Integrated Deep Learning and Attention Mechanisms for Accurate License Plate Recognition

*Xinhong Wei, Hongyun Song, Yibu Chang, Wenxin Jiang, Laixiang Xu**

School of Computer and Data Science, Henan University of Urban Construction, China
29813257@qq.com, songhongyunhuuc@yeah.net,
changyibu@huuc.edu.cn, jiangwenxinhuuc@163.com, xulaixiang@huuc.edu.cn

Abstract

With the rapid increase in the number of motor vehicles, the related issues of vehicle management are also increasing, and license plate recognition has become one of the most important means of managing vehicles. However, traditional license plate recognition methods suffer from complex feature modeling and low recognition efficiency in complex scene conditions. Therefore, we propose a deep learning license plate recognition method based on improved ZFNet and a fused dual attention mechanism. First, the classic ZFNet is simplified to promote feature extraction efficiency. Then, the channel and spatial dual attention mechanisms are used to enrich the feature extraction for license plate character positions. Finally, the residual network module is utilized to improve license plate recognition accuracy. The experimental results demonstrate that the proposed approach achieves an average recognition accuracy of 98.92%, which is superior to other excellent deep learning models, such as EfficientNet and ResNet. The suggested strategy has high promotional significance for enhancing the artificial intelligence enabled automated vehicle scheduling and control in smart city transportation.

Keywords: License plate recognition, Deep learning, ZFNet, Dual attention mechanism

1 Introduction

License plate recognition [1] is an important part of traffic control, widely used in various aspects such as traffic system sorting, parking fees, and law enforcement evidence collection. However, most license plate recognition methods still have certain limitations in accuracy and speed. Therefore, high-precision license plate recognition relies on the ability to accurately locate and recognize license plates, which remains a challenging problem in real-world dynamic traffic scenarios.

Traditional number plate recognition methods usually extract the inherent attributes of the license plate, such as edges [2] and color [3], as manual image features [4] for recognition. For example, Islam et al. [5] proposed an efficient ROI detection algorithm and obtained an accuracy of 83.52%. Shen et al. [6] used an SVM classifier for license plate recognition with an accuracy of 99%. Wen et al. [7] designed a model based on the improved Bernsen algorithm and Gaussian filter for license plate number recognition and obtained an accuracy of 93.54%. Although traditional methods can complete the task of license plate recognition, the processing of complex license plate data is complicated and cumbersome, making it difficult to adapt to the complex and ever-changing license plate recognition tasks.

In recent years, deep learning [8-10] has achieved remarkable results in several fields by virtue of its powerful feature extraction and learning capabilities. Convolutional neural network (CNN) [11], as one of the core algorithms of deep learning, is capable of automatically extracting rich feature information from images through multi-layer convolution and pooling operations to achieve accurate recognition of number plates. Rao et al. [11] proposed a new image correction scheme and an improved CRNN number plate recognition system and obtained an accuracy rate of 98.8%. Lin et al. [12] developed a real-time ALPR system based on edge AI and obtained 96% accuracy. Habeeb et al. [13] created a deep learning-based number plate recognition model and obtained an average accuracy of 87.21%. Tung et al. [14] used RetinaFace, a convolutional neural network with a single level of detection, and MobileNet as a method for license plate localization and obtained 98.60% accuracy. Saidani et al. [15] presented an automatic license plate recognition method improved on Faster R-CNN and obtained 97.46% accuracy. However, as the application scenarios continue to expand and become more complex, it poses higher challenges to the accuracy and robustness of the license plate recognition technique.

Based on the above analysis, we propose a deep learning license plate recognition method based on improved ZFNet and a fused dual attention mechanism. ZFNet [16], as a classical convolutional neural network structure, performs well in image classification and recognition tasks. Therefore, in this paper, by improving the ZFnet network structure, fusing the channel and spatial attention mechanisms, and introducing the setup residual network, we achieve more effective extraction of license plate character location features, improve the accuracy of license plate recognition, and enhance the robustness and recognition efficiency of the model. The main contributions

*Corresponding Author: Laixiang Xu; Email: xulaixiang@huuc.edu.cn
DOI: https://doi.org/10.70003/160792642025072604010

of this paper are as follows:

(1) Simplified ZFNet: By reducing the number of parameters and computational complexity of the model, the inference speed and efficiency of the network are improved. This can make the model more lightweight in practical applications and adapt to devices with limited computational resources.

