Received: 8 November 2022 | Revised: 30 December 2022 | Accepted: 17 January 2023 | Published online: 8 February 2023

RESEARCH ARTICLE

Artificial Intelligence and Applications
2023, Vol. 00(00) 1–8
DOI: 10.47852/bonviewAIA3202526

Spatiotemporal Edges for Arbitrarily Moving Video Classification in Protected and Sensitive Scenes

Maryam Asadzadehkaljahi[1], Arnab Halder[2,*], Umapada Pal[2] and Palaiahnakote Shivakumara[3]

[1]*PromiseQ, Germany*

[2]*Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India*

[3]*Computer Science and Information Technology, University of Malaya, Malaysia*

Abstract: Classification of arbitrary moving objects including vehicles and human beings in a real environment (such as protected and sensitive areas) is challenging due to arbitrary deformation and directions caused by shaky camera and wind. This work aims at adopting a spatiotemporal approach for classifying arbitrarily moving objects. The intuition to propose the approach is that the behavior of the arbitrary moving objects caused by wind and shaky camera is inconsistent and unstable, while, for static objects, the behavior is consistent and stable. The proposed method segments foreground objects from background using the frame difference between median frame and individual frame. This step outputs several different foreground information. The method finds static and dynamic edges by subtracting Canny of foreground information from the Canny edges of respective input frames. The ratio of the number of static and dynamic edges of each frame is considered as features. The features are normalized to avoid the problems of imbalanced feature size and irrelevant features. For classification, the work uses 10-fold cross-validation to choose the number of training and testing samples, and the random forest classifier is used for the final classification of frames with static objects and arbitrarily moving objects. For evaluating the proposed method, we construct our own dataset, which contains video of static and arbitrarily moving objects caused by shaky camera and wind. The results on the video dataset show that the proposed method achieves the state-of-the-art performance (76% classification rate) which is 14% better than the best existing method.

Keywords: moving objects detection, vehicles movements detection, shaky camera detection, subtraction approach, arbitrarily moving objects detection

1. Introduction

Automation is common for all the fields to make a system cost-effective, accurate, and to prevent human error during night especially for the protected and sensitive areas. The main application of the proposed work is monitoring and protecting sensitive areas irrespective of day and night. For instance, in the case of big outlets, industrial areas, and even parking vehicles on the street, it is necessary to watch and monitor to protect the goods, material, and vehicles from the intruder, robber and thieves. In the same way, there are some dangerous situations, like electric transformer, suicide spots, etc. In these situations, the same surveillance system can be used for saving the lives of people. Therefore, the developed system can be used for both protecting areas and saving people's life. There are several methods for object detection, object tracing, intruder detection in the video, etc., in the literature (Arjun et al., 2020; Chandra and Panda, 2021; Kim et al., 2020; Lenac et al., 2021; Ramesh et al., 2021; Vndana et al., 2022). However, most of the existing methods focus on detecting particular object in the day video for particular applications but not generalized methods and for different applications. In addition, none of the methods considers object movements caused by wind and shaky camera for detection as well as tracking. When the objects including leaves and branches of the trees are moving arbitrarily due to shaky camera and wind, the existing methods may not perform well because the features of object tracking overlap with the features of leaves and tree branches. Furthermore, most methods may not target developing real-time systems. Therefore, classifying arbitrary moving vehicles in the video caused by shaky camera and wind irrespective of day and night is an open challenge for the computer vision and image processing community. This is the motivation to develop a new simple and effective system based on spatiotemporal features for classifying arbitrarily moving objects and static objects in the video.

*Corresponding author: Arnab Halder, Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India. Email: arnabhalder1997@gmail.com

Figure 1
The challenges of frames captured by shaky and non-shaky camera for classification

Frame-0 Frame-1 Frame-2

(a) Examples of frames captured by shaky camera

(b) Examples of frames captured by non-shaky camera

It is evident from Figure 1 that sample frames captured by shaky and non-shaky camera are shown in Figure 1(a) and (b), respectively. It is observed from Figure 1(a) and (b) that the frames are suffering from poor quality and the objects are not visible properly. In the case of Figure 1(a), due to shaky camera, the frames are blurred while in Figure 1(b), due to indoor scenes, the objects in the frames are not visible. There are number of methods proposed in the literature for static and moving object detection (Boufares et al., 2021; Wang et al., 2021). However, these methods focus on the video captured in daytime with high quality. In addition, the methods are good for detecting objects which move in particular direction and speed but not the video containing arbitrary movements and directions. Therefore, there is dearth to develop a new method for classification of static and arbitrary movements objects, such as leaves in a tree and objects movements due to shaky camera in the video in real-time environments.

