IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 December 2023, accepted 23 January 2024, date of publication 21 February 2024, date of current version 29 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3368521

RESEARCH ARTICLE

Research on Parking Space Detection and Prediction Model Based on CNN-LSTM

ZHUYE XU[1,2,3], XIAO TANG[1,3], CHANGXI MA[4], AND RENSHUAI ZHANG[1,3]

[1]School of New Energy and Power Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China
[2]National Engineering Research Center of Highway Maintenance Technology, Changsha University of Science and Technology, Changsha, Hunan 410114, China
[3]Intelligent Building and Building Automation Laboratory, School of New Energy and Power Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China
[4]School of Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

Corresponding author: Zhuye Xu (xuzhuye@mail.lzjtu.cn)

This work was supported in part by the Natural Science Foundation of Gansu Province under Grant 22JR5RA343, in part by the Open Fund of the National Engineering Research Center of Highway Maintenance Technology under Grant kfj220108, in part by Gansu Provincial Science and Technology Plan (Key Research and Development Plan-Industrial) under Grant 22YF7GA142, in part by Gansu Provincial United Scientific Research Fund General Project under Grant 23JRRA1486, and in part by Gansu Provincial Science and Technology Major Special Project-Enterprise Innovation Consortium Project under Grant 22ZD6GA010.

ABSTRACT With the continuous acceleration of urbanization, the parking problem is becoming increasingly serious. How to better manage parking resources has become an urgent problem to be solved in urban development. In this context, according to the historical data and real-time video data collected by the parking camera, this paper proposes an algorithm for parking space detection and state recognition. Through image preprocessing, region of interest selection, Hough line detection, and parking information recognition of the input test image, an intelligent parking space detection model is constructed, which improves the utilization rate of parking space and reduces the management cost. On this basis, according to the free parking space data obtained by the detection algorithm, a short-term demand prediction algorithm for on-road parking based on Convolutional Neural Network (CNN) and Long Short-Term Memory Neural Network (LSTM) was proposed. Through the preprocessing of input parking space data, time vector transformation, data separation, model training, and prediction, the parking demand data is predicted and analyzed. By comparing the prediction results of multiple models, it was found that the CNN-LSTM prediction model had the best model stability and goodness of fit, the lowest Mean Error (MAE) and Root Mean Square Error (RMSE), the errors of working days were 13.301 and 21.156, and the errors of rest days were 12.573 and 20.739, respectively. It shows that CNN-LSTM can effectively capture the time and spatial feature information of parking lot free parking space data, and the prediction accuracy is good, which can be used to predict the number of free parking spaces in parking lots.

INDEX TERMS Key words deep learning, parking space detection, image processing, convolutional neural network, long short-term memory neural network.

I. INTRODUCTION

A. MOTIVATION

With the continuous growth of the national economy, people's consumption level is constantly improving. In the pursuit of a higher quality of life, cars have become an essential means of transportation in Chinese families, which has also led to the rapid growth of the total number of cars in our country. According to the statistics of the Ministry of Public Security of China [1], In 2022, the number of motor vehicles in China reached 417 million, including 319 million automobiles, 34.78 million newly registered motor vehicles, and 23.23 million newly registered vehicles. According to the statistics on car ownership in various cities across the country, 84 cities have more than 1 million cars, 39 cities have more than 2 million cars, and 4 cities such as Beijing and Shanghai have more than 5 million cars.

As car ownership has grown dramatically, the pace at which parking infrastructure has been built has not kept pace. This leads to a slow increase in the supply of parking Spaces, forming a sharp supply-demand contradiction with the increasing number of cars. According to the data of China

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

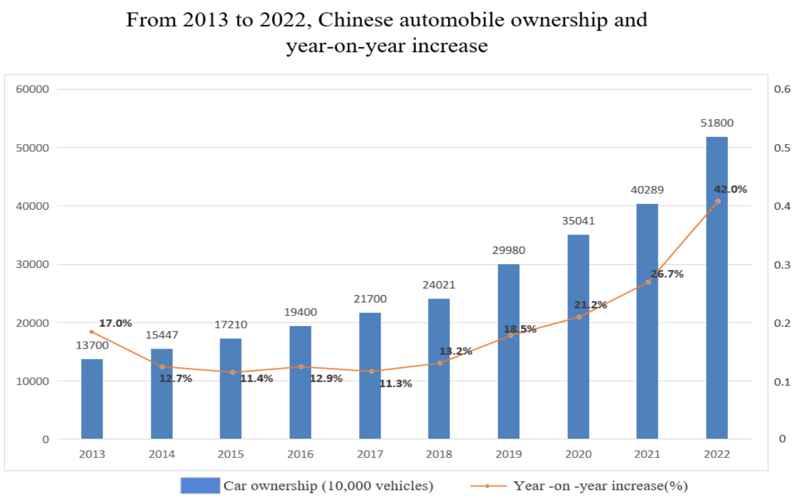


FIGURE 1. Automobile ownership and year-on-year growth in China from 2013 to 2022.

Urban Transportation Planning and Design Research Institute [2], there are only 0.6 parking Spaces for each car in big cities and 0.8 in small and medium-sized cities in China. This is far from the standard of 1.5 parking Spaces per car in developed countries, showing a serious shortage of parking Spaces in our country.

This imbalance between the supply and demand of parking Spaces has significantly affected the development of urban traffic. The difficulty of parking has become a common concern of urban residents, which not only aggravates traffic congestion, but also limits the convenience of residents' travel. To solve the problem of difficult parking, a series of targeted measures need to be taken, including but not limited to increasing the number of parking Spaces to be built, optimizing the management and utilization of parking lots, and encouraging the development of intelligent parking systems. Relevant departments should take action to alleviate the situation of insufficient parking Spaces through effective means, so as to ensure the smooth improvement of urban traffic development and residents 'quality of life.

B. RELATED WORKS

To solve the problem of imbalance between the supply and demand of parking resources, it is necessary to predict parking spaces in parking lots. Common prediction methods mainly include the following:

(1) Statisics-based methods: This method uses historical data for analysis and prediction, usually using techniques such as regression analysis [3], time series analysis [4], and cluster analysis [5]. The advantages of this method are low cost, fast solution speed, and easy implementation. However, the prediction accuracy of statistical prediction methods is related to many factors such as data quality, assumption limitation, complexity problems, and time problems. Among them, the data quality may be affected by many factors, the limitation of assumptions may lead to differences between historical data and future data, the complexity problem may not fully capture the diversity of parking demand, and the time problem may ignore the relationship between other factors, which also leads to the prediction accuracy cannot meet the actual needs.

(2) Machine learning-based methods: This method uses machine learning algorithms to analyze and predict historical data, such as support vector machines [6], fuzzy neural networks [7], gradient-boosted decision trees [8], etc. Compared with the methods based on statistics, the methods based on machine learning for parking demand forecasting have the advantages of automatic feature extraction, processing high-dimensional data, high prediction accuracy, and strong generalization ability. However, at the same time, there are also challenges such as high data quality requirements, algorithm selection and model parameter adjustment, and poor interpretability of prediction results. Therefore, it is necessary to comprehensively consider various factors for scientific evaluation and decision-making in practical applications.

(3) Methods based on sensor technology [9]: This method uses sensor technology to monitor the parking lot in real-time, and predicts the parking demand in real-time by collecting data such as parking space occupancy. The method based on sensor technology for parking demand forecasting has the advantages of high real-time performance, high data accuracy, and easy deployment and maintenance. There are also challenges such as high sensor installation costs, limitation of the number of sensors, and sensor accuracy and reliability.

(4) Deep Learning-based methods [10]: In the aspect of parking demand prediction, the deep neural network is used to learn and predict historical data, which has the advantages of automatic feature extraction, strong high-dimensional data processing ability, and high prediction accuracy. However, this method requires a large amount of data and computing resources and requires complex model training and tuning, and the interpretability of the prediction results is poor.

