

Identifying Vehicles Dynamically on Freeway CCTV Images through the YOLO Deep Learning Model

Shin-Hung Pan¹ and Shu-Ching Wang^{2*}

¹Asia University, 500, Lioufeng Rd., Wufeng, Taichung, Taiwan ROC

²Chaoyang University of Technology, 168 Jifong E. Rd., Wufeng, Taichung, Taiwan ROC

(Received December 26, 2020; accepted March 22, 2021)

Keywords: YOLO, deep learning, object detection, freeway, CCTV

This study focuses on object detection in computer vision research. The object detection process often encounters many uncertainties, such as the uncertainty of the number of objects in the image, the different conditions of the objects including their appearance, the current driving speed, the obstruction between vehicles, sunlight in the daytime, the lack of light at night, the irreversible factors related to the CCTV lens, and other factors, which make object detection and image preprocessing difficult. Taiwan's freeways are all equipped with CCTV to monitor real-time road conditions, and all CCTV images are available to the public via the internet. However, in freeway segments and tunnels, and even on traffic-prone roads, traffic jams and accidents are only judged by “human power.” Therefore, in this study, we use existing CCTV streaming video as a vehicle sensor data source and the You Only Look Once (YOLO) algorithm to perform object detection as well to tune adjustable parameters to achieve the desired results. From the preliminary results of this study, the current model based on the YOLOv3 algorithm and the Common Objects in Context (COCO) image dataset has an accuracy of 44% during the daytime and 41% during the nighttime for CCTV cameras installed outdoors. In the future, we will analyze larger amounts of CCTV video streaming data to detect whether a road is congested and even detect the occurrence of traffic accidents.

1. Introduction

Nowadays, the traditional way to control traffic flow is to use ramp meters to control the flow of traffic on national freeways. One of the important questions to consider in advance is which ramp, section, or CCTV is needed to effectively relieve traffic congestion and predict traffic jams and even report traffic accidents.

In the past, we had to rely on empirical rules of thumb, such as the traffic flow in Taoyuan is often the main cause of frequent traffic congestion in Yangmei, so if traffic congestion occurs on the Yangmei section of the road, the Freeway Bureau will limit the traffic flow on the south off-ramp of the Yangmei interchange in Taoyuan to alleviate traffic congestion. However, the actual traffic flow that causes traffic congestion may come from Linkou and Wugu, and incorrect

*Corresponding author: e-mail: scwang@gm.cyut.edu.tw
<https://doi.org/10.18494/SAM.2021.3236>

control will not only fail to relieve traffic congestion, but also increase the likelihood of traffic congestion near the ramp.

Knowing where a vehicle comes from is not an easy task. To understand the history of a vehicle's movement on a freeway, it is necessary to analyze the origin and destination table of each vehicle's trip through the start and end interchanges. However, an origin and destination table is not easy to create. Not only does it require a lot of labor, but it also has a high error rate. The complexity of a traditionally produced trip departure table is sufficient for approximately 30 people to be required to keep track of a single interchange. As there are different interchanges, one person is responsible for reporting the license plates of passing vehicles and another person is responsible for transcribing them, in three shifts a day. For the many interchanges on the freeway, it takes more than 1000 people to do a meter survey for each trip. However, even with so many people employed, the accuracy rate is usually only 30%, which is too low to be usable.

To help the Freeway Bureau reduce the incidence of traffic jams back to gridlock incidents with limited resources, three important points must be considered in the conventional thinking:

1. Criteria for determining the definition of traffic congestion

There is no standard definition of a traffic jam, and experts usually rely on their experience in the past to determine when a traffic jam is occurring. When should the traffic flow be restricted? In some road sections, only traffic lights are installed at interchange ramps for traffic flow control, which is the main cause of traffic congestion during off-peak hours.

2. CCTV images

The quality of CCTV images provided by the Freeway Bureau varies, and views of many road sections are obstructed by tree branches, cobwebs, or dirty lenses, making it more difficult to judge the quality of the images.

3. Data collection for trip departure tables

Using manual monitoring according to conventional thinking will not only greatly increase the difficulty of judgment, mental fatigue, and human costs, but also indirectly affect the accuracy of the trip departure table, which will reduce the overall accuracy to only 30%, despite the large number of people employed and the high cost.

To address the above issues, the methodological approach proposed in this study has the following aims:

1. Increase the accuracy of data collection by reducing the use of labor and the decrease in reliability due to manual surveillance.
2. Use CCTV streaming video as a vehicle sensor data source and an object detection algorithm to identify vehicles.
3. Use image preprocessing to deepen the color of object borders and increase color contrast to improve CCTV image quality and increase recognition.

In recent years, the use of deep learning methods to detect objects and humans in figures has become a hot research topic in machine learning. The task of object detection is to locate objects of interest in images or videos, and at the same time to detect their location and size. As shown in Fig. 1, there are three main application scenarios in a study that focuses on computer vision research: image classification, object localization, and object detection.⁽¹⁾

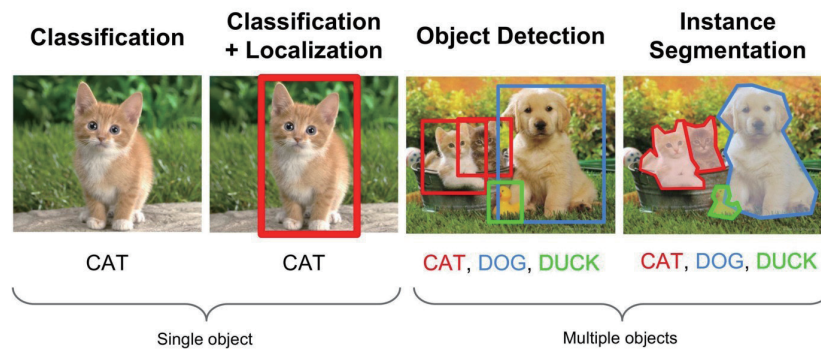


Fig. 1. (Color online) Applications of computer vision research.⁽¹⁾

A classification problem is a problem that classifies and predicts objects in the input image. If the input image contains more than one object, then the categorization question can be turned into asking if an object is in the input image.

