ORIGINAL RESEARCH

Intelligent traffic analysis system for Indian road conditions

Balaji Ganesh Rajagopal[1]

Abstract Now-a-days, security and surveillance has become an integral part of our everyday life. There is a need for an Intelligent Surveillance systems for the being developed Smart Cities to ensure safety at all levels. The Objective of this paper is to demonstrate an Integrated framework for Vehicle detection and classification from real-time video captured from the road traffic. This work proposed a complete framework for Surveillance System for Indian smart cities with an aim to improve the security and surveillance of vehicles in varying weather conditions. To realize road traffic flow surveillance under various environments which contain poor visibility conditions, this paper provides a solution to extract the required information from surveillance video under different weather condition like day, night and rain. Also proposed system will dynamically choose the respective algorithm based on identified nature of the weather. In Vehicle count and classification, algorithm which is used based on image segmentation using a Laplacian of Gaussian edge detector (LoG), morphological filtering of the edge map objects and classification into small, medium and large vehicles on the basis of size using a nearest centroid minimum distance classifier. The proposed approach can be used for both stationary and fast moving traffic in contrast to motion detection based approaches. The algorithm was implemented in Python and average detection and classification accuracies of 96.0% and 89.4% respectively were achieved for fast moving traffic, while for slow moving traffic, 82.1% and 83.8% respectively were achieved.

1 Introduction

Traffic surveillance is a fast growing area of research. There are many methods, which address different problems in traffic surveillance. There are references dealing not only with traffic surveillance, but also including considerations about weather conditions. It was found that the cost of maintenance and installation of the detectors could become uneven.

Cutting-edge studies of traffic surveillance deal with components of intelligent surveillance systems and high-level modules of computer vision. However, advances have also been made in low-level processing. The proposed system consists of several modules, which can be grouped into low-level and high-level. At high level, symbolic reasoning is introduced. At low level, two groups of modules are used: for daylight and nightlight. During the day, motion detection must be corrected with luminance variation detection. At night, morphology and headlight pairing are used.

This paper present about road traffic surveillance experiments performed by using few of the benchmark videos which was used in many papers as standard experiment videos for having different weather condition, making it reasonable and mandatory to examine and investigate the impact of weather conditions on the accuracy of traffic surveillance systems.

The efficiency of a video surveillance detection system depends on not only the accuracy of the sampled information but also the coverage over the road surface. On the

✉ Balaji Ganesh Rajagopal
  rbalaji@nitt.edu

[1] Department of Computer Applications, National Institute of Technology, Trichy, India

1734

Int. j. inf. tecnol. (June 2022) 14(4):1733–1745

other hand, the use of the camera at a different angle doesn't a perfect choice to get rid of climatic condition and it incurs high cost both in implementation and maintenance. The major aspects of Traffic surveillance systems are.

   (i)   Environment identification
   (ii)   Vehicle detection
   (iii) Vehicle classification
   (iv) Vehicle count

Instead of detectors at frequent intervals, a high-resolution video camera which captures the image continuously can result in the traffic scenario at all road sites including junction points. The algorithm analyses the continuously captured image and outputs the traffic situation by comparing the relative motion of the vehicles in the first image with the second successive image. This method reduces the cost of hardware components. Therefore, specialized algorithms need to be proposed by combining various algorithms in literature and to deal with various sub problems like object segmentation, vehicle classification, Environment Identification etc.

Big data concepts for storing and retrieving vehicle data along with its semantic features were described in [1, 2]. The system can store and retrieve data up to 1 TB of vehicle information. The system also categories vehicles based on its semantic feature for vehicle classification. Combining vehicle surveillance along with semantic features annotates related vehicle retrieving and classifications.

2 Related works

Recent development in video surveillance deals with an advanced system for vehicle identification and classification based on different climatic condition. The work initiates road surveillance system using single camera that classifies vehicle-based on linearity feature. Later, the works in literature deal with the concepts of traffic density estimation, vehicle speed, and vehicle count, Traffic rules violation monitoring (lane change, red light violation etc.).

In order to perform road traffic video surveillance and a dedicated video classification algorithm for Indian roads is required. Most of the existing algorithms [3–5] are divided into three parts. The first part addresses the problem and solutions in vehicle detection after background subtraction from the video frames, the second part describes vehicle tracking and the third part of the chapter deals with classification of vehicles into small, medium and large vehicles.

The system proposed in [6, 7], separates out luminance and chrominance from each other on the RGB colour space

by generating a new colour model. Thus, the pixel value should be kept in an expected chromaticity line.

Both colour and edge information were used [8] to perform the foreground segmentation. With the subtraction of the incoming current frame, $I^t(\varphi)$, for each channel, confidence maps were generated for both colour and edge information. Finally, a hysteresis threshold was used to obtain the foreground mask.

Detection and tracking of the vehicles [9–11] based on the similarity of the pixels (particles) in the vehicle images. The particles are initially grouped according to their spatial positions and motion vectors. The shapes of the particle groups are analysed to detect convex groups (vehicles), and the detected vehicles are tracked using histogram similarity. Finally, the vehicles are counted when they intersect the user-defined virtual loops.

The vehicle detection method [12] based on multi-scale edge fusion was proposed. Multi-scale images are obtained from the decomposition of the DoG pyramid. In this work, experiments with traffic images in different weather conditions were verified for the practicability of the proposed method.

A vision-based daytime brake light detection system [13, 14] uses a driving video recorder was proposed. At daytime, visual features, motions, and appearances of vehicles were highly visible.

