CTU Journal of Innovation and Sustainable Development
ISSN 2615-9422 | e-ISSN 2815-5602

DOI:10.22144/ctujoisd.2023.041

Face and number plate recognition for car anti-theft

Quoc Bao Truong[1*], Tan-Loc Tran[2], Tan-Kiet Thanh Nguyen[3], and Huu-Cuong Nguyen[1]
[1]*College of Engineering, Can Tho University, Viet Nam*
[2]*Soc Trang Vocational College, Viet Nam*
[3]*National Political Publishing House Can Tho Branch, Viet Nam*
**Corresponding author (tqbao@ctu.edu.vn)*

Article info.

Received 16 Jul 2023
Revised 10 Sep 2023
Accepted 25 Sep 2023

Keywords

Anti-theft car, convolution neural networks (CNN), object detection, optical character recognition (OCR)

ABSTRACT

Smart parking systems along with that is continuous development of new technologies, are widely applied to improve our lives. It can also add new technologies with advanced functions making it a multi-functional management system. Thanks to the anti-theft technologies that are installed, modern cars are significantly more difficult to steal than they once were. However, electrical systems can still experience issues, though and malfunction at some point. This paper suggests using video image recognition technology, at car parks and parking lots as an anti-theft solution, alerting the presence of the non-owner of vehicle in the driver's seat. The system automatically predicts whether the driver is valid with the registration number plate. The image of the car is captured by the camera at the entrance gates of the parking lot. The proposed algorithm includes face recognition in images, building a deep learning convolutional network that classifies faces (subscriber's images); using Cascade trainer to train number plate object recognition, vehicle number recognition through character recognition technique. The system can recognize reality through a personal computer connected to the camera at the scene or through photos and video files. Result, the model can face recognition and match to license plate in at a moment.

1. INTRODUCTION

1.1. The smart system using IoT

An array of smart parking system technologies using IoT have been developed based on: ultrasonic, electromagnetic field detection and infrared: The sensor is detectable small vary in the magnetic field when a metal thing is near; Infrared: This is a type of sensor, it measures ambient temperature changes and detects motion. In the context of a smart urban ecosystem, everything input from elements (vehicles, roads and users) has to be linked and analyzed jointly in order to supply the best service in a secure and fast manner (Moneeb et al., 2018).

Anti-theft systems play a crucial role in protecting cars from theft. In order to dissuade thieves from taking the vehicle, modern anti-theft systems are sophisticated and complex, incorporating GPS tracking devices, immobilizers, alarms, and more. The most popular type of anti-theft system currently used in cars is GPS tracking equipment. These in-car gadgets use satellites to continuously determine the precise location of the car. If the car is stolen, police authorities can get access to this data, which will help them locate and recover the vehicle swiftly.

Another form of anti-theft equipment used in modern cars is the immobilizer. In order to prevent the engine from starting without an approved key.. Numerous immobilizers also have an alarm system that will go off if someone tries to hot-wire the vehicle or otherwise tamper with it.

As mentioned above, this paper proposes using video image recognition technology as a solution to anti-theft, warning when the non-owner of a vehicle is in the driver's seat. The system can automatically identify whether the driver matches the registration number plate.

1.2. Face detection

American mathematician Bledsoe (1964) experimented with face detection by computer in 1964. His team used a scanner to receive people's face images and find matches to recognize faces by program computers. However, it was unsuccessful because he used rudimentary scanner and the computer's difficulty with facial expressions, lighting and pose.

Face detection was improved in methodologies came in 2001 when Viola and Jones (2001) - the computer vision researchers at the Mitsubishi Electric Research Laboratories proposed a model in real time to detect faces with high accuracy. The Viola-Jones framework is on the basis of training a model to acquire knowledge of what is a face and what is not a face. The model extracts particular features once trained, they are collected so that they can be compared with various stages features from new images. If the image tested passes through the feature contrast, then a face has been detected and operations continued.

The Viola-Jones framework has limitations, although it is used to recognize people's faces in real-time applications. For example, it possibly not work if a person's face with a face covering or scarf, or if the face is not in accordance with orientation, it might not find people's face. In recent years, face detection using deep learning has brought advances, it outperforms traditional computer vision methods.

