Engineering and Technology Journal 40 (04) (2022) 617-626

Engineering and Technology Journal
Journal homepage: https://etj.uotechnology.edu.iq

Comparative Analysis of GMM, KNN, and ViBe Background Subtraction Algorithms Applied in Dynamic Background Scenes of Video Surveillance System

Maryam A. Yasir *, **Yossra H. Ali**

Computer Science Dept., University of Technology-Iraq, Alsina'a street, 10066 Baghdad, Iraq.
*Corresponding author Email: cs.19.86@grad.uotechnology.edu.iq

HIGHLIGHTS

- Implementing and comparing the results of GMM, KNN, and ViBe background subtraction algorithms.
- Applying algorithms on dynamic background scenes from a well-known CDnet 2014 benchmark dataset.
- A wide range of evaluation metrics has been used (Accuracy, Precession, Recall, F1, FPR, FNR, and PWC).
- ViBe background subtraction algorithm shows the best overall performance.

ARTICLE INFO

Handling editor: Rana F. Ghani

Keywords: Background Subtraction; Dynamic Background; CDnet 2014; Foreground Objects; Moving Object; GMM; KNN; ViBe Video Surveillance Systems.

ABSTRACT

Background subtraction is the most prominent technique applied in the domain of detecting moving objects. However, there is a wide range of different background subtraction models. Choosing the best model that addresses a number of challenges is still a vital research area.

Therefore, in this article we present a comparative analysis of three promising algorithms used in this domain, GMM, KNN and ViBe. CDnet 2014 is the benchmark dataset used in this analysis with several quantitative evaluation metrics like precession, recall, f-measures, false positive rate, false negative rate and PWC. In addition, qualitative evaluations are illustrated in snapshots to depict the visual scenes evaluation. ViBe algorithm outperform other algorithms for overall evaluations.

1. Introduction

Detecting moving objects is one of the remarkable topics in the field of computer vision due to increasing security concerns and the spread of CCTV(Closed-Circuit Television Cameras) and CCD (Charge-Coupled Device) image sensors[1]. A wide range of real-time applications is employed in different environments with many occurring challenges. Extracting a foreground object from a particular region of interest can be devoted to the synopsis, tracking, and anomaly detection [2]. Many techniques have been used in the domain of moving object detection, like optical flow, frame differencing, and temporal differencing, while the background subtraction technique is the most popular one. It is authorized by the ease of implementation in a vast scope of video surveillance systems [3]. Numerous video surveillance environments are appealing. It is comprised of: a surveillance system of human activities like (transportation, warehouses, sports, and the military), a surveillance system of nature like (animals, insects, nature scenes), etc. [4]. In this paper, we shed light on three well-known techniques used in the domain of background subtraction. The rest of the paper contains the background subtraction models and process, the discussion of algorithms, the experiments, and analytical comparison with evaluation metrics. Eventually, we present a concrete conclusion supported by a summary table of the average foreground detection of the three models.

http://doi.org/10.30684/etj.v40i4.2154
Received 5 May 2021; Accepted 26 June 2021; Available online 25 April 2022
2412-0758/University of Technology-Iraq, Baghdad, Iraq

2. Background subtraction models

Background subtraction models have been extensively developed through the decades. The following are the main model approaches:

2.1 Basic models

In the basic models, the classification of background and foreground is done according to a threshold difference between the current frame and the background model [5]. Therefore, basic models are threshold dependable and work better on single model background distributions [6]. There are many examples of basic models like the mean model [7], median model [8], and histogram analysis model [9].

2.2 Mathematical models

These models are classified into ***parametrical statistical models*** and ***non-parametric statistical models.*** Gaussian Mixture Model (GMM) [10], Visual Background extractor (ViBe) [11], and Substance Sensitivity Segmenter (SuBSENSE) algorithm are examples of the parametrical statistical models [12]. Where kernel density estimation (KDE) [13] and Pixel-Based Adaptive Segmenter (PBAS) [14] are examples of non-parametrical statistical models.