The adaptive model [15, 16] can assess the real-time vehicle counts on urban roads using computer vision technologies which proposed an automatic real-time background update algorithm for vehicle detection and an adaptive pattern for vehicle counting based on the virtual loop and detection line methods.

A vehicle detection based on vehicle light colour classification [17], was developed based on the concept of a vision system in reduced visibility conditions. The system utilized a self-adaptive stereo vision extractor of 3D edges or obstacle and a colour detection of vehicle lights.

In order to separate the foreground object with a visually plausible boundary, several bi-layer separation methods are available which assume that the camera is mostly stationary and the background is known or can be modelled.

In the stereo video sequences method [18, 19], the static background, the object colour, gradient, and displacement information are integrated to infer the foreground layer in real time.

If the camera undergoes arbitrary translational and rotational motions and the background has complex geometry structures, the foreground object cannot be accurately extracted due to the motion estimation and dynamic foreground definition.

In the bilayer segmentation [20–22], using the camera parameters, the multi-view geometry constraint is incorporated to generate the foreground object. A novel

appearance and structure consistency constraint in 3D warping is introduced to model the essential difference between the moving object and the background in the video. The foreground is extracted by solving an optimization problem combining all these constraints and considering the temporal-spatial smoothness in a video.

Having witnessed the success of web camera applications and the appearance of high definition digital video cameras Object tracking becomes easier. Unlike still images, video sequences provide more information [23–25] about how objects and scenarios change over time, but at the cost of increased space for storage and wider bandwidth for transmission.

3 Methodology

The different weather condition in video traffic surveillance was addressed and it has been found that the results percentage is increased by 1% from the observations that used FNDR and TFDR. The System mainly focuses on traffic prediction in various weather conditions with speed variation.

In the proposed system, the main focus to identify the environment from captured frames. Then, the system will automatically choose the best approach for vehicle detection and classification based on the identified nature of the weather.

Environment tracking identifies various environmental factors like a day, night, rainy, and snowy condition. The preprocessing phase deals with traffic surveillance for enhancing the video frame under surveillance. The next phase is vehicle identification and classification for classifying the vehicle based on its size like (small, medium and large vehicles). The final phase counts the vehicle and predicts the traffic load. The system consists of four major modules like environment identification, Preprocessing, the core of the system and output display.

3.1 Environment identification module

The captured clip is divided into a number of frames. In this work, the number of frames per second (FPS) is 35 FPS for the video dataset under observation. The day and night environment is identified based on the number of dark pixels and the mean value of the intensity value. Mean(), Sum() has been calculated to find the image matrix intensity average value. The Environment Identification Module is used to identify the type of the video as Daytime or Night-time and the process is shown in Fig. 1. A frame representing the type of video is extracted from the video. The top half of the Image is cropped as it mostly contains the pixels related to the Environment. The colour model
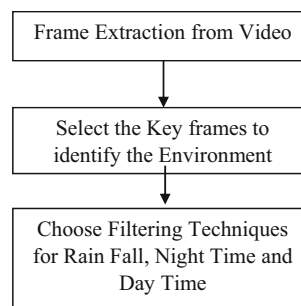

Fig. 1 Environment identification process

has been converted from RGB to HSV since luminance is the dominant feature to identify the type of video. The histogram is calculated for H and V components. It has been also experimented by considering the threshold value as 90 since most of the day image pixels' average value will be greater than 90. Hence frame environment will be identified as day if the threshold is greater than or equal to 90 and the image environment will be night if the threshold value is less than 90.

3.2 Night time image—hue, value histograms

It has been inferred from Figs. 2 and 3, that the histogram of Hue for the images has been accumulated with a PDF of less than 0.3, whereas the histogram for value for the images is accumulated with a PDF of more than 0.6. These measurements are indicative in nature and it is being experimented with more than 1000 video frames of different video sets and found that the Hue and Value Histograms are more inferential in identifying the Environment of the video frames. But the range of frequencies for the Histograms of Daytime and Night Time images are varying abruptly and it has been taken as a
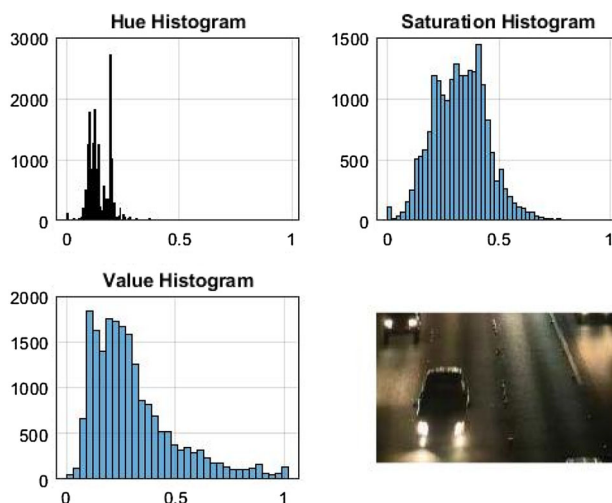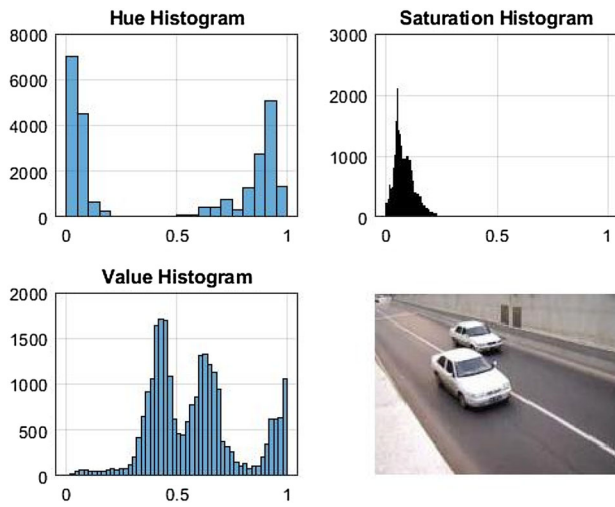

