International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Segmentation of Moving Objects using Numerous Background Subtraction Methods for Surveillance Applications

Supriya Agrawal, Prachi Natu

Abstract**: *Background subtraction is a key part to detect moving objects from the video in computer vision field. It is used to subtract reference frame to every new frame of video scenes. There are wide varieties of background subtraction techniques available in literature to solve real life applications like crowd analysis, human activity tracking system, traffic analysis and many more. Moreover, there were not enough benchmark datasets available which can solve all the challenges of subtraction techniques for object detection. Thus challenges were found in terms of dynamic background, illumination changes, shadow appearance, occlusion and object speed. In this perspective, we have tried to provide exhaustive literature survey on background subtraction techniques for video surveillance applications to solve these challenges in real situations. Additionally, we have surveyed eight benchmark video datasets here namely Wallflower, BMC, PET, IBM, CAVIAR, CD.Net, SABS and RGB-D along with their available ground truth. This study evaluates the performance of five background subtraction methods using performance parameters such as specificity, sensitivity, FNR, PWC and F-Score in order to identify an accurate and efficient method for detecting moving objects in less computational time.***

Keywords**: *Background subtraction, video surveillance, foreground detection, object detection, illumination changes, occlusion, human activity tracking, background modeling, Performance***

I. INTRODUCTION

Smart video surveillance systems have become more popular nowadays for human safety, road traffic control, crime control and many more. It is not practically possible to watch each and every events at every locations using man power so these surveillance cameras plays vital role to detect the events happened with its time and locations without human intervention. In India, Automatic surveillance system through CCTV (closed circuit television) cameras has rapidly increased. According to industry estimates, the security & surveillance market in India is growing 20-25% annually, Industry source estimate was worth Rs 8,200 crore in year 2017, reached Rs 11,000 crore in year 2018 and is expected to touch Rs 20,000 crore in year 2020. Not every security cameras perform better so developing an intelligent security system with less computational cost is a great challenge in surveillance field [1].

Revised Manuscript Received on January 30, 2020.
***** Correspondence Author
Supriya Agrawal*, Department of Computer Engineering, NMIMS University, Mumbai, India.
Dr. Prachi Natu, Department of Computer Engineering, NMIMS University, Mumbai, India.

Video surveillance is the most critical task to analyze the real events for security purpose. The most common applications of analytics of videos include: Motion detection, object tracking, People Counting, Traffic analysis, crowd analysis, gait analysis and many more. Over the years, detecting and localizing the objects in video sequences using object detection algorithms are becoming more popular and emerging research area in the field of computer vision. Object detection is a way to verify the presence of an object like Person, Tree, Pedestrian, Vehicle and more from the images or videos and then classify the objects for correct recognition. According to literature of object detection, there are mainly two types of objects need to identify: 1) Stationary objects and 2) Moving Objects from the video sequences. Each video is divided in to number of frames for extracting the required data for static or moving objects and consecutive frames contain the correlated information about that. So the initial step to identify the interested objects from the video is Background modeling [2][3][4][5].

Background modelling is a way to separate the moving objects called *foreground* from the *background* scene. Many algorithms and methodologies have been proposed focusing on the moving object detection using background modelling for different applications. There are three main methods of background modelling as discussed in figure 1.

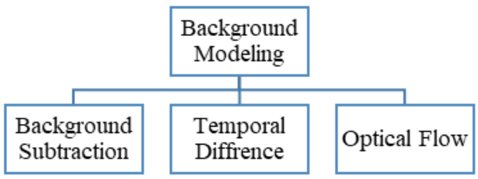


Fig 1. Types of Background Modeling Techniques

1) Background Subtraction: To detect the region of interest (objects) from the scene by subtracting the current (foreground) frame pixel by pixel from a reference (background) frame. Here one threshold value is fixed. If resultant values are less than threshold value it is background and otherwise it is considered as foreground. And finally, update the referenced image for updating the model [4] [5] [6].

2) Temporal Difference: It is also known as frame differencing method. It is most common method to detect the object when camera is moving.

It takes the difference of two consecutive frames but this method fails when two frames have similar regions [4] [5].

3) Optical Flow: It uses flow vectors over time to find out interested regions in the images. Apparent velocity with its direction of very pixel in the frame is calculated. It is effective but time consuming method to detect the objects [7] [8]

The aim of this study is to give comprehensive study of traditional and recent methods involved in background modelling specially in *Background Subtraction* with their limitations and challenges. Many surveyed have been done in literature but none is addressing the overall review of background subtraction methods along with challenges, applications, performance measures and available benchmarks datasets at one place. With this motivation, this paper consists of various sections such as Section 1 explains the need of video surveillance system and basic information of background model. Section 2 describes working of background model and categorize them in to traditional and advanced approaches with their limitations. Section 3 illustrates wide range of application areas of moving object detection. In section 4, we discuss various challenges of object detection using background subtraction methods. Section 5 shows available benchmarks datasets and their limitations. Section 6 and 7 present the performance parameters in selection of background subtraction methods for different datasets for accurate result. Section 8 discusses about the experimental analysis and performance evaluation of background subtraction models. Section 9 and 10 discusses the overall conclusion and future research direction in video surveillance field.

II. BASIC OF BACKGROUND SUBTRACTION TECHNIQUE

Background subtraction method is commonly used for moving object detection using static and moving camera. It is the initial step to separate the foreground image from the background image called referenced image and finally update the background model for moving object detection. The background subtraction method could be represented mathematically as

$$B(i,j) = |CF(i,j) - PF(i,j)| > T \quad (1)$$

where B(i,j) represents object after removal of background, CF(i,j) represents current frame, PF(i,j) represents previous frame and T is threshold value (0-255) which is to be set for detecting objects.

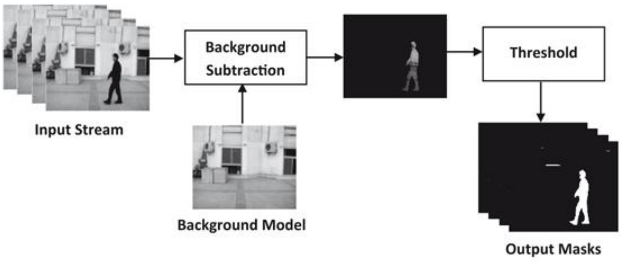


Fig 2. Basic Background Subtraction Model [4]

A. Traditional Methods of Background Subtraction Techniques

Researchers are applying Background subtraction methods for moving object detection from early 90's but still many improvements need to be required for different applications. In this section, we discuss various basic models, statistical models and cluster models of background subtraction techniques.

