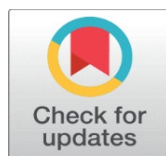


SMART CAR DOOR LOCK SYSTEMS USING DEEP LEARNING

Wen-Kung Tseng ¹, Yue-Xun Yang ¹

¹ Graduate Institute of Vehicle Engineering, National Changhua University of Education, Taiwan R. O. C.



Received 15 October 2024
Accepted 20 November 2024
Published 14 December 2024

Corresponding Author

Wen-Kung Tseng,
andy007@cc.ncue.edu.tw

DOI
[10.29121/ijetmr.v11.i12.2024.1519](https://doi.org/10.29121/ijetmr.v11.i12.2024.1519)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2024 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

This study investigated the performance of the smart car door lock system using deep learning. In the era of increasing automation, the advancement of science and technology has made people's lives more convenient, but it must be convenient and safe at the same time. The method used in this study combined facial recognition and fingerprint recognition as the verification method of the car door lock system. Also the research method is mainly divided into three parts: the application of YOLO face recognition, fingerprint recognition, and the integration of the two recognition methods, which are applied to the car door lock system. The face recognition part used the YOLO convolutional neural network, and one thousand facial images of the three people were used to train the recognition system. The weights were obtained after tens of thousands of training. After adjusting the various parameters, facial recognition could be operated. Fingerprint recognition used an optical fingerprint sensor to store the fingerprints of three people in the memory, and used software to control the signal output of the fingerprint recognition. Lastly, the facial recognition signal and fingerprint recognition signal were integrated with the software. The purpose is to achieve that the car door lock will be opened when the two verification methods are correct at the same time. On the contrary, when any verification method is wrong, the car door lock will still be closed.

Keywords: Biometrics Technology, Facial Recognition, Fingerprint Recognition, YOLO, Convolutional Neural Network

1. INTRODUCTION

According to statistics from the National Police Agency [Cahyaningtiyas et al. \(2016\)](#), there were a total of 1,456 vehicle thefts in 2019 in Taiwan. If the traditional access control system is replaced by biometrics, the theft rate may be effectively reduced. With the improvement of people's awareness of security, there are many ways to verify identity, such as magnetic cards, IC cards, passwords, etc., but this information may be copied, stolen, or forgotten. The biometric method uses the characteristics of the human body for identification, and there is no need to carry any items, which improves the convenience. The main focus of the study was on the application of deep learning using YOLO v3 proposed by Joseph Redmon et al. in 2018 [Hendry & Chen \(2019\)](#), which was an improved version of YOLO v2. It solved many previous problems and speeded up in terms of the AP index. YOLO v3 is 1.9

times faster than v2 [Joseph et al. \(2016\)](#). Pu Lia et al. proposed the research of comparing several different object detection algorithms in 2019 [Joseph et al. \(2016\)](#), which compared the detection data of Faster-RCNN, R-FCN, SSD and YOLO v3 for fire images. The results showed that the average of YOLO v3's accuracy was as high as 83.7%, and the frame rate reached 28 FPS, which was higher than other algorithms.

In the actual application part, Hendry et al. proposed the application of YOLO deep learning to automatically recognize license plates in 2019 [Joseph & Ali \(2017\)](#), using only 7 convolutional layers to detect a single category of targets, and also tested the system with different complexity conditions, for example, rainy background, dark or dim, and different tones and image saturation. The system achieved a 98.22% accuracy rate of license plate detection. In terms of face unlocking, Zhiguo Zhu and others proposed a face smart door lock system using OpenCV in 2020 [Joseph & Farhadi \(2018\)](#). They mentioned that the current smart door locks usually used fingerprints to unlock, but when the fingerprint was wrong, a password was still used to unlock. The password was relatively insecure, so face recognition was used to replace the password. It also mentioned that when the amount of data in the database increased, the error rate would decrease and the accuracy would increase. They proposed an efficient posture tracking algorithm that allows the accuracy to reach the level of 94.5%. In recent years, people's demand for security has increased, and various access control methods are gradually being developed. Cahyaningtiyas published a fingerprint access control system with integrated attendance records and user privileges in 2016 [Lu, R. \(2019\)](#), mentioning that many spaces held many important assets that were at risk when left unattended. The study used Arduino and optical fingerprint reader to build a database of privileged users and added an attendance record system to simplify the management of lab access records, taking into account both security and convenience.

