



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 6, June 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

An Ingrained System for Pedestrian Detection in Scotopic Vision

N. Musrat Sultana¹, Janipella Kiran²

Assistant Professor, Department of CSE, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad, India¹

B.Tech Student, Department of CSE, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad, India²

ABSTRACT: Convenient vision-based solutions for the driver broaden the means of human vision and backing safe travel. In the model, a scotopic-vision system for pedestrian detection in severe lighting conditions can be build using state of the art algorithms and widely available hardware. The proposed scotopic-vision system for pedestrian detection processes thermal images using a proprietary ODROID XU4 microcomputer under Ubuntu MATE operating system. A cascade object detector is applied for the task of human contour detection in context of thermal imagery and contrasted the results with the state-of-the-art deep learning approach. The experiments conducted prove effectiveness of the proposed solution. The aim is to detect pedestrians on the road during night times. The proposed scotopic - vision system for pedestrian detection processes thermal images and outputs the detected pedestrians in the thermal images. The model helps to create a scotopic - vision system supporting the driver during a night driving. Here, two models are using for pedestrian detection. They are Yolov2 and Haar Hog descriptor with Adaboost algorithm.

KEYWORDS: Yolov2, Adaboost, Haar-Hog, Support Vector Machine.

I. INTRODUCTION

Now a days the road traffic, one of the main problems is road capacity. There are four major causes of that phenomenon: poor planning of transport routes, existence of bottlenecks, lack of existing infrastructure adaptation to the current traffic load, accidents. The last cause is the most dependent on a human being. In an analysis of more than 2000 cases, humans have been identified as responsible for about 92% of accidents. In a study conducted by the Virginia Tech Institute of Transport in 2013, it was shown that driver fatigue was the cause of 20% of accidents. The fatigue causes loss of concentration and prevents the driver from making quick and correct decisions.

To prevent accidents new solutions in the vehicle construction are introduced. However, it is the driver who becomes the weakest link in the security chain on the road despite the increasing development devoted to car safety. To help the deconcentrated driver many manufacturers equip their vehicles with the on-board driver assistant systems known as the ADAS acronym. It stands for Advanced Driver Assistance Systems. It must be observed that some authors use the DAS acronym [5].

The main problem during the night driving is limited visibility of the road (especially in bad weather conditions), dazzling the driver by the lights of cars coming from the opposite direction, lack of the reflective elements on the clothes of cyclists and pedestrians who are frequently dressed in clothing made of dark materials that absorb light. The limited possibility of using traffic lights in case of heavy traffic is an additional difficulty [10]. Another limitation is the poor condition of the headlamps and bulbs, especially in older vehicles. The reflectors in the headlamps are burned out and the reflector glass is often matted over time, which significantly shortens the range of road illumination. In modern vehicles, the driver is assisted by adaptive lights, which illuminate the road much better and at the same time do not dazzle drivers coming from the opposite direction. However, these solutions are expensive and inaccessible for everyone. It is worth to consider what happens when the weather is not perfect, it is raining, and snowing or the fog is present. If the driver in such conditions continues driving the investigated. They are applied to human silhouette detection in thermal spectrum images. Finally, the proposed system is evaluated and the results are analysed in Section V.

II. RELATED WORK

“Pedestrian tracking for infrared image sequence based on trajectory manifold of spatio-temporal slice” is developed by S. P. T. Yang, D. Fu in 2017 [1]. He proposed his works on Trajectory manifolds both in horizontal and vertical directions are obtained to compute the height and width information of the target and to capture the position location. Methodology is used Spatio-temporal slice trajectory manifolds analysis and accuracy is 97.38. The results

achieved are “The proposed tracking method based on spatio-temporal slices can provide relatively accurate results and robust pedestrian bounding boxes with a low computational cost.”

“YOLO9000: Better, Faster, Stronger” is proposed by Joseph Redmon and Ali Farhadi in 2016 [2]. He works on increasing fast, accuracy and better object detection. Methodology used is Convolution Neural Networks (CNN) and accuracy is 78.6. The results achieved are YOLOv2 is state-of-the-art and faster than other detection systems across a variety of detection datasets.

