Design of a Real-time Pedestrian Detection System for Autonomous Vehicles

Harshitha R† and Manikandan J‡
†Department of Electronics and Instrumentation, Bangalore Institute of Technology,
V.V.Puram, Bangalore 560004, Karnataka, INDIA.
‡Crucible of Research and Innovation (CORI), PES University,
100-Feet Ring Road, BSK Stage III, Bangalore 560085, Karnataka, INDIA
{ email : harshithaharsh0104@gmail.com, manikandanj@pes.edu }

***Abstract*— Pedestrian detection is considered as an active area of research and the advent of autonomous vehicles for a smarter mobility has spearheaded the research in this field. In this paper, design of a real-time pedestrian detection system for autonomous vehicles is proposed and its performance is evaluated using images from standard datasets as well as real-time video input. The proposed system is designed using Histograms of Oriented Gradients feature extraction and Support Vector Machine classifier. 98.31% accuracy is obtained on using the proposed system, whereas 100% accuracy is obtained on non-detected pedestrians nearing the vehicle.**

Keywords*— Autonomous Vehicles; HOG Features; Pedestrian Detection.

I. Introduction

Real-time human detection from videos is one of the most active areas in computer vision due to its widespread applications such as intelligent surveillance and home security in [1], personal protection and kidnapping detection in [2], automatic detection of crimes in [3] and human computer interfaces in [4]. The successful progress towards design of autonomous vehicles such as autonomous cars [5,6], self-driven rider less bicycles [7], autonomous robots [8] and drones [9,10] have spearheaded research in the area of pedestrian detection. Pedestrian detection using HOG and neural networks is reported in [11], pedestrian detection for advance driver assistance system using HOG and Adaboost is reported in [12], pedestrian detection using Bayesian and Edgelet detectors in [13], using local binary patterns (LBP) in [14], using Motion and Appearance patterns in [15], using Shapelet features in [16] and using Deep Networks in [17].

In this paper, a real-time pedestrian detection system is designed and its performance is evaluated using offline images from standard datasets as well as real-time video input. The functionality of proposed system is illustrated in Fig. 1. A camera is used to capture the front scene and each frame is processed to extract features using Histograms of Oriented Gradients (HOG) followed by Support Vector Machine (SVM) classifier to differentiate between pedestrians and background. The pedestrians are identified and their locations are marked. The proposed system is helpful for smarter mobility and would serve the purpose of a driver assistance system too, wherein the identification of pedestrians is displayed on the dashboard for assisting the driver. Section II explains the design of proposed system followed by experimental results, conclusion and references.

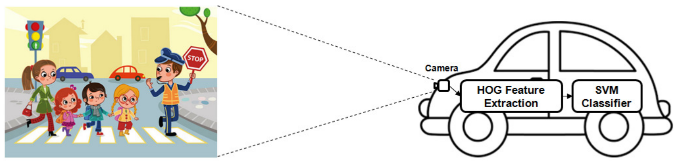


Fig. 1. Proposed Real-time Pedestrian Detection System

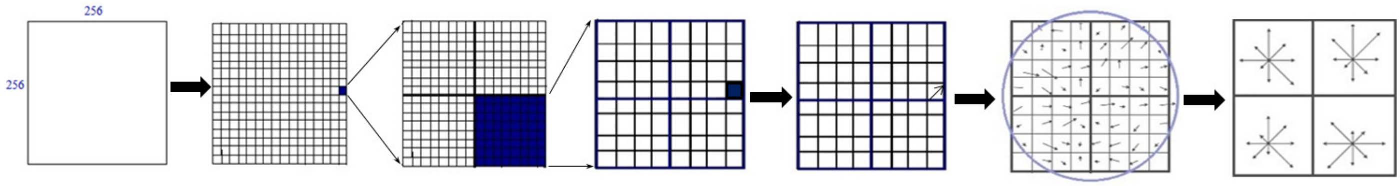
II. Design of Proposed Pedestrian Detection System

In this section, the design of proposed real-time pedestrian detection system is explained with feature extraction and classifier being the two main stages of the system.

A. HOG Feature Extraction

The input frames captured from camera are first processed to obtain Histograms of Oriented Gradients (HOG) features as shown in Fig. 2. It may be observed from Fig. 2(b) that the HOG features do give an idea about the existence of pedestrian in the image as traced for illustration in Fig. 2(b). The sequence of steps followed to obtain these HOG features from an input image is shown in Fig. 3, considering an input image of size 256 x 256 pixels. The input image is divided into 256 cells as shown in Fig. 3(b) with a cell size of 16 x 16 pixels and each cell is divided into four sub-cells as shown in Fig. 3(c) with a sub-cell size of 8 x 8 pixels.

