УДК 004.93

ОБНАРУЖЕНИЕ АВТОМОБИЛЬНЫХ ПАРКОВОЧНЫХ МЕСТ НА ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ МОДИФИЦИРОВАННОЙ МОДЕЛИ YOLOv5 С ПОЛУКОНТРОЛИРУЕМЫМ ОБУЧЕНИЕМ

ЧЖАН ШУАЙ[1), 2)]***, ГУАНГДИ МА***[3)]***, ЯН ВЕЙЧЕН***[3)]***, ФАН ЖУ***[4)]***, С. В. АБЛАМЕЙКО***[2), 5)]

[1)]*Лоян скорпио информейшн технолоджи, 471000, г. Лоян, пров. Хэнань, Китай*
[2)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*
[3)]*Эрсвью имэйдж корпорейшн, ул. Кейан роуд, 11, 313200, г. Хучжоу, Китай*
[4)]*Хэнаньский университет, ул. Минлун, 85, 475004, г. Кайфэн, Китай*
[5)]*Объединенный институт проблем информатики НАН Беларуси, ул. Сурганова, 6, 220012, г. Минск, Беларусь*

Образец цитирования:
Чжан Шуай, Гуангди Ма, Ян Вейчен, Фан Жу, Абламейко СВ. Обнаружение автомобильных парковочных мест на изображениях с использованием модифицированной модели YOLOv5 с полуконтролируемым обучением. *Журнал Белорусского государственного университета. Математика. Информатика.* 2023;3:72–81 (на англ.).
EDN: XVDRSN

For citation:
Zhang Shuai, Guangdi Ma, Yang Weichen, Fang Zuo, Ablameyko SV. Car parking detection in images by using a semi-supervised modified YOLOv5 model. *Journal of the Belarusian State University. Mathematics and Informatics.* 2023;3:72–81.
EDN: XVDRSN

Авторы:
Чжан Шуай – менеджер по исследованиям и разработкам[1)], аспирант кафедры веб-технологий и компьютерного моделирования механико-математического факультета[2)]. Научный руководитель – С. В. Абламейко.
Гуангди Ма – главный инженер.
Ян Вейчен – главный управляющий.
Фан Жу – директор Хэнаньской международной объединенной лаборатории теорий и ключевых технологий разведывательных сетей.
Сергей Владимирович Абламейко – доктор технических наук, академик НАН Беларуси, профессор; профессор кафедры веб-технологий и компьютерного моделирования механико-математического факультета[2)], главный научный сотрудник отдела интеллектуальных информационных систем[5)].

Authors:
Zhang Shuai, research and development manager[a], postgraduate student at the department of web-technologies and computer modelling, faculty of mechanics and mathematics[b].
coulturing@gmail.com
Guangdi Ma, chief engineer.
magd@ev-image.com
Yang Weichen, general manager.
yangwch@ev-image.com
Fang Zuo, director of Henan International Joint Laboratory of Theories and Key Technologies on Intelligence Networks.
zuofang@henu.edu.cn
Sergey V. Ablameyko, doctor of science (engineering), academician of the National Academy of Sciences of Belarus, full professor; professor at the department of web-technologies and computer simulation, faculty of mechanics and mathematics[b], and chief researcher at the department of intelligent information systems[c].
ablameyko@bsu.by
https://orcid.org/0000-0001-9404-1206

Прямоугольные, непрерывные парковочные места довольно сложно идентифицировать на любых изображениях городской территории при различных погодных условиях, низкой освещенности и низкой стоимости системы, обеспечивая при этом высокую точность обнаружения. Для решения этой проблемы предлагается использовать модифицированную версию модели YOLOv5, дополненную полуконтролируемым обучением (полуавтоматическим обучением или частичным обучением), которая позволяет обнаруживать парковки в любой сложной сцене независимо от линий парковочных мест и условий парковки. Благодаря сочетанию характера полуконтролируемого обучения и высокой точности моделей обучения с учителем модифицированная версия модели YOLOv5 позволяет использовать очень мало размеченных данных и большой объем неразмеченных данных. Это значительно сокращает время обучения, сохраняя при этом точность распознавания. По сравнению с другими моделями нейронных сетей модифицированная версия модели YOLOv5 обладает такими характеристиками, как высокая скорость обучения, небольшой размер модели и данных, а также высокая точность параметров распознавания.

Ключевые слова: обнаружение парковок; полуконтролируемое обучение; нейронная сеть YOLOv5.

CAR PARKING DETECTION IN IMAGES BY USING A SEMI-SUPERVISED MODIFIED YOLOv5 MODEL

ZHANG SHUAI[a, b]***, GUANGDI MA***[c]***,***
YANG WEICHEN[c]***, FANG ZUO***[d]***, S. V. ABLAMEYKO***[b, e]

[a]*Luoyang Scorpio Information Technology Ltd., Luoyang 471000, Henan, China*
[b]*Belarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus*
[c]*EarthView Image Inc., 11 Keyuan Road, Huzhou 313200, China*
[d]*Henan University, 85 Minglun Street, Kaifeng 475004, China*
[e]*United Institute of Informatics Problems, National Academy of Sciences of Belarus, 6 Surganava Street, Minsk 220012, Belarus*