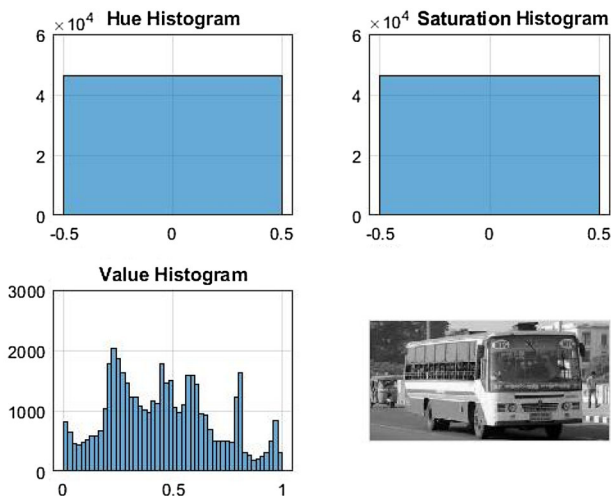Fig. 2 Hue and value histograms for the night-time image

Fig. 3 Hue and value histograms for the daytime image

Fig. 5 Nighttime capture of vehicles and their hue and value histograms

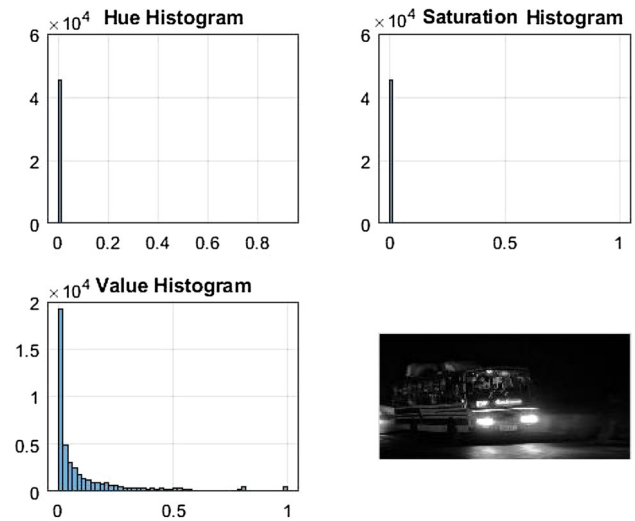quantitative measure to identify the type of Environment (Figs. 4, 5).

Also the daytime frames are taken into further processing to identify whether drizzles/rain particles are available or not; since night frames already have less information and having more dark pixels. Raindrop identification is done by using Edge based image segmentation like canny edge detection and Laplacian of Gaussian (LoG) has been applied, which was followed by Morphological Erosion and Closing.

3.3 Rain fall environment

Preprocessing the image and aim to remove visual effects of rain (i.e.) separate the rain layer and de-rained image layer from the rain image. The single frame image extraction uses a dictionary learning algorithm that focuses

on different weather condition. The mutual exclusive property is applied to image frames to remove patches that have large discriminative codes. Such separation of code implies nonlinear composition between two images for providing accurate results.

As an example, Figs. 6 and 7 are showing both frames with rain and rain particles removed frame.

3.4 Night vision environment

New incremental histogram equalization approach is developed in which the image histogram is equalized at each intensity level starting from the maximum intensity found in the original image up to 255, but in which, for each step, the histogram equalized output image of the previous step used as input. This new incremental approach is showed maximum PSNR values than the conventional

Fig. 4 Daytime capture of vehicles and their hue and value histograms

Fig. 6 Frame with raindrop

Fig. 7 Frame without raindrop

histogram equalization. This system is achieved good output image quality even for dark video frames enabling a viewer at the client's end to see under poor visibility conditions. The results of the new incremental histogram are shown for night frames in Figs. 8 and 9. It is evident that the dark spots of the night vision image have been improved which leads to detection of moving vehicles.

3.5 Day light environment

In this environment, assume that frames will be captured in well light intensity, hence there is no major preprocessing involved due to good weather condition.

3.5.1 Vehicle detection and classification using ROI modelling

The core system consists of two main stages: vehicle detection and vehicle classification. It combines selected image processing and computer vision algorithms to obtain the traffic density and uses a pattern classifier to classify the vehicles into small, medium and large classes. First, the collected color video data is broken down into constituent frames which are then converted into the grey scale to simplify and facilitate subsequent processing. The vehicles are then extracted from the video frames and their negatives using a Laplacian of Gaussian (LoG) edge detection method and mathematical morphology. The vehicles obtained are counted and it is used to calculate the traffic density as the number of vehicles per unit area of the road section at any given time. The size dimensions of the vehicles are also extracted and used as inputs to the classifiers at the classification stage.

The video frames are converted into the grey scale to simplify and facilitate subsequent processing. A negative of the image is calculated followed by Laplacian of Gaussian (LoG) edge detection method and mathematical morphology. The key stages of the core system and its algorithms are explained in Fig. 12.

