

# Vehicle Surveillance and Tracking using Background Segmentation



Bhaggiaraj. S, Ranjeeth Kumar. C, Rahul Vijay. K. S, Vignesh Prabhu. A

**Abstract:** A significant initial step for video investigation is Background Subtraction and it is utilized to find the objects of enthusiasm for additional prerequisites. Foundation deduction approach is a general technique for movement recognition strategy, which proficiently utilizes the distinction of the current picture and the foundation picture to recognize moving articles. Here the proposed calculation is known as Mixture of Gaussian (MOG) process. This goes under a quality investigation calculation for pictures, which could be handled in the recordings and casings. A methodology is utilized alongside the Kalman channel for outline by outline identification. At that point the MOG is utilized naturally to gauges the quantity of blend parts required to display the pixels foundation shading dissemination. Here executes the foundation concealment for static and dynamic foundation pictures without utilizing any reference foundation pictures, and furthermore smother the clamor out of sight picture's shadows. Kalman channel is a channel that contains strategies portrayed by inferior computational expense and depends on a strong factual model, on a heartiness level. At long last, the fragmented foundation picture is acquired with acceptable execution. At that point the key of this technique is the instatement and update of foundation picture and recognition of moving article, which is likewise exact.

**Keywords:** Background subtraction, motion detection, Kalman filter, Mixture of Gaussian, video analysis.

## I. INTRODUCTION

Computerized picture preparing is appropriately managed control of advanced pictures through an advanced PC. It is an assigned subfield of signs and frameworks yet center especially around pictures. Building up a PC framework that can perform preparing on a picture is the thing that DIP centers around. A particular computerized picture and the framework can process that picture utilizing viable calculations and can give a picture as a yield is the contribution of the framework. Foundation deduction of item is followed by utilizing this Digital Image Processing. [4] Background deduction with Dirichlet process is given the static foundation by utilizing middle and mean channel. Middle channel is a sublevel stockpiling and calculation necessities to the detriment of a more vulnerable adjusting foundation. Here the yield isn't exact with shadow/clamor. The issue is overwhelmed by the Background Subtraction with Mixture of Gaussian models. Two fundamental methodologies rule the writing which are Kalman separating and furthermore Mixture of Gaussians (MoG).

Revised Manuscript Received on July 10, 2020.

\* Correspondence Author

**Bhaggiaraj. S\***, Assistant Professor, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.

**Ranjeeth Kumar. C**, Assistant Educator, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.

**Rahul Vijay. K.S**, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.

**Vignesh Prabhu. A**, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

[15] MoG is increasingly powerful, as it can deal with multi- modular dispersions. Kalman channels can successfully follow a solitary Gaussian and are along these lines considered uni-modular so they can sift through just a leaf or sky, however not both. In MOG, first read a video as image and then it is converted into grayscale. Means and weights are initialized and then frame by frame process is continued with gaussian component for each pixel. The frame is updated with segmented background. In Kalman filter, some algorithm steps are read by same as in MOG. After read and computed number of frames the steps are different. That is difference of current and previous frame/image in individual RGB panel based on the threshold. The shadow/noise is removed by the Kalman filter. The frame is updated with Kalman filter segmented background. [14] The MOG and Kalman filter results are given into the performance analysis and produced better subtracted foreground. It is given the high performance & accurate segmented result with reduced noise and shadow.

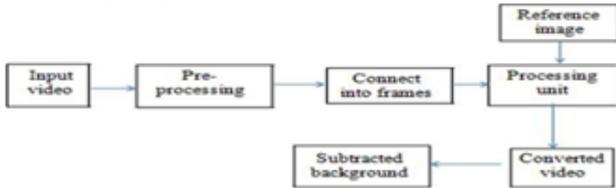
## II. MOTIVATION

The following section provides an introduction of the existing method and the various steps that are involved in the process.