Corresponding author: S. V. Ablameyko (ablameyko@bsu.by)

The problem of car parking detection in images attracts the attention of many researchers. In this task, it is quite difficult to identify rectangular, continuous parking spaces in all kinds of city images under different weather conditions, combining the low-light environment and the system's low cost with high detection accuracy. In this paper, we propose a modified version of the YOLOv5 model joined with semi-supervised learning that allows us to detect parking lots in any complex scene, independent of parking space lines and parking environments. Due to the combination of the nature of semi-supervised learning and the high accuracy of supervised learning models, the modified version of YOLOv5 model permits to use very little labeled data and a large amount of unlabeled data. It can significantly reduce training time while maintaining recognition accuracy. Compared with other neural network models, the modified version of YOLOv5 model has the characteristics of fast training speed, persistent operation, small model size, and high model precision and recall values.

Keywords: car parking detection; semi-supervised learning; YOLOv5 neural network.

Introduction

With the increase of the number of cars in cities, the task to make effective use of parking lots, reduce the time people spend in the parking process has become more and more important in smart cities. Over the last decade numerous intelligent systems have been developed for parking space detection with different lighting conditions and restrictions [1].

In this task, it is quite difficult to identify rectangular, continuous parking spaces with high accuracy using the detection of parking monitoring images, and it is not possible to effectively identify irregularly shaped parking lots and irregularly arranged parking lot vehicles. However, there is a need to realise a low-cost parking vehicle detection system which can be applied in all kinds of complex parking spaces, different weather conditions, combining the low-light environment and the low cost with high detection accuracy [2].

In terms of vehicle detection, there are two common approaches: the conventional computer vision approach and image analysis based on convolutional neural networks. The conventional image analysis approach is a vehicle detection method based on image feature analysis and parking space localisation [3]. The core question is what kind of features are extracted to detect a vehicle and what kind of classification tools are applied [4]. Parking spaces localisation allows us to locate a parking space on a video frame (segmentation) and record these positions. Input image perspective transformation is used to facilitate the possibility of describing the parking space as being rectangle and of easier detecting the separating parking lines. The parking spaces classification subsystem solves the following problems: co-ordinates of the parking space extraction on the video frame based on a given location for further processing the extracted images area (regions of interest), which corresponds to the parking spaces; calculation of the features of the parking space, generation of the features vector describing this state, and use of the features classification to separate the vacant and the occupied parking spaces for every region of interest; visual demonstration of the occupied parking spaces [5].

In parking space vehicle detection, object detection methods based on neural network models have recently started to be used. The most important part of this solution is a neural network model used to detect vehicles. The existing methods of identifying parking spaces through neural networks often use multi-layer neural networks. After a specific parking space is segmented, neural networks are used to detect whether there are vehicles in the parking space, such as Fast R-CNN, Faster R-CNN, Mask R-CNN. This is more difficult to use in an irregularly laid out parking lot, especially if the parking space markings are not clear. Another approach is to use the single-layer neural network and use the first-level target detector neural network model to directly return the target across the target candidate area without the need to segment the target. In this case, the detection speed is fast, and it has become the mainstream model category for vehicle identification in parking lots, such as SSD, YOLO series [6]. However, the accuracy of identifying object positions in this approach is not sufficient, and the recall rate is quite low.

Having this short analysis, one can see that typically neural network models focus on accuracy, detection speed, and model architecture complexity. However, the cost incurred by the large datasets required to train the model is also significant. How to train a neural network model suitable for various parking lots with less datasets? It has become an important problem.

In this paper, we propose a modified version of the YOLOv5 (you only look once version 5) model joined with semi-supervised learning that allows us to detect parking lots in any complex scene, independent of parking space lines and parking environments. Due to the combination of the nature of semi-supervised learning and the high accuracy of supervised learning models, the model allows us to use very little labeled data and a large amount of unlabeled data. It can significantly reduce training time while maintaining recognition accuracy.

Overview of neural network approaches for car parking detection

Let us consider some recent approaches for car detection in parking lots based on neural networks. D. Acharya, et al., used a deep CNN and a binary support vector machine classifier to detect spaces in outdoor parking lot occupancy images, and determined parking space occupancy from the images obtained from surveillance [7]. J. Nyambal, et al., used the Caffe and Nvidia DiGITS frameworks for predictive detection of vacant and occupied parking spaces [8].

A. Naufal, et al., proposed a preprocessed region-based convolutional neural network (Mask R-CNN) to mark the parking position on the input image of a full parking lot [9]. At the first stage, preprocessing is performed that combines contrast enhancement using the exposure fusion framework. At the second stage, each parking position is examined whether the position is vacant or not, using mAlexNet. A series of trials on images with varying light conditions indicate that the preprocessed Mask R-CNN can improve marking the parking positions with an accuracy of intersection over union reaching 85.80 %. The result of marking the parking position is then used in the trial of the availability of the parking space on video data using mAlexNet, and an accuracy attains 73.73 %.

The use of the Mask R-CNN method for the detection of the parking space has been proposed by J. Ahmad, et al. [10]. The bounding box from the detection is compared with the manually annotated bounding box to determine the classification of the parking space status. This research scenario is still limited because it is necessary to initialise manual annotations for each parking space. This research can detect parking spots correctly, with an accuracy of about 90 %. Missclassification occurs because of camera angles.