“Pedestrian detection based on faster R-CNN in night time by fusing deep convolutional features of successive images” is proposed by J. H. Kim, G. Batchuluun, and K. R. Park in 2018. They work on Improved the performance of our pedestrian detection by using CNN to generate each deep convolutional feature from successive video frames acquired during night time and then by fusing the deep convolutional features at the feature level. Methodology used is R-CNN and Accuracy is 75.3 [3]. Their results are When the presence or absence (according to Y) of a pedestrian in successive images is given, the feature z2 fused in successive frames (weighted summation) and the feature z1 of a single frame had the same expectation, and the variance of z2 was lower.

“Deep learning approach to detection of preceding vehicle in advanced driver assistance” is proposed by Powel Forczmanski and Adam Nowosielski in 2016 [4]. Methodology used is Convolution Neural Networks (CNN) and Accuracy is 83.5. They work on The successful detection of preceding cars is a part of the analysis of current road situation including emergency and sudden braking, unintentional lane change, traffic jam, accident, etc. The results are the vehicle detection based on the proposed assumptions is possible and quite successful.

“Pedestrian detection with convolutional neural networks” is proposed by M. Szarvas, A. Yoshizawa, M. Yamamoto [5]. Their methodology used is Convolution Neural Networks (CNN) and Accuracy is 90. The results are false positive (FP) rate of the CNN was less than U5-a of the FP-rate of the standard Haar wavelet + SVM combination. The convolutional neural network provides both higher accuracy and needs significantly less computation.

III. METHODOLOGY

The scotopic - vision system for pedestrian detection in severe lighting conditions can be build using state of the art algorithms and widely available hardware. The proposed scotopic - vision system for pedestrian detection processes thermal images using a proprietary ODROID XU4 microcomputer under Ubuntu MATE operating system. A cascade object is applied for detector for the task of human silhouette detection in context of thermal imagery and contrasted the results with the state-of-the-art deep learning approach.

3.1 Image Features

For any object detection and classification techniques, image features carries the object information in the image required for its detection. This particular image features should not carry any redundant or irrelevant information from the extraction. Here we are using a large class of data which needs a rapid extraction of features and its computation should also be easy enough. Besides, it should also robust to the background, illumination changes etc. For these reasons instead of going for a direct raw intensity images or gradients, one can go for a local feature descriptor of an image. These local features can be depends on gradients, points, intensities, blobs, colour and its combinations. Depends on the application, the features of the images were determined for detection and classification.

There are different types of approaches for representing image features, which in general divided into four categories. In that one is of distinct approach which is based on points and the remaining three is of dense approach based on image intensities, wavelets and gradients.

3.2 Pedestrian Sensor

Many solutions have focused on increasing human vision capabilities. Until now, no perfect system has been obtained and research on improving are still ongoing. Most of the elaborated scientific solutions, apart from products implemented in premium cars, were designed in the form of a system operating on a laptop, with quite large system resources and computing power.

Due to the fact that miniaturization is ubiquitous nowadays, and the ever-smaller equipment has enormous computation capabilities, authors of the article have attempted to create a light, simple, operating on a microcomputer system that could be embedded in the dashboard and to which a touch screen can be connected. The general system concept is presented in Fig.1.

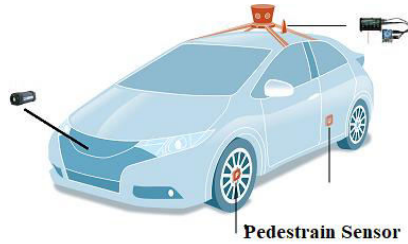


Fig.1: Main System Components

The main objective of the system is pedestrian detection in severe lighting conditions. Generally, we assume that the system should detect pedestrians in front of the car after dusk. With a limited visibility of the road the system will support the driver by displaying a grayscale thermal image with the detected objects surrounded by a frame and optionally offers the acoustic signal. Our initial experiments indicated that the detection rate of thermal imagery- based system surpasses the human perceptual abilities in low light conditions. The system should operate without delays offering the results online, preferable in real- time. We also assume that system components should be widely available enabling the development of the final solution.