In addition to classifying the main objects in the input image, it is necessary to locate their positions and the absolute positions of the output objects in the image. The problem with object positioning is that, compared with image classification, it requires not only an image category label, but also the absolute position of the object in the image to be positioned for supervised learning. For the object positioning problem, the input image may contain more than one object, and classifying and positioning multiple objects at the same time falls under the scope of object detection.

Object detection is used in a wide range of applications, including image annotation, product yield identification (coffee beans, tea leaves, etc.), computer vision question and answer, lip-reading, etc. These applications rely on object detection technology to detect single or multiple objects in the input image. There are many uncertainties in the object detection process, such as the number of objects in the image; objects having different appearances, shapes, and posture; and interference from lighting and blocking factors during object imaging, which makes detection difficult. Since the advent of deep learning, the development of object detection has mainly focused on two directions: two-stage algorithms such as the R-CNN series and one-stage algorithms such as You Only Look Once (YOLO),⁽²⁾ SSD, and so forth. The main difference between the two directions is that the two-stage algorithms need to form a preselected box that may contain the object to be inspected, and then carry out fine-grained object inspection, while the one-stage algorithms extract features directly from the network to predict the object classification and location.

In this study, we use CCTV images provided by the Freeway Bureau, Ministry of Transportation and Communications (MOTC), Taiwan, as the vehicle sensor data source, and the YOLO algorithm with a large amount of data training with the aim of identifying more vehicles in CCTV images more accurately both during the daytime and nighttime and under different weather conditions. We then collect data to estimate whether the road is starting to become congested to provide the public with more real-time and accurate traffic information.

2. Related Works

Traffic flow control and monitoring is an essential part of life, and even during festivals and holidays, we must rely on the system to maintain the control and operation of the freeway traffic flow. The traditional traffic flow control and monitoring system still has some deficiencies and over-reliance on “manual” monitoring, and through the personnel who monitor for a long time, it is inevitable that there will be some mental shortcomings. In this section, the literature of research on freeway traffic recognition and the application of the YOLO algorithm to image object detection will be discussed.

2.1 Research on freeway traffic recognition

Nurhadiyatna *et al.* provided video data and actual vehicle speeds in their research.⁽²⁾ In their experiment, three types of driving participants with different types of vehicles were used. The speedometer of each vehicle also verified the actual speed, and a mobile phone with a GPS function was placed near the speedometer in order to verify the speed calculated by the application using the GPS data received from the mobile phone; the GPS information could only be used to calculate the speed of the driver or the vehicle when the actual driver or the vehicle passed the camera. This allowed them to compare the actual hourly speed with the estimated hourly speed.

Historically, most of the accidents on highways are caused by human negligence, and the surveillance images and information are found only after the problems occur. To overcome these problems, Desai *et al.*⁽³⁾ developed an existing intelligent CCTV system to create a video image capture system that calculates and generates alarms in a timely manner, which means that no other sensors are needed for assistance and, accordingly, it can detect incidents in real time and send alerts to medical, fire, and police authorities simultaneously to inform them of the situation and request support to ensure that adequate life-saving resources are available. They also proposed a way of detecting drivers who violate the rules of the highway in an area, such as speed limit and traffic sign violations, and reporting them to the relevant authorities for tracking and punishment. In addition, they also classified vehicles by size and appearance to further improve road construction and traffic flow. They utilized background subtraction, shape shifting, and many other conceptual techniques to maximize the application and integration into an intelligent system. The authors' philosophy in their research was that the cost of road monitoring can be reduced through complete automation of the road system.

Christopher and Shanna⁽⁴⁾ suggested that the accuracy of k-means clustering depends on the object characteristics and the number of samples available in each category, and the longitude depends on the problem under consideration. The support vector machine has been used for many different problems and can be used for video surveillance monitoring and analysis.

Regardless of the algorithm used, the analysis of CCTV is used for object tracking. In order to perform dynamic behavior tracking, it is necessary to identify the object in the video, and once the detected object is identified, behavior recognition can be performed to identify the object.

Background features also play an important role in behavior analysis. Deleting background features is an effective way to detect objects and eliminate them by using different background models, but the accuracy of deleting background eigenfeatures is poor. The method uses multiple deformation features to train the behavior analysis model. Since artificial neural networks (ANNs) have many precedents in object classification, ANNs have a significant impact on a variety of scientific problems and can also be applied to behavior analysis in video surveillance.

Hardjono *et al.*⁽⁵⁾ pointed out that highway traffic management and planning are not common in developing countries because they need to bring together multiple intelligent transport systems to collect vehicle-related data, and although some methods do exist, in India, they can only be obtained from existing CCTV, radio road news, and police Twitter feeds. They also pointed out that if deep learning and YOLO are suitably combined, good and accurate performance can be achieved.