3.5.2 Negative transformation

For each grey frame extracted from the video, its negative is computed as shown in Fig. 10. This is to ensure that as much relevant edge detail as possible is extracted in the segmentation stage, thus minimizing spurious edge discontinuities.

Fig. 8 Unprocessed night frames

1738

Int. j. inf. tecnol. (June 2022) 14(4):1733–1745

Fig. 9 Histogram preprocessed frames
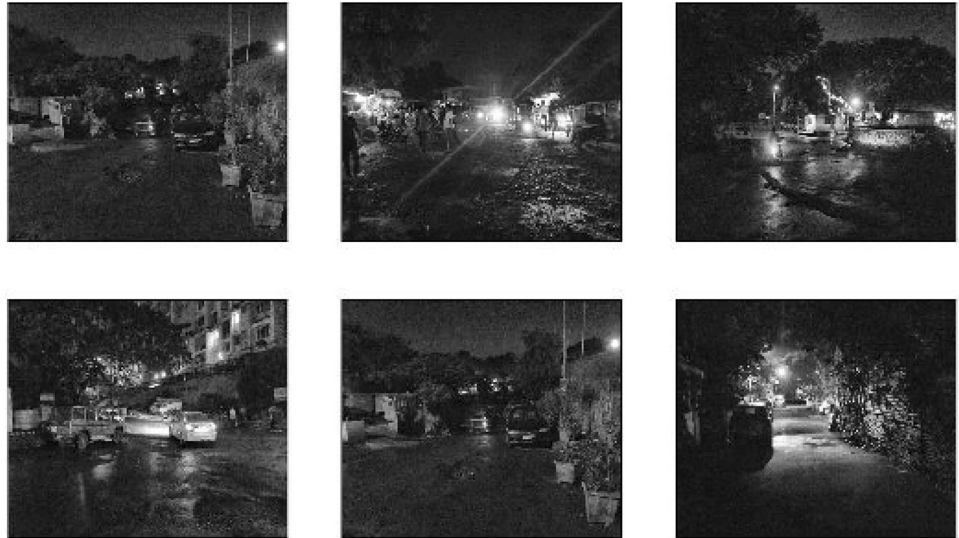
Fig. 10 Negative transformation for a given video frame

3.5.3 ROI mask modeling

One of the extracted frames is used to model a Region of Interest (ROI) polygon. This polygon is ultimately used to mask the processed binary frames so as to limit the counting and classification of vehicles to those found only within the region of interest. The size of this polygon is chosen empirically such that the vehicle intra-class variations are minimized. Figure 11 shows a traffic scene and the modeled ROI mask for the free-flowing traffic (Fig. 12).

3.5.4 Top-hat transformation

The segmentation performance is improved by compensating for non-uniform illumination of the scene using the morphological top-hat transformation prior to the segmentation stage. This is computed as given in Eq. 1

$$g(x,y) = f(x,y) - f(x,y)ob(x,y) \qquad (1)$$

where g(x, y) is the uniform background frame, f(x, y) is the input frame and $f(x, y)\, ob(x, y)$ is the morphological opening of $f(x, y)$ using a structuring element (SE), $ob(x, y)$.
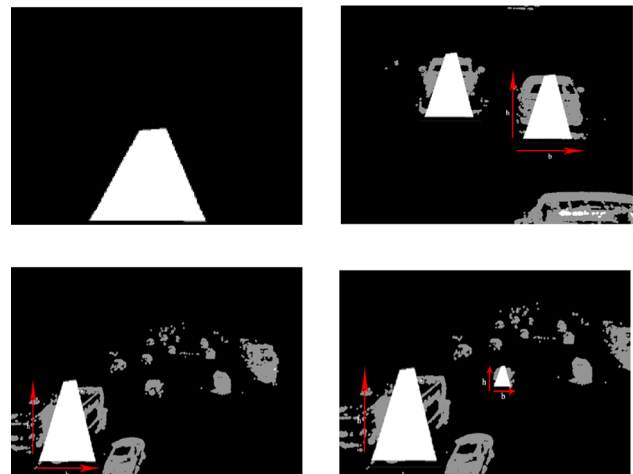
Fig. 11 ROI mask for vehicle tracking (small, medium and large vehicles)

The size of the structuring element is chosen such that it is larger than any object of interest in the scene so as to avoid deletion of any vehicle in the subtraction process. This transformation also helps to minimize the effects of shadows.

```
                              Input video
                                  │
                                  ▼
                    ┌─────────────────────────────┐
                    │ Frame extraction and Colour to│
                    │       Grey Conversion         │
                    └─────────────────────────────┘
                         │                    │
              ┌──────────┘                    └──────────┐
              ▼                                           ▼
     ┌──────────────────┐                    ┌──────────────────────┐
     │  Positive Frame  │                    │   Image Negative      │
     └──────────────────┘                    │   Transformation      │
         │          │                        └──────────────────────┘
    ┌────┘          └────┐                              │  Negative Frame
    ▼                    ▼                              ▼
┌──────────────┐  ┌──────────────┐            ┌──────────────────┐
│  ROI Mask    │  │ Pre-Processing│            │  Pre-Processing  │
│  Modelling   │  └──────────────┘            └──────────────────┘
└──────────────┘         │                              │
    │                    ▼                              ▼
    │            ┌──────────────┐            ┌──────────────────┐
    │            │ Segmentation │            │   Segmentation   │
    │            └──────────────┘            └──────────────────┘
    │                    │                              │
    │                    └──────────┬───────────────────┘
    │                               ▼
    │                    ┌──────────────────────┐
    │                    │        Summer         │
    │                    └──────────────────────┘
    │                               │  Binary Frame
    │                               ▼
    │                    ┌──────────────────────────┐
    │                    │ Post-Segmentation processing│
    │                    └──────────────────────────┘
    │   ROI Mask                    │
    │                               ▼
    └──────────────────►  ┌──────────────────────┐
                          │     Frame Masking     │
                          └──────────────────────┘
                                    │  ROI Frame
                                    ▼
                          ┌──────────────────────┐
                          │ Morphological Opening │
                          └──────────────────────┘
                                    │
                                    ▼
                          ┌──────────────────────────┐
                          │ Connected components and   │
                          │    feature extraction      │
                          └──────────────────────────┘
                              │                    │
                   ┌──────────┘                    └──────────┐
                   ▼                                           ▼
          ┌──────────────────┐                    ┌──────────────────┐
          │  Vehicle counter │                    │ Vehicle Classifier│
          └──────────────────┘                    └──────────────────┘
                   │                                           │
                   ▼                                           ▼
          ┌──────────────────────┐                   Vehicle class label
          │ Traffic delivery      │
          │    Estimator          │
          └──────────────────────┘
                   │
                   ▼
          No. of Vehicles/Area
```