With the evolution of the times, fingerprint recognition technology has flourished since the 1960s [Pu & Wangda \(2019\)](#), and the number of places where fingerprints are used has increased gradually in modern times, such as cell phones, laptops, customs clearance, and even identification at crime scenes. There are many types of fingerprint readers, and they are divided into capacitive and optical types according to the extraction principle. Capacitive sensors are many tiny capacitive arrays, and the fingerprint bumps and depressions differ in capacitance to achieve the recognition effect by using the difference in electrical capacity. The advantages are its small size and easy to install in small devices. The disadvantage is that the degree of wet and dry fingers will affect the results and the cost is higher. The optical type is based on an optical transmitter that projects light through a triangular prism or other reflective surface to reflect the fingerprint, and then allows the photoreceptor to capture the fingerprint image. The advantages are high durability and good adaptability to dry and wet fingers, but the disadvantage is its large size. In this study, the optical type was used as the fingerprint reader in the experiment.

When searching for related information or literature review, it is seldom to see the research that combines the two methods of face recognition and fingerprint recognition. Therefore, the integration of the two recognition methods was used in this study. In order to improve the security of the car door lock system, two verification methods were used to improve the reliability of the car door lock: facial recognition and fingerprint recognition. An embedded computer (NVIDIA AGX Jetson Xavier) was used as the identification system. A camera was used to capture the face images. The fingerprint recognition part uses an optical fingerprint sensor.

After registering the fingerprint of the authorized person, a fingerprint database can be established. When the finger is pressed on the fingerprint sensor, it will determine whether the fingerprint matches the fingerprint in the database. Finally, the two signals are integrated with the door lock to achieve the purpose of double verification.

2. METHODS

This section introduces the principles of YOLO convolutional neural network for face recognition and fingerprint recognition.

2.1. YOLO (YOU ONLY LOOK ONCE) OVERVIEW

YOLO (You Only Look Once) was first proposed by Joseph Redmon et al. in 2016 [Zhu & Cheng \(2020\)](#), and compared with other previous object detection methods. YOLO v1 was the first time that the task of detection was done in an end-to-end form, increasing the speed of many computations. However, one of the drawbacks was the large error in the position of the box, and the fact that v1 divided the image into 7x7 grids, each grid only detected 2 box candidate areas, which made the speed much higher, but at the same time the recall rate was very low, and many smaller targets would be missed. As shown in [Figure 1](#) [9], the framework of YOLO v1 first adjusts the image to 448x448, and then the image is imported into the neural network. The structure consists of 24 convolutional layers, 4 max pooling layers, and 2 fully connected layers.

Figure 1

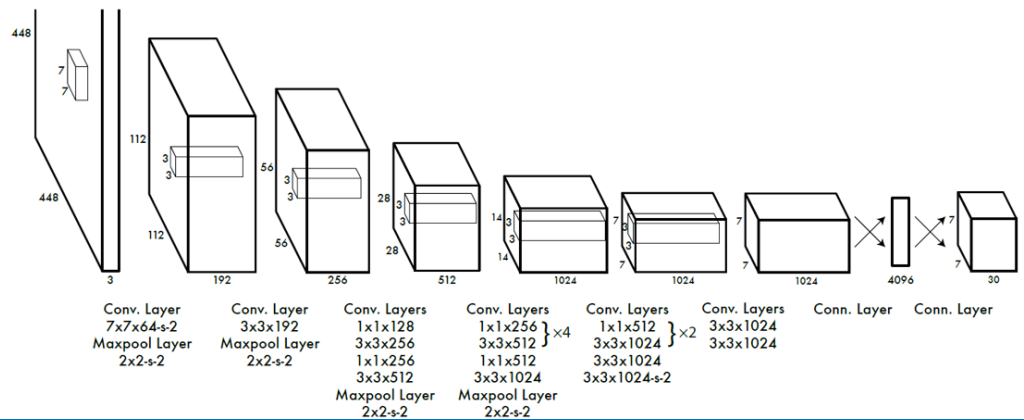


Figure 1 YOLO v1 framework

In 2017, YOLO v2 has been proposed [10], and based on v1, three parts of training, grid structure and database have been enhanced [9], the authors proposed the Darknet-19 architecture by removing the fully connected layer from the v1 version and replacing it with a fully domain-averaged pooling layer, which consists of 19 convolutional layers and 5 maximal pooling layers in the whole neural network, and adjusting the input size from 448x448 to 416x416 and increasing the feature size to 13x13, improved the recall rate. As shown in [Figure 2](#), after various enhancements, YOLO v2 was able to achieve 76.8% mAP at 76FPS and 78.6% mAP at 40FPS in the VOC2007 database, which improved the speed.