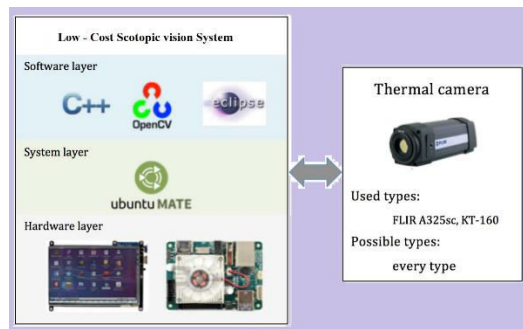


Fig.1.1: Multi-layer architecture of low-cost scotopic vision system.

Based on the above assumptions, we developed a multi-layer architecture consisting of hardware and software layers. It is presented in the above figure and described in details in the following paragraphs.

The main goal of the research, effective and fast algorithms are indispensable. The methods used should be capable of operation in a reasonable time under limited resources resulting from the microcomputer power. To be a low cost, the system should draw on existing Open-Source solutions. We opted for the OpenCV library which is a renowned standard for real- time image processing and recognition tasks. In this model, two models are used to detect the pedestrian on the road in severe lighting conditions. They are

- Yolov2.
- Haar Hog Descriptor with adaboost Algorithm.

3.3 Yolov2

YOLO is one of the best models in object recognition, able to recognize objects and process frames at the rate up to 150 FPS for small networks. However, In terms of accuracy mAP, YOLO was not the state of the art model but has fairly good Mean average Precision (mAP) of 63% when trained on PASCAL VOC2007 and PASCAL VOC 2012. However, Fast R-CNN which was the state of the art at that time has anmAP of 71% [3]. The previous YOLO architecture has a lot of problems when compared to the state-of- the-art method like Fast R-CNN. It made a lot of localization errors and has a low recall.

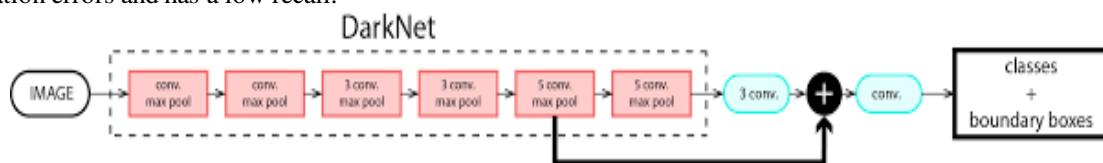


Fig.2 DarkNet architecture of Yolov2.

These sections should be organized to best present the material.

There are some incremental improvements that are made in basic YOLO.

3.3.1 High Resolution Classifier

The previous version of YOLO uses 224 * 224 as input size during training but at the time of detection, it takes an image up to size 448*448. This causes the model to adjust to a new resolution that in turn causes a decrease in mAP. The YOLOv2 version trains on higher resolution (448 * 448) for 10 epochs on ImageNet data. This gives network time to adjust the filters for higher resolution. By training on 448*448 images size the mAP increased by 4%.

3.3.2 Anchor Boxes for Bounding Boxes

YOLO uses fully connected layers to predict bounding boxes instead of predicting coordinates directly from the convolution network like in Fast R-CNN, Faster R-CNN. In this version, the fully connected layer is removed and instead add the anchor boxes to predict the bounding boxes.

3.3.3 Direct Location Problem

The previous version of YOLO does not have a constraint on location prediction which makes it unstable on early iteration. YOLOv2 predicts 5 parameters (t_x, t_y, t_w, t_h, t_o (objectness score)) and applies the sigma function to constraint its value falls between 0 and 1. This direct location constraint increases the mAP by 5%.

