Smart evaluation of videos for unusual-event identification in automated vehicle monitoring systems

Surjeet Yadav | Anu Sharma | Ananta Ojha

*Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Mechanical Engineering.

**Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science And Information Technology.

In this study, we look at various local monitors that may become unusual, as monitoring systems have different sensor types and perhaps propagating trustworthy information across cameras to improve detection accuracy. In this study, we look at various important issues and provide a new method for spotting certain abnormalities. We suggest an automated vehicle monitoring system (AVMS) to enhance detection accuracy. If the current reading from a certain local monitor is out of the ordinary, that monitor will issue a notification, and all of those alerts will go toward deciding whether or not an abnormal event has occurred. The effective implementation of any large-scale surveillance system relies on our algorithm meeting several parameters. It simply takes a few minutes to set up and is completely hands-free. The experimental findings demonstrate the effectiveness and simplicity of the suggested approach, with the assembled image quality being on the level with that of state-of-the-art techniques.

**Keywords:** unusual events, video analysis, automated vehicle, monitoring system, license plate recognition

1. Introduction

The use of video surveillance systems is becoming more widespread, and nowadays, you can find them installed in a wide variety of critical locations, such as airports, banks, public transit, or bustling city centres. People often have a positive attitude towards the enhanced feeling of safety provided by video surveillance, yet, they frequently have a negative attitude towards the associated loss of privacy. This valid fear often causes a delay in the implementation of video monitoring systems. The method uses an object-oriented representation of the environment as its foundation (Zhu et al 2020). The system re-renders a changed movie in response to the access control authorizations provided by the end user. Whenever an image is re-rendered, some parts will have their content removed. Thus, the information that is significant to the scene is maintained, but the details that are sensitive to privacy are not provided. In the same way, a privacy buffer employs privacy filters to operate on incoming sensor data to either prohibit access to sensitive information or change data to delete private communication (Nasaruddin et al 2020).

Figure 1 depicts a system diagram for the video surveillance system. Multiple inexpensive security cameras, either wired or wireless, are linked to a central computer to create this system. The server is then responsible for processing the video. To begin, a video analysis module will locate ROIs. The video is then compressed to be stored and sent with minimal resources. At the same time, the ROIs are subjected to a scrambling process (Hidayat et al 2020). Consequently, the persons in the scene are recognizable, but the background is unobscured, solving the problem of invasion of privacy. Finally, the system allows for Internet-based access from various client devices so that both live and archived videos may be viewed. Automatic number plate recognition is a method of reading a vehicle’s plate number from an image or video of the car in motion or at rest (Aslam et al 2022). This method works well for keeping an eye on moving objects. Powerful license plate readers have a wide range of applications, including thwarting theft, classifying illegal vehicles, collecting electronic tolls in a way that is unique to each location, tracking the flow of traffic inside a building, identifying speeders, identifying authorized cars in a parking lot, saving time by removing the need for human verification of parking passes, and much more. Issues such as noisy picture inputs, occlusion, vehicle orientations, various number plate kinds, additional images on number plates, non-standard sizes, low quality of the camera, and so on make number plate identification and recognition a difficult challenge. Existing systems often assume too simple conditions compared to real-world situations, such as only functioning with stationary cameras at a fixed viewing angle and resolution and only with a fixed license plate template (Mudgal et al 2021). As a result, automated activity identification is a need in many video surveillance uses. A proper automated vehicle
identification algorithm may significantly lessen the burden for people, making automatic vehicle recognition an integral aspect of video monitoring. Nonetheless, no universal vehicle identification system can reliably and automatically identify cars of interest across all surveillance applications. Vehicles of interest might vary depending on the application. It is challenging to develop a suitable general vehicle recognition algorithm because many surveillance applications need the activity recognition algorithm to deal with unusual events for which no training data and to be flexible enough to add new activities of interest to the system (Sharma and Gangadharappa). In this article, we go deep into the state of the art in video surveillance and provide solutions to various pressing issues. To improve detection precision, we propose using AVMSS. The recorded footage or pre-processed data is sent to the central monitoring facility for storage, analysis, and dissemination.

![Figure 1 Framework of a video monitoring system.](https://www.malque.pub/ojs/index.php/msj)

The rest of the sections of the paper are structured as follows. In Section 2, we give a literature review. Our detection system, is described fully in Section 3. In Section 4, we detail the experiment and its findings. Finally, we conclude the task in Section 5.

