PROJECT REPORT ON

"PEDESTRIAN DETECTION AND COUNTING USING HOG + SVM AND

PEDESTRIAN DETECTION USING OPTICAL FLOW METHOD"

SUBMITTED TO-

Dr. Dushyant Kumar Singh
Assistant Professor
CSE Dept.(MNNIT Allahabad)

SUBMITTED BY-

Raushni Kumari(1507004)
Priya Kumari(1507042)
B.tech(6^{th} semester)
NIT,PATNA

DECLARATION

We hereby declare that this project work entitled "PEDESTRIAN DETECTION AND COUNTING USING HOG+SVM AND PEDESTRIAN DETECTION USING OPTICAL FLOW METHOD" has been carried out by us in the department of COMPUTER SCIENCE & ENGINEERING MNNIT Allahabad under the guidance of Dr. *Dushyant Kumar Singh*, Assistant professor.

CERTIFICATE

This is to certify that………NAME……… **(**….ROLL NO….**)** have carried out the project entitled "PEDESTRIAN DETECTION AND COUNTING USING HOG+SVM AND PEDESTRIAN DETECTION USING OPTICAL FLOW METHOD" under the supervision of Dr. *Dushyant Kumar Singh*, Assistant professor of Department COMPUTER SCIENCE & ENGINEERING MNNIT Allahabad .This project reports the bonafide work done by them for the complete fulfillment of the requirement for the award of summer research internship in Computer Science & Engineering.

Dr. *Dushyant Kumar Singh*

(Project Guide)

Assistant professor

COMPUTER SCIENCE & ENGINEERING

MNNIT Allahabad

ACKNOWLEDGEMENT

We hereby take the privilege to express our gratitude to all the people who were directly or indirectly involved in the execution of this work, without whom this project would not have been a success.

We extend our deep gratitude, respect and obligation to our project supervisor, Dr. Dushyant Kumar Singh , Assistant Professor, Department of Computer Science and Engineering, MNNIT Allahabad for his timely suggestions and guidance.

Our heartiest thanks to our projectmates who have supported us in all possible ways. Words are inadequate to express our gratitude to our parents and friends who were supportive all the time. We would also like to thank MNNIT Allahabad to provide us opportunity to do such project.

Above all we bow before The Almighty for his immense blessings throughout the life.

PLACE: ALLAHABAD

PRIYA KUMARI

RAUSHNI KUMARI

CONTENTS

1. Introduction

Image detection and recognition technology is an important branch in computer vision and pattern recognition, which is used in pedestrian detection, intelligent monitoring and other fields. The nature of the image detection is a classification process mainly based on its feature, so good classifier for improving detection result has an important influence. In a mixed transportation system, pedestrian and vehicle are a major component of traffic. To detect the presence of pedestrian ahead of the vehicle and promptly inform the driver of the vehicle to adopt safety forecast for reducing or avoiding a collision with a pedestrian is of great significance. Thus pedestrian detection is very important. The key of pedestrian detection is the design of classifier, the performance of which determines the accuracy of pedestrian detection. Generally, after the sample feature and classifier type selected, the key of design is to establish a suitable pedestrian and non-pedestrian sample library. We adopt support vector machine to detect the pedestrian in video, focusing on the design of pedestrian detection classifier. Firstly, to build a pedestrian sample library, extract the Histogram of Oriented Gradient feature of the sample and then utilize support vector machine to obtain the pedestrian classifier, finally test the performance of the classifier, for the shortcoming of the classifier, we improved the sample library, retrained to get better performance classifier and we also used optical flow method to detect the pedestrians. Optical flow can be used to segment a moving object from its backgrounds and track it. In this paper, an enhanced Lucas Kanade optical flow technique was used to improve human detection in terms of speed and accuracy. We combined object segmentation output with a human detector using an optical flow algorithm. The proposed technique used the optical flow to find the area
of interest to complete object segmentation and use those results as an input for the human detector. This technique has been developed to be used in surveillance systems.

2. Problem Statement

An application of a pedestrian-detection system and counting pedestrians passing through a line in video sequences based on HOG feature + SVM and Optical flow method that aimed at localizing potentially dangerous situations under specific urban scenarios.

1. A pedestrian is a person travelling on foot, whether walking or running.
2. We intend to implement a system based on supervised learning of HoG.
3. Feature extraction from the Pedestrian database and classified using a linear SVM.
4. Detection of the pedestrian in a video and count the number of pedestrian.
5. Detection of the pedestrian through Optical flow method.