978-1-5090-6255-3/17/$31.00 ©2017 IEEE



(a) Input Image (b) Split into Cells (c) Sub-cells (d) Pixel Location (e) Gradient Vector (f) Gradient Vector of Subcell (g) Key Descriptor

Fig. 3. Procedure to obtain HOG Features of Input Image

(a) Input Image (b) HOG Features

Fig. 2. HOG Features of Input Image

The gradients G_x and G_y of a particular pixel location is computed using 1D masks in X and Y direction as

$$G_x = M_x \cdot I_x \quad (1)$$

and $$G_y = M_y \cdot I_y \quad (2)$$

where $M_x = [-1\ 0\ 1]$ and $M_y = [-1\ 0\ 1]^T$ are the masks used on I_x in X-direction and I_y in Y-direction respectively as shown in Fig. 4. The gradient magnitude, G_{mag} and orientation angle, G_{dir} are computed as

$$G_{mag} = \sqrt{G_x^2 + G_y^2} \quad (3)$$

$$G_{dir} = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (4)$$

The gradient vector for each pixel location in the sub-cells are drawn using G_{mag} and G_{dir} values as shown in Fig. 3(f). The gradient vectors in each sub-cell are normalized to obtain a key descriptor as shown in Fig. 3(g). Similarly, key descriptors are computed from each cell to obtain HOG features of input image as shown in Fig. 2(b).

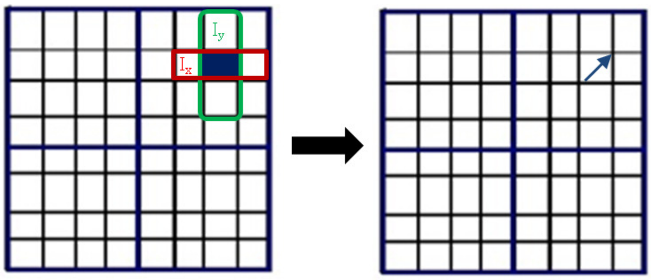


Fig. 4. Computation of Gradient Vector at a Pixel Location using Masks

B. *SVM Classifier*

SVM is considered to be the simplest and fastest classifier for both linear as well as non-linear classification problems [18]. SVM learning aims at finding a good hyperplane in a higher dimensional feature space, that best separates two classes as shown in Fig. 5. The equation of SVM hyperplane that classifies two classes is given as

$$\begin{aligned} Y(z) &= \sum_{sv=1}^{NSV} \alpha_{sv} y_{sv} \left(\mathbf{\Phi}^{\mathrm{T}}(\boldsymbol{x}_{sv})\ \mathbf{\Phi}(z)\right) + b \\ &= \sum_{sv=1}^{NSV} \alpha_{sv} y_{sv}\ \mathrm{K}(\boldsymbol{x}_{sv}, z) + b, \end{aligned} \quad (5)$$

where, z is the test data, $\mathrm{K}(\boldsymbol{x}_{sv}, z)$ is the kernel function, NSV denotes the number of support vectors, $\boldsymbol{x}_{sv}$ is the input, y_{sv} is the target output and α_{sv} is the Lagrange multiplier associated with each training data.

