

RESEARCH

Open Access



# Integration of new moving object segmentation and classification techniques using optimal salp swarm-based feature fusion with linear multi k-SVM classifier

Jemilda G.<sup>1</sup>  and Baulkani S.<sup>2\*</sup>

\* Correspondence: [ramabaulkani@yahoo.co.in](mailto:ramabaulkani@yahoo.co.in)

<sup>2</sup>Department of Electronics and Communication Engineering, Government College of Engineering, Tirunelveli, India  
Full list of author information is available at the end of the article

## Abstract

The feature extraction technique is applied on least enclosing rectangle (LER) of the segmented object to increase the processing speed. The main intuition of this salp swarm algorithm relays on reducing the computational load of the proposed classifier by removing the repetitive and unrelated features from the feature vector. Also, increased training samples of similarly shaped classes when applied on the classifier can generate the misclassification results. Thus, a new layered kernel-based support vector machine (k-SVM) classifier is developed by means of integrating the k-neural network classifier and layered SVM classifier. Because of the high dimensional features, a difficulty occurs in the application of a single classifier. In order to ease the computational load, this multi classifier is integrated with a shadow elimination technique to classify the object categories of intelligent transportations system such as motorcycles, bicycles, cars, and pedestrians.

**Keywords:** Image segmentation, Object classification, Feature extraction, Subtraction techniques, Image detection

## 1 Introduction

In current trend, intelligent transport system has received more attention in the research and commerce area. Smarter transportation system is generated through minimizing crowding, accident, and injury. To enhance the reliability, efficiency, and safety of transportation subsystem further, an improved transportation system management technique was developed. Presently, the intelligent transportation system is managed effectively, using one of the key technologies called wireless traffic video surveillance system. However, tasks such as vehicle detection, vehicle tracking, vehicle classification, and vehicle recognition are considered to be significant factors in the design of efficient traffic video surveillance system [1–3]. The first step invoked to develop a traffic video surveillance system is that designing an automated vehicle detection process. Essentially, this can be achieved by extracting necessary details about moving vehicles and applying these details for correct classification and recognition. Traffic conditions can

be monitored and analyzed accurately by means of classifying moving objects into categories such as bicycles, motorcycles, pedestrians, and cars. Also, these categories support a lot in accurate analysis of traffic conditions and retrieval of an object from the video frames. In general, the performance of an overall classification system is affected with two significant factors such as feature extraction from candidate objects and the classifier model. From the past decades, many studies have been proposed for the detection of intelligent transportation application categories, namely, cars, bicycles, motorcycles, and pedestrians [4–10]. In the last decade, extensive research has been done on moving object detection and tracking. Difficulty in object tracking was observed during unstructured objects structure and cameras, scenes, sudden movements of objects, and quick object changes. However, detecting the object from the video sequence and also tracking the object remains a challenge for researches. Histograms of oriented gradients (HoG), local binary patterns (LBP), Haar-like features, and Haar wavelets are the common features included in [5]. In previous works of [11, 12] two sets of feature descriptors HoG-LBP combination incurs the goodness of each feature descriptor; hence, the detection performance is highly improved. However, computational loads in the classifier and feature dimension cost are increased with this HoG-LBP combination. In order to rectify all these issues, a new automatic moving object segmentation and classification system is proposed. This novel approach includes a new feature descriptor, feature selection using a new optimization algorithm, and a new layered k-SVM classifier incorporating the shadow elimination technique which reduces the complexity effectively.