T. Agrawal and S. Urolagin used Mask R-CNN to solve the multi-angle parking problem [11]. This research proves that the Mask R-CNN model can run well in low angle closed circuit television areas with the highest accuracy value of 86 %. The model has tended to adapt to various parking scenarios. However, a problem found in this research still has issues related to lighting, and car detection failure occurs if there are other objects in front of it (such as trees or light poles, or even other cars).

Recently YOLO CNN started to be widely used for this task. Yucheng Guo and Hongtao Shi proposed the improved convolution neural network algorithm and image recognition technology to identify and locate parking spaces [12]. Then, the vehicle license plate information is identified using the improved YOLO model, the parking system model is established, and the parking selection algorithm is used.

Y. Miao, et al., proposed an effective nighttime vehicle detection approach [13]. First of all, the original nighttime images were enhanced by an optimal multi-scale retinex algorithm. Then, a pretrained YOLOv3 network was used and fine-tuned by the enhanced images. Finally, the detection network was used to detect vehicles from the nighttime images and outperformed two widely used object detection methods, namely the Faster R-CNN and SSD, on the precision and detection efficiency. The average precision of the proposed method reaches 93.66 %, which is 6.14 and 3.21 % higher than that of the Faster R-CNN and SSD, respectively.

Based on YOLOv3, Wang Hai, et al., combined visual images and lidar to improve detection accuracy and real-time performance [14]. Firstly, the obstacles are detected by the grid projection method using the lidar point cloud information. Then, the obstacles are mapped to the image to get several separated regions of interest (ROIs). After that, the ROIs are expanded based on the dynamic threshold and merged to generate the final ROI. Finally, a network YOLO is applied on the ROI to detect vehicles. The experimental results on the KITTI dataset demonstrate that the proposed algorithm has high detection accuracy and good real-time performance. Compared with the detection method based only on the YOLO, the mean average precision (mAP) is increased by 17 %.

D. Carrasco, et al., used YOLOv5 and a multi-scale mechanism to learn deep discriminative feature representations of different scales and automatically determined the scale most suitable for detecting objects in the scene, reducing the number of trainable parameters [15].

H. Bura, et al., proposed an edge based smart parking solution using camera networks and deep learning [16]. The object detection and tracking were performed using the YOLO model. On the contrary, the parking lot occupancy detection was carried out using the custom AlexNet: 1 input layer, 1 convolution layer, 1 ReLu, 1 max poolong, and 3 fully connected layers. The dataset was trained and evaluated using 150 000 images and could achieve 99.51 % better than the model called mini AlexNet-10 and outperformed as well in terms of inference time in order to perform in real time with the result of 7.11 ms.

M. Uzar, et al., analysed YOLO architectures in terms of performance assessments of vehicle detection in parking lots [17]. The labeling process was performed for three classes (car, bus, and minibus) using the Visual Object Tagging Tool. The labeled dataset was trained via transfer learning in YOLOv4-CSP, YOLOv4-tiny, YOLOv4-P5, YOLOv4-P6, YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x architectures. The weights of YOLO versions were implemented to the parking lots and the results were compared. To assess the performance of YOLO-based vehicle detection, mAP and F_1-score values were computed.

According to the results, when analysing the dataset with a limited graphics processing unit support, it is seen that large-scale models could not be trained properly. Thus, to determine the real performances of YOLOv4-P5, YOLOv4-P6, YOLOv5l and YOLOv5x models, it is recommended to train models with an unlimited graphics processing unit support and more training epochs. Moreover, the number and diversity of the dataset should be increased with the use of high-capacity processors.

Based on this analysis, we have chosen the YOLOv5 model that is the most popular detection model: it is fast enough to train, easy to deploy and expand, and has a small project code.

Materials and methods

YOLOv5 and supervised learning. Our model is based on YOLOv5, combined with the semi-supervised learning concept, to reduce the model size and the training time while ensuring the model's detection accuracy.

YOLOv5 is a highly popular single-layer neural network model in computer vision. YOLOv5 is known for its exceptional speed, top-tier performance, and user-friendly interface. These features make it ideal for real-time object detection tasks, such as autonomous driving, surveillance, and robotics. While YOLOv5 maintains the YOLO model's core architecture, it introduces a groundbreaking PyTorch-based training and deployment framework. This integration simplifies the model development and customisation, making it accessible to researchers and developers. YOLOv5 is a dynamic advancement in object detection technology. The network structure of the model is shown in fig. 1.

In the network structure diagram of YOLOv5, it can be seen that it is divided into four parts: input, backbone, neck, and prediction [18].

Semi-supervised learning is a machine learning algorithm used in the training data mixed with labeled data plus unlabeled data [19]. It combines a small amount of labeled data with a large amount of unlabeled data during training. Semi-supervised learning algorithms can be divided into groups: self-training, graph-based semi-supervised learning, semi-supervised support vector machine.

The self-training classifier is the most common self-training tool in semi-supervised learning [20]. The simple self-training process is to train a classifier with labeled data, and then use this classifier to classify unlabeled data, which will produce a pseudo-label [21] or soft label. The selected unlabeled samples are used to train the classifier. Self-supervised learning only needs a small number of low-cost characteristics of labeled datasets. It can achieve a highly scalable, highly robust vehicle identification model.