### Basic Video Analysis

Video investigation regularly starts with foundation deduction. A foundation model followed by a regularization plot is the way this issue is drawn nearer. [11] An exact model of the foundation plainly permits it to be recognized based on per-pixel from the closer view, while the regularization can consolidate data from adjoining pixels. [4] Estimating per-pixel foundation conveyances is a technique dependent on Dirichlet process Gaussian blend models. It is trailed by probabilistic regularization. Use of a Bayesian strategy that is non-parametric permits programmed surmising of per-pixel mode tallies, keeping away from over-/under-fitting. Image capturing process is known to cause noise in the Gaussian mixture models. Due to photon noise and varying brightness, it can vary over images. Now and again, for example, low light/warm is ruled. [19] Where a closer view object that looks especially like the foundation is Camouflage, e.g., an a ghillie suit expert marksman. Camera shake will regularly exist, that is an indication of mount focuses that are inclined to wind or vibrations, viewed as a kind of for the most part absorbed worldwide clamor. Moved item (essentially class changes), where the frontal area converges on to the foundation, or the other way around;

e.g., a vehicle could be left in the very scene, and after abundant time thought about piece of the foundation, possibly to later become forefront again when driven off Bootstrapping. Since it is beyond the realm of imagination to expect to get an edge with no closer view, a methodology ought to be fit for being initialised with frontal area objects in the very scene. It needs to then start learning the right foundation model after some time. Shadows can be thrown by the frontal area objects, however later handling is commonly not a piece of them. Brightening changes that are both progressive and comparably quick, for example, from a light switch being flipped on or off.



**Fig.1. Block diagram for background subtraction using reference**

The block diagram is described in the following steps. First, the input video is given to the pre-processing unit. Then the pre-processing result is given to connect into frames. The video is divided frame by frame process. The reference image is given to the processing unit. And same way the frames are given to the processing unit. This result is converted again into video. Finally, the subtracted background image is gotten with static background video.

- Reference background image is found the moving image and given to processing unit. [4] Here median filter and Dirichlet process algorithm is used in existing method.
- It produced the static background by done these processes. Then in this method isn't given the accurate performance of moving object.
- Then the shadow and noise aren't reduced

### Per-Pixel Background Model

Every pixel will have its own multi-modular thickness gauge [15] which is utilized to demonstrate. [10] The Dirichlet procedure (DP) Gaussian blend model is utilized all the while. [4] It is seen as the Dirichlet appropriation which is stretched out to an endless number of parts. It can permit to get familiar with the genuine number of blends that are required for speaking to the information. [11] For every pixel, a flood of qualities show up and one with each edge the model must be persistently refreshed utilizing steady learning.

### Probabilistic Regularization

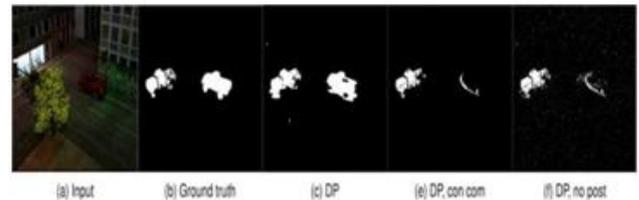
The per-pixel foundation model disregards data from the area of a pixel, leaving it defenseless to locally happening clamor and disguise. [11] Additionally, technique for Gibbs testing presents a lot of clamor which implies that the limit among frontal area and foundation is inclined to a vacillating impact. To kill these issues a Markov irregular field is made, which contains a hub for every pixel, associated utilizing a four-way neighborhood.

### GPU Implementation

Designs Processing Unit of foundation deduction calculation has constrained down to earth use when it can't be run progressively on a video feed as it is gotten since it has chosen an equal structure for the adopted strategy making it to run on

a GPU is direct i.e, Each pixel has an autonomous model, where conviction engendering regularization can be executed through message passing. The accompanying stages are available:

- Count of the likelihood of the current pixel that can be drawn from every blend segment in the model. Likewise ascertain the new segment likelihood. These calculations are completely viewed as autonomous.
- Probabilities are arranged, to make the likelihood of every pixel originating from the current model. Updation of the model utilizing Gibbs inspecting is done where age of irregular number is done through the Philox calculation. This is viewed as free on a for each pixel premise.
- The expenses for conviction spread is determined. This is normally equal.
- Message disregarding is done different various leveled levels. Red-dark update plan is utilized implies that it can refresh a large portion of the pixels at each progression, which is appeared in fig.2.
- Last veil is removed which is an autonomous estimation for every pixel.