Figure 2

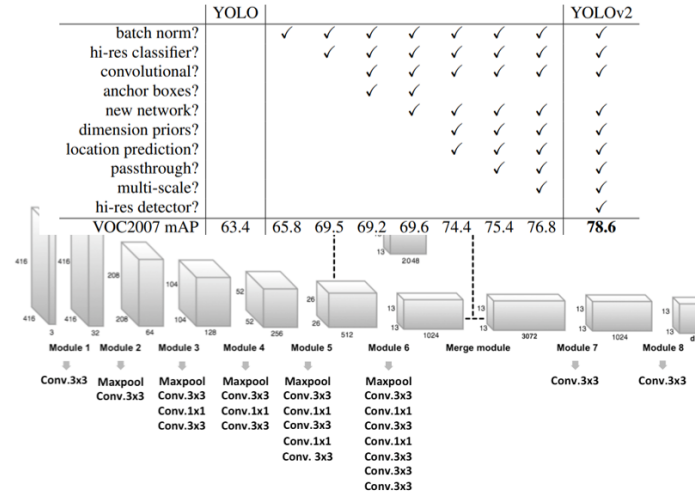


Figure 2 Comparison between YOLO v1 and YOLO v2

YOLO v3 version was proposed in 2018, this version is also the last one proposed by the original author Joseph Redmon and others, v3 was changed from the Darknet-19 architecture of v2 to the Darknet-53 architecture, starting with a 32-filter convolutional kernel, followed by five sets of repeated residual units (resblock_body), each consisting of a single convolutional layer and a repeated convolutional layer, with the repeated convolutional layers for a total of 52 layers, and the last fully concatenated layer is changed to a 1x1 convolutional layer for a total of 53 layers. Compared with Darknet-19, Darknet-53 adds more layers, so the speed is much slower, but it can still process 78 images per second [3]. YOLO v3 maintains a higher performance compared with other ResNet of the same precision.

During deep learning training, an equation is needed to evaluate the various losses in detection, such as the center marker, the width and height of the prediction boundary, the category, and the confidence index [9]. The first row represents the offset prediction of the model Bbox (bounding box), where S is the number of grids, B is the box, 1_(i,j)^{obj} means if there is a target in the box, its value is 1, and vice versa is 0, b_{*} represents prediction true, b_{*}^{*} represents ground true, the second and third rows are the confidence score prediction and class prediction of the model. The category prediction is changed to a binary cross entropy (BCE) loss function, where 1_(i,j)^{noobj} means that if there is no target in the box, its value is 1, and vice versa is 0. The loss function can be expressed as follows.

$$\begin{aligned}
 \text{loss} = & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{obj} [(b_x - \hat{b}_x)^2 + (b_y - \hat{b}_y)^2 + (b_w - \hat{b}_w)^2 + (b_h - \hat{b}_h)^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{obj} \left[-\log(p_c) + \sum_{i=1}^n BCE(\hat{c}_i, c_i) \right] \\
 & + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{noobj} [-\log(1 - p_c)]
 \end{aligned}$$

The performance of the model after YOLO training is mainly evaluated by common metrics such as IoU and mAP [9], IoU (Intersection over Union) is the intersection of the predicted range and the union of the object marker range during image detection, in other words, it is the overlap between the marker position and the predicted position of the model in the training image, and its value is from 0 to 1. The most common threshold value is 0.5, and when IoU > 0.5 it means that the

model can correctly detect the objects. and the formulation of IoU can be expressed as follows.

$$IoU(A, B) = \frac{A \cap B}{A \cup B} \quad (2)$$

mAP (Mean Average Precision) is a common metric for evaluating models, where AP (Average Precision) is the average precision of a target, and mAP is the average of the APs of all targets in the model.

