ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijsem.org
Vol 15, Issue.1 Jan 2021

Use of a deep learning system for independent accident recognition in caves with poor CCTV video

P Nageswara Rao, SK Mastanbi, Subba Rayudu

Abstract—

The foundation of this study's deep learning technology is a combination of the Object Detection and Tracking System (ODTS) and the widely-used Faster Regional Convolution Neural Network. Accidents such as (1) Wrong-Way Driving (WWD), (2) Stopping, (3) People Exiting Vehicles in Tunnels, and (4) Fires may be automatically detected and tracked using a (Faster R-CNN) for Object Detection and a Conventional Object Tracking method. Based on the Bounding Box (BBox) findings from item Detection, ODTS assigns ID numbers to each moving and detected item by comparing the Boombox from the current and previous video frames. Typical object detection frameworks are able to do what our system does—follow an item as it moves over time. We obtained average accuracy (AP) values of 0.8479 for autos, 0.7161 for individuals, and 0.9085 for flames after training a deep learning model in ODTS using datasets of tunnel event images. Then, using four films, one for every accident, the ODTS-based Tunnel CCTV Accident Detection System was evaluated using the trained deep learning model.

I. INTRODUCTION

Object detection technologies have allowed for the precise measurement and localization of specific items in both static and moving media. Cancer diagnostics, autonomous vehicles, and closed-circuit television (CCTV) security systems are just a few of the many new uses that have recently surfaced. Object tracking is another subfield of image processing that makes use of distinct identifiers to monitor the movement of objects over time. Be that as it may, tracking moving things begins with recognizing them in a static picture. Therefore, it stands to reason that the object detection's execution quality should significantly impact the object stacking outcomes. This object tracking system has been successful in many different applications. Countless instances abound, including but not limited to: protecting sensitive locations, tracking a moving target (such as a person or car), monitoring crime scenes using traffic cameras, and many more. In the framework of traffic management, this research investigates and controls traffic conditions using automated object identification as an example. Outlined below are the key aspects of every

Assistant Professor[1,2,3]
(Department of Computer Science Engineering)
priyadarshini institute of technology

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

section. The authors of [1] assert in their paper that they have created a system to identify automobiles that driverless vehicles might use. Within a car, this technology can detect and classify many items. employing a convolutional neural network to identify the vehicle type. Using the vehicle object tracking approach, the tracking center point is moved around the picture to follow the identified vehicle object. The technology then displays a bird's-eye picture of the region with the visible vehicle objects and estimates the distance between the driver's automobile and them. The system's approach aids the autonomous driving system by allowing an outside perspective of the vehicle's location. Consequently, the camera is able to precisely pinpoint the location of the moving vehicle to within 1.5 meters in length and 0.4 meters in width. In order to keep an eye on traffic on city streets and highways from above, the researchers at [2] created a deep learning detection system using Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs). The system takes in the satellite picture, analyzes it using a CNN to extract features, and then uses the binary classification of a support vector machine to locate the vehicle's BBox. Additionally, Arinaldi, Pradana, and Gurusinga [3] created a method to assess traffic, categorize vehicle types, and predict travel times. This method makes use of BBox data that is derived from object recognition in videos and images. We evaluated the system's methods against two popular options: the quicker RCNN and the more conventional Gaussian Mixture Model + SVM. It seems that R-CNN was successful in identifying the position and kind of vehicle at a quicker rate.

II. DEEP LEARNING-BASED OBJECT DETECTION AND TRACKING SYSTEM

A. Concept

Figure 1[7] shows how the ODTS identifies and tracks things in space and time. By now, ODTS should have learned enough to identify items in a given photo frame with reasonable accuracy. A trained object detection system determines the number of things in the scene at time T. The ODTS learns the coordinates of n BBoxes from video frames that it receives at intervals of c. The corresponding class of each noted item