(2) Fused dual attention mechanism: By introducing spatial and channel attention at the same time, it enhances the ability of the model to pay attention to key features, improves the feature representation effect of the model, and improves the accuracy of license plate recognition in complex scenes.

(3) Introduced a residual module: Mitigating gradient vanishing and explosion in deep network training by jumping connections, accelerating network convergence, and improving the model's expressiveness at a deep level.

The rest of this paper is organized as follows: Section 2 analyzes the literature review on number plate recognition. Section 3 focuses on the proposed methodology. Section 4 reports the experimental results and analyses. Finally, Section 5 gives conclusions and future research directions.

2 Literature Review

Accurate license plate recognition technology not only effectively improves the intelligence level of traffic management and monitoring and achieves fast and accurate acquisition of vehicle information but also improves the rapid control of traffic and is widely used in parking lot management, vehicle tracking, and security theft prevention. High precision and efficient license plate recognition play an important role in the rapid development of smart cities. There are currently many methods for license plate recognition, which can be roughly divided into traditional manual feature design and machine learning classification and excellent deep learning methods.

Traditional license plate identification algorithms primarily use the license plate's background, contour, texture, color, and other information to build manual feature extraction methods such as directional gradient histogram and scale-invariant feature transformation. The license plate is then classified and recognized using classifiers like SVM, high-speed learning machines, boosting models, random forests, and boosting algorithms. For instance, Corneto et al. [17] proposed a real-time number plate detection and segmentation method based on image analysis and processing techniques and obtained an average recognition accuracy of 91.5%. Zhu et al. [18] took a decision tree to achieve the recognition of number plate digits and obtained an accuracy of 90.00%. Hwan et al. [19] proposed an SVM method that can improve the pixel image of the license plate, and obtained the recognition accuracy of 97.57% of the license plate characters. Huallpa et al. [20] combined the OpenCV technique and obtained an average recognition accuracy of 95.5%. Islam et al. [21] proposed a region of interest number plate recognition algorithm based on the region of interest from the input image and obtained 92.8% accuracy. Wang et al. [22] provided a number plate recognition model combining color filtering and SVM and obtained an accuracy of 97.95%. Although traditional methods can also recognize license plate numbers, most methods rely on machine learning to manually extract features and design classifiers to complete the recognition task. These complex feature extraction processes greatly affect work efficiency and have a high degree of subjectivity. In addition, factors such as different numbers and letters, colors, sizes, shooting angles, and weather conditions of license plates can also pose certain recognition difficulties.

Deep learning-based number plate recognition approaches are widely used in practical applications as they show good accuracy and speed. For instance, Sultan et al. [23] constructed a new LPR framework and achieved an average recognition accuracy of 97.9%. Avianto et al. [24] exhibited a multitask learning method based on a convolutional neural network for number plate recognition and obtained an overall accuracy of 97.69%. Castro et al. [25] offered an improved single-shot detector and a depth-separable convolutional neural network and obtained an average recognition accuracy of 98.01%. Khokhar et al. [26] designed an integrated YOLOv8 for license plate recognition with an overall recognition accuracy of 99.02%. Huang et al. [27] offered a feature pyramid-based recognition method for license plate recognition and obtained 98.57% accuracy. Aqaileh et al. [28] proposed a deep learning technique using a transfer learning-based approach and obtained 99.6% accuracy. Wang et al. [29] reported combining the attention mechanism with YOLO-v5 and LPRnet and obtained 97.2% recognition accuracy. Although deep learning methods have made excellent progress in license plate recognition, there are still some shortcomings in the processing and extraction of important information, such as the adequacy of feature information extraction in key areas and the differentiation of background differences between similar numbers and letters. In addition, deep learning not only uses visualization methods to represent the extracted images in the feature extraction process but also finds it difficult to display the parts with significant internal differences, which brings certain un-explainability to the basis for selecting key features.