To address this challenge, the proposed work introduces a new spatiotemporal feature-based approach. This is new for classifying arbitrary moving objects in the day and night video compared to the state-of-the-art methods. In addition, the developed method should fit for the real-time environment. Hence, the proposed work focuses on simple and effective methods rather than heavy deep learning models which require a large number of samples and a greater number of computations. It is also true that deep learning-based models lack generality, and hence, the models may not perform well for the unpredictable situations. To the best of our knowledge, this is the first work for addressing challenges of shaky, non-shaky camera as well as wind effect.

The proposed method works based on the fact that normal moving objects exhibit regular patterns like uniform direction, speed, and shapes while arbitrary movements objects exhibit irregular patterns like non-uniform direction, speed, and deformable shapes. These observations motivated us to explore edge-based features extracted from foreground (moving objects) and background (static objects) in this work. This makes sense because the number of static and dynamic edges changes according to the direction of objects and movements. Thus, the contributions are follows: (i) this is the first work for classifying arbitrary movement objects caused by shaky camera and leaves caused by wind; (ii) the spatiotemporal information for extracting edge-based features is new compared to the existing methods, and (iii) the proposed method is tested on real-time environment like day and night videos.

The structure of the paper is as follows. The review of state-of-the-art methods for classification of static and moving objects is presented in Section 2. In Section 3, the method for foreground and background separation, edge-based features, and the approach for classification of static and arbitrary movement objects in video are described. Section 4 provides a discussion of several experiments to validate the proposed and existing methods. The conclusion and summary are presented in Section 5.

2. Related Work

A number of methods are developed in the past for moving object detection and foreground object separation in the video. Since these methods are relevant to arbitrary object movements detection, we review the latest papers. Boufares et al. (2021) proposed a method for moving object detection using temporal difference and OTSU thresholding technique. The approach uses input frame difference at pixels level for detecting moving objects in the videos. Wang et al. (2021) developed a model for moving object detection using frame difference and algorithm for teaching video. The method uses OTSU thresholding approach and median filter for moving object detection. Sadkhan et al. (2021) aimed to detect moving objects and tracking moving objects in the video. The approach uses subtraction model for moving object detection, texture, shape, and color-based features are extracted for classification, and detected moving object tracking is done by kernel- and point-based approaches. Rahiminezhad et al. (2022) explored adaptive coefficient and background subtraction for detecting moving objects in the video. The method focuses more on hardware implementation to fix it in the real-time environment. Sultana et al. (2021) developed a model based on adversarial regularization for moving object detection in complex scenes. The approach is capable of handling partial occlusion and poor-quality images.

Tang and Liu (2022) proposed a method for moving object detection using ghost and shadow. The key idea of the method is to explore visual background extraction. The OTSU thresholding is used to detect moving objects in the video. Shu et al. (2021) focused on developing a method for small moving object detection in the video. The method involves an event-based moving object detection and for tracking, the work uses registration and foreground enhancement models. Wang et al. (2022) presented a model for multi-scale moving object detection based on spatiotemporal online matrix factorization. The approach uses temporal difference motion difference and partial spatial motion information. Kim et al. (2022) developed a model for moving object detection based on instant background modeling. The approach uses spatiotemporal information and instant background modeling which consists of inpainting and super pixels to enhance the fine details in the images. Huang et al. (2021) developed a method for moving object detection based on independent component analysis. The approach uses frame difference model for foreground moving object detection. The features extracted by frame difference are combined with spatial information to achieve the best results. Deng et al. (2022) used super-resolution and optical image data for moving object detection. The approach uses CNN for feature extraction and fusion. Kovalenko et al. (2021) used image sequence for moving object detection based on thresholding technique. The approach estimates the deformation field using stochastic gradient procedure.

It is observed from the above review that none of the methods aim at detecting the objects which move in arbitrary directions in the video. Most methods focus on normal moving object detection-based subtraction model. Since the existing methods use specific properties of moving object detection, the methods may not be suitable for arbitrary movements object detection in the video. Furthermore, the scope of the existing methods is limited to day video but not night video where one can expect enormous degradations. Therefore, detecting arbitrary moving object detection in both day and night video is open challenge. Thus, this work aims at proposing a new method for detecting arbitrary moving object detection.

