Parking Space Recognition Based on Deep Convolutional Neural Network

Zhuowen Chen, Zijun Gao, Jiaqi Li, Junjie Zhang, Yanan Dai, Wenbo Hu, Changmao Li
{1176655862@qq.com , gaozj@dlpu.edu.cn , 1499842219@qq.com}

School of Information Science and Engineering, Dalian Polytechnic University, Dalian, China

Abstract. Automatic real-time recognition for lots of spots is a crucial and challenging task, which can significantly improve the intelligence level of a city and make it that much more convenient for people to travel. Deep learning models have been applied in various fields, and significant development has also been achieved. However, taking into account interference in recognition images and occlusions of lane lines, deep learning still lacks accuracy and real-time.Therefore, the paper argues that combining OpenCV processing and deep learning should be utilized. This method can improve recognition for identifying parking spaces and increase the intellectualization of a city. In this paper, a network model based on VGG19 has been proposed, and the SENet network module has been added to its output along with a performing secondary enhancement on training images. Advanced pre-processing to obtain accurate point coordinates has finally obtained a satisfactory effect. It also [1] predicts images with severe interference and images with line occlusion. These two experimental results show that the network can effectively identify even with severe interference. At the same time, a large number of experimental evaluations also show that this method can also be applied to small target recognitions in many fields and can also be used as a basic product for small target detections and recognitions in the future.

Key Words: Pretreatment, Deep Learning, Point Transformation, SENet Network, Feature Enhancement

1 Introduction

Parking space recognition refers to capturing images through a fixed camera, using the captured images to determine the changes in parking spaces and the total number of vehicles in the parking lot in real time. This is a key task in the field of target recognition. Because of its wide application from video surveillance to target tracking, it has attracted widespread attention in the industry.

Although true [1,2] proposed a new method of detecting parking spaces in 2007, in which empty parking spaces in static images are used as markers to identify parking spaces frame by frame. This method uses the feature point distribution and color histogram of each car, and uses a combination method to identify and predict parking spaces. However, due to its large amount of calculation and error, it couldn't be accurately identified in real time, which couldn't meet the requirements of intelligent parking space identification. In 2009, Ichihashi et al. [3] proposed a new camera system (ParkLotD), which is based on C-means fuzzy clustering and

---

particle swarm optimization to adjust the hyperparameters, improving the accuracy of indoor vehicle detection and outdoor space, and the success rate of parking space detection. In 2018, Karakaya et al. [4] suggested another method of parking sp。ace detection based on deep learning, this method obtains a model by constructing a deep convolutional neural network to process parking lot vehicles, simply collect useful information and make predictions. However, The prediction accuracy and real-time results of this model are poor, and it is unable to predict the parking situation of the parking lot, and it has not conducted in-depth research on the network.

Real-time detection of parking spaces is essential for the development of an intelligent society. but there are still deficiencies in this part of the research. This paper proposes a real-time detection and recognition method for parking spaces based on deep convolutional neural network VGG19[5,6,7] combined with SENet[7,8] module. The premise of this method is pre-processing by OpenCV. Compared with other methods, this method does not have interference-free images for initialization, but uses single-frame images extracted from the video for processing. Through the prediction of the model, interference is effectively eliminated, and the prediction accuracy reaches 95%. Since it is based on single-frame image processing, our method has good generalization and can be extended to more fields.

2 Related Work

Completing the intelligent recognition of parking spaces is one of the key steps to promote urban modernization, and there are several representative studies on it. The first is the shortest search method based on genetic algorithm by intercepting video frames as a management medium. The second method is to recognize the license plate and extend it to the vehicle. However, due to the pollution and corrosion of the license plate, and the camera installation specifications, the license plate recognition system is prone to failures, recognition errors and high costs, which is not universal [9]. The third method is to conduct research on the basis of wireless communication technology and image processing technology. The above methods still have disadvantages such as low degree and slow speed.

The core technology of finding vacant parking spaces in parking lots through video frames captured by surveillance cameras is to accurately identify vacant parking spaces in parking spaces. In this paper, through a lot of analysis of the existing domestic intelligent parking lot management system parking space detection technology, comprehensive advantages and disadvantages, found that vehicles detection algorithm based on the image features of vehicles and empty parking spaces can be well applied to large area of outdoor parking lots. An image processing system based on video frame objects is proposed. Use the deep convolutional neural network VGG19-SENet to enhance vehicle target features, improve accuracy, and reduce loss. The information collection function can help drivers find free parking spaces quickly and accurately, save time to look for vacant parking spaces and occupy as little traffic resources as possible.

