Monitoring the behaviours of pet cat based on YOLO model and raspberry Pi

Rung-Ching Chen*, Vani Suthamathi Saravanarajan, Hsiu-Te Hung

Deptartment of Information Management, Chaoyang University of Technology, Taichung City, Taiwan

ABSTRACT

With the progress of the times and the rapid development of science and technology, machine learning and artificial intelligence are increasingly used in transportation, logistics, and homes. In terms of pets, pet monitoring has also become very popular in recent years. In this study, a real-time monitoring system for home pets using raspberry pie is developed. The proposed method consists of a raspberry Pi based YOLOv3-Tiny identification system for rapid detection and better boundary frame prediction of the cat behavior. Based on the YOLOv3-Tiny method, the following fine-tuning is implemented: (1) The images collected using the mobile phones are uniformly cropped to 416*416 pixels. (2) The images were captured in different periods and randomly rotated to plus or minus 20 degrees to make the dataset more robust for training. (3) One thousand four hundred pictures of the cat's movements in the room for marking and training are used. The dataset is also used to train the YOLOv3 model. According to the input image categories, the results of the output are categorized into six cat actions. They were sleeping, eating, sitting down, walking, going to the toilet, and search for a trash can. The average accuracy of both models was 98%. Based on environmental constraints like speed and memory consumption, any one of these models can be used for image recognition. The raspberry Pi system continuously captures the images and instantly sends a message to the registered mobile phone to achieve an instant preventive measure if the cat goes to the toilet for too long or flips through the trash can.

Keywords: Raspberry Pi, Pattern recognition, YOLO model, Deep learning.

1. INTRODUCTION

Nowadays, most people keep pets, and the ratio of dogs to cats is the highest among pets. In recent years, more and more people have held pets. According to the Taiwan agricultural commission (2021), the population under 15 years of age has fallen by 4% every year. In contrast, the number of cats and dogs increased from 1.789 million in 2011 to 2.51 million in 2017. The growth rate is quite fast. It is worth noting that in 2017, although the number of people who raised dogs was still higher than that of people who raised cats, from the growth data in recent years, cats grew by 27%, while dogs only grew by 2%. The data estimated based on the number of dogs and cats and the average annual rate of increase and decrease in the child population from 2011 to 2017. Trends (2021) company predicts that the number of dogs and cats in Taiwan will exceed the number of children under 15 for the first time in the second half of 2020. It's estimated that 78 million dogs and 85.8 million cats are owned in the United States. According to the American Pet Products Association (APPA, 2021), approximately 44% of all households in the United States have a dog, and 35% have a cat. Based on the above analysis, we know that people keeping pets have increased in many countries, so the business opportunities of the pet industry will gradually rise.

Many families keep pets, how to take care of pets has become an important issue for modern people. Globally, the proportion of households raising dogs is higher than that

of cats. Most people keep dogs outside their homes, but some of them keep them inside their home. Untrained pets may urinate anywhere and bite the furniture in their home, causing distress to the owner, environmental factors such as poor air circulation, light time, and excessive odds and ends in the home may cause pets to have pica and urinary tract diseases. Although there are many pet surveillance products on the market that can check what the pet is doing through the camera, it is impossible for the owner to pay attention to the pet surveillance video anytime. If the owner leaves the pet surveillance video, the pet will eat or bite by mistake. Wires or trash cans can be harmful to pets.

Summarizing, the main contribution of this paper is fourfold: (1) This research collects photos of various behaviour of cats at home and uses Raspberry Pi and YOLO deep learning to develop a deep learning model to analyze their behaviour. (2) Through the model, the cat behaviour like when the cat search on a trash can or goes to the toilet for too long, the message is transmitted to the owner's mobile phone. (3) To present a detailed model performance by comparing YOLOv3-Tiny with YOLOv3. (4) Finally, proposed an end to end real-time pet monitoring system. The content of this article is organized as follows. Section 2 introduces related work, and section 3 describes the system architecture and YOLOV3-Tiny algorithm. Section 4 discusses the relevant experimental settings and the experimental results. The last section shows the summary and future of this article.

2. RELATED WORK

2.1 Pet Care

According to World population review (2021), Taiwan has the lowest fertility rate among the world countries in 2019, with a fertility rate of only 1.218. According to the experts, Taiwan will experience negative population growth in 2022. Because of this, the current society has entered an ageing and declining birthrate. Most modern people don't want to get married and have children and want someone at home to accompany them, so they choose to keep pets instead of having children, and the cost will not cause too much burden. At present, the proportion of keeping dogs is higher than that of cats. But the growth rate of cats is higher than that of dogs in recent years.

According to the Animal Protection Department (Taipei city animal protection office, 2021), the most common cause of death in cats is kidney failure. The cause of kidney failure in cats is chronic water shortage and urinary tract infection. If a cat has a urinary tract infection, the most important symptoms are frequent visits to the toilet and no urine discharge. If there are too many items in the house, it will cause the pet to have pica, which indirectly increases the risk of kidney failure in the cat. Thus, monitoring the six

normal day to day behaviour of the pets is effective in preventing further damage to the household and infection or injury to the pet. In this paper, the normal day activities like sleeping, eating, sitting down, walking, going to the toilet and search on a trash can are used for analysis.

2.2 Raspberry Pi

Initially, the Raspberry Pi is a Linux system single-board computer designed for education. In recent years, Raspberry Pi is improved continuously and made very small but has excellent performance. Many people use Raspberry Pi in various aspects, such as education, medical care, transportation, security, and home. Chen et al. (2020) combined the raspberry Pi and the gesture control board and proposed a dual authentication gesture recognition architecture, the preprocessing is performed on the collected gesture features and then transmits the data to the server for deep learning prediction. It has good accuracy in four and six gestures. In terms of home, (Nadafa et al., 2020) proposed a smart mirror with a home intrusion detection system, using Raspberry Pi as a model and using (Viola and Jones, 2001) classifier to detect faces and eyes. When an intruder breaks into the home, the system will record the face of the intruder and send a message notifying the user, In addition, the smart mirror is the same as a normal mirror, so it is not easy to be noticed by intruders. Chakraborty et al. (2020) proposed Facial Biometric System using Raspberry Pi 3 with the client server model and running a Local Gradient Hexa Pattern parallel algorithm.