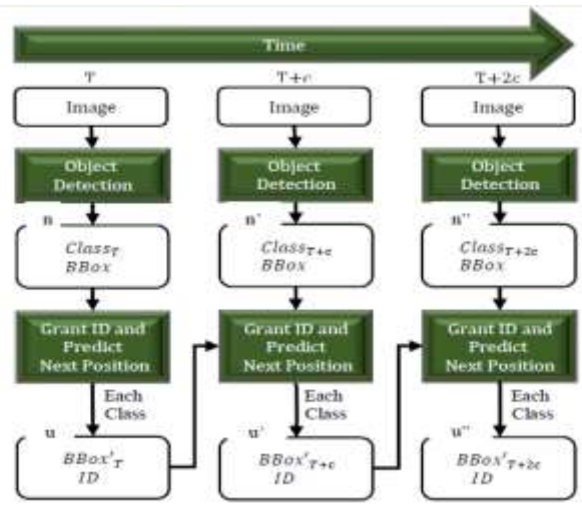


Figure 1: A timeline of the object identification and tracking technique implemented by the system. Once object detection has acquired class and BBox, the object tracking algorithm assigns an identification and makes predictions. place in time dependent on both the here and now and the history BBox.

The object detecting system quickly assigns a label to an item. Afterwards, a module that tracks dependent items is configured to utilize the observed object data to identify each item and anticipate its future move. There are less than or equal to n monitored boxes in your system at the moment. Assuming no objects have been previously monitored, the total number of tracking BBoxes will match the number of objects. At time T+c, when u is equal to zero, u' is equal to n'. The current tracking BBox is

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

generated from the items observed in each class, as the previous tracking BBox was unavailable. The SORT method[5] was used to build this object tracking module. It follows several objects with the same ID number using the principle of Intersection Over Union (IOU). To see where the identified items are moving next, the Kalman filter and Hungarian algorithm were used.

Time T+c follows a similar pattern to time T, where the object detection module was used to extract and C from the freshly given picture.

The item is considered to have recently entered the RoI if there is no object pair with an IOU value greater than 0.3 for any item at T+c. A unique identifier that is not already in use will be allocated to the freshly created object. A quicker RCNN learning algorithm is used for object detection in this system, while a SORT is used for ID assignment and tracking. Using SORT for multi-object tracking at a degree speed of 100-300 frames per second is definitely doable [6]. The object tracking capabilities were affected by the video frame interval c[7], as the system used the SORT[6] algorithm based on the IoU value. You may reduce the overall processing burden while watching a movie by changing the interval between the object detection network's detections. Experimental results corroborated this by monitoring things for as long as six frames, proving the capability to track objects over frame intervals [7].

Because object tracking performance degrades dramatically with increasing frame intervals, the video frame interval has to be adjusted based on the total number of linked cameras. using deep learning to communicate with a server.

B. A Method for the Prompt Prediction of Tunnel Crash Due to the reduced space for manoeuvre in the case of an emergency, driving via a road tunnel is inherently riskier than driving on an open road [7]. This is why four things—individuals, fires, stops, and WWD—must be tracked according to South Korean legislation. Tunnel driving is riskier than driving on a standard roadway since there is less room to get out of the car. That is why it is crucial to inform drivers as soon as possible in case of a tunnel emergency [7]. The national legislation of Korea mandates that all sensors in the country must record and identify "Person, Fire, Stop, and WWD." [4].

Concurrently, targets and any suspicious behavior are monitored by subsurface CCTV.

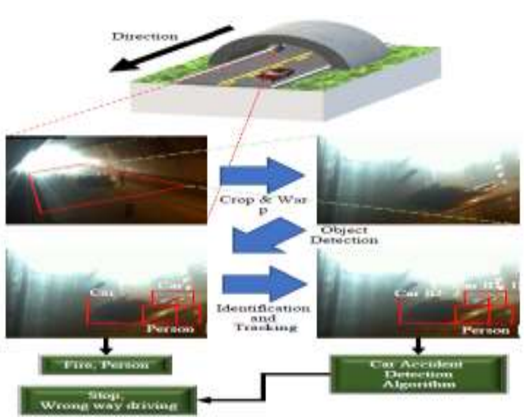
The use of an automated object detection system for the targets will help accomplish this aim. The system's performance outside of tunnels has been outstanding. However, the apparatus is completely useless when used inside a tunnel. Because (1) the tunnel was quite dark, the film was negatively impacted by the tail light or warning light of the driving car. A foreboding atmosphere pervaded the movies filmed in the tunnel. So, it's not the same hue as the roadway that runs outside the tunnel. The video surveillance system was not expected to work effectively once within the tunnels for the reasons mentioned above; it was designed for use on roadways outside the tunnels. Consequently, it is crucial to have an automated system that can identify accidents in road tunnels.

