# A Comprehensive Survey on Computer Vision Based Approaches for Moving Object Detection

Sourav Dey Roy and Mrinal Kanti Bhowmik*
Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India
souravdeyroy49@gmail.com and mrinalkantibhowmik@tripurauniv.in*

*Abstract*— **Moving object detection has achieved a noticeable attention in many computer vision applications. The research community have contributed lot of works for dealing with major challenges of moving object detection in real-world scenarios. The paper presents a comprehensive review on different moving object detection techniques classified into four categories: Background Modeling Based techniques; Frame Difference Based techniques; Optical Flow Based techniques and Deep Learning Based techniques. Moreover, detailed descriptions of various methods in each of this category are also provided.**

***Keywords— Computer Vision; Moving Object Detection Techniques; Comprehensive Survey.***

## I. Introduction

During the last few decades, automated video analysis has become a potential research area in computer vision due to its numerous applications to video based intelligent systems. There are three fundamental stages for analysis of video sequences: detection of salient moving objects, tracking of these salient objects on frame basis, and analysis of object tracks to predict the activity or behaviour of this objects. Furthermore, surveillance system have significant role in the defence against criminality and terrorist threats in both public and private sectors. It rely on the ability to detect moving objects in outdoor and indoor scenes which is considered as an efficient step for information extraction in computer vision applications. The term 'object' usually refers to its generalized form, including pedestrians and man-made objects (e.g. vehicles, ships, buildings, etc.) that have sharp boundaries and are independent of background environment.

However moving object/foreground object detection has been a challenging task because of the following reasons [1]: multiple moving objects may present in the scene; small and poorly textured moving objects; poor and rapid change in illumination conditions and shadows and multiple occlusions exist. During the last decades, considerable efforts have been made to develop various methods for the detection of different types of moving objects depicting vehicles and pedestrians in indoor or outdoor scenes objects. While enormous methods exist, a deep review and experimental analysis of the literature concerning generic object detection is still lacking. Depending upon this phenomenon, the paper presents a critical review on existing models for detection of moving objects in indoor or outdoor conditions. In the present survey, we categorized the object detection techniques depending on the approaches they used to detect the salient moving objects. This survey will be especially beneficial for the researchers to have better understanding of this research field and select the most suitable algorithm for their particular needs.

**Paper Outline.** Section II describes the methodological review on different moving object detection techniques based on background modeling. In Section III, methodological review on temporal difference based moving object detection techniques are illustrated. Section IV provides review on optical flow based moving object detection techniques. In Section V, review on deep learning based moving object detection techniques are described. And finally, Section VI concludes the paper.

## II. Methodological Review on Background Modeling Based Techniques for Moving Object Detection

Background modeling and subtraction is one of the simplest method for detecting moving objects in scene appearances. Each of the video sequences consist of background and foreground. If we are able to remove background data from the video sequences, then only necessary data are remained in the foreground, which contains the object of interest [2]. A better accuracy can be achieved if we already have prior information related to the appearance of background. Basically, background subtraction algorithm needs a stable background which is very complicated in real time application. In practical application, static background is not easy to be acquired and background should be updated selectively due to the dynamic change of background image. A considerable number of works have been done on background modeling for moving object detection, i.e. building a proper representation of the background scene. The objective of background modeling is to generate a reference model i.e. generated based on running average express or temporal average. After generating the background model, the video frames are subtracted from a reference or background model. Then the pixel informations in the current incoming frame that is different from the background model are considered to be the moving object. Several researchers make use of the concept of background modeling technique for moving object detection. The summarized description of these papers are shown in TABLE I. In [3], M.F. Savas et.al. proposed a block based adaptive method for detection of moving objects in dynamic environment conditions. To set an adaptive threshold parameter, they have used counter structure as Casares et al. [4]. This parameter was used for foreground determination there by background updation. During updation of background model, each 2×2 non-overlapped block is continually involved and normalization process was performed to equal the area under the model's probability function. In [5], R. Kalsotra et.al. proposed a morphological based approach combined with background subtraction technique and thresholding for moving object detection. In [6], Y. Zhou et.al. proposed an approach using a monucular visual odometry with background subtraction and motion compensation for detection of moving objects. They have used visual odometry for motion compensate within the image. After background subtraction, the detections are projected to a global coordinate frames and used GM-PHD filter [7] for tracking. In [8], J.M. Guo et.al. proposed a multilayer codebook-based background subtraction (MCBS) model for moving object detection in video sequences. The proposed method can cope up with background dynamics and further pixel-based classification is done for enhancing the results from the block-based background subtraction. In [9], J. Ye et.al. proposed a background subtraction method in RGB