**Fig.2 Frame 990 from the noisy night sequence**

The foundation is bound to show up in a scene that is appeared in condition 1 and it can utilize the middle in the past n outlines as the foundation model:

$$B(x, y, t) = \text{median}\{I(x, y, t - i)\}$$

$$\downarrow$$

$$|I(x, y, t) - \text{median}\{I(x, y, t - i)\}| > Th$$

$$i \in \{0, \dots, n - 1\}.$$

In this case the background is the mean of the previous n frames are shown in equation2:

$$B(x, y, t) = \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t - i)$$

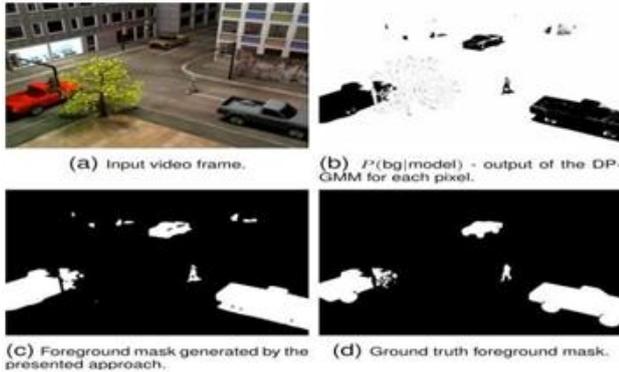
$$\downarrow$$

$$|I(x, y, t) - \frac{1}{n} \sum_{i=0}^{n-1} I(x, y, t - i)| > Th$$

In spite of being engineered it reproduces, decently precisely, nine real true issues, and has the benefit of having ground truth for all edges. The nine certifiable issues are

- Bootstrap has no preparation period such that a perfect foundation plate is never observed.
- Darkening is accomplished for the sun setting reenactment.
- Light switch has the road at evening time and any store turns its light on then off.
- Noisy night is the scene around evening time which contains overwhelming commotion.

- Camouflage is the place the vehicles and individuals have been hued comparatively to the foundation, so they mix in.
- Not having cover which is same as disguise, but at the same time is anything but difficult to see hues purposing for correlation.
- H264 technique packs the video vigorously can create average pressure ancient rarities are appeared in fig.3.



**Fig.3 Casing 545 from the bootstrap grouping of the SABS informational index**

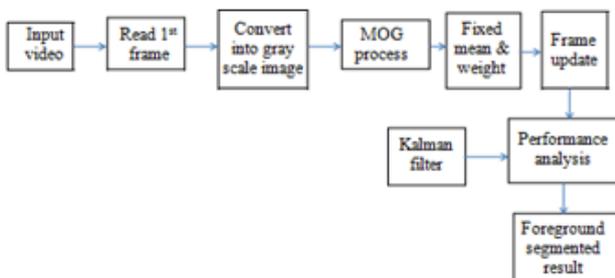
The frame differencing is used for background subtracting. The model:

- $B(x, y, t) = I(x, y, t-1)$
- $|I(x, y, t) - I(x, y, t-1)| > Th$

The escalator used heavily on the right hand side gets a lot of components since the method forms mainly for the people walking past, in case they stall and become a system of the non-foreground.

### III. PROPOSED SYSTEM

The Mixture of Gaussian process and Kalman filter are important algorithm. Static and dynamic background image are used in this method. The straightforward most path for foundation demonstrating is to get a foundation picture, which doesn't have a reference image shown in fig.5.1. The frame by frame is read in MOG process and Kalman filter algorithm, which is given to performance analysis. Finally, it produced the high performance & accurate segmented result with reduced noise and shadow.



**Fig.4 Block diagram for background subtraction with Kalman filter**

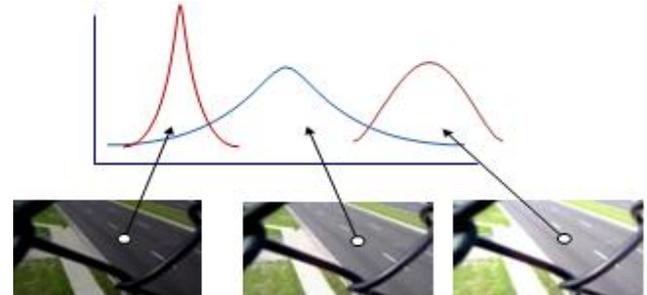
The Kalman filter result and MoG process result is given to the performance analysis. The performance analysis is compared the MOG and Kalman filter results and produced the better foreground segmented output with reduced noise/shadow.