2.3 Filter models

These models are single processing models that predict the background according to a pixel's previous information value, whether it is orientation or intensity value [4]. Filters like Tchebychev filter [15], Wiener filter [16], Kalman filter [17], Correntropy filter [18] and optical flow [19] are examples of this model.

2.4 Clustering models

These models use the intensity of color for each pixel to identify the foreground mask. Then, the current sample frame pixels are compared to the background and foreground clusters, and a decision is made on which cluster the pixel belongs to. Various methods lie under this model, such as background reconstruction [20], K-means [21], and the Codebook method [22].

2.5 Machine learning models

Machine learning models are the cutting edge models that comprise subspace learning with reconstructive and discriminative techniques [25,26], support vector machines (SVM) [25], robust subspace tracking [26], neural networks, and deep learning [29,30]. They have been broadly experimented with mainly because of the huge development of hardware processing power and the availability of an adequate collection of video surveillance training datasets [29]. Recently the convolutional neural network (CCN) has been implemented in various applications in the domain of computer vision. It was first introduced in the background subtraction field by Braham and Droogenbroeck [30]. CNN was originally developed for image classification tasks, and it is well-known for its robust performance in computer vision. However, the lack of datasets and the need for high computation restrict the use of CNNs in real-time applications [33,34]. The machine learning methods perform well with illumination changes but not with abnormal dynamic background movements or non-static shadows challenges. In addition, these models suffer from extensive processing time [33].

A lot of research has been conducted to study background subtraction, but no single model can overcome all the challenges regarding real-time video surveillance applications. Therefore, many articles presented a fusion of multiple models and strategies for a robust performance like in real-time semantic background subtraction (RT-BSB) [34].

3. Background subtraction process

Various background subtraction models have been built to extract the background from the scene resulting in moving foreground object detection. However, all these models share the same pipelined process for the background subtraction. The sequence of stages is as follows:

3.1.1 Background initialization

This stage represents the process of generating the first background scene by taking several video frames.

3.1.2 Background modeling

This stage represents the process of describing the background representation scene to be compared with the current frame.

3.1.3 Background maintenance

This stage represents the process of updating the background model according to the frequent changes.

3.1.4 Foreground detection

It is the process of classifying the pixel into background or foreground. This is achieved by comparing the background model scene with the recent frame scene.

Certain pre-processing operations could involve video framing and color space altering. On the other hand, post-processing might be taken by applying specific algorithms to overcome a particular challenge in the process of background subtraction [35]. Figure1 depicts an overview of the background subtraction process.