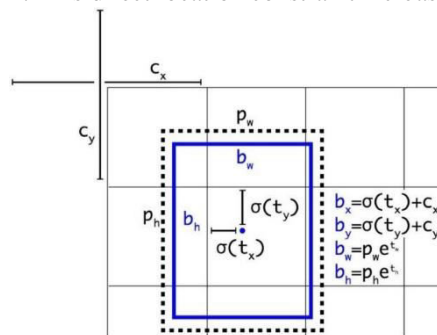


Figure 2.1: Anchor Boxes.

The following are the terms used in anchor boxes

$$b_x = \sigma(t_x) + c_x$$

$$b_y = \sigma(t_y) + c_y$$

$$b_w = p_w e^{t_w}$$

$$b_h = p_h e^{t_h}$$

$$\text{pr}(\text{object}) * \text{IOU}(b, \text{object}) = \sigma(t_o)$$

where, t_x, t_y, t_w, t_h , are predictions made by YOLO,

c_x, c_y is the top left corner of the grid cell of the anchor,

p_w, p_h are the width and height of the anchor,

c_x, c_y, p_w, p_h are normalized by the image width and height,

b_x, b_y, b_w, b_h are predicted boundary box

$\sigma(t_x)$ is the box confined score.

3.3.4 Fine Grained Features

YOLOv2 which generates 13 * 13 is sufficient for detecting large objects. However, if we want to detect finer objects, we can modify the architecture such that the output of previous layer 26 * 26 * 512 to 13 * 13 * 2048 and concatenates with the original 13 * 13 * 1024 output layer making our output layer of size.

3.3.5 Multi-Scale Training

YOLO v2 has been trained on different input sizes from 320 * 320 to 608 * 608 using step of 32. This architecture randomly chooses image dimensions for every 10 batches. There can be a trade-off established between accuracy and image size. For Example, YOLOv2 with images size of 288 * 288 at 90 FPS gives as much as mAP as Fast R-CNN. The algorithm automatically generates mask image without user interaction that contains only text regions to be inpainted.

IV. ARCHITECTURE

Figures shows the results of text detection from an image and inpainting by using exemplar. The following figures shows the original image, where the image is obtained by applying first set of criteria. All objects whose area greater than 10000 and filled area greater than 8000 are eliminated and major axis lengths are in between 20 to 3000 are considered to be text. Still, some small non-text objects are detected.

YOLO v2 is trained on different architectures such as VGG-16 and GoogleNet. An architecture called Darknet-19. The reason for choosing the Darknet architecture is its lower processing requirement than other architectures 5.58 FLOPS (as compared to 30.69 FLOPS on VGG-16 for 224 * 224 image size and 8.52 FLOPS in customized GoogleNet).

For detection purposes, we replace the last convolution layer of this architecture and instead add three 3 * 3 convolution layers every 1024 filters followed by 1 * 1 convolution with the number of outputs need for detection.

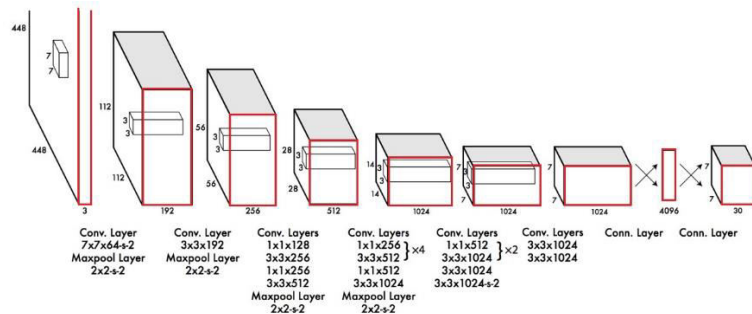


Fig.3. Architecture of YOLOv2.

YOLOv2 gives state-of-the-art detection accuracy on the PASCAL VOC and COCO. It can run on varying sizes offering a trade-off between speed and accuracy. At 67 FPS, YOLOv2 can give anmAP of 76.8 while at 40 FPS the detector gives an accuracy of 78.6 mAP, better than the state-of-the-model such as Faster R-CNN and SSD while running significantly faster than those models.