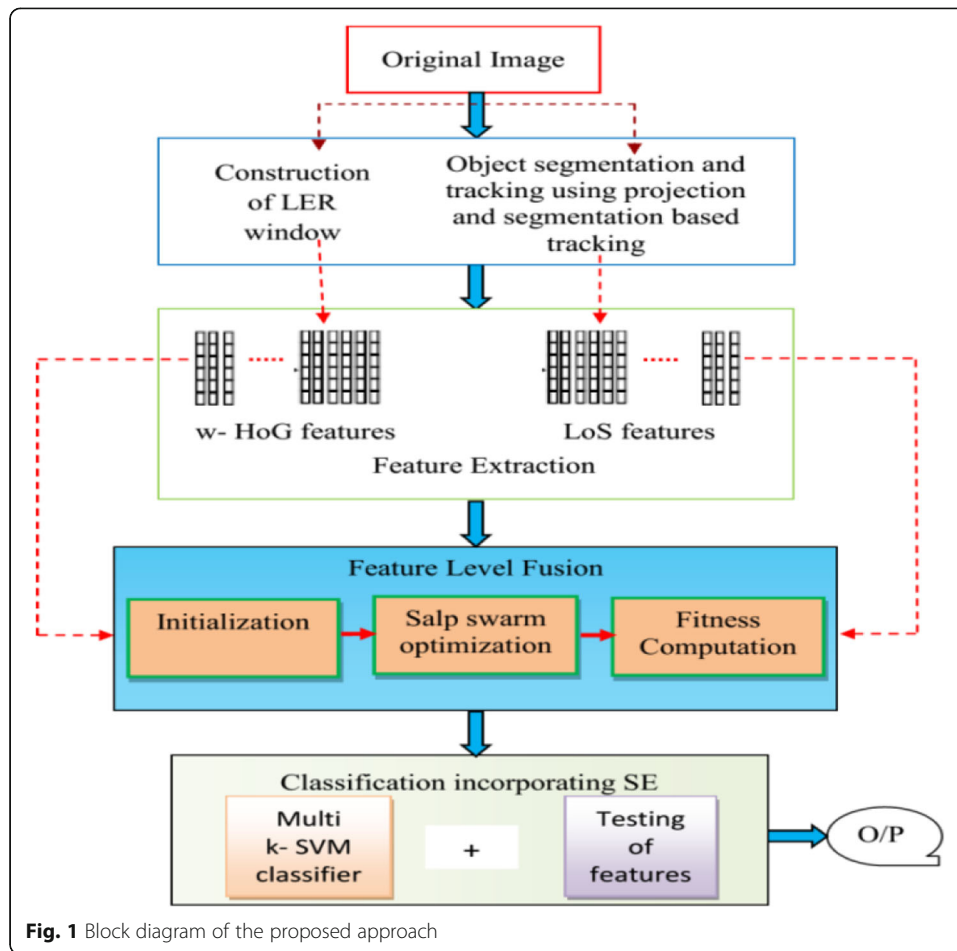
The purpose of this paper can be described in Section 1. The use of LER of a segmented object reduced the time consumption while on extracting the high dimensional feature descriptors such as the LBP and HoG. The use of a new layered k-SVM classifier and shadow elimination (SE) technique increases the classification accuracy. Applying the classification technique alone to a segmented object reduces the processing time and makes them feasible for performing real-time operations. Section 2 explains the proposed method including its design idea and practical implementation approach. Section 3 provides experimental results where the effectiveness of the proposed work is compared to the existing methods. Finally, Section 4 concludes this paper by summarizing our results, significance, and future possibilities of the work.

## 2 Methodology

To develop a new automatic moving object segmentation and classification system from the level-1 and level-2 sub bands, the local shape (LoS) and the HoG features are extracted. These extracted features are then fused at the feature-level fusion using salp swarm optimization (FFSSO) algorithm. For convenience, the fused features are now called w-LoSHoG descriptor hereafter. The proposed research work focuses on the construction of integrated moving object detection. Also, it is focused on the classification of system for better discrimination of real-time applications (i.e., intelligent transportation systems and human motion capture) is shown in Fig. 1.

### 2.1 Construction of LER window

Initially, the RGB color space incorporating the shadow elimination is considered to implement the proposed object segmentation technique. Five basic steps of this process



are as follows: At first, moving pixels are identified through determining the frame difference between the current and the previous frames. Secondly, the composing pixels are updated for the registered background regions. Thirdly, the moving objects from the background region are distinguished effectively by following the background difference calculation. Beyond the color-based modifications used in gray images, the initial three aforesaid steps of the proposed object segmentation technique also characterize the new function for registering a new object as a background region. Further, shadow effect of the segmented object is reduced in the fourth step. Ultimately, in the fifth step, vertical and horizontal histograms for the segmented image are determined to obtain the position of the LER window of an object. However, after a perfect segmentation of an object, its complete LER window is acquired. Subsequently, a tracking algorithm is employed to obtain the LER window of the moving object.

