

Bio-motion visual analysis for minimizing the death rate of human life in accidents using road side video surveillance.

Maha Vishnu VC*, Rajalakshmi M

Department of Computer Science and Engineering and Information Technology, Coimbatore Institute of Technology, Coimbatore, India

Abstract

Video surveillance and biomedical research have received a great attention in most of the active application-oriented research areas of computer vision, artificial intelligence, and image processing. Visual analysis of human motion is presently one of the active research areas in video surveillance system. Human motion analysis relates the detection, tracking and recognition of activities of the people, and more generally, the understanding of human behaviours, from image sequences involving humans. The road side traffic video surveillance aims at using several image processing methods to obtain better traffic and road safety, which in turn provides direct solution for to reducing death rate of accident victims. The distributed computing process has an efficient solution for this scalability issues in traffic video surveillance system. In this paper, a road traffic video surveillance system is proposed which can automatically identify road accidents from live video files. The system alerts neighbouring hospitals and highway rescue teams when accidents occur. This paper proposes a methodology to process the video file using map-reduce framework for better network service and better scalability solution in surveillance system. Then, the distributed video files are enhanced with Gaussian filtering. Efficient object classification and detection is carried out using Linear Discriminant Analysis (LDA) with Support Vector Machine (SVM) for traffic monitoring using video files from surveillance camera. Foreground object segmentation is a vital task which is carried out by Markov Random field (MRF) with Bayesian estimation process. This proposed method efficiently track and classify the foreground objects with use of object classification and detection and this will improve traffic monitoring in traffic scenes. The results obtained from the experiments on the proposed research shows the efficiency of traffic monitoring using traffic scenes.

Keywords: Traffic video surveillance, Map-reduce frame work, Linear discriminate analysis (LDA) with support vector machine (SVM), Gaussian filtering method, Markov random field (MRF) with Bayesian estimation process.

Accepted on August 31, 2016

Introduction

Video surveillance system has been widely used to monitor activity across large application environments such as: transport systems, banks, shopping malls, car parks, and hospitals. Security has become a critical problem in recent decades and there is a specific need to enable video surveillance system in most of the public places. Modern video surveillance systems integrate image analysis techniques for efficient image transmission, event-based attention, and model-based understanding of sequence. This is a contributing factor to the increasingly widespread deployment of intelligent video surveillance systems [1]. One among the most significant applications of the video surveillance systems is traffic surveillance, which has achieved growing interest in last few years. The increasing number of vehicles in the past decade has made it necessary for a good traffic surveillance system. In this field, the learning of the behaviour of traffic is observed to be

the most complicated task, particularly in largely changing environments [2]. Video surveillance in traffic scenes has made it.

The other complexity in traffic video surveillance system is that, it must handle large amount of data in real time. Thus, handling large data in video surveillance system has become an active research area in recent years in distributing computing environment [3]. Video surveillance in a distributing computing environment requires constant handling of huge data in order to recognize and track an object in traffic scenes. The system requires an efficient strategy to recognize and track an object by extracting the feature of the object. However, most object tracking techniques that are based on feature matching pose a challenge due to greater computational complexity and/or weak reliability in different environments.

To overcome the aforementioned complexities in video surveillance in traffic scenes, this work proposes a smart video

surveillance system which is capable of enhancing situational awareness across multiple scales of space and time [4]. The proposed work also focuses on most important field in traffic video surveillance applications namely moving object detection and classification [5]. When surveillance cameras are taking video file as an input into the surveillance systems, background subtraction method [6] is applied for segmenting the background and foreground objects for efficient object detection and classification process. Thus, initially, the background subtraction processes applied into map reduce frame work [7] in distribute file system environment, to reduce the scalability problem and also an efficient noise removal process is carried out to enhance the video files from the distributing frame work for further process.

This work is organized as the following sections. Section 2 provides a related work based on the proposed technique, in Section 3 the proposed methodology of video surveillance system in traffic scenes is discussed in detail. Section 4 evaluates the experimental results. Finally, Section 5 makes a conclusion about this proposed method.

Related Work

Huang et al. proposed a real-time object detection algorithm for night-time visual surveillance in [8]. This algorithm is based on contrast analysis wherein the contrast in local change over time is used to detect potential moving objects. After that motion prediction and spatial nearest neighbour data association are used to suppress false alarms. The experimental results on real scenes show that new algorithm is effective for night-time object detection and tracking.

Javed et al. discussed about the problems which are required to be solved prior to the development of completely automated outdoor surveillance systems, and presented solutions to solve the surveillance based problems [9]. Any kind of outdoor surveillance system should be capable of tracking objects that are moving within its area of view, and then perform the classification of these objects and detection of few of their activities. The authors designed a technique for tracking and classifying these objects in practical scenarios. Object tracking with a single camera is carried out employing background subtraction, which is then followed by the region correspondence. This considers multiple cues that are inclusive of velocities, sizes and the distances of surrounding boxes. Objects can be divided on the basis of their kind of motion. This characteristic may be brought into use for labelling the objects in the form of a single person, vehicle or a group of persons. This method is used to classify the objects which is in accordance with the detection of repetitive motion for every tracked object and developed a particular feature vector known as a 'Recurrent Motion Image' (RMI) for calculating the recurrent motion of objects. It is shown that the different kinds of objects render diverse RMI's and segregated into various categories based on their RMI. This approach is very resourceful both with respect to the criterion of computation and space. RMI's are moreover utilized for detecting the objects carried in the surveillance system.