### Mixture of Gaussian (MOG)

Here, each pixel has a blend of Gaussian appropriations model so it is sufficiently basic to deal with any essential non-frontal area varieties.

Two unique assignments are in foundation deducting, both waiting be performed constant having the video outlines as info..

- Background model learning
- Background or closer view pixel order



**Fig.5 In MoG based foundation deduction every pixel is demonstrated with a lot of Gaussian dispersions**

Then the algorithm is given below,

- Read a video
- Read 1st frame as a background
- Convert grayscale and find height and width
- Define MoG variables
- Initialize component means and weights
- Frame by Frame process
- Update Gaussian components for each pixel
- Calculate foreground
- Frame update
- Final foreground segmented results

### Kalman Filter

For refreshing the foundation picture inside any type of video successions, Kalman separating is introduced. There are procedures that can check estimation of uproar contrast. They moreover deal with the issue of pixels that are doused. It is to improve the precision and healthiness of the used count which also has been nicely adequately attempted in a busy time gridlock surveillance work by standing out it from an establishment invigorating figuring, considering Kalman filtering, taken from composing. Kalman channel strategies are portrayed by low computational expense and, being founded on a strong measurable model, by a decent power level. At that point the methodology is given underneath,

- Read a video
- Read all video frames and compute no of frames
- Kalman filter initialize
- Find difference between current and previous frame in individual RGB panel based on threshold
- Remove shadow/noise
- Frame by Frame process
- Find connected components
- Measure properties of image region
- Frame update

- Final foreground segmented results Final step:
- Both methods' step 10 will be updated as final results

### Kalman Filter an Application

States and observations are vectors and  $o \in \mathbb{R}^m$   
Both transition model and measurement model are linear

$$x_{i+1} = A \cdot x_i + \epsilon, \quad o_i = B \cdot x_i + \delta$$

A and B are matrices shown in equation 3, and are process noise and observation noises Assumption: at the first time point Prediction is convolutions of two Gaussians shown in equation 4:

$$p_{i+1}(x_{i+1}) = \int p_i(x_i) \cdot p(x_{i+1}|x_i) dx_i \sim$$

$$\sim \int \exp[-(x_i - \bar{x}_i)^T \Sigma_i^{-1} (x_i - \bar{x}_i)] \cdot$$

$$\cdot \exp[-(x_{i+1} - Ax_i)^T \Sigma_\epsilon^{-1} (x_{i+1} - Ax_i)] dx_i$$

$$p(x_0) = \mathcal{N}(\bar{x}_0, \Sigma_0)$$

→ the result is again a Gaussian [14]  $\mathcal{N}(\bar{x}'_{i+1}, \Sigma'_{i+1})$ .  
The correction is component-wise multiplications of two Gaussians shown in equation 5,

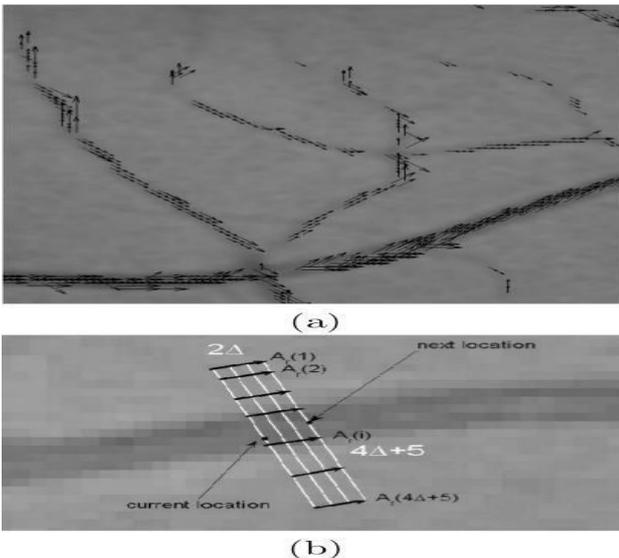
$$p_{i+1}(x_{i+1}|o_{i+1}) = p_{i+1}(x_{i+1}) \cdot p(o_{i+1}|x_{i+1}) \sim$$

$$\sim \exp[-(x_{i+1} - \bar{x}'_{i+1})^T \Sigma'_{i+1}{}^{-1} (x_{i+1} - \bar{x}'_{i+1})] \cdot$$