Basic Models use mean, average filter and histogram methodology for background subtraction from video sequences. Zhang and Ding [9] proposed adaptive background updating model to segment the moving foreground objects from background which continually changes. So we need to update the background scene continuously. They have used morphology operations to avoid noise from video sequences. Pai et al [10] proposed a method for vehicle and pedestrian detection with their walking style by taking the difference between referenced image with current image pixel by pixel and then apply median filter on output image for noise removal. They used ellipse shape contour extraction method for shadow removal. This method fails if pedestrian wear same color of dress matches with background images and found one out of 9 pedestrian is detected as non-pedestrian. Even vehicles with same color of road surface did not pass the entropy limit. Surendra et al [11] proposed a method using adaptive modelling using dynamic update of threshold value to detect vehicles from the monocular images of traffic video sequences. Authors used grey level intensity wise segmentation approach to detect vehicles from video sequences. Vehicles have same intensity as road surface were not correctly classified. Also if relative motion between vehicles is small which leads to wrong counting of vehicles. This method also fails when moving vehicles are overlapped in successive videos scenes. Qin et al [12] proposed a method using self-adaptive background updation with histogram analysis to track the moving vehicles from the video sequences size 320*240. They used twice time morphology operation for noise removal. This method is useful to detect the objects even if high illumination and intensity. The computation time per frame ranges from 0.085 to 0.135s, depending on the image quality and the number of vehicles. Wein et al. [13] focuses on to extract static objects like left behind objects from the video sequences using approximate median filter with ostu threshold method. In this paper, 22 videos of 30 seconds were tested. Average accuracy rate of stationary object detection is 92.8% which could meet real time requirements. This method effectively works on light changes and shadow condition.

Statistical Models for background removal or subtraction are mostly used in complex and cluttered environment. The most famous statistical model is Gaussian model and then many researchers modified the model and named as Gaussian Mixture Model, Mixture of Gaussian, double Gaussian model [4] [14] [15]. In addition, these papers have given a deep study about statistical methods with their challenges and issues. Wren et al. [16] proposed a first statistical model using Gaussian distribution to detect person in the indoor video sequences. It detect the object using blob detection with YUV color space feature.

This method fails with dynamic changes and sudden illumination changes in the background. It is also not able to track multiple human in a scene. In improvement of dynamic background object detection Stauffer and Grimson [17] proposed a multimodel Gaussian distribution where every pixel of the image is processed with mixture of K Gaussian function. The method focuses on two significant parameters – learning rate and proportion of background. Authors found the object detection algorithm worked well on considering k between 3-5 but more or less objects were not clearly identified. In improvement of that Zivkovic [18] proposed an adaptive Gaussian mixture model (AGMM) with consideration of static camera where every pixels are efficiently updated with Gaussian parameter 4 and evaluated the results on dynamic scene. Few authors worked on region based segmentation using statistical model as discussed in [19] [20]. Elgammal et al. [19] proposed a method named as Kernel Density Estimation in which group of pixels are processed instead of single pixel. Authors used normal kernel function with GMM for Density estimation, where each single sample of the N samples is considered to be a Gaussian distribution itself. This method is good to detect the objects in dynamic scene, cluttered and slow motion scene with color information. Picardi [20] gave good comparative analysis of statistical and basic background subtraction techniques in terms of their computational time, speed and cost. More research have been carried out using support vector models for separating the foreground from the background [20] [21] [22] in indoor and outdoor scene.

Clustering Models are also giving new direction to the researchers in the field of background modeling to overcome the problem of statistical methods using GMM. Sometimes Gaussian model is required high computation power and more memory. To limit this problem, Xia et al. [23] considered a pixel as a cluster of 3D feature vector where feature vector can be color space or other multimodal space. Each cluster is assumed to follow Gaussian distribution but sometimes it is not possible in practical applications. This method was implemented on the outdoor dynamic scenes with static and moving camera. A new approach Codebook has also been introduced by Kim et al. [24] with good performance of object detection in dynamic background, illumination changes and complex environment. Other clustering methods have given new insights to the researchers in the field of background modelling [24] [25] [26] [27][28].

Good quality plagiarism software/ tool (Turnitin / iThenticate) will be used to check similarity that would not be more than 20% including reference section. In the case of exclusion of references, it should be less than 5%.

B. *Recent Methods of Background Subtraction Techniques*

Many researchers mixed traditional methods and proposed new methods for trying to achieve the need of accurate moving object detection using background subtraction methods. In this section, the following methods of Background subtraction techniques have been discussed.

Hybrid models combine some traditional models based on pixel and region approach for background subtraction. These models achieve good results in dynamic background and illumination changes in the scene during object detection. Sun et al. [29] combined codebook with Gaussian Mixture Model (GMM) for human detection by finding out the outliers based on mean vectors and covariance matrix. This method works well in shadow removal condition also. Huang et al. [30] proposed pixel based RGB color space with optical flow to detect the objects in dynamic scene.

Non Parametric models are able to detect the object with high sensitivity. These models are quicker to accept the changes in the background subtraction process and provides the correct results in the field of moving object detection. A novel approach for universal background modelling was proposed by Barnich and Droogenbroeck [31] in which random sample based estimation for background modeling has been used for every pixel. This algorithm works only with color pixels which are very sensitive to noise and light changing condition. This noise sensitivity limitation has been resolved by [32].

In this section, wide varieties of background subtraction techniques have been surveyed and categorized them from basic model to non-parametric model under traditional and advanced approaches as shown in Table I.

Table-I: Classification of Background Subtraction Techniques

Approaches	**Background Subtraction Model Categories**	**Background Subtraction Techniques**	
Traditional	Basic Models	Median, Adaptive Median, Approximate Median, Histogram Analysis [9] [10] [11] [12] [13]	
	Statistical Models	Gaussian Distribution	Gaussian Model, Gaussian Mixture Model (GMM),Mixture Of Gaussian (MOG), Kernel Density Estimation (KDE) [14] [15] [16] [17] [18] [19]
		Support Vector	Support Vector Machine (SVM) [20], Support Vector Regression (SVR) [21]
	Cluster Models	K-Means [23] Codebook[24] [25] [26] [27]	
Advanced	Hybrid Model	Sun et al. [29] Huang et al. [30]	
	Non Parametric Model	Video Background Extractor (ViBe) [31] [32]	

III. APPLICATIONS OF BACKGROUND SUBTRACTION METHODS

Background subtraction methods play vital role in the field of background modeling for computer vision applications.