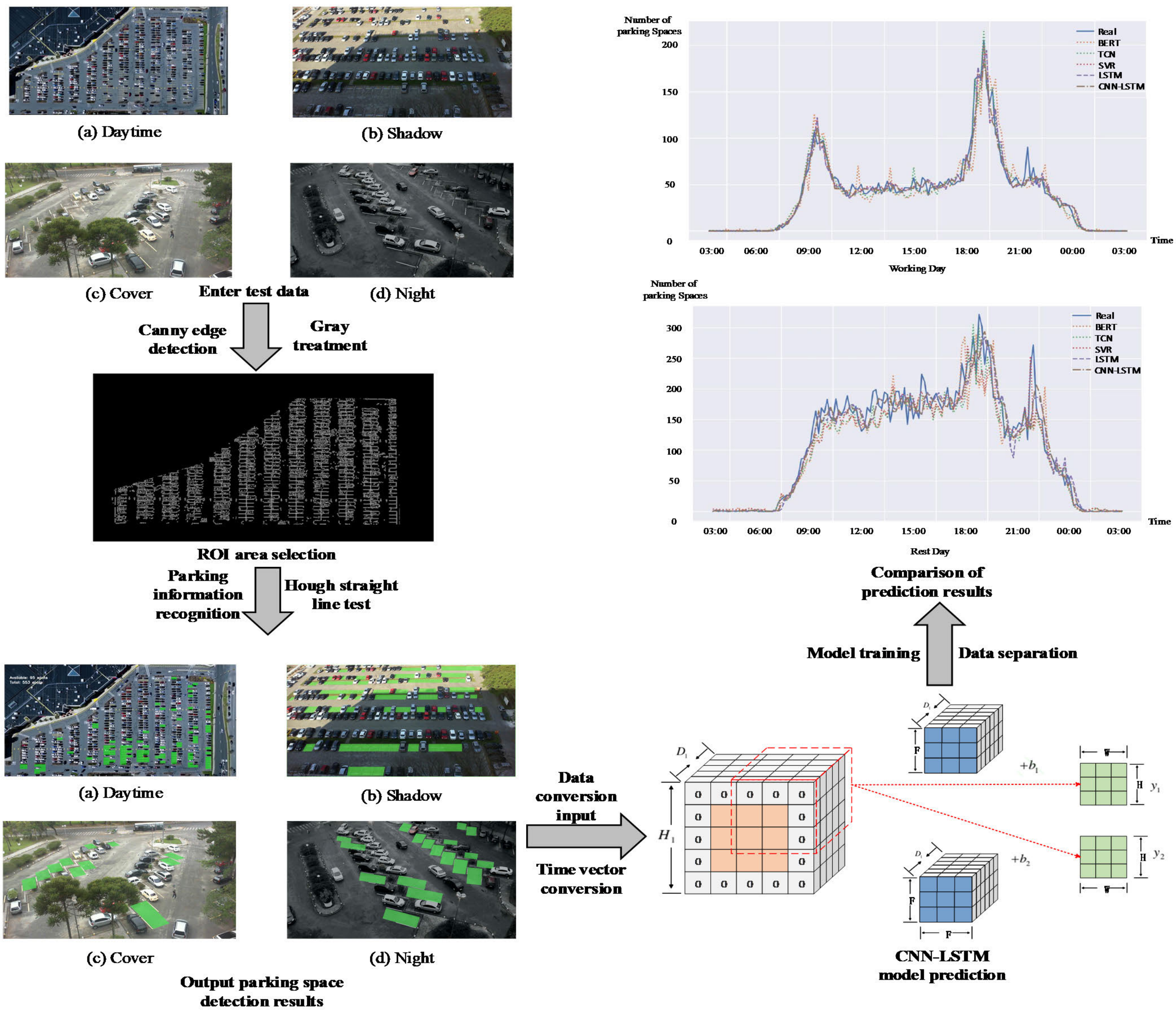


FIGURE 2. General structure diagram.

In summary, combined with the rapid development of current computer technology and equipment and the continuous improvement of computing power, this paper chooses the method based on deep learning. Through deep learning methods, features can be automatically extracted from a large amount of data, patterns in the data can be identified, and these patterns can be used to predict future parking demand, which is more feasible and practical in the field of parking demand prediction. In addition, the deep learning method can continuously optimize the model, improve prediction accuracy and generalization ability, and provide more accurate and reliable decision support for parking management and planning. Although prediction methods based on deep learning can achieve excellent prediction accuracy in the field of parking demand prediction, current studies often only consider historical time series data and ignore the real-time characteristics of parking demand [11], which will weaken the model's ability to mine and learn the change law of parking demand. The demand for parking spaces is often constantly changing, and the prediction model should also be constantly changing according to real-time. Therefore, if the real-time characteristics can be incorporated into the prediction model, the law of parking demand will be more accurately captured, and the prediction accuracy and practicability will be improved.

Based on the above research, this paper considers the historical data and real-time data of parking lots, constructs an intelligent detection model for parking spaces by edge detection of regions of interest, and identifies the occupancy of parking spaces. Combined with Convolutional Neural Network (CNN) and Long Short-Term Memory neural network (LSTM), a hybrid prediction model based on CNN-LSTM is proposed. Then the parking space can be effectively predicted. The structure of this paper is shown in Figure 2:

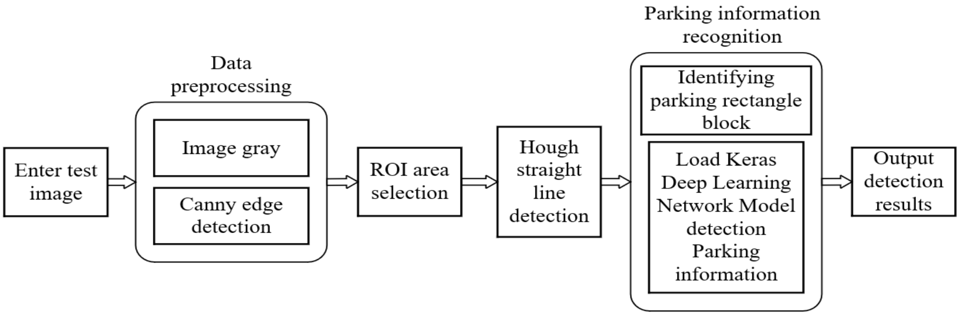


FIGURE 3. Working flow chart of intelligent detection and status recognition algorithm for parking spaces.

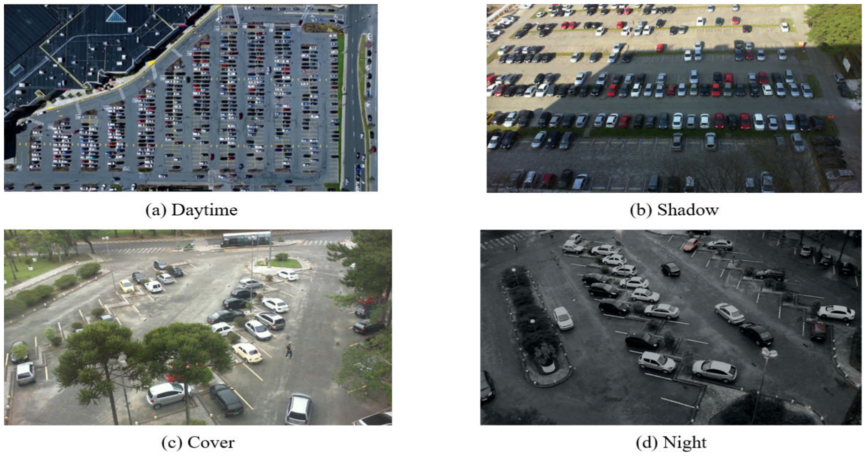


FIGURE 4. Enter test image.

II. INTELLIGENT PARKING SPACE DETECTION AND STATE RECOGNITION ALGORITHM BASED ON DEEP LEARNING

The premise of parking demand prediction is to accurately detect parking information and recognize the status information of parking spaces. In this paper, we propose a deep learning based intelligent parking space detection and status recognition algorithm. The algorithm uses the camera to collect parking video images as test data and then performs color selection, gray conversion, edge detection, ROI (Region of Interest) selection, line detection, rectangle recognition, and other operations on the test data to determine the location information of the parking space. A deep learning model is constructed to determine whether a parking space is occupied or not, which enables the state identification of parking spaces. Figure 3 shows the workflow of the proposed parking space intelligent detection and state recognition algorithm.

A. IMAGE DATA PREPROCESSING

1) THE IMAGE IS GRAYED OUT

In computer vision, image graying is the process of converting a color image into a grayscale image. For a color image, it consists of many pixels [12], and each pixel consists of the values of three color channels: red (R), green (G), and blue (B). Each color channel is divided into 256 levels, representing values ranging from 0 to 255, where 0 represents the minimum value of the channel, which is the brightest (whitest), and 255 represents the maximum value of the channel, which is the darkest (all black) [13]. The principle of image grayscale is to convert the RGB value of each pixel into a grayscale value, to obtain a single-channel grayscale image. When the three values are equal, that is, R=B=G, the image output color is gray. The process of graying is to convert the RGB value of each pixel to a gray value. When converting, we need to assign a weight to each channel to determine the proportion of each channel in the gray value. Typically, green has the highest weight, blue has the lowest weight, and red is in the middle. This is because the human eye has the highest sensitivity to green and the lowest sensitivity to blue. Considering the rationality of the image, the weighted average of the grayscale in this paper is shown in Equations (1) and (2):

$$Gray = 0.299R + 0.587G + 0.114B \tag{1}$$

$$Gray = R = G = B \tag{2}$$

Among them, Gray represents the gray value of the pixel in the image, 0.299, 0.587, and 0.114 are the most appropriate weight values of the three channels derived from the theory [14], and the sum of them is 1, which ensures that the gray value ranges from 0 to 255. The input test image before grayscale processing is shown in Figure 4, where A represents the parking lot picture unobstructed during the day; B represents a parking lot image produced by the shadow due to the sunlight of the day; D represents the parking lot image at

FIGURE 5. Image after Grayscale processing.

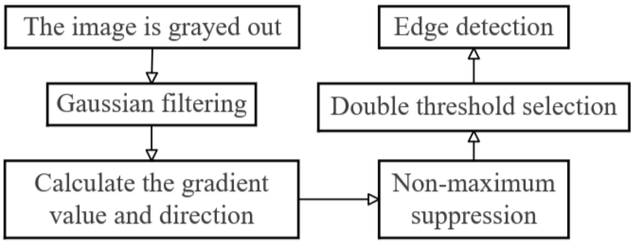


FIGURE 6. Flow chart of Canny edge detection algorithm.

FIGURE 7. Canny edge detection results.

night. Among them, the image ash processing of the parking lot image is shown in Figure 5.

2) CANNY EDGE DETECTION

Canny edge detection is a classical image processing algorithm, which was proposed by Qiang in 2020 [15]. The algorithm aims to detect the edge information in the image, which can reflect the information of the whole image to the greatest extent with the least amount of data. Edge can not only find the contour of the object but also segment different objects in the image. In a grayscale image, which consists of discrete pixels, we can binarize it to classify the pixels in the image into two classes: black and white. For binary images, the edge refers to the boundary line between two adjacent regions with different values [16], so the difference can be used to express the rate of change instead of the derivative. First, the image is smoothed by a Gaussian filter, and then the gradient strength and direction of each pixel in the image are calculated. Second, the non-maximally suppressed algorithm is used to remove the edge blurring effect. In summary, the detection algorithm can effectively identify the edge information in the image, and the edge of the parking space marker line can also be identified Producing a good recognition effect. Figure 6 shows the working flow chart of the Canny edge detection algorithm.

