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 11, Issuse. 1 Jan 2020

Utilizing a deep learning system to identify accidents on its own in tunnels with insufficient CCTV footage

Dr P RAMESH BABU , J APOORVA

Abstract—

In this study, the widely used deep learning network Faster Regional Convolution Neural Network is combined with the Object Detection and Tracking System (ODTS). The usage of a (Faster R-CNN) for Object Detection and a Conventional Object Tracking approach will be used to automatically identify and track unexpected events including (1) Wrong-Way Driving (WWD), (2) Stopping, (3) People Getting Out of Vehicles in Tunnels, and (4) Fires. Based on a comparison of the Bounding Box (BBox) from the current and previous video frames, item Detection's Object Tracking and Recognition System (ODTS) assigns ID numbers to each moving and recognized item. Our approach succeeds where traditional object detection frameworks fail because it can track a moving object over time. When a deep learning model in ODTS was trained with datasets of event images in tunnels, it was able to attain Average Precision (AP) values of 0.8479, 0.7161, and 0.9085 for target objects of automobiles, people, and flames, respectively. The trained deep learning model was subsequently used to an evaluation of

the ODTS-based Tunnel CCTV Accident Detection System, which included four videos featuring each accident.

I. INTRODUCTION

With the use of object identification tools, we can now precisely size up and pinpoint objects in both static and dynamic images and videos. Several applications have emerged, particularly in areas like autonomous vehicle technology, security and surveillance through CCTV, cancer diagnostics, etc. One such use of image processing that requires unique identification and monitoring of the positions of identified objects throughout time is object tracking. To begin following a moving target, however, object identification in a single static image must be accomplished. The results of object stacking, it might be argued, should thus be highly dependent on the efficacy of the object detection used. Successful uses of this object tracking technology include monitoring an accident scene with a traffic camera, keeping tabs on criminal activities or defending a vulnerable place, and following a moving target like a person or an automobile. As a case study in traffic management, this study analyzes and manages traffic conditions via the use of automated object recognition. The following are synopses of the various chapters. The authors of [1] claim that they have developed a vehicle detection system that may be utilized in conjunction with an autonomous vehicle. This system can identify a car as a distinct object type. vehicle classification using a convolutional neural network (CNN). The vehicular thing

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

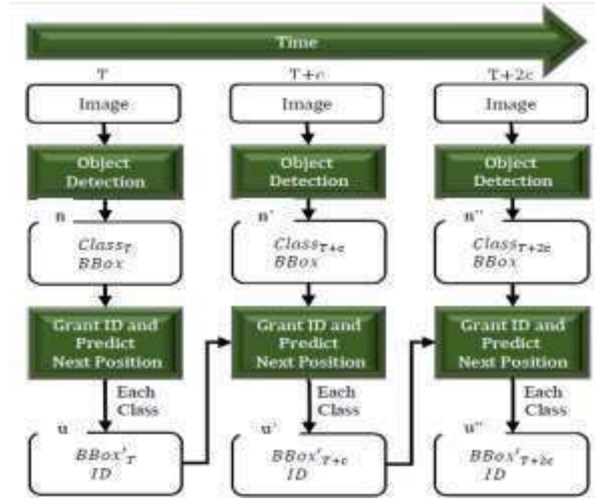
ISSN2454-9940
www.ijsem.org
Vol 11, Issuse. 1 Jan 2020

When a vehicle is detected in a photograph, the tracking algorithm spins the tracking center to follow the item. After calculating the distance between the driver's car and the visible vehicle objects, the screen will provide a zoomed-in, bird's-eye view of the area including the seen vehicles. The technique of the system allows for an outsider's view of the vehicle's position, which is helpful to the self-driving system. To an accuracy of 1.5 meters in each direction and 0.4 meters in any horizontal plane, the camera can track a moving vehicle. Researchers at [2] developed a CNN and Support Vector Machine (SVM)-based deep learning-based detection system to monitor urban street and highway traffic from above. The satellite image is used as an input, fed into a convolutional neural network (CNN) to extract features, and then subjected to a binary classification by a support vector machine in order to pinpoint the vehicle's BBox. In addition, Arinaldi, Pradana, and Gurusinga [3] developed a method to estimate the amount of traffic, classify the sorts of vehicles on the road, and foretell how long trips would take. This technique employs BBox data obtained via visual product identification. The method used by the system was compared against both the more time-efficient RCNN and the more traditional Gaussian Mixture Model + SVM. A faster-processing R-CNN seems to have correctly detected the location and vehicle type.