3. Proposed Method

This work considers video capturing at the rate of 30 frames per second as input for classification of video captured by shaky camera and non-shaky camera. As mentioned in the previous section, the proposed work exploits the observation that as direction of object changes, it affects the number of edges and their direction edges. To extract this observation, the proposed work separates foreground from the background using frame difference approach. The median of the 30 frames is computed, and it is considered as a background of the input video. Each frame is subtracted from the median frame, resulting in foreground regions. To study the effect of arbitrarily moving and static moving objects, we focus to get the Canny of respective input of 30 frames to get all the edges present in the frame, and foreground images or motion maps consisting of the objects in motion. So here we separated the edge pixels of an object into two domains. In the first domain, there exist all the edges of objects with static edge pixels as well as moving edge pixels, and in the other domain, there exist pixels of the moving objects only. The proposed method performs logical AND operation on these two distinct domains to get the specific pixels which present in both domains, in another word the output will be the edge pixels of the moving objects only, as it is present in both domains. This step outputs the number of moving edges and all edges for each frame from the two domains. With the number of moving edges and all edges, the method obtains the ratio, which is defined as the number of all edges divided by the number of moving edges. We believe that for the frame with static objects, the ratio yields a high value, as there exists a very smaller number of moving edge pixels, while a low value for the frame with moving objects, as there exists a high number of moving edge pixels. This makes a clear difference in classifying the frame with static and arbitrarily moving objects. Since for each frame the method outputs a ratio, 30 ratio features are extracted for the input video.

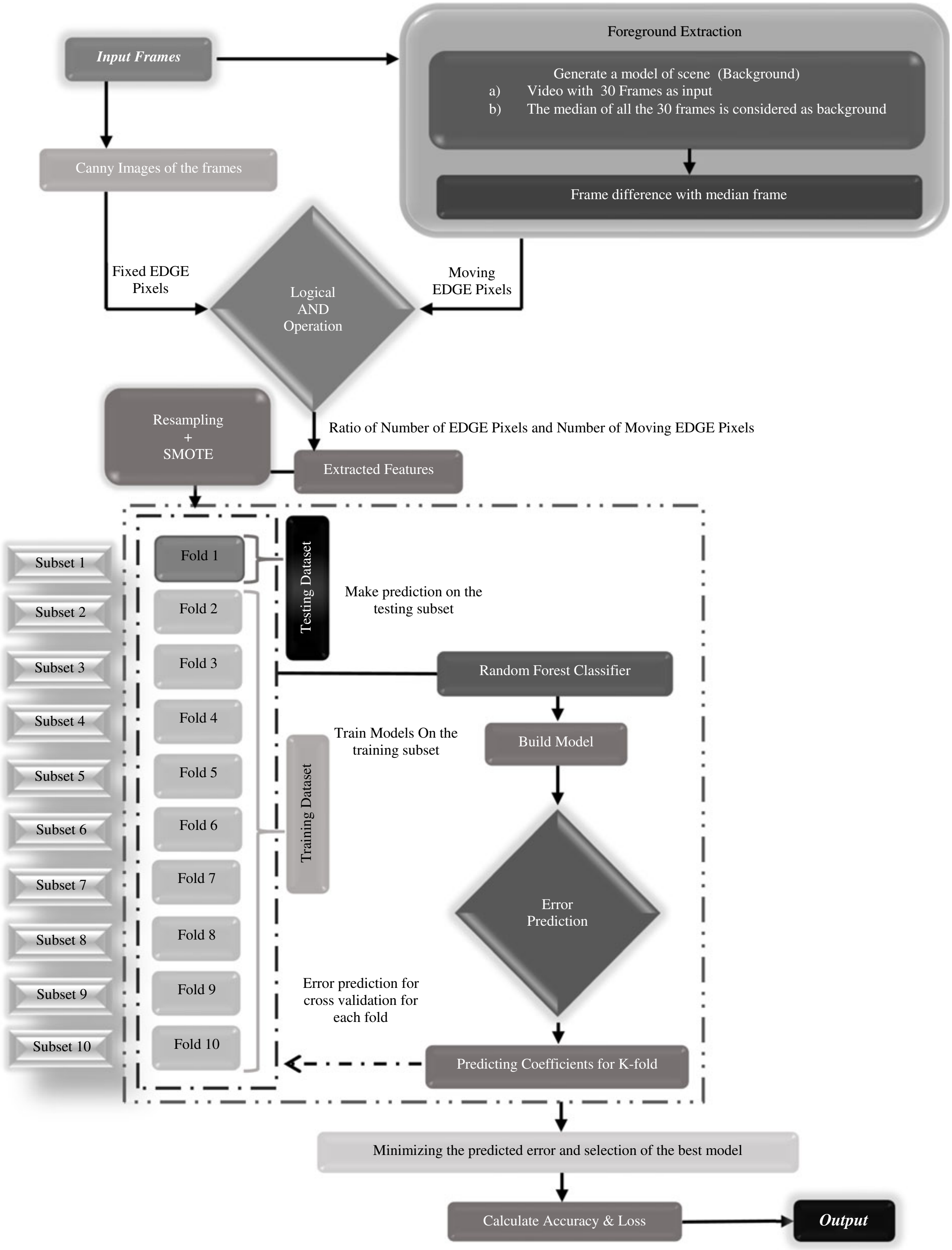
It is true that since there are no constraints on the number of objects in the frames and size of the frames, there is high chance of including the features that do not contribute much for classification. Therefore, the proposed work employs the Synthetic Minority Oversampling Technique (SMOTE) for resampling the dataset to balance the feature set and reduce the effect of irrelevant features (Chittroa et al., 2021; Solanki et al., 2021). For classification, the proposed work uses 10-fold cross-validation which provides the number of training and testing samples automatically. Since our goal is to develop simple and effective model which can fit for real-time environment, the proposed works adapt conventional machine learning models (Chhabra et al., 2022; Chithaluru et al., 2022; Kaushik et al., 2022) as classifier rather than deep learning models. For the final classification, the proposed work uses random forest (RF) classifier, which is a well-known technique for classification (Bharadwaj et al., 2022; Solanki et al., 2021). The reason to use random forest classifier is that this has the ability to handle imbalanced features and noise features, and it can avoid overfitting problems. The pipeline of the proposed method can be seen in Figure 2.