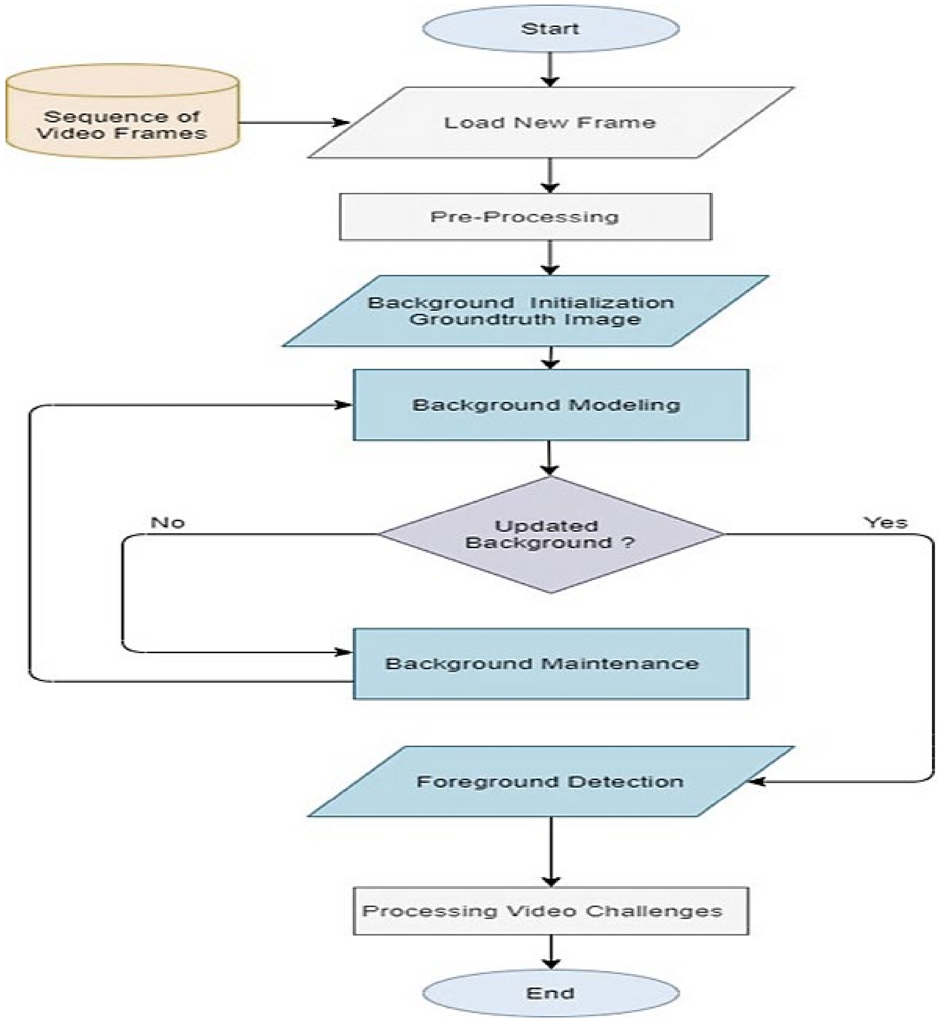

Figure 1: The overview of the background subtraction process [36]

4. Background subtraction algorithms

4.1 Gaussian mixture model (GMM)

The GMM was first introduced in 1999 by Stauffer and Grimson [10] as a breakthrough approach in the field of background modeling, where each pixel is modeled as a mixture of Gaussian using a real-time approximation for updating the model [37]. GMM is a statistical parametrical and unsupervised model that uses three main parameters: mean vectors, covariance matrices, and mixture weights from all component densities [38]. GMM is famous for its robust performance concerning gradual illumination changes and dynamic background challenges. Although, the GMM tends to have a poor performance with unexpected illumination changes and unusual background motions[39]. In addition, the efficiency of GMM is prone to be reduced due to the parametrical nature of the algorithm, like in selecting inaccurate parameters or the time consumed in selecting these parameters [40].

Many studies have been performed to enhance the GMM against different background subtraction challenges, such as Boosted Gaussian Mixture Model (BGMM), where the performance is boosted using a color space classifier and dynamic learning for updating the background model [41]. Improved Gaussian background modeling (GBM) is also an example of the enhanced model using wavelet denoising applied to the foreground object. This model performs better with respect to shadow and illumination changes challenges [42]. GMM has been used frequently in the field of foreground detections and is still applied in many optimized versions. The following are the general steps of the GMM algorithm [38].

4.2 K-nearest neighbor (KNN)

KNN is one of the well-known non-parametrical supervised algorithms where K denotes the number of neighbors to be considered in voting for the class detection. It is early introduced in the 1970s as a statistical model. Since then, KNN has been used in classification and regression problems [41]. KNN is influenced by the concept that similar things are close. In this context, the algorithm finds the closest k samples (neighbors) to the query using distance and then determines the query label with the same class label as the closest sample. The following are the general steps of the KNN algorithm [43].

5. Visual background extractor (vibe)

The viBe was introduced first by Barnich and Droogenbroeck in 2009 as a background subtraction algorithm, which adopts a new technique reliant on the pixel value in a multicolor space to be constrained to a diameter of the neighborhood [11]. The classification in ViBe is achieved by searching the neighbors to find the closest value to a pixel and updating the model without using the probability density function (pdf) to gain better results [44]. The viBe model has been enhanced over the years in different works [45]. The following are the general steps of the ViBe algorithm [46].