2. Related Work

The article (Singhet et al. 2019) presented a strategy for recognizing uncommon vehicles by first establishing a model for identifying more common cars. The issue of insufficient training data may be addressed using this approach, which suggests how to do it. However, they generate all unusual event models by altering the basic ordinary event model. This is even though, in nature, normal vehicle recognition and odd occurrences might be highly distinct. The detection, localization, and categorization of various objects in images have each been the subject of several strategies presented in the research literature (Tamilmani et al. 2022). The issue with automated LPR (ALPR) is that it recognizes the number of a vehicle’s number plate based on photographs of the plate. Extensive image processing, including object identification and pattern recognition, is required to implement the solution. Plate variances from country to country, the plate’s position inside the image, varied widths, diverse color combinations, plate contamination owing to other frames, and photographs are some of the issues associated with proper number identification. The study (Ahmed and Jeon 2021) presented an extensive and completely automated system for the detection and identification of license plates. It was constructed using a series of deep CNNs in conjunction with algorithms. The CNNs underwent training and optimization to improve their resilience in various situations, including changes in position, lighting, occlusion, and others. These CNNs are versatile enough to operate with a broad range of license plate templates, each with a unique size, background, and font combination. This article discussed several facets of LPR and offered solutions to all of the issues that were brought up. The research, however, is limited to still photographs and includes no videos. The article (Çengönül et al. 2023) investigated the effect that these external influences had on the accuracy and consistency of the LPR procedure. Some works have expressly concentrated on tackling these issues as their primary emphasis. One of the ways to make ALPR resistant to these challenges is to recreate all of the different sorts of challenges in a manufactured database and then test a classification algorithm on this dataset to come up with a solution that has a good chance of working. The automatic Number Plate Recognition System (Hatırcan et al. 2020) used a pipeline based on image processing. They used morphological approaches together with edge detection, histogram manipulation, and plate localization for character segmentation and plate localization. They used a synthetic neural network for character identification and categorization. However, the work’s character identification and detection procedures are simplistic and do not account for restrictions such as angle, distance from the vehicle, or poor picture quality. However, this technology is only applicable for use in offline detection, and no information is provided on how to adapt it for use in real-time. The study (Ghoniem et al. 2022) examined the challenge of LPR when faced with various designs and fonts. A demanding synthetic dataset with different font styles and number plate designs was used for testing after multiple templates matching for character identification and noise reduction. It has been shown that both the false positive and false negative rates are rather low.
3. Proposed Methodology

In this study, we investigate the problem of AVMS resolution in multi-camera video surveillance systems. In this study, we examine the issue of AVMS in multi-camera methods used for security purposes. Each screen is a separate object that takes local, low-level observations from the input video. The monitor’s observation may be the amount of local optical flow or the flow direction at present. The automated monitoring system performs real-time analysis of the video streams, searching for any abnormalities or unexpected occurrences using powerful computer vision algorithms. The system also checks the car’s actions against normative driving characteristics and models.

3.1. Identify the suspect’s vehicles in a video frame

To speed up the process, we should eliminate the impossible frames, which contain no relocating objects. Background estimate in video sequences is the starting point for this effort. Often, dirty backgrounds may be removed by averaging k frames from a video clip. By averaging, we can remove all of the clutter from the foreground and be left with a clear backdrop. Here’s how we can write out the equation (1):

\[ P(y, x) = \left( \sum_{l=1}^{r} l(y, x) \right) / r \] (1)

Where \( l(y, x) \) = input video sequences,
\( P(y, x) \) = estimation background,
\( r \) = number of average frames,
\( r \) = set 50.

After obtaining an estimated backdrop, we compare it frame by frame using the difference image \( l(y, x) \) to get images of moving objects, in this case, potential images of a car driven by a suspect. The following equation (2) gives its numerical value:

\[ L(y, x) = \begin{cases} 
|l(y, x)| & |P(y, x) - l(y, x)| \geq Th \\
0 & |P(y, x) - l(y, x)| < Th
\end{cases} \] (2)

Where \( Th \) = judgment threshold for determining whether or not the pixel is changing. In our study, we use a \( Th \) (30) for \( l(y, x) \) to calculate whether or not a video sequence consists of frames of a moving object; if it does, then the sequence is used as input for the next processing stage.

3.2. Removing a suspect’s car from a video frame captured by many cameras

In this phase, we suggest using a colored-based component technique to scan footage from various cameras to roughly ensure that suspicious vehicles are present in video frames from numerous cameras. The portions of the sequentially tone mapping \( l(e^*(y, x)) \) and \( l(p^*(y, x)) \) may be determined by equation (3):

\[ l(e^*(y, x)) = l(\text{sl}(e^*(y, x)) - \text{ls}(e^*(y, x))) \]
\[ l(p^*(y, x)) = l(\text{sl}(p^*(y, x)) - \text{ls}(p^*(y, x))) \] (3)

The contrast image and noise remove the image of the suspect’s car in the video frames corresponding to \( \text{ls}(e^*(y, x)) \), and \( \text{ls}(p^*(y, x)) \). The contrast image and noise removal image of the suspect’s car in video frames taken by various cameras are represented by the functions \( \text{ls}(e^*(y, x)) \), and \( \text{ls}(p^*(y, x)) \) accordingly.