Loce *et al.*⁽⁶⁾ suggested that the technology applied to automatic vehicle classification is rapidly becoming widespread owing to relatively reasonably priced sensors. For example, CCTV monitors, optical distance detection (LiDAR), and thermal imaging devices can assist us in vehicle detection, tracking, and classification of a single or multiple vehicles. For fixed CCTV, background segmentation can be implemented, which can handle a very limited number and variety of vehicles due to the lack of correct proportions and specifications for each type of vehicle. In many cases, hierarchical classification can still be used, and the classification process can be roughly divided into several categories. Previously, a coarse filter based on contour measurements was used to filter out the objects, which were then further sorted into more similar subcategories. The accuracy of vehicle classification can be improved by combining multiple sensors such as RGB cameras, light detection distance meters, and thermal imaging devices.

On the other hand, Ghosh *et al.*⁽⁷⁾ pointed out that traffic accidents are one of the important causes of death in India. Indeed, more than 80% of the time, the death is not caused by the accident itself but by the lack of timely help for the accident victim. On very busy, chaotic, high-speed highways, the main cause of accidents can be prolonged periods of unattended traffic. Therefore, the main purpose of their work⁽⁷⁾ was to create a system that will detect accidents based on real-time video footage from CCTV installed on the highway. The idea is to obtain each frame of the video and execute the video through a deep learning convolutional neural network (CNN) model that has been trained to classify frames of video images as accidental or non-accidental.

2.2 Research on application of YOLO algorithm

Computer vision and machine learning for object detection often have slow response times. Algorithms and ANNs such as YOLO not only solve this problem but also have no loss of accuracy. Čorović *et al.*⁽⁸⁾ pointed out that object detection is one of the important functions of software that will provide the next generation of automobiles with automatic driving. In their research, they trained a neural network on five object classes (cars, trucks, pedestrians, traffic signs, and lights), and the method was shown to be effective in a variety of driving conditions (sunny, cloudy, snow, foggy, and night).

However, if there are too many small tagged objects in the dataset, the accuracy and recall rate will not meet the expected target. When there are more than a certain number of tagged objects in the dataset, this will affect the expected results. Under busy traffic conditions, this may significantly reduce the chance of objects being detected, as a single component in the YOLOv3 detection frame is responsible for detecting more than three objects. In addition, if the object is obscured, the accuracy of object detection will be reduced. In both cases, if the camera is close to the target object (car), it can successfully detect and classify the object (Fig. 2). Under different weather conditions, objects (cars, traffic signs, etc.) can also be identified and classified successfully (Fig. 3).

Zhihao and Ying⁽⁹⁾ proposed a connectivity model that employs a YOLO-Tiny algorithm with a small number of convolutional layers and has a low feature utilization rate and low precision. However, since YOLO-Tiny is not good at detecting small objects, a YOLO-Tiny-based dense layer method was proposed by Zhihao and Ying⁽⁹⁾ to improve the performance by inserting a dense layer into the YOLO-Tiny neural network (Fig. 4). The improved neural network is tested on the Pascal VOC dataset, and the results show that the accuracy of the neural network is improved by 15% compared with the original algorithm. When compared with the original YOLO-Tiny model, the improved model is only 9.8 MB larger, and the size of the improved model is only about one-fifth of that of the original model.

Chen and Yeo⁽¹⁰⁾ proposed a framework with real-time image capture, vehicle detection, dynamic tracking, and alarm triggering mechanisms. In their study, YOLOv3 was used for vehicle detection and template pairing was used for motion tracking (Fig. 5). They also proposed a fault-tolerant strategy to adapt the frame to different camera angles, climate factors, lighting, and other conditions to make it more robust.



Fig. 2. (Color online) Object recognition and classification based on proximity detection.⁽⁸⁾

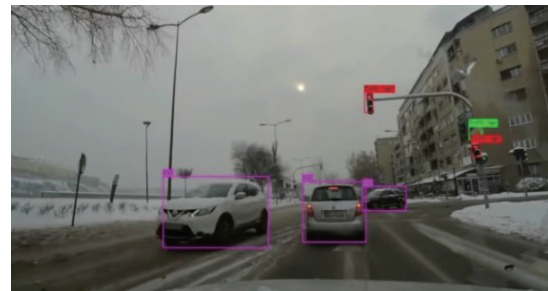


Fig. 3. (Color online) Identification of vehicles and traffic signs under different weather conditions.⁽⁸⁾

