IJITCE

ISSN 2347- 3657

# International Journal of Information Technology & Computer Engineering


Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 10,Issue 3,July 2022

# Use of a deep learning system for autonomous accident identification in tunnels with inadequate CCTV surveillance

Mr. M. NAGA RAJU, Mrs. SARAGADAM ANURADHA, MR G BHOOLOKA RAJU

***Abstract—***

***This research uses the Object Detection and Tracking System (ODTS) in tandem with the Faster Regional Convolution Neural Network, a popular deep learning network. Automatic identification and monitoring of unanticipated occurrences, such as (1) Wrong-Way Driving (WWD), (2) Stopping, (3) People Getting Out of Vehicles in Tunnels, and (4) Fires, will be implemented via the use of a (Faster R-CNN) for Object Detection and a Conventional Object Tracking method. Bounding Box (BBox) findings from Object Detection are accepted as input by ODTS, and ID numbers are assigned to each moving and detected object based on a comparison of the Boombox from the current and prior video frames. Conventional object detection frameworks often fail to accomplish the ability to follow a moving item over time, but our system does just that. Average Precision (AP) values of 0.8479, 0.7161, and 0.9085 were achieved for target objects of cars, people, and fires, respectively, when a deep learning model in ODTS was trained using a datasets of event photos in tunnels. The ODTS-based Tunnel CCTV Accident Detection System was then evaluated with four movies including each accident, using the trained deep learning model.***

## I. INTRODUCTION

The use of object detection technologies has resulted in the accurate measurement and localization of specific items in still and moving media. Several apps have surfaced largely in self-driving of automobiles, CCTV monitoring and security \system, cancer diagnosis, etc. Object tracking is another area \sin image processing to be done via unique identification \sand monitoring the locations of detected objects across time. However, object recognition in a single, static picture is an essential first step in tracking moving targets. Therefore, it can be claimed that the outcomes of object
stacking should be strongly reliant on the performance of \sthe object detection involved. This object tracking technology has been successfully used to a variety of applications, including tracking a moving target, such as a person or car, monitoring an accident scene with a traffic camera, keeping tabs on criminal activity or protecting a vulnerable location, etc. This research analyses and controls traffic conditions using automated object identification as a case study in the subject of traffic management. What follows are brief descriptions of each section. The authors of [1] claim to have created a vehicle detection system for the autonomous automobile that may be used while out on the road. A vehicle object is recognised and classified by this system. convolutional neural network to identify the vehicle type (CNN). The vehicle object

lohitnaga2@gmail.com, anuradhasaragadams@gmail.com
g.b.raju33@gmail.com

Department of cse

Gonna Institute of Information Technology and Sciences (GIITS)

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 10,Issue 3,July 2022

tracking method moves the tracking centre point around the picture to follow the vehicle object as it is identified. The system then determines the distance between the driver's automobile and the visible vehicle objects, and the screen displays a localized picture akin to a bird's perspective with the visualized vehicle objects. The system's procedure makes it possible to see the vehicle's location from an outside perspective, which aids the self-driving system. Therefore, the camera can pinpoint the location of the moving vehicle to within 1.5 meters up or down and 0.4 meters left or right. To keep tabs on traffic on city streets and highways from above, the researchers at [2] created a deep learning-based detection system that uses CNN and Support Vector Machine (SVM). This system takes the satellite picture as an input value, runs it through a CNN to extract features, and then uses a support vector machine to conduct a binary classification in order to locate the vehicle's BBox. In addition, Arinaldi, Pradana, and Gurusinga [3] designed a system to assess traffic volume, categorise vehicle types, and predict travel times. This method makes use of BBox information gleaned from video and image-based item recognition. The system's approach was evaluated next to the Gaussian Mixture Model + SVM and the quicker RCNN. Then, it seems that R-CNN with a higher processing speed effectively identified the location and kind of vehicle.