First, the image is preprocessed and a Gaussian filter is applied to smooth the noise and reduce the error of edge detection. The gradient and direction are calculated by the finite difference of the first-order partial derivatives, and two partial derivative matrices in the X and Y directions of the image are obtained. The convolution operator used by the Canny algorithm is as follows:

$$S_x = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} S_y = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \tag{3}$$

The mathematical expressions of gradient magnitude and gradient direction are shown in Equations (4), (5), (6), and (7):

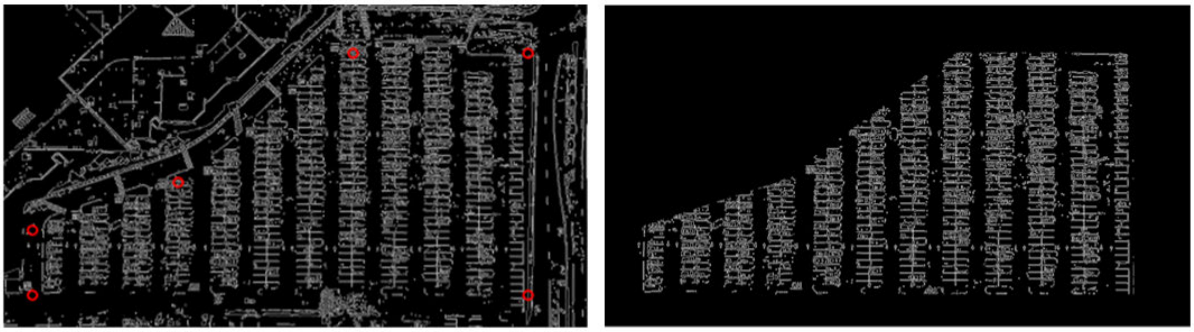
$$P[i,j] = \begin{pmatrix} f[i,j+1] - f[i,j] + \\ f[i+1,j+1] - f[i+1,j] \end{pmatrix} /2 \tag{4}$$

$$Q[i,j] = \begin{pmatrix} f[i,j] - f[i+1,j] + \\ f[i,j+1] - f[i+1,j+1] \end{pmatrix} /2 \tag{5}$$

$$M[i,j] = \sqrt{P[i,j]^2 + Q[i,j]^2} \tag{6}$$

$$\theta[i,j] = arctan(Q[i,j]/P[i,j]) \tag{7}$$

where f is the gray value of the image, P represents the gradient magnitude in the X direction, Q represents the gradient magnitude in the Y direction, M represents the magnitude of the pixel, and is the gradient direction. The larger the value of M at this point, the larger the gradient [17]. However, just calculating the gradient value does not prove that the point is an edge location in the image because the gradient value of a pixel point can be disturbed by noise and color changes. Therefore, non-maximum suppression is required, where the gray value of points whose gradient value is not a local maximum is set to 0, while the gray value of possible edge points is set to 128, to express the edge part of the image more accurately. Also, due to the presence of spurious edge phenomena caused by noise and color variations,



(a) Key points selection in ROI region (b) Generate ROI region

FIGURE 8. ROI region selection.

it is necessary to set a threshold to eliminate weak edge pixels. In Canny edge detection, an appropriate threshold is automatically selected according to the image content, and pixels below the threshold or above the high threshold are filtered out to retain the true edge information. Therefore, selecting an appropriate threshold is crucial for the accuracy of Canny edge detection. The results obtained by the Canny edge detection algorithm are shown in Figure 7.

From Figure 6, we can see that the Canny edge detection algorithm can efficiently detect the edges in the image and accurately detect and localize all the edges in the image under the condition that no false responses are guaranteed. In particular, the edge detection of the parking space marking line is closer to the real, with high accuracy and stability. It has been widely used in image processing due to its high sensitivity, its ability to capture the edge details of the target, and its ability to suppress noise interference and ensure accuracy and stability of edge detection through non-maximum suppression techniques.

B. ROI REGION SELECTION

ROI stands for Region of Interest, which refers to the part of the image or video that users pay special attention to or process [18]. In the field of computer vision and image processing, an ROI usually refers to an image region that needs to be extracted or analyzed, such as a region of a target object or a specific pattern region of interest. In this paper, according to the input test parking lot image, according to the percentage of the number of rows and columns of the image, six key points of the image are manually selected to form the trapezoidal part as the ROI area of this parking lot, while the street and building part outside the parking lot are filtered out. This trapezoidal ROI region can be viewed as a subset of the original image, containing a part of the pixels at the center position of the original image. Only the pixels inside the ROI are kept, while the pixels outside the rectangle are filtered out. This method is called mask-based filtering [19] because only the pixels within the ROI are retained, while the others are masked or filtered out. The next processing step will operate on this ROI region to obtain more accurate results. The manually selected positions of the image ROI key points and the generated ROI regions are shown in Figures 8 (a) and 8 (b), respectively

C. HOUGH TRANSFORM PARKING LINE DETECTION

The Hough Transform algorithm [20] is an image processing algorithm that can detect geometric shapes such as lines and circles in images. The principle of the algorithm is to transform the pixel representation on the image into the parameter representation in the parameter space, and then integrate and accumulate these parameters in the parameter space to finally determine the parameters of the line or circle [21]. This algorithm can efficiently recognize geometries in images and thus has a wide range of applications in computer vision. Specifically, first, a straight line in the image can be represented as the following equation in the Cartesian coordinate system x-y.

$$y = kx + b \tag{8}$$

where k is the slope of the line and b is the intercept of the line on the Y-axis. Straight lines can be represented by equations, and a straight line is composed of a series of points, an image is also composed of discrete pixels, so a straight line in an image is also composed of a series of discrete points. A point in the XY coordinate system can be represented by (x, y). However, we can deform the equation, so that point (b) can be viewed as a point in another coordinate system, namely in the KOB parameter space. As shown in Figure 9(a) and Figure 9(b).

Some points in the XOY frame will become straight lines when transformed to the KOB frame, and the line formed by two points in the XOY frame will become an intersection point in the KOB frame. When many points intersect at a point in the KOB coordinate system, this intersection point is the line to be detected. If the line is perpendicular to the X-axis in the XOY coordinate system, the slope does not exist, and the KOB coordinate system will lose its meaning. Therefore, Hough adopted the polar coordinate system and used distance and Angle to define the points in the coordinates, which can avoid this problem [22], as shown in Figure. 9(c).

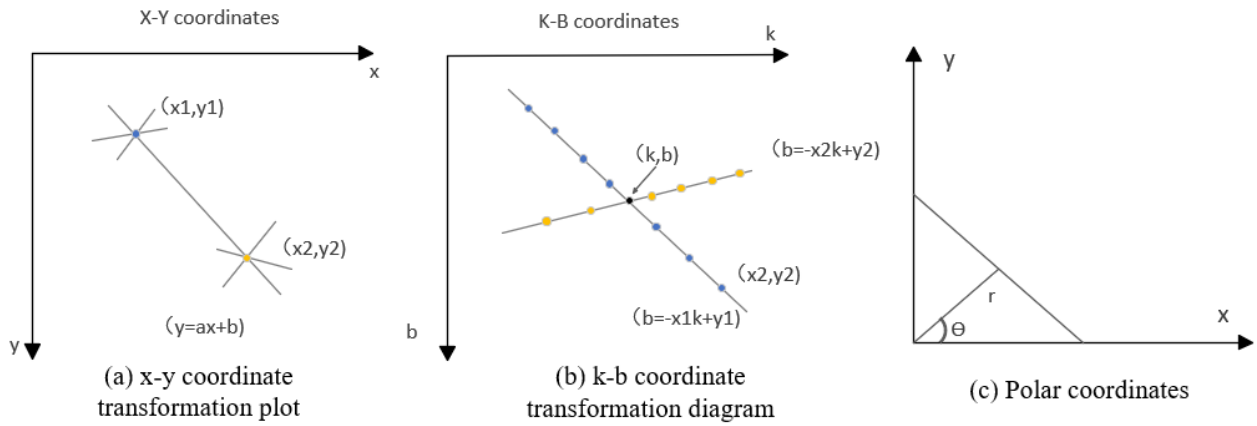


FIGURE 9. Conversion of polar coordinates parameters.

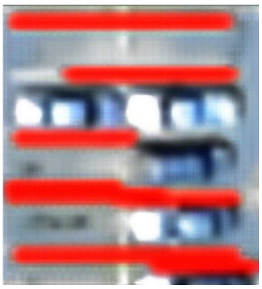


FIGURE 10. Hough line detection.

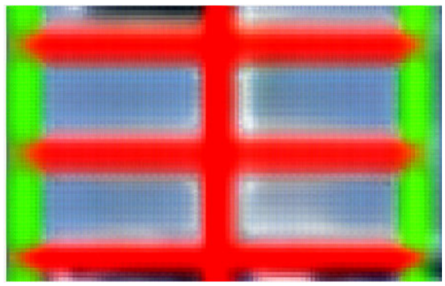


FIGURE 11. Drawing of the rectangular frame of the parking space.