## 2.2 Preprocessing

In this work, to better distinguish the features among the four classes of moving objects (i.e., pedestrians, cars, bicycles, and motorcycles), a weight mask for a LER window is introduced.

### 2.3 Feature extraction

In the feature extraction step, local shape (LoS) and HoG features are extracted effectively.

### 2.4 Feature selection using BSSA

In this approach, all solutions are constrained to the binary values [0, 1]. Further, optimal features are selected from each video frame through defining a solution as a one-dimensional vector (i.e., each cell having 0 and 1 values). Based on the number of w-LoSHoG features in a video frame, the length of the vector is defined. Value 1 indicates that the feature is selected; otherwise, the feature is not selected with value 0. These selected optimal features are sent to the new layered k-SVM classifier for object classification.

### 2.5 Layered K-SVM classifier with SE technique

In order to classify four classes of the moving objects such as cars, bicycles, motorcycles, and pedestrians, a newly developed layered k-SVM classifier is employed. Further, two classification stages are introduced in this classification. Initially, in the two-wheeled objects class, the bicycles and motorcycles are assigned due to their shape similarity. Basically, the LER window of an object is resized to obtain an accurate feature dimension while classifying the objects having different sizes. To the width and length of the LER window, the scaling factor is applied; thus, the maximum size of a rescaled LER (RLER) window obtained is  $128 \times 62$ . In case the original window of an object satisfied this constraint, then it is not necessary to perform resizing. Next, if determined that the classified object in the RLER window belongs to a two-wheeled object class, then, again, classification is performed to distinguish the object into a motorcycle or bicycle. However, the SE technique is incorporated with this classifier to classify three classes of objects in the initial stage itself. In other words, the SE technique can reduce the shadow effects on the segmented object. In this section, SE technique is used as a clue to distinguish the moving objects as fast as possible, instead of applying SE technique for segmenting the moving object, before classification. For instance, large shadow areas are generated by the cars than the motorcycles, and based on the shadow effect, it is easier to identify the moving object whether it is a bicycle or motorcycle. Further, the proposed multi-SVM classifier is trained using  $2N$  training samples. The output generated by the proposed multi-SVM classifier is lower than zero on testing, and then the object is recognized as a bicycle; otherwise, it is classified as a motorcycle (i.e., output  $> 0$ ).

## 3 Results and discussions

This section detailed the experimental outcomes and performance analysis of the proposed approaches.

### 3.1 Experimental setup

The performance of the proposed approaches is tested using the objects segmented from four videos under various scenes. Implementation is done using MATLAB. The experimental results are evaluated and performance is analyzed using the parameters,

like true positive rate (TPR), false positive rate (FPR), precision (P), recall (R), and accuracy (A). The pixel values for the size of each captured image in the video are fixed to  $740 \times 480$ . However, if the length or width of the object in captured image is smaller than 15 pixels, then they are difficult to distinguish. Furthermore, in this work, it is important to perform feature extraction process using the pixels of the object in an image; therefore, the width/length of the object in captured image should be large enough. Also, the number of interested pixels  $\hat{m}$  in an LER window was fixed to 18 as minimum pixels. Ultimately, the performances of several conventional features are used to analyze the classification performance and dimensionality reduction of the proposed FFSSO optimization approach and multi k-SVM classifier.

### 3.2 Framework validation

In this work, the pedestrian, car, bicycle, and motorcycle classes include  $M$  number of training samples to train the proposed multi k-SVM classifier. The  $M$  number of training samples for each class was fixed to 2000. Figure 2 depicts the training samples collected for each class under different scenes.

Further, the LER windows of the moving objects and the scenes of the four test videos which are different from the training video are shown in Fig. 3a–d. The four videos are taken under the duration 1550 s, 1450 s, 1807 s, and 1365 s. Different backgrounds can be observed in these four videos and moving objects are captured from side-to-side view of the image. Using the proposed segmentation approach, the number of segmented objects selected from these four videos were 1323, 1244, 988, and 300, respectively. The proposed classification approach is tested using these segmented objects. Using the update and background registration step, the background registered in different frames is depicted. In case the object remains stable for a certain period of time, then their background is registered as a new background.