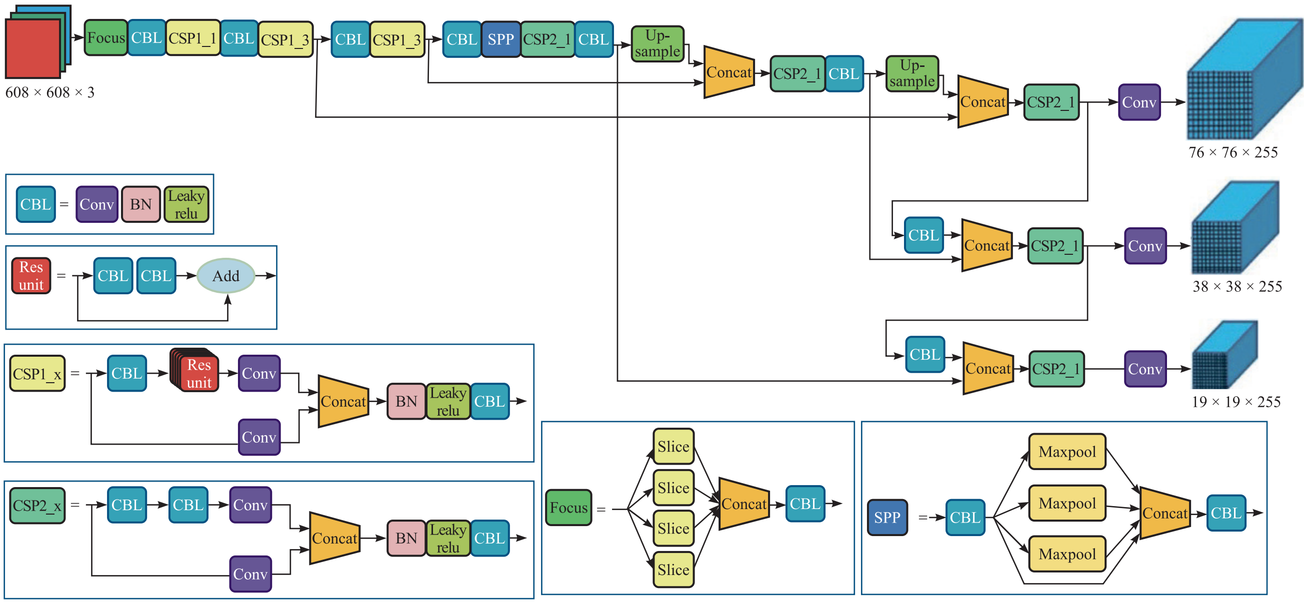

Fig. 1. The YOLOv5 network structure.
Source: [18]

Model design. The current mainstream approach using semi-supervised learning is to use the concept of semi-supervised learning models to train supervised learning models. Our idea is to combine the self-training idea, YOLOv5l model features, and parking lot usage scenarios to design a self-training learning method that adds a supervised learning model and a semi-supervised learning model using pseudo-labels.

We insert a full YOLOv5 training epoch in each loop of the self-trained model. In each YOLOv5 training cycle, we use the labeled data to train the YOLOv5 pre-trained model or the object detection model generated in the previous cycle. Then, we detect unlabeled data and generate pseudo-labels and use the generated pseudo-label loss data. The loss data modify the detection model produced by YOLOv5 training. Looping in turn, the system generates more pseudo-label and loss data, reduces unlabeled data, and generates object detection model files with higher accuracy. Figure 2 describes the basic process, which is as follows:

- use a small amount of labeled data to train a detection model;
- use the detection model to identify unlabeled data and obtain results;
- select high-quality detection results as pseudo-labels;
- use pseudo-labels and labeled data.

A new detection model is trained, recursively in turn, until it is manually stopped or the unlabeled data all become labeled data (see fig. 2).

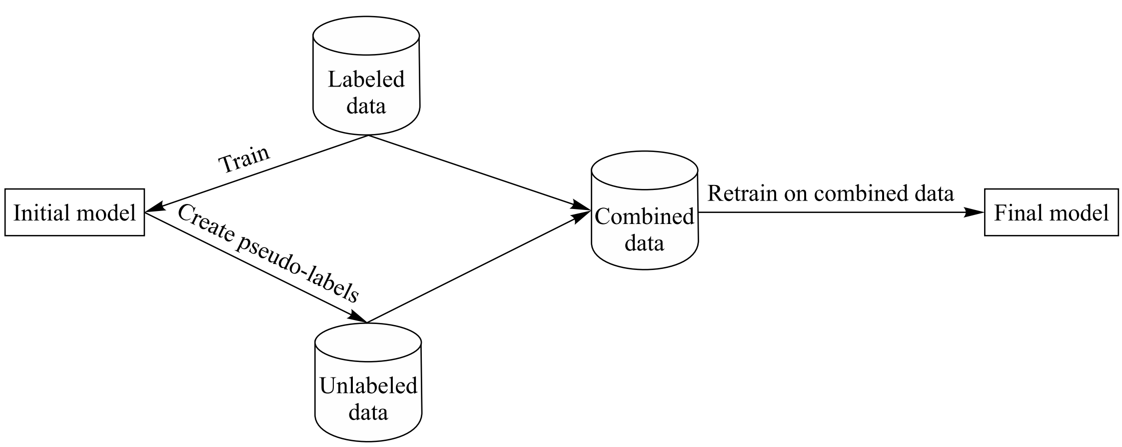

Fig. 2. The supervised learning model plus semi-supervised learning using pseudo-labels