Fig. 12 Flow graph for vehicle detection and vehicle class label identification

3.5.5 Image smoothing and blurring

The uniform background frame is smoothed using a median filter to remove random noise and then aggressively blurred using a Gaussian filter so as render 'noise' edges into the background and therefore reduce the chances of their detection. This also minimizes the effects of shadows in the traffic scene. Finally, the blurred frame is contrast enhanced linearly so as to emphasize the edges while preserving the mean intensity values of the frames using a contrast stretching algorithm prior to segmentation

3.5.6 Image segmentation

To extract objects in both the frame and its negative, the Laplacian of a Gaussian (LoG) edge detection method is used due to its excellent edge detection properties and relative simplicity. This preserves generality unlike the trial and error thresholds normally used in many of the reported approaches. The LoG of a two-dimensional image is computed as given in Eq. 2.

$$N^2G(x,y) = \left[\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}\right]e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

where ∇^2 is a Laplacian operating on the Gaussian smoothed image $G(x,y)$ and σ is the standard deviation of the image pixel intensities at point (x,y) (Fig. 13).

This approach enables the system to exploit the fact that shadows are semi-transparent and therefore appropriately enhancing and segmenting the frames and the effects of shadow can be greatly reduced. In this way, the complex and often ineffective shadow removal algorithms are avoided.

3.5.7 Summation

After segmentation, the positive and negative frames are added to eliminate double counts and to ensure that as many objects are detected as possible. This addition is possible since the frame and its negative are spatially registered and therefore the objects which occur simultaneously in both the frame and its negative reinforce each other. The output of the summer gives the complete edge map, and therefore the binary image of the frame.

3.5.8 Post-processing for accurate vehicle segmentation

The obtained binary frame is then subjected to morphological filtering. First, the segmented binary frame is closed so as to eliminate any spurious disjoints between connected components. Then the holes in the connected components are filled to ensure that true sizes of objects are used in subsequent stages. Then, the processed frame is masked using the modeled ROI mask so as to limit the counting and classification to the objects found in the region of interest only. In this way, objects that are not of interest such as roadside buildings and vehicles moving on lanes that are identified as in Fig. 14 are effectively eliminated as shown in Fig. 15. Finally, the irrelevant small objects within the ROI such as pedestrians are deleted using a morphological opening operation.

The consequence of morphological processing is that the shapes of the vehicles are not preserved, and therefore, cannot be used for classification. Instead, areas of the bounding boxes of the resulting connected components are extracted and used as inputs of the nearest centroid minimum distance classifier which assigns the vehicles appropriate class labels.

The resulting connected components in the fully processed frame represent vehicles on the road at that time. These components are counted to give the total number of vehicles on the given section of the road at the given time. Figure 17 shows the result of the count for the given input frame. Figure 14 shows the ROI polygon used for the slow-moving traffic scene.

From Fig. 16, it can be seen that the vehicles were well detected and that their shadows were rendered into the background.

It should be noted that object 5 in Fig. 16 consists of two vehicles which are represented as one vehicle as can be seen in Fig. 15. This is due to the fact that the car is occluded from the camera's view by the larger vehicle in front of it. This scenario highlights the importance of proper installation of cameras meant for traffic management systems.

3.5.9 Vehicle classification

A Euclidean distance based nearest centroid minimum distance classifier is used to classify the vehicles into three classes on the basis of their dimensions: small, medium and large vehicles.