Ahn et al. proposed effective object recognition and tracking technique, which employs a sophisticated feature matching for use on practical time environment [10]. This algorithm recognizes an object making use of invariant features, and minimizes the dimension of a feature descriptor to deal with the reorganization problems in video surveillance in cloud environment. The experimentation results obtained from the technique is observed to be faster and more reliable than the traditional techniques, along with that, the technique can do the detection and tracking of a mobile object accurately in various environments.

Zeng et al. presented an idea to avoid traffic accidents by the use of an object detection system that is abandoned for the road traffic surveillance systems that is on the basis of the information from three-dimensional image [11]. In this work, a Binocular Information Reconstruction and Recognition (BIRR) algorithm were used for implementing this concept. The suspected and abandoned objects were identified by the static foreground region segmentation algorithm that is based on the surveillance video obtained from a monocular camera. Once the suspected abandoned objects are detected, three-dimensional (3D) information about the suspected abandoned object is then reconstructed by means of the proposed theory regarding the reconstruction of 3D object information with images obtained from a binocular camera. In order to decide if the object detected is harmful to the actual road traffic, the road plane equation along with the suspected-abandoned object's height are computed as per the three-dimensional information. The results indicate that the system is capable of implementing rapid detection of the abandoned objects.

Proposed Methodology

Map-reduce framework for distributing environment for monitoring traffic scenes with video surveillance process

Map-reduce framework: Map-reduce frame work is simple and parallel programming concept which lets the users to make use of a well-designed style of programming for creating a map function which does the processing of a key-value pair that is associated with the input data, i.e. video files from the distributing storage environment to have a set of intermediate key-value pairs generated, in addition to a reduce function which combines all the intermediate values which are associated with the same intermediate key.

Distributed computing yields a potential concept which offers a virtualized infrastructure in the form of a service which is scalable and in which the complications of fine-tuned management of resource remains hidden from the end-user. Running data known as video files analysis in the distributing computing with huge data has become an active and trendy research topic to satisfy huge scalability demands. Typically, surveillance controls the map-reduce framework which can disintegrate the large computation video files into set of parallelizable smaller computations video files for further object reorganization in video surveillance process using the

traffic scenes. Video surveillance process used videos from the traffic surveillance cameras which are in distributing computing environment, in order to control traffics efficiently. Thus, the proposed map-reduce framework designed for traffic video surveillance system based on distributing computing environment to efficiently manage the large amount of video files and has the scalable video streaming component to smoothly manage the traffic video file from large number of networked traffic video cameras.

Video surveillance application can run on a local network or can be deployed on clouds that support Infrastructure as a Service (IaaS). Map-reduce framework employing the parallel pool key values for data in huge amounts across different surveillance cameras. To enhance classification and detection for traffic scene image retrieval from video files and processing speeds, the sequence video files are split into smaller units known as chunks and distributed over map reduce framework nodes. When jobs are submitted to map reduce framework, they are scheduled on machines that host the required data. The proposed surveillance system can be integrated with various traffic video files capturing from different surveillance cameras. The video files are in different file formats. The input video file format can be considered as a logical container that wraps compressed video files. The results presented influenced the use of map-reduce framework for proposed system [12]. The distributed map-reduce video processing framework is illustrated in Figure 1. The framework has two different phases such as map and reduces phases.

Map phase: At first the mapping phase, video file which is the input is separated into input splits by default size and then allocated to map tasks corresponding to the nodes responsible for processing in the cluster. The map task usually does the execution on the same node having its allocated data partition in the cluster. A map-reduce phase carries out computations that are user-specified on every input key value pair from the input data partition that is assigned for the task, and produces a set of intermediate results for every key. In map-reduce framework there is also used another phase such as the shuffle and sort phase. The shuffle and sort phase does the sorting of the intermediate data that is generated by every map task from the other nodes and partitions this data into regions that have to undergo processing by the reduce tasks. This phase also does the distribution of this data as required to nodes in which the reduce tasks will be executed.

Reduce Phase: In the reduce phase, the processing of the data partitioned by shuffle and sort phase is carried out. The reduce tasks perform extra user-mentioned operations on the intermediate data by possibly combining the values that are associated with a key to a smaller set of values to generate the output data. Every map task has to be completed before the shuffle and sort and reduce phases. It is not necessary that the number of reduce tasks be the same as the number of map tasks. For more sophisticated video data processing processes, several map-reduce calls might be connected together in sequential order.

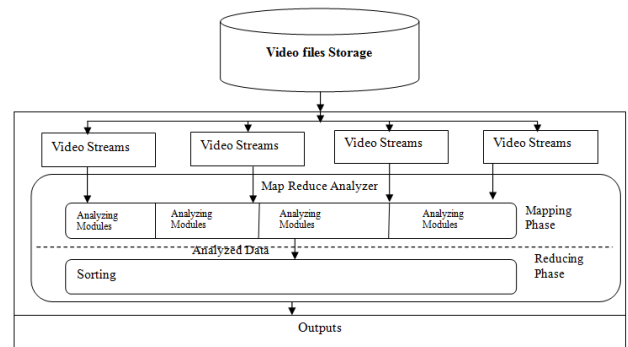


Figure 1. Map-reduce framework for surveillance video files in distributing computing environment.

In the first mapping phase of map-reduce, the cluster nodes are assigned data from the frames of a single video [13]. The input video frames are read and vehicles are detected using a machine-learning algorithm. For each frame processed by a map, a corresponding output frame is produced and is indexed for subsequent efficient access. The index contains pairs of key values which are used to connect a frame identifier with the corresponding data. The information extracted by the map-reduce frame including vehicle counts, accident data and average vehicle speed are taken into account for further object classification and object detection process in the proposed work.