Using polar coordinates, the point (x, y) distance becomes:

$$r = x\cos\theta + y\sin\theta \tag{9}$$

The points in the XOY frame become sinusoids in the polar frame, so the intersection of the sinusoids is the line to be probed. Due to the algorithm principle of the Hough transform, it is necessary to calculate the possibility of each point, which is a very time-consuming and labor-intensive task. Even after Canny edge detection, a large amount of data still needs to be processed. Therefore, to improve the efficiency of the algorithm, the probabilistic Hough transform can be used. This algorithm uses the mechanism of randomly selecting points to calculate, which can significantly reduce the amount of calculation, thus reducing the time complexity of the algorithm and improving the running efficiency. This function takes the binary image after edge detection as input and can control the distance and length of the line. HoughLinesP takes rho, theta, threshold, minLineLength, and maxLineGap. Here, rho represents the distance accuracy of the line segment in pixels and is set to 0.1. theta represents the Angle accuracy of the line segment in radians, set to pi/10; the threshold is the threshold parameter of the accumulation plane. Only when the length of the line segment exceeds this value, can it be detected. minlinelength is the minimum length of a line segment in pixels that is simply ignored and is set to 9. maxlineGap If the maximum separation between two lines is less than this value, it is considered a line and the value is set to 4. After the Hough transform, the result is shown in Figure 10:

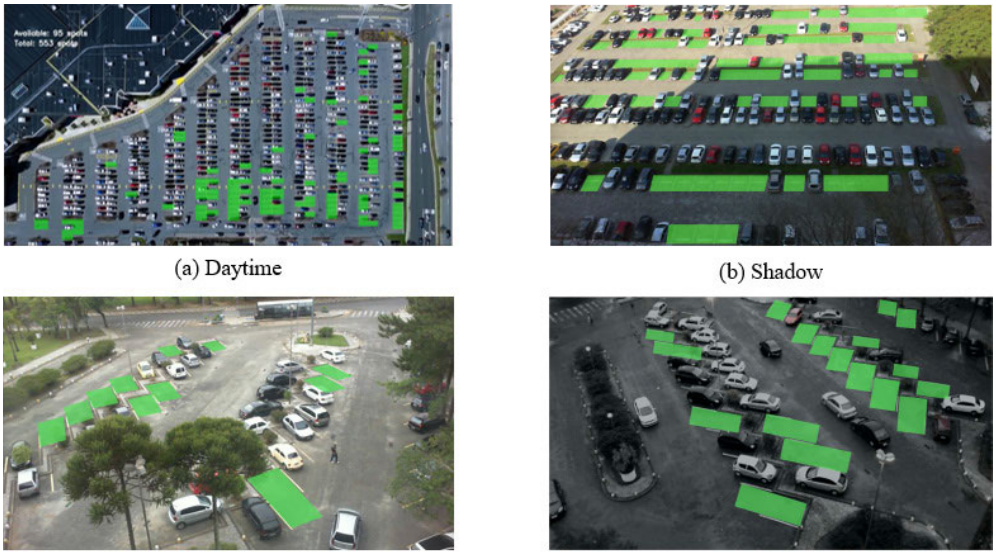
D. PARKING SPACE INFORMATION RECOGNITION

1) IDENTIFY THE RECTANGULAR BLOCK OF PARKING SPACE

In parking lot images, we need to identify the rectangular location of each parking space, which is crucial for subsequent parking state detection and prediction. After identifying the lines around the parking space by Hough transform, they are classified according to certain rules, and the park. The detected lines are converted into rectangular blocks using the identify_blocks function, and a new image and a list of rectangular blocks are returned. For each rectangular block, the approxPolyDP function is used to fit it to a rectangle to obtain the rectangular position of each parking space, which is convenient for further processing and analysis of each parking space in the future. The rectangular frames of the parking spaces are plotted in Figure 11:

2) A KERAS DEEP LEARNING NETWORK MODEL WAS CONSTRUCTED TO IDENTIFY EMPTY PARKING SPACES

After the parking space border recognition, we need to judge whether each rectangular parking space is parked, to detect the number of remaining empty parking spaces in the parking lot. We load the model from the weight file of the pre-trained deep learning model by taking the previously identified image of the parking space rectangle and importing it into the Keras load model function. Keras is a high-level neural network API written in Python, capable of running with TensorFlow, CNTK, etc as a backend. There is no separate configuration file in a specific format. Models are defined in Python code,



(a) Daytime (b) Shadow (c) Cover (d) Night

FIGURE 12. Recognition of parking space status.

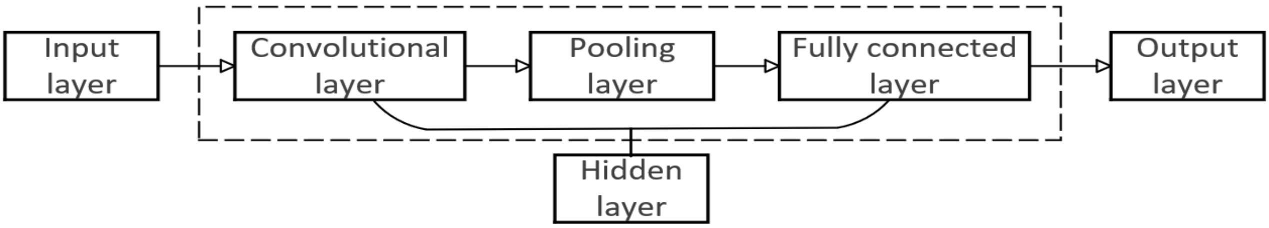


FIGURE 13. Convolutional neural network.

compact, easy to debug, and easy to extend. The core data structure of Keras is the neural network model. Here we define the Keras_model function, which takes a weight path parameter weights_path, loads the model from this path, and returns it via the load_model function. When the model is loaded, the function returns the model so that further prediction can be performed on it in the main program. Finally, the test images are normalized for prediction and judging, and the detection results are output.

The detection of the number of parking spaces in a parking lot is achieved by using a Keras deep learning pre-trained model, and the experimental results are shown in Figure 12:

As can be seen from the data in the upper left corner of the figure, there are a total of 553 parking spaces in the parking lot, and the number of empty parking spaces is 95 at the moment. Since this is only information about parking spaces at a particular moment, for different moments, it is necessary to identify parking spaces for each frame of the image, build a dataset, and implement parking space prediction for the parking lot.

III. PARKING SPACE DEMAND PREDICTION ALGORITHM BASED ON CNN-LSTM

A. EQUATIONS CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Network (CNN) is a kind of feedforward neural network that is outstanding in processing large image data. The convolutional neural network is a special neural network structure, which can effectively mine the spatial correlation in historical data [23]. Compared with the common multi-hidden layer neural network, CNN combines local sensing area, shared weight, and spatial or temporal down sampling, which greatly reduces the training parameters of the network and improves the efficiency of model operation [24]. The unique structure of CNN makes it better able to deal with tasks in image, speech, natural language processing, and other fields and becomes an indispensable part of the field of deep learning. CNN mainly includes a convolutional layer, linear rectifier layer, pooling layer, fully connected layer, and other structures [25]. Classical convolutional neural networks include a structure of convolutional layers, pooling layers, and fully connected layers, as shown in Figure 13:

1) CONVOLUTION OPERATION

Convolutional layers and nonlinear activation functions in convolutional neural networks compute the input and output by convolution operations. Convolutional layers extract features by sliding kernels over the input data. The weights in these kernels are multiplied and summed with the pixel values in the input data to produce an output feature map. At the same time, the nonlinear activation function performs

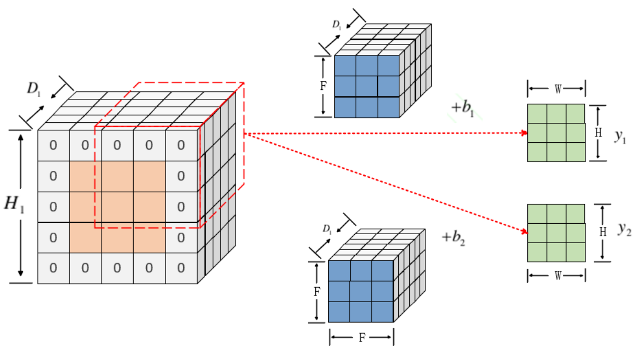


FIGURE 14. Convolution operation of a convolutional neural network.

a nonlinear transformation on the output feature map after the convolution operation, which further improves the expression ability of the model. The combination of convolutional layers and nonlinear activation functions enables convolutional neural networks to better capture the local features of the input data and form a higher level of abstract feature representation after stacking multiple layers. The calculation process is as follows:

$$y_i = \sigma(k_i \times x + b_i), i = 1, 2 \ldots K \tag{10}$$

where x is the input of a convolutional layer with a width W_1, height H_1 and depth D_i. The dimensions of each convolutional filter are F x F xD_i, where F is the width and height and i is the depth. Each convolutional filter has a biased term b_i. The nonlinear function is denoted by f. The output matrix of the ith convolutional filter is a matrix of size W $\times$ H $\times$ 1, and there are K convolutional filters. The step size of the convolution operation is S=1. For the output of the ith convolutional filter, the width is W_2, the height is H_2 and the depth is K. Each element in the ith output matrix is computed by taking the dot product of the input matrix and the ith convolutional filter k_i.

$$\begin{cases} W_2 = \dfrac{W_1 - F + 2P}{S} + 1 \\ H_2 = \dfrac{H_1 - F + 2P}{S} + 1 \end{cases} \tag{11}$$

In convolutional neural networks, the size of the output matrix can be changed by padding zeros around the input matrix. When the span of the convolution is S=1, we can make the input and output have the same size in space by setting the number of padded zeros on each side to P=(F-1)/2. Therefore, the width W_2 and height H_2 of the output can be calculated by the above formula.