The Raspberry Pi and camera used in this research is shown in Fig. 1, the hardware adopts Raspberry Pi 4 Model B V1.2 version, its hardware specification processor adopts Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz; GPU adopts Dual Core VideoCore VI® Multimedia CoProcessor; Open GL ES 3.0; hardware-accelerated OpenVG; H.265(4Kp60) high-profile decode; the memory is 2GB LPDDR4-2400 SDRAM (shared with GPU); ethernet uses Gigabit Ethernet; the wireless network adopts dual-frequency 2.4G Hz and 5G Hz IEEE802.11ac, Bluetooth 5.0, BLE; the display interface uses two micro-HDMI (4Kp60); the other interfaces are 2 USB 3.0 ports and 2 USB 2.0 ports; Storage adopts microSD 64G; the camera adopts Raspberry Pi Camera Module V2.1, its specification is 8 million pixels; 3280x2464 pixel still photos; SONY IMX219 sensor chip; camera rate supports 1080p 30FPS, 720p 60FPS, 640x480p 90FPS. The operating system uses NOOBS to install the Raspbian system.

Compared to Raspberry Pi 3, the fourth-generation Raspberry Pi has significant upgrades in all aspects. The CPU is three times faster. In terms of USB ports, for the first time, USB 3.0 is used. In terms of video output, the original Composite RCA and HDMI v1.3 upgraded to two sets of HDMI v2.0.

Fig. 1. Raspberry Pi 4 and raspberry Pi camera

2.3 Deep Learning

Deep learning is a new field developed from artificial neural networks in machine learning. With the recent development of deep learning, many scholars have achieved good accuracy in image recognition, natural language processing, and biomedicine. Hansen et al. (2018) used a self-trained CNN algorithm and original SVM model to recognize the face of 10 pigs on the farm. In the final result, the accuracy of the SVM model was only 91%, whereas the model by Hansen et al. (2018) the accuracy rate of the human using self-trained CNN algorithm is as high as 96.7%, and the lower detection error rate can reduce the cost of labour and time. Nguyen et al. (2017) proposed an automated wild animal detection system that uses a CNN model to detect wild animals. The accuracy of detecting most of the wild animals is as high as 90%, and the accuracy of domestic animals is as high as 96%. Khatri et al. (2020) used SSD networks to detect dog breeds. The SSD network uses multi-scale features for detection, which has a good performance in detection accuracy. In this research, the average accuracy of detecting toy poodle dogs is as high as 96.7%. Wu et al. (2018) used the traditional CNN algorithm to identify dog breeds. Kaggle data, 5000 dog-related pictures were extracted to conduct this experiment. The accuracy rate was as high as 85% for the 50 common breeds. In the other 120 uncommon varieties, the accuracy is significantly lower at the rate of 64%. Borwarnginn et al. (2019) believe that the most pets kept in the family are dogs. Because there are many types of dogs, each breed has unique diseases and health conditions, so to provide appropriate treatment, it is necessary to identify the breed. In this research article, regional binary pattern (LBP) and directional gradient histogram (HOG) improve the traditional CNN algorithm to identify dog breeds. The experimental results show that the traditional CNN algorithm has an accuracy rate of 79.25%. The improved method has significantly improved the accuracy rate to 96.75%. The researches mentioned in the literature review shows that deep learning has a valuable contribution to animal image recognition, so this research will also conduct experiments based on deep learning techniques. Demir et al. (2020) designed an energy-efficient image recognition system for multiple frames based on the environmental conditions, composed of a Raspberry Pi 3 Model B, a Pi NoIR Camera v2.1, and 850 nm LEDs combined with the convolutional neural network (CNN)-based animal recognition block used for monitoring marine animals.

2.4 YOLO

YOLO was first mentioned in 2015 by Joseph Redmon in the paper You Only Look Once: Unified, Real-Time Object Detection (Redmon et al., 2015), YOLO is a target detection system based on a single neural network. Unlike other algorithms, YOLO divides the picture into multiple segments and predicts the bounding box coordinates of each unit and the probability of its category. Afterwards, Joseph Redmon et al. (2015) improved YOLO and proposed YOLOv2 and YOLOv3. The full text of YOLOv2 is YOLO 9000: Better, Faster, Stronger (Redmon and Farhadi, 2016) and can detect 9000 kinds of objects. A new feature extractor called Darknet-19 is used in many types of research to improve the detection speed and accuracy of the model. To achieve better classification results, YOLOv3 has made improvements based on Darknet-19 proposed by YOLOv2 and adopted a more personal convolutional layer neural network named Darknet-53, which is mainly composed of 1X1 and 3X3 convolutional layers with a total of 53 layers. As the number of network layers continues to deepen, many researchers used the ResNet (He et al., 2015) structure to solve the gradient problem, which reduces the difficulty of training deep networks and improves the accuracy more obviously.

Because YOLO has good detection speed and detection accuracy in object detection, many scholars use the YOLO model for image recognition research. Liang et al. (2020) used the YOLOv3 model to detect litchi fruits in the natural environment at night. For better detection results for litchi at night, this study cross-validated the brightness level and distance range of night searchlights to determine the best combination of brightness and lighting. The final detection results also have good accuracy. Wu et al. (2019) improved YOLOv3, using DenseNet (Girshick et al., 2013) is advantages in model parameters to replace the backbone network of YOLOv3 for feature extraction, alleviating the problem of in-accurate detection and overlapping bounding boxes in the original network, and forming YOLO-Densebackbone convolutional neural network, and the traditional YOLOv3 reference, the improved algorithm detection accuracy increased by 2.44%. Hongwen et al. (2019) used pigs in the farming scenes as the research object, and proposed a DAT-YOLO model that combines channel attention module (CAB) and spatial attention module (SAB) to improve Tiny-YOLO, DAT-YOLO retains Tiny-YOLO's multi-scale feature extraction block to ensure strong detection performance for faces of different sizes. Due to

the fusion of CAB and SAB, it can be improved without significantly increasing the amount of calculation and parameter Model feature extraction performance, in the face pose category, the original Tiny-YOLO average accuracy rate is 73.99%, while the improved DAT-YOLO average accuracy rate is 82.38%.