For classifying arbitrary moving objects in indoor and outdoor videos, the proposed work comprises the median-based frame difference model for foreground segmentation and the RF model for classification. These are well-known models for object detection and classification. However, the way the proposed work adapts the models for addressing challenges of arbitrarily moving object detection in day and night videos is new compared to the existing methods. The proposed model works based on the intuition that the behavior of the direction of arbitrary moving objects is inconsistent and unstable while for the static objects, the behavior of the directions is stable and consistent. This makes sense because the force generated by the wind and movements caused by shaky camera are arbitrary.

3.1. Edge-based approach

For input of 30 frames video, the proposed method obtains background by performing median operation on 30 frames. The result of median over 30 frames is considered as the background of the input video as it is illustrated in Figure 3(a) for both shaky and non-shaky camera frames. It is noted from Figure 3(a) that the details in the images are enhanced compared to respective input frames. The individual frames are subtracted with the

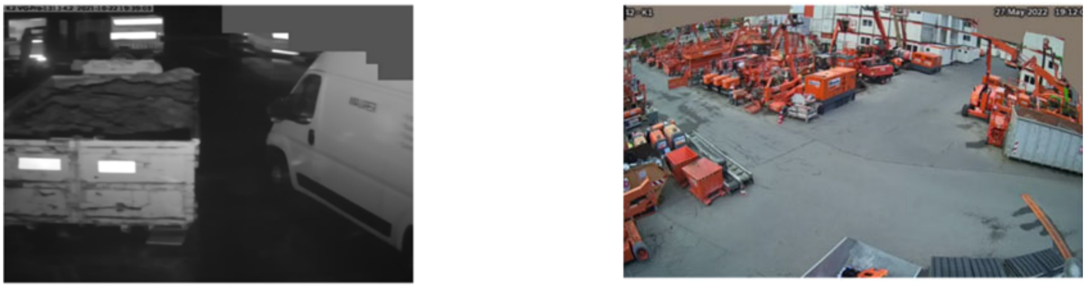
Figure 2
Block diagram of the proposed method



background information to obtain foreground information. The results are shown in Figure 3(b) for both shaky and non-shaky camera frames, where one can see moving pixels in the case of shaky camera frame while there are no moving pixels in the case of non-shaky camera frame. For estimating the number of moving and static edges, the method finds the difference between Canny of input frames and the Canny of foreground images. Sample Canny edge images of input of shaky and non-shaky camera frames can be seen in Figure 3(c). The ratio of the number of static edges divided by the total number of moving edges is