The next step is to use a fusion technique to roughly confirm the legitimacy of video frames featuring a suspicious car. The other images will be filtered to determine the true identity of the vehicle in question.

\[ T(y, x) = \begin{cases} 
\text{ls}(y, x) & (l(e^*(y, x) < D)(l(p^*(y, x) < D) \\
0 & (l(e^*(y, x) \geq D)(l(p^*(y, x) \geq D)
\end{cases} \] (4)

In this study, \( D \) was dependent on the color values of the suspect’s vehicle, and \( T(y, x) \) represents the outcomes; if it is smaller than \( D \), then \( \text{ls}(y, x) \) is the suspect vehicle frames.

3.3. License Plate location

To facilitate accurate plate localization, the image has been optimized. To begin, the image of the car is cleaned of noise to improve accuracy. Several morphological procedures were conducted once the image was converted to grayscale, which brought out sharper edges and more contrast. The next stage is to apply blurring to provide consistency throughout the image, and then adaptive thresholding is used to eliminate irrelevant details (things that aren’t license plates). Methods such as grayscale conversion, histogram equalization, sharpening, and OTSU masking are often used.
The license plates x and y coordinates must be determined once the image has been cleaned of background distractions. Algorithm 1 outlines the procedures that must be followed for this to occur. From the pre-processed image, contours are extracted. According to the Open CV docs, a contour is a line drawn across a set of points that shares its color or intensity with the whole group. The outlines may be used for several purposes, including shape analysis and identifying objects. Plate characters’ contour coordinates were utilized to extract regions of interest, and their respective bounds, areas, aspect ratios, widths, heights, and angles were recorded as Python class objects representing every contour character. The potential plate set is then utilized as input in the subsequent step. This plate detection from imagine process has been encapsulated in Algorithm 1.

**Algorithm 1: License plate detection**

**Input:** Image file name as imageOriginal

**Output:** sequence of lists, each of which provides the contours necessary to complete the license plate.

\[
\text{imageGrey, imgThresh ← Preprocess (imageOriginal)}
\]

\[
\text{Contours ← opencv.findContours (imageOriginal)}
\]

\[
\text{listofpossibleChars=[]}\]

For c in outlines:

\[
x,y,w,h←opencv.boundingRect (c)
\]

If \(z + u > 80 \text{ and } u > 2 \text{ and } z > 8 \text{ and } 0.25 < \left(\frac{u}{z}\right) < 1.0\):

\[
\text{listofpossibleChars. Append (c)}
\]

\[
\text{listOfListsOfMatchingCharsInScene← GroupMatchingPlates (listofpossibleChars)}
\]

\[
\text{listOfPossiblePlate← []}
\]

For plate in listOfListsOfMatchingCharsInScene:

\[
\text{Possible plate← extract plate (plate)}
\]

If possible, plate. imgplate not none:

\[
\text{listOfPossiblePlate.append (possible plate)}
\]

Return listOfPossiblePlate

Several processes, such as contour detection, geometric constraint setting, grouping, etc., are employed for the preprocessed data in Algorithm 1. These procedures take the original picture and remove all contours, filter out the ones that don’t seem like characters based on their geometry, and then extract the number plate. The method was applied to every conceivable dish, and the results were compiled into a single, convenient list. The most likely candidate license plates for this item have been selected.

The input image has been processed to remove the noise in Figure 2. When talking about an input image, "noise reduction" signifies the method of eliminating distracting and unwelcome visual artifacts and disruptions so that the final product is more pleasant to the eye (Hashmi et al 2019). Figure 2a depicts an automobile’s input image, whereas Figure 2b shows the output. The intensity values in a grayscale picture generally run from 0 to 255, with 0 denoting black (the lowest possible intensity) and 255 denoting white (the highest possible intensity). Shades of grey are represented by the intermediate values. Image contrast enhancement used to create Figure 2c. Enhancing contrast in a picture is making the distinction between one part of the image and another more apparent in terms of brightness and color. Its purpose is to enhance the picture such that it is either aesthetically pleasing or analytically more straightforward. The image in Figure 2d was blurred to get rid of the noise. In image processing, blurring a picture to eliminate noise is a typical method for smoothing out imperfections. Blurring a picture to get rid of noise smoothes out the blemishes and makes the image seem better overall. Image noise may result from several sources, including but not limited to sensor limits, transmission mistakes, and ambient variables. To improve image quality, noise reduction attempts to lessen the effect of noise while keeping crucial image information intact and preventing the introduction of distracting artifacts.