Pseudo-labeling is the SSL method that helps with untagged data [22]. First, we can annotate a small batch of data and train a model with it, and then use this model to detect objects on unlabeled data. The labels annotated by the model are called pseudo-labels. We can then combine the labeled and pseudo-labeled data and train the model again. The basic workflow is:

1) train the original model M with a small amount of the labeled dataset D;
2) identify unlabeled datasets using the M model;
3) filter out the detection results of precision >90 % and recall >90 %, and construct a new pseudo-label dataset D';
4) train a new M model with new labeled data and keep iterating.

Since in the parking environment, monitoring images can be continuously acquired as unlabeled data, the continuous operation of the neural network model can be realised until human intervention stops or the model data is infinitely close to 1. Because it is too time-consuming to complete the training in actual training, especially after the unlabeled dataset is gradually converted into a labeled dataset, a large number of training datasets will be generated, which will gradually increase the single training time and reduce the training efficiency.

Modifying the new pseudo-label dataset D' directly modifies the loss function of the model, that is, the loss value of the existing model is superimposed on the loss value of the prediction dataset D': loss = loss(labeled data) + alpha*loss(unlabeled_data) [23]. The specific flow chart and analysis are shown in fig. 3, which includes the following stages:
1) train the supervised model M using labeled data;
2) the supervised model M is used to predict unlabeled data, resulting in the predicted probabilities P and pseudo-labels;
3) update the loss calculation method of the model: use loss(labeled_data) + alpha*loss(unlabeled_data) as the loss value of the model;
4) train a new model M' using the labeled data and the pseudo-labeled data filtered according to the predicted probability P (the pseudo-label filtering method is step 3 in the self-training plus YOLOv5 model design);
5) replace M' with M and repeat the above steps until the model effect does not improve.

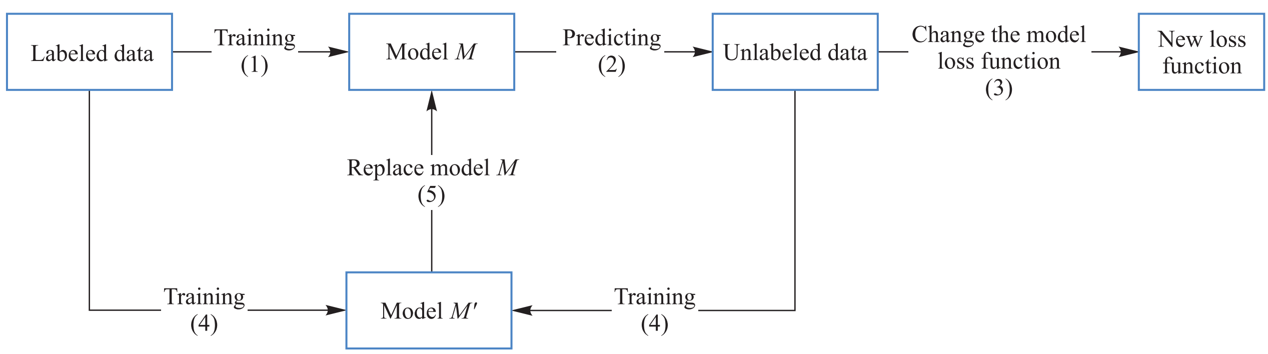

Fig. 3. The improved model of the training process

Due to the characteristics used in the production environment of a parking lot, the keyframes can be extracted regularly in the parking lot surveillance video. The unlabeled data can be added as a picture, and the results will be returned to the system after the detection is completed. The model continues to learn.

The advantages of the proposed model are:
• the possibility to use a small amount of labeled data to achieve similar training results for the same YOLOv5 model. It allows us to reduce the cost of the model dataset construction;
• thanks to the learning method of sustainable words, it can be persistently operated on the object detection settings in the parking lot;
• there is no need to consider that the dataset of vehicles is too small and difficult to train.

Datasets and training process

The database based on the CNRPark[1] open source parking lot dataset, which uses more than 200 photos of Italy's parking lots, consisting of 9 cameras, taken every 30 min, since 16 November 2015. All-weather parking information for the period 9 December 2015, including images of parking lots in different climates such as daytime, night, sunny, cloudy, rainy, etc., with a variety of environmental elements such as parking spaces, roads, sidewalks, trees, and green spaces to train vehicle identification in the parking lot. Using the filtered datasets «vehicles open images», we filter out top views, test maps, and have common vehicle types such as motorcycles and trucks to train multiple vehicle type identification. Ultimately, the dataset for this project contains 276 training data, 32 verification data, parking lot plus photos taken by drones.

[1]CNRPark + EXT is a dataset for visual occupancy detection of parking lots / G. Amato [et al.] // CNR Parking Dataset : website. July 2015. URL: http://cnrpark.it/ (date of access: 14.04.2022).