6. Experiments

This article compares foreground detection using the aforementioned background subtraction models (GMM, KNN, and ViBe). The models are tested using the CDnet 2014[47] benchmark dataset considering the dynamic background challenge videos. CDnet 2014 was chosen as the most well-known and preferable dataset due to providing a wide range of videos with several challenges and providing ground-truth scenes that help in statistical evaluation. In these experiments, we compare the background subtraction image with the correspondent ground-truth image to evaluate the performance of each method concerning quantitative evaluation metrics at the pixel level, and the background subtraction method classifies the pixels into background or foreground. Seven metrics are utilized for the performance evaluation as follows:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall(sensitivity) = \frac{TP}{TP+FN} \quad (3)$$

$$F - Measures(F1) = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$

$$FPR = \frac{FP}{FP+TN} \quad (5)$$

$$FNR = \frac{FN}{TP+FN} \quad (6)$$

$$PWC = \frac{FP+FN}{TP+FN+TN+FP} * 100 \quad (7)$$

TP is the sum of foreground pixels correctly classified, **TN** is the sum of background pixels correctly classified, **FP** is the sum of background pixels incorrectly classified as foreground pixels, and **FN** is the sum of foreground pixels incorrectly classified as background pixels. ***Accuracy*** indicates the correct classification for a pixel. Whether it is a foreground or a background pixel***, Precision*** indicates the proportion of truly detected foreground pixels to the number of all pixels classified as foreground pixels. R***ecall*** indicates the number of pixels correctly classified as a foreground of all the foreground pixels. The ***F-measure*** is the harmonic mean of recall and precision.

On the other hand, we have the following metrics; False Positive Rate (***FPR)*** is the sum of background pixels misclassified as foreground pixels. False Negative Rate ***(FNR)*** is the sum of foreground pixels misclassified as background pixels. Percentage weight loss ***(PWC)*** indicates the error rate, which is the percentage of misclassified pixels to the original pixels. Usually, a low recall is a sign of over-segmentation of the foreground objects, whereas a low precision is a sign of under-segmentation of the foreground objects. In addition, high F-measures are a sign of a robust background subtraction algorithm, while the lower FPR, FNR, and PWC are signs of a better performance. The following Tables 1, 2, and 3 illustrate the analytical metrics results of applying the GMM, KNN, and ViBe models, respectively, on the CDnet 2014 dynamic background videos dataset. Tables (1, 2, and 3) depict the results of applying each algorithm on the videos (overpass, fall, boats, canoe, fountain1, and fountain2) that contain dynamic background challenges. The comparative analysis uses the aforementioned metrics of accuracy, precession, recall, F-measure(F1), FPR, FNR, and PWC. Table 4 illustrates the average results of applying each algorithm on different dynamic background videos with the same evaluation metrics. The results show that the ViBe model performed the best precession and recall. Also, it achieved the lowest results for FNR, FPR, and PWC. On the other hand, KNN achieved the best harmonic mean of precession and recall (F1) with the highest accuracy.

Table 1: Performance metrics of GMM on CDnet 2014 dataset

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Overpass	0.987	0.151	0.512	0.423	0.008	0.488	1.302
Fall	0.979	0.146	0.781	0.461	0.018	0.219	2.097
Boats	0.993	0.134	0.271	0.371	0.003	0.729	0.698
Canoe	0.968	0.297	0.402	0.407	0.014	0.598	3.245
Fountain1	0.987	0.027	0.631	0.209	0.013	0.369	1.307
Fountain2	0.996	0.095	0.663	0.508	0.004	0.337	0.427

Table 2: Performance metrics of KNN on CDnet 2014 dataset

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Overpass	0.988	0.173	0.703	0.549	0.008	0.297	1.180
Fall	0.968	0.119	0.699	0.377	0.029	0.301	3.210
Boats	0.995	0.142	0.318	0.426	0.002	0.682	0.528
Canoe	0.980	0.375	0.615	0.585	0.011	0.385	1.983
Fountain1	0.987	0.023	0.497	0.180	0.013	0.503	1.296
Fountain2	0.997	0.109	0.587	0.531	0.002	0.413	0.306