2.5 Smart Home

The smart home is one of the well-known technologies. In today's society, more and more attention given to convenience, safety and privacy. Smart home brings better safety, comfort and convenience to families. With the help of a smart home system, homeowners can easily monitor home conditions and control household electronic products, such as lighting equipment, air conditioners, doors and windows, and sweeping machines. Smart homes have always been a hot topic, so many scholars are researching and developing a complete set of smart home systems, such as Botticelli et al. (2018) proposed a smart home that can monitor environmental parameters, control doors and windows, and security services. The home system, system is equipped with sensors on the doors and windows. When the homeowner is not in the house and detects that a stranger is moving the doors and windows, the system will send a warning message to the homeowner's mobile phone. The homeowner can also check the current state of the doors and windows of the house through the mobile phone. Al Rasyid et al. (2018) established a smart home system using Arduino as the node and Raspberry Pi as the controller. The system provides a complete web interface, and the homeowner can open and close each node from the web interface. It can monitor the temperature and humidity in the house, and control the doors and windows. The homeowner can enter his mailbox and password on the web interface. When the motion sensor in the house detects the movement in the area, it will send a notification to the registered cell phone.

At present, smart homes are gradually becoming popular, and many companies on the market give smart home platform services, such as Home Kit (2021), Apple's smart home platform. Home Kit can identify and operate various accessories and supports up to 20 types of smart devices, including Sockets, fans, doorbells, security, etc. Although many smart home platform services are already available on the market, many platforms are limited to using specific accessories, so that consumers purchase many accessories to integrate with the platform services, while IFTTT (2021) does not have such a problem, it can work with different APPs, connected devices and software services, and provides a built-in Applet, which can achieve smart home automation. IFTTT supports many devices such as home appliances, smart curtains, indoor environment control, Smart lighting, smart security, smart hubs and bridges, etc. The Applet that IFTTT can establish has a high degree of freedom and can trigger another smart device through time, place and sensors. Therefore, in this research, IFTTT is used Connect the Raspberry Pi to the mobile phone and the computer. When any abnormal behaviour in the cat is detected, it will immediately send a notification to the registered mobile phone. In recent years, the successful application of edge computing in the IoT is applied in many applications like smart cities, smart homes and smart industries with the integrated sensor systems. Yang et al. (2021) proposed a facial expression recognition system using the Raspberry Pi with multiple classification algorithms and can be deployed anywhere with network connections.

3. PROPOSED METHOD

3.1 System Structure

The architecture proposed in this study is as described in the abstract. We use the YOLOv3-Tiny model to train the collected cat pictures. According to the literature review, many types of research related to image recognition using Raspberry Pi and IoT are ongoing. Many Deep Learning algorithms combined with Raspberry Pi makes image recognition very effective. In the introduction to YOLO in the literature overview, we know that YOLO has good detection accuracy in image recognition. Fig. 2 shows the system architecture, designed considering the compatibility of the equipment by choosing the Raspberry Pi camera as the input device camera. Then used the model already established on the Raspberry Pi for image recognition and IFTTT network service Platform for transmitting the detected behaviour images to mobile phones. IFTTT is a cloud platform for people who need to design automated smart homes. The same design extended for pet identification, which can well connect APP and Internet by integrating the device and software. This research also uses VNC (Virtual network computing, 2021) to detect the images remotely. VNC consists of the client and server model. Install VNC Server on the Raspberry Pi to set the IP and password, and Install VNC Viewer on the computer and mobile phone as the monitoring terminal. So in this research, when the behaviour is detected, it will automatically send a message to the mobile phone and view the Raspberry Pi image through the computer and mobile phone.
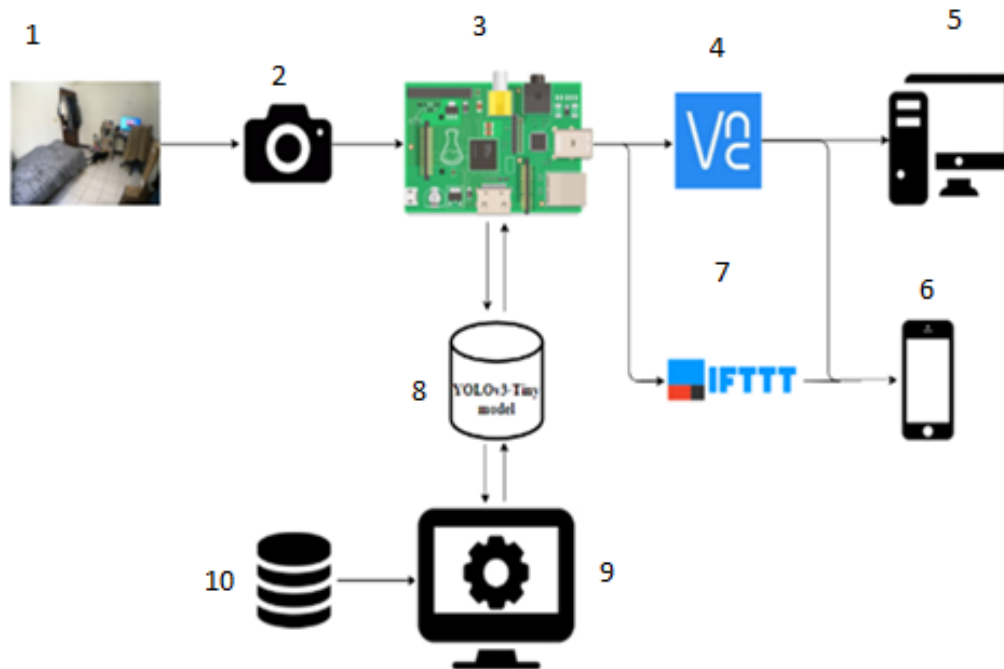
Fig. 2. The system workflow

The system workflow of Fig. 2 is shown as follows.
1. A real-time image is taken from a Raspberry Pi camera.
2. Connect the camera of the Raspberry Pi.
3. Raspberry Pi 4 Model B V1.2.
4. VNC is a free and open source remote control tool. VNC can be divided into the client (VNC Viewer) and server (VNC Server). Install VNC Server on the Raspberry Pi to set the IP and password, and then Install VNC Viewer as the monitoring terminal with the computer.
5. The computer monitors real-time images through VNC Viewer.
6. If the Raspberry Pi detects the behavior, the phone will instantly display the cat's behavior, accuracy, and show the image.
7. IFTTT can integrate different apps, connected devices, and software services. When the Raspberry Pi detects a behavior, it saves the behavior, accuracy, and photos through a written program and then sends these to the phone through IFTTT.
8. Pre-trained Yolov3-Tiny model
9. Considering that the Raspberry Pi training model will cause a burden and take too long a training time, we train the computer model in advance.
10. Pre-processed data

3.2 YOLOv3-Tiny

In this research, considering the configuration of the Raspberry Pi hardware, YOLOv3-Tiny is used as the object detection model. YOLOv3-Tiny is simply a simplified version of YOLOv3. Although it is not as good as YOLOv3 in detection speed and accuracy, here in this article, the experiment results have shown good performance. While compared to YOLOv3, the Tiny version compresses the network a lot, omits some feature layers and does not use the residual layer, and only uses two different scale output layers of 13*13 and 26*26. Since the backbone network of YOLOv3-Tiny is relatively shallow, the higher-level semantic features cannot be extracted. However, the advantage of YOLOv3-Tiny is that the network model is simple and the computational burden is small, which is convenient for use on mobile devices such as Raspberry Pi. Fig. 3 shows the YOLOv3-Tiny network model.