Through a large amount of research and analysis on the state and features of vehicle images, after the necessary noise filtering and other morphological preprocessing of the overall image of the parking lot [12], the key areas and irrelevant areas of the parking lot are specifically identified. Carry out image segmentation [13], so as to obtain the area of interest with less interference in the parking space, and then obtain the specific parking space image, achieve

the purpose of obtaining the spatial structure information of the parking space, and facilitate the analysis and integration of image information.

The image is input into the VGG19-SENet network structure[10], the important advantage is that the simple structure of VGGNet is adopted and the SENet image enhancement network is introduced. At the same time, it strengthens the learning ability of the network structure for parking features, and the training model can classify images more accurately. After all the fully connected layers are removed, computing resources are saved, and system performance is not significantly reduced.

3 OpenCV Pretreatment

3.1 Introduction to preprocessing

In order to improve the accuracy of parking space detection results, we need to perform image preprocessing on the parking lot video frames collected by the camera, so as to achieve the purpose of removing redundant information in the image and improving the reliability of the parking space image input to the network. In this paper, the preprocessing of the parking lot video image includes morphological operations to remove noise and to obtain the ROI area of each detection feature by fitting a straight line.

3.2 Image Denoising

Image denoising [14] is a process of reducing noise in digital images, sometimes called image denoising, to improve the image signal-to-noise ratio, so that the image quality in the input network is better. In this paper, due to the influence of the optical characteristics of the camera and the lighting changes of the captured video, there is a certain amount of noise in the collected parking map. In order to prevent noise from negatively affecting the final parking detection during the training model, the key is to eliminate noise. Noise removal is divided into two parts: coarse filtering and fine filtering. First, the dark spots and the extraneous color information in the top view of the parking lot of this paper are binarized according to the threshold, performing bitwise_and operation on the processed image and the original image. For the photos of the parking lot, median filtering is generally used for the first processing to filter out most irrelevant factors.

Fig. 1. Comparison of original image and coarse filtered image.

Fine filtering mainly uses the Canny operator in the OpenCV library to perform edge detection on the grayscale image of the parking lot. The working mode is mainly to detect the horizontal and vertical of the image and the diagonal pixels of the operator to obtain the horizontal direction G_x and the vertical direction G_y respectively. The purpose of the first-order derivative value is to obtain the edge detection binary image that can express the pixel gradient intensity G and the gradient direction θ, thereby obtaining the edge detection binary image of the overall range of the parking lot. The Canny algorithm is calculated as formulas (1) and (2):

$$G=\sqrt{G_{\mathrm{x}}^2 + G_y^2} \tag{1}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \tag{2}$$

The above work can complete the conversion of RGB images of the complex parking lot to the image denoising tasks that can clearly distinguish the edge of the parking space and part of the car outline, which is more suitable for machine processing [16].

Fig. 2. Edge binary image obtained by fine filtering.

3.3 ROI Area Selection

After obtaining a clearer parking space edge and contour image in formula (3), the main goal of this section is to select the fixed range of the parking lot by the method of straight line fitting, and use appropriate for the edge contour of the binarized image. The geometric feature extraction algorithm [17] draws the exact position of the edge of the parking space [18], then divides each train space into a cluster, and finally divides each parking space.

Firstly, the boundary corner points of the parking lot are selected to determine the general range of the parking lot, multiple boundary line equations are obtained by linear fitting through each corner point, the mask image of the parking lot is drawn, and the image after edge detection is matched through the mask operation to obtain the required ROI image.
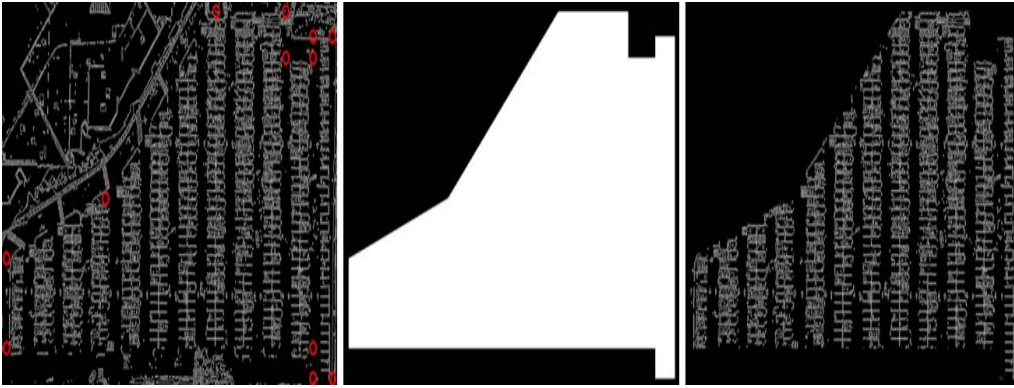

Fig. 3. ROI area edge image.