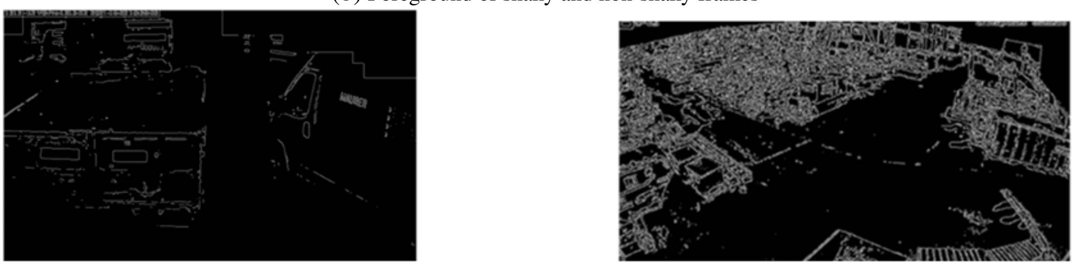
Figure 3
The steps for extracting features using foreground and background information

Shaky camera non-shaky camera



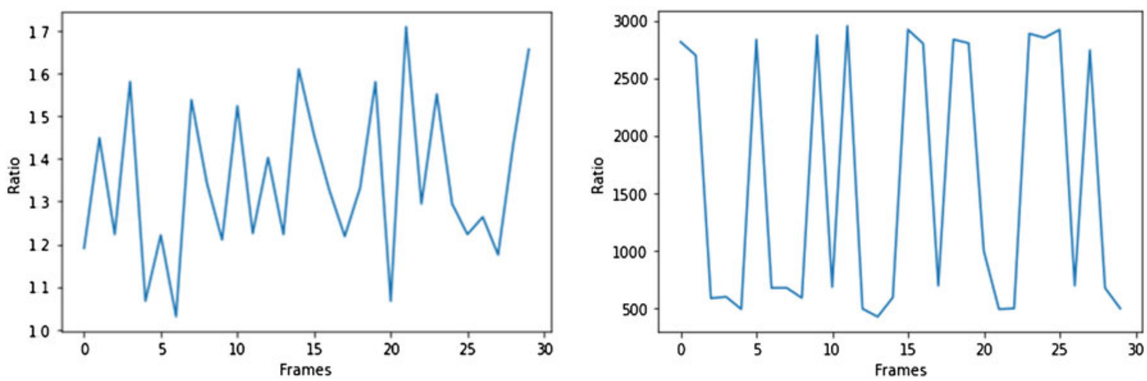
(a) The results of background model (median of 30 frames)

(b) Foreground of shaky and non-shaky frames



(c) Canny edge images of shaky frame-0 and non-shaky frame-0

Figure 4
The ratio features for every frame in input video of shaky (left side) and non-shaky (right side) camera



considered as a feature of the frame. This process outputs 30 features for 30 frames of video, which is considered as a feature vector for classification.

The effect of the ratio feature can be seen in Figure 4, where the ratio features are small for every frame in the case of shaky camera video while the ratio features are very large for every frame in the case of non-shaky camera video. This is the advantage of the ratio features for classification of video containing moving objects and static objects. For classification of video containing moving objects and static objects, the method is trained with samples provided by 10-fold cross-validation. At the same time, the same approach is used for choosing the testing samples. For the final classification, the method uses RF classifier, which considers the features corresponding to training and testing samples for classification. The reason to use RF classifier is that it is a well-known technique for handling imbalance feature vector sizes and if the feature vector contains noisy features.

The Random Forest is a uniquely designed decision tree created using randomly selected variables for a dataset extracted by bootstrap sampling; the classification is performed based on the majority of the decisions. Moreover, the contribution of each variable to data classification can be obtained using the created decision trees; furthermore, the importance of each variable can be determined. A combination of classifier trees represents RF classifiers, one of the finest approaches to represent input variables in the form of trees that makes a forest-like structure. Input data are represented in trees, and each tree specifies a class label. RF depends on its error rate. Error rate signifies two directions. First one is the correlation between trees, and the second one is the strength of the tree.