![Figure 2 Image preprocessing procedure.](https://www.malque.pub/ojs/index.php/msj)
3.4. Character Segmentation

License plates that have been cropped for OCR may now be seen here. The first step in using OCR is to isolate the characters on license plates using cropping. In most cases, the license plate outline is too small for accurate character recognition. This meant that any unnecessary symbols on the plate needed to be cleaned. The methods for image processing described in the second phase were used once more to eliminate the noise in the plate image. Figure 3 shows how they extracted objects and removed everything except characters from the plate's image that could be read as numbers or letters.

![Figure 3 OCR and Noise Reduction for License Plates.](image)

The images were then scaled up, made grayscale, and had their contrast raised. The entire image is then converted to black and white using dynamic thresholding. To produce black text on a white backdrop, use the binary inverse. The contrast made it easier to distinguish across character types. Individual characters were chopped out of the license plate together with the other characters except the 36 (0-9 and A-Z) characters. Each divided character has its height and breadth raised by factors of 1.3 and 1.6. Accordingly, a white border has been included.

To prepare each character picture for the character recognition module, cropping, and pre-processing are performed. Specified number plate outlines are read in as input at this stage. One of the plate's characters is represented by each method. Plate images have had thresholding performed on them to remove the contours of certain characters. Better OCR is achieved by adding borders and enlarging the clipped picture. The processed and cropped text image is then given into the character model, which determines which symbol this outline is and associates it with the license plate phrase.

4. Result and Discussion

Here, we provide experimental findings from several security cameras using our suggested AVR. At the moment, we only capture a single set of test data from many cameras on the same day using known cars. Precision and accuracy are the two most often used indicators. We evaluate our suggested technique in comparison to state-of-the-art approaches like deep neural networks (DNN) (Li et al 2018) and cognitive Internet of Vehicles (CloV) (Arooj et al 2022). Accuracy and precision are two characteristics of a highly effective system.

![Figure 4 Accuracy.](image)

The proportion of useful information included within the total quantity of data collected is shown and described in Figure 4. There is a pace at which accuracy can be assessed. It’s a measure of how likely the information being obtained is useful. In comparison to the 88 and 90 accuracy scores achieved by the DNN and CloV, respectively, our suggested technique achieved 99 accuracy. Compared to the other two approaches, the efficiency of the accuracy value is higher. The accuracy of the proposed and current approaches is compared in Table 1.

<table>
<thead>
<tr>
<th>Table 1 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>DNN</td>
</tr>
<tr>
<td>CloV</td>
</tr>
<tr>
<td>AVMS [Proposed]</td>
</tr>
</tbody>
</table>
Precisely how much useful information has been collected from the universe of available useful information is shown and described in Figure 5. The rate of precision is a measurable quantity. It's a measure of how likely useful information will be found. While the DNN achieved a precision of 85 and the CloV of 78, our suggested approach achieved a score of 92. Compared to the other two methods, the precision value is the most effective. The precision of the proposed and current systems is compared in Table 2.

<table>
<thead>
<tr>
<th>Table 2 Precision.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision (%)</td>
</tr>
<tr>
<td>DNN</td>
<td>85</td>
</tr>
<tr>
<td>CloV</td>
<td>78</td>
</tr>
<tr>
<td>AVMS [Proposed]</td>
<td>92</td>
</tr>
</tbody>
</table>

AVMS systems use deep learning algorithms to analyze and decipher several vehicle data sources, providing real-time monitoring and analysis for increased security, effectiveness, and decision-making. DNNs are computationally demanding models that need a lot of memory and computing power. It may not be easy to implement and operate these models in real-time applications, such as autonomous vehicle monitoring systems, and may need specialized hardware or distributed computer configurations. By using the capabilities of cognitive computing, CloV empowers AVMS to make wise and educated judgments. Large volumes of real-time data from several sensors and sources, such as vehicle sensors, cameras, GPS, traffic data, and meteorological conditions, may be analyzed by it. This improves safety and efficiency by enabling prompt reactions to possible dangers, traffic jams, and other urgent circumstances. Integrating numerous technologies, infrastructure, and communication networks is necessary to implement CloV systems. Costs for development, deployment, and maintenance may rise as a result of this complexity. It may also be difficult to guarantee compatibility and standardization across various cars, manufacturers, and infrastructure suppliers. By continually monitoring cars and their surroundings, AVMS contributes to increased road safety. Compared to other approaches, AVMS is superior.

5. Conclusion

In this study, we provide a new approach to AVMS for use with a network of cameras specifically designed for monitoring public spaces. It is possible to use connected component analysis and an adaptive LPR approach in a single license plate location algorithm. We employ shift filtering, a technique for segmenting text, to break down license plate data. For vehicle video sequences, you may also apply adaptive balanization for segmentation. To demonstrate the efficiency and resilience of the proposed approach, we also ran extensive simulations utilizing a variety of situations. Based on our experiments, we know that 99 percent of our AVMS detections are accurate. A particular LPR system's processing rates may be further increased by employing the parallelization strategy during different phases of development.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.
Reference