In the frame prediction model, YOLOv3 continues the practice of YOLOv2. The final prediction is shown by the bounding box. A bounding box is an imaginary rectangle drawn on the object detected, outlining the object within the image by defining the b_x, b_y, b_w, and b_h, representing the X coordinate of the centre, Y coordinate of the centre, the width and height of the bounding box respectively. The equations are given below:

$$b_x = \sigma(t_x) + C_x \tag{1}$$
$$b_y = \sigma(t_y) + C_y \tag{2}$$
$$b_w = p_w e^{t_w} \tag{3}$$
$$b_h = p_h e^{t_h} \tag{4}$$

Where σ is the sigmoid function, c_x and c_y are the coordinate offsets of the cells, p_w and p_h are the side lengths of the preset anchor box. t_x, t_y, t_w, and t_h are the predicted output of the model, t_x and t_y are the predicted coordinate offset values, t_w and t_h are scale scaling. Fig. 4 shows the anchor box.
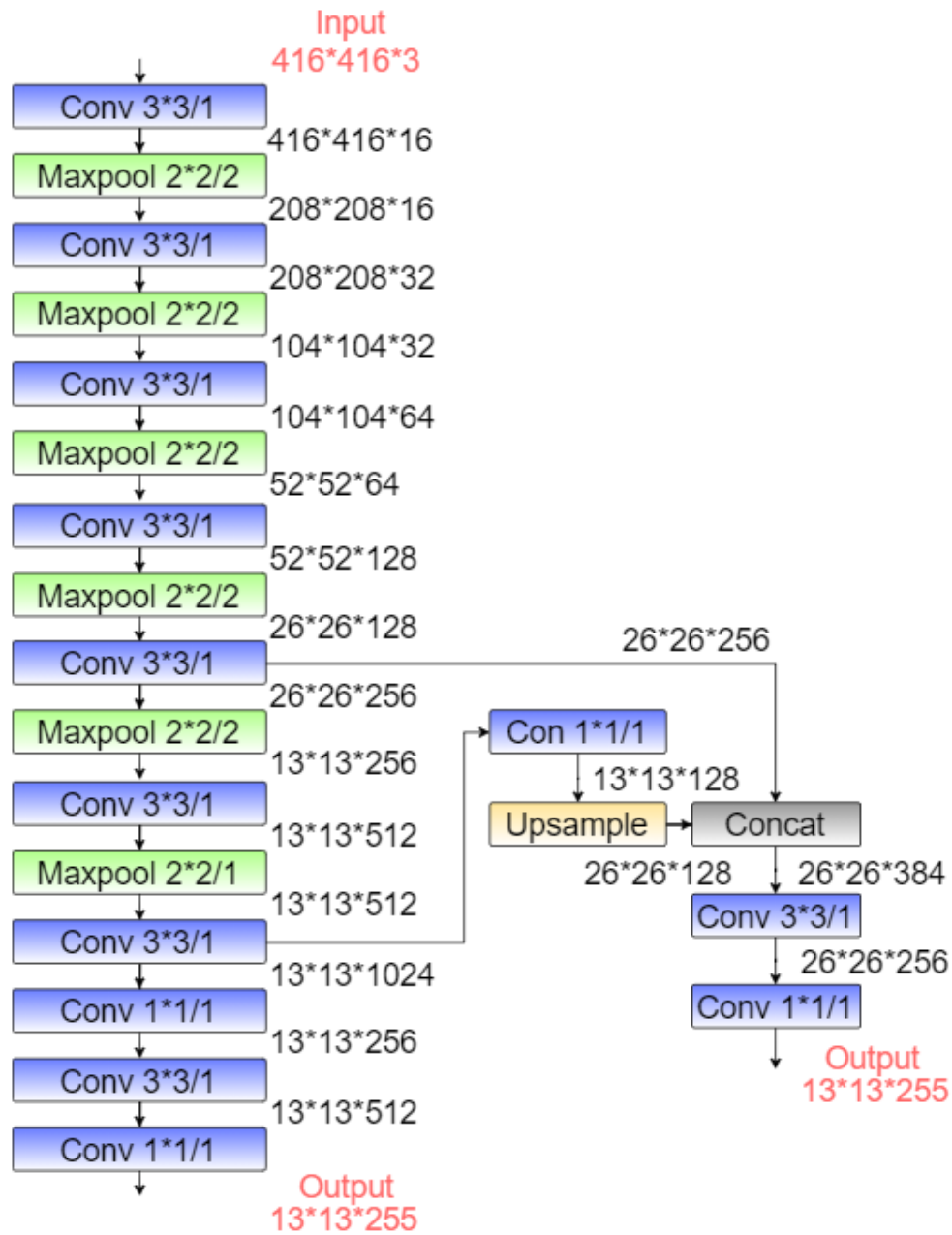
Fig. 3. YOLOv3-Tiny model