The first author is grateful to Council of Scientific & Industrial Research (CSIR), Government of India for providing the Senior Research Fellowship (SRF) under CSIR-SRF Fellowship Programme with Grant No: 09/714(0020)/2019-EMR-I, Dated: 01/04/2019.
978-1-7281-7366-5/20/$31.00 ©2020 IEEE



Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:56:09 UTC from IEEE Xplore. Restrictions apply.

TABLE I

METHODOLOGICAL REVIEW ON BACKGROUND MODELING BASED MOVING OBJECT DETECTION TECHNIQUES

| Author/ Year | Method Used | Dataset/ Type of Image Sequences | Accuracy/ Result | Observation |
|---|---|---|---|---|
| M.F. Savas et.al. [3]/ 2018 | Background subtraction with block-based adaptive threshold parameter. | Wallflower and CD.net 2014 datasets | F-Measure: 0.8551 | The method reduces frame processing speed and provides noticeable results at gray levels. |
| R. Kalsotra et.al. [5]/ 2017 | Background subtraction with Adaptive thresholding and Morphological processing. | Videos of MATLAB and own captured videos. | Not Provided | The proposed method is robust to noise and dynamic scenes but fails to cope up with shadow and camouflage effect. |
| Y. Zhou et.al. [6]/ 2017 | Background subtraction with Monocular Visual Odometry (VO) and motion compensation to detect moving objects. | Simulated video of a city and real outdoor videos. | Recall: 0.8290 Precision: 0.5675 | The method is robust to videos acquired by a motion camera in pronounced parallax. |
| J.M. Guo et.al. [8]/ 2013 | Multilayer codebook-based background subtraction (MCBS). | Campus 200, Water surface 480, Meeting room 1755, Indoor gttest1 342, and Intelligent Room 82 | F-Measure: 0.9265 | The proposed method can significantly handle the dynamic variation of scenes. |
| J. Ye et.al. [9]/ 2012 | Background modeling using Robust estimators and a fast test is utilized to detect foreground pixels in the evaluation stage. | Highway I and Campus | Shadow Detection Rate: 84.15% | The method shows noticeable results for shadow problems. |
| B. Karasulu et.al. [10]/ 2012 | Background subtraction based on entropy values and Simulated Annealing (SA). | CAVAIR dataset | F-Measure: 0.6325 | The method is adaptive to dynamic variation of occlusions. |

color space where a metrically trimmed mean is adopted to estimate the background model and mean absolute deviation is performed as a scale estimate. And finally, cascading chromaticity difference estimator, brightness difference estimator, and spatial analysis are used to identify the actual moving objects irrespective of shadows. In [10], B. Karasulu et.al. proposed a moving object detection by adopting the frame differencing method and tracking is also performed by learning an updated background model. Also, simulated annealing optimization technique is used with background subtraction method to estimate the optimal threshold value for bi-level segmentation (i.e. foreground-background pixels).