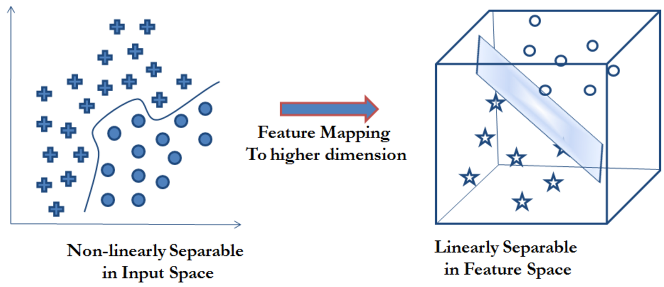


Fig. 5. Linear Hyperplane of SVM Classifier

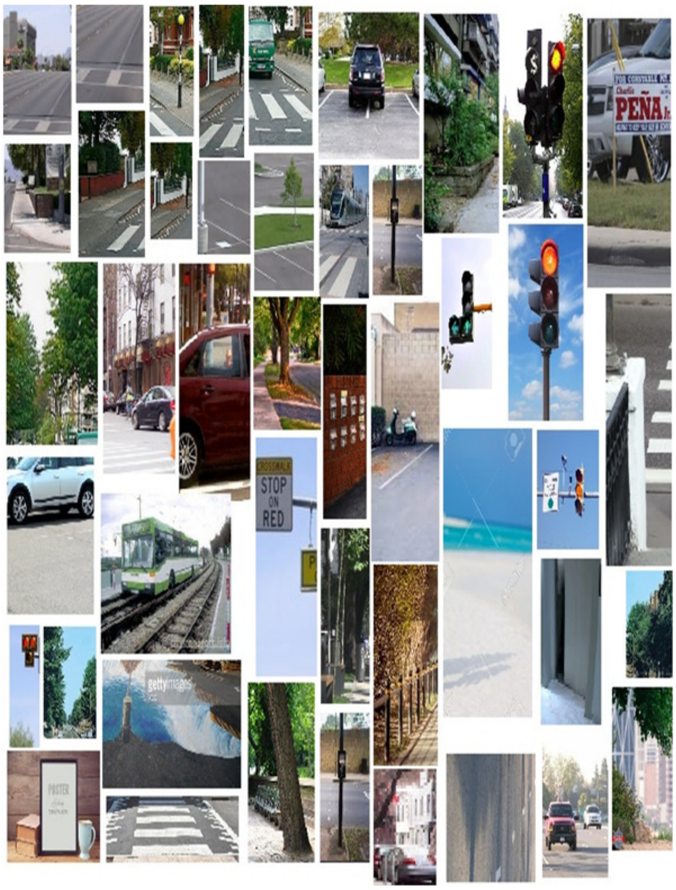
It is well-proved that Gaussian or Radial Basis Function (RBF) kernel deals efficiently and effectively with non-linear relationship between input vectors and target vectors over other kernel functions [19], and hence RBF kernel is used for the proposed system too. The kernel function for RBF kernel is defined as

$$\mathrm{K}(x_i\ .\ \mathrm{z}) = \exp(-\ \gamma\ \|x_i - \mathrm{z}\|^2) \quad (6)$$

where γ is the width vector, a parameter of RBF kernel that needs to be tuned for best accuracy. The HOG features extracted from each window are fed to Support Vector Machine classifier for pedestrian detection.

III. Experimental Results

The performance of proposed system is evaluated by carrying out various experiments using images collected online, using images from INRIA Person Dataset [20] and using real-time inputs. The photo collage of samples considered for training from two classes, background and



(a) Images for Training Background (Class 1)

(b) Images for Training Pedestrians (Class 2)

Fig. 6. Photo Collage of few images considered for training proposed system

pedestrians, is shown in Fig. 6(a) and (b) respectively for images obtained online. The performance of proposed pedestrian detection system using images obtained online and real-time input images are shown in Fig. 7 and Fig. 8 respectively, which is again a photo collage of few results obtained, wherein a blue box denotes the identification of a pedestrian in an image.

The performance result of proposed system is reported in Table I for all the experiments carried out using different datasets. It is observed that a maximum accuracy of 98.31% is obtained on using the proposed system with testing samples of

Fig. 7. Results obtained from the proposed Pedestrian Detection system

Fig. 8. Results obtained during Real-time testing of proposed system

TABLE I. PERFORMANCE EVALUATION OF PROPOSED SYSTEM

Dataset	# Images	# Pedes.	# Pedes. Detected	Accuracy
Images obtained Online	50	130	125	96.13%
INRIA Dataset (Testing Datasets)	1126	1126	1107	98.31%
INRIA Dataset (Full Image)	287	439	423	96.35%
Real-Time Input Images	35	59	55	93.22%

INRIA datasets, which is a mixture of cropped images of pedestrians and backgrounds. Pedestrian detection accuracy of 96.35% is obtained on testing the system using full scenery images provided in the INRIA database. It is observed during experimentation that the proposed real-time pedestrian detection system gave 93.22% accuracy based on the distance between pedestrian and camera, whereas 100% accuracy is obtained, the moment pedestrian comes closer towards the camera. This suffices the requirement of pedestrian detection for autonomous vehicles.