In order to address these issues, researchers created a Tunnel CCTV Accident Detection System using deep learning in [7]. The training process made use of a deep learning model known as Faster R-CNN. The algorithm was also based on a previous one

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

that had been trained using picture datasets that contained underground accident photos. The next step is for ODTS to use an object tracking mechanism that is particular to cars. As a result, CADA is periodically triggered to create Stop and WWD events depending on the position of the monitored item. The ability of CADA's approach to identify an accident condition is seen in Figure 2. After retrieving the ROI from the tunnel's CCTV screen, crop and distort the original screen photo to bring it into alignment with the ROI. This strategy seeks to standardize the detection of Stop and WWD events, similar to [1]. By resizing nearby and distant items to the same size and cropping the picture to show just the ROI, the image extract makes the image more trainable. Unlike [1], these are new ideas.Using a Faster RCNN that has already been trained to detect human, fire, and vehicle activity is the next stage [5].We manually specified the object class to create an additional 'No Fire' object in order to lower the frequency of the wrong response for the Fire object. Objects that can be deceiving, such as tunnel lights or automobile taillights, are indicated with the No Fire object. With the exception of the backdrop, the data's properties mirror the Faster R-CNN training object class classification. This strategy has the potential to reduce the occurrence of fire misdetection when utilizing untrained data.



Using deep learning, researchers developed a Tunnel CCTV Accident Detection System in [7] to tackle these problems. A deep learning model called Faster R-CNN was used for training purposes. Additionally, the model was built upon an earlier model that had been trained on photo datasets that included accidents that occurred in subterranean passages. The next step is for ODTS to use a Car-specific object tracking function. Every so often, it will use the data from the Car Accident Detection Algorithm (CADA) to generate Stop and WWD events. Figure 2 shows that CA DA's procedure could detect an accident state. In order to align the original CCTV screen shot with the recovered RoI from the tunnel's CCTV screen, crop and distort it. Similar to [1], this approach aims to standardize the detection of Stop and WWD events. The image extract makes the image more trainable by cropping it so that just the ROI is visible and by making all objects, whether close by and far away, the same size. These are novel concepts, in contrast to [1].

The next step is to use a Faster RCNN that has already been trained to identify human, fire, and vehicle activity [5]. To reduce the frequency of the incorrect answer for the Fire object, an extra 'No Fire' object was constructed by manually specifying the object class. Things like tunnel lights, car taillights, and other potentially misleading objects are marked with the No Fire object. Except for the background, the data's characteristics reflect the object class's categorization in Faster R-CNN training. Fire misdetection using untrained data might be mitigated using this approach.

$$\mathrm{IoL} = \frac{\textit{Overlapped Length of Vertical element of BBox}}{\textit{Union Length of Vertical element of BBox}} \quad (1)$$

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

In response to these issues, a deep learning-based Tunnel CCTV Accident Detection System was created by the researchers at [7]. During the training process, the Faster R-CNN deep learning model was used. An extra perk is that the model was based on another model that was trained using picture datasets that included tunnel-related events. Subsequently, the ODTS makes routine use of the target car's tracking data to trigger Stop and WWD events using the Car Accident Detection Algorithm (CADA). Car objects are the only ones that can benefit from this specialized object tracking feature. As shown in Figure 2, CADA's technique has the ability to recognize an accident condition. Reshape the first CCTV screen capture so it fits inside the newly-recovered ROI from the tunnel's CCTV video. Building on the work of [1], this method aims to standardize the detection of Stop and WWD occurrences. It is easier to train the picture once it has been "extracted," which means shrinking the image to show just the ROI and resizing both nearby and faraway objects to the same size. Unlike [1], these are novel ideas.