Using the frame difference technique, the movement of the object is identified as soon as the object starts moving. Then, the background is registered as a new background region. It is possible to identify that the two different objects, namely, pedestrian and bicycle are entering into the scene at the same time. In this case, the bicycle is occluded by a pedestrian. Initially, a pedestrian alone is covered by the LER window. However, when segmentation approach is used, the bicycle is segmented soon after the occlusion vanished and also a new LER window is created (indicated in second and third columns).

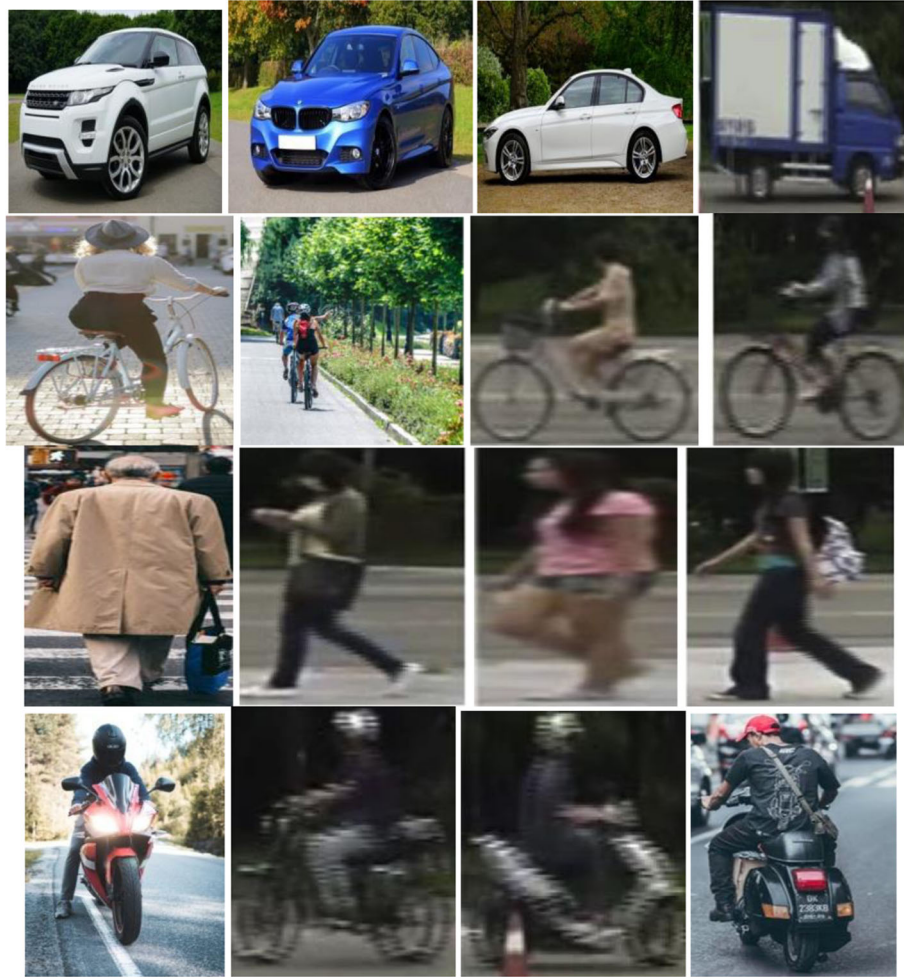
### 3.3 Evaluation metrics

In order to reveal the performance of proposed approaches, the evaluation metrics such as true positive rate (TPR), false positive rate (FPR), precision (P), recall (R), and accuracy (A) were adopted and they are defined in Eqs. 1, 2, 3, 4 and 5.