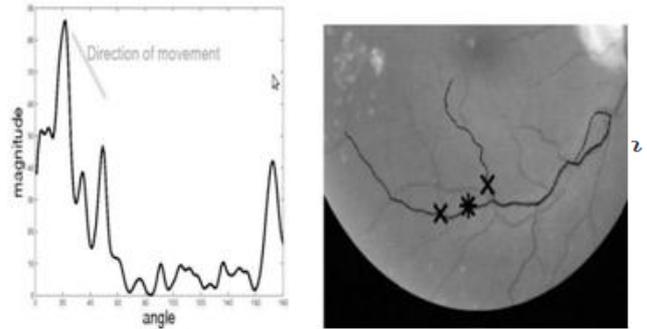
$$\cdot \exp[-(o_{i+1} - Bx_{i+1})^T \Sigma_\delta^{-1} (o_{i+1} - Bx_{i+1})]$$

the result is a Gaussian again  $\mathcal{N}(\bar{x}_{i+1}, \Sigma_{i+1})$ .

It is not necessary to propagate the probability distributions explicitly (i.e. to compute it for all  $\mathcal{F}$ ). Only the parameters need to be re-computed (i.e. the mean and the covariance matrix).



**Fig.6 Tracking object of blood vessels**



**Fig.7 Object moves along the blood vessels**

The tracking object of blood vessels is moved along the blood vessels shown in fig.5.3. Its state is made out of the position, speed, thickness, dim qualities watched so far and so forth., which is appeared in fig.5.4. The veins are followed by this Kalman channel. The guide has point Vs extent of heading of development.

### Sending the Data to the User

The data processed will be stored in a local database and the various forms of data stored in the database can be precariously sent to the designated user in the form of an email. The user can check the email to view the data and analyze the result accordingly. For example, if the user is in search of a particular colored vehicle, the surveillance system will have collected data about all the vehicles that match the description provided by the user and save them in the database. The same can also be mailed to the user if they are cleared to get the data as requested.

## IV. RESULTS AND DISCUSSIONS

This section shows the results and discussion of the proposed method.

### A. Experimental setup

The proposed framework is done in the framework with 8 GB RAM, Intel i7 processor, Windows 10 Operating System and a Nvidia 1060 GPU. This procedure is executed utilizing the product apparatus MATLAB.

### B. Data description

For the experimentation, the background image and input image are captured by the digital camera or web camera on the highway where there is ample vehicle traffic. Also for additional purposes, sample traffic video present in MATLAB directory has also been used.

### C. Experimental results

This section shows the sample results of the proposed method using the traffic video that contains Type 1 vehicles which are cars and Type 2 vehicles which are trucks and buses. The forged images are trained and tested in Fig.3 and Fig.5. Similarly, the original images are trained and tested as in Fig.4 and Fig.6.