3 Proposed Method

3.1 Improved ZFNet

Combining the sample size and data characteristics of license plate data, this paper chooses to use the classical convolutional neural network ZFNet for feature extraction and classification recognition of license plate numbers and finds that the recognition effect is not ideal. Therefore, this paper improves ZFNet. In this paper, one fully connected layer is removed on the basis of the classical ZFNet, and the improved model consists of five convolutional layers, three pooling layers, one fully connected layer, and the output layer. In the Figure, k, s, and p denote the convolution kernel size, step size, and padding, respectively, and n denotes the number of output categories. The specific structural parameter settings are manifested in Figure 1.

In the convolutional layer, the feature map of the previous layer is involved in the convolutional operation through the convolutional kernel, and the result obtained is processed by the activation function to form a new layer of output feature map. The output feature maps of each layer and the input feature maps of the previous layer establish a convolutional relationship with each other. It can be defined as follows.

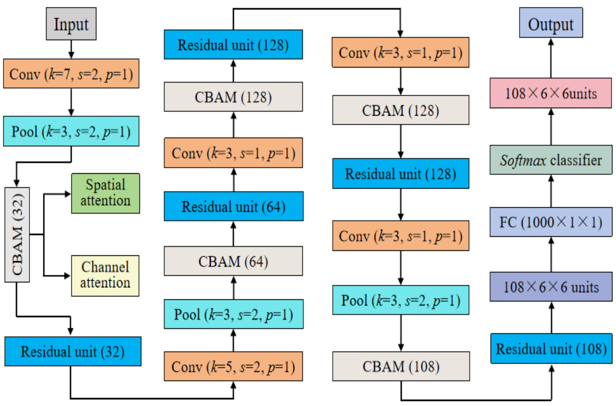

Figure 1. Improved ZFNet

$$x_j^l = f(\sum_{i\in m} x_i^{l-1} + b_j^i) \tag{1}$$

where l is the number of network layer order, k represents the convolutional kernel, m means the size of the feature map's sensory region on the initial image, and b denotes the offset.

The pooling layer is usually designed after the convolutional layer to reduce feature dimensionality and parameters and accelerate network training. In this paper, the maximum pooling method is used to process features. The expression of the feature map for the pooling operation is as follows.

$$x_j^l = f\left(\rho_j^l down\left(x_j^{l-1} + b_j^i\right)\right) \tag{2}$$

where down() is the down-sampling function, ρ and b represent the different constants used when the feature map is sampled.

Convolutional neural networks are usually paired with a fully connected layer in the last few layers in order to reduce the original high-dimensional feature vectors to low-dimensional vectors, removing redundant and noisy information in the high-dimensional feature vectors, and improving the recognition accuracy.

3.2 Added Convolutional Block Attention Module

The previous improved ZFNet convolutional layer may introduce redundant information, although it extracts multi-level features. Therefore, the convolutional block attention module (CBAM) is added to suppress unimportant features, reduce the interference of redundant features, and improve the computational efficiency of the network. CBAM adaptively weights different channels and spatial locations through channel and spatial attention, enabling the network to more efficiently learn and utilize features.

Let the input feature map be X, its number of channels be C, and global average pooling formula be (AvgPool). We operated on each channel c. The average AvgPool(X) of the C channels can be formulated as follows.

$$AvgPool(X)c = \frac{1}{H\times W}\sum_{i-1}^{H}\sum_{j-1}^{W} X_{c,i,j} \tag{3}$$

Where H is height and W represents width.

Global maximum pooling (MaxPool) operates on each channel c, and the mathematical expression for the maximum value MaxPool(X) of C channels can be expressed as follows.

$$MaxPool(X)c = \max_{i,j} X_{c,i,j} \tag{4}$$

The results of global average pooling and global maximum pooling are fed into a shared multilayer perceptron (MLP) or convolutional layer for feature transformation, respectively. Let the transformed features be FAvg and FMax. The MLP can be represented in the form of a linear transformation plus an activation function.

$$MLP(x) = \sigma(Wx + b) \tag{5}$$

where W is the weight matrix, b represents the bias term, and σ denotes the activation function.

The results of the feature transformations are summed, and the weight values Mc are generated by the Sigmoid function. It can be calculated as follows.

$$M_c = \sigma(F_{Avg} + F_{Max}) \tag{6}$$

The generated weight values are multiplied with each channel of the original feature map to achieve weighting of channel attention.

$$F_c(X) = M_c \odot X \tag{7}$$

where ⊙ is the element-by-element multiplication.