According to the feature difference between the straight line feature of the parking space and the irregular shape of the vehicle contour, the Houph transform that can filter the curve to determine the straight line can be used to obtain the boundary of each parking space [19]. The main principle of Houph straight line detection is to detect each edge, a straight line that may contain pixels in the Cartesian coordinate system is mapped to the polar coordinate system. The straight line is extracted by voting, the mapping relationship is as follows:

$$\rho = x\cos\theta + y\sin\theta \qquad (3)$$

However, the shortcoming of Houph transform that it is easy to lose the line segment length information. As a result, there is still some redundant information in the parking space boundary image, and no clear and distinct parking space segmentation map is obtained. In view of this situation, the bubbling sorting of each Houph line with the boundary conditions of the parking space can filter out these redundant parking space boundary lines, and get a clear parking space distribution boundary [20]. By calculating the starting parking point according to the space line spacing, the boundaries of all parking spaces can be determined, and the coordinates of each parking point can be obtained, then the initial state of each parking space can be obtained, which is convenient for subsequent segmentation and sampling of the parking space image.

4 Network model structure

This paper designed a deep convolutional neural network. The purpose is to accurately predict the situation of parking spaces in the parking lot and mark them in real time in the original image. Because the actual situation in the later stage is affected by complicated and special circumstances such as vehicle blurring and parking space occlusion, the final judgment result shows that the VGG19-SENet [21,22,23] network must have strong suppression ability against interference and noise, and have strong robustness. The network can still get better results in various complex situations.

4.1 SENet

Figure 4 shows the overall structure of the SENet network, which mainly contains the Squeeze module and the Excitation module. The Squeeze module performs convolution operations in a local space, and encodes all spatial features on a channel into a global feature. This paper uses global average pooling as the first step, complex aggregation methods can also be used. The Excitation module is an operation that integrates the relationship between channels after obtaining the global description characteristics. It has two characteristics. First, it is very flexible and can obtain the nonlinear relationship between each channel; the second is the learning relationship can be fused, allowing multi-channel features to be learned. Therefore, it can reduce the complexity of the model and improve anti-interference ability [25,26]. The basic mapping relationship is as follows:

$$Z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} U_c(i,j) \tag{4}$$

$$s = F_{ex}(z,w) = \sigma(g(z,w)) = \sigma(W_2 \operatorname{Re}LU(w,z)) \tag{5}$$

$$X_c = F_{scale}(u, S_c) = S \cdot U_c \tag{6}$$

The F_{sq} and F_{ex} in formulas (1) and (2) respectively represent the feature compression and extraction of the input feature maps [27,28] to achieve the purpose of reducing the dimensionality of the data. The flexibility of the SENet network is that it can easily connected with the existing network to achieve the effect of enhancing feature information and speeding up training. In this paper, the network that VGG19 and SENet are connected together is used.
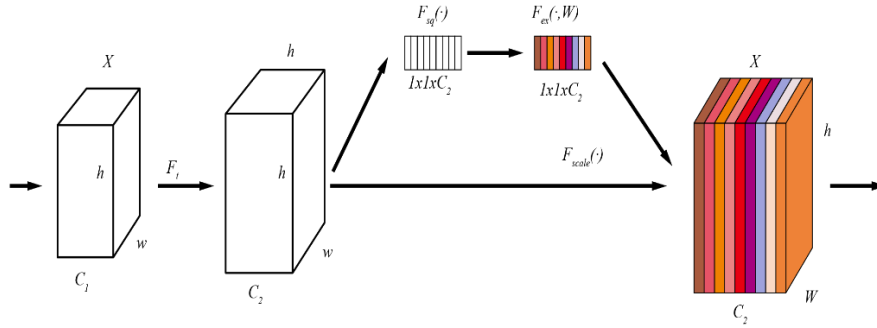
$F_{sq}(\cdot)$ $F_{ex}(\cdot,W)$

$1x1xC_2$ $1x1xC_2$

$F_{scale}(\cdot)$

Fig. 4. SENet network structure.

