using camera which may fit in the vehicle dashpot or CCTV footage fitted in home or industry. The videos are converted into frames. The frames are analyzed and further pedestrian is detected using image processing techniques. In this paper, the review is carried out using the recent research work. Around 20 numbers of papers are taken for the review. The various techniques used in the pedestrian detection such as histogram oriented gradient method, SURF, SIFT, LDA are gathered using the review. The various classifiers required to identify the person or detecting the pedestrian are analyzed. The various hardware implementations used in the recent research is discussed. The recently achieved accuracy and error rate are analyzed using the review.

General Terms
Advanced Driver Assistance System (ADAS), Histogram Oriented Gradient (HOG), Multilayer neural network.

Keywords
Security survilence, Modulating neural network, YOLO, Caltech, KITTY, INRIA

1. INTRODUCTION
A pedestrian detection plays an important role on the Advanced Driver Assistance System (ADAS) used in the vehicles, and Security surveillance system. It is useful for fully automated driverless vehicles, and supports the driver in terms of pedestrian alert in semi automated vehicles. It is also useful for investigation, pedestrian identification, and analysis in security footage.

In this paper, a review is performed on pedestrian detection using the recent research work. The various stages such preprocessing, segmentation, feature extraction techniques used by the various researchers are discussed. An around 20 research papers are taken for the analysis. The review covers the research work in the years 2018, 2017, 2016, and 2015.

2. REVIEW ON THE YEAR 2018
GPU based pedestrian activity detection was proposed by tripathi [15]. A UT and ATM data sets are used in the proposed algorithm. The algorithm comprises of block sampling, descriptor generator, distributed of extracted features, and classification. The feature extraction is carried out using MHI and HOG descriptor. The feature centroid and distance calculation is carried out using GPU.

3. REVIEW ON THE YEAR 2017
A pedestrian count in the crowd is computed by Li Yin [7]. The big data analytics is introduced in this method. The pedestrian data is captured using Google street view images. The images are initially trained by using the LSVM method and pedestrian model is created. The street image is captured and multi scale segmentation is performed. The fine segmentation is performed by comparing with the trained image. The method is tested with 2D images and the increased result may achieve by means of using 3D images. The two draw backs in this method are usage of meta data and un measurable images in the google data base.

A combination of Controlled Convolution Neural Network (CCNN) and Modulating Neural Network (MNN) based method was proposed by song tang [8]. The training is performed to collect positive samples by using you tube and NICTA data base. The negative samples are collected by using INRIA, Stanford back ground, and VOC 2007 data sets. Then further generic detector and random trimming is performed to attain negative samples. The positive and negative samples are applied on CCNN and MNN techniques. The method does not keep the samples permanently. The method achieved the precision in the rate of 91.4%.

A pedestrian detection using YOLO neural network was proposed by molchanov and his team [9]. There are three different method of neural architecture is followed in this method such as YOLO, YOLO-small, YOLO-tiny. The YOLO provides the better quality result but takes more resource. YOLO tiny gives the optimum performance result. Caltech and KITTY data bases are used to test and train the algorithm. The training is performed initially using new convolution layers of Caltech and KITTY data sets. Fine tune is performed in the stage two using the small images.

An efficient pedestrian detection using low resolution sparse representation was proposed by Fang and team [10].The experiment is carried out using Caltech, INRIA, TUD, and ETH data sets. In the LR images initially sparse code is applied. The high resolution and low resolution dictionary is created. The mapping takes place between LR, HR dictionary and sparse code. Finally high resolution image is obtained from low resolution image. It confirms that the processing speed is 11times higher than the existing method.

An Aggregated Channel Features (ACF) with optimum parameters for pedestrian detection was proposed by bastian [11]. There are four steps are followed in ACF such as training data set size, sliding window stride, sliding window size, and number of boot strapping stages. The Caltech sata set is used for training and testing algorithm. The size of the sliding window is 64x32. A number of data sets are 4250. Sliding stride and number of boot strapping stages are 4. The method achieved 0.2 sec detection with the lesser filter value

of 10. It is comparatively best performance as compared with the checker board method.