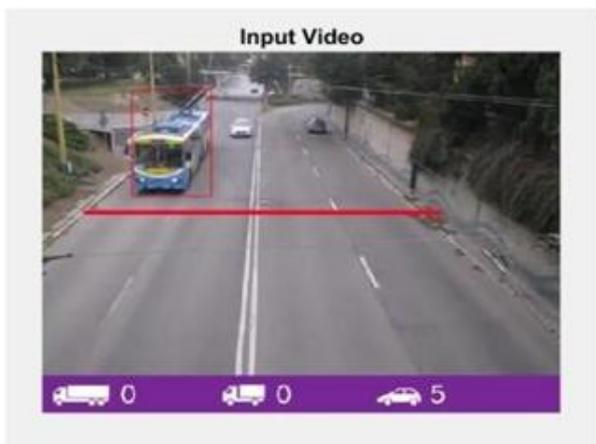


Fig.8 Input video with Type 1 and Type 2 vehicles

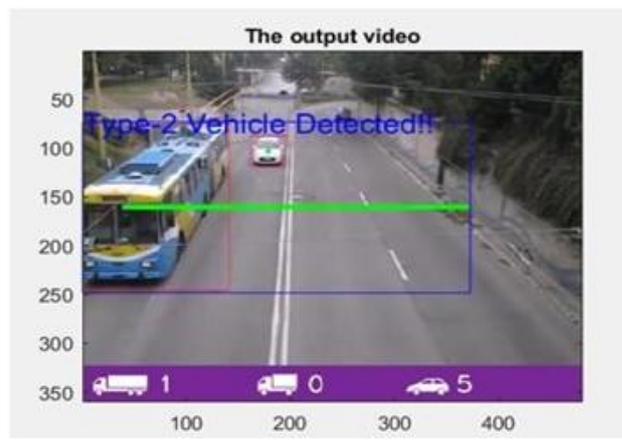


Fig.12 Output video where the type 2 vehicles and type 1 vehicles have been detected

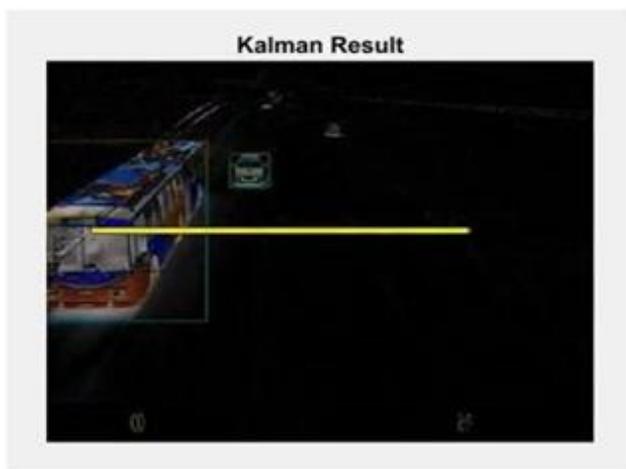


Fig.9 Kalman filter applied on the input video

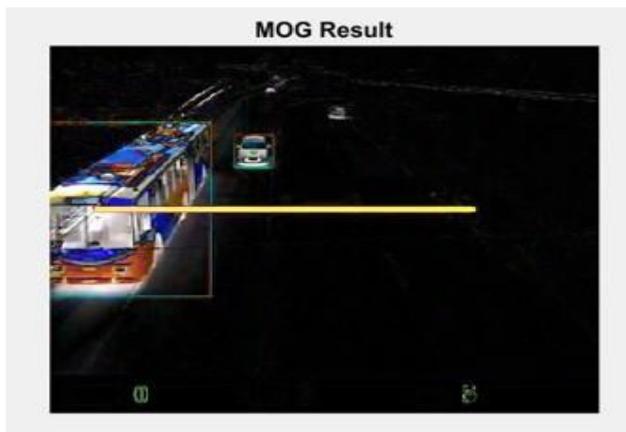


Fig.10 MoG filter applied on the input video

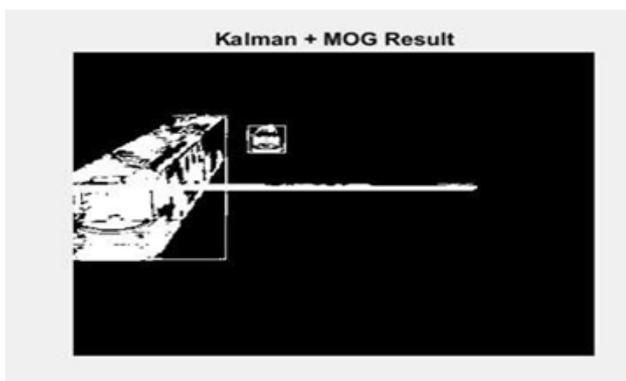


Fig.11 Kalman + MoG filters applied together on the same input video

#### D. Performance Results

The following graphs shows the various performance results the algorithm has produced in MATLAB. These data play a huge role in how well the system will be able to detect the vehicles.