Table 3: Performance metrics of ViBe on CDnet 2014 dataset

Video	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
Overpass	0.989	0.164	0.564	0.157	0.006	0.436	0.111
Fall	0.967	0.125	0.570	0.129	0.026	0.430	3.311
Boats	0.992	0.129	0.371	0.088	0.004	0.629	0.824
Canoe	0.975	0.455	0.512	0.360	0.007	0.482	2.516
Fountain1	0.991	0.055	0.750	0.079	0.009	0.250	0.899
Fountain2	0.996	0.126	0.780	0.133	0.003	0.220	0.355

Table 4: Average performance metrics of GMM, KNN and ViBe on CDnet 2014 dataset

Model	Accuracy	Precession	Recall	F1	FPR	FNR	PWC
GMM	0.985	0.141	0.543	0.396	0.010	0.456	1.512
KNN	0.986	0.156	0.569	0.441	0.011	0.430	1.417
ViBe	0.985	0.175	0.591	0.157	0.009	0.407	1.336

Moreover, Figure 2 depicts the visual results' comparison of the foreground results of applying the three models on the CDnet 2014 dynamic background scenes. Where (a) is the original scene, (b) is the ground truth provided by the dataset, (c) is the foreground mask created by the GMM model, (d) is the foreground mask created by the KNN model, and (e) is the foreground mask created by ViBe model.

7. Conclusions

This article experimented with an effective comparative analysis of three robust and well-known background subtraction algorithms. The CDnet 2014 dataset was employed as a benchmark focusing on the dynamic background scenes. Seven different quantitative evaluation metrics were used in the experiments on the video surveillance dataset. These metrics are accuracy, precession, recall, F1, FPR, FNR, and PWC. The aforementioned performance metrics show that all three models perform well in terms of accuracy, while KNN performs the best in terms of F1 measures. But ViBe algorithm performs the best in terms of precession, recall, and the best in terms of lowest results for FPR, FNR, and PWC, making it the most reliable model among all three models in this comparative analysis of dynamic video surveillance scenes.

Algorithm 1: GMM algorithm
Input: Mean $\mu \mathbf{c}$
Covariance matrix $\sum \mathbf{c}$
Mixing coefficients $\boldsymbol{\pi}\mathbf{c}$
Output: foreground mask of the moving object in the video scene.

1. Calculate $\boldsymbol{\Upsilon}\mathbf{c}$ for all $\mathbf{k}$.
2. Recalculate parameters ($\mu\mathbf{c}, \sum\mathbf{c}$ *and* $\boldsymbol{\pi}\mathbf{c}$) based on $\boldsymbol{\Upsilon}\mathbf{c}$ values.
3. Calculate log-likelihood function.
4. Set the stop criterion for the convergence.
5. If the log-likelihood value converges to some value
   (or if all the parameters converge to some values)
   then stop,
   else return to Step 2.

Algorithm 2: KNN algorithm

Input: unknown example(query) **x**, the dataset **S** and the distance **d**

Output: foreground mask of the moving object in the video scene.

1. for(x',l') ϵ **S** do
   compute the distance d(x',x)
   end for
2. Sort the |**S**| distance by increasing order.
3. Count the number of occurrences of each class lj among the K nearest neighbor.
4. Assign to **x** the most frequent class.

Algorithm 3: ViBe algorithm

Input: Given samples **N**
The values of a sample **x v(x)**
The distance threshold **R**
Classification threshold **λ**
Subsampling rate **t**

Output: foreground mask of the moving object in the video scene.