## III. METHODOLOGICAL REVIEW ON FRAME DIFFERENCE BASED TECHNIQUES FOR MOVING OBJECT DETECTION

Temporal or Frame differencing approaches are basically used to detect the presence of moving objects on a static background scenarios. Because objects are moving on a frame by frame basis with respect to time, the location of the object in two consecutive frames are different. Therefore, calculating the pixel difference between the two consecutive incoming frames, one can determine the exact position/ location of the object on that particular frame. After the concept of temporal/ frame differencing approach was proposed, the techniques for moving object detection by utilizing frame differencing developed promptly. The brief summarization of these papers are shown in TABLE II. In [11], Z. Xu et.al. used three frame differencing in combination with Gaussian Mixture Model (GMM) for background subtraction. In order to increase the robustness of the proposed method, depending on the three frame difference method, they have used XOR and OR operations as an alternative of AND operation to detect the foreground pixels. Moreover, GMM is used to extract the whole foreground from the image including shadows. Finally, two generated foreground masks are combined by using AND operation to acquire a comprehensive foreground mask. In [12], Y.Wang et.al. proposed a pixel based non-parametric method which contains both spatial and temporal features for moving object detection. In this method, background model is generated by using the first n number of frames and sampling them into m number of times there by considering 3×3 neighbourhood blocks. Moreover, a new background update strategy is proposed that enables the generated background model to fit on the dynamic scenes. In [13], J. Guo et.al. combines the idea of improved frame-difference and Gaussian mixture background subtraction for moving object detection. For accurate and reliable detection of moving objects they have used image repair and morphological operations in the proposed method. In [14], J. Zhang et.al. proposed a hybrid method i.e., using non-parametric background model and frame difference for traceability video analysis. Basically they have used hybrid method to approximate the pixel difference of two incoming frames and predicts whether an object is movable. Thereafter the frames that predicts the presence of changes are selected out as a foreground mask. In [15], S.S. Senga et.al. proposed a technique for moving object detection in static background conditions. In this method, first pre-processing is done for noise removal and calculate the difference between the pre and post consecutive frames of the currently frames. After that the algorithm chooses the maximum pixel intensity values between both the difference frames and divide the calculated difference into several number of non-overlapping blocks and computes the average and mean of pixel intensities of each consecutive block. Consequently it then detects the foreground and background pixels from each considered block and post processing using morphological operation are used to detect the presence of objects. The method has been applied on video datasets comprising real world scenarios and shown satisfactory results for objects with deformable sizes and numbers.

## IV. METHODOLOGICAL REVIEW ON OPTICAL FLOW BASED TECHNIQUES FOR MOVING OBJECT DETECTION

Optical flow represents the pattern of deceptive motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene [16]. It approximates a velocity for each pixels in an image and determines where this pixel is located in the next image. The key theory behind this approach is that the motion of moving objects retains changes in the magnitude of pixel intensities and are significant indication for allocating the moving objects. Thus these methods involves approximating the optical flow field and performs clustering processing according to the optical flow distribution characteristics of image. Some optical flow based moving object detection algorithms have been proposed in the literature. The description of these papers are shown in TABLE III. In [16], J. Huang et.al.

TABLE II

METHODOLOGICAL REVIEW ON FRAME DIFFERENCE BASED MOVING OBJECT DETECTION TECHNIQUES

| Author/ Year | Method Used | Dataset/ Type of Image Sequences | Accuracy/ Result | Observation |
|---|---|---|---|---|
| Z. Xu et.al. [11]/ 2018 | Three frame difference and mixed Gaussian model method. | Real time surveillance and pedestrian videos | Not Provided | The proposed algorithm is adaptive to shadow and noise. |
| Y. Wang et.al. [12]/ 2017 | Pixel wise non-parametric spatiotemporal moving object detection method | CD .net 2014 dataset | F-Measure: 0.8881 | The proposed method is robust to dynamic background and ghost artifacts. |
| J. Guo et.al. [13]/ 2017 | Improved frame differencing, Gaussian mixture background subtraction and Morphological post processing. | CAVAIR dataset | Not Provided | The proposed method is able to handle noise and fill cavities/ holes in the detected objects. |
| J. Zhang et.al. [14]/ 2016 | Hybrid non-parametric frame difference method to approximate the pixel difference between consecutive frames. | Outdoor videos of vehicles and pedestrians | Not Provided | The proposed method can effectively remove noise from the scenes. |
| S. S. Sengar et.al. [15]/ 2016 | Non-overlapping blocks based improved three frame difference approach. | Complex indoor and outdoor sequences | Detection Error Rate: 10.36% | The proposed method achieves better local performance. |



Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:56:09 UTC from IEEE Xplore. Restrictions apply.