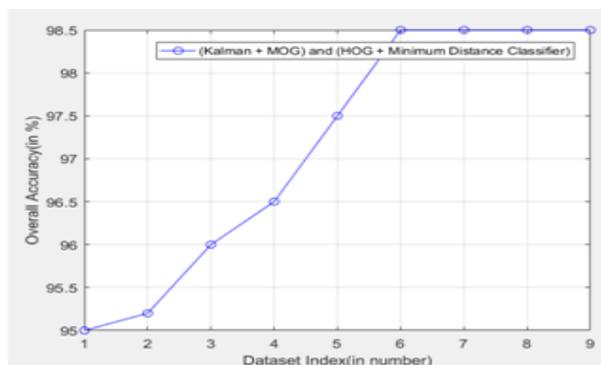


Fig.13 Accuracy Results

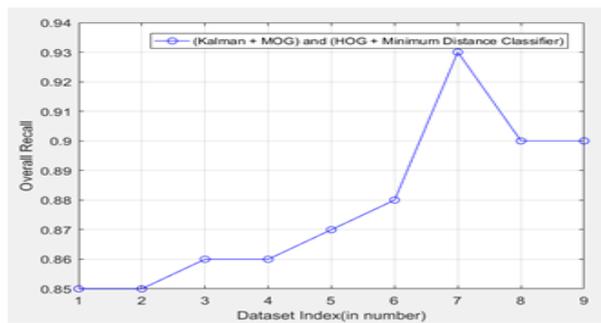


Fig.14 Overall Recall Results

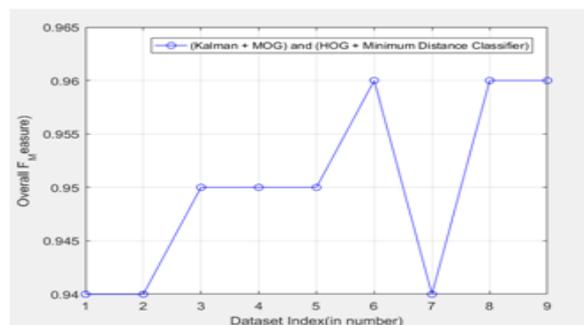


Fig.15. Overall FMeasure Results

The algorithm is fairly effective as it has the capabilities of Kalman + MoG algorithm which also makes use of minimum distance classifier. Thus, from the performance results it is crystal clear that the Kalman + MoG algorithm is considerably effective to similar methods that are around.

## V. CONCLUSION

The background subtraction is dealt by using the Mixture of Gaussian models and Kalman filter. Then the quality of performance is improved. Here the video or moving object is converted into the frame by frame detection. It is not taken the direct video as input. In future, it will be modified as the direct video will be given and gotten the segment value. Here the frame by frame detection isn't needed. So, the system will be given reduced blocks. The process to take time of frame detection will be reduced. And afterward unexpected complex lighting changes are not dealt with, which implies it neglects to deal with some indoor lighting changes. Moreover, an increasingly advanced model of the frontal area and an unequivocal model of left items could additionally improve our technique.

## REFERENCES

1. Z. Luo, F. Branchaud-Charron, C. Lemaire, J. Konrad, S. Li, A. Mishra, A. Achkar, J. Eichel, and P.-M. Jodoin, "MIO-TCD: A new benchmark dataset for vehicle classification and localization," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5129–5141, Oct. 2018.
2. R. Theagarajan, F. Pala, and B. Bhanu, "EDeN: Ensemble of deep networks for vehicle classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2017
3. P. L. St-Charles, G. A. Bilodeau, and R. Bergevin, "SubSENSE: A universal change detection method with local adaptive sensitivity," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 359–373, Jan. 2015.
4. Tom S.F. Haines and Tao Xiang, "Background Subtraction with Dirichlet Process Mixture Models", *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol.36, No.4, April 2014
5. Z. Dong and Y. Jia, "Vehicle type classification using distributions of structural and appearance-based features," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 4321–4324.
6. Y. Peng et al., "Vehicle type classification using data mining techniques," in *The Era of Interactive Media*. New York, NY, USA: Springer, 2013.
7. Yonghong Tian, Yaowei Wang, Zhipeng Hu, and Tiejun Huang. "Selective Eigen background for Background Modelling and Subtraction in Crowded Scenes", *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 23, No. 11, November 2013
8. Florian Seidel Clemens Hage Martin Kleinstueber, "pROST: A Smoothed p-norm Robust Online Subspace Tracking Method for Realtime Background Subtraction in Video", March 29, 2013
9. D.Neelima, D. Santhi Priya "An Extended And Multifeatures Background Subtraction With Svm", *[Ijesat] International Journal Of Engineering Science & Advanced Technology Volume-2, Issue-5, 1316 – 1326, IJESAT | Sep-Oct 2012*
10. Ruben Heras Evangelio, Michael Patzold and Thomas Sikora. "Splitting Gaussians in Mixture Models", *IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*, 2012.
11. M. Hofmann, P. Tiefenbacher, and G. Rigoll, "Background Segmentation with Feedback: The Pixel-Based Adaptive Segmenter", *IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, 38-43, 2012.
12. Lucia Maddalena, Alfredo Petrosino, "The SOBS algorithm: what are the limits?" *IEEE CVPR2012*
13. Xinyi Cui, Junzhou Huang, Shaoting Zhang, and Dimitris N. Metaxas, "Background Subtraction Using Low Rank and Group Sparsity Constraints", Fitzgibbon et al. (Eds.): *ECCV 2012, Part I, LNCS 7572*, pp. 612–625, 2012, Springer-Verlag Berlin Heidelberg 2012.
14. Pushkar Gorur, Bharadwaj Amrutur, "Speeded up Gaussian Mixture Model Algorithm for Background Subtraction", *8th IEEE International Conference on Advanced Video and Signal-Based Surveillance*, 2011".
15. Benjamin Langmann, Seyed E. Ghobadi, Klaus Hartmann, Otmar Loffeld, Pierrot-Deseilligny M., Mallet C., Tournaire O. (Eds), "Multi-modal background subtraction using Gaussian mixture models", *IAPRS*, Vol. XXXVIII, Part 3A – Saint-Mandé, France, September 1-3, 2010.