After that, use a Faster RCNN that can distinguish between people, flames, and automobiles [5]. An additional 'No Fire' object was created by explicitly defining the object class in order to decrease the frequency of the wrong response for the Fire object. The No Fire label is used on indicators that might deceive drivers, like tunnel lights, automobile taillights, and so on. Aside from the backdrop, Faster R-CNN training classifies objects based on data attributes.

Applying this approach to untrained data may lead to a reduction in false fire alarms.

III. EXPERIMENTS

The built-system experiments in this study are divided into two parts: one that evaluates the performance of the system in detecting accidents, and another that evaluates the performance of deep learning. whole building. An important factor influencing ODTS SORT is the accuracy of object recognition. Hence, getting the deep learning object identification network's learning right was crucial for the system's completion. The system was tested to see whether it could identify the four distinct accident situations after the deep learning model had been trained. Given the need for both the object identification performance of the deep learning model and the discriminating skills of the CADA, it was important to test the system for each photo to see whether it could detect each situation. Introducing Deep Learning: A Primer Instead of training on a moving video, the deep learning network was shown a series of static images and told to identify items in them. The study here uses the concept of an epoch to denote a single cycle of the training operation throughout the whole dataset. Some of the data collection that will be examined includes images taken at the site of the accident. Training was conducted using the more efficient R-CNN [5].

TABLE I. THE STATUS OF USED IMAGE DATASET

Number of Videos	Number of images	Number of objects		
		Car	*Fire*	*Person*
45	70914	427554	857	44862

Table.1 shows the training data set as it is right now.

These 70,914 video stills were assembled from the frames of 45 different videos. The training technique of deep learning deviates from the standard deep learning procedure in that it does not split the data into learning and inference sets. Unlike publicly available datasets, the data set used in this study has

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

continuous visuals throughout each movie, which helps explain why. You might argue that the videos all have the same basic structure, but the pictures inside them change depending on what's there or not. Whether the training and inference data were separated for each image would have no effect on the object identification network's inference performance.

Because the object identification stability on the complete video might drop, which adversely influences the accident detection performance, evaluating the detection strategy of the whole tunnel CCTV image accident detection system is tough. Consequently, we trained on all available data and now we're testing how well deep learning does object recognition using the data we learned. There is a dearth of Fire things since tunnel fires are so rare.

A low false alarm rate is crucial for the tunnel control room since both missing and false fire alarms are conceivable. Notifying an installed system several times that a false detection has occurred when none has is very unlikely to make it dependable. On the other hand, the enriched data from the time-lapse dataset is often added to the training data set, which might automatically increase the detection performance. The number of objects marked as "No Fire" was much higher than the number of objects marked as "Fire" in this experiment that aimed to explore ways for reducing erroneous detection. The inclusion of 10 epochs expedited the training of R-CNN. As a deep learning framework, Linux was used with Tensor-flow 1.3.0 [7]. Faster R-CNN training makes advantage of NVIDIA GTX 1070 graphics cards. The whole training period lasted for 60 hours, and the mean (AP) was used to measure the correctness of each class's inference evaluation.

TABLE II. INFERENCE RESULT OF DATASET

Number of images	Average Precision (AP)		
	Car	*Person*	*Fire*
70914	0.8479	0.7161	0.9085

Table displayed the AP values for the three things that needed to be detected. The training data set is dominated by cars, which also happen to have the greatest AP value of any item class. Car produces better outcomes than other object kinds.
The video's Car was expected to have reliable deep running object detection performance, in other words. The AP for the Person object is low, as shown in Table.2, due to the item's small size and long, thin shape. Due to the small number of training items (857), erroneous identification was possible, despite the Fire object's relatively strong accuracy percentage (AP) of 0.9085.
The amount of false positives may be reduced, however, by training deep learning on non-fire related data. However, in order to use them to detect the fire in the tunnel control center, more images of fire occurrences required to be gathered and included into training.

B. Testing the Capability of a Full-Tunnel CCTV System to Detect Accidents The deep learning model that was trained provides the basis for the performance. The efficiency of the Tunnel CCTV Accident Detection System, which is based on deep learning, has to be evaluated. Table 3 shows that four films were selected for this objective. A piece of software was created so that the detection results could be shown on the video. The discovery was made after just 10 seconds of visual observation when the video frame interval was set to 6 fps at 30 fps[7]. You may see a brief synopsis of the

ISSN2454-9940
www.ijsem.org
Vol 15, Issuse.1 Jan 2021

video's running length, occurrence time, and detection time in Table 3.