2.2. PRINCIPLES OF FINGERPRINT RECOGNITION

Fingerprint recognition technology is a biometric identification technology, which is a set of fingerprint image acquisition, processing, feature extraction and comparison modules of the pattern recognition system. Fingerprints are lines formed by the bumpy skin on the end of the fingers of primates, and these lines are also imprinted on objects. Fingerprints are characterized by details such as starting points, ending points, triangular points, and bifurcation points. Since fingerprints are different for each person, and different fingers of the same person have different fingerprints. Fingerprint identification is done by comparing these detailed features. The human fingerprint contains a large number of messages, called fingerprint features. There are many feature points, which provide separate identification information and are the basis for fingerprint recognition. There are general features, local features, and the general features in the fingerprint as shown in Figure 3. These include the center point (the progressive center of the fingerprint), the triangle point (the point where two or three lines meet), the bifurcation point (the point where one line bifurcates into two lines), and the number of lines (the number of lines in the fingerprint); the local features are the details of the fingerprint, the direction, curvature, and location of the nodes at the feature points, which are important indicators to distinguish different fingerprints.

Figure 3



Figure 3 Fingerprint Features

There are many ways of fingerprint recognition, such as optical, capacitive, etc. These recognition technologies have their own advantages and disadvantages. The optical type mainly relies on light sources, triangular prism. CMOS and other units use light to irradiate the finger to make fingerprints appear, and then through the CMOS to capture fingerprint images for processing. The advantages are lower cost, high durability, and better adaptability to wet and dry fingers. The disadvantage is the larger size. The capacitive type is scanned by the charge change of fingerprints or pressure difference, and the sensors have many tiny capacitive arrays. The difference in capacitance between the raised and depressed areas of the fingerprint is used to identify the difference in capacitance, the advantage is that it is small and

easy to install in a small device. However, the disadvantage is that it affects the result when the finger is wet or dry, and the cost is higher. [Table 1](#) shows the comparison of different fingerprint recognition devices.

Table 1

Table 1 Comparison of Different Fingerprint Recognition Devices		
	Optical	Capacitive
Principle	Fingerprint reading based on light reflection	Capacity change according to concave and convex difference
Advantages	More durable, good wet and dry finger adaptation, low cost	Smaller size
Disadvantages	Larger size	Less durable, poor adaptation to dry and wet fingers, high cost

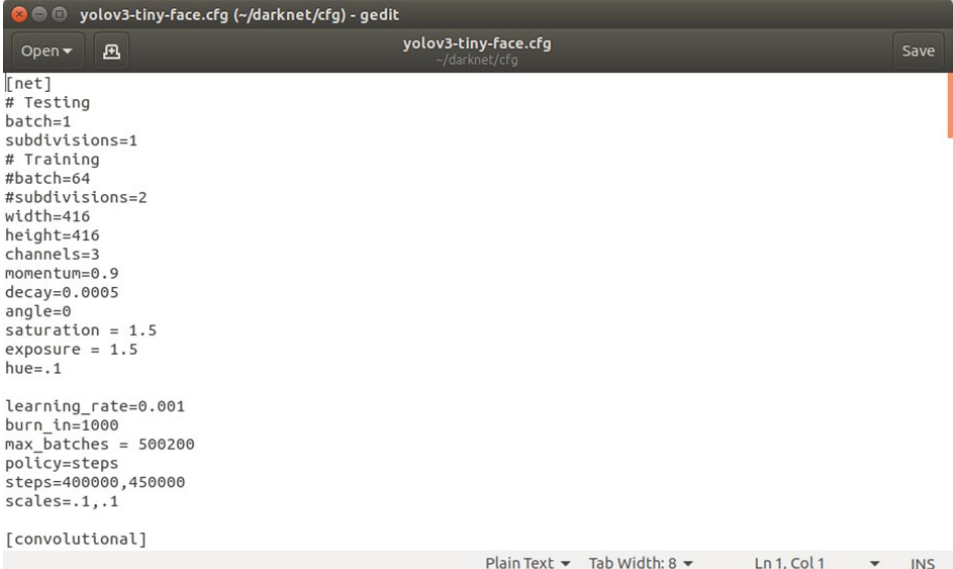
3. RESULTS AND DISCUSSIONS

This section introduces the experimental setup, experimental procedures and results.

3.1. FACE RECOGNITION

In this experiment, the face recognition part has been done using deep learning with a large amount of image data needed for training. First, the face images of three people were taken in the experiment for different angles of the face. The total number of the face images is 3000. After finishing the image collection, each image needs to be manually labeled. LabelImg software was used for labeling in this experiment. Then corresponding txt files were generated for each image. There are five values in the file, namely class, x, y, w, and h, where x and y are the center coordinates relative to the image width and height, and w and h are the ratio relative to the image width and height. Finally the cfg files for training were set up as shown in [Figure 4](#). The deep learning controller is used for training, and the detailed specifications for the deep learning controller are shown in [Table 2](#).