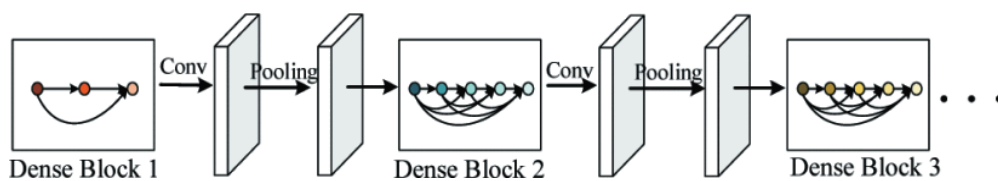


Fig. 4. (Color online) Dense connection layer structure.⁽⁹⁾

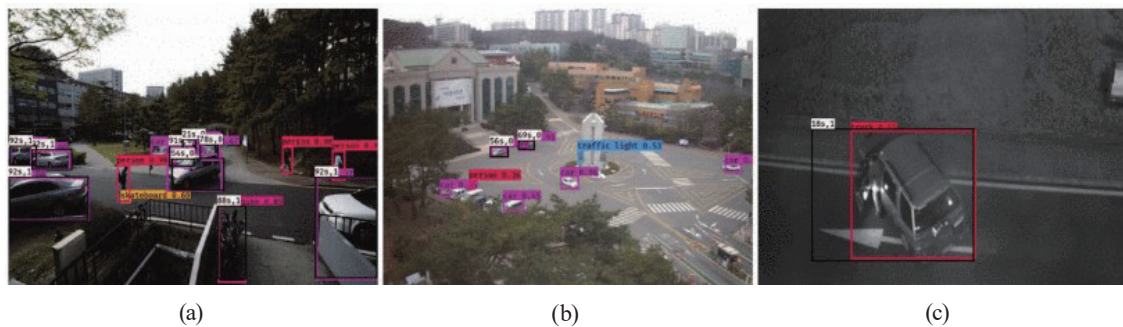


Fig. 5. (Color online) Identification results of Chen and Yeo.⁽¹⁰⁾ (a) Multiple vehicles. (b) Far distance. (c) Dark scene.

3. Materials and Methods

The data analyzed in this study is from the CCTV open dataset provided by the Taiwan Freeway Bureau, MOTC,^(11,12) and by combining the CCTV video data with the YOLOv3 model object detection technology, we can perform vehicle detection, traffic flow calculation, and even predict traffic jams in advance.

3.1 CCTV datasets from Freeway Bureau, MOTC

According to the 2020 CCTV open data collection provided by the Freeway Bureau of MOTC, there are 1554 CCTV camera locations and network connection locations on national freeways and expressways in Taiwan as of April 2020. The roads covered and the number of CCTVs on each road are shown in Table 1.

The CCTV open data was originally provided in XML format, which was divided into static data and dynamic data by the Freeway Bureau. In this study, we converted the data into Excel format, as shown in Figs. 6 and 7 and described below:

1. The CCTV static data (Fig. 6) has 11 fields: CCTV ID, road section, road name, starting position (km), ending position (km), longitude, latitude, data version, data name, update time, and update interval.
2. The CCTV dynamic data (Fig. 7) has seven fields: CCTV ID, camera network URL, status, version, data name, update time, and update interval.

3.2 Common Objects in Context (COCO) image dataset

The full version of the YOLOv3 neural model was used in this study, which was paired with the COCO image dataset⁽¹³⁾ for training the deep learning model. The COCO dataset was developed by Microsoft and contains a large number of images. The version of the dataset released in 2017 is currently mainly used. Its features include the following:

1. provides information on physical cutting of objects,
2. can be used for context recognition,
3. provides Superpixel object segmentation information,

Table 1
Roads covered by CCTV open data and numbers of CCTVs.

Type	CCTV Number	Type	CCTV Number
Freeway No.1	349	Expressway No.72	20
Freeway No.2	25	Expressway No.74	29
Freeway No.3	574	Expressway No.76	25
Freeway No.4	12	Expressway No.78	15
Freeway No.5	99	Expressway No.82	22
Freeway No.6	75	Expressway No.84	28
Freeway No.8	18	Expressway No.86	19
Freeway No.10	40	Expressway No.88	10
Expressway No.66	15	Others	165
Expressway No.68	14		

cctvid	roadsection	roadname	startlocationpoint	endlocationpoint	px	py	version	listname	updatetime	interval
nfbCCTV-N1-S-0-M	Freeway No. 1	Keelung Top - Keelung Interchange	167	168	121.735560	25.122778	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-0.1-M	Freeway No. 1	Keelung Top - Keelung Interchange	167	168	121.734535	25.121284	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-0.3-M	Freeway No. 1	Keelung Top - Keelung Interchange	167	168	121.733088	25.119885	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-0.7-M	Freeway No. 1	Keelung Top - Keelung Interchange	167	168	121.735556	25.117608	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-1-M	Freeway No. 1	Keelung Top - Keelung Interchange	167	168	121.728615	25.115833	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-3.9-M	Freeway No. 1	Badu Interchange - Dahua System	169	11585	121.708829	25.100252	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-4.57-M	Freeway No. 1	Badu Interchange - Dahua System	169	11585	121.703611	25.097500	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-6.85-M	Freeway No. 1	Wudu Interchange - Xizhi Interchange	170	172	121.686940	25.086945	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-8-M	Freeway No. 1	Wudu Interchange - Xizhi Interchange	170	172	121.675550	25.085556	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-8.8-M	Freeway No. 1	Wudu Interchange - Xizhi Interchange	170	172	121.668653	25.081948	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-11.55-M	Freeway No. 1	Xizhi Interchange - Xizhi System	173	174	121.643890	25.072779	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-22.9-M	Freeway No. 1	Neihu Interchange - Yuanshan Interchange	176	177	121.534376	25.072825	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-23.29-M	Freeway No. 1	Yuanshan Interchange - Taipei Interchange	177	178	121.532000	25.072971	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-25.06-M	Freeway No. 1	Yuanshan Interchange - Taipei Interchange	177	178	121.514509	25.078191	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-26.9-M	Freeway No. 1	Taipei Interchange - Sanchong Interchange	178	179	121.496282	25.076417	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-32.7-M	Freeway No. 1	Wugu Interchange - Wugu System	0	180	121.440780	25.069700	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-32.98-M	Freeway No. 1	Wugu Interchange - Wugu System	0	180	121.436830	25.070390	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-33.8-M	Freeway No. 1	Wugu Interchange - Freeway Bureau	180	0	121.429980	25.068610	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-34.02-M	Freeway No. 1	Wugu Interchange - Freeway Bureau	180	0	121.427710	25.068138	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-34.4-M	Freeway No. 1	Freeway Bureau - Taishan Interchange	0	0	121.424830	25.066692	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-35.16-M	Freeway No. 1	Freeway Bureau - Taishan Interchange	0	0	121.417910	25.063253	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-36.12-M	Freeway No. 1	Taishan Interchange - Linkou Interchange	0	182	121.410473	25.058725	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-38.25-M	Freeway No. 1	Taishan Interchange - Linkou Interchange	0	182	121.393300	25.062868	1.1	CCTV Static Info	2020/4/13 13:04	86400
nfbCCTV-N1-S-40.83-M	Freeway No. 1	Taishan Interchange - Linkou Interchange	0	182	121.368850	25.065650	1.1	CCTV Static Info	2020/4/13 13:04	86400