TABLE III. DETCTIED TIME OF THE EACH ACCIDENT BY ACCIDENT DETECT SYSTEM

Accident video information	Item on video time		
	Video length	*Occurrence time*	*Detected time*
Stop	126s	5s	7s
Wrong Way Driving	29s	4s	12s
Fire	64s	29s	29s
person	72s	50s	50s

Table 3 displays the lag time that exists between the occurrence of the Stop and WWD events and their subsequent detection. Because this is a hallmark of CADA, we find that it occurs once every Repeats every 2.4 seconds. However, the algorithm could tell that there was a delay of 8 seconds between Stop and WWD. Images, such as Person and Fire, on the other hand, revealed speedy identification right after the incident. Table 3's graphics, however, were only utilised for training, which means the final product may look different if put in the field. Hence, the tested was put to use, and further test movies were shot.

IV. CONCLUSION

Combining an object tracking algorithm with a deep learning-based network is the novel approach to object detection and tracking (ODTS) that is suggested in this article. The authors demonstrate how this technology may be used to track the motion of an object in real-time. You could come across and put to use an item that falls into a certain category. The speed of object detection is critical in ODTS object tracking as SORT relies only on BBox information rather than a picture. It is arguable if continuous object detection performance is necessary unless the object tracking method relies heavily on precise item identification.

And for tunnels, there's a CCTV accident detection system that uses ODTS. An experimental environment was created to identify system-wide mishaps, and a deep learning object identification network was tested in it. This system employs CADA to distinguish between cycles by using the vehicles' real-time data. We reduced the time it took to identify an accident from 10 seconds to under a minute by playing with their visuals. A trustworthy vehicle object's object identification performance was well-protected by deep learning training, but a person's performance was severely lacking.

However, the amateur film lacks fire objects, thus there's a good chance of false detection. Conversely, further training on non-fire items may reduce false detection. Future work on securing the fire image could lead to an improvement in the deep learning object recognition network's performance while recognizing fire objects. As an example of a CCTV accident detection system in a tunnel, the ODTS is helpful. However, it may also be used to areas where monitoring the motion of an item is crucial, such as predicting vehicle speeds or preventing unlawful parking. It is critical to protect many images in addition to Fire and Person objects to boost the system's dependability. It is also possible to increase system dependability via the installation and continuous monitoring of the tunnel management website.

REFERENCES

[1] Robotics and Autonomous Systems, 2019, vol. 112, pp. 178-189, "Bird's eye view localization of surrounding vehicles : longitudinal and lateral distance estimation

ISSN2454-9940
www.ijsem.org
Vol 15, Issue.1 Jan 2021

with partial appearance" (E. S. Lee, W. Choi, D. Kum).
The paper "Robust vehicle detection by combining deep features with exemplar classification" was published in Computerizing in 2016, volume 215, pages 225-231 and was written by L. Cao, Q. Jiang, M. Chang, and C. Wang.
the third In 2018, the authors Arinaldi, Pradana, and Singur published an article titled "Detection and classification of vehicles for traffic video analytics" in the Proceeding of the Computer Science, volume 144, pages 259-268.

"Development of a deep-learning based automatic tunnel incident detection system on CCTV," published in 2018 in the Proceedings of the Fourth International Symposium on Computational Geo dynamics, was written by K. B. Lee, H. S. Shin, and D. G. Kim and can be found on pages 140–141.

,

[5] In the 2015 Proc. Neural Information Processing Systems, the authors Ren, He, Girshick, and Sun discuss "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks" (pp. 91–99).

In the 2016 IEEE International Conference on Image Processing, the authors Beckley, Pyongyang, Otto, Ramos, and Upcroft presented a paper titled "Simple Online and Real-time Tracking," which can be found on pages 3464–3468.

Korean Tunneling and Underground Space Association, 2018, vol. 20, K. B. Lee, H. S. Shin, and D. G. Kim, "Development of a deep-learning based automatic tracking of moving vehicles and incident detection processes on tunnels."