Figure 4



```

yolov3-tiny-face.cfg (-/darknet/cfg) - gedit
yolov3-tiny-face.cfg
~/darknet/cfg
Save
[net]
# Testing
batch=1
subdivisions=1
# Training
#batch=64
#subdivisions=2
width=416
height=416
channels=3
momentum=0.9
decay=0.0005
angle=0
saturation = 1.5
exposure = 1.5
hue=.1

learning_rate=0.001
burn_in=1000
max_batches = 500200
policy=steps
steps=400000,450000
scales=.1,.1

[convolutional]

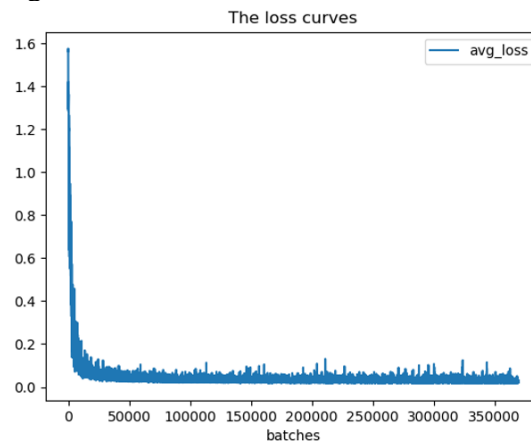
```

Figure 4 CFG File

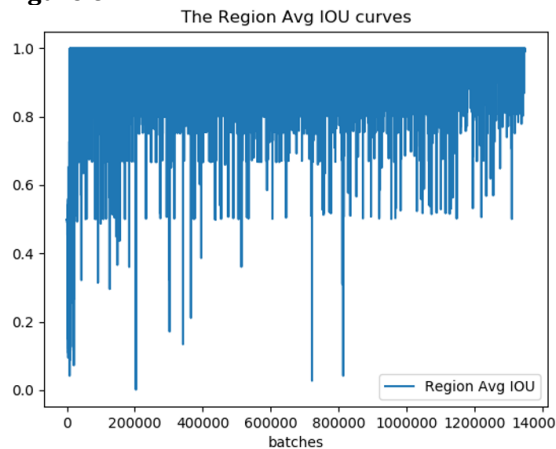
Table 2**Table 2 Detailed Specifications for Deep Learning Controller**

OS	Ubuntu 18.04
CPU	Intel Core i7-8700K 3.2GHZ
RAM	32 GB
GPU	GeForce GTX 1080 Ti

The deep learning model was set to train 400000 times. After the training, the Loss and IoU would be output in graphical way. The Loss value is the loss function of the model, and the smaller value means the less loss as shown in Figure 5. The Loss value is very large at the beginning of the training, but it starts to decrease obviously after 10,000 times, and finally it is close to 0.

Figure 5**Figure 5** The Loss Value During Training

IoU (Intersection over Union) is another index to judge the training of the model, which is to measure the degree of overlap between the predicted position of the model and the actual position of the object. When the degree of overlap is high, the IoU value should be close to 1 as shown in Figure 6. As the number of training times increases, the IoU gets closer to 1 and stays mostly above 0.8 by the end of training.

Figure 6**Figure 6** The IOU Value During Training

After the training, the deep learning model was used for photo, video, and real-time verification. In the real-time verification, the faces of three people were detected as shown in Figure 7. It can be seen that the three people could be detected with 100% confidence.