3.1 Haar Hog Descriptor with Adaboost Algorithm

Haar-like features are mainly used in object recognition, since it notes the detailed information of the object. These features are used in human face recognition system where it needs very fine details should be noted. This feature got the name due to its similarity with Haar wavelets.

The Haar-like feature adds the intensity values in each pixel location of a subsection image and computes the difference between the sums. Those differences can be used to identify the particular subsections of the image. For example, while recognition of Haar-like features, if the region of lips is darker than the region of nose. As a result, Haar-like features of a face is obtained by placing two rectangles over the lips and nose.

According, a window of certain size is slides over the input image, in each and every subsection of an input image, Haar-like features is computed. Every time the difference of subsection image is compared and simultaneously learned to a detector for classification of desired object /non-object.

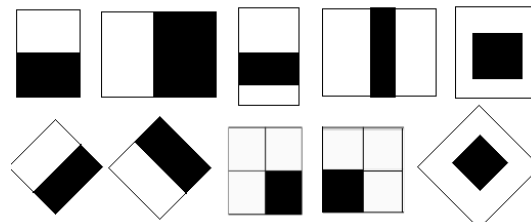


Fig. 3.1: Different types of Haar-like features.

However, this type of scheme works little bit better than random guessing, but an iterative scheme is introduced to a large group of Haar-like features of similar object is trained by a classifier cascade to produce a strong classifier which can classify an object with high accuracy.

The main advantage of Haar-like features is computation speed. It has a high computation speed compared to others which makes the system competitive in real time object recognition. The fast computation of Haar-like features is determined by summed area tables, also called integral images. The rotated Haar-like features are also implemented for

performance improvisation in object recognition. Figure 3.1 shows the different types of Haar- like features with different orientations.

3.2. Histogram of Oriented Gradients

HOG has been used as a key descriptor in the area of object detection, where a gradient orientation plays a major role. This HOG technique is almost similar to scale-invariant feature transform but differs in few situations where it performs on dense grid of uniform spaced cells. HOG feature descriptors which was first described, where it shows a great success in detection of pedestrians in both images and videos, and even tested on few kind of common animals and vehicles in static images.

The essential thought behind HOG is that the appearances and shapes of nearby objects inside an image might be decently portrayed by the dissemination of intensity gradients as the votes for predominant edge directions.

Such feature descriptor might be gotten by first separating the image into little coterminous locales of equivalent size, called cells, then gathering all the histogram of gradient orientations for every pixel inside each one cell, and finally joining all these histograms. With a specific end goal to enhance the recognition precision against illumination variation and shadowing, neighborhood contrast normalization could be connected by computing the values of an intensity over a bigger space of an image is called as a block, and utilizing the resultant quality to normalize all cells inside the block.

Thus, HOG descriptors withstand greater advantages over other techniques. These features were more robust to the movement of the pedestrian which is suitable for detection of humans. In general, there are four types of HOG-block techniques are available: Rectangular HOG (R-HOG), Fig 3.2 shows the Circular HOG(C-HOG), Bar HOG and Centre-surround HOG.

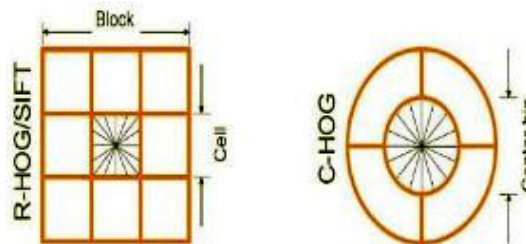


Fig. 3.2: Left: R-HOG block. Right: C-HOG block.