3. Methodology

In our approach for Pedestrian detection we have used two type of method.Which are given below-

3.1 HOG+SVM Method

Our implementation will be closely based on the works on Dalal and Triggs with little modification. Figure 2 shows the process flow we intend to follow.

3.1.1 Image Acquisition

Images acquired from standard dataset were used for testing. Other than that picture captured from mobile phone cameras were used to validate the system.

3.1.2 Preprocessing

We will be working on processing the images in the gray scale space as the above cited paper suggests and gives no distinct advantage of using the RGB or LAB color spaces. Apart from that we apply gamma normalization to improve the intensity of the image. This is done as images taken from the mobile device had low illumination.

3.1.3 Dataset

We will use the MIT and (Figure 3)INRIA Pedestrian data set for Training and Testing of our SVM. Instead of the Hold-Out method used in the cited approach we plan to use k-cross validation. We have a data set of approx 2100 images.(900 MIT +1200 INRIA). We try all combinations of the datasets to identify the best one.

3.1.4. HOG Extraction

We will be implementing the HoG Feature Extraction procedure from scratch following the implementation given by Dalal and Triggs:

1. Gradient Computation: The most common method to compute gradient is to simply apply the point discrete derivative mask in both horizontal and vertical direcd tions. This method requires filtering the intensity data of the image with the kernels [-1 0 1] in both horizontal and vertical directions.

2. Orientation Binning: A cell histogram is created by weighted quantization of the orientation of each pixel of the the cell into pre-defined orientation-based bins. The cells are usually square in shape (for convenience we will stick with rectangular), but they can be rectanw gular or circular. The weighting of the orientations can be either by using the gradient magnitude itself.

3. Block Division and Normalization: The cells must be grouped together in order to factor in changes in ilg lumination and contrast. The complete HOG descripl tor is then the vector of the components of the nort malized cell histograms from all of the block regions. Two main block geometries exist: rectangular R-HOG blocks and circular C-HOG blocks. R-HOG blocks are generally square grids, represented by three parameg ters: the number of cells per block, the number of p pixt els per cell, and the number of channels per cell hise togram. C-HOG has two variants: ones with a single, central cell and the ones with angular divided cells, which can be described by number of angular/radial bins, radius if center bin and expansion factor for rab dius of additional radial bins. These blocks are normald ized by four prominent methods: L1 norm, L1 norm square root, L2 norm and L2 norm followed by clips ping (L2 Hys). We will experiment and choose the one that works best.