4. Experimental Results

Since there is no standard dataset for experimentation, we construct our own dataset for evaluating the proposed method. Our dataset consists of a total 2959 videos, of which 817 shaky camera samples and 2142 non-shaky camera samples. Our dataset includes outlets of industrial areas and factories where material and goods can be parked in open environments and a huge warehouse. The videos of 1–5 s are captured by CCTV camera mounted on the roof if it is indoor and poles if it is outdoor scenes. Since videos are captured at different time, situation, and weather conditions, our dataset includes diversified video, such as video with low, high quality, degradation, distortion, and can have any object including human and vehicles. If it is indoor video, it suffers from poor resolution and quality. If it is outdoor video, it suffers from external factors of weather conditions and illumination effect of lights. In the case of outdoor video, the presence of trees and leaves makes the problem much more complex. In summary, our created dataset is complex and challenging compared to normal object detection and tracing video.

It is noted that the size of classes is not balanced, and hence, the method generates some synthetic points to equalize both classes using SMOTE. As a result, our dataset consists of total 4284 points, of which 1325 synthetically generated points of minority class. Furthermore, the dataset includes the video captured in day and night of protected and sensitive areas which include indoor and outdoor scenes. Therefore, the dataset is complex and challenging for classification of arbitrary moving video and static video.

To show effectiveness of the proposed method, we implemented three state-of-the-art methods, namely Boufares et al. (2021), Wang et al. (2021) that use temporal difference and OTSU thresholding for moving object detection in the video, and Rahiminezhad et al. (2022) which propose subtraction approach for moving object detection in video, for comparative study. The reason to choose the above methods is that the objective of the methods is the same as the proposed method. In addition, the methods focus on detecting moving and non-moving objects for classification of video, which is similar to the idea proposed in this work.

For evaluating performance of the proposed and existing methods, we consider the following standard measures: precision, recall, F1 score, and accuracy.

Accuracy: In a given dataset consisting of (TP+TN) data points, the accuracy is equal to the ratio of total correct predictions (TP + TN + FP + FN) by the classifier to the total data points. The model's accuracy can be calculated as defined in equation (1).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad 0.0 < \text{Accuracy} < 1.0 \tag{1}$$

where TP is the true positive; TN is the true negative; FP is the false positive; and FN is the false negative.

Precision: This is equal to the ratio of the true positive (TP) samples to the sum of true positive (TP) and false positive (FP) samples, which are defined as in equation (2).

$$\text{Precision} = \frac{TP}{\text{TP} + \text{FP}} \tag{2}$$

Recall: Recall is the evaluation metrics equal to the ratio of the TP data samples to the sum of TP and FN data samples, which are defined as in equation (3).

$$\text{Recall} = \frac{TP}{\text{TP} + \text{FN}} \tag{3}$$

F1 score: F1 score is equal to the harmonic mean of recall value and precision value. The F1 score gives the perfect balance between precision and recall, thereby providing a correct evaluation of the model's performance. F1 score can be calculated as defined in equation (4).

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4.1. Ablation study

In this work, we used the RF classifier for classification of arbitrary moving video and static video. To test the contribution of the RF classifier, we compare the performance of the proposed method-random classifier with other well-known classifiers. For this experiment, the proposed work calculates accuracy for replacing the RF classifier with the following classifiers (Chittroa et al., 2021; Solanki et al., 2021) on our dataset, and the results are reported in Table 1.

Decision tree classifier: A Decision Tree is a tree-structured classifier represented as follows: the internal nodes represent the features of a dataset, branches represent the decision rules, and each leaf node represents the outcome. There are two nodes in the decision trees: decision node and leaf node, Decision nodes are

Table 1
Accuracy for the proposed method with different classifiers

Classifiers	Random forest	Decision tree	SVC	Logistic regression	Gradient boosting	KNN
Accuracy	0.747236	0.646782	0.337555	0.230515	0.612973	0.679468

used to make any decision and can have multiple branches, whereas leaf nodes are the output of those decisions and do not contain any further branches. The decisions or the test are performed on the basis of features of the given dataset and are a graphical representation to get all the possible solutions to a problem based on given conditions.

Gradient boosting classifier: It is one of the boosting algorithms used to minimize bias error of the model. Gradient boosting algorithms can be used for predicting not only continuous target variables (as a regressor) but also categorical target variables (as a classifier). When it is used as a regressor, the cost function is mean square error, and when it is used as a classifier, then the cost function is log loss.

K-Nearest Neighbor(KNN): K-NN classifier comes under a lazy learning process in which training and testing can be realized on the same data. In the process, the data of interest are retrieved and analyzed depending upon the majority value of the class label assigned as per k, where k is an integer. The value of k is based on the distance calculation process. The choice of k depends on data. Larger value of k minimizes the noise on classification. Similarly, parameter selection is also a prominent technique to improve the accuracy in classification.