$$\text{False positive rate (FPR)} = \frac{\text{FP}}{(\text{TN} + \text{FP})} \quad (1)$$

$$\text{True positive rate (TPR)} = \frac{\text{TP}}{(\text{FN} + \text{TP})} \quad (2)$$

$$\text{Precision (P)} = \frac{\text{TP}}{(\text{FP} + \text{TP})} \quad (3)$$



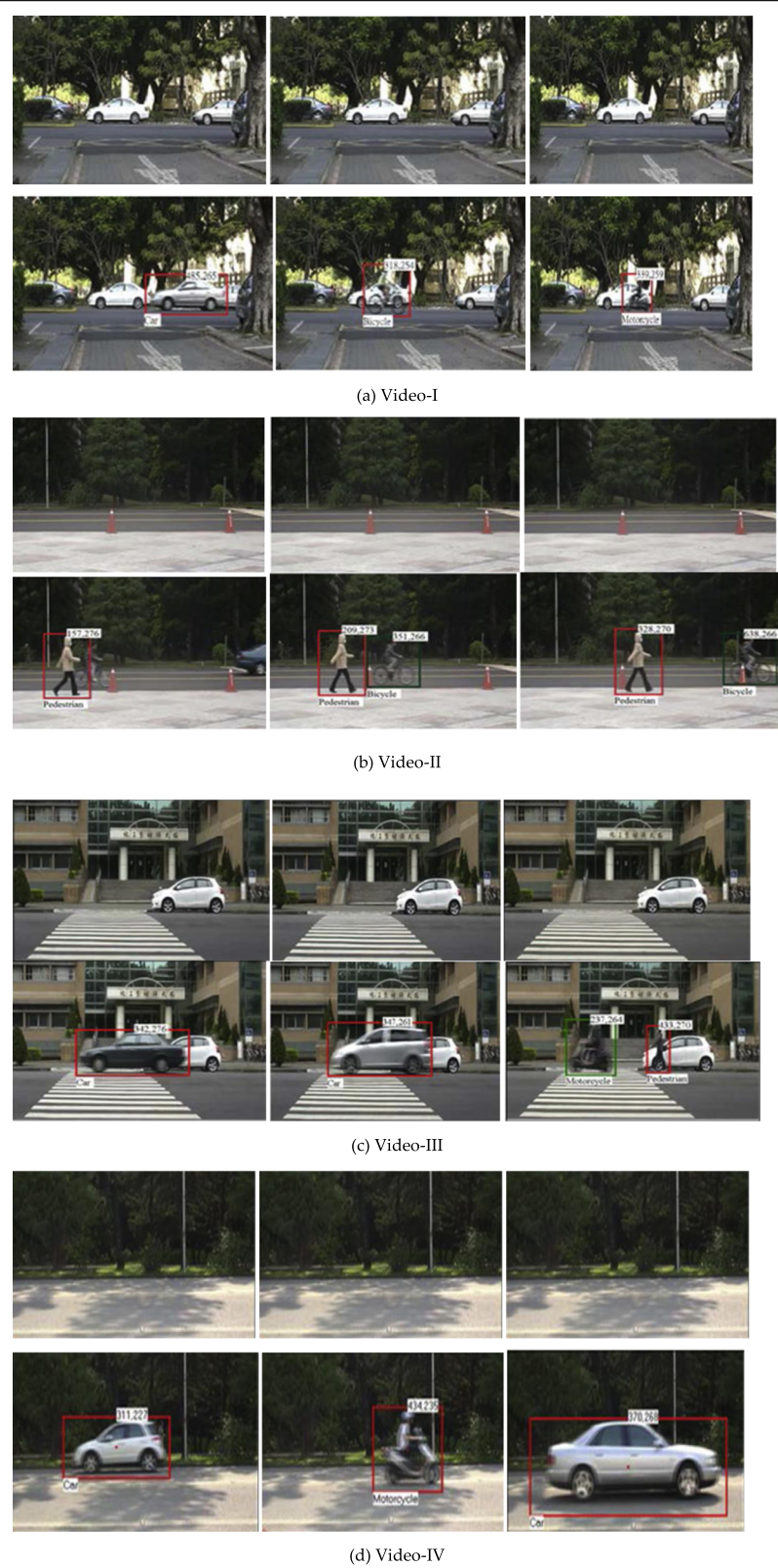
**Fig. 2** Training samples collected for each class in different scenes

$$\text{Recall (R)} = \frac{\text{TP}}{(\text{FN} + \text{TP})} \quad (4)$$

$$\text{Accuracy (A)} = \frac{\text{TN} + \text{TP}}{(\text{TN} + \text{TP} + \text{FP} + \text{FN})} \quad (5)$$