Fig. 6. (Color online) CCTV static data sheet of freeway.

cctvid	url	status	version	listname	updatetime	interval
nfbCCTV-N1-S-0-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=160,A40,0	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-0.1-M	https://cctvn11.freeway.gov.tw/vStream.php?pm=210,A57,1	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-0.3-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=162,A42,24	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-0.7-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=168,A32,16	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-1-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A44,4	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-3.9-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=168,A32,18	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-4.57-M	https://cctvn11.freeway.gov.tw/vStream.php?pm=210,A57,26	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-6.85-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A44,5	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-8-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A40,2	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-8.8-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=168,A32,19	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-11.55-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A44,6	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-22.9-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=169,A33,29	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-23.29-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=161,A41,18	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-25.06-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=151,A36,20	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-26.9-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=161,A41,19	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-32.7-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=159,A47,3	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-32.98-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A44,30	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-33.8-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=159,A47,4	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-34.02-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=160,A40,9	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-34.4-M	https://cctvn11.freeway.gov.tw/vStream.php?pm=210,A57,72	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-35.16-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=159,A47,5	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-36.12-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=151,A36,9	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-38.25-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=164,A44,14	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300
nfbCCTV-N1-S-40.83-M	https://cctvn01.freeway.gov.tw/vStream.php?pm=161,A41,23	0	1.1	CCTV Dynamic Info	2020/4/13 13:55	300

Fig. 7. (Color online) CCTV dynamic data sheet of freeway.

4. contains 330000 images, of which over 200000 are labeled,
5. includes 1.5 million objects,
6. has 80 object categories,
7. contains 91 stuff categories,
8. text descriptions are provided for each image, and
9. provides 250000 key points of portrait data.

Since the main detection target of this study is vehicles, the data classes selected for training are the three image data types numbered 3: Car, 6: Bus, and 8: Truck.

The YOLO model uses the Darknet-53 network architecture as the basis of the classical neural network. To facilitate test set training, we also convert the trained weights into Keras weights suitable for use with TensorFlow to conduct actual recognition tests in a Python environment.

Figure 8 shows a screen captured during the actual recognition testing. It shows the recognition results in a freeway CCTV image, including the identified objects, the category name (e.g., car), and the confidence level (e.g., 0.9104), and draws a box around objects.

4. Results

Table 2 shows the specifications of the experimental equipment used in this research. The average recognition speed is about 13–15 frames per second (FPS) when using the YOLOv3 algorithm to recognize freeway CCTV streaming images.

Figure 9 shows CCTV images of four freeway sections taken during the daytime, on which the identification results are superimposed. These four road sections were chosen because, at the time of writing, they were some of the sections with the most serious congestion during the daytime (average speed of 20–39 km/h) according to the Freeway Bureau, as shown in Fig. 10.

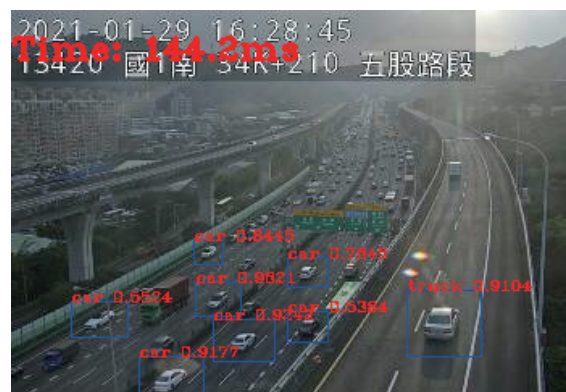


Fig. 8. (Color online) Recognition results in freeway CCTV image.

Table 2
Specifications of experimental equipment.

Central processing unit	Intel® Core™ i7-8750H@2.20 GHz
Graphics card	Nvidia Geforce GTX 1070
Memory	16.0 GB
Operating system	Windows 10 1909