The terms face recognition and face detection are used switchably, and they both belong to face identification. However, facial recognition is an application of face detection. Facial recognition software is implemented for biometric verification, for mobile apps and unlocking phones. Facial recognition to anti-crime and avoid violence in the banking, retail and transportation industries. Face recognition goes over detecting a person's face. It can specify whose face. The process captures a digital image of a person's face by a computer application or extracts frame from a video, then confronts it with images in storage.

Several methods were used for face detection, The following are advantages and disadvantages:

Viola-Jones algorithm. It uses a machine learning model to recognize what is a person's face and what is not. The problems recognized faces that are covered or not in accordance with oriented, although this method is still popular in real-time applications for recognizing faces.

Knowledge or rule based. This method characterizes a face based on rules. However, setting up clear knowledge based rules can be a difficult task.

Feature-based or feature invariant Features like a person's nose or eyes are used to recognize a face in these methods. Those can be bored by light and noise.

Template matching. To detect a face, it confronts images with stored defined face patterns or features and similar two. However, this method struggles to resolve variations in scale, pose, and shape.

Appearance-based. It uses machine learning and statistical analysis to find similar features of face images. Changes in lighting and orientation fight this method.

Convolutional neural network based. A type of deep learning ANN is a convolutional neural network (CNN) used in image validation and processing on each pixel data. A region-based CNN is an R-CNN, that restricts to a particular place or part of the objects and arranges in classes or categories according to characteristics of objects in images. This method is mainly used on areas or regions, in an image that are the same to other places, for example, the pixels at a region of an eye, a nose, or a mouth. If this region of them matches up with another place, then they know CNN has found a match. However, this method (CNN) can become very compound when they "overfit," which means they match places of noise in the training data and not the test patterns.

Single shot detector (SSD). This approach based on a region network like R-CNN requires two images, one for generating place proposals and the other for finding the object of each proposal, SSDs only claim one image to find many objects inside the image, so that SSDs perform faster than R-CNN. However, SSDs have difficulty finding small faces and farther away in image.

1.3. Automatic number plate recognition (ANPRR)

Automatic number plate recognition (ANPR) was created in 1976 at the Police Scientific Development Branch in the United Kingdom (ANPR International Ltd, 2018). Original systems were put in by 1979, and produced industrial systems were awarded contracts, first at EMI Electronics, and then at Computer Recognition Systems (CRS, now part of Jenoptik) in Wokingham, UK. On the A1 road and at the Dartford Tunnel, early trial systems were deployed. In 1981, the first arrest was made a stolen of car by detection (ANPR International Ltd, 2018). Although, ANPR was not widely interested and used until there were new developments and cheaper during the 1990s. However, the ANPR data collected for future use (i.e., in handling then undetected crimes) was restored in the early 2000s (Taylor, 2005).

In terms of software, the system works on standard computer hardware and can be linked to databases or other applications. First, it uses a lot of image processing methods to detect, normalize, and segment the image of the license plate, and then optical character recognition to show the character item of the license plate.

ANPR uses optical character recognition (OCR) on photos captured by cameras. When vehicle number plates were used a style more legible, the font was changed, inserting a small gap in some letters to them more plain and therefore more readable to such systems. Some number plates use transformation in font sizes and positioning. Such differences must be able to cope by ANPR systems to be truly effective. Complete systems can process with distinct tailored to each country.

In this paper, we propose an outside car anti-theft warning of the non-owner of the vehicle in the driving seat by confirming the driver image matches the owner of the vehicle (subscriber's images). This system can detect car theft crime in parking or internal areas.

2. MATERIALS AND METHOD

We propose a model to use an image of the front of the car. We extract the frame in a video, then we will detect face and number plate. The research hypothesis is that each registered car number has a limited number of people allowed to drive (subscribers' images). That means each registered car number has a pretraining model for driver recognition. The proposed algorithm includes automatic number plate recognition and face recognition in images by a deep learning convolutional network that classifies faces. If the result face in the frame matches subscriber's images then pass, or else alert if a stranger is driving.

2.1. Detect object Cascade of Boosted Classifiers

Viola and Jones advanced a face detection algorithm grounded on AdaBoost learning algorithm and using Haar-like features Viola and Jones (2001). The input to the algorithm is a digital image. The algorithm searches for Haar-like features of the face on each small image region of the original image. Each such small region is rectangular in shape and is called a candidate or sub-window. The algorithm will scan through the entire image using this sub-window. The size of this sub-window can be expanded to search for faces of various sizes. To evaluate whether or not a candidate is a face, the algorithm uses classifiers called Haar-like Feature Classifier to find the Haar-like features of the candidate.