Here, TP indicates the total number of true positive pixels, TN denotes the total number of negative pixels, FP indicates the total number of false positive pixels, and FN denotes the total number of false negative pixels and so on. Precision defines the percentage of all identified pixels corresponding to the moving object. Recall defines the percentage of all pixels corresponding to moving object which is correctly identified. Accuracy defines the percentage of all pixels in RLER window which is correctly rejected and detected. To accurately detect the objects in the background, the value of precision, recall, TPR, and accuracy should be high and at the same time, the value of FPR should be low.





**Fig. 3** Classification results of test videos. **a** Video-I. **b** Video-II. **c** Video-III. **d** Video-IV

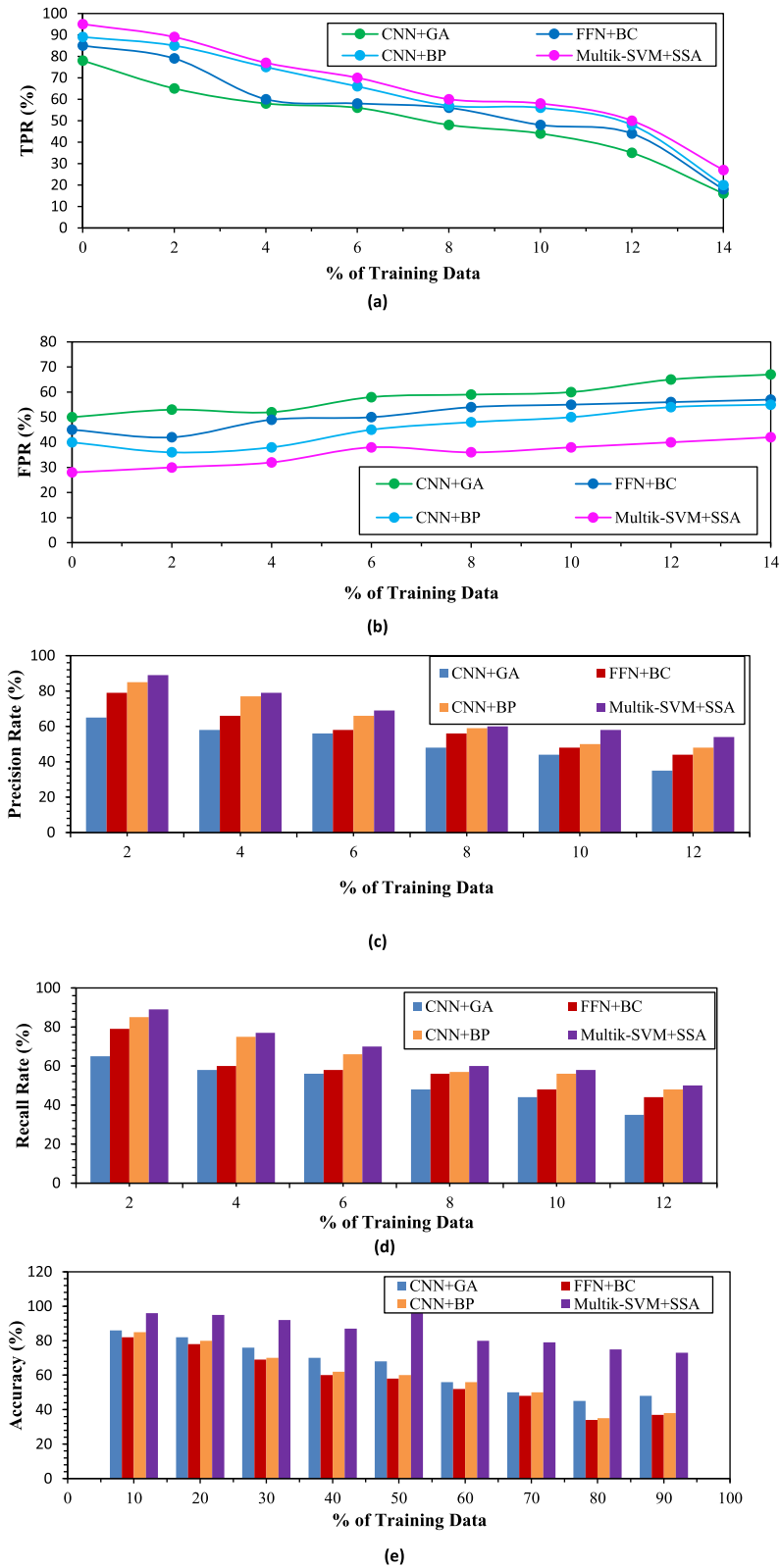


Fig. 4 a–e Performance of classifiers



### 3.4 Comparison with conventional classifiers

Here, TP indicates The performance of automatic moving object classification system (i.e., multi k-SVM with salp swarm algorithm (multi k-SVM + SSA)) is analyzed by comparing the efficiency of hybrid classifiers such as convolutional neural network and genetic algorithm (CNN + GA), feed-forward neural with Bayesian classifier (FFN + BC), and conventional neural network with back propagation algorithm (CNN + BP). The performance of the proposed system is analyzed by increasing the training data. Figure 4a–e depicts the TPR, FPR, precision (P), recall (R), and accuracy (A) of the proposed system on moving object classification.

For increasing the training data, the proposed classification system achieved better performance than the other hybrid classifiers. This goodness is observed because the multi k-SVM classifier is developed by integrating the k-NN and the SVM classifiers. For achieving multi-classification, k-NN classifier is the best choice because it performs classification wholly based on the distance among the training data and test sample. Further, in this work, the high dimensional features are extracted using the newly developed w-LoSHoG feature descriptor where the SVM classifier has the ability to behave better on the high dimensional data. Due to this advantage, the two classifiers namely k-NN and SVM are integrated to develop the layered k-SVM classifier. Also, the SE technique is incorporated with this developed classifier to avoid misclassification of the similar training samples having similar images.