In our study, we employed a dedicated validation dataset to evaluate the performance and accuracy of the model. This validation dataset is created by randomly sampling data from our dataset, making sure it does not overlap with the training data. The validation dataset accounts for 1 % of the total dataset. The validation process involves evaluating whether the model can correctly detect objects in the validation dataset. This approach is taken to ensure that the performance of the model is evaluated on the unseen data, thereby preventing over-reliance on the training dataset.

The training process follows the training cycle of semi-supervised learning, and the training method of YOLOv5 is used in each training cycle.

Each training cycle follows the training, generates a pre-training model, identifies unlabeled data, generates pseudo-labels, and uses pseudo-labeled data to modify model parameters. It is divided into two steps in a single training cycle [24]:

1) pre-training. We use ImageNet data to train the first 20 convolutional layers of the YOLOv5 network plus 1 average pooling layer plus 1 fully connected layer. We train the image resolution resize to 224 × 224;

2) initialisation. The network parameters of the first 20 convolutional layers of the YOLOv5 model are initialised with the first 20 convolutional layers of the VOC model, and then the VOC-20 class annotation data is used for YOLOv5 model training. To improve the image accuracy, when training the detection model, resize the input image resolution to 448 × 448.

Such a training mode has the characteristics of a quick, low background false detection rate, strong versatility.

Experimental results

In our experiments, we used the database CNRPark. We analysed images taken from stationary cameras and from drones. Images were taken in day time, evening and night time, and in different weather conditions: sunny and rainy.

We set three classification categories: car, motorbike, traffic cone. The definition of precision >0.7 and recall >0.8 is considered a hit target. Precision, recall, mAP are mainly used as indicators of detection accuracy, detection hit rate, and detection ability.

Training model is represented as follows: YOLOv5 plus self-training model.

Dataset includes 18 images, selected from the standard dataset as labeled datasets, and the remaining images, used as unlabeled datasets.

Number of training is limited to 400 times.

The experiment results for car detection are shown in table 1. Semi-supervised learning is affected by dataset changes and the model is affected by pseudo-labels generated by the previous detection. In order to ensure the stability of the experimental result data, the experimental results are obtained by taking the average after three experiments. The precision – recall (PR) curve and F_1 curve of the training process are shown in fig. 4.

Table 1

Experimental results for car detection

Experiment	Precision	Recall	mAP	F_1	Labeled images	Train time cost
1	0.999	0.941	0.601	0.969	18	20 min 29 s
2	0.998	0.941	0.595	0.970	18	20 min 20 s
3	0.999	0.941	0.607	0.971	18	20 min 39 s

Figure 5 shows the car parking detection results: the image from a drone on a sunny day (see fig. 5, *a*) and from a stationary camera in different conditions: evening (see fig. 5, *b*), night (see fig. 5, *c*), and rain (see fig. 5, *d*). Due to the influence of scene and environment on the image brightness and contrast, the detection results are different in different environments. However, the obtained results show that the proposed models can effectively identify parking lot vehicles.

From the obtained results, it can be seen that with 18 labeled data and training time of 20 min, a better precision and recall can be obtained. Although the mAP is lower than YOLOv5, the results are still satisfactory.

It can be seen from the number of training sessions used, labeled data, and training use: only 18 labeled data are used, indicating that the model can be trained with a very small amount of data labels, and because there is less labeled data, the training speed is extremely fast. It can be seen from table 1 that the values of precision, recall, and F_1 are extremely high, particularly the precision is close to 1. Figure 4 shows that the overall curve of the model is stable, and fig. 5 shows that the model can effectively identify vehicles in the parking lot.

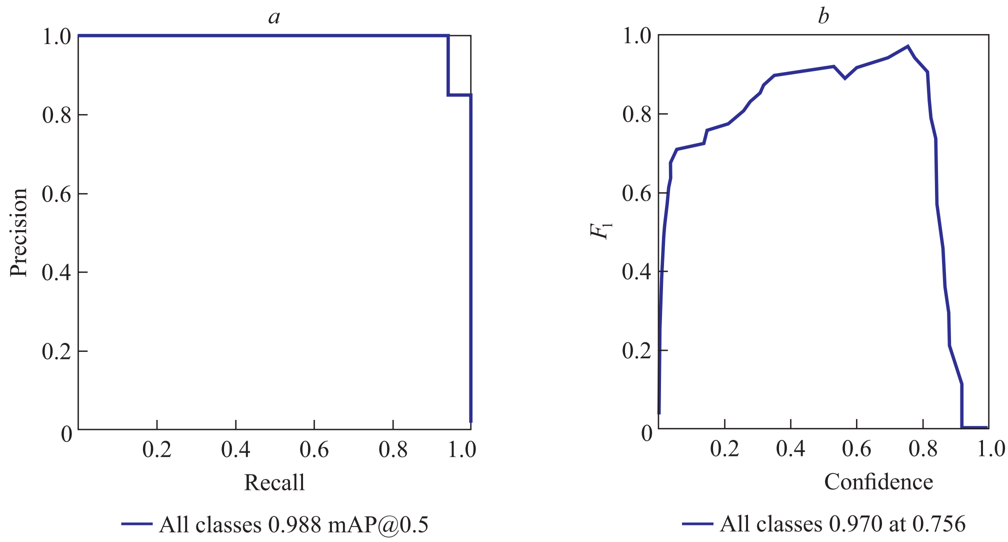

Fig. 4. The PR curve (*a*) and the F_1 curve (*b*)

Fig. 5. Car detection results on a sunny day (*a*),
in the evening (*b*), at night (*c*), on a rainy day (*d*)

From the above experiments, it can be concluded that for the model using YOLOv5 plus semi-supervised learning, the training speed can be improved by adding cache, and it is also verified that the longer the training time, the better the training accuracy and accuracy of the model. For the multi-scale parameter, since the modified model precision has tended to be close to 1 there is no need to use this parameter to promote precision and recall.