It is the initial step to detect the objects (static or moving) from the video sequences. In past decade, researchers focused on various real applications of background modeling supported background subtraction methods. In this section, various applications of background subtraction methods for moving objects detection have been discussed.

Human Actions Surveillance: In need of making smart surveillance system, background subtraction methods are useful for detecting human actions/activities from the video sequences. Human actions tracking have become an open research area to design automatic surveillance system for safety and security purposes. Automatic pedestrian detection system to helps in counting total number of pedestrian on the road as well as Pedestrian behavior [10][16][29]. In particular of crimes like murder, snatching, kidnapping, fighting, terrorist attack and many more applications are discussed in [33] [34].

Vehicle detection and traffic analysis: Various background subtraction methods models were implemented to detect vehicles on the road according to their shape and speed. These models also help in monitoring total number of vehicles and types of vehicles like car, bike, truck, cycle etc to monitor the traffic in cities [11][12]. Background subtraction also helps in detecting free parking slot as well as illegal parking slot in shopping malls.

Gesture Recognition: Background subtraction methods help in detecting movements of hands and body of the human to monitor the human activities. This model is also useful in prediction of score in sports like Soccer or Hockey [35]. Gesture recognition applications also helpful to detect eyes, face expressions and body positions of human.

Miscellaneous Applications: There are many areas in which background modelling is essential. Many researchers are working on Medical image analysis, virtual reality, aerial view monitoring using drone, animal detection, marine object detection and many more using background modeling**.**

IV. CHALLENGES OF BACKGROUND SUBTRACTION METHODS

Object detection using background subtraction method remains an open research area in computer vision applications. Many methods and their improvements have been proposed but correct detection of objects with high speed and less computational cost is still a great challenge. In particular, six challenges of background subtraction techniques in surveillance area are discussed here.

Illumination Changes

Background model should adapt gradually changes in appearance of the objects in a video. Intensity of the light changes during day and night time. Sudden illumination can also occur indoor and outdoor scene. Weather condition like cloudy, rainy and sunny also effect on appearance of the objects. Sometimes false detection happen due to sudden changes in illumination. Many researchers tried to solve this challenge for correctly detect moving objects [10] [24] [31] [32].

Dynamic Background

In a video scene, some part of the background may contain regular or irregular movement is called dynamic background and handling this challenge is a difficult task. Fountain, water waves, flash of traffic light, waving trees are examples of dynamic background which were solved in [16] [23] [24] [29] [30].

Shadow Appearance

Objects may be detected with their shadow so it is difficult to detect correct object in video sequences. It is also difficult to segment foreground from background with overlapped shadows of the objects. Shadow appearance is a great challenge and many researchers proposed methods to remove shadow from the scene [10] [20] [37].

Occlusion

When one object is hidden behind the other objects is called occlusion. It occurs in real life situation like two people are walking, car is under the bridge or object is appear and disappear again. In this challenge, we have some information about the object on the basis of that entire information will be extracted. Occlusion may be partial (when small part of the object is overlapped with other objects) or full (when entire object is overlapped). Occlusion may affect the computation time of the surveillance system. Researched proposed various method to solve this challenge still lot to be done [10] [40] [41] [42] [43].

Speed of Moving Objects

The speed of the objects for background modeling is an important factor. If speed of moving objects in video sequences is slow then few methods fail to detect the correct objects and if object's speed is too fast then a trail of ghost region remains in the detected background [20] [38].

Clutter

When foreground contains lots of noise and unwanted things so it is difficult to segment the objects from the scene. It is difficult to design a system which is reliable to detect the object from the clutter scene. Many methods were proposed which require lots of improvements [19] [36] [39].

This journal uses double-blind review process, which means that both the reviewer (s) and author (s) identities concealed from the reviewers, and vice versa, throughout the review process. All submitted manuscripts are reviewed by three reviewer one from India and rest two from overseas. There should be proper comments of the reviewers for the purpose of acceptance/ rejection. There should be minimum 01 to 02 week time window for it.

V. DATASETS FOR BACKGROUND SUBTRACTION TECHNIQUES

Even though no single dataset available which covers all the challenges and applications of background modeling. There are varieties of datasets with different applications available for same context [3] [51]. In this section, we try to summarize available video datasets with ground truth and background subtraction techniques for object detection and activity analysis. Ground truth in dataset is used to check the accuracy of the techniques in real world. Here, we discussed 8 video datasets and techniques which used these dataset particularly for moving object detection.

Wallflower

This is the benchmark dataset to detect the moving objects from video sequences. It contain test images which is to be used for detecting the foreground objects. Wallflower dataset is generally used for video surveillance applications. In paper [44], wallflower dataset were used with 7 videos of indoor/outdoor scenes containing 16,158 frames and 10 different BGS methods were evaluated. Ground truth is provided in the form of binary mask to segment foreground and background pixels.

PETS (Performance Evaluation of Tracking and Surveillance) Dataset

PETS 2000 dataset contains outdoor scene to track human and vehicle. It contains 3672 frames for training and 1452 frames for testing. No ground truth is available. PETS 2001 is more significant than PETS 2000. It contains multi view data for tracking vehicle and human. The available ground truth is bounding box in the image. Likewise we have PETS 2006, PETS 2009 to PETS 2017. In PET 2017, total 15043 frames for training and 18221 frames for testing. In paper [45], PETS dataset is available on which various BSG methods were tested.

BMC Dataset

Background modeling challenge (BMC) dataset is designed to extract moving objects as foreground from the background. This benchmark dataset is widely used to evaluate various BGS algorithms with different performance parameters. The ground truth of BMC dataset is pixel label to test the accuracy of the methods. In paper [46], 20 urban video sequences were captured which describes weather conditions like cloudy with noise or without noise, foggy with noise etc. they have tested 6 basics BGS algorithms on captured video of BMC dataset with different performance parameters as discussed in next section.

SABS Dataset

Stuttgart artificial background subtraction (SABS) datasets is used to detect human activities at pixel by pixel label. This dataset was designed for people detection and people counting in indoor/outdoor scene. A method was proposed in [47] with 9 artificial video of outdoor scene with some background subtraction challenges. This method was tested on SABS dataset and found good result when background is static.