3.3 Monitoring Pet Behaviour

The pet behaviour monitoring system captures the pet's day-to-day activity and triggers a notification to the owner in case of any abnormal behaviour. The monitoring system captures the pet's sleeping, eating, sitting down and walking behaviour, and the owner can monitor the change in their daily activities. As mentioned in the literature review, there is a need to monitor the two most critical behaviour to prevent illness or accidents. Thus the monitoring system sends a notification to the owner when the pet urinates for more than 30 seconds and searches on a trash can. This way, the owner is not overburdened with messages, but at the same time, an owner can monitor the pet's subsequent activities.
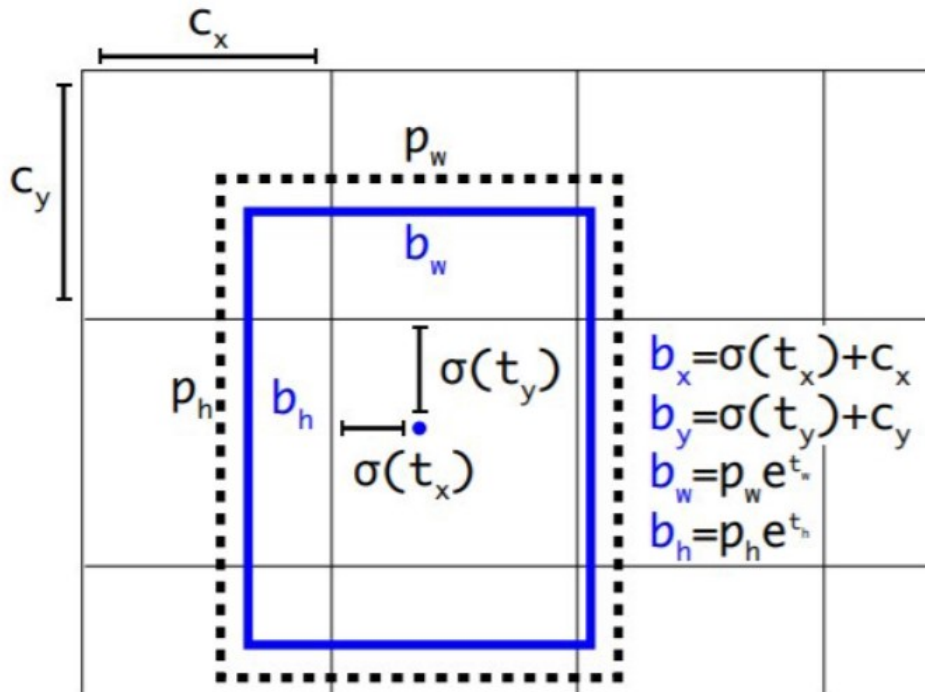
Fig. 4. The anchor box

Fig. 5. Mark pictures

4. EXPERIMENTAL RESULTS

4.1 Data Set and Experimental Environment

In the data set, we collected 246 photos of walking, 180 photos of eating, 232 photos of sleeping, 210 photos of sitting, 260 photos of search on a trash can, and 282 photos of going to the toilet. The training set and the test set are 7 to 3, and the pictures cropped into 416*416. Use a mobile phone to shoot in the morning, noon and night to make the data set more representative. The data set collected using the mobile phones and uniformly cropped the collected pictures to 416*416 pixels. To improve the accuracy of the dataset, the angle of the images is randomly rotated within plus or minus 20 degrees, and the shots are taken in different periods to make the data set more representative in training.

The hardware devices used in this research are Raspberry Pi 4 model B v1.2, Raspberry Pi Camera module v2.1, and their specifications are 8 million pixels. The software uses the Yolov3-tiny model and uses the Bbox label tool **(**2021) to mark the coordinate position and category label. In the

yolov3-tiny parameter, set the learning rate to 0.001, batch to 64, subdivisions to 16, max_batches to 12000, and step to 9600 and 10800, finally get the best accuracy.

4.2 Training and Results

In the training process, to verify the performance of the YOLOv3-Tiny model, we also train the YOLOv3 model for comparison, using the same training set and test set. Tested 432 images in a test set of 6 categories after training the model. Fig. 6 shows the average loss and mAP of YOLOv3-Tiny. The smaller the loss value, the better the model performance is. This experiment used a total of 12,000 training steps to analyze the training process for better accuracy. In the first 1200 training steps, the loss function decreased rapidly. After 6000 training steps, the value of the loss function gradually stabilized and finally less than 0.09, mAP is as high as 98.1%. The average loss and mAP of YOLOv3 as shown in Fig. 7, the same 12,000 training steps are used to analyze the training process, the final loss value is less than 0.03, while mAP is as high as 98.3%, the overall accuracy is quite close, the difference between mAP is only

0.02%. Although YOLOv3-Tiny has a shallow backbone network, still its performance seems to be quite good. The average frame number of the YOLOv3 model is found considerably higher than that of the YOLOv3-Tiny model.

Table 1 lists the test results of six behaviors based on YOLOv3-Tiny and YOLOv3. The above test shows that although the overall accuracy of YOLOv3 is slightly higher than that of YOLOv3-Tiny, YOLOv3-Tiny still gives a good performance, the accuracy of the dataset detection is as high as 97% or more. The sleep category is as high as 99% and searching the trash is the lowest with 97.7%. We implement the recognition system on YOLOv3-Tiny so it can process in time and just only deliver the behavior information. So, it will be more efficient than delivers the image to the website for YOLOv3 to processing.

The final detection result for the various cat behaviour is shown in Fig. 8. When going to the toilet for more than 30 seconds and the cat search on a trash can, a message will be sent to the owner's mobile phone, as shown in Fig. 9. The owner can remotely view the cat's current behaviour through a mobile phone or computer, as shown in Fig. 10.
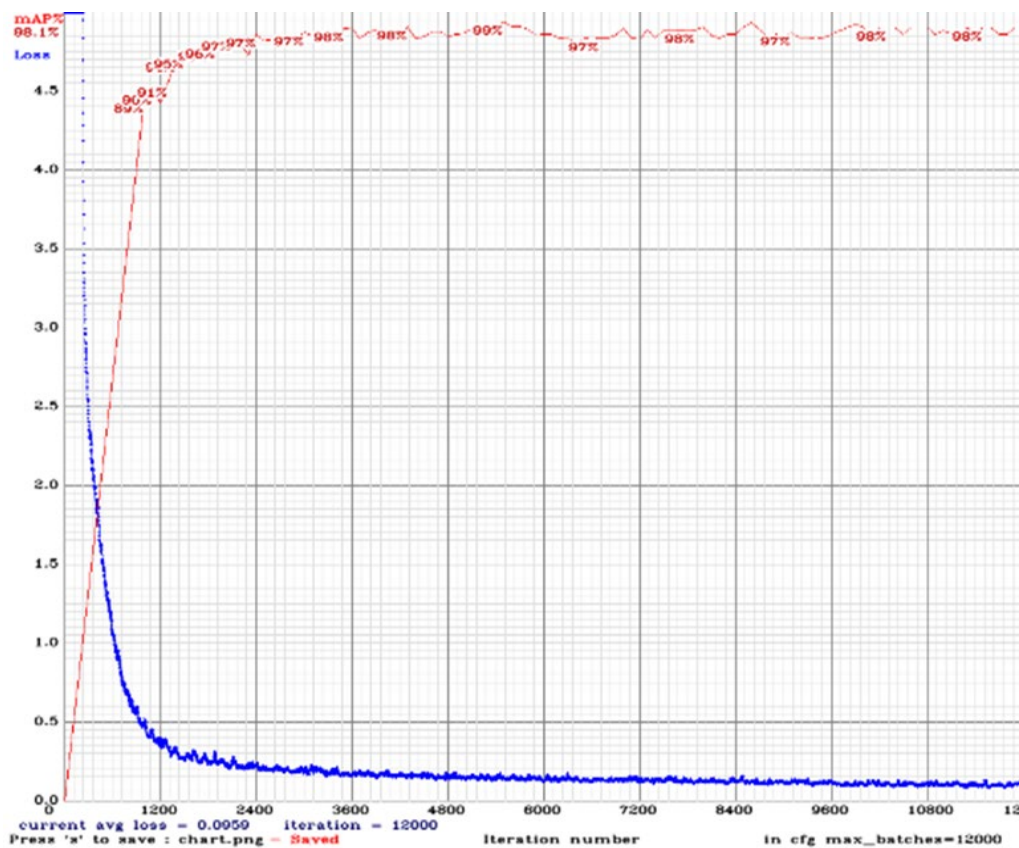