The algorithm uses an integrated image to support the processing of the candidate's Haar-like feature. The algorithm performs candidate evaluation through several stages. Each stage can use many Haar-like features to evaluate a candidate. Each Haar-like feature is classified by a Haar Feature Classifier. Each classifier Haar Feature Classifiers will calculate and produce a result. Next, the algorithm will sum the results of the Haar Feature Classifiers and compare this total value with a threshold value of the period under consideration to conclude that the candidate has passed this stage and continued to the next stage. next or eliminated. If a candidate can pass all the evaluation stages, then it is considered a face.

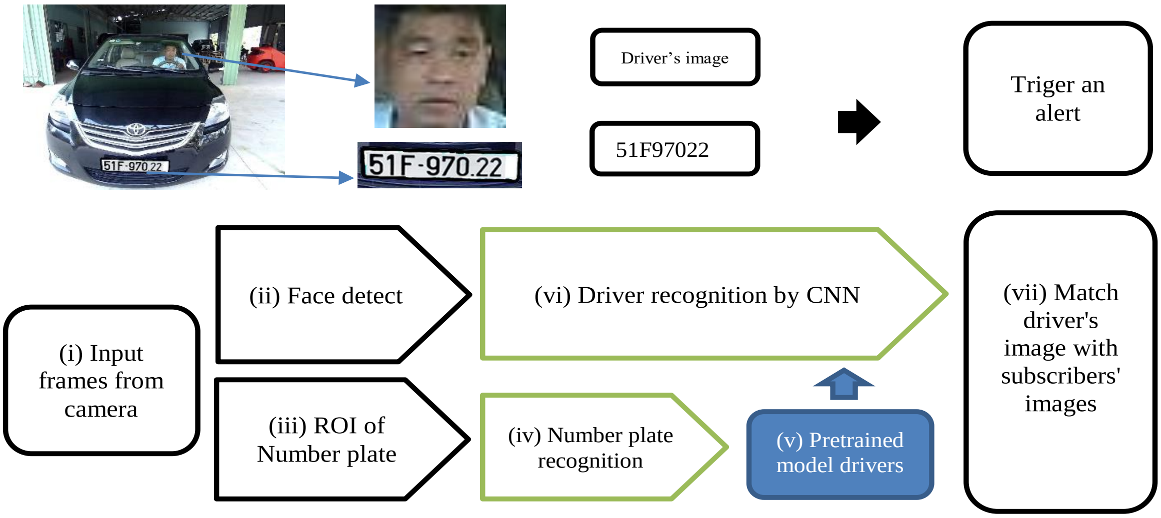


Figure 1. The architecture of the proposed system

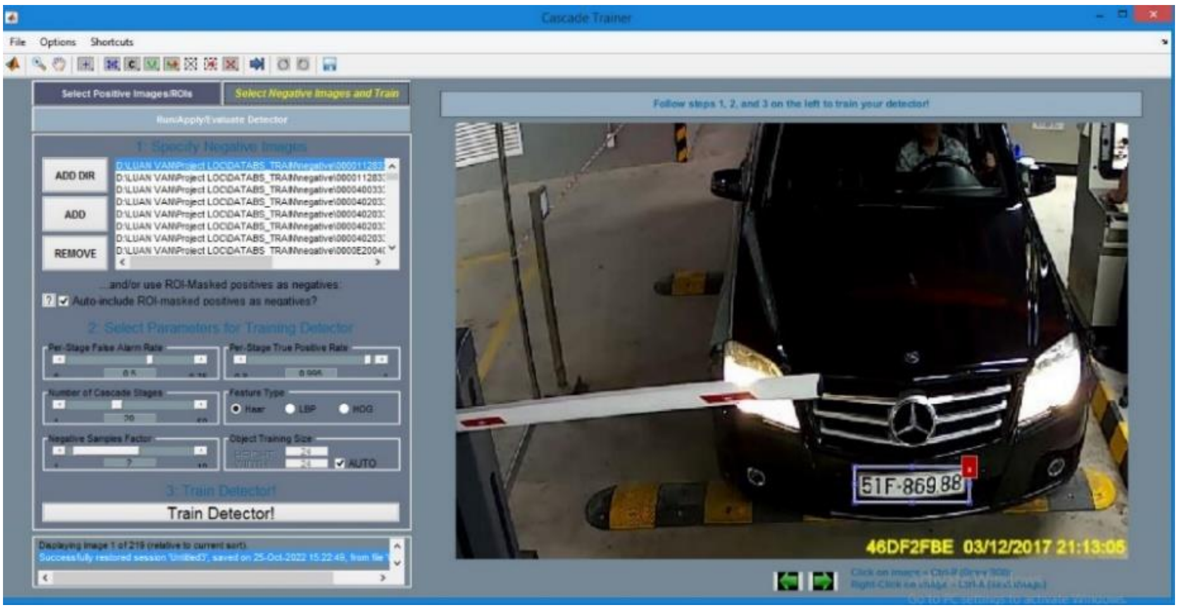


Figure 2. Choose regions of interest (ROI) for positive images and setup parameters in MATLAB's Cascade trainer tool

In this paper, we use MATLAB including the vision.CascadeObjectDetector System object go with a variety of pre-trained classifiers for finding eyes, noses, frontal faces, profile faces, and the upper body and a tool *Cascade Trainer* (Brett, 2023) to train the detection of long license plate regions. A set of positive samples along with that is a set of negative images requires training by Cascade classifier. We must determine a set of positive images with regions of interest (ROI) pointed out to be used as positive samples. The Image Labeler outputs a board to use for positive samples. We also must determine a set of negative images from which the function in *Cascade Trainer* generates negative samples.

2.2. Optical Character Recognition (OCR)

Optical character recognition was developed in 1974 by Ray Kurzweil, who started Kurzweil Computer Products, Inc. This document-transforming technology could recognize text with any font. He determined that his method will be developed into machine learning device the best used for those who are blind. A reading machine was made by him, that could read text out and translate text into a text-to-speech format. Xerox bought his company to in 1980, as They were interested in commercialize paper-to-computer text transformation.