Noises removal by using Gaussian filtering

The Gaussian function has critical applications in several fields of mathematics, inclusive of image processing. This approach uses a Gaussian filtering to enhance the video streams obtained from the map-reduce framework. The Gaussian features which render it to be useful for enhance the video frames or the detection of edges once after enhancing. A Gaussian function containing one variable along with spread σ is of the following form, in which d is scalar.

$$f(x) = de^{-x^2/2\sigma^2} \rightarrow (1)$$

A Gaussian function having two variables defined as follows,

$$f(x, y) = de^{-\frac{(x^2 + y^2)}{2\sigma^2}} \rightarrow (2)$$

in which the constant d was fixed such that the area lying under the curve will be 1. In order to make a video mask for the purpose of filtering, usually make d is a large number in such a manner that all video masks are integers. The Gaussian is defined centered on the source video file and therefore does not require any location parameter μ as the normal distribution does. Video surveillance processing algorithm will translate Gaussian function while it will be applied into noise removal process of video files. The area lying under function $f(x)$ is 1; this means that it is greatly desirable to be a noise removing filter which does not have any effect on static regions. $f(x)$ is basically a positive even function; $f'(x)$ is just the multiplication of $f(x)$ with an odd function $f(x)$ and then it

scaling down done by σ^2 . More structure is revealed in $f''(x)$. Equation 5 shows that $f''(x)$ is actually the difference between two even functions.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \rightarrow (3)$$

$$f'(x) = \frac{-1}{\sqrt{2\pi}\sigma^3} x e^{-\frac{x^2}{2\sigma^2}} = \frac{-x}{\sigma^2} f(x) \rightarrow (4)$$

$$f''(x) = \left(\frac{x^2}{\sqrt{2\pi}\sigma^5} - \frac{1}{\sqrt{2\pi}\sigma^3} \right) e^{-\frac{x^2}{2\sigma^2}} = \frac{x^2}{\sigma^4} f(x) - \frac{1}{\sigma^2} f(x)$$

$$= \left(\frac{x^2}{\sigma^4} - \frac{1}{\sigma^2} \right) f(x) \rightarrow (5)$$

The second derivative of a Gaussian $f''(x)$ as reduce negative area of the video masks and two enhance side mask of positive area: the zero crossings are located at $-\sigma$ and $+\sigma$, corresponding to the modulation points of $f(x)$ and the extreme points of $f'(x)$.

Thus the noise reduction process using Gaussian function in a Gauss filters in the x-direction then followed by a filter in a direction that is non-orthogonal. The decomposition is σ shown to be highly resourceful from the perspective of computing.

Background Subtraction process using the Markov random field (MRF)

The MRF background subtraction method is used to detect a moving object as a foreground by segmenting it from a traffic scene of a traffic surveillance camera. The surveillance camera could be static or dynamic [14]. Background subtraction process tries to detect the objects that are moving such as moving vehicles in traffic scenes from the difference observed between the current frame and the reference frame in a pixel-by-pixel or block-by-block scheme. The reference frame of the video file is generally referred to as background image or background model. There is a change in dynamic traffic scenes; a good background model adapts dynamic scenes. The background information updating in periodic intervals is able to do this [15], though this can also be performed without updating the background information [16]. Foreground objects in a video stream are identifying with use of background subtraction method. It is a widely used real-time method. It is the most important stage in surveillance application. This work forced the design of a background algorithm like Markov Random Field (MRF) that targets to enhance the performance pertaining to the objects classification and detection process. In this section discussed in detail about the MRF based background subtraction in traffic scenes.

Background subtraction model: Let consider the current noise free video file image as O with s pixels from the Gaussian filtering method with background representation as O^b and foreground representation as O^f , and Γ_{SB} , Γ_{DB} and Γ_F

denote as static and dynamic background and foreground corresponding regions respectively. The current noise free video file image as O can be denoted in terms of noise-free foreground pixel value m , mean of significant static background cluster μ_{SB} and significant dynamic background cluster μ_{DB} , as well

$$O(s) = \begin{cases} \mu_{SB}(s) & s \in \Gamma_{SB} \\ \mu_{DB}(s) + v(s) & s \in \Gamma_{DB} \\ m(s) & s \in \Gamma_F \end{cases} \rightarrow (6)$$

Let the measurement D represents the difference between current noise free image Img and background image B , and correlated-cluster-distance dt presents distance between two means of correlated clusters. The resulted measurement d is expressed as,

$$d_o(s) = \begin{cases} dt(s) & s \in \Gamma_{SB} \\ dt(s) + v(s) & s \in \Gamma_{DB} \\ dt(s) - m(s) - \mu_B(s) & s \in \Gamma_F \end{cases} \rightarrow (7)$$

Where correlated-cluster-distance dt_s is defined as

$$dt(s) = \begin{cases} 0 & s \in \Gamma_{SB} \cup \Gamma_F \\ |\mu_{DB}(s) - \mu_B(s)| & s \in \Gamma_{DB} \end{cases} \rightarrow (8)$$

The distribution of pixel associated with static background can be regarded as completely overlapped with distribution of maintained background representation at the same position, i.e. $\mu_{SB} = \mu_B$, whereas the distribution of pixel associated with dynamic background can be regarded to be partial overlapped with distribution of maintained background representation. As to the pixel associated with foreground, its distribution is considered to be uncorrelated with that of maintained background representation; accordingly the correlated-cluster-distance is zero. Given the measurement $d_o(s)$, and make the following two hypothesizes: $H_0: s \in \Gamma_{SB} \cup \Gamma_{DB}$; $H_1: s \in \Gamma_F$.