2) POOLING AND FULLY CONNECTED LAYERS

In convolutional neural networks, the pooling operation is a down-sampling process, which can extract certain attributes from a set of related input data as low-dimensional output [26]. The most common pooling method is Max pooling, which selects the maximum value within the pooling window as the output result. This pooling operation can effectively reduce the dimension of the feature map, thereby reducing the computational burden and improving the generalization ability of the model. The downsampling operation of the pooling layer is shown in Figure 15 for the pooling operation of the convolutional neural network. The Max pooling operation is performed independently on each slice of the depth dimension of the input matrix, taking the maximum value in each slice as the output, so that the depth dimension of the output matrix is the same as the input matrix. In addition to this, the width and height of the output matrix can be calculated in the same way as the convolutional layer, which is usually achieved using the same stride as the width or height of the pooling window. By reducing the size of the feature map, the pooling operation can effectively reduce the computational complexity and memory consumption of the model, while also avoiding the problem of overfitting.

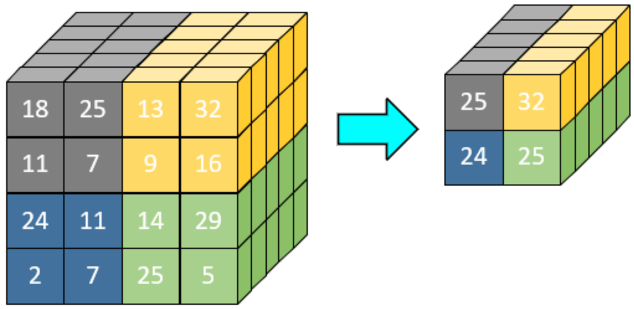


FIGURE 15. Pooling operation of convolutional neural networks.

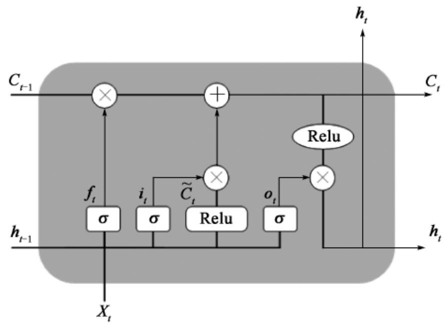


FIGURE 16. The working mechanism of the LSTM unit.

B. LONG SHORT-TERM MEMORY NEURAL NETWORKS

To better deal with time series problems, scholars have proposed a Recurrent Neural Network [27](RNN). By transferring the information processed at the current moment to the next moment, the network has a certain memory function, which can be used to solve the problems of language recognition, language modeling, machine translation, and so on. However, conventional recurrent neural networks suffer from the problem that they do not handle long-range dependencies well. After a long period of information transmission, useful information may be lost for reasons such as vanishing gradients or exploding gradients, resulting in the inability to remember long-term information. Therefore, based on the traditional recurrent neural network, scholars have proposed the Long Short-Term Memory network [28](Long

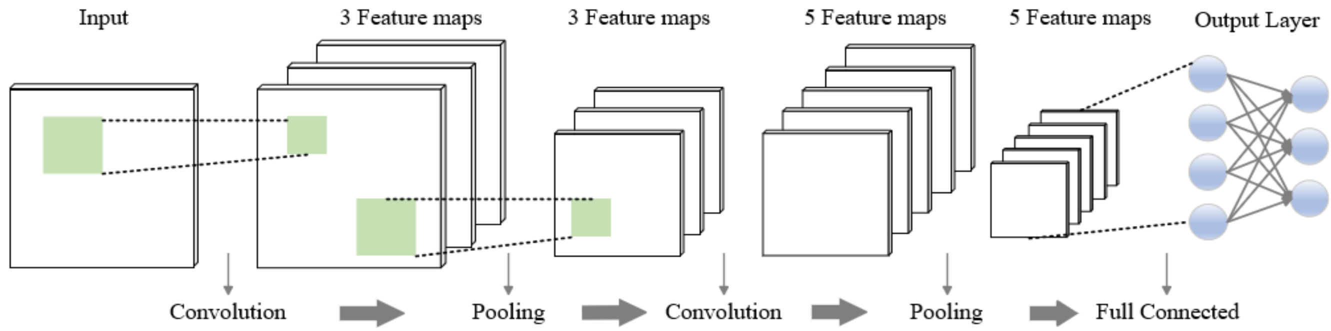


FIGURE 17. CNN neural network structure schematic diagram.

Short-Term Memory network, LSTM), which is an improved algorithm of RNN [29]. This network is well-suited to deal with long-term dependencies.

The biggest difference between RNN and LSTM is the difference in the structure of neurons in the hidden layer. Compared with the traditional RNN model, the LSTM model adds a cell state in the hidden layer to store long-term state information [30]. The key point of the LSTM is that it introduces three gating units including forget, input, and output gates to control the updating and discarding of historical information. The working mechanism of LSTM is shown in Figure 16:

1) FORGET THE GATE

The first LSTM step is used to decide what information can go through the cell state. This decision is controlled by the forget gate via sigmoid, which passes or partially passes based on the output from the previous moment. Its expression is:

$$f_t = \sigma(W_f \cdot [h_{x-1}, x_t] + b_f) \tag{12}$$

where W_f and b_f are the weight parameter matrix and bias of f_x, respectively. h_{x-1} is the hidden state at the previous time. x_t is the input information at the current time.

2) INPUT GATE

The second step consists of two parts. The first is the "input gate" which uses the sigmoid to determine which values to update; The second is the tanh layer used to generate new candidate values to add to get the candidate values. The combination of these two steps is the process of discarding information that is not needed and adding new information, which is expressed as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{13}$$

where f_t and i_t are the weights of C_{t-1} and $\tilde{C}_t$, and $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$, W_i, b_i are the weight matrix and bias of i_t, respectively. $\tilde{C}_t = tanh(W_t \cdot [h_{t-1}, x_t] + b_t$, W_t is the weight matrix of $tanh$, b_t is the bias; C_{t-1} and C_t are the cell states at the previous time and the current time.

3) OUTPUT GATE

The final step is to decide the output of the model, the output gate. We first pass the sigmoid layer to get an initial output, then use tanh to scale the value between -1 and 1, and multiply the output of the sigmoid pair by pair to get the output of the model. Its expression is:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{14}$$

$$h_t = o_t \cdot tanh(C_t) \tag{15}$$

where W_o is the weight matrix of the output layer; b_o is the bias.

C. CNN-LSTM PREDICTION MODEL

Parking lot data is a typical time series data, characterized by dense data and large volumes. Using CNN or LSTM model alone in the long-term prediction process may fall into the overfitting state or cause gradient disappearance and explosion problems [31]. In this paper, we propose a CNN-LSTM based parking demand prediction algorithm by combining CNN and LSTM to more accurately predict the number of free parking spaces in a parking lot at different time periods. In this algorithm, a CNN network is used to extract temporal features, and an LSTM network is used to process time series information. By combining the advantages of the two algorithms, more information can be learned in the spatiotemporal feature maps of parking, and the over-fitting and gradient problem of the single algorithm can be solved to improve the prediction accuracy and efficiency. As shown in Figure 17, this is a simple graph of the CNN structure. This is a diagram of the structure of a simple convolutional neural network. The first layer enters information about the graph data. Feature map, with depth 3. The feature maps of the second layer of feature maps are then further pooled to obtain the feature maps of the third layer. Again, the depth of this feature map is 3. Finally, repeat the previous operation to obtain the feature map of the fifth layer. At this point, the characteristic number of the feature map is 5. Finally, the fully connected layers.

In the CNN-LSTM model constructed in this paper, the CNN network is responsible for extracting the correlation features of on-road parking time, and the LSTM network is responsible for predicting the time series distribution of

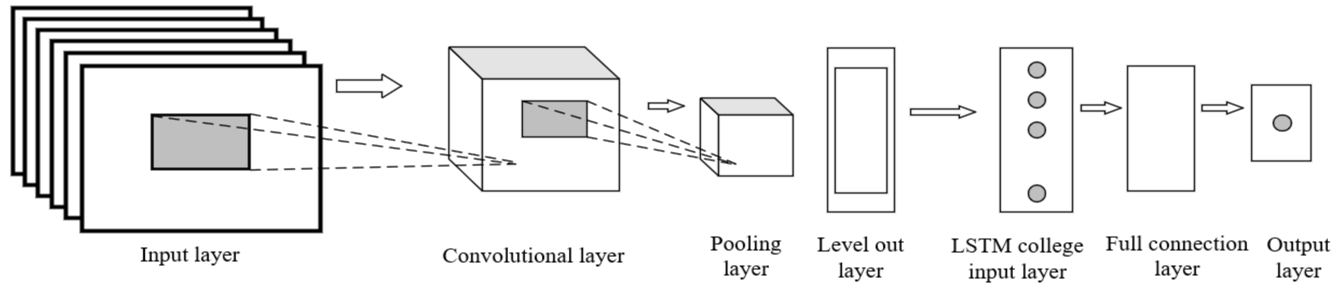


FIGURE 18. Structure of CNN-LSTM combined model.

on-road parking demand. For the multi-dimensional data characteristics, the data of n×m features were composed of n rows and m columns matrix, and the matrix data of t moments were constructed into a time distribution layer to obtain more step information of time series data. The dimension of the input layer of the CNN model is n×m×t. The output of the CNN convolutional layer feature map is:

$$\overline{X}_{i,j} = f_{cov}\left(\begin{array}{c}\sum_{n=0}^{k}\sum_{m=0}^{k} w_{n,m} \\ X_{i+n,j+m} + b_{n,m}\end{array}\right) \tag{16}$$

where $X_{i+n,j+m}$ is the value of the NTH row and m columns of the input matrix, $X_{i+n,j+m}$ is the selection activation function, $w_{n,m}$ is the weight of the convolution kernel n rows and m columns, $b_{n,m}$ is the convolution kernel bias, and k is the sliding window size. Usually, the input matrix is convolved using multiple convolution kernels. The CNN pooling layer uses 2×2 or 3×3 filters and sliding window step (stride = 2) to reduce dimension and down-sampling, maximize the activation of features with significant weights, remove interference and noise information, and then flatten the data into a one-dimensional array $\overline{X}_Z$ $\left(\overline{X}_Z = \left[X_{Z1}, X_{Z2}, \cdots, X_{Z(t-1)}\right]\right)$ through the flattening layer, which is input to the LSTM layer as a time series. Finally, the prediction results of parking space demand are obtained through the fully connected layer and the output layer, as shown in Figure 18.