Fig. 13 Real-time traffic flow scene
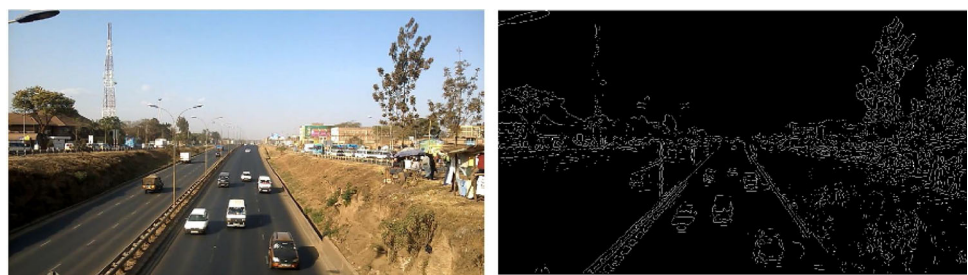
Original   Segmented
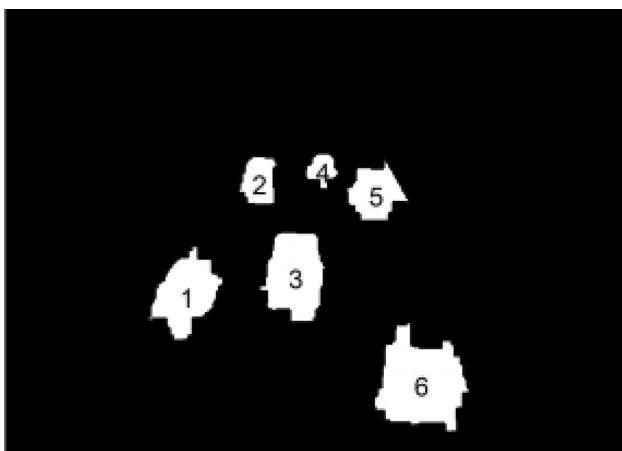
Fig. 14 ROI polygon

Fig. 15 Processed frame

Fig. 16 Post-processed frame

For both convenience and practical reasons, the five-fold cross-validation technique was used. Using this method, the dataset of the extracted vehicles is split randomly into five approximately equal subsets for cross-validation. Each

subset contains all the three classes, but not necessarily in equal portions. At each of the five validation trials, one subset is used for testing while the other four are used for training. Classification accuracies for the five trials are averaged to obtain the classification accuracy of the algorithm for a given dataset.

In order to use the nearest centroid minimum distance classifier, the feature vectors of the vehicles present in the four training subsets are averaged for each class at each trial. Therefore, in the training set, each class is represented by its mean vector

To classify a given unlabeled vehicle, the Euclidean distance between its feature vector and each of the vectors representing the three classes is calculated. Then the vehicle is assigned to the class of the nearest centroid. This can be simplified by evaluating the decision functions of all the three classes for this classifier as given in Eq. (3)

$$d_j(x) = x^T m_j - \frac{1}{2} m_j T m_j \quad j = 1, 2, 3\ldots \qquad (3)$$

where $d_j(x)$ is the decision function of class w_j, x is the unknown feature vector and m_j is the mean vector representing class w_j; x is assigned to class w_j if one of the three decision functions, $d_j(x)$ yields the largest numerical value.

4 Results and discussions

In order to assess the performance of the algorithm for different classes of vehicles like small, medium and large vehicles under various illumination levels across the day, the data was collected at 06.30 h before the sun is up; 12.30 h when the sun is overhead and the shadows are negligible, and 1630 h when both reflections from the road surface and shadows are strongest. Data was also collected from a traffic scene that involved very slow moving traffic
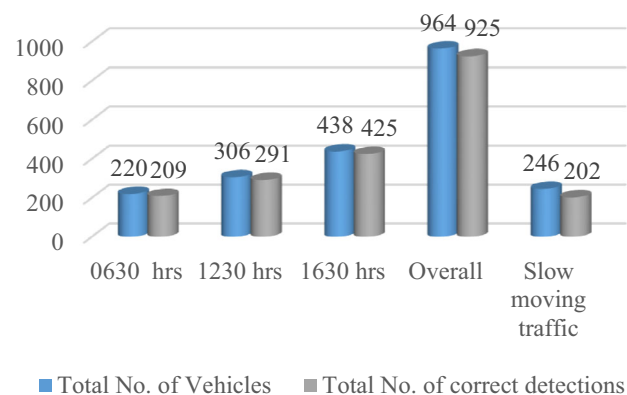

Fig. 17 Number of vehicles detected during different times of the day

so as to assess the performance of the proposed system on such traffic scenes or on stationary ones. Each collection period lasted 20 min, resulting in at least 36000 frames each time. The comparative results obtained for vehicle class deduction has been presented in Fig. 17. It is clearly evident that the number of truly detected vehicles are more in numbers in the day time with good illumination and decreases while the illumination levels drop-off in the early morning and in the late evening.

At 0630–0650 h, there were a total of 220 vehicles in the video data, 209 of which were correctly detected. This translates to 95% detection accuracy. In order to obtain as many vehicles as possible for classification, manual adjustments were done to the vehicle detection algorithm for frames whose vehicles were not correctly detected and as a result, 216 vehicles of the 220 were extracted for classification, while the other 4 were over segmented and were therefore not included in the classification. For this dataset, a classification accuracy of 81.7% was achieved. The same was done for the other two datasets from the other two time periods. The results are summarized in Table 1. The three datasets from the three time periods of the day were added to form an overall dataset and then fivefold cross-validated has been done. Using this method, the dataset of the extracted vehicles is split randomly into five approximately equal subsets namely subset1 to subset5 for cross validation. Each subset contains all the three classes, but not necessarily in equal portions. At each of the five validation trials, one subset is used for testing while the other four are used for training. Classification accuracies for the five trials are averaged to obtain the classification accuracy of the algorithm for a given dataset. Table 2 through Table 6 shows the confusion matrices for each of the 5 subsets of the overall dataset used as a testing set. From Table 3, it has been inferred that Medium class vehicles are mostly misclassified into Small and Large vehicles.