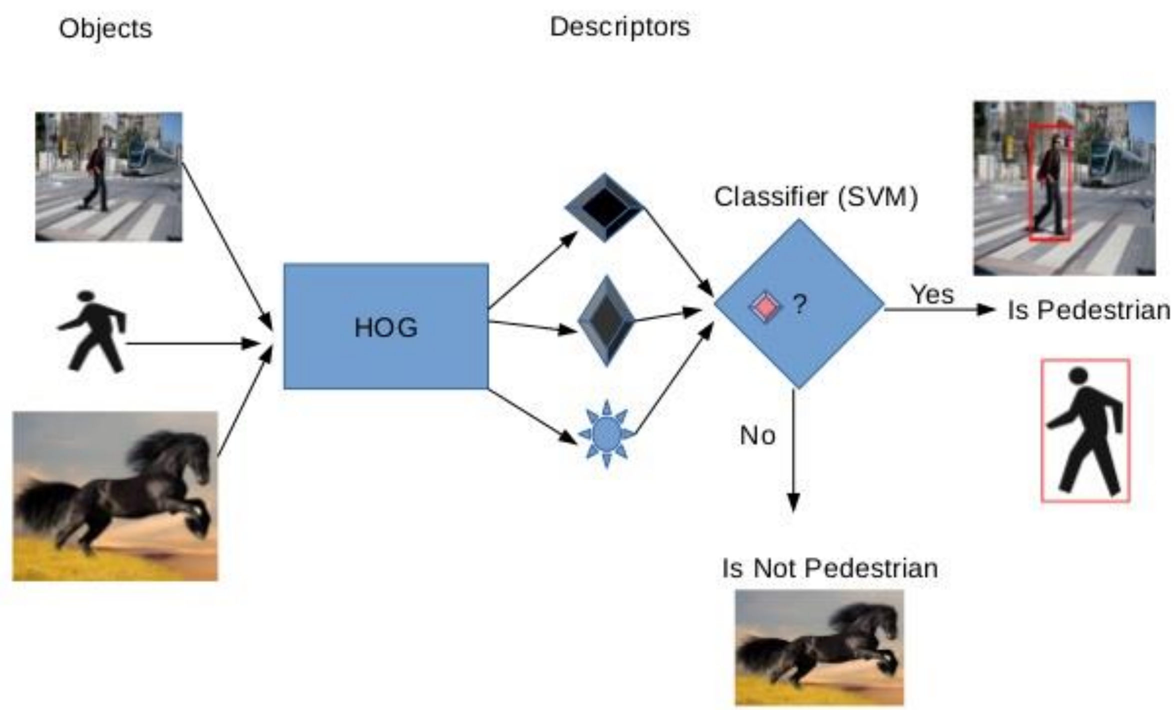
3.1.5. Training the Classifier

Based on the literature survey done, we chose a linear kernel SVM in our project.We kept the error cost to the default value of 1. We attempted using weighted SVMs.

3.1.6. Weighted SVM

We observed that pedestrians are actually being classified based on their silhoutte boundaries rather than the whole image itself.On the contrary ,rest of the image is actually proving to be counter productive as the svm is being trained on them as well.The variation in the texture found on the clothes and the random texture in the background can be dealt with by giving them lower weights and giving them more weight to regions around these silhoutte.We implement this by sampling the gradient image into the same number of blocks as in the case of

HOG extraction.All the histogram of input from one block are given the corresponding weight as in the sampled image.



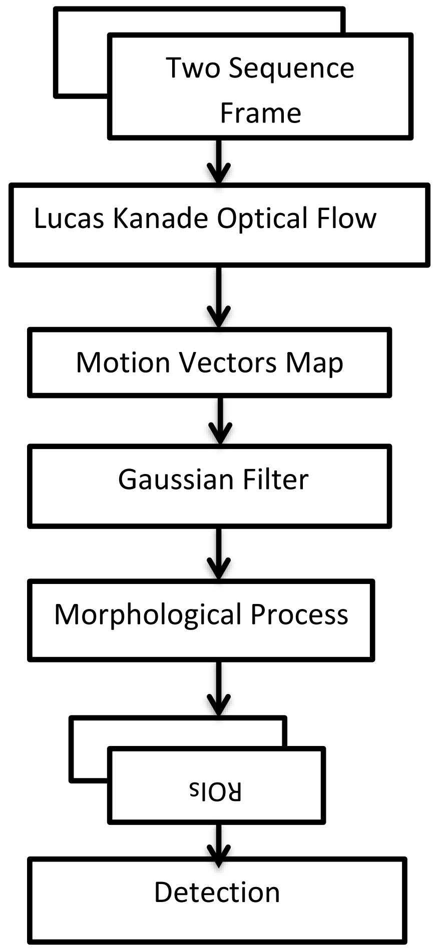
Pedestrian detection Process

4

3.2 Optical Flow

Optical flow is the pattern of apparent motion of image objects between two consecutive frames caused by the movement of object or camera.It is a 2D vector field where each vector is a displacement vector showing the movement of points from first frame to second.In our implementation for pedestrian detection we have used Lucas Kanade method .

- The Lucas Kande optical flow technique is enhanced and used to improve the human detector speed and accuracy.The Lucas Kanade method is commonly used as a differential method for optical flow estimation in computer vision. It uses least squares criterion to solve basic optical flow equations for all the pixels in a neighbourhoods . It is a purely local method assumes that the flow is constant in local neighbourhoods of the pixel under consideration. Its advantages are that it can make very fast calculations and accurate time derivatives. The disadvantage of the Lucas Kanade method is the creation of errors regarding the boundaries of moving object.
- Then Gaussian filter was used to reduce unwanted motion vectors to avoid wrong segmentation Every two frames were processed through the enhanced optical flow technique to get motion vectors (map of vectors).
- In order to use this new map to do segmentation and obtain the Region of Interest (ROI), the new map of motion vectors were converted to binary blobs.
- Furthermore, some morphological processes were performed to enhance segmentation stage by removing isolated noises. Finally, segmentation for each blob was performed and processed to human detector.
- To achieve an efficient and effective human detection, detector is used to recognize human motion among all motion objects in the scene.