Convolutional Neural Network (CNN) is a kind of deep learning network that can learn temporal local correlation and extract features [32]. We construct a deep learning network containing three convolutional layers with 16, 32 and 64 neurons set to extract temporal features.

Next, the output of CNN was reshaped into two-dimensional time series data, and then it was input into the Long Short-Term Memory (LSTM) network for time series feature extraction. The LSTM network can learn information about the time dimension in the features. To improve the performance of the model, we added the Dropout function after the LSTM network. We employ a two-layer LSTM network with 64 and 128 neurons, respectively, to obtain the temporal memory information in depth. The memory network capability of the LSTM network is able to fully learn the parking demand information in the temporal dimension for efficient training. Inside the LSTM network, the feature map is processed through three gating units, and then the desired feature information is output by the output gate.

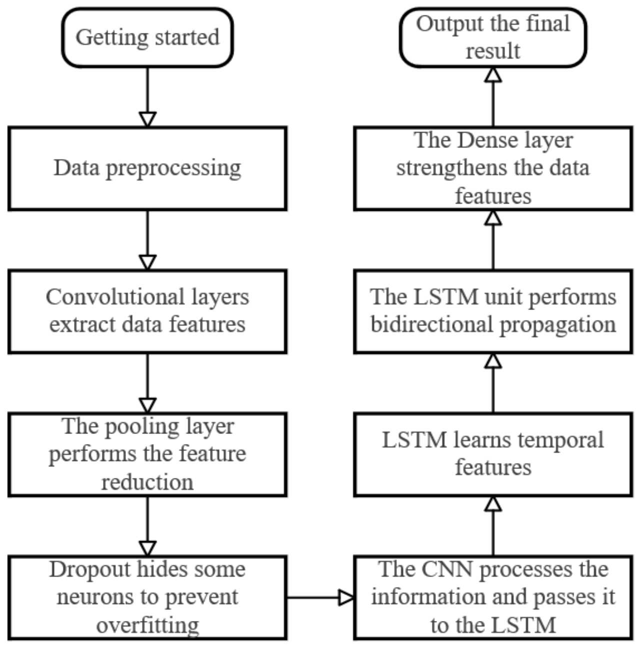


FIGURE 19. Structure of CNN-LSTM combined model.

Finally, a fully connected layer is used to connect all nodes of the previous layer to the current layer to synthesize the previously extracted features and output the result.

The flow of data processing using this model is shown in Figure 19:

IV. EXPERIMENTAL SIMULATION

A. EXPERIMENTAL DATA

The dataset used in this paper is parking data from July 1 to August 31, 2022 for a large parking lot in a city. The data period is 62 days. The number of idle parking spaces in the parking lot is collected per hour, which is around 1200, which is around 1200 article data. After data pre-processing, a total of 1158 data were obtained. During the modeling process, the dataset is divided into training and test sets according to a 9:1 10-fold cross-validation method. To obtain high-quality model parameters, the training set mainly uses the trained model to predict another part of the data. Because of the large distribution of traffic in the parking lot on work days and rest days, we will separate the work days from the rest days. We selected the data from August 22 to August 26 as the test

TABLE 1. Prediction error values.

Time	Day	Weather	Shadow	Status	Number of free
2022/7/1 12:00	1	3	0	1	289
2022/7/1 13:00	1	3	0	1	351
2022/7/1 14:00	1	3	0	1	329
2022/7/1 15:00	1	3	2	1	274
…	…	…	…	…	…
2022/7/1 20:00	1	3	10	0	89
2022/7/1 21:00	1	3	10	0	155
2022/7/1 22:00	1	3	10	0	201

TABLE 2. Partial data on parking spaces.

Sample Date	18:00		19:00		20:00		21:00		Weather	Activity	Traffic Flow
	Original	Normal ization	Original	Normal ization	Original	Normal ization	Original	Normal ization			
2022-7-13	189	0.6473	153	0.2942	81	0.1752	183	0.0289	0.5	1	0.65
2022-7-14	143	0.1765	139	0.1422	77	0.0215	112	0.1801	1	1	0.45
2022-7-15	152	0.2942	98	0.1698	29	0.0101	83	0.0322	0.5	0	0.25
2022-7-16	144	0.1771	112	0.0301	101	0.0000	227	0.2207	1	0	0.65
2022-7-17	151	0.2941	161	0.0389	124	0.0000	176	0.0101	1.5	1	0.55
2022-7-18	167	0.4529	99	0.4706	78	0.2941	197	0.1322	0.5	0	0.45
2022-7-19	156	0.2950	101	0.2948	44	0.2452	121	0.2031	1	1	0.35

set for the weekdays, and the data from the remaining days of the weekdays were used as the training set; The data is used as a test set. Table 1 shows some of the data. Of these, the weather is the state of the weather. The shadow is an area of shadow. The 1 value of Status is during the day and the 0 value is at night.

B. CNN-LSTM MODEL PARAMETER SETTING

In this paper, we choose to use the Keras framework to train and build a CNN-LSTM parking space demand prediction model. First, based on the above input data, the number of free parking spaces in different periods is determined as the input variable by taking into account the periodic changes of the number of free parking spaces and the temporal topology of different dates and different periods. Second, a prediction model is constructed using the deep learning framework PyTorch, and the optimal parameters of the network are determined by adjusting hyperparameters to predict the number of free parking spaces. The specific modeling process is as follows:

(1) Determine the input data characteristics. Based on the data shown in Table 1, the number of free parking spaces at different times throughout the day is used as the input feature of the CNN-LSTM model to track the spatiotemporal correlation features of the number of free parking spaces at different times.

(2) Normalization. To simplify the parameter learning process and improve the operation speed of the algorithm, the input vector is normalized according to formula (17), and the normalized input feature values all fall in the range of [0,1].

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{17}$$

where: x_i^* is the number of free parking spaces after normalization; x_i is the actual number of free parking spaces; x_{min}, x_{max} are the maximum and minimum values of free parking spaces, respectively. The processed data is shown in Table 2:

(3) Description of training parameters. The model is trained using the training and validation sets, and the model is optimized using a combination of forward and backward propagation algorithms. In the training process, we will set the input sequence length to "3", the feature dimension of the input data to "3", the number of layers of the CNN to "6", the number of neurons in each hidden layer of the CNN to "12", the number of samples used in each training model to "64", the total number of iterations of the model training to "5000". The learning rate is set to "0.001", the feature dimension of the model output is set to "1", and the ratio of splitting the dataset into training and validation sets is set to "0.9". Here, 90% of the data in the dataset is used to train the model and 10% of the data is used to validate the model. The prediction model will be obtained after the training.

(4) Evaluation of prediction results. After completing the training of the deep learning model, the test set is used to evaluate and predict this model. Using the data in the test set as input, the trained model is used to predict, and the corresponding prediction results are obtained. The prediction

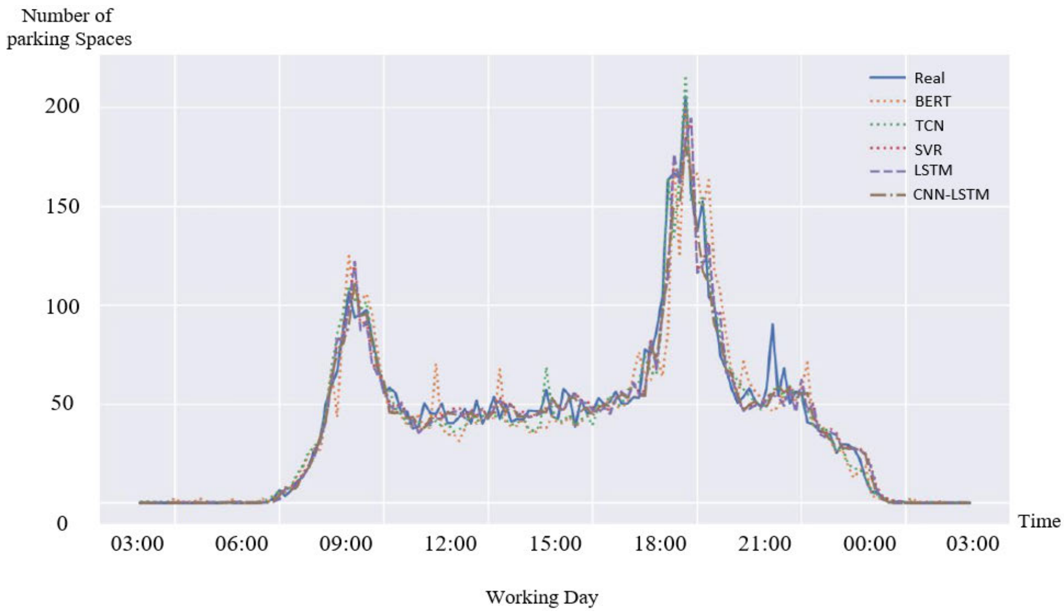


FIGURE 20. Each model vehicle traffic prediction comparison chart on the working day.