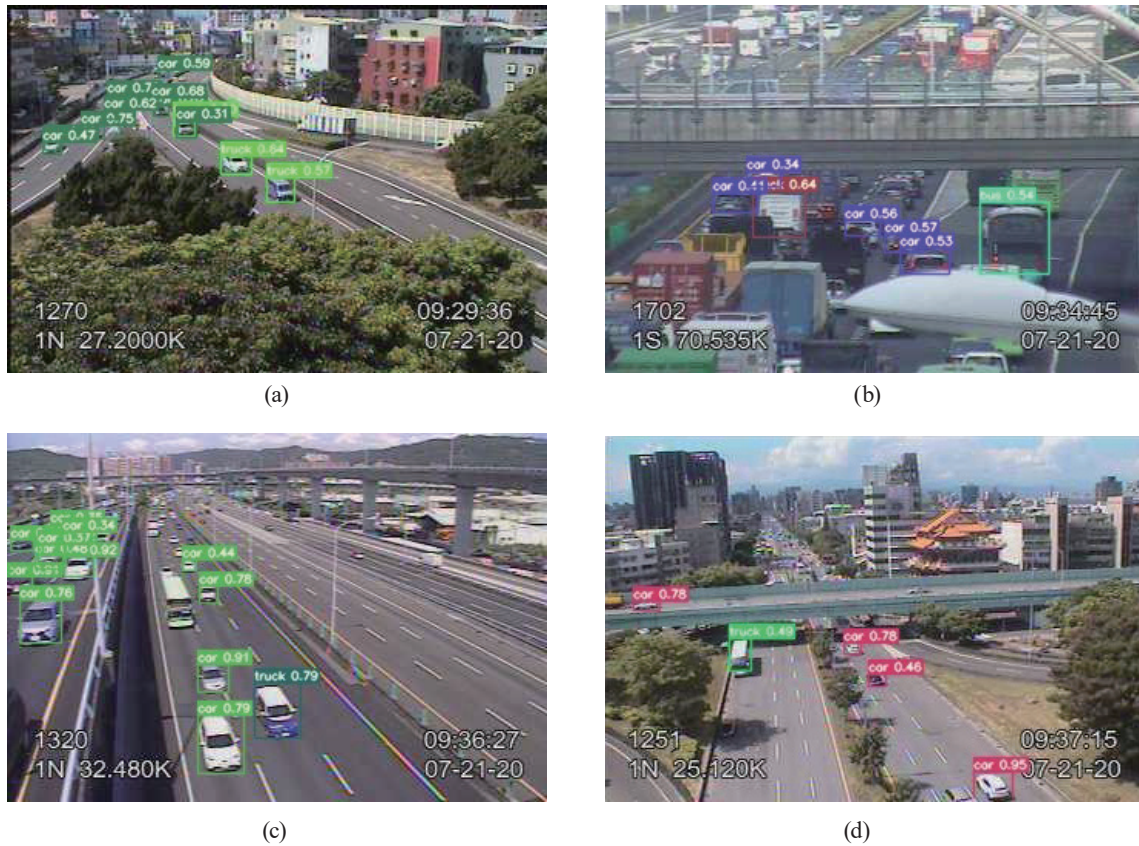


Fig. 9. (Color online) CCTV images of four freeway sections taken during the daytime, on which the identification results are superimposed. (a) National Freeway No.1 northbound at Wugu (32 km)–Sanchong (27 km). (b) National Freeway No.1 southbound at Yangmei (70 km)–Yangmei (71 km). (c) National Freeway No.1 northbound at Wugu (32 km)–Sanchong (33 km). (d) National Freeway No.1 northbound at Sanchong (27 km)–Taipei (25 km).

Rank	Road Section (km)	Direction	Avg speed of vehicles
1	Wugu(32) - Sanchong(27)	North	22
2	Yangmei (70) - Yangmei (71)	South	23
3	Wugu(32) - Sanchong(33)	North	25
4	Sanchong (27) - Taipei(25)	North	26
5	Container Terminal No.3~5 (4) - Qijin(3)	West	30
6	Jhonghe (26) - Sindian (28)	East	31
7	Siihih(13) - Donghu(15)	South	37
8	Tucheng (42) - Jhonghe (35)	North	38
9	Ruilong Rd.(369) - Wujia(370)	South	39
10	Taipei (25) - Yuanshan (23)	North	39

Fig. 10. (Color online) Congestion ranking of freeways during the daytime⁽¹⁴⁾ (captured on 2020/7/21 09:28).

Table 3 shows the recognition rate and misjudgment rate during the daytime. It can be seen that on a normal freeway, the recognition accuracy is between 58 and 71%. The reason for the low recognition rates for rows 2 and 5 is that at the time of recognition, the vehicles were in

Table 3
Recognition accuracy rate and false detection rate during the daytime.

Congestion ranking	Ramp (km)	Direction	Travel speed (km/h)	Vehicles	Correct number	Misjudgment number (%)	False detection rate (%)
1	Wugu (32)–Sanchong (27)	North	22	131	86	45	34.4
2	Yangmei (70)–Yangmei (71)	South	23	383	64	319	83.3
3	Wugu (33)–Wugu (32)	North	25	331	235	96	29.1
4	Sanchong (27)–Taipei (25)	North	26	145	84	61	42.1
5	Container Terminal No.3–5 (4)–Qijin (3)	West	30	46	3	43	93.5