II. A DETECTION AND FOLLOWING SYSTEM BASED ON DETAILED LEARNING OF OBJECTS

A. Concept

Figure 1[7] depicts how the ODTS finds and tracks things throughout time. It is assumed that ODTS has been trained enough to correctly identify items in a given image. The ODTS is fed video frames at regular intervals of c, from which it learns the positions of n BBoxs, where T is the total number of objects in the picture. The correct grouping for every item that has been seen



Timeline of the item identification and tracking process used by the item identification- Tracking system. class and BBox are gathered from object detection, the object tracking algorithm assigns an identification and generates predictions about the item's future location based on the current and historical BBox.

When an item is detected, it is instantly given a label by the system. In the future, a dependant object tracking module will be triggered utilizing the observed object data to uniquely identify each item,, and predict its next position, ′. U is keeping tabs on more than n number of boxes. Unless there are no previously tracked BBoxes, the number of objects is equal to the number of tracking BBoxes. At time T+c, u' = n' if and only if u = 0. When the previous tracking BBox was missing, the current tracking BBox was calculated based on the objects seen inside each class. This object tracking module was developed using the Kalman filter and Hungarian algorithm to predict the next locations of the detected objects, as well as the SORT algorithm[5], which uses the concept of Intersection Over Union (IOU) to track multiple instances of the same object with the same ID number.

Similar to how and C were retrieved from the newly given image at time T, this is done again at time T+c using the same object detection module.

At T+c, we will do the same for any item in which there is no object pair with an IOU higher than 0.3; we will consider it to have just joined the RoI. The newly produced entity will be given a special name that isn't being used by anything else. In this system, the faster RCNN learning algorithm[5] is used for object detection while a SORT[6] is used for ID assignment and tracking. SORT[6] allows for multi-object tracking at a degree speed of 100-300 frames per second. Since the system carries out object tracking utilizing the SORT[6] method based on the IoU value, the video frame interval c[7] had an effect on the object tracking capabilities. The time it takes for the object detection network to make a detection may be changed to lower the video's processing needs. This has been verified experimentally, with results demonstrating that objects may be tracked for as much as six frames[7].

To maximize object tracking performance, the video frame interval should be adjusted based on the number of connected cameras. approaching a deep learning-powered server.

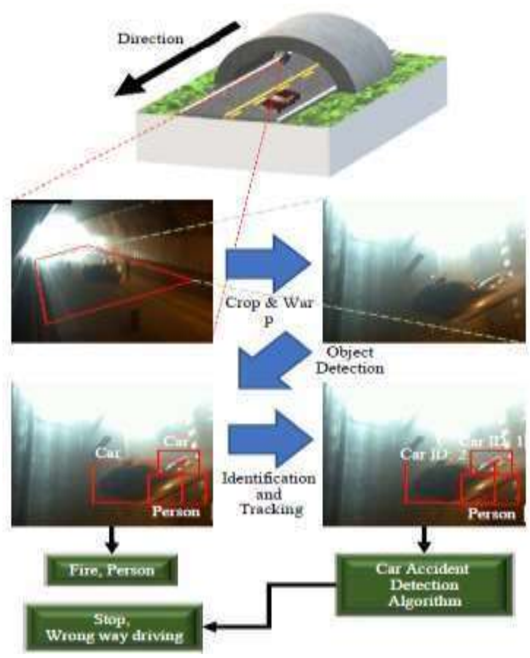
B. A Predictive System to Prevent Tunnel Disasters

Driving in a road tunnel is riskier than on a standard highway because of the lack of escape routes; drivers should be made aware of any potential emergencies in the tunnel. This being the case, the laws of South Korea mandate the recording of data pertaining to human beings, fires, stops, and WWD. Risk is increased while driving through a road tunnel since there is less space to exit the vehicle. Therefore, motorists need to be informed promptly of any tunnel emergencies[7]. South Korean law [4] mandates that all sensors in the country be able to detect and report on "Person, Fire, Stop, and WWD."