## II. DEEP LEARNING-BASED OBJECT DETECTION AND TRACKING SYSTEM

### *A. Concept*

For an illustration of how the ODTS detects and follows objects over time, see Figure 1[7]. It is expected that ODTS has received enough training to recognize objects accurately. on a certain frame of a picture. The ODTS receives video frames at intervals of c, from which it learns the coordinates of n BBoxs. number of items in the scene at time T, as determined by a trained object detection system. Each identified object's matching category

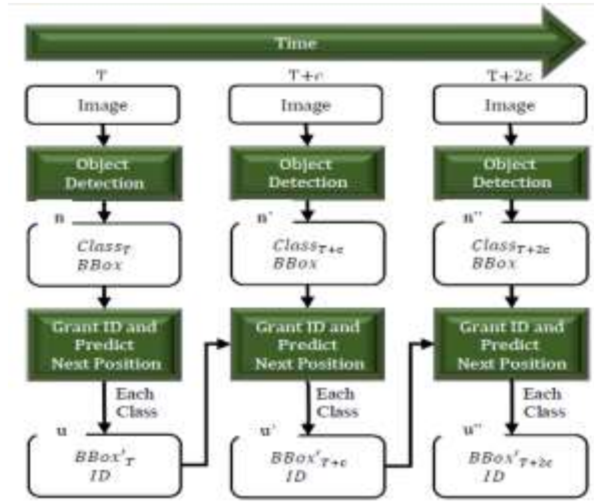



Figure 1: Timeline of the Object Detection- Tracking system's object detection and tracking procedure. The object tracking algorithm assigns an identifier and makes predictions after class and BBox have been collected from object detection. position in the future based on the present and the past BBox.

The object detecting system immediately labels object. Future, a dependent object tracking module is activated using the observed object data to uniquely identify each item,, and anticipate its next location, ′. Number of Boxes u is monitoring is not equal to n. The number of objects identified is equal to the number of tracking BBoxes unless the number of previously tracked BBoxes is zero. If u = 0 at time T+c, then u' = n' at that instant. In other words, the current tracking BBox is derived from the items observed inside each class while the previous tracking BBox was unavailable. With the help of the SORT algorithm[5], which employs the concept of Intersection Over Union (IOU) to track multiple instances of the same object with the same ID number and the Kalman filter and Hungarian algorithm to predict where the detected objects will go next, this object tracking module was crafted.

At the subsequent time step T+c, the same object detection module is used to retrieve and C from the freshly provided picture, just as was done before at time T.

Similarly, at T+c, we will treat as freshly entered the RoI any item in for which there is no object pair with an IOU value greater than 0.3. A unique identifier that does not already belong to anything else will be given to the newly formed object. This system uses a SORT[6] for ID assignment and object tracking in addition to a quicker RCNN learning algorithm[5] for object detection. Using a degree speed of 100-300 frames per second, multi-object tracking is possible with SORT[6]. The object tracking ability was influenced by the video frame interval c[7], since the system performs object tracking using the SORT[6] algorithm based on the IoU value. By varying the duration between detection made by the object detection network, the total amount of processing required may be reduced throughout the course of a video. Experiments on the object tracking capabilities across the frame interval confirmed this, showing that objects could be followed for as much as six frames[7].

Object tracking performance degrades dramatically with increasing frame interval; hence, the video frame interval has to be tuned for the total number of linked cameras. towards a server using deep learning.

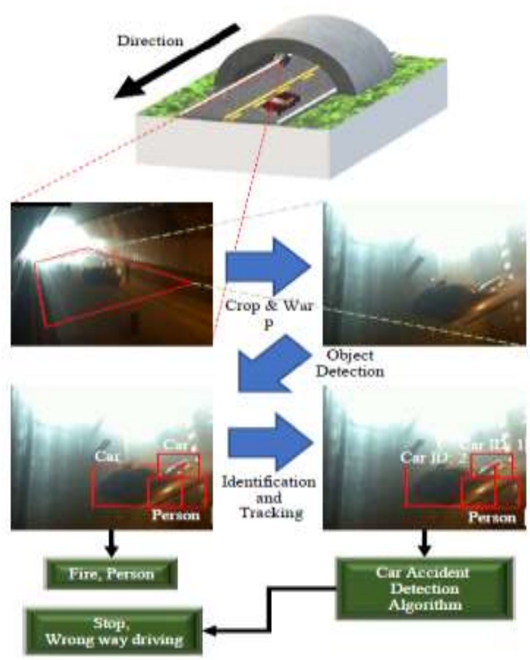
B. An Early Warning System for Tunnel Accidents