CD.Net (Change Detection) Dataset

CD.Net 2012 dataset is used for video surveillance applications where 31 large scale videos (approx. 90,000 manually annotated frames) of different categories are captured with pixel label ground truth [48]. CD.Net 2014 is also for real time video surveillance system containing 21 additional videos (approx. 70,000 manually annotated frames) with pixel label ground truth as discussed in [49].

RBG-D Dataset

This dataset is specially designed for moving object detection with dense depth information. It is used to extract color information during background modeling using depth sensor. Depth sensor generally provides geometric information about the video scene with pixel label ground truth. In paper [50] 4 outdoor scene with depth and color information were used to detect moving object detection.

CAVIAR Dataset

CAVIAR dataset is used for monitoring human behavior in real life. The objective of designing this dataset is to track human activities like walking, fighting, snatching, shopping, entering to and exit from a shop and many more indoor/outdoor activities. In paper [52], there were more than 20 videos captured for tracking pedestrian with ground truth is bounded box. There were two sets of recorded video in which initial 8000 frames are labeled as pedestrian and rest frames for testing.

IBM Dataset

IBM dataset were created by IBM smart surveillance system for human tracking. This dataset contain indoor/outdoor video scenes of human activities. It is specially designed to crowd analysis with ground truth is bounded box. A background subtraction method was proposed [53] and compared with 4 video datasets like IBM, PET, CAVIAR and FUDAN for estimating crowd density. After experiment, it was observed that proposed method worked well on IBM dataset with low density crowd. Different datasets were created for solving different challenges in context of object detection using background subtraction methods as depicted in Table II. In this context, this table has been designed as per the literature survey done by the authors particularly in video surveillance system.

Table-II: Available Benchmark Datasets for Background Subtraction Challenges

Datasets / Challenges	Wallflower	PET	CD.Net	CAVIAR	IBM	SABS	RGB-D	BMC
Illumination Changes	√	√			√	√		√
Dynamic Background	√	√	√	√	√	√	√	√
Shadow Appearance		√	√	√		√	√	√
Occlusion		√		√	√	√	√	
Motion of Objects	√		√					
		√						

In past years, many background subtraction methods were proposed for moving object detection to solve discussed challenges of background subtraction techniques and these methods were also tested on different benchmark datasets to test the accuracy of the surveillance system. In this paper, video dataset contains verities of real or synthetic recorded scenes to solve few challenges of object detection but not all. Most important no such benchmark datasets are available till now which can solve all discussed challenges of moving object detection using background subtraction techniques.

VI. PERFORMANCE MEASURES

Background subtraction plays most important role to segment the foreground region from the background scene in video surveillance system.

Many background subtraction algorithms have been proposed with different conditions and applications. All algorithms have different features performance parameters to evaluate their accuracy and performance. Performance measures are important to identify the correct pixel in segmentation of foreground and background from images/videos. In past years [55] [56], researchers evaluated wide variety of BSG algorithms with various performance parameters but choosing the correct parameter for particular algorithm is a great challenge. Cheung and Kamath [57] evaluated 6 Background subtraction techniques on urban traffic scene with weather condition but BSG methods fail to identify color information because their dataset contains gray images. Toyama et al [28] identified several video surveillance challenges with well-known dataset: wallflower. They considered True Positive and False Positive parameters for evaluation. In papers [20] [58] speed, memory and accuracy are also important factors for evaluation of BGS algorithms. In this study, we compiled many evaluation parameters on the basis of pixel level as suggested in the literature.

A. Parameter Description

The performance measures of Background subtraction algorithms for moving object detection are based on the following parameters:

True Positive (TP): Number of pixels correctly detected as foreground

True Negative (TN): Number of pixels correctly detected as background

False Positive (FP): Number of pixels falsely detected as foreground

False Negative (FN): Number of pixels falsely detected as background

However, these parameters are highly dependent on each other like when FN decreases, FP always increases or vice-versa. Additionally, researchers [3][51][58][59] also worked on some other evaluation parameters to evaluate the performance of algorithms for correct pixels segmentation. These parameters are shown in Table III.

Table- III: Other Performance Parameters for Background Subtraction Techniques

Parameters	Description	Formula
Specificity	Precision provides the information about validity of segmented results.	TP/(TP+FN)
Sensitivity	Recall provides the information about correctly identified the pixels.	TP/(TP+FN)
True Positive Rate (TPR)	This is used to detect correct pixels	1-Specificity
False Negative Rate (FNR)	This is used to detect number of pixels which are missed to detect as foreground	FN/(TP+FN)
Percentage of Wrong Classification (PWC)	To measure incorrectly classified pixels as foreground	((FN+FP)/(TP+FN+FP+TN))*100
F-measure	F-measure is used to evaluate efficiency of the method based on sensitivity and specificity into a single metric	$F-measure = 2\frac{recall \times precision}{recall + precision}$
ROC Curve (Receiver Operating Characteristic) Analysis	The curve is plotted between sensitivity and specificity for different cut-off points to a particular decision threshold	

VII. INFERENCES OF DATASETS AND PERFORMANCE PARAMETERS FOR BACKGROUND SUBTRACTION METHODS

After studying various background subtraction methods along with datasets and performance parameters particularly for video surveillance have been discussed in this section. The following table illustrates the 6 datasets and their performance parameters on which particular background subtraction technique works and inferences have been discussed in last column.