<u>Steps for enhancing motion pedestrians detection performance</u>

4. Algorithms used –

4.1. HOG(Histogram of Oriented Gradient)

The basic idea behind this approach is capturing the object appearance and shape by Characterizing it using Local intensity Gradients and Edge Direction.
The image is densely divided into small spacial regions called cells.
For each cell a 1-D histogram of gradient directions/edge directions is coputed and later all cell data is combined to give a complete HoG descriptor of the winc dow. The variety of colors and illumination in the surroundd ing makes normalization inevitable. We further describe the normalization technique as a part of our approach later in the report. In their work, Triggs and Dalal[1] use the MIT and Inria Dataset(1805).They generate a sufficiently large negative set by sampling out patches from person-free images. Later in their work, they provide a detailed compari ison with other state-of-the-art methods- Generalized Haar Wavelets, PCA-SIFT and Shape Context methods showing great superiority of the HOG implementation over the rest. Besides the usual square R-HOG Blocks, comparison has even been shown after carrying out the training and teste ing with vertical descriptors (2x1 cell), horizontal descripi tor (1x2 cell) and C-HOG geometry Blocks. Learning from their inferences we look to implement the proposed system.

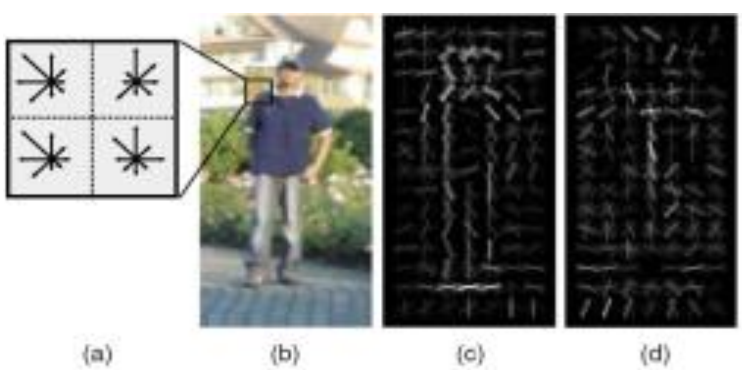
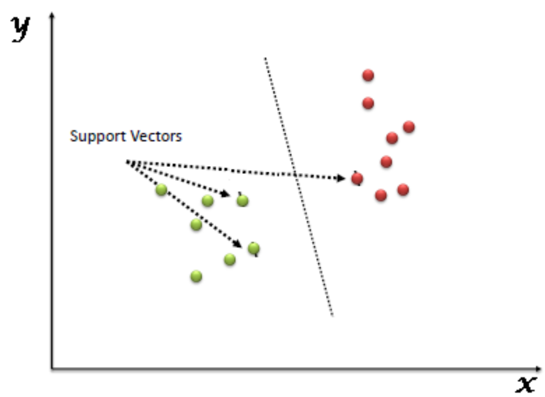


Fig 1. Histogram of Oriented Gradient

4.2. SVM(Support Vector Machine)

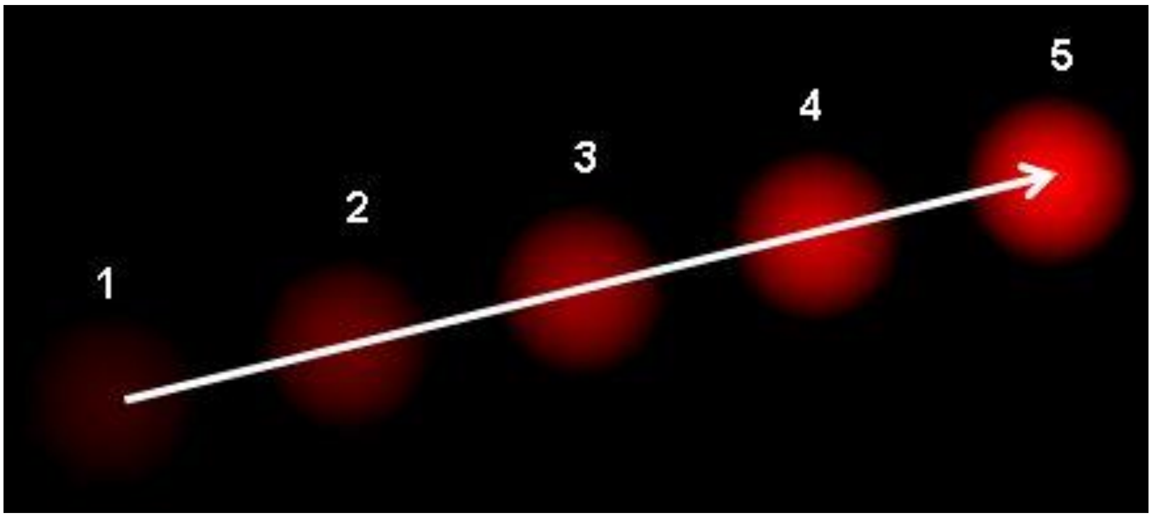
“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well.



Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes (hyper-plane/ line).

4.3. Optical Flow

Optical flow is the pattern of apparent motion of image objects between two consecutive frames caused by the movement of object or camera. It is 2D vector field where each vector is a displacement vector showing the movement of points from first frame to second. Consider the image below:



It shows a ball moving in 5 consecutive frames. The arrow shows its displacement vector. Optical flow has many applications in areas like:

- Structure from Motion
- Video Compression
- Video Stabilization

Optical flow works on several assumptions:

1. The pixel intensities of an object do not change between consecutive frames.
2. Neighbouring pixels have similar motion.

Consider a pixel $I(x, y, t)$ in first frame. It moves by distance $(dx\ dy)$ in next frame taken after dt time. So since those pixels are the same and intensity does not change, we can say,

$$I(x, y, t) = I(x+dx, y+dy, t+dt)$$

Then take Taylor series approximation of right-hand side, remove common terms and divide by dt to get the following equation:

$f_x u + f_y v + f_t = 0$

Where:

$$f_x = \partial f/\partial x; \; f_y = \partial f/\partial y$$

$$u = dx/dt; v = dy/dt$$

Above equation is called Optical Flow equation. In it, we can find fx and fy, they are image gradients. Similarly ft is the gradient along time. But (u,v) is unknown. We cannot solve this

one equation with two unknown variables. So several methods are provided to solve this pand one of them is Lucas-Kanade.

LUCAS-KANADE METHOD

We have seen an assumption before, that all the neighbouring pixels will have similar motion. Lucas Kanade method takes a 3x3 patch around the point.So all the 9 points have the same motion. We can find (fx ,fy, ft) for these 9 points. So now our problem becomes solving 9 equations with two unknown variables which is over-determined. A better solution is obtained with least square fit method.

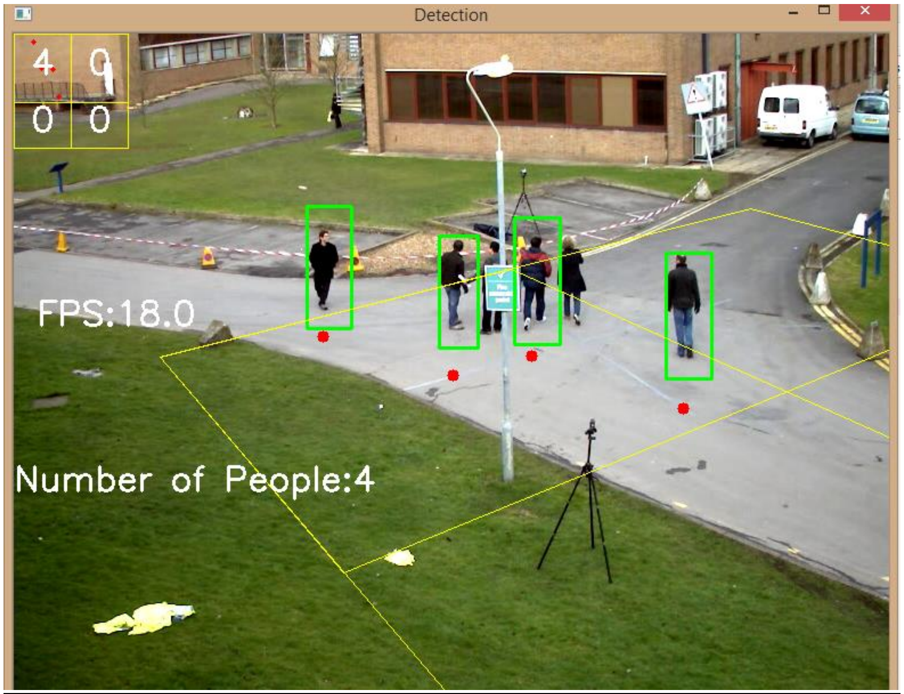
So from user point of view, idea is simple,we give some points to track, we receive the optical flow vectors of those points. But again there are some problems. Until now, we were dealing with small motions. So it fails when there is large motion. So again we go for pyramids.When we go up in the pyramid,small motions are removed and large motions becomes small motions. So applying Lucas Kanade there, we get optical flow along with the scale.

5. Experimental Results and Analysis

Result obtained from HOG+SVM method:-

On testing over a wide set of videos collected ,the system works reasonably robust.