4.2 VGG19-SENet

Based on Karen Simonyan's VGG network, this paper constructs the VGG19-SENet network model. VGGNet explores the relationship between the depth of the convolutional neural network and its performance, with multiple 3×3 small convolution kernels and 2×2 maximum pooling layer, VGGNet successfully constructed a 16-19 layer deep convolution. The entire network uses the same size as the convolution kernel (3×3) and the maximum pooling size (2×2). Convolutional concatenation uses a larger convolution kernel alone, has fewer parameters, and has more nonlinear transformations than a single convolution layer. However, because VGG has three fully connected layers, it uses more parameters, consumes more computing resources, and reduces training speed. In order to optimize the network structure and reduce the number of parameters, the three full connection layers of VGG are removed, and the SENet network is added before OUTPUT to improve the training speed of the model and improve the accuracy of prediction. Finally, after the feature enhancement of SENet, the softmax function is used for output, and better results can be obtained.

Network structure, the VGG19-SENet network structure is shown in Figure 5. After the input image [48,48,3] is entered, first two layers of convolution are performed to obtain [48,48,46], and one layer of pooling output is performed [48,48,128]. The output image obtained after five rounds of convolution and pooling operations in the figure is [1,1,512]. Then enter the SENet network for feature enhancement, finally add a full connection layer, and use the Softmax function for output.
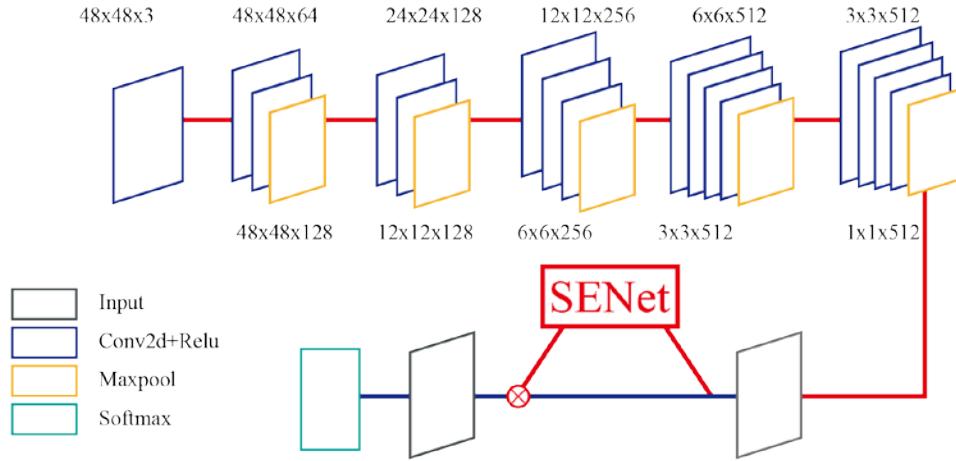
Fig. 5. Final network structure flow chart.

5 Experiment

In this paper, the VGG19-SENet network model is designed using the Keras framework. The optimizer selects SGD. Set Lr = 0.0001, momentum = 0.9, and learning rate 0.002. Using categorical_crossentropy as the loss function, and set other hyperparameters to default values. When generating a model, ImageDataGenerator is used to generate images in batches and enhance image features. As an image generator, it can iteratively loop and feed images into the network in batches.

5.1 Datasets and Expeerimental Settings

Data set: The data set used in this paper is composed of two parts, one uses the image obtained after OpenCV preprocessing, then put the coordinates of each parking space together. As shown in Figure 6, save the image of each car, the image size is 48*48. The other part uses Python crawlers to get top views of different types of cars on the Internet, which is also saved as 48*48. In the end, training set is 6000 images, and test set is 2000 images. This paper mainly sets the ROI area of the predicted images, removes the severely disturbed area, and reconstructs the place where the lane line is blocked, to show the results of OpenCV preprocessing and the excellent characteristics of the network.

Evaluation scheme: In the experiment, in this paper, each frame of parking lot image is predicted, and the accuracy of the prediction is used as an index to evaluate the performance of the VGG19-SENet model. At the same time, it can still perform well for videos with low definition and strong interference. In the case of severe interference, predicting objects with lower definition is a great challenge for our model and data set. We also evaluate the performance of the model through this experimental result.