A pedestrian detection in the abnormal crowd was proposed by Aguilar [12]. The motion detection analysis process is initially carried out. It processed by comparing current frame with the previous frame. A contour is formed between the frame movements. A background is subtracted using template collection. The algorithm was tested under UMN, UCSD ped-1, and avenue data sets. The average frame processing rate is achieved as 72.6 sec under I3, 4GB RAM configuration machine.

A scale aware Fast R-CNN method for pedestrian detection was proposed by Jianan li [13]. The algorithm was tested using Caltech, INRIA, ETH, and KITTI data sets. The pedestrian is classified as smaller and larger size in the initial stage. The ROI pooling is carried out in large and small scale sub network. The weight is calculated in small and large scale. The difference of weight is calculated between small scale and large scale. The average precision obtained in KITTI data base is 77.93%. The author concludes that this method is suitable for detection of smaller height pedestrian.

A pedestrian detection in the agricultural form was proposed by pezzimenti and his team [14]. The camera is fixed in the agricultural tractor vehicle. An apple and orange growing path pedestrian was detected in this algorithm. A data base is created with the name of national robotics engineering agriculture center. The created data base is compared with the KITTI, Caltech and Daimler data bases. The various features such as colours, textures, poses, occlusion are extracted from the pedestrians.

4. REVIEW ON THE YEAR 2016

The shadow is the obstacle which reduces the detection rate pedestrian. The shadow removal process was carried out by Lingxiang Zheng and his team [2]. The image back ground and fore ground is separately computed. The pedestrian region is extracted from the back ground region. The difference in value is calculated and applied on the histogram manipulations. The shadow in the image is removed and compared with the data set. The method achieved the lower error rate as 0.6503.

A simple non-linear motion filter is used by yi wang and his team [3]. This method drastically improves the accuracy of pedestrian detection. The principle of this method initially extracts the background of the image using past frames. The background means image without pedestrians. It is the effective non-linear filter element. The back ground is suppressed from the moving elements (pedestrian). The location of pedestrian is identified. The identified mask is applied on the original which consists of both back ground and moving pedestrian. The result is analyzed under various stages of pedestrians such as base line, small pedestrian, stand still, different position, thermal, and crowded scene. The method achieves the reduced false prediction without increasing miss rate.

An optimization based techniques is proposed by tome and his team in pedestrian detection[4]. The Caltech pedestrian data set is used in this algorithm for training and testing. The pedestrian is initially identified. The particular frame is analyzed in horizontal and vertical direction. A coarse filter is applied and feature is extracted at various levels. It also includes locally de-correlated features. The each part of feature is analyzed and histogram padding is adopted. Individual region score is computed. The deep convolution is

carried out by using neural network. The method shows the considerable improvement in accuracy.

The forgery detection in front of camera based method was proposed by Fahmy[5]. The feature is extracted using new Zernike method. It is the clustering method. A silhouette coefficient is extracted. The difference between forgery and Zernike moment image is computed. Further Euclidian distance is measured among the dissimilarity. The method achieves the true prediction rate of 98.7%, and false prediction ratio of 0.2%.

A detection of image in the smart environment increases the hit rate. The smart environment based on infra red and visible fusion was proposed by Castillo and his team [6]. The people are sample in two types of environment such in day visible light and night mode infra red. The frames are converted at the rate of 5/second. The fusion based segmentation is applied. The image is captured at various environmental conditions such as -2° to 38°, low, medium, ambient light, foggy, snow covered, sunny, and cloudy. The result shows the precision of 0.98, and F-Score metric of 1%.

5. REVIEW ON THE YEAR 2015

Shihong yao and his team uses INRIA and Daimler data base for testing and training the pedestrian samples [1]. There are two types of features were extracted from image such as Histogram Oriented Gradient (HOG) and Local Self Similarity (LSS) descriptor. The gradient feature in x-direction, y-direction, and complete image gradient is extracted in HOG based descriptor.