The procedure assigns label q_s to each pixel s from the binary label-set: $w = \{Fg, Bg\}$ corresponding to two possible classes: foreground (Fg) and background (Bg). Each label q_s either takes the value Bg if the observed measurement $d_o(s)$ supports the null hypothesis H_0 , or the value Fg if the observed value of $d_o(s)$ does not support this assumption, i.e. alternative hypothesis H_1 . Therefore, the background subtraction is equivalent to a global labelling $\phi = \{q_s | s \in S\}$. Using Bayesian estimation, it estimates the label set such that it's a posterior probability $P(\phi | d)$ given the measurement object d is maximized. In this proposed MRF background subtraction process, the Bayesian estimation is applied for context distinct regions similar to local region processing techniques. The context distinct region is denoted as sub-image which contains significant measurement d to be observed. The following section discussed about the background subtraction using MRF with Bayesian estimation.

Markov random field (MRF): Assume labelling sets $\phi_{Bg}^s = \{q_i = Bg | i \neq s, i \in S\}$ and $\phi_{Fg}^s = \{q_i = Fg | i \neq s, i \in S\}$ from Bg

and F_g are known, the Bayesian likelihood ratio [17] of measurement d at pixel s can be determined as,

$$\frac{P(\bar{d}_0(s)|\phi_{Bg}^s)_{Bg}}{P(\bar{d}_0(s)|\phi_{Fg}^s)_{Fg}} \leq Th \cdot \frac{P(\phi_{Bg}^s)}{P(\phi_{Fg}^s)} \rightarrow (9)$$

where Th is a decision threshold, and (ϕ) , which is the measurements lying inside the structural window w , is denoted $(\phi)(s) = \{d_o(s) | s \in w_s\}$, $P(\cdot | \Phi)$ is the conditional pdf of the observed measurement image d given global labelling ϕ . $P(\phi_{Bg}^s)$ and $P(\phi_{Fg}^s)$ are a priori probabilities for Bg and Fg respectively. Since Gaussian distribution has been made on background model for each pixel, measurement can, therefore, be regarded as to obey zero mean Gaussian distribution after removing the contribution of correlated-cluster-distance dt . The information from two distinct models is combined using Bayesian estimation. The local visual observations at each node to be labelled yield label likelihoods. The resultant label likelihoods are merged with a priori spatial knowledge about the neighbourhood denoted in the form of an MRF.

Consequently, $P(\cdot | \Phi)$ results in conditional Gaussian pdf $\eta(\cdot)$ in structural window formulation by the conditionally independent assumption,

$$\eta(\bar{d}_0(s)|H_j) = \left(\frac{1}{\sqrt{2\pi\sigma_j^2}} \right)^{w_s} \exp \frac{\sum_{s \in w_s} (d_o(s) - dt(s))^2}{-2\sigma_j^2} \rightarrow (10)$$

The a priori probability is modelled by Markov random fields.

Taking above equation, the likelihood test Equation 9 can be further extended to regional form. Here, pixels in the same local region share the same variance and correlated-cluster distance for the likelihood test. Exploiting the fact that background associated distributions inside a local region can be retrieved from the multivariate Gaussian background model. The domain-specific information can be utilized by encoding any relevant parameters about that region to select proper regional variance and correlated-cluster distance.

Let σ_R^k and dt_R^k represent the selected variance of null hypothesis and correlated-cluster-distance of k^{th} local region. Considering the variance related to foreground is much larger than the variance related to the camera noise or the dynamic background variation, the Bayesian likelihood test can be approximated by taking the logarithm on both sides, the derived equation depends on a priori measurement distributions been observed in temporal domain. The σ_R^k further serves as a multiplication factor which adaptively increases/decreases the threshold corresponding to the context of observed background variation in a local region.

The MRF-based Bayesian estimation model for background subtraction initialization strategy allows the training video file sequence to contain foreground objects. This work have indicated that the outcomes of a good background model initialization in significantly enhanced foreground object detection that results in better object recognition in surveillance system.

Proposed object classification and object detection process using Linear discriminant analysis (LDA) with support vector machine (SVM) in traffic scene video surveillance system.

In traffic scene video surveillance application, determining the type of object is critical. For example in traffic video surveillance system, detecting a pedestrian at zebra crossing may not be an alarm condition, whereas spotting a moving vehicle at pedestrian crossing point when there is a red signal. Generally there are two methods avail to object classification process such as an image-based method and video surveillance based method. In this proposed work focused on a video surveillance approach to object classification using traffic scenes. This work assumes that the objects of interest have been detected and tracked by an object surveillance system. Video surveillance systems use information about the appearance, shape, and motion of moving objects to quickly distinguish people, vehicles, carts, animals, doors opening/closing, trees moving in the breeze, etc. in traffic scenes which is from the surveillance camera.

Object classification and detection processes: This proposed work uses a combination of Support Vector Machine (SVM) classifier [18] and Linear discriminant analysis (LDA) for object classification and detection process. SVM is a method for machine learning from the examples that is well-founded in the theory of statistical learning. If a large amount of video files of the target objects is available, SVM is a very useful tool for summarizing the input video files in terms of the support vectors. Moreover object detection or recognition has to be performed at many locations and scales in a given image.

Since object-specific context characteristics clearly reflect the objects properties in the image scene, they could be employed for distinguishing the objects. This method is not sufficient to perform moving object recognition by its own because of its lack of robustness against changes in the background. This paper describes a new method for object classification and detection from image sequences from the video files. The system uses a hierarchy of SVMs for object detection and LDA for classification.

Object classification process: The proposed classification module has two processes, a feature extraction process and a scene classification process.