Inadequate room for evacuation makes driving in a road tunnel more hazardous than on a regular highway; vehicles should be alerted if an emergency develops in the tunnel[7]. Since this is the case, the national legislation in South Korea specifies that the following items and events are to be tracked: persons, fires, stops, and WWD. Driving through a road tunnel is risky since there is less room to get out of the car than on a standard roadway. Therefore, drivers need to be made aware of any tunnel emergencies as soon as possible[7]. According to Korean national rule [4], sensors in South Korea must be able to detect and track "Person, Fire, Stop, and WWD."

Meanwhile, CCTV in tunnels are used to keep an eye on targets and any unexpected happenings.

For this aim, we will use an automated object detection system for the targets, which has shown to function quite well in environments other than tunnels. However, the technology is completely useless inside of a tunnel. The reason for this is because (1) the tunnel video has poor illumination, so the video was heavily impacted by the tail light of the driving vehicle or the warning light of the automobile in operation. (2) There was a foreboding quality to the tunnel footage. That is to say, its hue contrasts with that of the tunnel's outside roadway. Due to the aforementioned factors, it was expected that the video surveillance system designed for use on roadways outside of the tunnels would not function well once within. As a result, there has to be an automated accident detection system developed specifically for use in road tunnels.

In [7], researchers created a deep learning-based Tunnel CCTV Accident Detection System to address the aforementioned issues. Faster R-CNN, a deep learning model, was utilised for training. Furthermore, the model's foundation was another model trained on picture datasets that included incidents in underground passageways. Then, ODTS employs a specialised object tracking function, applicable exclusively to Car objects, and periodically employs the Car Accident Detection Algorithm (CADA) to utilise the tracking information of the target Car object to create Stop and WWD events As can be seen in Figure 2, CADA's process may identify an accident condition. Crop and distort the original CCTV screen picture to match the extracted region of interest (RoI) from the tunnel's CCTV screen. While this method is similar to [1], it is designed to provide a consistent norm for recognizing Stop and WWD occurrences. By cropping the picture so that just the desired region of interest (ROI) is visible, and by scaling nearby and faraway objects to the same scale, the image extract makes the image more amenable to training. As opposed to [1], these are new ideas.Then, use a Faster RCNN that has been taught to recognize the presence of automobiles, fires, and people[5].Then, an additional 'No Fire' object was built by explicitly defining the object class to lessen the number of times the erroneous response for Fire object was returned. For potentially deceptive items, such as tunnel lights, automobile taillights, etc., the No Fire object is designated. In Faster R-CNN training, the characteristic of the data is mirrored in the designation of the object class, with the exception of the backdrop. This strategy has the potential to reduce Fire misdetection with untrained data.




In [7], researchers created a deep learning-based Tunnel CCTV Accident Detection System to address the aforementioned issues. Faster R-CNN, a deep learning model, was utilised for training.

Furthermore, the model's foundation was another model trained on picture datasets that included incidents in underground passageways. Then, ODTS employs a specialised object tracking function, applicable exclusively to Car objects, and periodically employs the Car Accident Detection Algorithm (CADA) to utilise the tracking information of the target Car object to create Stop and WWD events. As can be seen in Figure 2, CA DA's process may identify an accident condition. Crop and distort the original CCTV screen picture to match the extracted region of interest (RoI) from the tunnel's CCTV screen. While this method is similar to [1], it is designed to provide a consistent norm for recognizing Stop and WWD occurrences. By cropping the picture so that just the desired region of interest (ROI) is visible, and by scaling nearby and faraway objects to the same scale, the image extract makes the image more amenable to training. As opposed to [1], these are new ideas.

Then, use a Faster RCNN that has been taught to recognize the presence of automobiles, fires, and people[5]. Then, an additional 'No Fire' object was built by explicitly defining the object class to lessen the number of times the erroneous response for Fire object was returned. For potentially deceptive items, such as tunnel lights, automobile taillights, etc., the No Fire object is designated. In Faster R-CNN training, the characteristic of the data is mirrored in the designation of the object class, with the exception of the backdrop. This strategy has the potential to reduce Fire misdetection with untrained data.