Letters on the image seceded out by OCR software, then the letters joined into words and the words joined into sentences and it enabled means of entry to and editing of the first content. Thanks to that we don't need data entry manually. The texture of a document image was also analyzed by an OCR program. It construes the document page into items such as images, tables, or blocks of text. The lines are construed into words and from there separate into characters. Once the characters have been separated from words, the program will match characters with a set of pattern images. They present you with the recognized text, after processing all likely matches (IBM, 2022). Optical character recognition (OCR) is occasionally also called by name text recognition. It has developed for almost 50 years, but the more technology has been applied to our lives, the more new technology has become in our daily lives.

2.3. Face recognition with Convolutional Neural Network (CNN)

2.3.1. Convolutional Neural Network (CNN)

One of the effective identification methods is Convolutional Neural Network. It developed a widespread practice in object classification, and recently, this method has been applied more widely in computer vision and image processing (Krizhevsky et al., 2012), (LeCun, 2015), and this method also performs better than traditional methods (Chatfield et al., 2011) through wide published documents. CNN includes the image passing through one or more layers with couples of convolutional followed by max pooling. A convolutional layer has many processes, therein set of kernels (filters) are used that scan and apply small local parts of the input, where these kernels are duplicated in the whole input space. A max-pooling layer built a lower-resolution version of the convolutional layer activations by giving the maximum kernel activation from different positions within a defined window. This holds translation invariance and tolerance to minor differences in places of object parts. There are parts of higher layers on lower resolution inputs that need to be processed more complexly, we use more map kernels so that work. The top fully connected layers finally do the classification of the overall inputs, they were combined from all positions of inputs. This architecture has good outcomes related to image processing.

2.3.2. Deep learning proposal

Convolutional neural network architectures were designed differently for each input image size. In this paper, we use the size of images at the input layer with 160×160 pixels. The architecture is defined in Figure 3. After each convolution layer (Conv) a Rectifier activation function (ReLU) is used, to come for each pooling layer, and MaxPooling is applied. Top convolutional layers with a filter size of 1×1 were defined by the fully connected layers as it is regular in MatConvNet (Vedaldi & Lenc, 2015). The final layer has n subclass matching to n classification of datasets. After all layers, a SoftMax loss is applied.