Meanwhile, in the tunnels, surveillance cameras keep an eye out for targets and unusual occurrences.

In order to do this, we will use a tunnel-optimized automatic object detection system for the targets. The equipment, however, is totally worthless as one enters a tunnel. The reason for this is because (1) the tunnel footage was poorly lit, making it susceptible to artifacts from the vehicle's taillight or warning light. (2) The video shot in the tunnels looked eerily ominous. That is to say, its color is distinct from the asphalt outside the tunnel. It was anticipated that the video monitoring system built for use on roads outside the tunnels would not work properly once inside. Therefore, there must be research into and implementation of an automatic accident detection system tailored to highway tunnels.

To solve these problems, researchers in [7] developed a deep learning–based Tunnel CCTV Accident Detection System. In order to train, we used the deep learning model faster R-CNN. Another model trained on image datasets that contained occurrences in tunnels served as the basis for this model. Then, the Car Accident Detection Algorithm (CADA) is regularly used by ODTS to use the tracking information of the target Car object to generate Stop and WWD events. This specialized object tracking function is only relevant to Car objects. Figure 2 shows how the CADA mechanism could detect an accident state. Crop and distort the original CCTV screen image so that it fits inside the ROI you just got from the tunnel's CCTV footage. Similar to [1], but with the goal of standardizing the identification of Stop and WWD occurrences, this approach was developed. An image extract prepares the image for training by cutting out unwanted background and making foreground and background items seem the same size. These are brand new concepts, unlike those in [1].Then, use a Faster RCNN that has been trained to identify vehicles, fires, and humans[5].After that, we constructed an extra 'No Fire' object by formally defining the object class to reduce the frequency of the incorrect answer for the Fire object. Tunnel lights, car taillights, and other possibly misleading objects get the No Fire label. With the exception of the background, Faster R-CNN training uses data characteristics to assign an object class. Misdetection of Fire may be decreased when using this method with untrained data.



To solve these problems, researchers in [7] developed a deep learning–based Tunnel CCTV Accident Detection System. In order to train, we used the deep learning model faster R-CNN. Another model trained on image datasets that contained occurrences in tunnels served as the basis for this model. Then, ODTS uses the Car Accident Detection Algorithm (CADA) to regularly use the tracking information of the target Car object to generate Stop and WWD

ISSN2454-9940
www.ijsem.org
Vol 11, Issuse. 1 Jan 2020

events. This specialized object tracking function is only relevant to Car objects. Figure 2 shows how the procedure used by CA DA may detect an accident state. Crop and distort the original CCTV screen image so that it fits inside the ROI you just got from the tunnel's CCTV footage. Similar to [1], but with the goal of standardizing the identification of Stop and WWD occurrences, this approach was developed. An image extract prepares the image for training by cutting out unwanted background and making foreground and background items seem the same size. These are brand new concepts, unlike those in [1].

Then, use a Faster RCNN that has been trained to identify vehicles, fires, and humans[5]. After that, we constructed an extra 'No Fire' object by formally defining the object class to reduce the frequency of the incorrect answer for the Fire object. Tunnel lights, car taillights, and other possibly misleading objects get the No Fire label. With the exception of the background, Faster R-CNN training uses data characteristics to assign an object class. Misdetection of Fire may be decreased when using this method with untrained data.