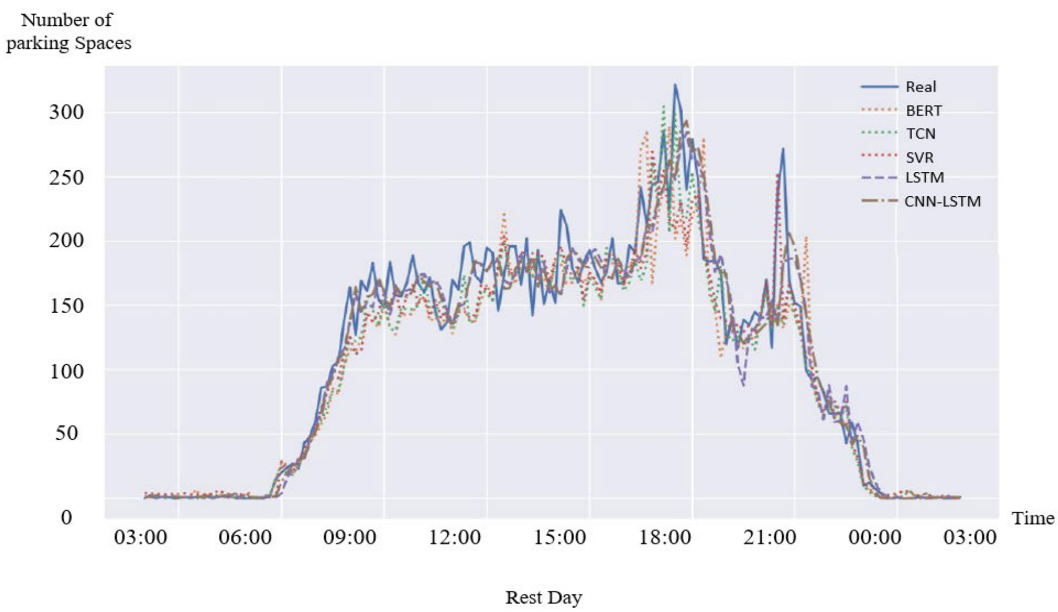


FIGURE 21. Each model vehicle traffic prediction comparison chart on the rest day.

TABLE 3. Prediction error values.

Dataset	Working Day	Rest Day
MAE	0.058	0.35
MSE	0.064	0.38
RMSE	0.076	0.42
R^2（Goodness of fit）	96.33%	98.42%

results are used to evaluate the model's performance and accuracy, to further optimize the model, or to apply it to real-world scenarios.

(5) Model evaluation metrics. The time prediction research in this paper is aimed at the regression problem of time series, so the mean Absolute error (MAE) is selected as the loss function, while the mean square error (MSE) is used as the evaluation standard, and the root mean square error (RMSE) is used when comparing the prediction results, and the results are shown in Equations (18), (19) and (20).

$$MAE = \frac{1}{m}\sum\nolimits_{i=1}^{m}\left|\overline{X}_i - X_i\right| \tag{18}$$

$$MSE = \frac{1}{m}\sum\nolimits_{i=1}^{m}\left(\overline{X}_i - X_i\right)^2 \tag{19}$$

TABLE 4. Comparison of model performance.

Model	Working Day				Rest Day			
	MAE	RMSE	Model stability	R^2（Goodness of fit）	MAE	RMSE	Model stability	R^2（Goodness of fit）
BERT	23.456	33.344	0.0721	0.8626	21.845	29.631	0.0667	0.8911
TCN	19.337	25.123	0.0431	0.9126	17.744	25.741	0.0519	0.9216
SVR	18.541	25.421	0.0655	0.9417	16.121	23.713	0.0572	0.9542
LSTM	16.599	23.728	0.0376	0.9543	14.389	22.712	0.0233	0.9639
CNN-LSTM	13,301	21.156	0.0216	0.9633	12.573	20.739	0.0199	0.9749

$$RMSE = \sqrt{\frac{1}{m}\sum\nolimits_{i=1}^{m}\left(\overline{X}_i - X_i\right)^2} \tag{20}$$

C. COMPARISON MODEL INTRODUCTION

1) BERT MODEL

Based on the Transformer structure, pre-training language models, obtaining language representation by reading a large amount of text data, with the ability to understand context information in two-way, perform excellent performance in natural language processing tasks, but more computing resources and data are required [33]. The advantage is the learning ability of the context and the complex language structure of the context and complex language structure.

2) TCN MODEL

A time sequence model based on convolutional neural networks, through convolution operations and multi-layer stacked structures to capture the patterns and characteristics in the time series. Compared with the traditional RNN and LSTM, long-term dependencies are more effective and have parallel to parallel Calculate advantages [34]. As a comparison model for CNN-LSTM models, TCN has the advantage of capturing long-term dependencies and training efficiency, and can be flexibly applied to sequence data of different lengths and types.

3) SVR MODEL

A machine learning model is used to regain problems, and the relationship between input data and target variables is established by looking for the best ultra-flat plane. The SVR is mapped to the high-dimensional space, and the best ultra-flat plane is found in this space to achieve regression forecasts [35]. In contrast to CNN-LSTM models, SVR is usually used for non-sequential data. The advantage is that small datasets have better returns and can handle nonlinear relationships, but time series data and complex patterns have poor capture capacity.

D. COMPARISON MODEL INTRODUCTION

Through the introduction of commonly used time sequence predictive models, the BERT, TCN, and SVR models are applied to the idle parking space data concentration of this article to concentrate the changes in traffic, and compare the analysis of the LSTM, CNN-LSTM models established in this article to verify the performance of the model of this model Essence The flow prediction results of each model on the working day are shown in Figure 20. The results of passenger flow prediction on the off-peak days are shown in Figure 21.

As can be seen from the above figure, for the basic rule of passenger flow, whether it is the traditional statistical algorithm or the algorithm based on a neural network, the model can learn certain characteristics of the passenger flow on weekdays and rest days and has a good prediction of the passenger flow in the future. Among them, the prediction accuracy of the BERT model is poorer than that of other prediction models, the prediction effect of the SVR model is relatively poor for peak hours, the prediction effect of the CNN-LSTM passenger flow prediction model is the best, and the fitting degree between the predicted value curve and the real value is the best.

According to Table 3 according to the error-index given, it can be concluded that the CNN-LSTM model performs well on this dataset and has good prediction and fitting ability. Lower values of mean absolute error (MAE) and mean square error (MSE) indicate better prediction performance of the model on the test set, while lower values of root mean square error (RMSE) indicate less difference between the predicted and actual results of the model. A high value of goodness-of-fit indicates that the model has a good degree of fit to the actual data, and has good predictive and fitting ability.

The stability of the model is measured by the standard deviation of the mean squared error of training the model five times:

$$S_{MSE} = \sqrt{\frac{1}{5}\sum\nolimits_{i=1}^{5}\left(MSE_i - \overline{MSE}\right)^2} \tag{21}$$

The effectiveness of each model can be obtained by training the parking lot data set, as shown in Table 4.

As can be seen from the above table, the CNN-LSTM combined prediction model is the best and has the smallest error values compared to the evaluation metrics of the five models. It has a MAE of 13.301, RMSE of 21.156 and model stability of 0.0216 on weekdays. The goodness of fit value is 0.9633. The MAE is 12.573 on weekdays, the RMSE is 20.739, the model stability is 0.0199 and the goodness of fit is 0.9749. It further demonstrates the superiority of the convolutional long short-term memory network CNN-LSTM model developed in this paper for short-term prediction of the number of free parking spaces.

Experimental results show that CNN is able to capture the temporal dependence in the prediction problem of the number of free parking spaces in a parking lot, while LSTM is able to capture the temporal dependence in the prediction problem of the number of free parking spaces in a parking lot. Passenger flow features are constructed from temporal and spatial features in two dimensions..The CNN-LSTM combined model is built to help predict the number of free parking spaces in a parking lot from two dimensions, time and space, to achieve optimal prediction.

V. SUMMARY AND PROSPECT

A. SUMMARY

With the increasing number of motor vehicles in our country, the problem of parking in urban areas has become more and more prominent. There are two main reasons for the difficulty of parking. On the one hand, the number of parking spaces is limited and cannot meet the demand for parking. On the other hand, even if there are spare parking spaces in the parking lot, it is difficult for drivers to know the number of spare parking spaces in a timely manner. Based on the above background, this paper focuses on the following:

(1) Analyze and summarize the research status and main research results at home and abroad. We objectively summarize and analyze the detection methods for valid parking spaces in domestic and foreign parking lots. At the same time, in terms of prediction and current traditional methods and deep learning, the advantages and disadvantages of each scheme are summarized from the combination of traditional prediction methods and mainstream neural network models. The introduction and multi-view analysis of existing technical research programs help to capture current research trends and pave the way for follow-up studies.

(2) Based on the real-time parking video data detected by the camera, the algorithm of parking space detection and state recognition is proposed by using gray processing, Canny edge detection, line detection, ROI region selection, Hough transform, parking space fitting detection, and other methods. Experimental results show that the proposed parking space detection method is highly accurate and robust.

(3) The relevant theories of parking space time series demand forecasting are introduced in detail. It has been shown that parking space sequence data is essentially a time series with the basic features of trend, periodicity and chaos. Therefore, the basic theory of time series, convolutional neural networks, long short-term memory neural networks, and other related knowledge are detailed in this paper to provide a theoretical basis for the design and construction of the model.

(4) Realization of parking space demand prediction scheme. In this paper, we propose a CNN-LSTM based model for parking space demand prediction, which combines the advantages of convolutional neural networks and long short-term memory neural networks to provide more accurate parking space demand prediction. Based on the detection data of the number of free parking spaces presented in this paper, experiments are performed by feeding the proposed model into the model. The results show that the proposed parking demand prediction model performs better than the existing parking demand prediction methods, and the prediction stability of the proposed model is greatly improved.