Support vector classifier: Support vector classifier (SVC) is usually preferred for data analysis because of its computational capability within a very less time frame. This classifier works on the decision boundary concept recognized as hyperplane. The hyperplane is used to classify the input data into the required target group. SVC is not affected with over fitting problems and makes it more reliable.

Logistic regression classifier: This classifier is based on the probability of outcome from the input process data. Binary logistic regression is generally preferred in machine learning techniques for dealing with binary input variables. To categorize the class into specific category, a sigmoid function is utilized.

It is noted from Table 1 that the proposed method with RF classifier reports the best accuracy compared to all other classifiers. Therefore, one can infer that the proposed method with RF classifier is suitable for this work. When we compare the results of the proposed method with other classifiers, decision tree and KNN approaches are better than other classifiers. This is because the decision tree and KNN have the ability to cope with the imbalanced feature vectors, and it avoids the overfitting problem. However, other methods are good when the data are simple but not for the nonlinear data.

4.2. Experiments on classification of arbitrarily moving objects

Quantitative results of the proposed and existing methods on our dataset are reported in Table 2, where it is noted that the proposed method is better than two existing methods in terms of accuracy. The reason for achieving the best accuracy by the proposed method is the step of separating background and foreground regions in the images; edge-based features extraction and use of RF classifiers are generalized steps and hence work well for diversified videos. But the existing methods are developed for the day images; the methods do not work well for the night images. Therefore, both the existing methods report poor results compared to the proposed method.

5. Conclusion and Future Work

For classification of arbitrary moving video, we have proposed a new method based on background and foreground separation, edge-based ratio features, and the RF classifier. For defining background in the video frames, the proposed work uses a median of 30 frames as background for the input video. The frames are subtracted with the median image to obtain foreground information. The number of static and moving edges is computed using Canny of the input frames and Canny of foreground images of the respective frames. Similarly, for classification, the proposed work uses 10-fold cross-validation, which provides the number of training and testing samples. For classification, the proposed work uses the RF classifier by feeding the ratio feature as input. Experimental results on our dataset and the comparative study with the existing methods show that the proposed method outperforms the existing methods in terms of accuracy. However, when the input video contains the static, arbitrary moving objects and uniform moving objects, the performance of the proposed work degrades. This will be a three-class classification problem and beyond the scope of the work. We plan to address such limitation in near future by exploring randomness of direction, speed, deformable shape, and content of the video.

Conflicts of Interest

Palaiahnakote Shivakumara is an editor-in-chief and Umapada Pal is an advisory board member for *Artificial Intelligence and Applications*, and were not involved in the editorial review or the decision to publish this article. The authors declare that they have no conflicts of interest to this work.

References

Arjun, D., Indukala, P. K., & Menon, K. A. U. (2020). Development of a framework for effective surveillance and human intrusion detection in border regions covered with dry leaves. In *Proceedings of ICISC*, pp 106–109.

Bharadwaj, A., Dagar, V., Khan, M. O., Aggrawal, A., Alvarado, R., Kumar, M., Irfan, M., & Proshad, R. (2022). Smart IoT and machine learning-based framework for water quality assessment and device component monitoring. In *Environmental Science and Pollution Research*. Springer, pp 46018–46036.

Table 2
Performance of the proposed and existing methods for classification of arbitrary moving videos

Methods	Proposed						Boufares et al. (2021)						Wang et al. (2021)						Rahiminezhad et al. (2022)					
Classes	Shaky			Non-shaky			Shaky			Non-shaky			Shaky			Non-shaky			Shaky			Non-shaky		
Measures	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
Results	0.75	0.78	0.76	0.77	0.74	0.76	0.60	0.65	0.63	0.63	0.58	0.60	0.59	0.73	0.65	0.65	0.50	0.57	0.58	0.64	0.60	0.62	0.56	0.59
Accuracy	**0.76**						0.62						0.62						0.60					