Table-IV: Inferences of Dataset along with Performance Parameters of Background Subtraction Method

Datasets	BGS Methods	Performance Parameters	Inferences
Wallflower	Toyama et al. [44]	FP, FN	Proposed method was compared with 10 other BGS methods and found better with 10509 total number of wrong pixel segmentation for object detection
PET	Young and Ferryman [45]	FNR, FPR, PWC	Performed 7 BGS methods on PET dataset and concluded that wallflower linear prediction filter given better result with FPR 0.278, FNR 0.083 and PWC 0.181. Here computed values are less which shows better performance of method
SABS	Rezaei and Ostadabbas [47]	F-measure, PWC	Here 9 different BGS methods are evaluated on different SABS video datasets with different background subtraction challenges and found that method proposed in [18] gives better result with less wrong classification of pixels. GMM method [17] is also better in dynamic scene here
CD.Net 2012	Goyette et al.[48]	TP, TN, FP, FN, Recall, Specificity, FPR, FNR, F-Measure, PWC, Precision	Here 18 different BGS methods have been tested using CD.Net 2012 and found FPR value lies between 0.33 to 0.66 for shadow removal
BMC	Vacavant et al.[46]	Recall, Precision, F-measure	Evaluated 6 different BGS methods on outdoor video sequences and found F-measure range varies from 0.5 to 0.95

RGB-D	Camplani and Salgado [50]	TP, TN, FP, FN, Recall, Specificity, FNR, F-measure, PWC	Proposed method guarantees good object detection based on the combination of algorithm with classifier

VIII. EXPERIMENTAL SETUP AND RESULT DISCUSSION

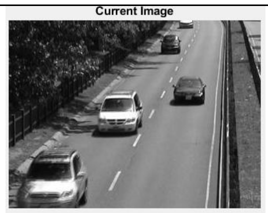
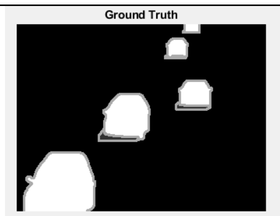
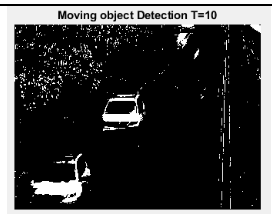
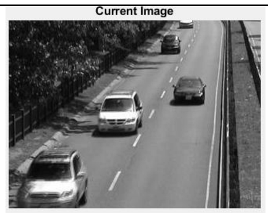
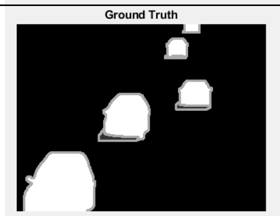
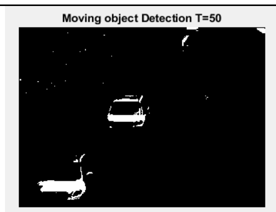
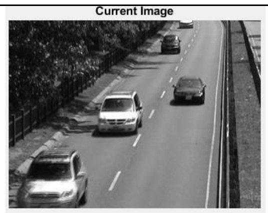
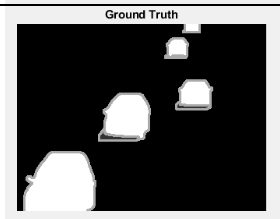
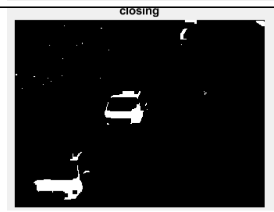
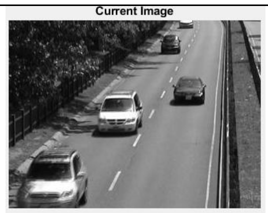
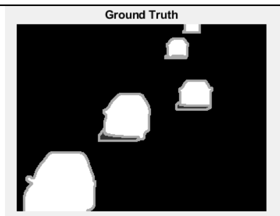
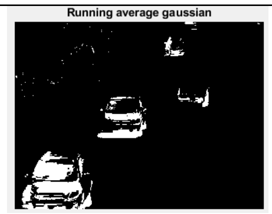
In this study, we compared five methods of background subtraction on CDNet 2012 Baseline(highway) dataset [48] to segment vehicles from the scene. There were 1700 frames of size 320×240 resolution in which first 100 frames were used for background initialization and rest of the frames were used for background updation to detect the object. Ground truth is also available for each frame. Background subtraction techniques which have been compared are (i) frame difference with mean filter (ii) median filter with adaptive thresholding (iii) adaptive median filter with morphology operation (iv) average running Gaussian model (iv) Gaussian mixture model. We used closing morphology operation of rectangle shape with structure element size 5×5. For GMM implementation, we have arbitrarily choosen K= 3 number of Gaussian with learning rate 0.01.

The basic information about the input frame, background initialization image and ground truth image of the dataset as shown below:

(a) (b) (c)

Fig 3. (a) Input frame (b) Background image (c) Ground truth

A. Segmented results of various background subtraction techniques

Sr.No.	*BGS Methods*	*Original Frame*	*Ground Truth*	*Segmented Result*
(I)	Frame difference with mean filter			
(2)	Median Filter with adaptive thresholding			
(3)	Adaptive Median Filter with Morphology Operation			
(4)	Average Running Gaussian			

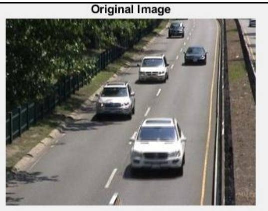
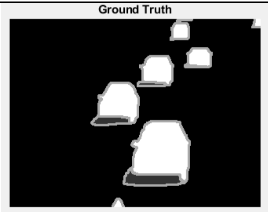
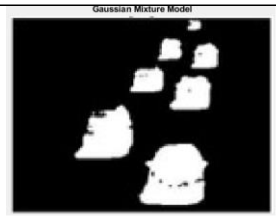
(5)	Gaussian Mixture Model			

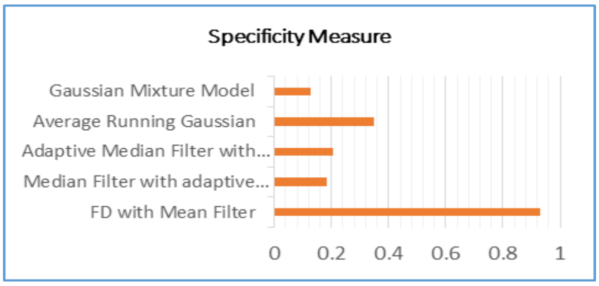
Fig 4.Segmented results from top to down (i) Frame difference with mean filter (ii) Median filter with adaptive thresholding (iii) Adaptive median filter with morphology operation (iv) Average running Gaussian (v) Gaussian mixture Model

B. Performance Analysis

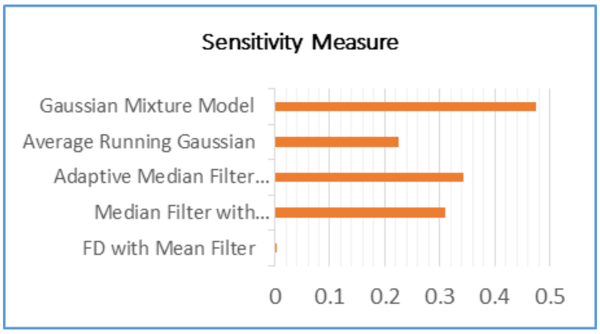
The background subtraction methods were compared against the ground truth given in the dataset for segmenting foreground objects from the scene. The performance measures such as Specificity, Sensitivity, False Negative Rate (FPR), Percentage of Wrong Classification (PWC) and F-score were calculated for background subtraction techniques. The comparative analysis has been shown in Table V