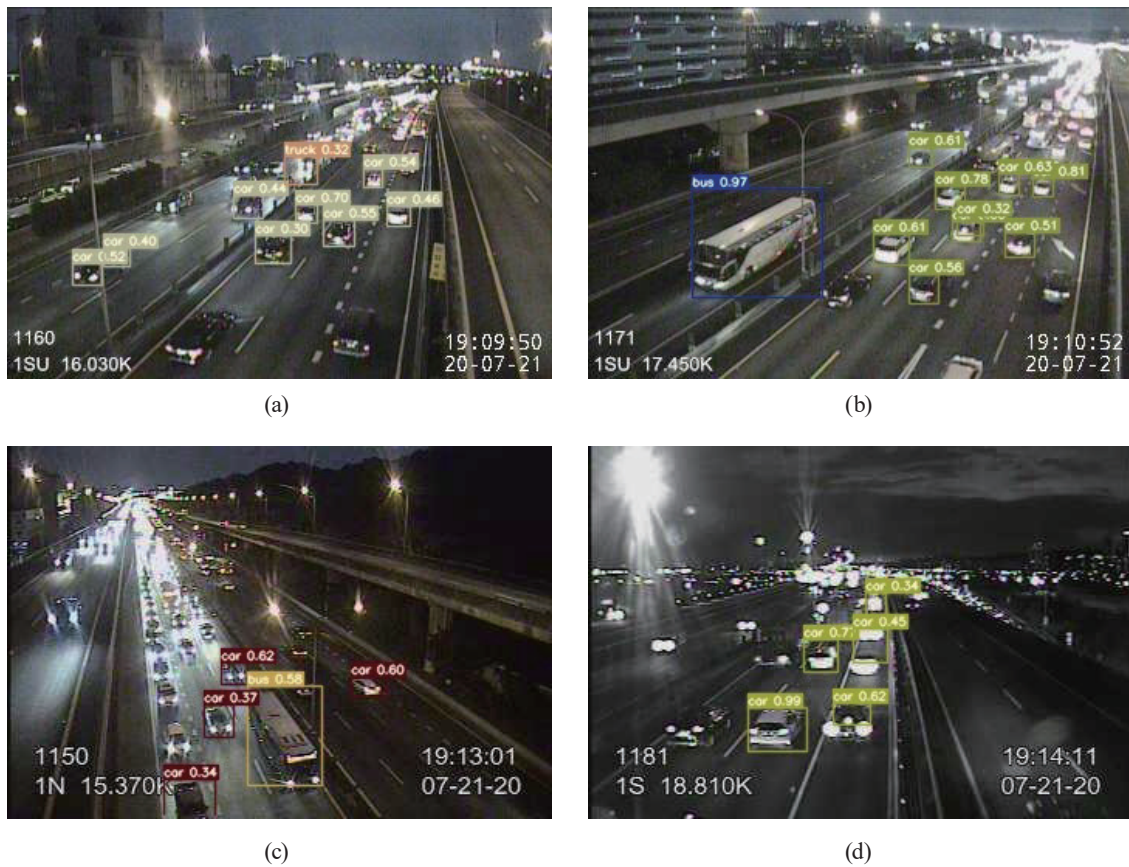


Fig. 11. (Color online) CCTV images of four freeway sections taken during the nighttime, on which the identification results are superimposed. (a) National Freeway No.1 southbound at Donghu (15 km)–Neihu (17 km). (b) National Freeway No.1 southbound at Neihu (17 km)–Yuanshan (23 km). (c) National Freeway No.1 northbound at Neihu (17 km)–Donghu (15 km). (d) National Freeway No.1 southbound at Tiding (18 km)–Xiatayou (20 km).

traffic jams, and their proximity caused YOLO to recognize several neighboring vehicles as a single vehicle.

In contrast to the daytime, Fig. 11 shows CCTV images of four freeway sections taken during the nighttime, on which the identification results are superimposed. These four road sections were chosen because, at the time of this writing, they had the some of the most serious congestion at night (average speed of 10–30 km/h) according to the Freeway Authority, as shown in Fig. 12.

Rank	Road Section (km)	Direction	Avg speed of vehicles
1	Nangang(16) - Nanshen Rd.(16)	South	16
2	Donghu (15) - Neihu (17)	South	17
3	Neihu(17) – Yuanshan(23)	South	18
4	Chongde (16) – Tanzi (18)	East	19
5	Yilan(38) – Touching(30)	North	20
6	Neihu (17) - Donghu (15)	North	22
7	Tiding Blvd(18) – Xiatayou Off-Ramp(20)	South	23
8	Xizhi (17) - Donghu (15)	South	23
9	Container Terminal No.3~5 (4) - Qijin(3)	West	27
10	Donghu (15) - Xizhi(13)	North	28

Fig. 12. (Color online) Congestion ranking of freeways at night⁽¹⁴⁾ (captured on 2020/7/21 19:07).

Table 4
Recognition accuracy rate and misjudgment rate at night.

Congestion ranking	Ramp (km)	Direction	Travel speed (km/h)	Vehicles	Correct number	Misjudgment number	Accuracy rate (%)	False detection rate (%)
2	Donghu (15)–Neihu (17)	South	17	218	105	113	48.2	51.8
3	Neihu (17)–Yuanshan (23)	South	18	380	168	212	44.2	55.8
6	Neihu (17)–Donghu (15)	North	22	360	99	261	27.5	72.5
7	Tiding (18)–Xiatayou (20)	South	23	286	119	167	41.6	58.4
8	Xizhi (13)–Donghu (15)	South	23	224	100	124	44.6	55.4

Table 4 shows the recognition rate and misjudgment rate at night. It can be seen that on a normal freeway, the recognition accuracy is between 28% and 48%. Rows 1 and 4 in Fig. 8 are excluded because there is no CCTV installed at these two locations, and row 5 is excluded because the CCTV image is an interchange down to a flat road, which is not consistent with the objective of this study. The low recognition accuracy was due to the low resolution of the original CCTV images. The accuracy is also lower than that for the daytime images because the lack of illumination increases the difficulty of recognition. The low recognition rate for congestion rank 6 is because the camera is directed at the front of the vehicle, and at night, vehicles have their headlights on, making it more difficult to identify the vehicles in the image correctly.

The YOLO algorithm enables real-time recognition and with an average recognition accuracy of 44% during the daytime and 41% during the nighttime for the CCTV images streamed back. The tunnel CCTV images provide a higher vehicle recognition rate than the other outdoor CCTV images because the cameras are closer to the vehicles and are less prone to light changes due to the time of day or dirty lenses caused by weather, as well as tree branches and other obstacles.