1. **for** each pixel **x do**
2. **while** neighbours < λ **and** index < N **do**
3. Compute distance between v(x) and v(i)
4. **if** distance < distance threshold (R) **then**
   neighbours= neighbours + 1
   end if
5. counter=counter + 1
   end while
6. **if** neighbours ≥ λ **then**
7. Save the pixel ϵ background
8. Update current pixel background model with probability 1/t
9. Update neighbouring pixel background model with probability 1/t
   else
10. Save the pixel ϵ foreground
    end if
    end for

(a) Original scene (b)Ground-truth scene (c) GMM detection (d) KNN detection (e) Vibe detection

Figure 2: Comparison of foreground detection results

Author contribution

All authors contributed equally to this work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

[1] N. K. Jain, , R.K. Saini, A review on traffic monitoring system techniques, Soft Comput. Theories and Appl.,742(2019) 569–577. https://doi.org/10.1007/978-981-13-0589-4_53

[2] D. Gutchess, M. Trajkovics, E. C. Solal, D. Lyons, A. K. Jain, A background model initialization algorithm for video surveillance, Proc. Eighth. IEEE Int. Conf. Comp. Vision., 1 (2001) 733-740 .https://doi.org/10.1109/iccv.2001.937598

[3] D. Zeng, M. Zhu, A. Kuijper, Combining background subtraction algorithms with convolutional neural network, J. Electron. Imaging., 28 (2019) 13011. https://doi.org/10.1117/1.JEI.28.1.013011

[4] K. Goyal ,J. Singhai, Review of background subtraction methods using Gaussian mixture model for video surveillance systems, Artif. Intell. Rev., 50 (2018) 241–259. https://doi.org/10.1007/s10462-017-9542-x

[5] T. Bouwmans, Recent advanced statistical background modeling for foreground detection-a systematic survey, Recent .Patents. Comput. Sci., 4 (2011) 147–176. https://doi.org/10.2174/2213275911104030147

[6] S. Lee, D. Kim, Background subtraction using the factored 3-way restricted Boltzmann machines, arXiv (2018)1802-01522. https://doi.org/10.48550/arXiv.1802.01522

[7] P. W. Power, J. A. Schoonees, Understanding background mixture models for foreground segmentation, Proc. image vision comput. New Zealand, 2002 .

[8] M. Hadiuzzaman, N. Haque, F. Rahman, S. Hossain, M. R. K. Siam, T. Z. Qiu, Pixel-based heterogeneous traffic measurement considering shadow and illumination variation, Signal. Image. Video. Proc., 11 (2017) 1245–1252. https://doi.org/10.1007/s11760-017-1081-z

[9] Z. Zivkovic , F. V.D. Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, Pattern Recognit. Lett., 27 (2006) 773–780. https://doi.org/10.1016/j.patrec.2005.11.005

[10] C. Stauffer, W. E. L. Grimson, Adaptive background mixture models for real-time tracking, in Proceedings, IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2 (1999) 246-252 . https://doi.org/10.1109/cvpr.1999.784637

[11] O. Barnich , M. Van , Droogenbroeck, ViBe: a powerful random technique to estimate the background in video sequences, IEEE Int. Conf. Acoust. Speech . Signal. Proc. , (2009) 945-948. https://doi.org/10.1109/icassp.2009.4959741

[12] P. L. St-Charles, G. A. Bilodeau, R. Bergevin, Subsense: A universal change detection method with local adaptive sensitivity, IEEE Trans. Image Proc., 24 (2014) 359–373. https://doi.org/10.1109/tip.2014.2378053

[13] M. Hua, Y. Li, Y. Luo, Robust Background Modeling with Kernel Density Estimation., Online Eng. Innov. based Intell. Proc., 11 (2015)13-16. https://doi.org/10.3991/ijoe.v11i8.4880

[14] S. Javed, S. H. Oh, S. K. Jung, An Improved Pixel-Based Adaptive Background Segmenter for Visual Surveillance System, 2014. Munteanu, Oana et al, The detection of moving objects in video by background subtraction using Dempster-Shafer theory. (2015). F. El Baf, T. Bouwmans, B. Vachon, Type-2 fuzzy mixture of Gaussians model: application to background modeling, Int. Sym. Vis. Comput., (2008) 772–781. https://doi.org/10.1007/978-3-540-89639-5_74