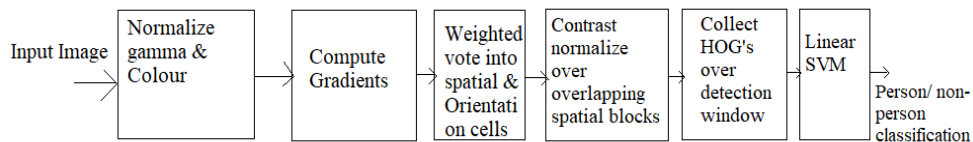


Fig. 3.3: Hog Descriptor for Person classification

3.3. Gradient Computation

The first step in extracting HOG features is of computing image gradients. A 1D cantered derivative mask is applied both in the direction of horizontal and vertical for calculating Image gradients. The predominant filter kernel used in calculating gradients is given below

$$[1 \ 0 \ 1] \text{ and } [1 \ 0 \ 1]^T$$

Actually, there are many tedious masks, such as Canny, Sobel, Prewitt and diagonal masks, but the performance of these masks were poorer compared to the proposed window and Fig 3.2 and fig 3.3 were given by [4]. The computation of magnitude and orientation at each pixel location $I(x,y)$ is calculated by

$$\begin{cases} G_{mag}(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)} \\ \theta(x,y) = \arctan(G_y(x,y)/G_x(x,y) + \pi/2 \end{cases} \quad (1)$$

Where $G_x(x, y)$ and $G_y(x, y)$ are the gradients at each pixel location in horizontal and vertical direction, respectively

In the case of colour images, the channel with larger magnitude gives the magnitude and orientation of the dominant pixel. While finding the angle it is necessary that to add $\pi/2$ since the angle value results in the range between $-\pi/2$ to $\pi/2$. It gives better performance when it ranges from 0 to π for an unsigned orientation.

3.4. Orientation Binning

Here in this orientation binning procedure, Histogram of each cell is calculated and the cells which are group of pixel regions are either rectangular or radial in shape, and those bins were evenly distributed from 0° to 180° , so every histogram bin will consist of a spread of 20° . Each and every pixel in a cell will consider its weighted voting to assign one of 9 bins according to its orientation. The weight of votes mainly considers its gradient magnitude but some other methods are also available for example the square or square root of the gradient magnitude etc.

3.5. Block Descriptors

Illumination and contrast variations makes the system more vulnerable to noise. To attain robustness the strength of the gradient should be normalized locally. This will make the blocks with larger group of pixels or combination of few cells. These blocks overlaps with others in its neighbouring hood to contribute its orientation more than once. As mentioned earlier there are four kinds of geometric blocks used among those rectangular blocks is commonly used. In general R-HOG block consists of square grids and the optimal parameters are said to be 2 X 2 cell blocks of 8 X 8 pixels each.

3.6. Block Normalization

Normalization of blocks can be done by four different ways. Let u denotes the non-normalized feature vector which collects the histogram of all cells in a block, $\|u\|_k$ denotes k-norm where $k=1, 2$ and ϵ is a small constant.

In this contest, a normalized scheme which is given below Normalization scheme is used.

$$\frac{v}{\sqrt{\|v\|_2^2 + \epsilon^2}} \quad (2)$$

The final feature vector of HOG is containing the elements of normalized histograms from all of its block regions.

3.7. Classification Methods

There are many classification methods which will classify the objects based on different types of learning algorithms and with different features of the objects. Classification algorithms generally consists of two modules, one is of training module and another is testing module. In the introductory training stage, an extraordinary depiction of each one class is made by learning with normal features removed from the training samples and dividing them in the peculiarity space. In the subsequent testing stage, these feature space partitions are utilized to group recently include characteristic vectors removed from the testing dataset.

The most popular classification techniques are AdaBoost and SVM which performs outstandingly. AdaBoost classification technique is used in object recognition and classification and SVM is a binary classifier which classifies a set of two.

AdaBOOST

AdaBoost is nothing but Adaptive Boosting. This technique is used for improving the performance of many algorithms. AdaBoost learning algorithm applies sequentially by its input training data and each prediction adds to an ensemble. When each ensemble is added the learning is weighted according to its accuracy. Even the dataset is also reweighted by samples which classifies correctly loses weight and samples classifies wrongly gains weight.

Support Vector Machine

Support Vector Machine (SVM) is a popular technique for the classification. The main criteria in this technique is dealing with the training data and testing data. Here we have to train the system by using training data which should make learn the system to classify the object. This technique can be used very easily. Few examples where the classification can be done by using SVM is in hand written digit recognition, spam filtering, text classification and mainly used in pedestrian detection.