TABLE III

METHODOLOGICAL REVIEW ON OPTICAL FLOW BASED MOVING OBJECT DETECTION TECHNIQUES

| Author/ Year | Method Used | Dataset/ Type of Image Sequences | Accuracy/ Result | Observation |
|---|---|---|---|---|
| J. Huang et.al. [16] / 2018 | Homography matrixes are used for background modeling in the form of optical flow. | Real time surveillance and pedestrian videos | F-Measure: 0.747 | The method can deal with dynamic background but is sensitive to shadow. |
| L. Kurnianggoro et.al. [17]/ 2016 | RANSAC homography method. | Dataset taken from [18] | F-Measure: 0.510 | The method can detect moving objects from the moving camera. |
| S. Kamate et.al. [18]/ 2015 | Detection of moving objects using adaptive background subtraction and meanwhile Lucas Kanade method based tracking is performed. | Real time outdoor videos of aerial vehicles | Not Provided | The performance of the algorithm is noticeable when applied to static camera views. |
| K.P. Risha et.al. [20]/ 2015 | Moving objects are detected by utilizing Lucas Kanade method along with morphological based post processing. | Real time outdoor videos of vehicles | Not Provided | The proposed method is capable for detection of moving objects in static background conditions. |
| X. Li et.al. [21]/ 2015 | TV-L optical flow with SLIC super pixel segmentation. | Freiburg-Berkeley Motion Segmentation Dataset (FBMS-59) | Center local error: 5.87 | The proposed method are able to detect moving objects in dynamic background conditions. |
| J. Hariyono et.al. [22]/ 2014 | Optical flows are obtained using Affine transformation and finally recognized based on histogram of oriented gradients (HOG). | Real time outdoor videos | Detection Rate: 0.988 | The proposed method can deal well detect pedestrians from moving vehicles in the scenes. |

modelled the background using homography matrixes in the form of optical flow. For detecting the foreground moving pixels, a dual mode judging mechanism has been purposed to improve the performance of the proposed method in challenging situations of real world. In [17], L. Kurnianggoro et.al. modeled the background using zero optical flow vectors. In this method, the previous frames are aligned using homography matrix and dense optical flow was estimated between result of aligning and the current frame. Finally a simple optical-flow magnitude threshold was utilized to determine the foreground points. In [18], S. Kamate et.al. used Unmanned Aerial Vehicles (UAVs) dataset for moving objects detection and there tracking. In this method, moving salient objects are detected using adaptive background subtraction technique and meanwhile Lucas Kanade optical flow tracking [19] algorithm is used for tracking these detected objects. In [20], K.P. Risha et.al. proposed a hybrid method for moving target detection utilizing a method of optic flow along with morphological operations. In [21], X. Li et.al. proposed a technique for moving object detection using TV-L optical flow and SLIC (Simple Linear Iterative Clustering) segmentation. At first, SLIC is used to segment the current incoming frame and the flow vectors are determined using TV-L method. Then optical flow gradients are calculated from the boundaries of segmented super pixels and the pixels with the larger optical flow gradients are selected by empirically setting a threshold value. In [22], J. Hariyono et.al. proposed a pedestrian detection method from moving vehicles based on optical flow and histogram of gradients (HOG). Finally moving regions are detected as transformed objects and morphological post processing is done to acquire the reliable human regions. Moreover, HOG features are extracted from the objects and classified using linear support vector machine.