Comparison and discussion

In order to compare the characteristics of our approach with the existing ones, we tested the Fast R-CNN, YOLOv4, YOLOv5l and modified YOLOv5 models to detect cars at parking lots. We tested the models to define the speed of 93 parking spaces in 16 pictures on a low-performance device (take Raspberry Pi 4B as an example).

The results show that the training results are very accurate. We use test data (test data from the CNRPark dataset, independent of training and validation data). The test results show that it can be used with high accuracy and run on low-performance devices. The comparison results of different models are shown in table 2.

Table 2

Comparison results of different models

Model	Precision	Recall	Detection model size, Mb	Training time, min
Fast R-CNN	0.977	0.893	341	173
YOLOv4	0.988	0.916	254	134
YOLOv5l	0.983	0.927	92	118
Proposed model	0.998	0.941	31	20

The detection accuracy of the Fast R-CNN model is affected by the arrangement of parking spaces, and it is best detected only in consecutively arranged parking lots. The YOLO model is not affected by the arrangement of parking spaces, and it is more suitable for parking lots with irregular arrangements.

The precision and recall of different models are similar, and the trained precision and recall are both close to 1, indicating that they can complete the task of object detection well.

The size of the target detection files of different model training results varies greatly. The proposed model has only 18 labeled data, only 5 % of the datasets available to other models, and is modified based on the YOLOv5 model, so the target detection file size is extremely small and the inference speed is faster. It is also more suitable for continuous training with low performance equipment and high training efficiency. The detection results show that it can be well applied in low-light environments.

Conclusions

In this paper, a new modified YOLOv5 model is proposed by adding semi-supervised learning ideas to the standard YOLOv5 model. Through experimental comparison, the model can use a small amount of labeled data and a large amount of unlabeled data to train a good neural network model. Compared with other neural network models, the model has the characteristics of fast training speed, persistent operation, small model size, and high model precision and recall values. Through the deployment in the production environment, the feasibility and effectiveness of the model are fully verified.

The experimental results show that the modified YOLOv5 model is fully capable of vehicle detection in a parking lot environment, and has the ability to operate in an industrial environment. Its extremely fast training speed is also very suitable for training on low-performance equipment.

In the future, we will achieve a better target detection application technology from the six directions of lightweight object detection, object detection in combination with AutoML, domain adaptive object detection, weak supervision target detection, small object detection, and information fusion object detection. With the application of the concept of adaptive learning in the field of neural networks, attempts will be made to realise multi-machine collaborative identification, increase vehicle entry and exit behaviour detection, and improve detection accuracy.

References

1. Alam M, Moroni D, Pieri G, Tampucci M, Gomes M, Fonseca J, et al. Real-time smart parking systems integration in distributed ITS for smart cities. *Journal of Advanced Transportation.* 2018;2018:1485652. DOI: 10.1155/2018/1485652.

2. Faheem F, Mahmud SA, Khan GM, Rahman M, Zafar H. A survey of intelligent car parking system. *Journal of Applied Research and Technology.* 2013;11(5):714–726. DOI: 10.1016/S1665-6423(13)71580-3.

3. Yusnita R, Norbaya F, Basharuddin N. Intelligent parking space detection system based on image processing. *International Journal of Innovation, Management and Technology.* 2012;3(3):232–235. DOI: 10.7763/IJIMT.2012.V3.228.

4. Huang Ching-Chun, Vu Hoang Tran, Chen Yi-Ren. A multiclass boosting approach for integrating weal classifiers in parking space detection. In: *2015 IEEE International conference on consumer electronics; 2015 June 6–8; Taipei, Taiwan.* [S. l.]: IEEE; 2015. p. 314–315. DOI: 10.1109/ICCE-TW.2015.7216918.

5. Bohush R, Yarashevich P, Ablameyko S, Kalganova T. Extraction of image parking spaces in intelligent video surveillance systems. *Machine Graphics and Vision.* 2018;27(1–4):47–62. DOI: 10.22630/MGV.2018.27.1.3.

6. Wang Hai, Yu Yljie, Cai Yingfeng, Chen Xiaobo, Chen Long, Liu Qingchao. A comparative study of state-of-the-art deep learning algorithms for vehicle detection. *IEEE Intelligent Transportation Systems Magazine.* 2019;11(2):82–95. DOI: 10.1109/MITS.2019.2903518.

7. Acharya D, Yan W, Khoshelham K. Real-time image-based parking occupancy detection using deep learning. In: Peters S, Khoshelham K, editors. *Research@Locate-2018. Proceedings of the 5th annual conference; 2018 April 9–11; Adelaide, Australia*. [S. l.]: [s. n.]; 2018. p. 33–40 (CEUR workshop proceedings; volume 2087).