Boufares, O., Boussif, M., & Aloui, N. (2021). Moving object detection system based on the modified temporal difference and OTSU algorithm. In *Proceedings of SSD,* pp 1378–1382.
Chandra, N., & Panda, S. P. (2021). A human intruder detection system for restricted sensitive areas. In *Proceedings of ICORT*.
Chhabra, M., Ravulakollu, K. K., Kumar, M., Sharma, A., & Nayyar, A. (2022). Improving automated latent fingerprint detection and segmentation using deep convolutional neural network. *Neural Computing and Applications*.
Chithaluru, P., Stephan, T., Kumar, M., & Nayyar, A. (2022). An enhanced energy-efficient fuzzy-based cognitive radio scheme for IoT. In *Neural Computing Applications*, *34*, 19193–19215.
Chittroa, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasiski, M., Jasinski, L., Gono, R., & Bolshev, V. (2021). Prediction of chronic kidney disease-A machine learning perspective. *IEEE Access*, *9*, 17312–17334.
Deng, Z., Cui, Z., & Cao, Z. (2022). Super resolution detection method of moving object based on optical image fusion with MMW radar. In *Proceedings of IGARSS*, pp. 1900–1903.
Huang, Y., Jiang, Q., & Qian, Y. (2021). A novel method for video moving object detection using improved independent component analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*, 2217–2230.
Kaushik, K., Bharadwaj, A., Kumar, M., Gupta, S. K., & Gupta, A. (2022). A novel machine learning-based framework for detecting fake Instagram profiles. Wiley, pp. 1–12.
Kim, H., Kim, P., & Kim, H. J. (2020). Moving object detection for visual odometry in a dynamic environment based on occlusion accumulation. In *Proceedings of ICRA*, pp. 8658–8664.
Kim, W. J., Hwang, W., Lee, J., Woo, S., & Lee, S. (2022). AIBM: Accurate and instant background modelling for moving object detection. *IEEE Transactions on Intelligent Transportation Systems*, *23*, 9021–9036.
Kovalenko, R., Tashilinskii, A., & Tsaryov, M. (2021). Using threshold processing to moving object detection in the image sequence. In *Proceedings of ITNT*.
Lenac, K., Cuzzocrea, A., & Mumolo, E. (2021). A novel genetic scan-matching-based registration algorithm for supporting moving objects tracking effectively and efficiently. *IEEE Access*, *9*, 91741–91753.
Ramesh, B., Zhang, S., Yang, H., Ussa, A., Ong, M., Ochard, G., & Xiang, C. (2021). e-TLD: Event-based framework for dynamic object tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*, 3996–4006.
Sadkhan, A. S. B., Talebiyan, S. R., & Farzaneh, N. (2021). An investigate on moving object tracking and detection in image. In *Proceedings of BICITS*, pp. 69–75.
Shu, Y., Sui, Y., Zhao, S., Cheng, Z., & Liu, W. (2021). Small moving object detection and tracking based on event signals. In *Proceedings of ICCC*, pp. 792–796.
Solanki, Y. D., Chakrabarti, P., Jasinski, M., Leonowicz, Z., Bolshev, V., Vinogradov, A., Jasinska, E., Gono, R., & Nami, M. (2021). A hybrid supervised machine learning classifier system for breast cancer prognosis using feature selection and data imbalance handling approaches. *MDPI, Electronics*, pp. 1–16.
Sultana, M., Mahmood, A., & Jung, S. K. (2021). Unsupervised moving object detection in complex scenes using adversarial regularization. *IEEE Transactions on Multimedia*, 2005–2018.
Tang, M., & Liu, W. (2022). A moving object detection algorithm for removing Ghosting and shadow. In *Proceedings of CCC*, pp. 7207–7212.
Rahiminezhad, A., Tavakoli, M. R., & Sayedi, S. M. (2022). Hardware implementation of moving object detection using adaptive coefficient in performing background subtraction algorithm. In *Proceedings of MVIP*.
Vndana, G. S., Paradhasaradhi, B., & Srihari, P. (2022). Intruder detection and tracking using 77GHz FMCW radar and camera data. In *Proceedings of CONECCT*.
Wang, J., Zhao, Y., Zhang, K., Wang, Q., & Li, X. (2022). Spatio-temporal online matrix factorization for multi-scale moving object detection. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*, 743–757.
Wang, Z., Wang, J., & Wang, N. (2021). Moving object detection and marking based on frame difference and train algorithm for teaching video. In *Proceedings of ASID*.

How to Cite: Asadzadehkaljahi, M., Halder, A., Pal, U., & Shivakumara, P. (2023). Spatiotemporal Edges for Arbitrarily Moving Video Classification in Protected and Sensitive Scenes. *Artificial Intelligence and Applications* https://doi.org/10.47852/bonviewAIA3202526