For the slow moving traffic scene, there were a total of 246 vehicles in the video data, 202 of which were correctly detected. This translates to 82.1% detection accuracy. After manual manipulations on the frames whose vehicles were not correctly detected as explained above, 224 vehicles were extracted for classification. On this dataset, a classification accuracy of 83.8% was achieved. These results are generally poorer than those for the free flowing traffic scene as shown in Table 1. The results of the vehicle class prediction are presented in Fig. 18. The reason for this is that occlusions were more severe in the slow moving traffic scene than for the free flowing traffic scene. Consequently, at times, two or even more vehicles could be detected as one vehicle rather than as separate vehicles.

The relatively low camera position was the main cause of detection errors since it was difficult to deduce the spaces between the vehicles on the same lane when the vehicles involved were very close together. It was also the main cause of misclassification. For small vehicles, due to their size, their entire tops were visible while this was not true for the other classes where only the fronts and some part of the tops were visible; resulting in the usage of different dimensions for different vehicles in classification. This greatly reduced the ability of the classifier to distinguish between small and medium vehicles as is evident from the given confusion matrices in Table 2. From the confusion matrix in Tables 4, 5 and 6, the number of misclassified medium sized vehicles depends on the classification features of the small vehicles at a maximum. Since the camera calibration parameters are not considered for object detection, which leads to the higher misclassification rate among the small and medium classes of vehicles.

From the confusion matrices, it is being inferred that for the overall dataset, only two large vehicles were misclassified as small and medium vehicles as shown in Table 2, and there was no small vehicle which was misclassified as

Table 1 Summary of the results for classification of Vehicle classes

Time of the day	Total no. of vehicles	Total no. of correct detections	Total no. of wrong detections	Training set for vehicle detection and classification				Detection accuracy (%)	Classification accuracy (%)
				No. of vehicles	No. of Small vehicles	No. of Medium vehicles	No. of large vehicle		
0630 **h**	220	209	11	216	132	56	28	95.0	81.7
1230 h	306	291	15	298	182	65	51	95.1	88.0
1630 h	438	425	13	438	280	78	80	97.0	93.8
Overall	964	925	39	952	594	199	159	96.0	89.4
Slow moving traffic	246	202	44	224	114	73	37	82.1	83.8

Table 2 Confusion matrix for vehicle class label detection for subset 1	Subset 1		Predicted class labels		
			Small vehicle	Medium vehicle	Large vehicle
	Actual class label	Small vehicle	117	1	0
		Medium vehicle	8	23	8
		Large vehicle	0	0	31
	Classification accuracy		91.0%		

Table 3 Confusion matrix for vehicle class label detection for subset 2	Subset 2		Predicted		
			Small vehicle	Medium vehicle	Large vehicle
	Actual class label	Small vehicle	114	4	0
		Medium vehicle	10	21	8
		Large vehicle	1	1	29
	Classification accuracy		87.2%		
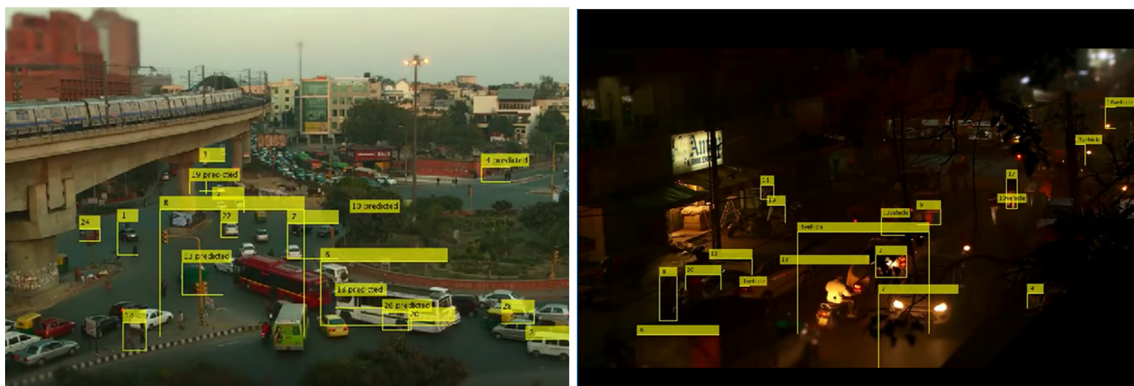
Fig. 18 Samples of vehicle class prediction at daytime and nighttime

Table 4 confusion matrix for vehicle class label detection for subset 3	Subset 3		Predicted class labels		
			Small vehicle	Medium vehicle	Large vehicle
	Actual class label	Small vehicle	115	3	0
		Medium vehicle	9	23	8
		Large vehicle	1	0	31
	Classification accuracy		89.0%		

Table 5 Confusion matrix for vehicle class label detection for subset 4	Subset 4		Predicted class labels		
			Small vehicle	Medium vehicle	Large vehicle
	Actual class label	Small vehicle	112	8	0
		Medium vehicle	10	25	5
		Large vehicle	0	0	32
	Classification accuracy		88.0%		

a large vehicle. This was due to the fact that the visible parts of large vehicles were generally larger than the visible parts of most the small vehicles, making it easier for the classifier to distinguish between the two classes. Attempts

Table 6 Confusion matrix for vehicle class label detection for subset 5

Subset 5		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class label	Small vehicle	117	4	0
	Medium vehicle	8	29	4
	Large vehicle	0	0	33
Classification accuracy		91.8%		

to raise the camera position were not successful due to practical limitations in video capture.