The convolutional neural network architectures of the model defined with 5 layers include:

[Convolutional + Rectifier activation function + MaxPooling] x 2 → [Convolutional + Rectifier activation function] → [Convolutional + Fully connected] → Softmax.

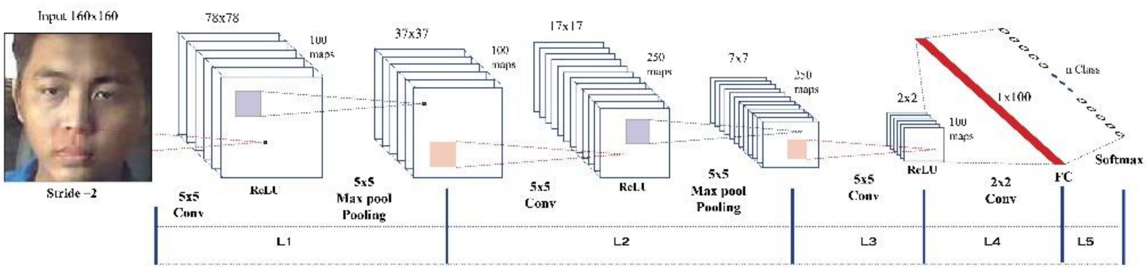


Figure 1. The proposed CNN architecture

3. RESULTS AND DISCUSSION

3.1. Experimental dataset

The images long license plates dataset for training by Cascade trainer was collected from cameras in parking lots. Includes over 500 "positive" (possible) images that contain license plates and 800 "negative" images that are random images that do not contain license plates.

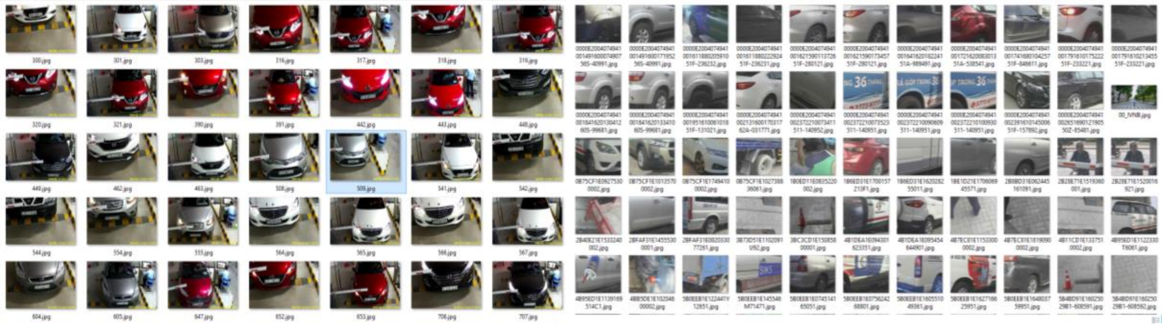


Figure 4. The images dataset "positive" and "negative" for training with Cascade trainer

The image of the driver in this experiment is extracted from the video image when the driver is sitting at the steering wheel, then using the *vision.CascadeObjectDetector* System in MATLAB to extract the face of the driver to be identified. In this paper, the data set is collected to identify several drivers used for testing. Photos are collected in different light and dark time periods.

Figure 5. The results of detecting number plate and face by Cascade Trainer algorithm

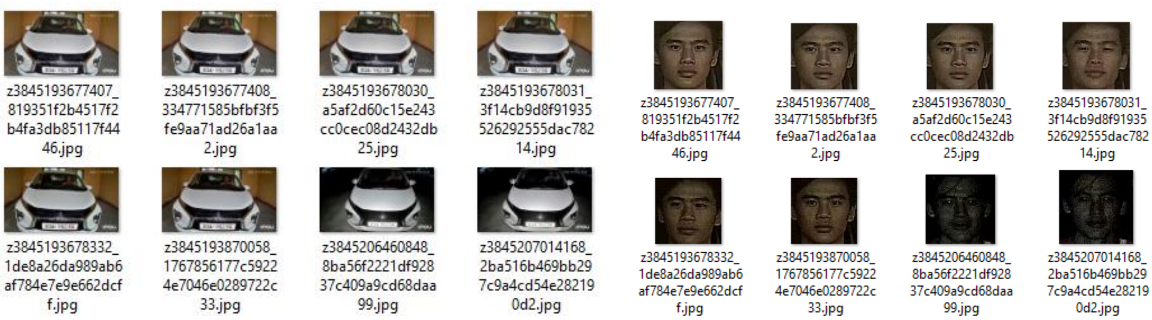


Figure 6. The images dataset face drivers then using the *vision.CascadeObjectDetector* System

The first dataset 10-labels for training is described in Table 1. The data is randomly divided into 3 sets of train/validation/test according to the percentage of 60%/20%/20%.