Feature extraction method: The surveillance system uses a feature set which are got by the convolution of the local mask patterns with a provided object from video file. These masks have been introduced for position invariant object detection [19].

Let m^i , i =different number of mask patterns and V_s the 3×3 patch at the pixel s in an object. In this system, the i^{th} feature f_i , is given by $\sum_s m^i \cdot V_s$, where the total is over the image pixels. In this proposed system there are also used as a non-linear convolution. The non-linear convolution operates chiefly on the areas on the edges in an image. The non-linear operation has been motivated by the recent research [20].

$$f_i = \sum_s c_{s,i},$$

Where,

$$c_{k,i} = \begin{cases} V_s m^i & \text{if } V_s \cdot m^i = \max_j (V_s \cdot m^i) \\ 0 & \text{otherwise} \end{cases} \rightarrow (11)$$

This proposed system extracts different number of features based on the 3×3 patch at pixel s in an object. Within the standard supervised machine learning methods, large number of hand-labelled training images is required in each scene for surveillance. The object classification and detection process is vital to implement surveillance scheme in real-time with more number of cameras. Hence this work first apply the object classifier to unlabelled samples from each scene, and then use the obtained soft label to update a LDA based scene-specific classifier, here this work used SVM classifier to recognize the objects from the video file.

Scene context features: Scene context features, which are extracted by using above mentioned methods are defined as those are useful for classification in any single scene from surveillance video files, but fail when training in one scene and testing in another scene. Further, context features are dependent on scene that cannot be transferred across the scenes due to their diverse distributions in multiple scenes. For the purpose of the classification task of isolating the pedestrians from vehicles, features like object position and direction of motion are shown in the form of scene-context features [18]. In traffic scenes from the surveillance camera where roads are static, by simply using spatial information of objects, this work may classify and differentiate the moving objects such as pedestrians from the moving vehicles. This proposed work used some scene-independent features, which can be used to separate or classify one moving object (pedestrian) from another moving object (vehicle) in traffic scene. There are different features are utilized in this work including scene-context features such as x - and y - object coordinates and the aspect ratio. Simple scene-independent features such as the bounding-box area, percentage of occupancy and speed of motion are merged with the above said scene context features in the classifier updating process.

Scene classification: LDA Based Scene Specific Classifier: Linear Discriminate Analysis (LDA) [21] is mostly used technique for dimensionality reduction in many classification problems, here this LDA used to classify or differentiate the foreground object in traffic scene. The objective is to find a projection W that maximizes the ratio of between-class scatter Sc_B against within-class scatter Sc_W (Fisher's criterion), which is denoted as follows:

$$\arg \max_W W^T Sc_B W / W^T Sc_W W \quad (12)$$

This proposed work used different number of samples in each scene to train LDA based scene-specific classifier to differentiate the foreground object in video file. With different distribution of training images, the number of support vector in the SVM classifier used to detect the foreground objects, thus this proposed object classification and detection method

making time cost beyond control. To achieve efficient performance in real-time surveillance system using in traffic scenes, this work employ LDA with SVM as the classifier.

Object detection of a scene using SVM classifier: After classification of the foreground images from the given input traffic scene video file, the objects such as vehicles detection is vital process for surveillance process, in this purpose of detecting the object here this proposed work employs the Support Vector Machine (SVM) method, in which the training images are taken into one class. In SVM, each class is denoted by a set of N vectors; each vector consists of the features, which are extracted by using feature extraction method in the proposed system. The system uses a linear SVM [22] that determines the hyper plane, $u \cdot y + a$ which isolates the two classes best. The u refers to the weight vector, the y stands for features vector, and a refers to a constant. This hyper plane is the one that increases the distance or margin observed between the two image classes. The margin, equivalent to $2 \|u\|^{-1}$, is a vital geometrical quantity as it renders an estimate of the similarity between the two image classes. An important contribution by the SVM approach, the multiclass problem, which is transformed into binary-tree architectures without any reduction in performance. Also, a considerable improvement in the speed of object recognition can be accomplished with the aim of maximizing the number of classes.

Experimentation Results

The following section describes about the dataset used for surveillance process and the experimental result of the proposed work is present in detail. The proposed work has been implemented in MATLAB 2015 map-reduce frame work with parallel pool key value input and output.

Datasets

There are few publicly available datasets suitable for the proposed video surveillance system. Hence this work employs to take on many videos from real-world traffic scenes to evaluate proposed methods. To demonstrate the effectiveness of motion pattern learning, the proposed surveillance system on many videos from different traffic scenes. In this work, the results are shown in traffic scene which includes the primary traffic scene from surveillance cameras shown in Figure 2.

For object classification, different number of videos from different traffic scenes is collected to evaluate the performance of proposed machine learning classifier. The object detection process is also evaluated by using these videos collections from different traffic scenes. The proposed work showed the results of learning paths and improvement of object detection on i-LIDS parked vehicle detection dataset [23]. The effectiveness of proposed methodology is evaluated based on the object detection, path detection and motion detection processes and is clearly shown in Figures 3 and 4. And these results are compared with some existing methods such as LDA-based classifier and Ada-boost classifier [24]. For object detection, the results are compared with Generalized Method

of Moments (GMM) [25] approach. All these experimentation results demonstrate the efficiency of proposed method.