Table-V: Performance Evaluation Metrics for BGS Methods

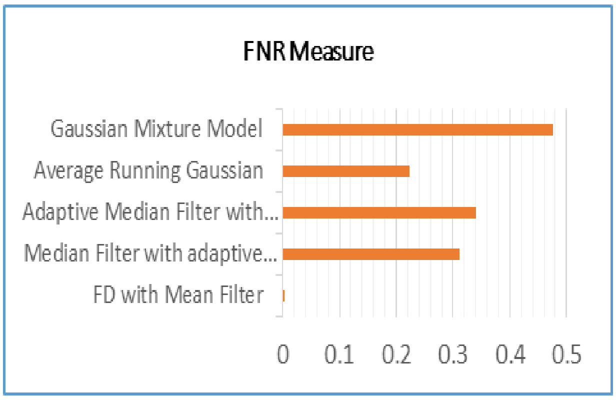
Performance Metrics / Techniques	Specificity	Sensitivity	FNR	PWC	F-Score
FD with Mean Filter	0.0732	0.9289	0.0071	0.1353	0.0015
Median Filter with adaptive thresholding	0.9925	0.1837	0.8163	0.8896	0.3103
Adaptive Median Filter with Morphology operation	0.9934	0.2060	0.7940	0.8926	0.3416
Average Running Gaussian	0.9823	0.3506	0.1685	0.7501	0.2240
Gaussian Mixture Model	0.9994	0.1263	0.1370	0.7010	0.4750



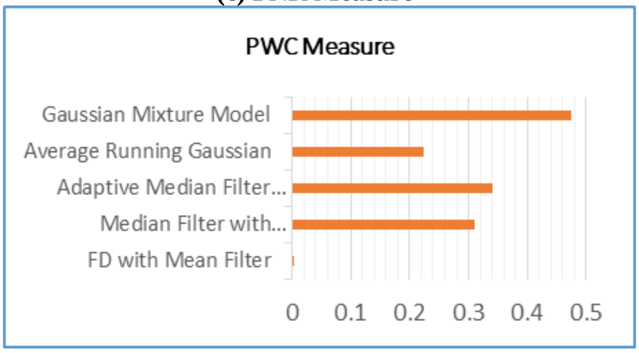
(a) Specificity Measure



(b) Sensitivity Measure



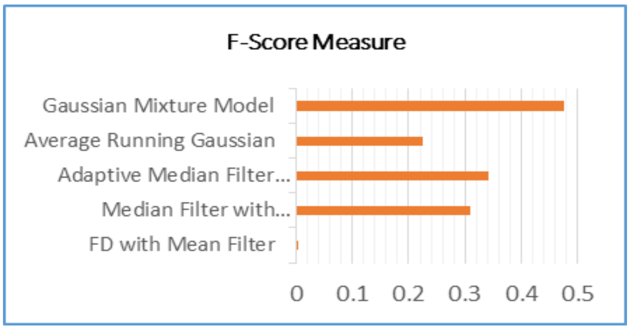
(c) FNR Measure



(d) PWC Measure

Retrieval Number: C8811019320/2020©BEIESP
DOI: 10.35940/ijitee.C8811.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication



(e) F-Score Measure

The specificity of Gaussian Mixture Model (GMM) are 0.9994 which is maximum amongst other methods which means GMM method segments the white pixels from the background to identify the objects correctly. The percentage of wrong classification (PWC) of GMM is 0.7010 which is lesser than other methods such as Median Filter with adaptive thresholding (0.8896), Adaptive Median Filter with Morphology operation (0.8926) and Average Running Gaussian (0.7501). The PWC value of FD using mean filter is 0.1353 and specificity value is 0.0732 due to not classify white pixels more accurately for background subtraction. F-Score of GMM is 0.4750 as compared to other techniques which means GMM continuously updating the background model to segment the objects from the scene.

IX. CONCLUSION

Video surveillance system plays key role in our life for detecting an event for security purpose. For video analysis, background subtraction is initial part to detect the objects from the videos. In this work, we discussed about wide varieties of Background subtraction techniques and categorized them in traditional to advanced approaches. We also studied various applications of background subtraction methods especially for smart surveillance system. Then the most important, challenges of BGS methods were discussed followed by various benchmark datasets with their available ground truth. While reviewing many background models, performance is the most important factor. Hence, we also discussed various performance parameters to evaluate the accuracy of the background subtraction methods. Table 2, 3 and 4 discuss overall inferences of this study where few background subtraction methods along with datasets were discussed with selected performance parameters to evaluate the performance of the methods. This paper shows the performance of five background subtraction methods such as FD using mean filter, Median Filter with adaptive thresholding Adaptive Median Filter with Morphology operation, Average Running Gaussian and Gaussian Mixture Model on CDNet 2012 Baseline(highway) dataset. Performance of background subtraction techniques were evaluated using five parameters like Specificity, Sensitivity, Percentage of wrong classification, False Negative Rate and F-Score. FNR (0.0071) and PWC (0.1353) values of FD using mean filter are less desirable but its specificity value (0.0732) is very less which means method is not suitable in dynamic environment. The specificity of GMM method is 0.9994 which is better than other methods to detect the moving objects correctly by continuously updating the background. Hence, we can say GMM works well for segmenting the objects from background scene but developing a smart surveillance system with less computational cost is still required.

X. FUTURE RESEARCH DIRECTION

There are many open research areas in video surveillance applications that need to be explored such as underwater activity detection, marine life detection, food detection, plastic detection and aerial surveillance using drone, wild life surveillance, and many more. Now a days sports can also be an emerging area where predicting scores with high accuracy by visualizing player actions could be a great help in sports field. Camera positions in surveillance applications face lots of challenges due to static and moving state of the camera. Though this study does not cover camera position conditions but there are variety of background subtraction methods available which were implemented to detect moving objects. Video surveillance systems are not able to recognize the objects which are very much similar and overlapped. This situation is called occlusion where lots of work can be done. We can develop such methods which can detect hidden part of the object on the basis of available information with good accuracy with human intervention. There are lots of methods to solve various challenges in moving object detection but developing a novel approach that fits in various applications and solve many challenges in once will bring new aspect in surveillance field.