## 4 Conclusion

In this module, an effective moving object segmentation and classification approaches were presented. Initially, the projection-based segmentation method was proposed for object segmentation. The LoS and HoG features are extracted from the segmented object using Haar DWT feature extraction process. A new feature descriptor called w-LoSHoG was developed by FFSSO optimization approach. The salp swarm algorithm (SSA) was imposed to find an optimal weight score to fuse the extracted LoS and HoG features; hence, the dimensionality issue and increase in processing time was gradually decreased. Finally, a new multi k-SVM classifier was developed by means of integrating the k-neural network classifier and layered SVM classifier. In order to ease the computational load, this multi classifier has been developed to classify the object categories of intelligent transportation systems such as motorcycles, bicycles, cars, and pedestrians. The experimental results proved the effectiveness of the proposed methods when compared to other existing conventional single and hybrid classifiers in terms of TPR, FPR, precision rate, recall rate, and accuracy

As a future work, degradation of video frames can be reduced by means of using improved lossless video surveillance techniques. Moreover, instead of doing classification with large objects, small-sized objects and its shadow can be applied as an input for classification.

### Abbreviations

BC: Bayesian classifier; BP: Back propagation; CNN: Convolutional neural network; FFSSO: Feature-level fusion using salp swarm optimization; GA: Genetic algorithm; HOG: Histogram of oriented gradients; K-SVM : Kernel-based support vector machine; LBP: Local binary patterns; LER: Least enclosing rectangle; LoS: Local shape; MATLAB: Matrix laboratory; NN: Nearest neighbor; RGB: Red-green-blue; RLER: Rescaled least enclosing rectangle; SE: Shadow elimination; SVM: Support vector machine

### Acknowledgements

We would like to acknowledge our sincere thanks to the Management of our Colleges, our Colleagues, Research Scholars of Anna University, Chennai, and our family members who have supported and helped us in different stages of this Research work.

### Authors' contributions

The authors have tried a new automatic moving object segmentation and classification system from the level-1 and level-2 sub bands, the local shape (LoS) and the HoG features are extracted. The authors read and approved the final manuscript.