The image is further is divided in smaller blocks at various scales. The HOG feature is extracted at each level. The feature is classified using SVM and Adaboost classifiers. The experimental result is compared with the various feature extraction techniques such as SIFT, SURF, HOG, and LDA. The author confirms that the HOG-LSS provides the better result in pedestrian detection.

6. REVIEW ON THE YEAR 2014

Traffic Sign Recognition (TSR) is an important feature of ADAS and this work was proposed by Fatin and his team [16]. The traffic signs are used to alert the driver speed limits and possible dangers. The shape based pattern detection is used in this method. It includes the segmentation, shape detection, and classification.

Victoria and his team proposed pedestrian detection sub system [17]. The video are captured and converted in to frames. The HOG feature is extracted from the image and pedestrians are detected. The pedestrian tracking collision detection are performed. It is one of subsystem used in the semi automated vehicles.

7. REVIEW ON THE YEAR 2013

Cheng [18] and his team proposed a pedestrian detection in visual surveillance video footage. It composes cascaded two multiplexer using Haar and shaplet transforms. Haar features use contrast value of the adjacent pixels. Haar provides faster detection The classification is carried out using Adaboost classifier. The method also improve the detection time.

A survey on intelligent video analysis was performed by Honghai [19] and his team. The key discussion on this paper is described below. Recent technologies demanded the significant need for the video surveillance. The manual visualization of human eye gets tired with different screens. Alert conditions may be set with real time algorithms to

deliver information. In order to reduce the computational complexity in the computer vision the low-level image processing of sensor acquisition is performed.

Video system in poor environment provides poor video quality for diagnosis and system management. It helps in target overlapping, target detection and tracking of target. It is used in 3-D image reconstruction from 2-D image in solving the problem of computer vision fields. Smart cameras works on embedded sensors form real time computer vision which is useful in the person tracking.

The main motive for developing the security of the military applications. It helps in people tracking, vulnerable driver assistance and focuses on low cost and high accuracy for practical implementation of security system. It can be used for data transfer, authentication, encryption, coding, and data packing for security in the networks or either centralized or distributed.

Hunjae and his team proposes algorithm for lane detection [20]. Lane detection in important in many advanced driver assistance systems (ADAS). Vision based lane algorithm mainly used gradient information as a lane features. In this method it produces a gray level image from RGB color image based on linear discriminate analysis. Many accidents have been caused due to carelessness of the driver and increased vehicles in the surrounding.

Lane detection can be considered as the basic module of the ADAS and can be used to detect the approaching vehicle and obstacles. In certain illumination condition the difference is not large so they are converted into gray level values and its difficult to extract proper gradient features. The Histogram gradients can be used to enhance the gradients for contrast adjustment. The information lost by the gray level image cannot be recovered. The Hough and edge detection are used for initial lane detection.

Performance of vision based algorithm is needed to be evaluated by ADAS according to illumination conditions for constructing video database. The result of the histogram equalization contains the unwanted edge detection and it does not work well on multi-illumination conditions of water reflections like heavy rain at night. The gradient enhancing produces a color image of maximized intensity. In addition to that it does not provide high computational complexity. It takes 50ms per frame and can be used in real time applications.

Alberto Broggi and his team presents vision of the artificial vision intelligent system laboratory(VisLab) for future automation and sensor selection for extensive testing [21]. He used the BRAiVE protocol which is mainly based on vision. This implementation is based on the extensive and invasion sensor and the importance of extensive tests to validate the design. Driverless vehicle will reshape the future mobility of people and goods.

BRAiVE which is short for brain drive and it uses Advanced driver assistance system (ADAS). It is further classified into the following three categories 1) Development , 2)Test and 3)Demonstration. It is used for sensing and vehicle detection from one of the four front sensing cameras and all the controllers connected to the processor for proper power supply to all the devices and the perception of obstacles detected from both the vision and laser.