The second dataset 12-labels collection for training is described in Table 2.

Table 1. The dataset 12-labels partitioned

Label	Number of samples	Train 60%	Val 20%	Test 20%
1	150	90	30	30
2	50	30	10	10
3	130	78	26	26
4	150	90	30	30
5	150	90	30	30
6	50	30	10	10
7	50	30	10	10
8	100	60	20	10
9	50	30	10	10
10	50	30	10	10
11	50	30	10	10
12	70	42	14	14
Total	1050	630	210	210

Table 2. The dataset 10-labels partitioned

Label	Number of samples	Train 60%	Val 20%	Test 20%
1	150	90	30	30
2	50	30	10	10
3	130	78	26	26
4	150	90	30	30
5	150	90	30	30
6	50	30	10	10
7	50	30	10	10
8	100	60	20	20
9	50	30	10	10
10	50	30	10	10
Total	930	558	186	186

3.2. Experimental results

3.2.1. Number plate recognition or identify the vehicle

After determining the number plate area, the image is segmented with filters Prewitt, Sobel, to detect the border in the image, the purpose is to find the boundary of the number plate character, then identify the characters in the frame, by correlating the image region containing the characters with the built in character pattern matrix. Here, we setup MATLAB code *n*umber plate *recognition* based on Pankaj (2018) for our work.

Convert the original image to binary image, use edge detection, clean image. Then separate the containers (objects) that the same set of points, here are 0 and 1. Thus, we get the objects that are images of digits and characters on the image.

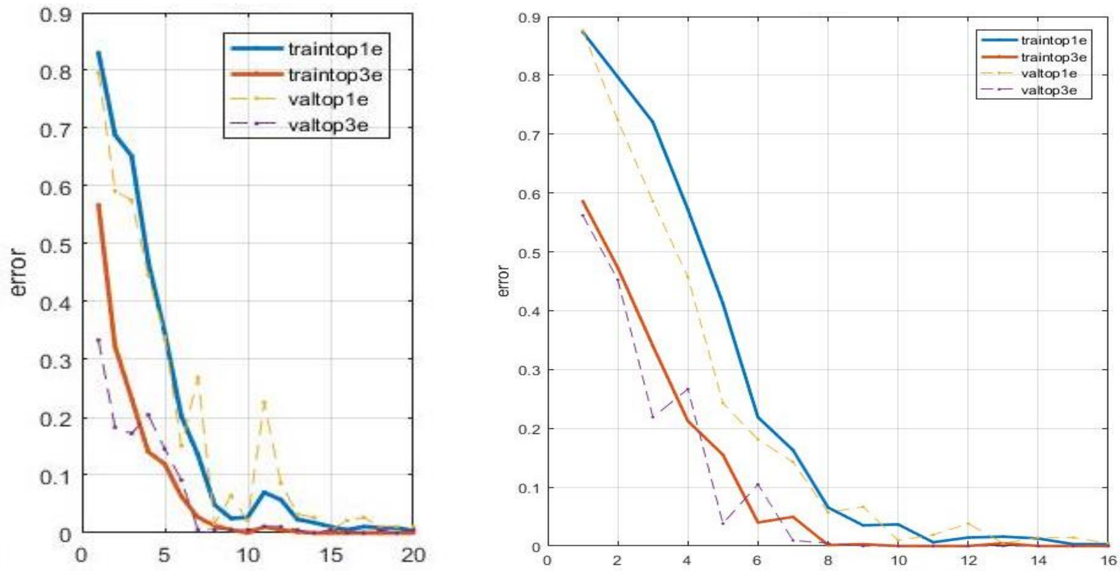
→ 51F97022

Indeed, number plate recognition can be replaced by RFID (Radio Frequency Identification) technology. This technology allows unique identification an object, which is a form of wireless transmission that includes the handle of electrostatic or electromagnetic coupling in the radio frequency portion of the electromagnetic spectrum. They are also widely used to automate non-stop toll collection through electronic toll collection. Facial recognition can be combined with license plate recognition technology or RFID technology to detect crooks and issue warnings to the owner of a vehicle or the police.