### Authors' information

G. Jemilda received the B.E degree in Computer Science and Engineering from Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, in 1999 and the M.Tech. degree in Computer Science and Engineering from Dr. M.G.R. Educational and Research Institute, Chennai, in 2006. She has been working as a faculty in Computer Science and Engineering in Jayaraj Annapackiam CSI College of Engineering, Nazareth, since January 2007. She is currently a research scholar in Information and Communication Engineering in Anna University, Chennai. Her research interests include image processing, data structure, and mobile computing. She has published many papers in reputed journals. Dr. S. Baulkani received the B.E degree in electronics and communication engineering from Madurai Kamaraj University, Madurai and M.E. degree in computer science and engineering from Bharathiyar University, Coimbatore, in 1986 and 1998, respectively. She received her Ph.D. degree in Information and Communication Engineering from Anna University, Chennai, in 2009. Presently, she is working as an associate professor in the Department of Electronics and Communication Engineering, Government College of Engineering, Tirunelveli, Tamil Nadu, India. Her areas of interest are digital image processing, network security, web mining, and soft computing. She has published many papers in reputed journals and conferences. Under her guidance, 3 members are awarded with Ph.D. and many of them are pursuing Ph.D.

### Funding

No funding agency but the authors' own money.

### Availability of data and materials

The authors' own data and materials are used and no third party material is used.

<https://motchallenge.net/>

### Competing interests

The authors do not have any competing interests.

### Author details

<sup>1</sup>Department of Computer Science and Engineering, Jayaraj Annapackiam CSI College of Engineering, Nazareth, India.

<sup>2</sup>Department of Electronics and Communication Engineering, Government College of Engineering, Tirunelveli, India.

Received: 14 November 2019 Accepted: 12 May 2020

Published online: 25 May 2020

### References

1. W. Pawlus, H.R. Karimi, K.G. Robbersmyr, Data-based modeling of vehicle collisions by nonlinear autoregressive model and feed forward neural network. *Inf. Sci.* **235**, 65–79 (2013)
2. B.H. Chen, S.C. Huang, A novel moving vehicles extraction algorithm over wireless internet. *IEEE Int Conf Sys Man Cybernetics (SMC)*, 2505–2509 (2012)
3. Y. Zhang, X. Du, Automatic field data analyzer for closed-loop vehicle design. *Inf. Sci.* **259**, 321–334 (2014)
4. Q. Wang, T. Shuangshuo, Z. Dingding, X. Hu, Saliency based object tracking in complex scenes. *Neurocomputing*. **314**(7), 132–142 (2018)
5. A. Mondal, A. Ghosh, S. Ghosh, Scaled and oriented object tracking using ensemble of multilayer perceptrons. *Appl. Soft Comput.* **73**, 1081–1094 (2018)
6. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge. *Int J Comp Vision*. **115**, 211–252 (2015)
7. Q. Ye, J. Liang, J. Jiao, Pedestrian detection in video image via error correcting output code classification of manifold subclasses. *IEEE Transac Intel Transoratat. Syst.* **13**(1), 63–71 (2012)
8. Om Prakash, Manish Khare, Chandra Mani, Alok K. Singh "Moving object tracking in video sequences based on energy daubechies complex wavelet transform", National Conference on Communication Technologies & its impact on Next Generation Computing CTNGC, 2012.
9. G. Jemilda, S. Baulkani, Moving object detection and tracking using genetic algorithm enabled extreme learning machine. *Int J Comp Commu Control*. **13**(2), 161–173 (2018)
10. Ah.E. Hegazy, M.A. Makhoul, Gh S El-Tawel, "Improved salp swarm algorithm for feature selection", Journal of King Saud University – Computer and Information Sciences, 2018, pp. 1–10.
11. P. Jing, Y. Su, L. Nie, H. Gu, J. Liu, M. Wang, A framework of joint low-rank and sparse regression for image memorability prediction. *IEEE Transac Circuits Syst Video Technol.* **29**(5), 1296–1309 (2019)
12. S. Mirjalili, A.H. Gandomi, S.Z. Mirjalili, S. Saremi, H. Faris, S.M. Mirjalili, Salpswarm algorithm: a bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **114**, 163–191 (2017)

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.