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

To deal with these problems, researchers at [7] developed a Tunnel CCTV Accident Detection System based on deep learning. The deep learning model Faster R-CNN was used for. training. As an added bonus, the model was built upon another model that was trained using image datasets that contained occurrences that occurred in tunnels. Then, the ODTS uses the Car Accident Detection Algorithm (CADA) to regularly use the tracking information of the target Car object to generate Stop and WWD events. This specialised object tracking function is only relevant to Car objects. The potential for CADA's procedure to detect an accident state is shown in Figure 2. Crop and distort the original CCTV screen image such that it fits inside the ROI that was just recovered from the tunnel's CCTV footage. Similar to [1], but with the goal of standardising the recognition of Stop and WWD occurrences, this approach was developed. The image is more easily trained after being "extracted," which involves reducing the image to reveal just the targeted area of interest (ROI) and scaling local and distant objects to the same scale. These are original concepts, in contrast to [1].

Next, use a Faster RCNN that has been trained to identify vehicles, fires, and humans[5]. To reduce the frequency of the incorrect answer for the Fire object, an extra 'No Fire' object was constructed by explicitly declaring the object class. Indicators that may mislead drivers, such as tunnel lights, car taillights, etc., are given 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.

# III. EXPERIMENTS

The experiments conducted in this research using the built system are separated into two sections: the assessment of deep learning's learning performance and the evaluation of the system's performance in detecting accidents. whole structure. Object recognition accuracy has a significant impact on ODTS SORT. Therefore, in order to finish this system, high performance of object detection was necessary via the correct learning of deep learning object detection network After the deep learning model was trained, the whole system was put to the test to see whether it could recognize the four specific accident scenarios that were being looked for. In this scenario, testing the system for each picture to see whether it is able to identify each circumstance was necessary due to the need of both the object recognition performance of the deep learning model and the discriminate capabilities of the CADA. The Fundamentals of Deep Learning The deep learning network was taught to recognize objects in a sequence of still photos rather than a moving movie. In this research, one epoch represents one cycle of the training procedure for the whole data set. Images from accident scenes make up a portion of the data set to be studied. Training was done with the faster 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 |

The current state of the training data set may be seen in Table.1.

ISSN 2347–3657

Volume 10,Issue 3,July 2022

The frames from 45 videos were used to create this collection of 70,914 video stills. In contrast to the typical deep learning procedure, In deep learning's training procedure, learning data and inference data were not partitioned. This is due to the fact that, in contrast to publicly accessible datasets, the data set utilised in this work has continuous pictures throughout each movie. That is to say, all of the videos have the same visual backdrop, and the images inside them vary based on the presence or absence of various things. Inference performance of the object detection network would be same whether the training data and inference data were split for each picture.

However, it is challenging to evaluate the detection technique of the full tunnel CCTV picture accident detection system because the stability of the object identification on the whole video may decline, which negatively impacts the detection performance of the accident. Therefore, all accessible data was used for training, and the learnt data is being used for evaluating the performance of deep learning in object identification. Due to the rarity of fires in the tunnel, only a small amount of Fire items exist.

Since both false and missed fire alarms are possible, it is very important for the tunnel control room to have a low rate of false alarms. It is very unlikely that an installed system will be reliable if it is repeatedly alerted that false detection has happened when none has occurred. In contrast, the detection performance might be improved automatically using the time-lapse dataset's enriched data, which is regularly added to the training data set. As a result, the number of items labeled "No Fire" was much greater than the number of objects labeled "Fire" during this experiment designed to test methods for decreasing false detection. R-CNN training was sped up by the addition of 10 epochs. Tensor-flow 1.3.0 on Linux was used as the deep learning framework[7]. NVIDIA GTX 1070 is used for the hardware in Faster R-CNN training. The total training duration was 60 hours, and the accuracy of each class's inference evaluation was determined using the mean (AP).

***TABLE II. INFERENCE RESULT OF DATASET***

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

Values of AP for the three items to be detected were shown in Table. Cars make up the most numerous object type and have the highest AP value in the training data set. compared to other object types, Car's results are superior.