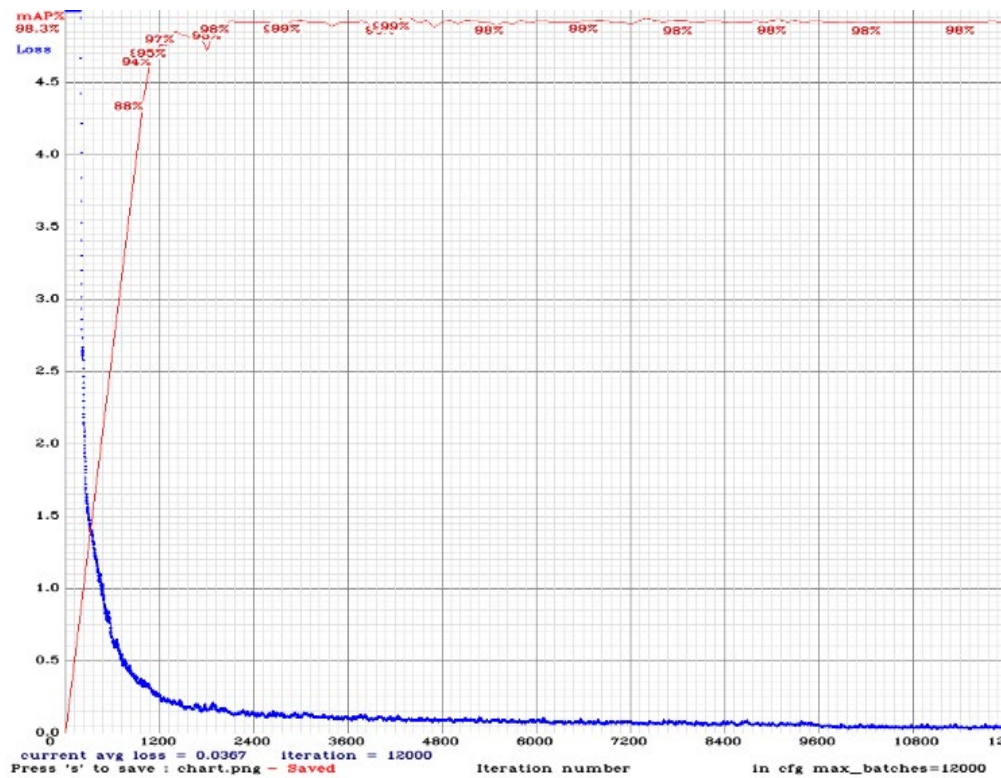**Fig. 6.** Average loss and mAP of YOLOv3-Tiny model

Fig. 7. Average loss and mAP of YOLOv3 model

Table 1. The detection accuracy of testing data

Category	Training Data	Testing Data	Accuracy of YOLOv3-Tiny	Accuracy of YOLOv3
Moving	172	74	98.78%	98.81%
Eating	126	54	98.33%	98.36%
Sleeping	162	70	99.57%	99.57%
Sitting	140	70	97.81%	98.02%
Search on a trash can	182	78	96.51%	97.69%
go to the toilet	196	86	98.24%	98.56%

(a)

(b)