From the results of the experiment, we find the following.

1. The average recognition rate is 65% when there is enough illumination in the daytime and no weather-related interference.
2. The average recognition rate of the night CCTV images is 41% when there is enough illumination and no weather-related interference.

3. For the CCTV images in tunnels, because of the shorter distance between the lens and the vehicle, the images of the vehicle are larger and clearer, so better and more stable recognition is obtained continuously.
4. The camera can recognize vehicles on different roads and in different directions of traffic.
5. Outdoor CCTVs are subjected to strong winds, heavy rain, and sunshine for many years, resulting in dirty and damaged lenses, making the job of people involved in surveillance more difficult. Therefore, the recognition rate of some old and damaged CCTVs is only about 50%.
6. Because the dataset used in this study is trained on the COCO dataset, the recognition accuracy is poor for some vehicle models.

5. Conclusions

According to the preliminary results of this study, the current model based on the YOLOv3 algorithm and the COCO dataset has an accuracy of 44% during daytime and 41% during nighttime for CCTV cameras installed outdoors. However, if the picture quality is not clear, the lens is obscured, or objects are too small, the accuracy of vehicle recognition is not satisfactory, which should be improved in future work.

The main factors affecting recognition accuracy are as follows:

1. Low CCTV image resolution in some sections, indirectly affecting vehicle recognition accuracy,
2. Different road conditions, lighting (day or night), and weather, and
3. Insufficient training data.

In the future, we will attempt to develop an automated method of capturing images of vehicles from freeway surveillance cameras. After maintaining a high and stable recognition rate, it should be possible to also detect or predict whether traffic jams are occurring, or even to determine when traffic accidents have occurred as well as other situations to provide users or relevant authorities with the ability to respond quickly.

Acknowledgments

Our thanks go to Michael Burton of Asia University for English editing.

References

- 1 A Year in Computer Vision—Part 1 of 4. <https://towardsdatascience.com/a-year-in-computer-vision-part-1-of-4-caeb040b6f46> (accessed June 2020).
- 2 A. Nurhadiyatna, B. Hardjono, A. Wibisono, W. Jatmiko, and P. Mursanto: Int. Conf. Advanced Computer Science and Information Systems (ICACSIS, 2012) 179. <https://ieeexplore.ieee.org/document/6468754>
- 3 G. Desai, V. Ambre, S. Jakharia, and S. Sherkhane: Int. Conf. Smart City and Emerging Technology (ICSCET, 2018) 1. <https://doi.org/10.1109/ICSCET.2018.8537279>
- 4 C. S. Christopher and Shanna: Int. Conf. Recent Advances in Energy-efficient Computing and Communication (ICRAECC, 2019) 1. <https://doi.org/10.1109/ICRAECC43874.2019.8995084>
- 5 B. Hardjono, H. Tjahyadi, M. G. A. Rhizma, A. E. Widjaja, R. Kondorura, and A. M. Halim: IEEE 9th Annu. Information Technology, Electronics and Mobile Communication Conf. (IEMCON, 2018). <https://doi.org/10.1109/IEMCON.2018.8615085>

- 6 R. P. Loce, R. Bala, and M. Trivedi: Computer Vision and Imaging in Intelligent Transportation Systems (John Wiley & Sons Ltd, 2017) Chap. 3.
- 7 S. Ghosh, S. J. Sunny, and R. Roney: Accident Detection Using Convolutional Neural Networks. IconDSC (2019) 1. <https://doi.org/10.1109/IconDSC.2019.8816881>
- 8 A. Ćorović, V. Ilić, S. Đurić, M. Marijan, and B. Pavković: 26th Telecommunications Forum (TELFOR, 2018) 1. <https://doi.org/10.1109/TELFOR.2018.8611986>
- 9 P. Zhihao and C. Ying: IEEE 4th Advanced Information Technology, Electronic and Automation Control Conf. (IAEAC, 2019) 1558. <https://doi.org/10.1109/IAEAC47372.2019.8997591>
- 10 W. Chen and C. K. Yeo: IEEE Int. Conf. Smart Computing (SMARTCOMP, 2019) 459. <https://doi.org/10.1109/SMARTCOMP.2019.00088>
- 11 CCTV Open Data of Static Information: <https://data.gov.tw/dataset/37665> (accessed June 2020).
- 12 CCTV Open Data of Dynamic Information: <https://data.gov.tw/dataset/37666> (accessed June 2020).
- 13 COCO dataset: <http://cocodataset.org/#home> (accessed on June 2020).
- 14 Congestion Ranking Page of freeway: https://1968.freeway.gov.tw/tf_ranking (accessed July 2020).

About the Authors



Shin-Hung Pan received his M.S. and Ph.D. degrees from the Department of Information Management, Chaoyang University of Technology, Taiwan. Currently, he is an assistant professor with the Department of M-Commerce and Multimedia Applications, Asia University, Taiwan. His current research interests include cloud computing, edge computing, and IoT. (vincentpan@asia.edu.tw)



Shu-Ching Wang received her M.S. degree in electrical engineering from National Chen-Kung University and her Ph.D. degree in information engineering from National Chiao-Tung University, Taiwan. Currently, she is a professor with the Department of Information Management, Chaoyang University of Technology, Taiwan. Her current research interests include distributed computing, cloud computing, and IoT. (scwang@gm.cyut.edu.tw)