In other words, the video's Car's deep running object identification performance was anticipated to be quite dependable. Table.2 shows that the AP for the Person object is relatively low since the Person object has a long, petite form and is relatively modest in size. Although the accuracy proportion (AP) of the Fire object was rather good (0.9085), incorrect detection was likely since there were so few items available for training (857).

However, training on deep learning using things that aren't on fire might help lower the number of false positives. But additional photographs of a Fire occurrence needed to be collected and incorporated into training before they could be used to identify the Fire in the tunnel control centre.

B. Full-Tunnel CCTV Accident Detection System Test for Accident Detection The performance is based on the trained deep learning model. An assessment of the effectiveness of the deep learning-based Tunnel CCTV Accident Detection System is required. As shown in Table 3, four films were chosen for this purpose. For the purpose of seeing the detection results on the video, a software was developed. Within 10 seconds of visual observation, it was found when the video frame interval was adjusted to 6 frames per second at 30 fps[7]. Table 3 provides a concise summary of the video's duration, the occurrence time, and the detected time.

***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

International Journal of
Information Technology & Computer Engineering
ISSN 2347–3657
Volume 10,Issue 3,July 2022

may look different if put in the field. Hence, the tested was put to use, and further test movies were shot.

## IV. CONCLUSION

Combining a deep learning-based object detection network with an object tracking algorithm, the authors of this study suggest a novel technique of ODTS and demonstrate its application to the real-time analysis of an item's motion. item of a certain kind may be retrieved and used. Object identification speed is critical in ODTS object tracking due to the fact that SORT relies only on BBox information rather than a picture. Since this is the case, it is questionable if continuous object detection performance is necessary, unless the object tracking method relies heavily on object recognition accuracy.
Further, a CCTV accident detection system for tunnels was created using ODTS. Experiments were performed to test the effectiveness of a deep learning object identification network and to identify the presence of a system-wide accident. Using the dynamic information of the vehicles, this system incorporates CADA to make distinctions at each cycle. By playing around with the visual of each accident, we were able to reduce the time it took to spot an accident from 10 seconds down to under a minute. However, the object recognition performance of a trustworthy Car object was secured by deep learning training, but the object detection performance of a Person was rather poor.
However, since there aren't enough Fire objects in the unskilled films, there's a large chance of erroneous detection. However, by further training items that are No Fire, erroneous detection may be reduced. Securing the Fire image in the future could boost the deep learning object detection network's performance while detecting fire objects. In addition to its utility as an example of a Tunnel CCTV Accident Detection System, the ODTS has potential applications in domains that require monitoring of the dynamic movement of a particular item, such as vehicle speed estimate or unlawful parking. Securing multiple pictures, as well as Fire and Person objects, is important to improve the system's dependability. Additionally, system dependability might be enhanced by implementation of and constant monitoring of the tunnel administration website.

## REFERENCES

*[1] E. S. Lee, W. Choi, D. Kum, "Bird's eye view localization of surrounding vehicles :Longitudinal and lateral distance estimation with partial appearance," Robotics and Autonomous Systems, 2019, vol. 112, pp. 178-189.*
*[2] L. Cao, Q. Jiang, M. Chang, C. Wang, "Robust vehicle detection by combining deep features with exemplar classification," Computerizing, 2016, vol. 215, pp. 225-231.*
*[3] A. Arinaldi, J. A. Pradana, A. A. Singur, "Detection and classification of vehicles for traffic video analytics," Proceed computer science, 2018, vol. 144, pp. 259-268.*
*[4] K. B. Lee, H. S. Shin, D. G. Kim, "Development of a deep-learning based automatic tunnel incident detection system on CCTV," in Proc. Fourth International Symposium on Computational Geo mechanics, 2018, pp. 140-141.*
*[5] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks," in Proc. Neural Information Processing Systems, 2015, pp. 91-99.*
*[6] A. Beckley, Z. Pyongyang, L. Otto, F. Ramos, B. Upcroft, "Simple Online and Real-time Tracking," in Proc. IEEE International Conference on Image Processing, 2016, pp. 3464-3468.*
*[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," Korean Tunneling and Underground Space Association, 2018, vol. 20,*