(c)
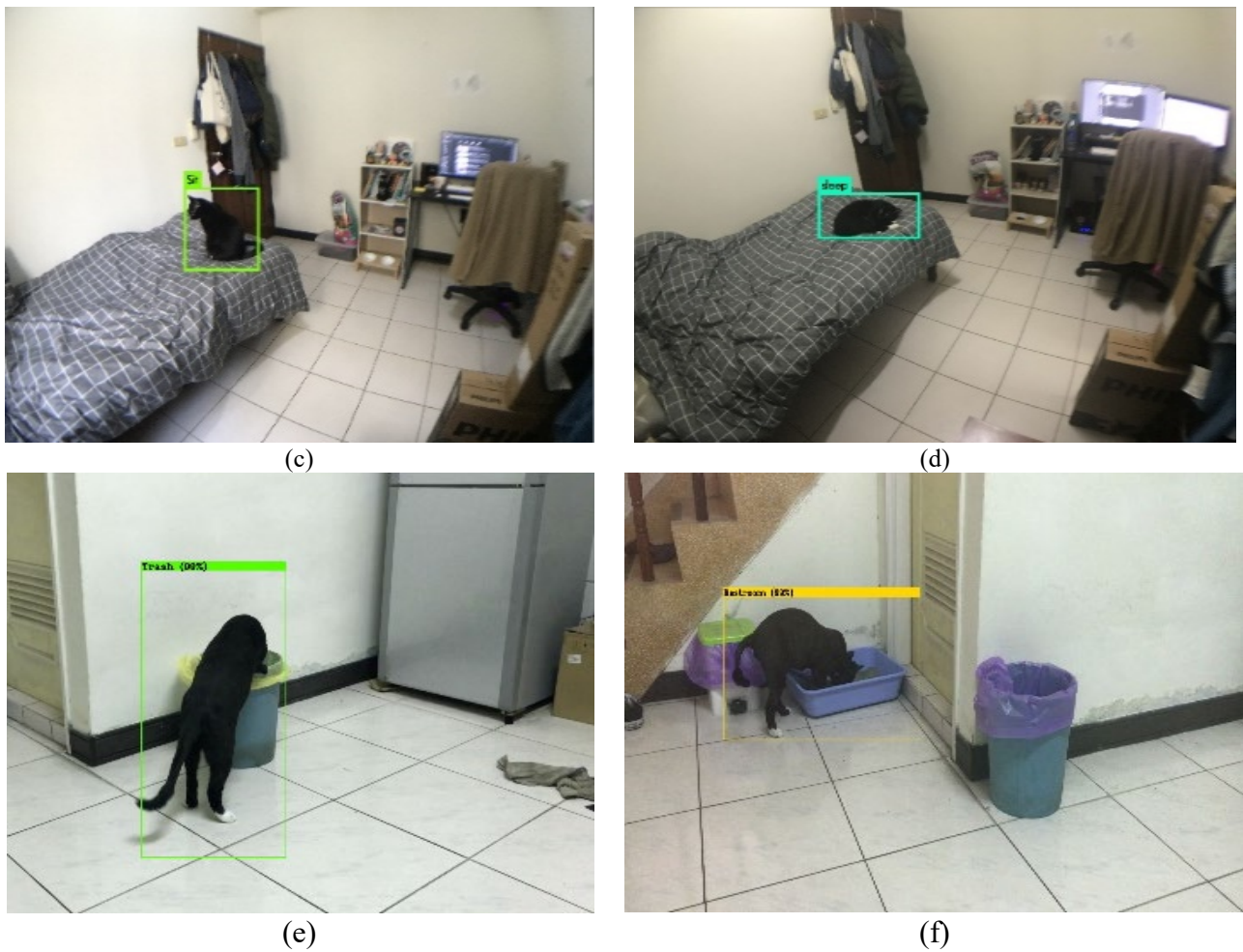

(d)

(e)

(f)
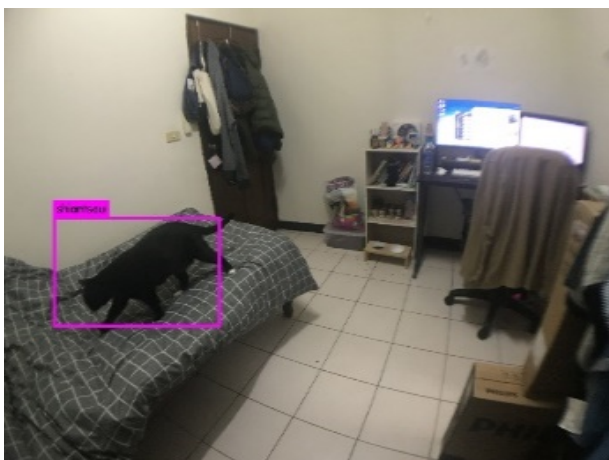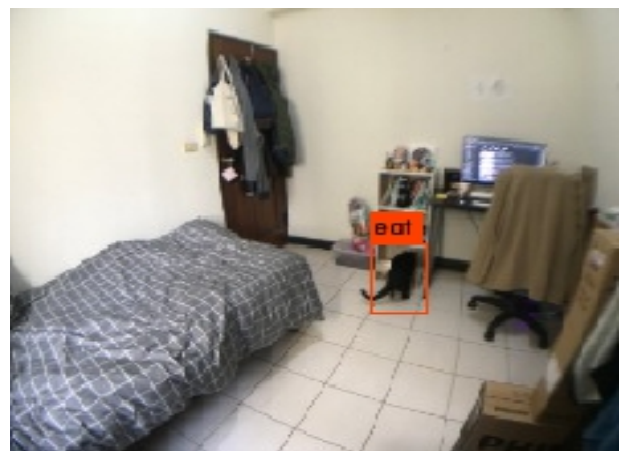
Fig. 8. Detection results: (a) moving, (b) eating, (c) sleeping, (d) sitting, (e) search on a trash can, (f) go to the toilet

(a)

(b)

Fig. 9. Detection behavior is displayed on the mobile: (a) cat searching on a trash, (b) cat go to the toilet for more than 30 seconds
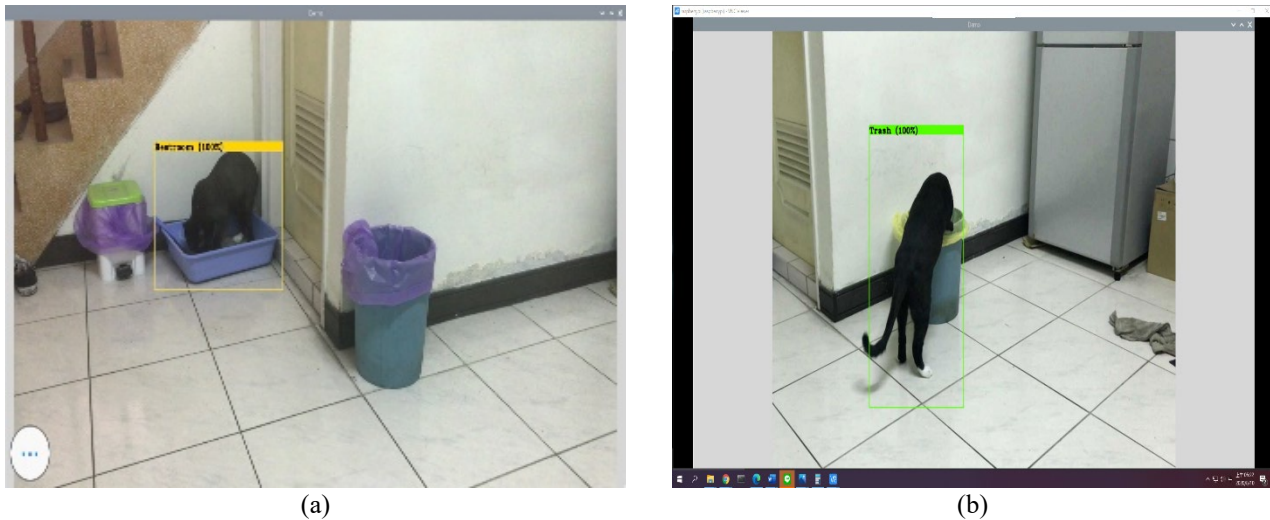
(a)   (b)