3.2.2. Driver recognition

With the dataset 10-labels in 20 training epochs, the model error is equivalent to 0. The learning process with good model error at epoch 16.

Experimental results the identification of driver with 2 datasets, accuracy is over 99%. The success rate of the system on the experimental 2 dataset is presented in Table 3 and Table 4.



a) dataset 10-labels *b) dataset 12-labels*

Figure 2. Curves of train and validation errors model's driver identification

We saw that, in each dataset, the true false positive was only 1 image, therefore, the proposed CNN model can respond well to the task of face recognition. This can be combined with license plate recognition technology or RFID technology to warn to owner of vehicle when there is a stranger at seat driver.

Table 3. Experimental results on the dataset 12-labels

Label	Number of images	True	False	Success rate (%)
1	30	29	**1**	96.66
2	10	10	0	100
3	26	26	0	100
4	30	30	0	100
5	30	30	0	100
6	10	10	0	100
7	10	10	0	100
8	10	10	0	100
9	10	10	0	100
10	10	10	0	100
11	10	10	0	100
12	14	14	0	100
mAP (mean average precision)				**99.72**

Table 4. Experimental results on the dataset 10-labels

Label	Number of images	True	False	Success rate (%)
1	30	30	0	100
2	10	10	0	100
3	26	26	0	100
4	30	30	0	100
5	30	30	0	100
6	10	10	0	100
7	10	10	0	100
8	20	19	**1**	95.0
9	10	10	0	100
10	10	10	0	100
mAP (mean average precision)				**99.5**

4. CONCLUSION

This paper outlines the development of a proposed technology for driver face recognition in images by deep learning convolutional network and number plate recognition together. This allows for the recognition of the driver (subscriber's images), which can give an early warning of theft at the security gate. This technique can be deployed in car parks or family garages.

However, the system still has some limitations that need to be addressed in the future. These include the inability to fully detect small-sized faces (distant from the camera) or objects that are occluded by shadows, photos taken through the glass. Therefore, it is necessary to study more about the angle, light mode, and what type of camera is appropriate to clearly see both the license plate and the driver's face.

REFERENCES

ANPR International Ltd, UK. (2018). *History of ANPR*. http://www.anpr-international.com/history-of-anpr/

Bledsoe, W. W. (1964). The model method in facial recognition. *Technical Report PRI 15.* Panoramic Research, Inc., Palo Alto, California.

Brett, S. (2023). *Cascade trainer: Specify ground truth, train a detector*. https://www.mathworks.com/matlabcentral/fileexchange/39627-cascade-trainer-specify-ground-truth-train-a-detector. MATLAB Central File Exchange.

Chatfield, K., Lempitsky, V.S., Vedaldi, A., & Zisserman, A. (2011). The devil is in the details: an evaluation of recent feature encoding methods. *BMVC, 2*(4), 8.

IBM. (2022). https://www.ibm.com/cloud/blog/optical-character-recognition.

Krizhevsky, A. Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems,* (pp. 1097-1105).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-44.

Moneeb, G., Muhammad, M., & Arif, U. R. (2018). SMART TSS: Defining transportation system behavior using big data analytics in smart cities. *Sustainable Cities and Society, 41,* 114-119, ISSN 2210-6707. https://doi.org/10.1016/j.scs.2018.05.008

Pankaj, K. (2018, November 26). Car number plate detection using MATLAB and image processing

https://circuitdigest.com/tutorial/vehicle-number-plate-detection-using-matlab-and-image-processing.

Taylor, J. A. (2005, April 7). Forensic person tracking method and apparatus. *U.S. Patent Application No. 10/957,999.* https://patents.google.com/patent/US20050075836

Vedaldi, A., & Lenc, K. (2015). MatConvNet: Convolutional neural networks for MATLAB. *Proceedings of the 23rd ACM International Conference on Multimedia, ACM,* (pp. 689-692).

Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1, pp. I-I). IEEE.