AvgPoolX performs average pooling and maximum pooling on the input feature map X in the channel dimension to obtain two spatial feature maps, AvgPool(X) spatial and MaxPool(X) spatial. The results of average pooling and maximum pooling are spliced together to form a new feature map F_{concat}. The spliced feature maps are fused by a convolutional layer to generate the spatial attention map M_s. They can be described as follows.

$$F_{concat} = concat\left(AvgPool(X)_{spatial}, MaxPool(X)_{spatial}\right) \tag{8}$$

$$M_s = \sigma\left(Conv\left(F_{concat}\right)\right) \quad \textbf{(9)}$$

where Conv is the convolution operation and σ denotes the Sigmoid activation function.

Multiply the generated attention map M_s with the original feature map X in the spatial dimension to achieve weighted spatial attention.

$$F_s(X) = M_s \odot X \quad \textbf{(10)}$$

The CBAM hybrid attention mechanism mainly combines channel attention and spatial attention to autonomously loop-learn features and make appropriate adjustments to channel and position parameters. Its mathematical expression is as follows.

$$F(X) = M \odot X \quad \textbf{(11)}$$

where M is the final attention map and F(X) denotes the weighted feature map.

These steps and formulas form the core of CBAM's hybrid attention mechanism, providing a powerful performance boost for convolutional neural networks in tasks such as image processing. The structure of the proposed CBAM is displayed in Figure 2.

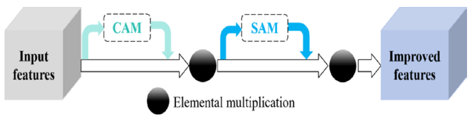

Figure 2. Proposed CBAM structure

3.3 Fused Residual Unit

Due to the proposed dual attention mechanism introduced earlier, the network training becomes unstable, which can easily lead to the gradient vanishing and exploding. Therefore, this paper introduces the residual connection module, which connects the input directly to the output jump, so that the gradient can flow directly, alleviating the problem of gradient disappearance or explosion and enabling the network to be trained to become more stable.

Residual learning is the process of transforming the mapping of network learning from X to H(X) into learning the skip learning connection difference from X to F(X)=H(X)-X, which can effectively solve the gradient explosion phenomenon that occurs during network training, namely residuals. The proposed residual unit is elucidated in Figure 3.

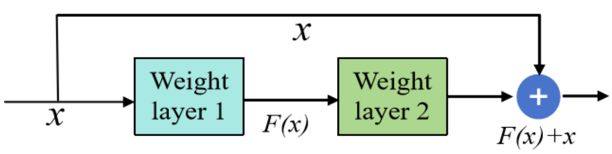

Figure 3. Proposed residual unit structure

Let the input be x and the desired output be H(x), the residual unit approximates the difference between H(x) and x by learning the residual mapping $F(x, \{W_i\})$.

$$F\left(x,\{W_i\}\right) = H\left(x\right) - X \quad \textbf{(12)}$$

where W is the weight parameter of the convolutional layer.

The final output y of the residual unit can be expressed as follows.

$$y = F\left(x,\{W_i\}\right) + x \quad \textbf{(13)}$$

$F(x, \{W_i\})$ represents the residual mapping learned through the main path convolutional layer, while x is the input passed directly through the jump connection.

Based on the above analysis, the proposed overall license plate recognition framework is presented in Figure 4. First, this article simplifies the ZFNet structure to improve feature extraction efficiency. Second, after simplifying the ZFNet pooling layer, a CBAM module is added to highlight the features of important positions in the license plate data through spatial and channel dual attention mechanisms, enabling it to fully extract rich key region features. Finally, a residual module is added after each CBAM module, and a skip connection feature extra extraction method is introduced to avoid the problem of gradient vanishing, allowing the network to learn the differences between input and output features more fully, accelerate convergence, improve optimization performance, and enhance the recognition performance of the model.

4 Experimental Results and Analysis

4.1 Experimental Setup

The experimental algorithm processing platform is a personal computer with an Intel Core i5-13500H processor and Intel Iris Xe Graphics Family graphics card. The software environment is Windows 11x 64bit, the programming language is Python 3.9, and the tool is PyCharm 2024.2.3. The deep learning framework Pytorch 2.4.1 is built on the server side to realize recognition training and testing.