REFERENCES

1. Kim, C., Lee, J., Han, T., & Kim, Y. M. (2018). A hybrid framework combining background subtraction and deep neural networks for rapid person detection. Journal of Big Data, 5(1), 22.
2. Zhou, X., Gong, W., Fu, W., & Du, F. (2017, May). Application of deep learning in object detection. In 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) (pp. 631-634). IEEE.
3. Shaikh, S. H., Saeed, K., & Chaki, N. (2014). Moving object detection using background subtraction. In Moving Object Detection Using Background Subtraction (pp. 15-23). Springer, Cham.
4. Bouwmans, T. (2014). Traditional and recent approaches in background modeling for foreground detection: An overview. Computer science review, 11, 31-66.
5. Benezeth, Y.; Jodoin, P.; Emile, B.; Laurent, H.; Rosenberger, C. Comparative study of background subtraction algorithms. J. Electron. Imaging 2010, 19, 033003
6. Xu, Y., Dong, J., Zhang, B., & Xu, D. (2016). Background modeling methods in video analysis: A review and comparative evaluation. CAAI Transactions on Intelligence Technology, 1(1), 43-60.
7. Huang, J., Zou, W., Zhu, J., & Zhu, Z. (2018). Optical Flow Based Real-time Moving Object Detection in Unconstrained Scenes. arXiv preprint arXiv:1807.04890.
8. Zhang, Y., Zheng, J., Zhang, C., & Li, B. (2018). An effective motion object detection method using optical flow estimation under a moving camera. Journal of Visual Communication and Image Representation, 55, 215-228.
9. Zhang, R., & Ding, J. (2012). Object tracking and detecting based on adaptive background subtraction. Procedia Engineering, 29, 1351-1355.
10. Pai, C. J., Tyan, H. R., Liang, Y. M., Liao, H. Y. M., & Chen, S. W. (2004). Pedestrian detection and tracking at crossroads. Pattern Recognition, 37(5), 1025-1034.
11. Surendra Gupte, Osama Masoud, F. K. Martin Robert, and P. Nikolaos, "Detection and Classification of Vehicles," IEEE Transactions on Intelligent Transportation Systems, pp.37-47, March vol.1, num.3, 2002.
12. B. Qin, C. Zhang, Z. Fang, W. Li, "A Quick Self-Adaptive Background Updating Algorithm Based on Moving Region", CVPRIP 2006, 2006.