As regards to the classification accuracies related to the times of the day, it may seem that they improve as the day matures till late evening. The number of training samples and the accuracy with which the data collected happened might also influences the accuracy of the results. But in majority of the cases, it seems that illumination conditions and the combinations of the hybrid algorithms for vehicle class detection plays a predominant role in the vehicle detection and classification. For the morning dataset, the set was much smaller than any of the other two. Also, the fact that the camera was removed after each event and installed a new the next time data was to be collected meant that the regions of interest (ROI) were not exactly the same for the different time periods. This could also cause errors and was the main reason for coming up with the overall dataset will enable to get a reasonable average of the classification accuracy of the algorithm across the given time periods. It is therefore, clear that the variations are related directly to object extraction rather than the classification itself.

5 Conclusion

This work presents computer-vision based Intelligent Traffic Surveillance System for Indian Road conditions. The system produces better results for nighttime traffic surveillance and in different weather conditions. The vehicle classes will be used to predict and estimate the traffic density at any time in a given road segment. The system also considers various climatic conditions to monitor video surveillance system the vehicle is being classified into small, medium and large vehicles after obtaining the vehicle count.

This work attempted to solve the problem of vehicle detection, counting and classification in natural traffic scenes using video surveillance systems for both free-flowing and slow-moving or stationary traffic scenes. Also, the ultimate aim to build robust system irrespective of different weather conditions has been achieved. Various preprocessing techniques for different environments has been presented and the system is designed to dynamically

choose the suitable algorithm. The shadows were also well handled with a good degree of success. The vehicle detection algorithm used a novel approach in which the vehicles were simultaneously extracted from the traffic video data frames and their negatives using the Laplacian of a Gaussian edge detector. Edge linking was achieved through mathematical morphology and summation of the positive and negative edge maps. However, despite the success of the algorithm, it was noted that over segmentation occurred for very large trucks: with cabins and their trailers being detected as separate vehicles. The algorithm also had problems with occluded vehicles.

To minimize these problems and classification errors, it is suggested to raise the position of the camera to be high enough with respect to the ROI. The camera and its parameters like Pan, Tilt, and Zoom quietly affect the performance of the moving object detection in terms of disturbing the region of interest under surveillance.

References

1. Gandomi A (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag 35(2):137–144
2. Xu Z (2015) Semantic based representing and organizing surveillance big data using video structural description technology. J Syst Softw 102(1):217–225
3. Choudhury S (2016) Vehicle detection and counting using haar feature- based classifier. IEEE Access 12(5):106–119
4. Ren S (2015) Faster R-CNN: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst 23(1):15–28
5. Guo JM (2015) Nighttime vehicle lamp detection and tracking with adaptive mask training. IEEE Trans Veh Technol 65(6):4023–4032
6. Balid W (2017) Intelligent vehicle counting and classification sensor for real-time traffic surveillance. IEEE Trans Intell Transp Syst 19(6):158–169
7. Vaddi RS (2015) Computer vision based vehicle recognition on indian roads. Int J Comput Vis Signal Process 5(1):1–13
8. Yang H (2017) Real-time vehicle detection and counting in complex traffic scenes using background subtraction model with low-rank decomposition. IET Intell Transp Syst 12(1):75–85
9. Liu W (2017) An ensemble deep learning method for vehicle type classification on visual traffic surveillance sensors. IEEE Access 5(1):24417–24425

10. Xiao Z (2017) Dense scene flow based coarse-to-fine rigid moving object detection for autonomous vehicle. IEEE Access 5(1):23492–23501
11. Barcellos P (2015) A novel video based system for detecting and counting vehicles at user-defined virtual loops. Exp Syst Appl 42(4):1845–1856
12. Kenana MU (2016) Multiscale edge fusion for vehicle detection based on difference of Gaussian. Optik Int J Light Electron Opt 1(17):4794–4798
13. Moutakki Z (2018) Real-time system based on feature extraction for vehicle detection and classification. Transp Telecommun 19(2):93–102
14. Chen HT (2016) Daytime preceding vehicle brake light detection using monocular vision. IEEE Sens J 13(1):120–131
15. Tian B (2017) Hierarchical and networked vehicle surveillance in ITS: a survey. IEEE Trans Intell Transp Syst 18(1):25–48
16. Liu F, Zeng Z (2017) A video-based real-time adaptive vehicle counting system for urban roads. Publ Libr Sci One 12(11):1–16
17. Chan S (2017) Fast tracking of hidden objects with single-pixel detectors. Electron Lett 53(15):1005–1008
18. Fang Y (2017) Performance evaluation of visual tracking algorithms on video sequences with quality degradation. IEEE Access 5(1):2430–2441
19. Sebastien C (2017) Track everything: limiting prior knowledge in online multi-object recognition. IEEE Trans Image Process 23(10):4669–4683
20. Datondji SRE (2016) A survey of vision-based traffic monitoring of road intersections. IEEE Trans Intell Transp Syst 1(3):1–18
21. Kamkar S (2016) Vehicle detection, counting and classification in various conditions. IET Intell Transp Syst 10(6):406–413
22. Taha M (2016) Day/night detector for vehicle tracking in traffic monitoring systems. Int J Comput Inf Eng 10(1):68–82
23. Chen J (2018) Image thresholding segmentation based on two dimensional histogram using gray level and local entropy information. IEEE Access 11(8):5269–5275
24. Yang Z (2018) Vehicle detection in intelligent transportation systems and its applications under varying environments: a review. Image Vis Comput 69(1):143–154
25. Dadi HS (2018) Face recognition and human tracking using GMM, HOG and SVM in surveillance videos. J Data Sci 5(157):2198–2213