Each device can be switch on and off by customer built when a device is absorbing current the corresponding light can be turned on and the main switch is located on the keypad of the complete switching of the system. All the systems are integrated to provide the normal look of the original device as it is focusing on the vision technologies to implement the perception layer to deliver a robust set of data about the surrounding.

In this paper proposed by wen-chang-cheng a cascade classifier combining Adaboost and support vector machine applied to pedestrian detection involved using window of fixed size to extract the area of interest [18]. So that weak classifiers are able to focus on higher degradation rate other than accuracy. The driver assistance system can avoid collision through pedestrian and route detection that guarantee the safety of the drivers.

Adaboost has higher detection rate on the other hand sliding window combined with various features in SVM provides better features when compared with Adaboost system. The feature extraction of HOG with grayscale pixel values based on the magnitude and direction of gradient in each cell of 3*3 the gradient is divided into 45 degree as a unit. This algorithm is able to self-adaptively increase the overall accuracy.

The sample videos are collected from the campus and the SVM provides the accuracy of 99.91% and cascade Adaboost provides the accuracy of 99.60% the results differ because of pedestrian and non-pedestrian samples in the videos. The size of the sample was normalized to 15*36 pixels were more complex arises and increases the difficulty of pedestrian detection. It improves the time consumption of SVM while applied to large-scale.

8. CONCLUSION

The review work cover the various researches from the year of 2013, 201, 2015, 2016, 2017, and 2018 in pedestrian detection. The researcher used the various data sets such as Caltech, INRIA, ETH, KITTI, ATM, UT, Stanford back ground, and VOC 2007. The various feature extraction techniques such as HOG, Haar, SIFT, SURF, and LDA performance and result achievements are discussed. The frame rate selection and frame size are discussed. The various feature classifiers such as SVM, ADA-Boost techniques are discussed. The precision of the recent paper reach up to 0.98. The closest accuracy achieved so far is 99.6%. The pedestrian detection is useful in the ADAS, agriculture farms, security surveillance system in semi automated and fully automated systems.

9. REFERENCES

[1] Wen-Chang Cheng, Ding-MaoJhan. 2015 A self-constructing cascade classifier with AdaBoost and SVM for pedestrian detection. Engineering Applications of Artificial Intelligence, Elseiver, 1016-1028.

[2] Lingxiang Zheng, Xiaoyang Ruan,Yunbiao Chen, Minzheng HuangFröhlich, 2016 Shadow removal for pedestrian etection and tracking in indoor environmentsConference on Human Factors in Computing Systems, Springer

[3] Yi Wang, Sebastien Piérard, Song-Zhi Su,Pierre-Marc Jodoin. 2016 Improving pedestrian detection using motion-guided filtering, Pattern Recognition Letters, Elsevier,1-7.

[4] D. Tome, F. Monti, L. Baroffio, L. Bondi, M. Tagliasacchi, S. Tubaro. 2016 Deep convolutional neural networks for pedestrian detection, Signal Processing : image communication, Elsevier.

[5] O. M. Fahmy, 2016 A new Zernike moments based technique for camera identification and forgery detection, Springer

[6] Jose´ Carlos Castillo, Antonio Fernandez-Caballero, Juan Serrano-Cuerda, Marıa T. Lopez, Arturo Martinez Rodrigo, 2016 Smart environment architecture for robust people detection by infrared and visible video fusion, J Ambient Intell Human Comput, Springer.

[7] Li Yin, Qimin Cheng, Zhenfeng Shao, Zhenxin Wang, and Laiyun Wu. 2016 'Big Data': Pedestrian Volume Using Google Street View Images, Springer, 461- 469.

[8] Song Tang,Mao Ye, Ce Zhu, and Yiguang Liu. 2017 Adaptive pedestrian detection using convolution neural lnetwork with dynamically adjusted classifier, Journal of Electronic Imaging, SPIE, 26(1), 013012-1 – 013012-9.