[15] R. Chang, T. Gandhi, M. M. Trivedi, Vision modules for a multi-sensory bridge monitoring approach, Proc. Int. IEEE Conf. Intell. Transport. Syst., (2004) 971–976. https://doi.org/10.1109/itsc.2004.1399038

[16] K. Toyama, J. Krumm, B. Brumitt, B. Meyers, Wallflower: Principles and practice of background maintenance, Conf. Comput. Proc. Seventh IEEE Int. Conf., 1(1999) 255-261. https://doi.org/10.1109/iccv.1999.791228

[17] F. Lei , X. Zhao, Adaptive background estimation of underwater using Kalman-Filtering, IEEE Int.Congress .Image Signal Proc., 1 (2010) 64–67. https://doi.org/10.1109/cisp.2010.5647080

[18] G. T. Cinar , J. C. Pr cipe, Adaptive background estimation using an information theoretic cost for hidden state estimation, Int. Jt. Conf. Neural. Netw., (2011) 489–494. https://doi.org/10.1109/ijcnn.2011.6033261

[19] X. Yu, X. Chen, H. Zhang, Accurate motion detection in dynamic scenes based on ego-motion estimation and optical flow segmentation combined method, Symp.Photo. Optoel., 2011, 1–4. https://doi.org/10.1109/sopo.2011.5780637

[20] M. Xiao, C. Han, X. Kang, A background reconstruction for dynamic scenes, Int. Conf. Infor. Fusion, (2006) 1–7. https://doi.org/10.1109/icif.2006.301727

[21] D. Butler, S. Sridharan, V. Bove, Real-time adaptive background segmentation, IEEE Int. Conf. Acoustics, Speech, Signal. Proc., 3 (2003) 349. https://doi.org/10.1109/icassp.2003.1199481

[22] L. Wang , C. Pan, Effective multi-resolution background subtraction, IEEE Int. Conf. Acoust. Speech . Signal .Proc., (2011) 909–912. https://doi.org/10.1109/icassp.2011.5946552

[23] E. J. Candès, X. Li, Y. Ma, J. Wright, Robust principal component analysis, J. ACM, 58 (2011) 1–37. https://doi.org/10.1145/1970392.1970395

[24] D. Farcas, C. Marghes, T. Bouwmans, Background subtraction via incremental maximum margin criterion: a discriminative subspace approach, Mach. Vis. Appl., 23 (2012) 1083–1101. https://doi.org/10.1007/s00138-012-0421-9

[25] N. Vaswani, T. Bouwmans, S. Javed, P. Narayanamurthy, Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery, IEEE Signal Proc. Mag., 35 (2018) 32–55. https://doi.org/10.1109/msp.2018.2826566

[26] I. Junejo, A. Bhutta, H. Foroosh, Dynamic scene modeling for object detection using single-class SVM, Proc. IEEE Int. Conf. Image Proc, 1 (2010) 1541–1544.

[27] L. Maddalena , A. Petrosino, Self-organizing background subtraction using color and depth data, Multimed. Tools Appl., 78 (2019) 11927–11948. https://doi.org/10.1007/s11042-018-6741-7

[28] T. Minematsu, A. Shimada, H. Uchiyama, R. Taniguchi, Analytics of deep neural network-based background subtraction, J. Imaging, 4 (2018) 78. https://doi.org/10.3390/jimaging4060078

[29] T. Bouwmans, S. Javed, M. Sultana, S. K. Jung, Deep neural network concepts for background subtraction: A systematic review , comp. Eval. Neural, Netw., 117 (2019) 8-66. https://doi.org/10.1016/j.neunet.2019.04.024

[30] M. Braham , M. V. Droogenbroeck, Deep background subtraction with scene-specific convolutional neural networks, int. Conf. Syst. Signal. Imag. Proc., (2016) 1–4. https://doi.org/10.1109/iwssip.2016.7502717