4.2 Experimental Data

The experimental dataset mainly consists of the original images taken by mobile phones and the images generated by generative adversarial network, which contains a total of 5050 samples covering four types of license plates, namely, yellow, blue, white, and green.

According to the 3:1:1 ratio principle, the data is randomly divided into a training set, a validation set, and a testing set. The specific distribution of the number of each class is sketched in Table 1.

Table 1. The division of each type for license plate

Types	Training	Validation	Testing	Total
White license plate	1557	519	519	2595
Yellow license plate	663	221	221	1105
Green license plate	339	113	113	565
Blue license plate	471	157	157	785
Total	3030	1010	1010	5050

Blue license plates are commonly used for small family cars, off-road vehicles, short cargo trucks, etc. Yellow license plates are usually used for long-haul trucks, fuel-burning motorcycles, driving school training vehicles, long-distance buses, new energy electric vehicles, etc. White license plates are usually used for government agency vehicles such as public security, courts, procuratorates, etc. Green license plates are usually for new energy vehicles. The samples are manifested in Figure 5.

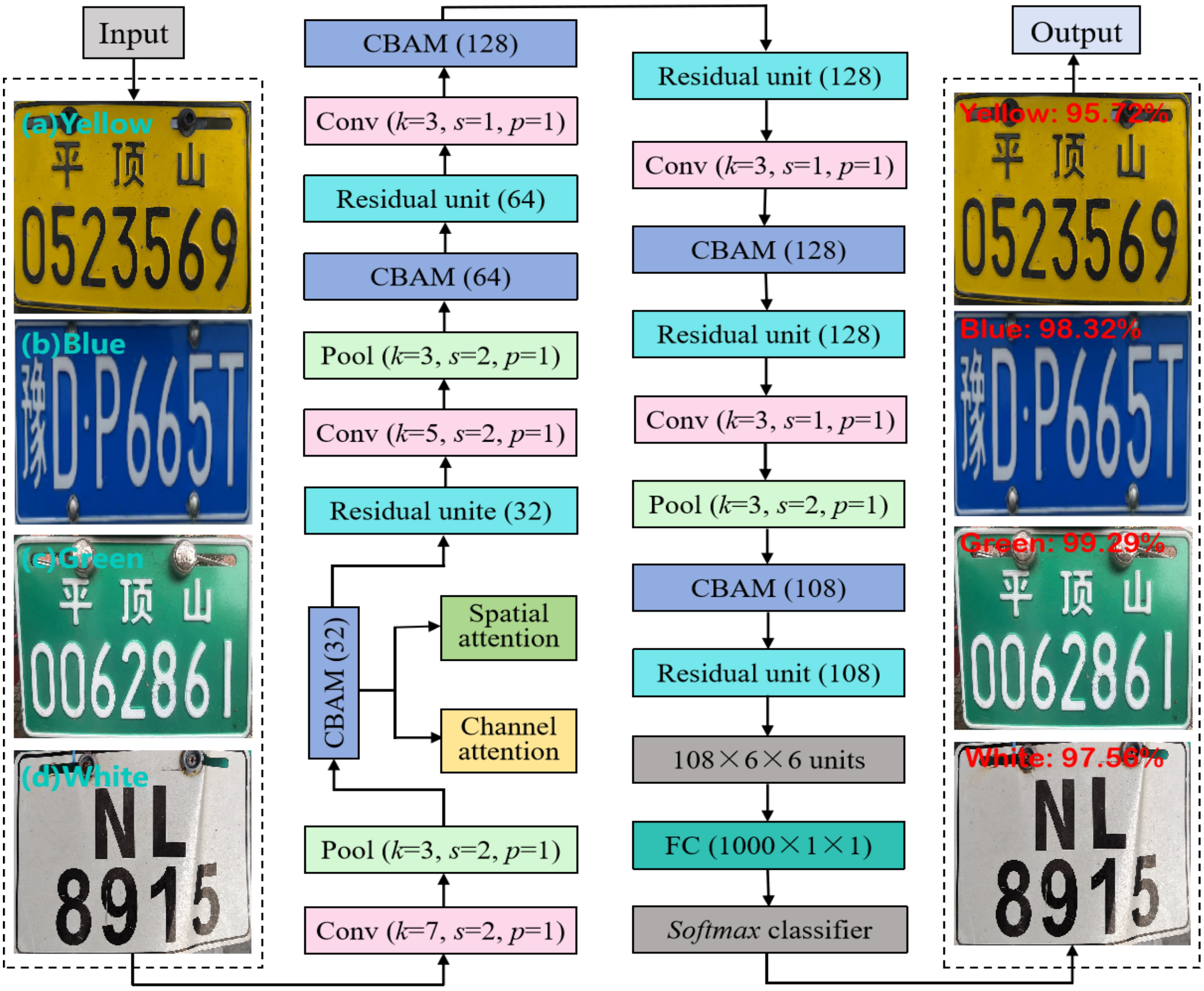

Figure 4. Proposed overall license plate recognition framework

(a) Blue license plate (b) Green license plate

(c) Yellow license plate (d) White license plate

Figure 5. The samples of license plates

4.3 Train and Test Results

We obtained the optimal parameters for this experiment by setting different experimental parameters. The results are shown in Table 2.

Table 2. Comparison of experimental parameters

Methods	Batch size	Learning rate	Epoch	Test loss	Test accuracy
Proposed model	16	0.001	20	0.210	0.982
	32	0.002	30	0.198	0.990
	64	0.002	40	0.232	0.972
	32	0.0002	50	0.081	0.998
	64	0.0002	50	0.100	0.994
Experimental settings	32	0.0002	1200	----------	----------------

As shown in Table 2, the recognition accuracy obtained by the experimental parameters used in this experiment is 2.6% higher than the lowest accuracy recognized by other parameters, which are all higher than other parameters. Overall, the recognition accuracy of the deep learning model constructed in this chapter is slightly higher than that of type recognition and can meet the requirements of license plate recognition.