5.2 Implementation Details

The prediction of parking spaces in this paper is based on two parts. The first part is to use the OpenCV tool to preprocess to obtain the point coordinates of each parking space, input the point coordinates into the model, and predict each parking space through the model in each frame of the image. The second part is to design a VGG19-SENet network structure, through SENet network and ImageDataGenerator class twice to enhance the image to improve accuracy. After basic preprocessing, the key is to classify the point coordinates into clusters, divide the parking lot by column, divide the parking lot line equidistantly, and accurately index each parking space. In the VGG19-SENet network, this paper uses different network models to make predictions and compares them with our network. The results are shown in Table 1. For our VGG19-SENet network, after experiments, we adjusted the convolution kernel to 3*3 and the step size to 1, and the predicted experimental results were the best, as shown in Figure 6.

Table 1. Accuracy of different networks and time used.

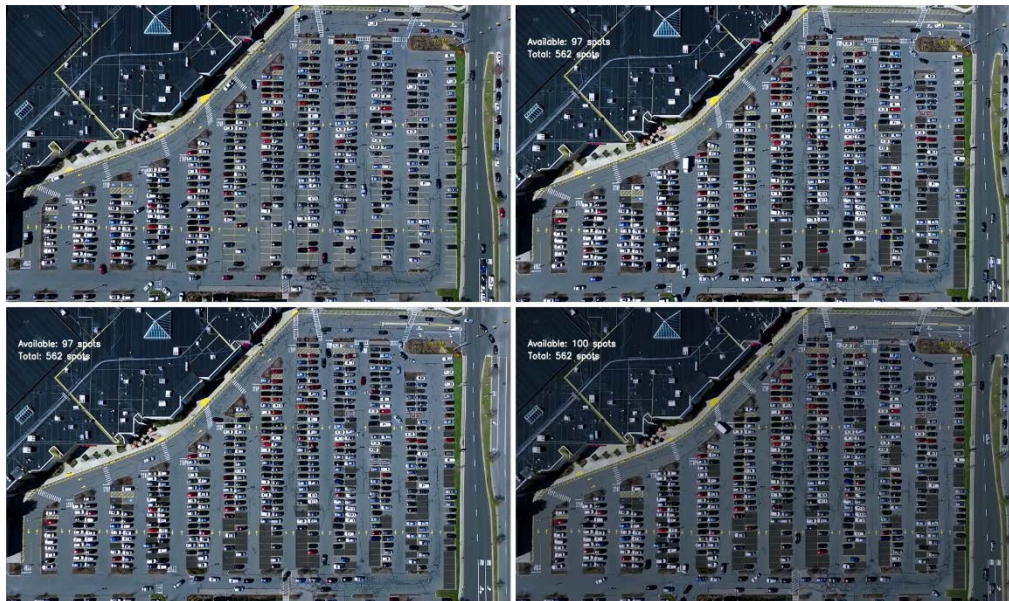| NET	Accuracy	Time
VGGNet	84.3%	12.4s
ResNet-50	78.2%	9.2s
ResNet-101	91.3%	15.1s
ResNet-152	90.3%	12.8s
BN-Inception	86.5%	14.2s
VGG19-SENet	95.6%	8.3s

Fig. 6. Predict experimental results

6 Conclusion

This paper proposes a method of VGG19-SENet network to improve the accuracy of parking space recognition. This method first uses OpenCV preprocessing to cut the image, obtains coordinates through point transformation, and then uses a new network model to predict the parking spaces. The key point is to use SENet network and ImageDataGenerator class to enhance the target feature twice to improve the success rate of small target detection. This is a method that has never been explored in the current technology. For this reason, we have added the SENet module to the VGG19 network. The function is to classify and strengthen the features after the output of the full connection layer of VGG19 to make the experimental results better. By comparing the experimental results of different networks and the prediction results under severe interference, the effectiveness of this method is verified, which is better than some current methods. In addition, experimental verification shows that our proposed network has good generalization and can be applied to different scenarios.

References

[1] Xu Lexian, Chen Xijiang, Ban Ya, Huang Dan.: Intelligent parking space detection method based on deep learning. China Laser. Vol. 46, pp. 230-241 (2019)

[2] True N. Vacant parking space detection in static images. pp. 17 (2007). San Diego:University of California

[3] Ichihashi H, Notsu A, Honda K, et al. Vacant parking space detector for outdoor parking lot by using surveillance camera and FCM classifier. IEEE International Conference on Systems, Man and Cybernetics. pp. 3603-3608 (2013)

[4] Karakaya M, Akıncı F C. Parking space occupancy detection using deep learning methods.2018 26th Signal Processing and Communications Applications Conference. pp. 1-4 (2018)