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

To address these issues, scientists at [7] designed a deep learning–based Tunnel CCTV Accident Detection System. Faster R-CNN, a deep learning model, was employed during training. To top it all off, the model was constructed on top of another model that was trained using picture datasets that included events that took place in tunnels. The ODTS then routinely utilizes the tracking information of the target Car object to create Stop and WWD events using the Car Accident Detection Algorithm (CADA). Only Car objects may benefit from this specialized tracking feature. Figure 2 depicts the possibility for CADA's technique to identify an accident condition. The original CCTV screen picture must be cropped and distorted to fit inside the ROI that was recently extracted from the tunnel's CCTV video. This method is quite similar to [1], however it was designed to standardize the detection of Stop and WWD events. After the picture has been "extracted," which entails cropping the image to expose just the ROI of interest and scaling local and distant objects to the same scale, it may be trained more effectively. These ideas are completely new, unlike [1].
Then, use a Faster RCNN that has been taught to recognize things like cars, fires, and people[5]. An additional 'No Fire' object was built by defining the object class directly to decrease the number of times the wrong response was given for the Fire object. The No Fire label is placed on potentially misleading indicators like tunnel lights, automobile taillights, etc. Faster R-CNN training employs data characteristics to assign a class to an object, with the exclusion of the backdrop.
Using this technique on untrained data has the potential to reduce false fire alarms.

III. EXPERIMENTS

This study splits its trials testing the developed system into two parts: one tests how well deep learning can pick up new information, while the other looks at how well the system can spot potential mishaps. whole framework. The performance of ODTS SORT is heavily dependent on how well objects are recognized. Therefore, accurate learning of a deep learning object detection network was crucial for the completion of this system. The system was put to the test after the deep learning model was trained to see whether it could identify the four distinct accident situations. Due to the need of 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 image to check whether it is able to recognize each condition. How Deep Learning Works, From the Ground Up Instead of being trained on a moving video, the deep learning network was shown how to identify things in a series of still photographs. An epoch in this study is a complete training cycle of the whole dataset. Some of the material to be analyzed consists of photographs taken at accident sites. R-CNN, a faster variant, was used for training[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

The current state of the training data set may be seen in Table.1.
The 70,914 images in this collection were taken from 45 different films. In deep learning's training approach, unlike the usual deep learning procedure, learning data and inference data were not separated. This is because, unlike publicly available datasets, the dataset used in this study consists entirely of still

images and does not cut away at any point throughout each film. That is to say, the exteriors of all the videos seem the same, however the contents change depending on whether elements are present or absent. Whether or not each image's training data and inference data were divided had no effect on the object detection network's inference performance.
However, it is difficult to assess the detection method of the entire tunnel CCTV image accident detection system, since the detection performance of the accident may suffer if the stability of the object identification on the whole video decreases. This means that all available data was utilized for training, and the results of that training are now being used to assess how well deep learning performs at object recognition. Fires are very uncommon in the tunnel, hence there are only a limited number of Fire goods.

It is crucial that the tunnel control room have a low false alarm rate given the possibility of both false and missing fire alarms. If a system is constantly triggered to report a false detection when none has occurred, it is very unlikely to be dependable once implemented. On the other hand, utilizing the enriched data that is routinely contributed to the time-lapse dataset, the detection performance might be increased automatically. Therefore, in this experiment meant to evaluate approaches for reducing false detection, the number of things labeled "No Fire" was much higher than the number of objects labeled "Fire." The training time for R-CNN was reduced by 10 epochs. The deep learning framework was Tensor-flow 1.3.0 running on Linux[7]. The hardware utilized in Faster R-CNN training is an NVIDIA GTX 1070. Each class's inference accuracy was measured using the mean (AP) after a total of 60 hours of training.