B. FUTURE PROSPECTS

As research on parking demand prediction methods in transportation becomes more in-depth, the overall operation efficiency of parking lots will become more efficient. Therefore, the study of parking demand prediction and its applications is an important research topic in transportation. At present, with the continuous application of science and technology in new theoretical and practical research, we are expected to effectively solve the problem of difficult parking, improve the service level of parking lots, and promote the development of big data, Internet, and automation in the parking field, and develop intelligent parking systems. However, the task of improving transport infrastructure construction is a long way off and we need to work hard. The related research work of this paper still has some shortcomings and needs to be improved, which is specifically shown in the following aspects:

(1) In this paper, computer vision technology is used to detect parking spaces based on the parking space video captured by the onboard camera. The quality of the video has a direct impact on the detection results, so in this paper we only consider weather conditions on sunny days in the experimental phase and do not conduct experiments in severe weather conditions such as cloudy, rainy and snowy days.

(2) The research in this paper is limited to a specific parking lot, and its parking changes are analyzed. However, there is a lack of research and discussion on surrounding or identical-type parking lots, which is also a limitation of this study. Moreover, while the analysis of the temporal dimension is important, the study of the spatial dimension cannot be neglected. Therefore, further analysis incorporating spatiotemporal features will certainly have a more accurate impact on the model predictions.

(3) Due to the strong volatility of parking space data, although the CNN-LSTM network can improve the long-term memory ability of time series, the prediction results usually show a smooth trend. In the future, further research on prediction methods is needed to more effectively capture the volatility of such data.

REFERENCES

[1] Internet Data and Information Center. Ministry of Public Security. *34.78 Million New Motor Vehicles Will be Registered in 2022*. [Online]. Available: https://baijiahao.baidu.com/s?id=1754692984352001313&wfr=spider&for=pc

[2] W. Zhuoran, S. Weihua, and Y. Yueyuan, "The current situation and innovation door of the parking lot under the environment of China's automobile ownership growth environment," *Ind. Technol. Forum*, vol. 6, no. 11, pp. 94–95, 2017.

[3] Z. Shengpai, "Fragrance thinking about the development of car ownership," *Shanghai Automobile*, vol. 12, pp. 24–26, 2018.

[4] W. Lei and Z. Zhongyi, "MLR-based bus itinerary prediction model," *J. Dalian Jiaotong Univ.*, vol. 36, no. 2, pp. 1–5, 2015.

[5] C. Shuhui, C. Lingshan, and Y. Jundian, "A radar maneuvering target detection algorithm based on K-means," *Agricult. Equip. Vehicle Eng.*, vol. 60, no. 7, pp. 100–104, 2012.

[6] S. Dong, M. Chen, L. Peng, and H. Li, "Parking rank: A novel method of parking lots sorting and recommendation based on public information," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Lyon, France, Feb. 2018, pp. 1381–1386.

[7] F. Mengyang, "Analysis of parking space detection method of automatic parking system," *Inf. Rec. Mater.*, vol. 22, no. 7, pp. 148–149, 2021.

[8] L. Di, G. Qihang, and J. Tinglin, "The parking space prediction supports the vector machine for the optimization of the bird search algorithm."

[9] L. Q. Chen, "Discussion on parking space monitoring system based on Zig Bee and infrared detection," *Inf. Rec. Mater.*, vol. 18, no. 3, pp. 7–8, 2017.

[10] A. Farley, H. Ham, and A. Hendra, "Real time IP camera parking occupancy detection using deep learning," *Proc. Comput. Sci.*, vol. 179, pp. 606–614, Jan. 2021.

[11] W. Peimao, Z. Congliang, and L. Zhengyue, "The design and implementation of the smart license plate recognition system," *Shanxi Electronic Technology*, vol. 2022, no. 1, pp. 38–40, 2022.

[12] H. Oshima, S. Yasunobu, and S.-I. Sekino, "Automatic train operation system based on predictive fuzzy control," in *Proc. Int. Workshop Artif. Intell. for Ind. Appl.*, May 1988, pp. 485–489.

[13] Q. Feng, L. Yong, and W. Shurong, "The key technologies of the panoramic parking auxiliary system," *J. Hubei Univ. Technol.*, vol. 35, no. 5, pp. 13–61, 2020.

[14] L. Eric and S. Mark, "Predicting pedestrian crosswalk behavior using convolutional neural networks," *Traffic Injury Prevention*, vol. 24, no. 4, pp. 6–8, 2023.

[15] Z. Qiang, "Design and implementation of an automatic parking system based on a panoramic camera," Ph.D. dissertation, Univ. Electron. Sci. Technol. China, Sichuan, China, 2020.

[16] L. Yongyi, M. Hongye, and Y. Wei, "Research on parking space status recognition method based on computer vision," *Sustainability*, vol. 15, no. 1, pp. 82–94, 2022.

[17] W. Feng, P. M. Parrot, and Y. R. Tie, "The design and implementation of an ultrasonic-based parking auxiliary system," *Car Electr.*, vol. 2020, no. 11, pp. 28–31, 2020.

[18] J. K. Suhr and H. G. Jung, "Sensor fusion-based precise obstacle localisation for automatic parking systems," *Electron. Lett.*, vol. 54, no. 7, pp. 445–447, Apr. 2018.

[19] W. Jinfen, "Based on the characteristics of time and space, the parking space on the road is a short-term demand prediction and its application research," Ph.D. dissertation, Ningbo Univ., Zhejiang, China, 2021.

[20] T. Xiaoyu, L. Sirui, and X. Qiuchi, "Facial expression recognition based on dual-channel fusion with edge features," *Symmetry*, vol. 14, no. 12, pp. 79–98, 2022.

[21] H. Yang, J. Lisheng, and W. Huanhuan, "Automatic ROI setting method based on LSC for a traffic congestion area," *Sustainability*, vol. 14, no. 23, pp. 118–124, 2022.

[22] L. Lei, "Parking space detection and recognition method based on panoramic visual automatic parking," Ph.D. dissertation, Xi'an Univ. Electron. Sci. Technol., Xi'an, China, 2018.

[23] W. Jinjiang and W. Pengfei, "A parking space detection method based on the visual vision system," *Anal. Instrum.*, vol. 2019, no. 1, pp. 71–77, 2019.

[24] X. Lexian and B. C. Xijiang, "Intelligent testing methods based on deep learning," *China Laser*, vol. 46, no. 4, pp. 230–241, 2019.

[25] G. Yucheng, S. Hongtao, and A. S. Hassan, "Automatic parking system based on improved neural network algorithm and intelligent image analysis," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1762–1766, Sep. 2021.

[26] Z. Xiaochen, "Research and implementation of computer vision-based parking space detection and reversing assist algorithm," Ph.D. dissertation, Northeast Univ., Boston, MA, USA, 2011.

[27] S. Fernando and B. Vinicius, "Image segmentation of breakwater blocks by edge-base Hough transformation," *J. Appl. Geodesy*, vol. 17, no. 2, pp. 11–15, 2022.

[28] C. Xianchang, "Deep learning algorithm and application research based on the convolutional neural network," Ph.D. dissertation, Zhejiang Univ. Technol., Zhejiang, China, 2014.

[29] J. Qu, F. Liu, Y. Ma, and J. Fan, "A neural-network-based method for RUL prediction and SOH monitoring of lithium-ion battery," *IEEE Access*, vol. 7, pp. 87178–87191, 2019.

[30] C. Wen, Y. Jie, and Y. Haima, "Research on the number of parking parts based on image recognition," *Intell. Comput. Appl.*, vol. 11, no. 4, pp. 82–87, 2021.

[31] W. Shiyu, L. Yan, and S. Yilai, "The application of an improved Canny operator in the robot vision system," *Comput. Syst. Appl.*, vol. 26, no. 3, pp. 144–149, 2017.

[32] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[33] K. T. Chitty-Venkata, S. Mittal, M. Emani, V. Vishwanath, and A. K. Somani, "A survey of techniques for optimizing transformer inference," *J. Syst. Archit.*, vol. 144, Nov. 2023, Art. no. 102990.

[34] Y. Yao, Z.-Y. Zhang, and Y. Zhao, "Stock index forecasting based on multivariate empirical mode decomposition and temporal convolutional networks," *Appl. Soft Comput.*, vol. 142, Jul. 2023, Art. no. 110356.

[35] L. P. A. Nathan, R. R. Hemamalini, R. J. R. Jeremiah, and P. Partheeban, "Review of condition monitoring methods for capacitors used in power converters," *Microelectron. Rel.*, vol. 145, Jun. 2023, Art. no. 115003.

ZHUYE XU received the Ph.D. degree from Lanzhou University of Technology, in 2021. He is currently a Teacher with Lanzhou Jiaotong University. His main research interest includes medical image processing.

XIAO TANG received the B.S. degree from Chengdu Technological University of Science and Technology, in 2022. He is currently pursuing the master's degree with the School of New Energy and Power Engineering, Lanzhou Jiaotong University, China. His research interests include intelligent transportation and image processing.

CHANGXI MA received the B.S. degree in traffic engineering from Huazhong University of Science and Technology, in 2002, and the Ph.D. degree in transportation planning and management from Lanzhou Jiaotong University, in 2013. He is currently a Professor with Lanzhou Jiaotong University. He is the author of three books and more than 100 articles. His research interests include ITS, traffic safety, and hazardous materials transportation.

RENSHUAI ZHANG is currently pursuing the B.S. degree with the School of New Energy and Power Engineering, Lanzhou Jiaotong University, China. His research interests include multi-object tracking and image processing.