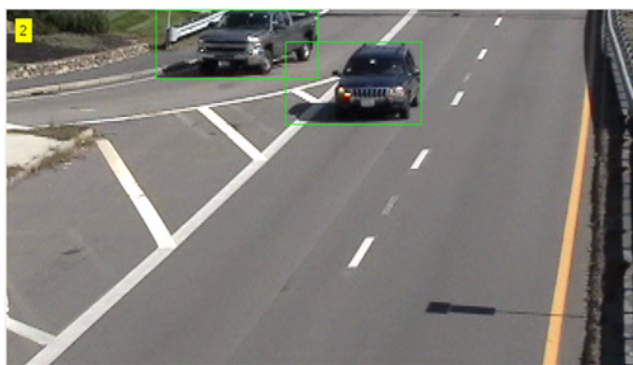


Figure 2. The input video file.



Figure 3. (a) and (b) The object detection result of the proposed method.

From the Figures 3a and 3b, the moving object detection efficiency of the proposed method is clearly illustrated. The proposed approach also calculates the number of moving vehicles in a particular video frame accurately. The proposed object detection method correctly detects moving vehicles in the given input video frame than the existing object detection methodologies in lesser time.



Figure 4. Path detection result of the proposed method.

The Figure 5 shows the motion prediction in the real traffic scene to help object tracking. The vehicle prediction is done to identify the direction of the moving vehicle. After efficient detection of moving vehicles and classification of foreground and background objects from given input traffic scene video frames, this work also concentrated on evaluation of the speed and direction of the moving vehicles. Initially, speed, direction

and motion vector of moving vehicles are calculated using the given input video frames, after that it will increase in more number videos. From the evaluation, eight moving vehicles are detected and further results are obtained from optical flow, direction and speed detection methods which are listed in Table 1.

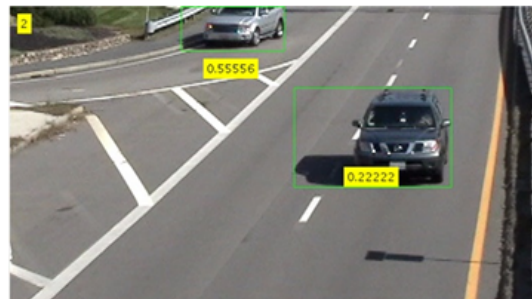


Figure 5. Motion prediction result of the proposed method.

Table 1. Direction, speed, and motion prediction values of the moving vehicles in the input videos.

car	direction	prediction	Speed in kmph
1	'left'	0.88889	55.019
2	'left'	0.88889	74.904
3	'Straight'	0.77778	2.081
4	'left'	0.88889	21.144
5	'right'	0.66667	3.441
6	'right'	0.55556	13.317
7	'right'	0.66667	3.5572
8	'Straight'	0.77778	3.4731

Table 1 shows the prediction value and the direction in which the vehicle moves.

The Figure 6 shows the results of clustering motion patterns.

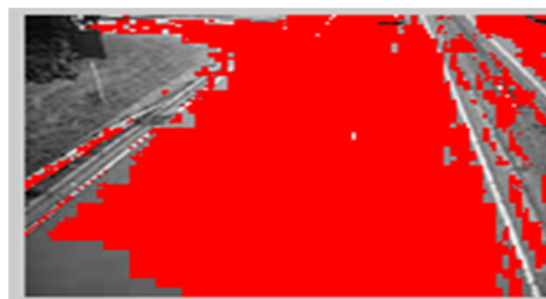


Figure 6. The result of clustering process performance evaluation.

The proposed video surveillance system's LDA with SVM classifier for object detection and classification performance is measured with the help of sensitivity, specificity and accuracy. These performance parameters are denoted as follows:

Sensitivity (SEN): Sensitivity also known as Recall (or) True Positive Rate (TPR) is defined as the probability of the actual positive classes that are correctly identified.

$$\text{SEN} = \text{TP} / (\text{TP} + \text{FN}) \times 100 \rightarrow (13)$$

Figure 7 shown below describes the comparison of sensitivity rate of the proposed LDA-SVM classifier with existing classifiers such as LDA, Ada-Boost and co-training classifiers. Table 2 shows the sensitivity values of the proposed and existing methods.

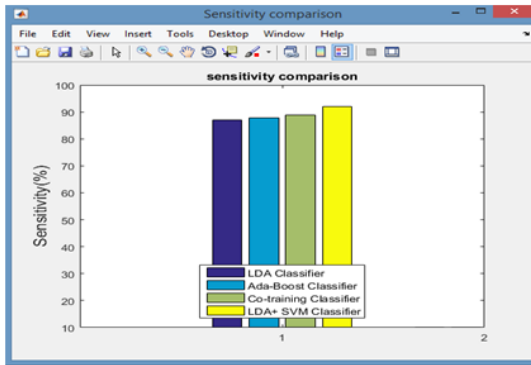


Figure 7. Sensitivity results of the proposed and existing method.

Table 2. Sensitivity value of the proposed and existing method.

Methods/Input	LDA classifier	Ada-Boost classifier	Co-training classifier	LDA + SVM classifier
1	88%	88.5%	89%	92%
2	84.9%	85.2%	87.4%	89.2%
3	86%	84.6%	88.1%	88.8%
4	87.7%	89.8%	90.8%	91.3%

Specificity (SPE): Specificity (or) True Negative Rate (TNR) is the probability of the actual negative classes which are correctly identified.

$$\text{SPE} = \text{TN} / (\text{TN} + \text{FP}) \times 100 \rightarrow (14)$$

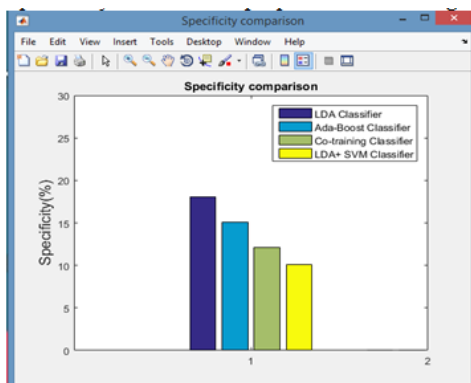


Figure 8. Specificity rates of proposed and existing system classifiers.