## V. METHODOLOGICAL REVIEW ON DEEP LEARNING BASED TECHNIQUES FOR MOVING OBJECT DETECTION

Deep Learning based approaches has modernized the computing landscape leading to a fundamental change in how applications are being created in various aspects of computer vision. One such field that has been affected by Deep Learning in a significant way is object detection [23]. After the concept of deep learning was proposed, the moving object detection techniques by utilizing the concept of CNNs developed rapidly as shown in TABLE IV. In [24], B. Heo et.al. proposed a novel method for moving object detection in dynamic background conditions using CNN. The proposed method composed of two deep learning networks entitled as appearance network (A-Net) and motion network (M-Net). The purpose of A-Net is to detect the appearance of the moving objects and M-net detects its motion. Finally, the two networks are combined to detect moving objects. In [25],Y. Chen et.al. proposed a deep sequence learning architecture to detect the moving objects. This method first extract pixel-wise semantic features through deep encoder-decoder network. Then to exploit the pixel wise changes, a novel attention long short-term memory (Attention ConvLSTM) has been proposed. Finally, spatial transformer network (STN) and a conditional random field (CRF) layer are combined to smooth the foreground boundaries. In [26], E.D. Tejada et.al. used Principal Component Pursuit (RCP) as a pre-processing step for video background modeling before Faster-RCNN model as proposed by S. Ren et.al. [27] so as to improve the detection and classification of moving objects. In [28], M. Babaee et.al. proposed a background subtraction method using deep CNN for moving object segmentation. Additionally, a new approach for background modeling is also proposed and finally post-processing using spatial-median filtering of the network outputs is done. In [29], T.N. Le et.al. proposed a framework based on spatio-temporal features for object detection. In this method, a novel SpatioTemporal Deep (STD) features has been proposed consisting of local and global features. The local features are extracted using a region-based Convolutional Neural Network (CNN) [30] whereas the global features are extracted using a block-based CNN [31] from temporal segments. Also a SpatioTemporal CRF (STCRF) is introduced for approximating the spatial relationship between frame regions. In [32], D. Zeng et.al. proposed a novel multiscale fully

TABLE IV

METHODOLOGICAL REVIEW ON DEEP LEARNING BASED MOVING OBJECT DETECTION TECHNIQUES

| Author/ Year | Method Used | Dataset/ Type of Image Sequences | Accuracy/ Result | Observation |
|---|---|---|---|---|
| B. Heo et.al. [24]/ 2017 | VGG Architecture based two Networks: Appearance network (A-net) that detects movable objects and motion network (M-net) that distinguishes the foreground objects. | Cycle, Fence, Campus 1, Campus 2, and Daimler | F-Measure: 0.738 | The proposed method effectively copes with the background contamination frequently. |
| Y. Chen et.al. [25]/ 2018 | Attention ConvLSTM for modeling pixel-wise transition over time, STN model and a CRF layer are integrated for motion and spatial smoothness. | CD.net 2014 dataset | F-Measure: 0.935 | The method effectively handle the spatio-temporal changes by learning high-level CNN features. |
| E.D. Tejada et.al. [26]/ 2017 | Robust PCA as a pre-processing step is used prior to the Faster R-CNN model. | CD.net 2014 dataset | F-Measure: 0.6539 | The proposed method increase classification accuracy for the dynamic objects. |
| M. Babaee et.al. [28]/ 2017 | CNN for extracting significant features from a image-background pair and median filtering based post-processing is done on the network output. | CD.net 2012 and CD.net 2014 dataset | F-Measure: 0.7548 | The proposed method increase the performance in static background condition. |
| T.N. Le et.al. [29]/ 2017 | Spatio Temporal Deep (STD) feature that utilizes local and global contexts over frames, spatio temporal Conditional Random Field (STCRF) to compute saliency from STD features. | 10-Clips dataset, SegTrack2 dataset and DAVIS dataset | F-Measure: 0.8858 (F-Adap and F-Max) | The proposed method can handle complex background and provides reliable saliency assignment. |
| D. Zeng et.al. [32]/ 2018 | Multiscale fully convolutional network (MFCN) architecture for background subtraction. | CD.net 2014 and SBM-RGBD dataset | F-Measure: 0.9814 | The proposed method could reduce the detection error in both static and dynamic variation of scenes. |
| M. Braham et.al. [33]/ 2016 | Background subtraction using spatial features learned with convolutional neural networks. | CD.net 2014 dataset | F-Measure: 0.9046 | The proposed method can handle dynamic variation of background. |



Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:56:09 UTC from IEEE Xplore. Restrictions apply.

convolutional network (MFCN) architecture for background subtraction. They have shown that the accuracy of foreground detection can be enhanced by utilizing the deep level features from MFCN. In [33], M. Braham et.al. proposed a background subtraction algorithm based on spatial features learned with CNN. The algorithm utilizes a background model decreased to a scene-specific background image that enables to train how to subtract the background from an input image patch.

## VI. CONCLUDING REMARKS

In this paper, the review of different moving object detection techniques proposed in the literature has been presented. Studies specifies that although approaches based on temporal/ frame difference provides noticeable results in a static environment but the performance usually depends on the similarity in visual scenes and also the speeds of the moving objects. On the otherhand, the approaches based on background modeling and optical flow are mostly affected by the noise and poor illumination due to atmospheric phenomenon. Moreover, background modeling approaches can almost provide the complete object information only if the background is known. Conversely, deep learning based approaches have sparked off thunder in the domain of object detection and have provided good performances in many complex situations and challenges but the models are computationally expensive as large amount of data are needed to train the models.

## ACKNOWLEDGMENT

The work presented here is being conducted in the Computer Vision Laboratory of Computer Science and Engineering Department, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India.

## REFERENCES

[1] X. Ji, W. Zhiqiang and Y. Feng, "Effective vehicle detection technique for traffic surveillance systems," Journal of Visual Communication and Image Representation, Elsevier, Vol. 17, No. 3, pp. 647-658, 2006.

[2] A. Yilmaz, O. Javed and M. Shah, "Object tracking: A survey," ACM computing surveys (CSUR), Vol. 38, No. 4, pp. 13, 2006.

[3] M.F. Savaş, H. Demirel and B. Erkal, "Moving object detection using an adaptive background subtraction method based on block-based structure in dynamic scene," Optik, Elsevier, Vol. 168, pp. 605-618, 2018.

[4] M. Casares, S. Velipasalar and A. Pinto, "Light-weight salient foreground detection for embedded smart cameras," Jounal of Computer Vision and Image Understanding, Vol. 114, pp. 1223–1237, 2010.

[5] R. Kalsotra and S. Arora, "Morphological based moving object detection with background subtraction method," 4th International Conference on Signal Processing, Computing and Control (ISPCC), IEEE, pp. 305-310, 2017.

[6] Y. Zhou and S. Maskell, "Moving object detection using background subtraction for a moving camera with pronounced parallax,". In Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE, pp. 1-6, 2017.

[7] B.N. Vo and W.K. Ma, "The gaussian mixture probability hypothesis density filter," IEEE Transactions on Signal Processing, Vol. 54,N. 11, pp. 4091–4104, 2006.

[8] .M. Guo, C.H. Hsia, Y.F. Liu, M.H. Shih, C.H. Chang and J.Y. Wu, "Fast background subtraction based on a multilayer codebook model for moving object detection," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 23, No. 10, pp. 1809-1821, 2013.

[9] J. Ye, T. Gao and J. Zhang, "Moving object detection with background subtraction and shadow removal," 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, IEEE, pp. 1859-1863, 2012.

[10] B. Karasulu and S. Korukoglu, " Moving object detection and tracking by using annealed background subtraction method in videos: Performance optimization," Expert Systems with Applications, Elsevier, Vol. 39, No. 1, pp. 33-43, 2012.

[11] Z. Xu, D. Zhang and L. Du, "Moving Object Detection Based on Improved Three Frame Difference and Background Subtraction," 2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), IEEE, pp. 79-82, 2017.

[12] Y. Yang, Q. Zhang, P. Wang, X. Hu, and N. Wu, "Moving object detection for dynamic background scenes based on spatiotemporal model," Advances in Multimedia, Hindwai, 2017.