Fig. 10. View on device: (a)view on mobile, (b)view on computer

5. CONCLUSION AND FUTURE WORK

Nowadays, many camera-based pet monitors on the market are mostly video recording system which captures every movement of the pets. All these monitoring systems record the behaviour of the pets then the owners have to keep watching the video in the frequent interval of time to check the safety of the pets. This continuous monitoring causes inconvenience to the pet owner as they have to keep checking the monitoring screen. The pets are likely to search on the trash can during this period or bite something cause life danger while the owner leaves the screen. Therefore, this research uses the problems encountered in raising cats to design a pet cat behaviour detection system based on the YOLO model. The system correctly detects the cat behaviours such as walking, sitting, sleeping, eating, going to the toilet, or search on the trash can. If the pet goes toilet for more than 30 seconds or flips the trash can, the system will send the detected behaviour and photos to the mobile phone through IFTTT. The system also uses VNC remote Monitoring, which allows the owner to remotely monitor the cat's current behaviour through a mobile phone or computer. In the experiment, compared YOLOv3-Tiny with YOLOv3. Although YOLOv3-Tiny is not very good in detection speed, its accuracy is still high compared to YOLOv3. The test results show that the accuracy rate of the six categories is as high as 98.2%.

In the future, there will only be more and more families raising pets. To apply the system to each pet, we hope to collect more kinds of pets and behaviour categories, such as licking, playing, tail flicking, etc., to enrich the data set and improve the usability and practicability of the experiment. The pet monitoring system will be combined with object monitoring technology to identify the kind of objects the pets are biting or picking up. Further, it is difficult to improve the detection speed due to the limitations of the Raspberry Pi hardware. So we would use different hardware for better detection performance.

ACKNOWLEDGMENT

REFERENCES

Al Rasyid, M.U.H., Saputra, F.A., Prasetiyo, A. 2018. I-ON smart controller: Portable smart home solution based on arduino and raspberry Pi, 2018 International Conference on Applied Science and Technology (iCAST), 161–164.

APPA. Available online: https://www.americanpetproducts.org/(accessed on 06/14/2021)

Bbox label tool, 2021. Available online: https://github.com/puzzledqs/BBox-Label-Tool/ (accessed on 06/14/2021).

Borwarnginn, P., Thongkanchorn, K., Kanchanapreechakorn, S., Kusakunniran, W. 2019. Breakthrough conventional based approach for dog breed classification using CNN with transfer learning, 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE), 1–5.

Botticelli, M., Ciabattoni, L., Ferracuti, F., Monteriù, A., Pizzuti, S., Romano, S. 2018. A smart home services demonstration: Monitoring, control and security services offered to the user, 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), 1–4.

Chakraborty, S., Singh, S.K., Kumar, K. 2020. Facial biometric system for recognition using extended LGHP algorithm on raspberry Pi, in IEEE Sensors Journal, 20, 8117–8127, doi: 10.1109/JSEN.2020.2979907.

Chen, R.-C., Dewi, C., Zhang, W.-W., Liu, J.-M., Huang, S.-W. 2020. Integrating gesture control board and image recognition for gesture recognition based on deep learning, International Journal of Applied Science and Engineering, 17, 237–248.

Demir, H.S., Christen, J.B., Ozev, S. 2020. Energy-efficient image recognition system for marine life, in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 39, 3458–3466, doi: 10.1109/TCAD.2020.3012745.

Girshick, R., Donahue, J., Darrell, T., Malik, J. 2013. Rich feature hierarchies for accurate object detection and semantic segmentation, arXiv preprint arXiv:1311.2524.

Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B. 2018. Towards on-farm pig face recognition using convolutional neural networks, Computers in Industry, 98, 145–152.

He, K., Zhang, X., Ren, S., Sun, J. 2015. Deep residual learning for image recognition, arXiv preprint arXiv:1512.03385.

HomeKit. Available Online: https://www.apple.com/tw/shop/accessories/all accessories/homekit/(accessed on 06/14/2021).

Hongwen, Y., Zhenyu, L., Qingliang, C., Zhiwei, H., Wen, L.Y. 2019. Detection of facial gestures of group pigs based on improved Tiny-YOLO, Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 169–179.

IFTTT. Available Online: http://www.ifttt.com/ (accessed on 06/14/2021).

Khatri, S., Rajput, A., Alhat, S., Gursal, V., Prof. Deshmukh, J. 2020. Image-based animal detection and breed identification using neural networks, Journal of Science and Technology, 05, 130–134.

Liang, C., Xiong, J., Zheng, Z., Zhong, Z., Li, Z., Chen, S., Yang, Z. 2020. A visual detection method for nighttime litchi fruits and fruiting stems, Computer and Electronics in Agriculture, 167, 105192.

Nadafa, R.A., Hatturea, S.M., Bonalam, V.M., Naikb, S.P. 2020. Home security against human intrusion using raspberry Pi, Procedia Computer Science, 167, 1811–1820.

Nguyen, H., Maclagan, S.J., Nguyen, T.D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E.G., Phung, D. 2017. Animal recognition and Identification with deep convolutional neural networks for automated wildlife monitoring, 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 40–49.

Redmon, J., Farhadi, A. 2016. Yolo9000: Better, faster, stronger, arXiv preprint arXiv:1612.08242.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A. 2015. you only look once: Unified, real-time object detection, arXiv preprint arXiv:1506.02640.

Taipei city animal protection office. Available online: https://www.tcapo.gov.taipei/(accessed on 06/14/2021).

Taiwan agricultural commission. Available online: https://animal.coa.gov.tw/(accessed on 06/14/2021).

Trends Insight Available online: http://www.trendsightinc.com/insight/(accessed on 06/14/2021).

Viola, P., Jones, M. 2001. Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, 1, I–I.

Virtual network computing. Available online: http://www.realvnc.com/ (accessed on 06/14/2021).

World population review. Available online: https://worldpopulationreview.com/(accessed on 06/14/2021).

Wu, F., Jin, G., Gao, M., He, Z., Yang, Y. 2019. Helmet detection based on improved YOLO V3 deep model, in 2019 IEEE 16th International Conference on Networking, Sensing and Control, 363–368.

Wu, F., Sinnott, R.O., Chen, W. 2018. A mobile application for dog breed detection and recognition based on deep learning, 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT), 87–96.

Yang, J., Qian, T., Zhang, F., Khan, S.U. 2021. Real-time facial expression recognition based on edge computing, in IEEE Access, 9, 76178–76190. doi: 10.1109/ACCESS.2021.3082641.