Where, TP (True Positive) refers to the number of rightly grouped abnormal event in traffic scenes, FP (False Positive) indicates to the number of wrongly grouped abnormal event in

traffic scenes, FN (False Negative) refers to the number of wrongly grouped normal event in traffic scenes, and TN (True Negative) indicates the number of rightly grouped normal event in traffic scenes. Figure 8 shows the comparison of specificity rate of the proposed method and existing classification methods. Table 3 shows the specificity values of the proposed and existing methods.

Table 3. Specificity value of the proposed and existing method.

Methods/Input	LDA classifier	Ada-boost classifier	Co-training classifier	LDA+SVM classifier
1	18%	15%	12%	10%
2	16%	16.50%	11.40%	8.90%
3	16.50%	15.80%	13.60%	9.10%
4	17.80%	15.30%	11.20%	10.80%

Accuracy (ACC): Accuracy is the percentage corresponding to the correctly done classification of normal and abnormal scenes in traffic monitoring process.

$$\text{ACC} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{FP} + \text{TN}) \times 100 \rightarrow (15)$$

The result shown in Figure 9 proves that the proposed object detection and classification process improves the tracking results in the traffic image analysis in surveillance process.

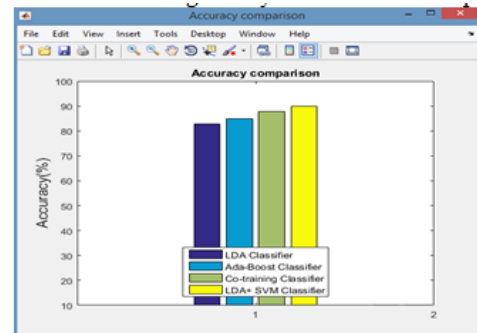


Figure 9. Accuracy rate of proposed system classifier and existing classification system classifier.

The Table 4 shows that the classification accuracy results comparison of the proposed and existing methods.

Table 4. Classification accuracy comparison value of the proposed and existing method.

Methods/Input	LDA classifier	Ada-boost classifier	Co-training classifier	LDA+SVM classifier
1	81.50%	85%	88%	90%
2	80.80%	82.30%	86%	92.60%
3	76.90%	81.80%	87.50%	91.25%
4	79.20%	79.50%	84.20%	89.50%

The overall accuracy of the proposed vehicle detection and classification LDA+SVM Classifier is observed to be more than 90%, whereas the accuracy of the LDA Classifier, Ada-

boost classifier and co-training classifier is around 85%. From observation of the accuracy result, the proposed SVM classifier is observed to perform well with LDA in traffic surveillance system [26]. The significance of the proposed approach is clearly seen with reduced false positive rate, which in turn increases the classification and detection accuracy.

Figure 10 shows the optical flow result, which is used for calculating the moving vehicle speed based on the motion vector value as well as the results of optical flow. For calculation of speed and direction, optical flow algorithm is used to calculate changes in the intensity of the pixels of the images from the video frames. These clear velocity components are then subjected to certain image processing methods to obtain centroid of the moving vehicle across the video frames. The coordinates of video frame's image centroid are mapped to world space. Based on this value, the velocity of the vehicle is estimated for calculating the speed of the moving vehicle in the input video frame.

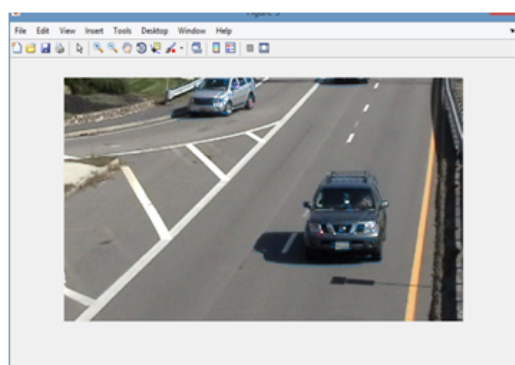


Figure 10. Optical flow results.

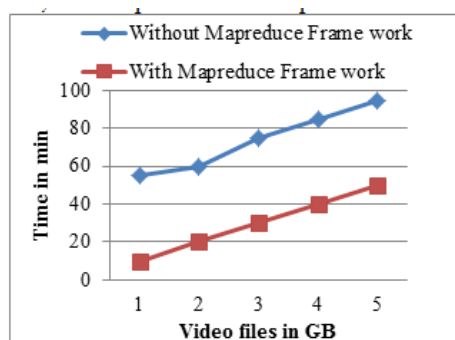


Figure 11. Comparison result of the input video file execution time with and without map reduce framework.

Finally the Figure 11 shows the efficiency of the proposed map-reducing frame work for moving object detection and classification using the large scale traffic scene video files and the Figure 12 shows the map-reduce result for detection, motion prediction and speed calculation of the proposed method. Figure 11 shows unexpectedly long runtime consumption for the object detection and classification without the map reduce framework in the moving object detection approach based on traffic surveillance camera video files. Without map reduce frame work, the process was slowed down

because it requires a long time for object detection and foreground background object classification with large input video file. The result from the Figure 11 is reliable and showed that the map reduce framework is essential for better performance with large scale input video files, because of its efficient map reducing process.