[31] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. neural inf. Proc. Syst., 60 (2012) 84–90. https://doi.org/10.1145/3065386

[32] P. Xu, M. Ye, X. Li, Q. Liu, Y. Yang, and J. Ding, Dynamic background learning through deep auto-encoder networks, Proc . ACM int.conf. Multimed., (2014) 107–116. https://doi.org/10.1145/2647868.2654914

[33] T. Bouwmans, Subspace learning for background modeling: A survey, Recent Patents Comput. Sci., 2 (2009) 223–234. https://doi.org/10.2174/2213275910902030223

[34] A. Cioppa, M. V. Droogenbroeck, M. Braham, Real-time semantic background subtraction, IEEE Int. Conf. Image Proc. , (2020) 3214–3218. https://doi.org/10.1109/icip40778.2020.9190838

[35] R. A. Hadi, L. E. George, M. J. Mohammed, A computationally economic novel approach for real-time moving multi-vehicle detection and tracking toward efficient traffic surveillance, Arab. J. Sci. Eng., 42 (2017) 817–831. https://doi.org/10.1007/s13369-016-2351-8

[36] M. Yasir , Y. Ali, Review on Real Time Background Extraction: Models, Applications, Environments, Int. J. Online , Bio.Eng., 17(2021) 2626-8493 . https://doi.org/10.3991/ijoe.v17i02.18013

[37] C. Lin, B. Yan, W. Tan, Foreground detection in surveillance video with fully convolutional semantic network, IEEE Int. Conf. Imag. Proc. (2018) 4118–4122. https://doi.org/10.1109/icip.2018.8451816

[38] D. A. Reynolds, Gaussian Mixture Models, Encycl. biometrics, 741 (2009) 659–663.

[39] T. Bouwmans, F. El Baf, B. Vachon, Foreground detection: a survey, Handbook , Pattern. Recognit. Computer Visi.,(2010) 181–199. https://doi.org/10.1142/9789814273398_0008

[40] B. White, M. Shah, Automatically tuning background subtraction parameters using particle swarm optimization, IEEE Int.Conf. Multimed . Expo, (2007) 1826–1829. https://doi.org/10.1109/icme.2007.4285028

[41] I. Martins, P. Carvalho, L. C. Real, J. L. A. Castro, BMOG: boosted Gaussian mixture model with controlled complexity, Pattern. Anal. Appl., 21 (2018) 641–654. https://doi.org/10.1007/s10044-018-0699-y

[42] X. Lu, C. Xu, L. Wang, L. Teng, Improved background subtraction method for detecting moving objects based on GMM, IEEE Trans. Electr. Electron. Eng., 13 (2018) 1540–1550. https://doi.org/10.1002/tee.22718

[43] S. T. Ali, K. Goyal, J. Singhai, Moving object detection using self adaptive Gaussian Mixture Model for real time applications, Int. Conf. Recent, Innov. Signal proc. Embed. Syst., (2017) 153–156. https://doi.org/10.1109/rise.2017.8378144

[44] O. Barnich , M. V. Droogenbroeck, ViBe: A universal background subtraction algorithm for video sequences, IEEE Trans. Image. Process., 20 (2010) 1709–1724. https://doi.org/10.1109/tip.2010.2101613

[45] L. F. Huang, Q. Y. Chen, J. F. Lin, H. Z. Lin, Block background subtraction method based on ViBe, Appl. Mech. Mater., 556-562 (2014) 3549–3552. https://doi.org/10.4028/www.scientific.net/amm.556-562.3549

[46] X. Zhou, X. Liu, A. Jiang, B. Yan, C. Yang, Improving video segmentation by fusing depth cues and the visual background extractor (ViBe) algorithm, Sensors, 17 (2017) 1177.https://doi.org/10.3390/s17051177

[47] Y. Wang, P.M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, P. Ishwar, CDnet 2014: An expanded change detection benchmark dataset, IEEE conf. comput. vision . pattern recognit. Wksp., 17 (2014) 387–394. https://doi.org/10.1109/cvprw.2014.126