While observing the pedestrian in any video or real time scenario, apart from the detection, we are also calculating the frame per second that explains about the frequency(rate) at which consecutive images called frames appear on the display,and counting of number of pedestrian at a particular time also takes place at the same time.The final detection of pedestrian through HOG+SVM the output is given below:

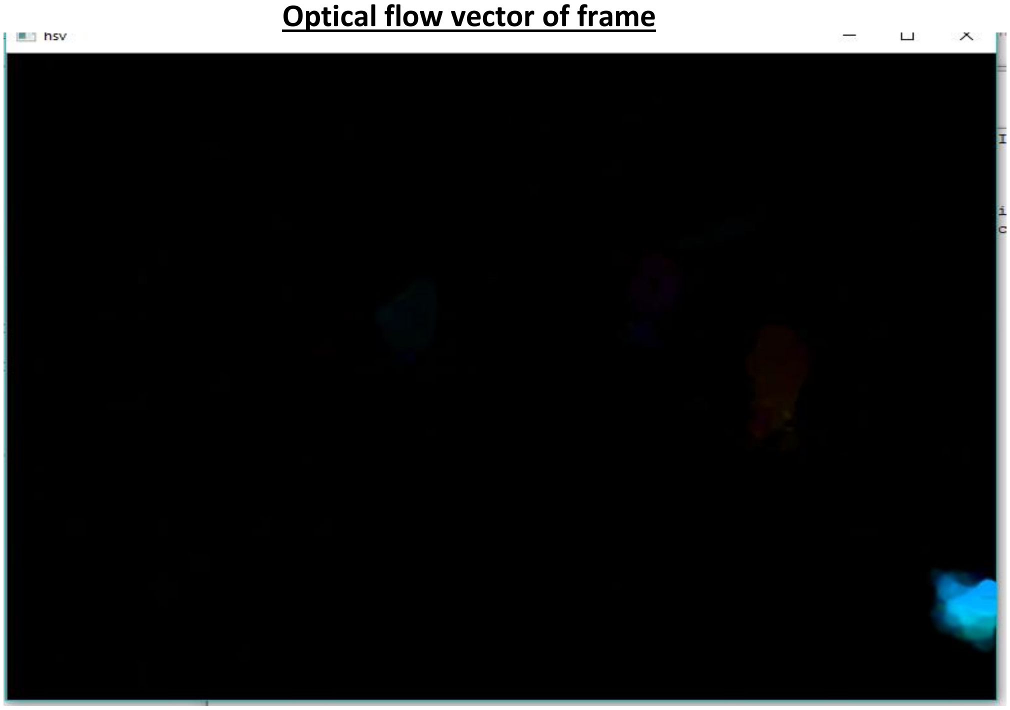


Result obtained from Optical flow method:

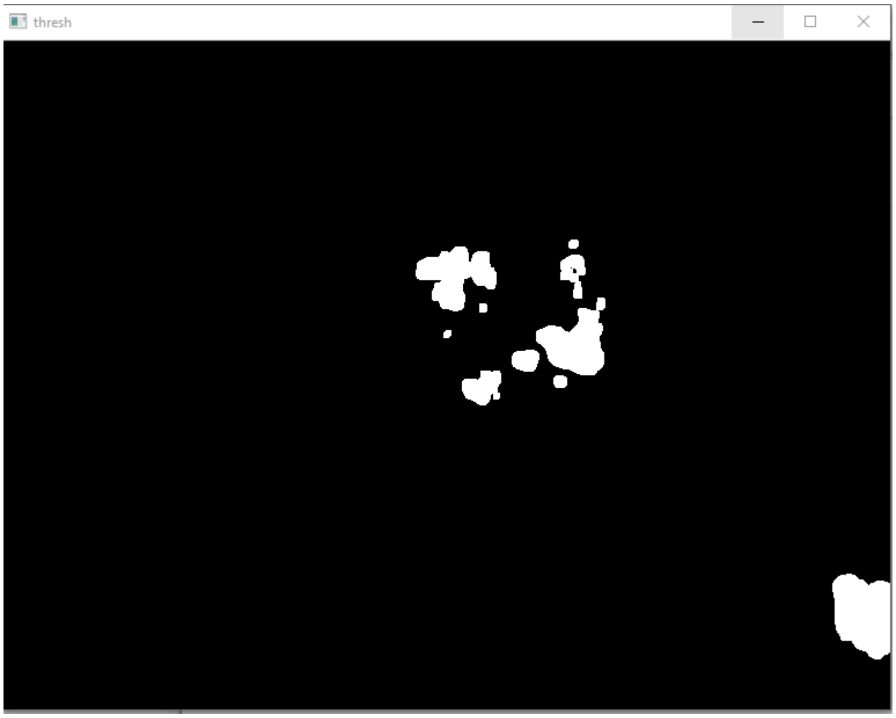
For pedstrian detection using Optical flow method we have got the following observations:

- Optical flow:It track the feature point in a video.
- HSV(Hue Saturation Value): we will create an application which extracts a colored object in a video.
- Thresholding of frame: If pixel value is greater than a threshold value, it is assigned one value (may be white), else it is assigned another value (may be black).
- Detection of pedestrian:final detection takes place i.e it identifies the pedestrian through the motion performed by the pedestrian.

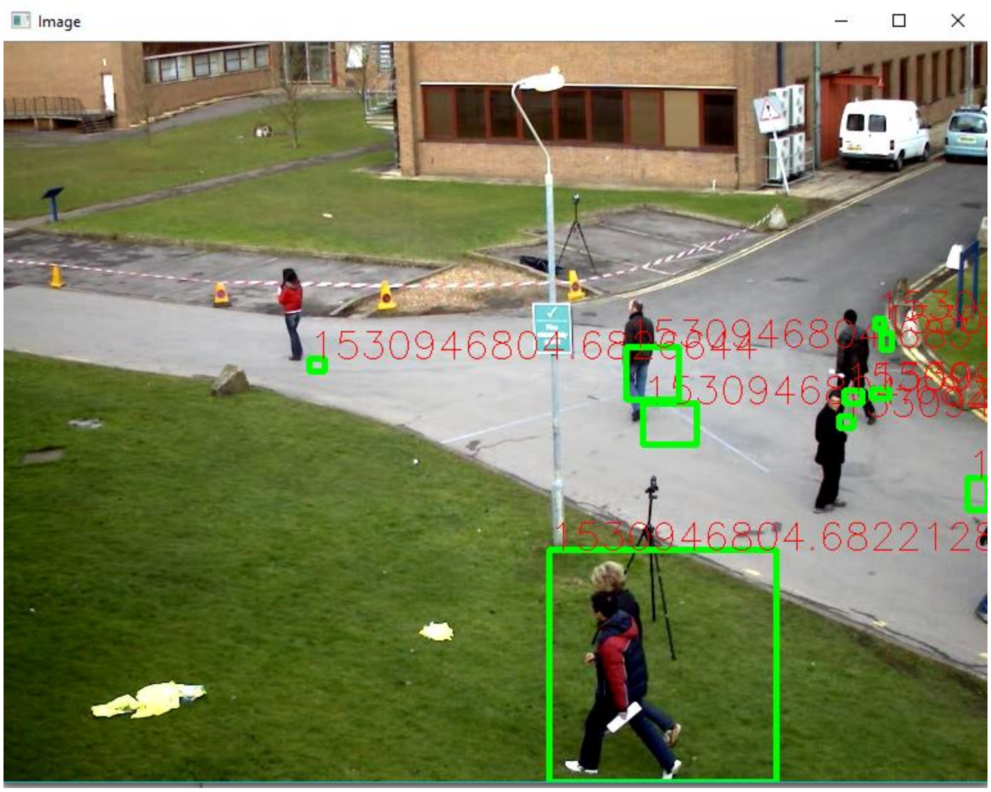
Optical flow vector of frame



HSV of frame



Thresholding of frames



Detection of pedestrian in frame

After detecting the pedestrian, main task was to calculate accuracy, rate and precision of the system. Accuracy, rate and precision of the system was calculated in the following manner:-

- **True positives (TP):** These are cases in which we predicted yes (they have the object of Interest), and they do have the object of Interest.
- **True negatives (TN):** We predicted no, and they don't have the object of Interest.
- **False positives (FP):** We predicted yes, but they don't actually have the object of Interest.
- **False negatives (FN):** We predicted no, but they actually do have the object of Interest

Where,

True positive rate=TP/ (TP+TN)
True negative rate=FN/ (FN+FP)
False positive rate=FP/ (FP+FN)
False negative rate=TN/ (TN+TP)
Precision=TP/ (FP+TP)
Accuracy = (TP+TN)/(TP+TN+FN+FP)

Format for Confusion Matrix:

Total	Predicted: NO	Predicted: YES
Actual: NO	TN	FP
Actual:YES	FN	TP

We have taken a video clip of 1 minute and 20 seconds, in which we have got 240 frames. From 240 frames we have considered 15 frames for the calculation of precision, rate and accuracy.