Conclusion

This paper proposes a systematic solution for roadside traffic analysis using map-reduce framework for reducing the death rate of the human life from accidents. This work focused on video surveillance in traffic scenes using road side surveillance camera devices, whereas object detection and classification is a complex problem in the real world, in order to overcome this a machine learning based Linear Discriminate Analysis (LDA) with Support Vector Machine (SVM) classifier with map-reduce framework is proposed, and also Gaussian filtering method based noise removal process, Markov Random Field (MRF) background subtraction processes are also used to improve the object detection and classification for video surveillance system. Initially the proposed work used a parallelizing map-reduce frame work for solve scalability problems, while taken as a large amount of video files as an input of traffic video surveillance process and to find an object in a video file as it is human or moving vehicle in traffic scenes in a distributing environment. Then noise reduction process employed Gaussian filtering method for enhancing the video files from map-reduce frame work. After that the foreground objects are segmented from video files with the Markov Random Field (MRF) method. Finally the object detection and classification have done by using LDA with Support Vector Machine (SVM) classifier. The existing surveillance system had good accuracy and divide object to group but the problem is system need good computing power. The large amount of data need more time to process, this will create a scalability problem in surveillance system. In order to overcome scalability problem and increase the performance efficiency in this proposed work used the map-reduce frame work to reduce the scalability problem and improve the performance of object classification and detection for traffic monitoring process using traffic scenes and this method also increase computation speed of surveillance system. This result is applied to human motion analysis system, namely human detection, tracking and activity understanding based on the video.

References

1. Snidaro L, Remagnino P, Ellis T. Active video-based surveillance system. IEEE Sig Proc Magaz 2005; 25:37.
2. Xiang T, Gong S. Beyond tracking: Modelling activity and understanding behaviour. Int J Comp Vision 2006; 67: 21-51.
3. Dean J, Ghemawat S. Map reduce: simplified data processing on large clusters. Commun ACM 2008; 51: 107-113.
4. Hampapur A, Brown L, Connell J, Ekin A, Haas N, Lu M, Pankanti S. Smart video surveillance: exploring the concept

- of multi scale spatiotemporal tracking. *Sig Proc Magaz* 2005; 22: 38-51.
5. Cucchiara R, Grana C, Piccardi M, Prati A. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Trans on Pattern Anal Mach Intel* 2003; 25: 1337-1342.
 6. Benezeth Y, Jodoin PM, Emile B, Laurent H, Rosenberger C. Comparative study of background subtraction algorithms. *J Electr Imag* 2010; 19: 033003-033003.
 7. Lee Y. Toward scalable internet traffic measurement and analysis with hadoop. *ACM SIGCOMM Comp Commun Rev* 2013; 43: 5-13.
 8. Huang K, Wang L, Tan T, Maybank S. A real-time object detecting and tracking system for outdoor night surveillance. *Pattern Recogn* 2008; 41: 432-444.
 9. Javed O, Shah M. Tracking and object classification for automated surveillance. *Comp Vis ECCV* 2002; 343-357.
 10. Ahn H, Lee YH. Performance analysis of object recognition and tracking for the use of surveillance system. *J Ambient Intel Human Comput* 2015; 1-7.
 11. Zeng Y, Lan J, Ran B, Gao J, Zou J. A novel abandoned object detection system based on three-dimensional image information. *Sensors* 2015; 15: 6885-6904.
 12. Kim M, Cui Y, Han S, Lee H. Towards efficient design and implementation of a Hadoop-based distributed video transcoding system in cloud computing environment. *Int J Multimed Ubiq Eng* 2013; 8: 213-224.
 13. Althebyan Q, Alqudah O, Jararweh Y, Yaseen Q. Multi-threading based map reduce tasks scheduling. *Int Conf Inform Commun Sys* 2014; 1-6.
 14. Hu W, Tan T, Wang L, Maybank S. A survey on visual surveillance of object motion and behaviours. *Sys Man Cybern Appl Rev* 2004; 34: 334-352.
 15. Lin HH, Liu TL, Chuang JH. Learning a scene background model via classification. *IEEE Trans Sig Process* 2009; 57: 1641-1654.
 16. Du Ming T, Shia-Chih L. Independent component analysis-based background subtraction for indoor surveillance. *IEEE Trans Imag Proc* 2009; 18: 158-167.
 17. Zhang W, Duan P, Lu Q, Liu X. A real time framework for video object detection with storm. *Ubiq Intell Comput* 2014; 732-737.
 18. Bose B, Grimson E. Improving object classification in far-field video. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recogn* 2004; 2: 2.
 19. Kurita T, Hotta K, Mishima T. Scale and rotation invariant recognition method using higher-order local autocorrelation features of log-polar image. *Proc ACCV* 1998.
 20. Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. *IEEE Trans Pattern Anal Mac Intell* 2007; 29: 411-426.
 21. Phillips PJ, Moon H, Rizvi SA, Rauss PJ. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans Pattern Anal Mac Intell* 2000; 22: 1090-1104.
 22. Guo G, Li SZ, Chan K. Face recognition by support vector machines. *IEEE Int Conf Autom Face Gest Recogn* 2000; 196-201.
 23. Stauffer C, Grimson E. Learning patterns of activity using real-time tracking. *IEEE Trans Pattern Anal Mach Intell* 2000; 22: 747-757.
 24. Zhang L, Li S, Yuan X, Xiang S. Real-time object classification in video surveillance based on appearance learning. *IEEE Int Vis Surveil Conjoint* 2007.
 25. A human computer interfacing application. *Int J Pharm Bio Sci* 201; 2: 29-31.
 26. Wang T, Snoussi H. Detection of abnormal events via optical flow feature analysis. *Sensors* 2015; 15: 7156-7171.

*Correspondence to

Maha Vishnu VC

Department of Computer Science and Engineering and Information Technology

Coimbatore Institute of Technology

India