[9] V. V. Molchanov, B. V. Vishnyakov, Y. V. Vizilter, O. V. Vishnyakova, and V. A. Knyaz. 2017 Pedestrian detection in video surveillance using fully convolutional YOLO neural network, Proc. of SPIE, Vol. 10334

[10] Wenhua Fang, Jun Chen, and Ruimin Hu. 2017 Efficient Pedestrian Detection in the Low Resolution via Sparse Representation with Sparse Support Regression, LNAI 10235, pp. 313–323.

[11] Blossom Treesa Bastian, C. Victor Jiji.2017 Aggregated Channel Features with Optimum Parameters for Pedestrian Detection, LNCS 10597, Springer, pp. 155–161

[12] Wilbert G. Aguilar, Marco A. Luna, Hugo Ruiz, Julio F. Moya, Marco P. Luna, Vanessa Abad, and Humberto Parra.2017 Statistical Abnormal Crowd Behavior Detection and Simulation for Real-Time Applications, LNAI 10463,Springer, pp. 671–682.

[13] Jianan Li, Xiaodan Liang, Shengmei Shen, Tingfa Xu, Jiashi Feng, and Shuicheng Yan. 2017 Scale-aware Fast R-CNN for Pedestrian Detection IEEE Transactions on Multimedia.

[14] Zachary Pezzementi, Trenton Tabor, Peiyun Hu, Jonathan K. Chang. 2017 Comparing apples and oranges:Off-road pedestrian detection on the National Robotics Engineering Center agricultural person-detection dataset, Wiley

[15] Vikas Tripathi, Durgaprasad Gangodkar, Samin Badoni and Sagar Singh Bisht. 2018 GPU Based Bag of Feature for Fast Activity Detection in Video, 2018 Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing, springer, 133-141.

[16] Fatin Zaklouta , Bogdan Stanciulescu. 2014 Real-time traffic sign recognition in three stages, Robotics and Autonomous Systems, Elsevier, 16-24.

[17] Victoria A. Banks, Neville A. Stanton, Catherine Harvey. 2014 Sub-systems on the road to vehicle automation: Hands and feet free but not 'mind' free driving, Safety Science, Elsevier, 505-514.

[18] Wen-Chang Cheng, Ding-MaoJhan. 2013 A self-constructing cascade classifier with AdaBoost and SVM for pedestrian detection, Engineering Applications of Artificial Intelligence, Elsevier, 1016-1028.

[19] Honghai Liu, Shengyong Chen, Naoyuki Kubota. 2013 Intelligent Video Systems and Analytics: A Survey, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 9, NO. 3, 1222-1233

[20] Hunjae Yoo, Ukil Yang, and Kwanghoon Sohn. 2013 Gradient-Enhancing Conversion for Illumination-Robust Lane Detection, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 14, NO. 3, 1083-1094

[21] Alberto Broggi, Michele Buzzoni, Stefano Debattisti, Paolo Grisleri. 2013 Extensive Tests of Autonomous Driving Technologies, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 14, NO. 3, 1403-1415.

[22] B Thiyaneswaran, R Kandiban, KS Jayakumar. 2012 Elimination of iris hazards intended for localization using visible features of iris region, Procedia engineering, Elsevier, 246-252, https://doi.org/10.1016/j.proeng.2012.06.032

[23] B Thiyaneswaran, A Saravanakumar, R Kandiban. 2016 Extraction of mole from eye sclera using object area detection algorithm, Wireless Communications, Signal Processing and Networking (Wispnet'16), IEEE Xplore, 1413-1417, DOI: 10.1109/WiSPNET.2016.7566369

[24] D. Sandhiya, B Thiyaneswaran. 2017 Extraction of dorsal palm basilic and cephalic hand vein features for human authentication system, International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET'2017), 2231-2235, IEEE Xplore, DOI:10.1109/WiSPNET.2017.8300156.