8. Nyambal J, Klein R. Automated parking space detection using convolutional neural networks. arXiv:2106.07228v1 [Preprint]. 2021 [cited 2023 January 10]: [6 p.]. Available from: https://arxiv.org/abs/2106.07228v1.

9. Naufal AA, Fatichah C, Suciati N. Preprocessed mask RCNN for parking space detection in smart parking systems. *International Journal of Intelligent Engineering and Systems.* 2020;13(6):255–265. DOI: 10.22266/ijies2020.1231.23.

10. Ahmad J, Lewis Z, Duraisamy P, McDonald T. Parking lot monitoring using MRCNN. In: *10th International conference on computing, communication and networking technologies (ICCCNT); 2019 July 6–8; Kanpur, India.* [S. l.]: IEEE; 2019. p. 10–13. DOI: 10.1109/ICCCNT45670.2019.8944394.

11. Agrawal T, Urolagin S. Multi-angle parking detection system using mask R-CNN. In: *BDET-2020. Proceedings of the 2nd International conference on big data engineering and technology; 2020 January 3–5; Singapore, China.* New York: Association for Computing Machinery; 2020. p. 76–80. DOI: 10.1145/3378904.3378914.

12. Yucheng Guo, Hongtao Shi. Automatic parking system based on improved neural network algorithm and intelligent image analysis. *Computational Intelligence and Neuroscience.* 2021;2021:4391864. DOI: 10.1155/2021/4391864.

13. Miao Y, Liu F, Hou T, Liu L, Liu Y. A nighttime vehicle detection method based on YOLOv3. In: *2020 Chinese automation congress (CAC); 2020 November 6–8; Shanghai, China.* [S. l.]: IEEE; 2020. p. 6617–6621. DOI: 10.1109/CAC51589. 2020.9326819.

14. Wang Hai, Lou Xinyu, Cai Yingfeng, Li Yicheng, Chen Long. Real-time vehicle detection algorithm based on vision and lidar point cloud fusion. *Journal of Sensors.* 2019;2019:8473980. DOI: 10.1155/2019/8473980.

15. Carrasco DP, Rashwan HA, García MÁ, Puig D. T-YOLO: tiny vehicle detection based on YOLO and multi-scale convolutional neural networks. *IEEE Access.* 2023;11:22430–22440. DOI: 10.1109/ACCESS.2021.3137638.

16. Bura H, Lin N, Kumar N, Malekar S, Nagaraj S, Liu K. An edge based smart parking solution using camera networks and deep learning. *2018 IEEE International conference on cognitive computing (ICCC); 2018 July 2–7; San Francisco, USA.* [S. l.]: IEEE; 2018. p. 17–24. DOI: 10.1109/ICCC.2018.00010.

17. Uzar M, Öztürk Ş, Bayrak OC, Arda T, Öcalan NT. Performance analysis of YOLO versions for automatic vehicle detection from UAV images. *Advanced Remote Sensing Journal*. 2021;1(1):16–30.

18. Detailed YOLOv5 model framework [Internet]. Beijing: Lezhi Network Technology Co.; 2021 March 3 [updated 2022 July 12; cited 2022 November 19]. Available from: https://blog.csdn.net/qq_16792139/article/details/114310670.

19. Zhu X, Goldberg AB. *Introduction to semi-supervised learning.* Cham: Springer; 2009. XII, 116 p. (Brachman RJ, Dietterich T, editors. Synthesis lectures on artificial intelligence and machine learning; lecture 6). DOI: 10.1007/978-3-031-01548-9.

20. Rosenberg C, Hebert M, Schneiderman H. Semi-supervised self-training of object detection models. In: *Seventh IEEE workshops on applications of computer vision (WACV/MOTION'05); 2005 January 5–7; Breckenridge, USA. Volume 1.* [S. l.]: IEEE; 2005. p. 29–36. DOI: 10.1109/ACVMOT.2005.107.

21. Lee Dong-Hyun. Pseudo-label: the simple and efficient semi-supervised learning method for deep neural networks. In: *ICML 2013 workshop: challenges in representation learning (WREPL); 2013 June 16–21; Atlanta, USA. Volume 3.* [S. l.]: [s. n.]; 2013. p. 2–4.

22. Sohn K, Zhang Z, Li C-L, Zhang H, Lee C-Y, Pfister T. A simple semi-supervised learning framework for object detection. arXiv:2005.04757 [Preprint]. 2020 [cited 2023 January 21]: [15 p.]. Available from: https://arxiv.org/abs/2005.04757v2.

23. Janocha K, Czarnecki WM. On loss functions for deep neural networks in classification. arXiv:1702.05659 [Preprint]. 2017 [cited 2023 January 21]: [10 p.]. Available from: https://arxiv.org/abs/1702.05659v1.

24. Zhang Haoting, Tian Mei, Shao Gaoping, Cheng Juan, Liu Jingjing. Target detection of forward-looking sonar image based on improved YOLOv5. *IEEE Access.* 2022;10:18023–18034. DOI: 10.1109/ACCESS.2022.3150339.

Received 11.01.2023 / revised 21.11.2023 / accepted 23.11.2023.