Figure 7

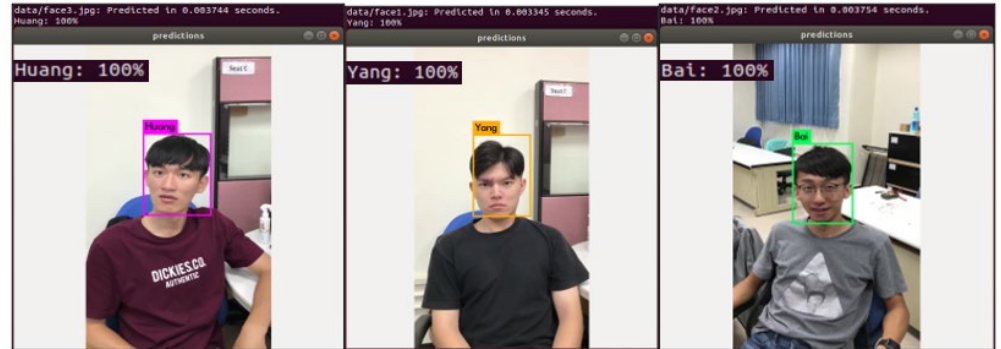


Figure 7 Real-time verification

3.2. FINGERPRINT RECOGNITION

This experiment used an optical fingerprint reader combined with an Arduino to control the desired device. The fingerprints of each person were registered. There were three people in this experiment, and each person registered two fingerprints. When an authorized fingerprint was detected, the ID and confidence level of the fingerprint would be displayed. However, when an unauthorized fingerprint was pressed on the reader, "unauthorized fingerprint" would be displayed.

The identification light was installed on the door panel. The white light is on when the identification is correct and the red light is on when it is incorrect for the user to observe as shown in Figure 8 and Figure 9.

Figure 8



Figure 8 White Light Turns on When Fingerprint is Correct

Figure 9**Figure 9** Red Light Turns on when Fingerprint is not Correct

3.3. INTEGRATING FACE AND FINGERPRINT RECOGNITION

After completing the face recognition and fingerprint recognition experiments, the two recognitions were integrated into the smart car door lock system. The signal output from the GPIO of NVIDIA jetson AGX Xavier was connected to the Arduino. The face recognition signal pin was set to facepin. The fingerprint recognition signal pin was set to fingerpin, and the door lock relay signal pin was set to lockpin. When both facepin and fingerpin are 1, the signal would be output to the lockpin to make the relay operate and open the door lock. The platform of the smart car door lock system is shown in Figure 10. After turning on the system, the current status of the camera could be displayed on the screen. When the face in the database was detected, the terminal window would show true and output the signal. On the other hand, when the fingerprint reader was pressed and detected correctly, the white light on the door panel would light up and send out a signal. When both signals were 1, the relay would open the smart car door lock.

Figure 10**Figure 10** The Platform of a Smart Car Door Lock System

4. CONCLUSIONS

This study was to design a smart car lock system that combines face and fingerprint recognition. Nowadays, the door lock system usually uses keys, cards, etc.; higher-level ones may have either fingerprints or face recognition to operate. The smart door lock systems integrating face and fingerprint recognition were rarely used. So based on biometric identification, a smart car door lock system was

developed. The experiment was mainly divided into several stages. First stage was to use the YOLO neural network for face recognition. To build the face recognition system, face images from three participants with various angles and different background were used for training the deep learning system. Second stage was to build the fingerprint recognition system. The final stage was to combine these two systems into the smart door lock system. The results showed that the smart door lock system could operate correctly.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Cahyaningtiyas, R., Arianto, R. & Yosrita E. (2016). Fingerprint for Automatic Door Integrated with Absence and User Access, 2016 International Symposium on Electronics and Smart Devices (ISESD). 26-29
<https://doi.org/10.1109/ISESD.2016.7886686>
- Hendry, R. & Chen, C. (2019). Automatic license Plate Recognition via sliding-window Darknet-YOLO Deep Learning, Image and Vision Computing. 87, 47-56. <https://doi.org/10.1016/j.imavis.2019.04.007>
- Joseph, R. & Ali, F. (2017). YOLO9000: Better, Faster, Stronger. Computer Vision and Pattern Recognition (CVPR).6517-6525.
<https://doi.org/10.1109/CVPR.2017.690>
- Joseph, R. & Farhadi, A. (2018). YOLOv3: An Incremental Improvement, arXiv,1804-1812.
- Joseph, R., Santosh, D., Ross, G., & Ali, F. (2016). You Only Look Once: Unified, Real-Time Object Detection, arXiv., 1506-1512.
- Lu, R. (2019). AP/mAP/IoU? [Transcription]
- Pu, L. & Wangda, Z. (2019).Image Fire Detection Algorithms Based on convolutional Neural Networks, Case Studies in Thermal Engineering., Article 100625.
<https://doi.org/10.1016/j.csite.2020.100625>
- Zhu, Z. & Cheng, Y. (2020). Application of Attitude Tracking Algorithm for Face Recognition Based on OpenCV in the Intelligent Door Lock, Computer Communications. 154,390-397.
<https://doi.org/10.1016/j.comcom.2020.02.003>