[13] J. Guo, J. Wang, R. Bai, Y. Zhang and Y. Li, "A New Moving Object Detection Method Based on Frame-difference and Background Subtraction," IOP Conference Series: Materials Science and Engineering, IOP Publishing, Vol. 242, No. 1, pp. 012115, 2017.

[14] J. Zhang, J. Cao and B. Mao, "Moving object detection based on non-parametric methods and frame difference for traceability video analysis," Procedia Computer Science, Elsevier, Vol. 91, pp. 995-1000, 2016.

[15] S.S. Sengar and S. Mukhopadhyay, "A novel method for moving object detection based on block based frame differencing," 2016 3rd International Conference on Recent Advances in Information Technology (RAIT), IEEE, pp. 467-472, 2016.

[16] J. Huang, W. Zou, J. Zhu and Z. Zhu, "Optical Flow Based Real-time Moving Object Detection in Unconstrained Scenes," arXiv preprint arXiv:1807.04890, 2018.

[17] J. Son, I. Jung, K. Park and B. Han, "Tracking-by-segmentation with online gradient boosting decision tree," IEEE International Conference on Computer Vision, pp. 3056-3064, 2015.

[18] S. Kamate and N. Yilmazer, "Application of object detection and tracking techniques for unmanned aerial vehicles," Procedia Computer Science, Elsevier, Vol. 61, pp. 436-441, 2015.

[19] C.M. Huang, Y.R. Chen and L.C. Fu, "Real-time object detection and tracking on a moving camera platform," 2009 ICCAS-SICE, pp. 717-722, IEEE, 2009.

[20] K.P. Risha and A.C.Kumar, "Novel method of detecting moving object in video," Procedia Technology, Elsevier, Vol. 24, pp. 1055-1060, 2016.

[21] X. Li and C. Xu, "Moving object detection in dynamic scenes based on optical flow and superpixels," 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, pp. 84-89, 2015.

[22] J. Hariyono, V.D. Hoang and K.H. Jo. "Moving object localization using optical flow for pedestrian detection from a moving vehicle," The Scientific World Journal 2014, 2014.

[23] K. Fragkiadaki, P. Arbelaez, P. Felsen and J. Malik, "Learning to segment moving objects in videos," IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 4083-4090, 2015.

[24] B. Heo, K. Yun and J.Y. Choi, "Appearance and motion based deep learning architecture for moving object detection in moving camera," 2017 IEEE International Conference on Image Processing (ICIP), IEEE, pp. 1827-1831, 2017.

[25] Y. Chen, J. Wang, B. Zhu, M. Tang and H. Lu, "Pixel-wise deep sequence learning for moving object detection," IEEE Transactions on Circuits and Systems for Video Technology, 2017.

[26] E.D. Tejada and P.A. Rodriguez, "Moving object detection in videos using principal component pursuit and convolutional neural networks," 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, pp. 793-797, 2017.

[27] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," Advances in neural information processing systems, pp. 91-99, 2015.

[28] M. Babaee, D.T. Dinh and G. Rigoll, "A deep convolutional neural network for video sequence background subtraction," Pattern Recognition, Elsevier, Vol. 76, pp. 635-649, 2018.

[29] T.N. Le and A. Sugimoto, "Video salient object detection using spatiotemporal deep features,." IEEE Transactions on Image Processing, Vol. 27, No. 10, pp. 5002-5015, 2018.

[30] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 580–587, 2014.

[31] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," IEEE international conference on computer vision, IEEE, pp. 4489–4497, 2015.

[32] D. Zeng and M. Zhu, "Background Subtraction Using Multiscale Fully Convolutional Network," IEEE Access, Vol. 6, pp. 16010-16021, 2018.

[33] T.H. Lin and C.C. Wang, "Deep learning of spatio-temporal features with geometric-based moving point detection for motion segmentation," 2014 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 3058-3065, 2014.



Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:56:09 UTC from IEEE Xplore. Restrictions apply.