IV. CONCLUSION

In this paper, design of a real-time pedestrian detection system for autonomous vehicles is proposed and its performance is evaluated by carrying out various experiments using offline images, images from standard dataset and real-time input. The system is capable of detecting pedestrians with an accuracy of 98.31% and it is observed that the non-detected pedestrians are also detected, once they come closer to the camera, thus achieving 100% recognition accuracy.

REFERENCES

[1] Yi G, Myoungjin K, Yun C and H Lee, “Design and Implementation of UPnP-based Surveillance Camera system for Home Security”, in Proc. Int. Conf. on Information Science and Applications, Suwon, South Korea, June 2013, pp. 1-4.

[2] Miyahara A and Nagayama I, “An Intelligent security camera system for kidnapping detection”, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 17, pp. 746-752.

[3] Goya K, Zhang X, Kitayama K and Nagayama I, “A Method for Automatic Detection of Crimes for Public Security by using Motion Analysis”, in Proc. Fifth Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, Kyoto, Sep 2009, pp.736-741.

[4] R. Moldovan, B. Orza, A. Vlaicu and C. Porumb, “Advanced human-computer interaction in external resource annotation,” in Proc. Int. Conf. on *Automation, Quality and Testing, Robotics,* Cluj-Napoca, 2014, pp. 1-6.

[5] K. Bimbraw, “Autonomous cars: Past, present and future a review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology,” in Proc. 12th Int. Conf. on *Informatics in Control, Automation and Robotics*, Colmar, 2015, pp. 191-198.

[6] C. Pozna and C. Antonya, “Issues about autonomous cars,” in Proc. *11th Int. Sym. on Applied Computational Intelligence and Informatics*, Timisoara, 2016, pp. 13-18.

[7] JA Brizuela-Mendoza, CM Astorga-Zaragoza, A Zavala-Río, F Canales-Abarca and J Reyes-Reyes, “Fault Tolerant Control for Polynomial Linear Parameter Varying (LPV) Systems applied to the stabilization of a riderless bicycle,” in Proc. *Control Conference (ECC), 2013 European*, Zurich, 2013, pp. 2957-2962.

[8] S. A. Miratabzadeh *et al*., "Cloud robotics: A software architecture: For heterogeneous large-scale autonomous robots," *2016 World Automation Congress (WAC)*, Rio Grande, PR, USA, 2016, pp. 1-6.

[9] S. Perkins, "From insects to drones," in *Engineering & Technology*, Vol. 10, No. 11, pp. 72-76, December 2015.

[10] K. Moskvitch, "Take off: are drones the future of farming?," in *Engineering & Technology*, Vol. 10, No. 7-8, pp. 62-66, August-September 2015.

[11] Gangqiang Shi and Weiwei Hu, "The video people detection based on neural network," *2015 IEEE 16th International Conference on Communication Technology (ICCT)*, Hangzhou, 2015, pp. 300-303.

[12] J. Rettkowski, A. Boutros and D. Gohringer, "Real-time pedestrian detection on a xilinx zynq using the HOG algorithm," *2015 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, Mexico City, 2015, pp. 1-8

[13] Wu B and Nevatia Ram, “Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors,” International Journal of Computer Vision, Nov 2007, Vol. 75, Issue 2, pp. 247-266.

[14] X Wang, TX Han and S Yan, "An HOG-LBP human detector with partial occlusion handling," *2009 IEEE 12th International Conference on Computer Vision*, Kyoto, 2009, pp. 32-39.

[15] P Viola, MJ Jones and D Snow, "Detecting pedestrians using patterns of motion and appearance," *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, Nice, France, 2003, pp. 734-741 vol.2.

[16] P Sabzmeydani and G Mori, "Detecting Pedestrians by Learning Shapelet Features," *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, MN, 2007, pp. 1-8.

[17] Anelia A, Alex K, Vincent V, Abhijit O and Dave F, “Real-time Pedestrian Detection with Deep Network Cascades,” in Proc. 26th British Machine Vision Conference (BMVC), Swansea UK, Sep 2015.

[18] Christopher Burges, “A tutorial on Support Vector Machines for Pattern Recognition,” Data Mining and Knowledge Discovery 2, 1998, pp.121-167.

[19] J. Manikandan and B. Venkataramani, “Study and evaluation of a multi-class SVM classifier using diminishing learning technique,” Neurocomputing, Volume 73, Issues 10–12, June 2010, pp.1676-1685.

[20] INRIA Person Dataset. http://pascal.inrialpes.fr/data/human/.