[5] Soft Computing; Findings from Xinxiang Medical University Provides New Data about Soft Computing (A Novel Multi-focus Image Fusion By Combining Simplified Very Deep Convolutional Networks and Patch-based Sequential Reconstruction Strategy). pp. 1275-1279 (2020). Computer Weekly News

[6] Saha S. Pagnozzi Alex, Bourgeat Pierrick, et al. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model. Neuroimage. pp. 215:116807 (2020)

[7] Hua Cui, Gege Yuan, Ni Liu, et al. Convolutional neural network for recognizing highway traffic congestion. Journal of Intelligent Transportation Systems. Vol. 24, pp. 279-289 (2020)

[8] Ze.L. Design of automatic parking space visual inspection and positioning system based on deep learning and OpenCV. pp. 15-85 (2019). Jiangsu University

[9] Fei.X. Research on Automatic Recognition Technology of Parking Space Based on Visual Inspection. pp. 1-72 (2015). Shenyang Polytechnic University

[10] Wujie.D.: Research and Application of Parking Space Recognition Method Based on Digital Image Technology. pp. 3 (2018). Guangxi Normal University

[11]Wenwen.W, Hongcai.T.: Research on Anomaly Detection Model based on Optimized VGG19 Convolutional Neural Network. pp. 253-258 (2020). Journal of Chengdu University of Information Technology

[12]Liangchao.J, Chuanyou.L, Fanqing.Y.: Lane Detection Based on OpenCV. pp. 41-44 (2018). Chang'an University

[13] Zhang Wan, Jin Fengxiang, Zhao Xiangwei, Ji Min, Li Ting.: Research on close-range image enhancement and segmentation algorithm based on OpenCV. pp. 881-886 (2018). Beijing Surveying and Mapping

[14] Hong.Y, Tong.Z.: Research on Classical Image Denoising Algorithms. Information and Computer. pp. 66-67. (2020)

[15] Lei.L.: Parking space detection and recognition method based on panoramic vision automatic parking. pp. 1-80 (2018). Xi'an Electronic TechnologyUniversity

[16] Yunong.W.: Research on Recognition Method of Parking Space Based on Image Similarity. pp. 1-68 (2018). Ji'lin University

[17] Guoxu.D.: Research and Practice of OpenCV Contour Recognition. Equipment manufacturing technology. pp. 101-103 (2020)

[18] Xiaochen.Z.: Research and Implementation of Parking Space Detection and Reversing Assistant Algorithm Based on Computer Vision. pp. 1-77 (2011). Northeastern University

[19] Yusong.H.: Research and Implementation of Parking Space Detection in Garage Based on Binocular Vision. pp. 1-90 (2019). Chong'qing University

[20] Xiaohang.L, Jia.G, Fulun.P, Jianjun.M, Cheng.W, Bo.S.: Improved algorithm of road detection based on Hough transform. Applied Optics. pp. 229-234 (2016)

[21] Daniel Octavian Melinte, Luige Vladareanu.: Facial Expressions Recognition for Human–Robot Interaction Using Deep Convolutional Neural Networks with Rectified Adam Optimizer. Sensors. pp. 1-18 (2020)

[22] Yang Ke, Wang Weibin, Xu Weiqing.: Parking space detection method based on small deep learning network. Automobile Practical Technology. pp. 155-158 (2020)

[23] Yongfei.X, Zhongming.Y, Yucheng.Y.: Outdoor parking lot empty space detection based on Faster R-CNN and ResNet. Fujian Computer. pp. 32-34 (2018)

[24] Zhifeng.Z, Jinqing.L, Wenzao.S.: Intelligent parking space query system based on deep learning. Computer System Applications. pp. 107-114 (2019)

[25] Sven Koitka, Lennard Kroll, Eugen Malamutmann, et al.:Correction to: Fully automated body composition analysis in routine CT imaging using 3D semantic segmentation convolutional neural networks. pp. 1-2 (2020)

[26] Zhang, C, Liu, S, Xu, X, Zhu, C.: C3ae: Exploring the limits of compact model for age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 12587-12596 (2019)

[27] Xie, J.C, Pun, C.M.: Deep and ordinal ensemble learning for human age estimation from facial images. IEEE Transactions on Information Forensics and Security 15. pp. 2361-2374 (2020)

[28] Shen, W, Guo, Y, Wang, Y, Zhao, K, Wang, B, Yuille, A.L.: Deep regression forests for age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2304-2313 (2018)