13. Wei, Q., Yuqiang, Z., Wei, G., & Hailan, L. (2015, September). Research on Stationary Object Detection Technique Based on Dual-Background. In 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC) (pp. 1783-1786). IEEE.
14. T. Bouwmans, F. El Baf, B. Vachon, "Statistical Background Modeling for Foreground Detection: A Survey", Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing, Volume 4, Part 2, Chapter 3, pages 181-199, January 2010.
15. Xu, Y., Dong, J., Zhang, B., & Xu, D. (2016). Background modeling methods in video analysis: A review and comparative evaluation. CAAI Transactions on Intelligence Technology, 1(1), 43-60.
16. Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: Real-time tracking of the human body. IEEE Transactions on pattern analysis and machine intelligence, 19(7), 780-785.
17. C. Stauffer, W. Grimson, "Adaptive background mixture models for real-time tracking", Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999, pages 246-252, 1999
18. Z.Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction." Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.Vol. 2. IEEE, 2004.
19. Elgammal, D. Harwood, L. Davis, "Non-parametric Model for Background Subtraction", 6th European Conference on Computer Vision, ECCV 2000, Dublin, Ireland, June 2000.
20. Piccardi, M. (2004, October). Background subtraction techniques: a review. In 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583) (Vol. 4, pp. 3099-3104). IEEE.
21. H. Lin, T. Liu, J. Chuang, "A probabilistic SVM approach for background scene initialization", ICIP 2002, Rochester, New York, September 2002.
22. H. Lin, T. Liu, J. Chuang, "Learning a Scene Background Model via Classification", IEEE Transactions on Signal Processing, Volume 57, Issue 5, pages 1641-1654, May 2009
23. J. Xia, J. Wu, J. Chen, Z. Cui, "Background Extraction based on the Adaptive Online Clustering", Journal of Computer Engineering, 2011
24. Kim, K., Chalidabhongse, T. H., Harwood, D., & Davis, L. (2004, October). Background modeling and subtraction by codebook construction. In 2004 International Conference on Image Processing, 2004. ICIP'04. (Vol. 5, pp. 3061-3064). IEEE.
25. M. Kumar, S. Venugopal, "Robust Technique for Background Subtraction using K-Means Algorithm", International Journal of Engineering and Science Research, Volume 5, Issue 5, pages 218-222, May 2015
26. S. Naganjaneyulu, N . Priya, "Background Modeling based Universal Multimode Background Subtraction using Pillars K-Means Clustering", International Journal of Research, October 2018.
27. J. Zhang, K Dai, G. Li, "HSV color-space and codebook model based moving objects detection", Systems Engineering and Electronics, Volume 30, Issue 3, pages 423-427, March 2008.
28. Y. Zhang, J. Bae, W. Kim, "Multi-Layer Multi-Feature Background Subtraction Using Codebook Model Framework", International Conference on Digital Image Processing and Vision, ICDIPV 2016, Singapore, 2016
29. Sun, I. T., Hsu, S. C., & Huang, C. L. (2011, October). A hybrid codebook background model for background subtraction. In 2011 IEEE Workshop on Signal Processing Systems (SiPS)(pp. 96-101). IEEE.
30. Lai, J. C., Huang, S. S., & Tseng, C. C. (2010, June). Image-based vehicle tracking and classification on the highway. In The 2010 International Conference on Green Circuits and Systems (pp. 666-670). IEEE.
31. O. Barnich, M. Van Droogenbroeck, "ViBe: a powerful random technique to estimate the background in video sequences", International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009, pages 945-948, April 2009
32. M. Van Droogenbroeck, O. Paquot, "Background Subtraction: Experiments and Improvements for ViBe", IEEE Workshop on Change Detection, CVPR 2012, June 2012.
33. Xu, X., Tang, J., Liu, X., & Zhang, X. (2010, October). Human behavior understanding for video surveillance: recent advance. In 2010 IEEE International Conference on Systems, Man and Cybernetics (pp. 3867-3873). IEEE.
34. Jacques, J. C. S., Braun, A., Soldera, J., Musse, S. R., & Jung, C. R. (2007). Understanding people motion in video sequences using Voronoi diagrams. Pattern Analysis and Applications, 10(4), 321-332.
35. B. Tamas. Detecting and analyzing rowing motion in videos. BME Scientific Student Conference, 2016.
36. Zhang, Z., He, Z., Cao, G., & Cao, W. (2016). Animal detection from highly cluttered natural scenes using spatiotemporal object region proposals and patch verification. IEEE Transactions on Multimedia, 18(10), 2079-2092.
37. Khan, S. H., Bennamoun, M., Sohel, F., & Togneri, R. (2015). Automatic shadow detection and removal from a single image. IEEE transactions on pattern analysis and machine intelligence, 38(3), 431-446.
38. Vondrick, C., Pirsiavash, H., & Torralba, A. (2016). Generating videos with scene dynamics. In Advances In Neural Information Processing Systems (pp. 613-621).
39. Kumar, S., & Yadav, J. S. (2016). Video object extraction and its tracking using background subtraction in complex environments. Perspectives in Science, 8, 317-322.
40. Wang, X., Han, T. X., & Yan, S. (2009, September). An HOG-LBP human detector with partial occlusion handling. In 2009 IEEE 12th international conference on computer vision (pp. 32-39). IEEE.
41. Wu, B., & Nevatia, R. (2006, June). Tracking of multiple, partially occluded humans based on static body part detection. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 1, pp. 951-958). IEEE.
42. Chan, K. C., Ayvaci, A., & Heisele, B. (2015, September). Partially occluded object detection by finding the visible features and parts. In 2015 IEEE International Conference on Image Processing (ICIP) (pp. 2130-2134). IEEE.
43. Sun, D. Partially Occluded Object Detection (Doctoral dissertation, University of Pittsburgh), 2016.
44. K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: Principles and practice of background maintenance," in Proc. 7th IEEE Int. Conf. Comput. Vis., vol. 1. Sep. 1999, pp. 255–261.
45. D. P. Young and J. M. Ferryman, "PETS metrics: Online performance evaluation service," in Proc. IEEE Int. Workshop Perform. Eval. Tracking Syst., Oct. 2005, pp. 317–324.
46. Vacavant, T. Chateau, A. Wilhelm, and L. Lequievre, "A benchmark dataset for outdoor foreground/background extraction," in Proc. 11th ACCV Workshops, 2012, pp. 291–300
47. Rezaei, B., & Ostadabbas, S. Background subtraction via fast robust matrix completion. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1871-1879), 2017.
48. Goyette, N.; Jodoin, P.; Porikli, F.; Konrad, J. Changedetection.net: A New Change Detection Benchmark Dataset. In Proceedings of the 25th IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 1–8. 4.
49. Wang, Y.; Jodoin, P.; Porikli, F.; Konrad, J.; Benezeth, Y.; Ishwar, P. CDnet 2014: An Expanded Change Detection Benchmark Dataset. In Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 393–400.
50. M. Camplani, L. Salgado, Background Foreground segmentation with RGB-D Kinect data: an efficient combination of classifiers, Journal of Visual Communication and Image Representation 25(1), 2014, Pages 122-136.
51. Kalsotra, R., & Arora, S. (2019). A Comprehensive Survey of Video Datasets for Background Subtraction. IEEE Access, 7, 59143-5917
52. Dadi, H. S., Pillutla, G. K. M., & latha Makkena, M. (2017). Human Tracking Under Severe Occlusions. i-Manager's Journal on Software Engineering, 12(1), 29.
53. P. Karpagavalli and A. V. Ramprasad, ''Estimating the density of the people and counting the number of people in a crowd environment for human safety,'' in Proc. Int. Conf. Commun. Signal Process., Melmaruvathur, India, Apr. 2013, pp. 663–667.
54. Negri, P., & Lotito, P. (2012). Pedestrian detection on CAVIAR dataset using a movement feature space. In Proceedings of 13th Argentine Symposium on Technology, AST 2012.
55. Brutzer, S., Höferlin, B., & Heidemann, G. (2011, June). Evaluation of background subtraction techniques for video surveillance. In CVPR 2011 (pp. 1937-1944). IEEE.
56. Liu, L., & Sang, N. (2011, August). Metrics for objective evaluation of background subtraction algorithms. In 2011 Sixth International Conference on Image and Graphics (pp. 562-565). IEEE.
57. S. Cheung and C. Kamath. Robust techniques for background subtraction in urban traffic video. Video communications and image processing, SPIE Electronic Imaging, 5308:881-892, 2004. 1938

58. M. Karaman, L. Goldmann, D. Yu, and T. Sikora. Comparison of static background segmentation methods. In Proc. SPIE, volume 5960, pages 2140-2151. SPIE, 2005. 1938
59. Yu, T., Yang, J., & Lu, W. (2018). Background Modeling with Extracted Dynamic Pixels for Pumping Unit Surveillance. Mathematical Problems in Engineering, 2018.

AUTHORS PROFILE

Supriya Agrawal has earned her Bachelor of Technology in Computer Science & Engineering from Uttar Pradesh Technical University, Lucknow in 2007 and Master of Technology in Computer Science from Banasthali University, Rajasthan in 2013. Currently, she is pursuing Ph.D in Computer Engineering from NMIMS University, Mumbai.
She has 2 years of industrial experience in IT Company. She is currently working as Assistant Professor in Computer Engineering department at Mukesh Patel School of Technology Management & Engineering, NMIMS University, Mumbai, India. Her research interest is Image Processing and Computer Vision. She has published many research papers in various conferences and journals. Ms. Agrawal is a Technical Program Committee member of ICACDS conference, Springer. She is also a member of IAENG society.

Dr. Prachi Natu has completed her B.E. in Electronics in 2004 and Telecommunication and M.E. in 2011 in Computer Engineering from Mumbai University. She has earned her Ph D in Image Processing from NMIMS University, Mumbai.
She has 13 years of experience in academics and currently working as Assistant Professor in Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai. Her area of interest are Image processing, Computer Vision, Machine learning and Deep learning. She has published many research papers in various journals and conferences.