16. T. Bouwmans, F. El Baf, B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey", *Laboratoire MIA, Université de La Rochelle, Avenue M. Crépeau, 17000 La Rochelle, France, 12 Nov 2008*
17. Massimo Piccardi, "Background subtraction techniques: a review", *The ARC Centre of Excellence for Autonomous Systems (CAS) Faculty of Engineering, UTS, April 15, 2004*
18. J. Yao and J.-M. Odobez, "Multi-layer background subtraction based on color and texture," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
19. P. Wayne Power Johann, A. Schooners, "Understanding Background Mixture Models for Foreground Segmentation", *Proceedings Image and Vision Computing New Zealand 2002, University of Auckland, Auckland, New Zealand 26–28th November 2002.S*
20. Sebastian Brutzer, Benjamin H'oferlin, Gunther Heidemann Intelligent Systems Group, Universität Stuttgart, Germany, "Evaluation of Background Subtraction Techniques for Video Surveillance", *IEEE*.

## AUTHORS' PROFILES



**Bhaggiyaraj. S** is as of now filling in as Assistant Professor (SI. Gr) in the Department of Information Technology at Sri Ramakrishna Engineering College, Coimbatore. He has a four year certification in Information Technology (2004) from Vellore Institute of Technology and a graduate degree in VLSI (2009) from GCT. He has thirteen years of instructing experience. He has distributed such a significant number of papers in diaries and meetings and is at present chipping away at his PhD on Cloud Computing in Anna University.



**Ranjeeth Kumar. C** is as of now functioning as Assistant educator (Sr. Gr) in the branch of Information Technology in Sri Ramakrishna Engineering College. He is right now seeking after Ph.D. degree in Anna University in the area Artificial Intelligence. He acquired his Master of Engineering qualification in Computer Science Engineering from Sri Ramakrishna Engineering school and finished his B Tech in Information Technology from Dr. Mahalingam school of Engineering and Technology, Pollachi. He has finished a consultancy venture for International Aerospace Manufacturing Private Ltd. Bengaluru. He has distributed papers in 8 International diaries, introduced papers in 3 National Conferences. He is a real existence individual from CSI, ISTE, IAENG and IACSIT. He has finished BEC with starter Certification and has NPTEL accreditations.



**Rahul Vijay. K.S** is at present seeking after his B-Tech in Information Technology from Sri Ramakrishna Engineering College in Coimbatore, India. His inclinations incorporate picture handling, Robotic Process Automation and Salesforce.



**Vignesh Prabhu. A** is right now seeking after his B-Tech in Information Technology from Sri Ramakrishna Engineering College in Coimbatore, India. His inclinations incorporate web structuring, Robotic Process Automation and cloud Salesforce.