- Working procedure of an SVM is given below
- Training vectors x_i where $i=1,2,\dots,l$
- Train SVM with training data x_i
- Consider a decision vector y with two classes which is a hyper plane separates all the data

$$y_i = \begin{cases} 1 & \text{if } x_i \text{ is in class 1} \\ -1 & \text{if } x_i \text{ is in class 2} \end{cases}$$



There are few problems regarding the separation of data in SVM. Few data sets which are given for classification can be linearly separable, whereas few data sets can't be linearly separable. A typical learning algorithm is achieved for separation of data for those type of cases.

V. EXPERIMENT RESULTS

Following table shows the results of text detection from an image and inpainting by using exemplar based Inpainting algorithm. Figs. 2, 3, 4 (a) shows the original image. (b) is the image obtained by applying first set of criteria. All Below report shows different test cases and the results obtained after performing those cases.




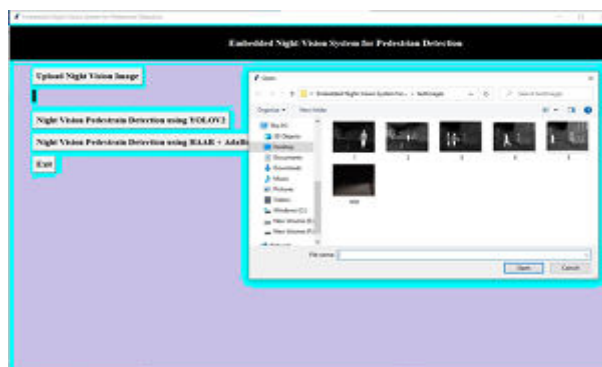
Cases	Input	Expected Output	Actual Output
Processing without uploading Image	Don't upload test image	"No image selected"	It shows "No Image selected"
Processing with test image-1	Uploaded test image-1 	Pedestrian detected by both models.	Pedestrian Detected by both models.
Processing with test image-2	Uploaded test image-2 	Pedestrian detected by both models.	Pedestrian detected correctly by HogDescriptor with adaboost. YOLOv2 failed to detect.
Processing with test image-3	Upload test image-3 	Pedestrian detected by both models.	Haar HogDescriptor with adaboost succeed to detect the pedestrian. YOLOv2 failed to detect.

Table 1. Test Report of Scotopic Vision System

Objects whose area greater than 10000 and filled area greater than 8000 are eliminated and major axis lengths are in between 20 to 3000 are considered to be text. Still, some small non-text objects are detected.



(a) Upload Window Scotopic-vision Image.



(b) Uploaded Image of Scotopic-Vision System.



(c) Pedestrian Detection using Yolov2 Model.



(d) Pedestrian Detection using HaarHog Descriptor with adaboost

VI. CONCLUSION

The system addresses topic related to the safety of pedestrians on the road. We investigated the ability to build a sensor for human body detection in severe lighting condition based on existing algorithms and widely available hardware. In contrast to existing solutions found on the market, installed generally in expensive cars, our proposal is low cost. Built on the state-of-the-art algorithms, it has demonstrated accuracy in difficult lighting conditions. Systems of this kind can offer the extension of the human visual senses and improve the safety on the road.

Regarding the future scope, there is a need of more research to reduce the false negatives and false alarm rate particularly in cluttering back ground, Occlusion and while crossing the road where the shape of the pedestrian was not clear. Efficiency should also be considered whenever needed to balance the robustness. These systems which I done in this research work are comprised to images, so that is enough to have a detection, but when coming to real time we need to work on videos where a tracking method also necessary to implement.