By using Optical flow method for pedestrian detection, we have got the following statistics that include true positive, true negative, false positive ,false negative ,their corresponding rates along with accuracy and precision for the consecutive frames. Confusion matrix for detection of pedestrian through optical flow method is shown below:-

Table: 1

Frame no.	TP	FP	TN	FN	precision	True positive rate	True negative rate	False positive rate	False negative rate	Accuracy(in %)
1.	2	1	0	2	0.66	1	0.66	0.33	0	40
2.	1	1	0	2	0.5	1	0.66	0.33	0	25
3.	1	1	0	2	0.5	1	0.66	0.33	0	25
4.	2	1	0	1	0.66	1	0.5	0.5	0	50
5.	0	0	0	0	1	1	undefined	undefined	0	100
6.	2	0	0	2	1	1	1	0	0	50
7.	3	0	0	1	1	1	1	0	0	75
8.	2	0	0	2	1	1	1	0	0	50
9.	2	0	0	2	1	1	1	0	0	50
10.	2	0	0	2	1	1	1	0	0	50
11.	3	2	0	2	1	1	0.5	0.5	0	42.85
12.	3	2	0	2	0.6	1	0.5	0.5	0	42.85
13.	3	1	0	2	0.6	1	0.66	0.33	0	50
14.	5	1	0	0	0.83	1	0	1	0	83.33
15.	3	2	0	2	0.6	1	0.5	0.5	0	42.85
										Average accuracy(in %)=51.79

By using HOG+SVM method for pedestrian detection, we have got the following statistics that include true positive, true negative, false positive ,false negative ,their corresponding rates along with accuracy and precision for the consecutive frames. Confusion matrix for detection of pedestrian through HOG+SVM method is shown below:-

Table: 2

Frame no.	TP	FP	TN	FN	Precision	True positive rate	True negative rate	False positive rate	False negative rate	Accuracy (in %)
1.	3	0	0	1	1	1	1	0	0	75
2.	5	0	0	1	1	1	1	0	0	83
3.	2	0	0	1	1	1	1	0	0	66.6
4.	4	0	0	1	1	1	1	0	0	80
5.	4	0	0	2	1	1	1	0	0	66.4
6.	3	0	0	1	1	1	1	0	0	75
7.	2	0	0	1	1	1	1	0	0	66.6
8.	4	0	0	1	1	1	1	0	0	80
9.	5	0	0	2	1	1	1	0	0	71
10.	5	0	0	2	1	1	1	0	0	71
11.	4	0	0	1	1	1	1	0	0	80
12.	5	0	0	0	1	1	1	undefined	0	100
13.	2	0	0	2	1	1	1	0	0	50
14.	3	0	0	1	1	1	1	0	0	75
15.	5	0	0	1	1	1	1	0	0	83
										Average accuracy(in %)=74.84

6. Conclusion

While computing the accuracy through optical flow method and HOG+SVM method we observe that the accuracy obtained using optical flow method is 51.79% although accuracy obtained through HOG+SVM is 74.84% that's why using HOG+SVM method for detection of pedestrian will be preferred over optical flow.

7. References

1. Pushkar P. Goswami, Diwakar Paswan, **D. K. Singh**, "Detecting moving objects in traffic surveillance video", International Journal of Control Theory and Applications, Vol 9.17, pp 8423-8430 (2016) [**Scopus]**
2. Histograms of Oriented Gradients for Human De[tection,Dalal, Navneet and Triggs, Bill,Proceedings of the 2005 IEEE Computer Society Conference on Computer Vi2 sion and Pattern Recognition (CVPR'05).
3. http://saliency.mit.edu/datasets.html
4. http://pascal.inrialpes.fr/data/human/
5. https://ieeexplore.ieee.org/document/5975165/
6. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1467353
7. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1336346
8. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4250466
9. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206724
10. https://scholar.google.com/scholar?as_q=Disparity+Statistics+for+Pedestrian+Detection%3A+Combining+Appearance%2C+Motion+and+Stereo%2C&as_occt=title&hl=en&as_sdt=0%2C31
11. https://thesai.org/Downloads/Volume7No9/Paper_55-Intelligent_Pedestrian_Detection_using_Optical_Flow.pdf
12. https://www.researchgate.net/publication/263844904_Moving_Object_Localization_Using_Optical_Flow_for_Pedestrian_Detection_from_a_Moving_Vehicle