TABLE II. INFERENCE RESULT OF DATASET

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

In Table, we saw the AP values for the three objects under investigation. The training data set is dominated by automobiles, which also have the highest average posterior score. When compared to outcomes with other object classes, Car emerges on top.
That is to say, it was expected that the video Car's deep running object detection performance would be quite reliable. Due to its long, slim proportions and small overall footprint, the Person object has a low AP as shown in Table.2. Incorrect identification was possible given the small sample size of training data (857 items), even though the Fire object's accuracy percentage (AP) was quite excellent (0.9085).
False positives might be reduced by deep learning training using data that isn't really flammable. Before they could be utilized to correctly identify the Fire in the tunnel control center, however, further images of a Fire event needed to be acquired and included into training.
B. Accident Detection System Testing Using Full-Tunnel CCTV The results are determined by the deep learning model's training data. The efficiency of the Tunnel CCTV Accident Detection System, which uses deep learning, must be evaluated. Table 3 displays the four films selected for this function. A program was written so that the detection outcomes could be seen in motion. When the video frame interval was sped up to 6 frames per second at 30 fps, it was seen in under 10 seconds[7]. The length, occurrence time, and detection time are summarized briefly 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 shows the delay that occurs between when Stop and WWD events occur and when they are detected. Repeats once every 2.4 seconds, which makes sense given that this is a defining feature of CADA. There was a lag of 8 seconds between Stop and WWD, but the algorithm caught it. Images of the occurrence, such as those depicting a person or a fire, allowed for prompt identification. Although the visuals in Table 3 were used for practice purposes, this does not guarantee that they will appear the same in the final result. As a result, we started using the results of the tests and shooting more test films.

IV. CONCLUSION

ISSN2454-9940
www.ijsem.org
Vol 11, Issuse. 1 Jan 2020

This research proposes a new method of ODTS for real-time analysis of an item's mobility by fusing a deep learning-based object identification network with an object tracking algorithm. object of a certain kind may be recovered and put to good use. Since ODTS object tracking uses SORT, which only uses BBox information and not an image, the speed with which objects may be identified is crucial. Unless the object tracking technique strongly depends on object identification accuracy, it is unclear whether continuous object detection performance is required in this context.

On top of that, ODTS was used to develop a CCTV accident detection system specifically for tunnels. In order to determine whether or not a system-wide accident might be detected by means of a deep learning object recognition network, experiments were conducted. This method utilizes CADA to make distinctions at each cycle by using data about the vehicles' dynamics. We were able to cut the time it takes to notice an accident from 10 seconds to under a minute by tinkering with the image of each incident. However, deep learning training ensured the object identification performance of a reliable Car object, but the object detection performance of a Person was relatively bad.

The problem is that there aren't enough Fire objects in the amateur videos for accurate detection to be possible. However, false alarms may be avoided by adding more training on objects that are No Fire. If the Fire picture can be made more secure in the future, it could help the deep learning object detection network better identify fires. The ODTS may be used as an example of a Tunnel CCTV Accident Detection System, but it also has potential uses in areas where the dynamic movement of a single object has to be monitored, such as vehicle speed estimation and illegal parking. Multiple image encryption, together with Fire and Person object encryption, is crucial for system reliability. The deployment and frequent monitoring of the tunnel management website may further improve the system's reliability.

REFERENCES

"Bird's eye view localization of surrounding vehicles:Longitudinal and lateral distance estimation with partial appearance," by E. S. Lee, W. Choi, and D. Kum, will appear in Robotics and Autonomous Systems, volume 112, pages 178-189 (2019).

[2] "Robust vehicle detection by combining deep features with exemplar classification," Computerizing, vol. 215, pp. 225-231, 2016, L. Cao, Q. Jiang, M. Chang, C. Wang.

[3] "Detection and classification of vehicles for traffic video analytics," Proceed computer science, vol. 144, pp. 259-268, 2018, A. Arinaldi, J. A. Pradana, A. A. Singur.

[4] "Development of a deep-learning based automatic tunnel incident detection system on CCTV," by K. B. Lee, H. S. Shin, and D. G. Kim, published in the proceedings of the fourth international symposium on computational geo mechanics (2018), pages 140–141.

Neural Information Processing Systems, 2015, pp. 91-99; "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks," by S. Ren, K. He, R. Girshick, and J. Sun.

[6] "Simple Online and Real-time Tracking," by A. Beckley, Z. Pyongyang, L. Otto, F. Ramos, and B. Upcroft, in Proc. IEEE International Conference on Image Processing, 2016, pp. 3464-3468.

For example, see [7] K. B. Lee, H. S. Shin, D. G. Kim, "Development of a deep-learning based automatic tracking of moving vehicles and incident detection processes on tunnels," 2018;vol.20,no.6;pp.1161-1175 Korean Tunneling and Underground Space Association.