REFERENCES

- [1] S. P. T. Yang, D. Fu, "Pedestrian tracking for infrared image sequence based on trajectory manifold of spatio-temporal slice," *Multimedia Tools and Applications*, vol. 76, no. 4, p. 11021, 2017.
- [2] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," arXiv preprint arXiv:1612.08242, 2016.
- [3] J. H. Kim, G. Batchuluun, and K. R. Park, "Pedestrian detection based on faster r-cnn in nighttime by fusing deep convolutional features of successive images," *Expert Systems with Applications*, vol. 114, pp. 15 – 33, 2018. [Online]. <http://www.sciencedirect.com/science/article/pii/S0957417418304354>
- [4] P. Forczmanski and A. Nowosielski, "Deep learning approach to detection of preceding vehicle in advanced driver assistance," in *Challenge of Transport Telematics - 16th International Conference on Transport Systems Telematics, TST 2016, Katowice-Ustron, Poland, March 16-19, 2016, Selected Papers, 2016*, pp. 293–304. [Online]. Available: https://doi.org/10.1007/978-3-319-49646-7_25
- [5] M. Szarvas, A. Yoshizawa, M. Yamamoto, and J. Ogata, "Pedestrian detection with convolutional neural networks," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, June 2005, pp. 224–229
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, ser. NIPS'12*. USA: Curran Associates Inc., 2012, pp. 1097–1105. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2999134.2999257>.
- [7] R. Burduk, "The adaboost algorithm with the imprecision determine the weights of the observations," in *Intelligent Information and Database Systems, N. T. Nguyen, B. Attachoo, B. Trawinski, and K. Somboonviwat, Eds. Cham: Springer International Publishing, 2014*, pp. 110–116.
- [8] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of online learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997. [Online]. <http://www.sciencedirect.com/science/article/pii/S002200009791504X>.
- [9] R. Sun, G. Zhang, X. Yan, and J. Gao, "Robust pedestrian classification based on hierarchical kernel sparse representation," *Sensors*, vol. 16, no. 8, 2016. [Online]. Available: <http://www.mdpi.com/1424-8220/16/8/1296>.
- [10] X. Fu, R. Yu, W. Zhang, J. Wu, and S. Shao, "Delving deep into multiscale pedestrian detection via single scale featuremaps," *Sensors*, vol. 18, no. 4, p. 1063, 2018. [Online]. Available: <https://app.dimensions.ai/details/publication/pub.1101899151> and <http://www.mdpi.com/14248220/18/4/1063/pdf>.
- [11] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004. [Online]. Available: <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>.
- [12] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, June 2005, pp. 886– 893 vol. 1.
- [13] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Proceedings of 12th International Conference on Pattern Recognition*, vol. 1, Oct 1994, pp. 582–585 vol.1.
- [14] P. Forczmanski, "Human face detection in thermal images using an ensemble of cascading classifiers," in *Hard and Soft Computing for Artificial Intelligence, Multimedia and Security, S.-y. Kobayashi, A. Piegat, J. Pejas, I. El Fray, and J. Kacprzyk, Eds. Cham: Springer International Publishing, 2017*, pp. 205–215.
- [15] D. E. King, "Max-margin object detection," 2015.
- [16] K. Matecki, P. Forczmanski, A. Nowosielski, A. Smolinski, and D. Ozga, "A New Benchmark Collection for Driver Fatigue Research Based on Thermal, Depth Map and Visible Light Imagery." Springer, 01 2020, pp. 295–304.
- [17] Yuxi Chen and Chongzhao Han, "Night-time pedestrian detection by visual-infrared video fusion," in *2008 7th World Congress on Intelligent Control and Automation*, June 2008, pp. 5079–5084.
- [18] Hangil Choi, S. Kim, Kihong Park, and K. Sohn, "Multi-spectral pedestrian detection based on accumulated object proposal with fully convolutional networks," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, Dec 2016, pp. 621–626.
- [19] Y. Hou, Y. Song, X. Hao, Y. Shen, and M. Qian, "Multispectral pedestrian detection based on deep convolutional neural networks," in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Oct 2017, pp. 1–4.
- [20] R. S. Ghiass, O. Arandjelovic, H. Bendada, and X. Maldague, "Infrared face recognition: A literature review," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, Aug 2013, pp. 1–10.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

Impact Factor:
7.512

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details