To compare and validate the feature extraction ability of the proposed model with the traditional ZFNet for number plate data, we perform iterative training and observe the trend between the accuracy and loss values, and the results are exhibited in Figure 6.

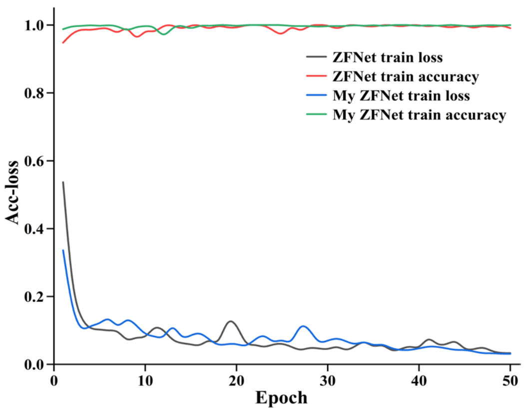

Figure 6. The accuracy and loss curves of the classic and improved ZFNet

The concept of ACC-loss is that accuracy is a measure of the proportion of correctly predicted samples to the total samples, which reflects the classification performance of the model, while loss is a measure of the gap between the model prediction and the true value, which is usually used to optimize the model. During the training process, a decrease in the loss value indicates a decrease in the model error, while an increase in the accuracy indicates an increase in the predictive accuracy of the model. The two generally change inversely, i.e., a decrease in loss is usually accompanied by an increase in the accuracy, but in some cases (e.g., imbalanced data or overfitting), the relationship between the two changes may be more complicated.

As illustrated in Figure 6, the loss function of the traditional ZFNet always decreases slowly from 0.55 downward and ultimately stays between 0.1 and 0.2, while the loss function of the improved ZFNet always decreases slowly from 0.39 downward and stays between 0.1 and 0.2. The fluctuation of the traditional ZFNet fluctuates more. Both networks' accuracies are steadily improving until they reach a stable convergence state. The upgraded ZFNet approach in this study offers a stronger feature extraction capability for license plate recognition, as seen by its increased accuracy compared to the more stable traditional ZFNet.

Metrics such as specificity, precision, sensitivity, f1-score, and accuracy are commonly used to evaluate the performance of deep learning models. Their expressions are as follows.

$$Specificity = \frac{TN}{TN + FP} \tag{14}$$

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{16}$$

$$F1\_score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{17}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

Specificity refers to the percentage of negative cases found among all negative cases. Accuracy depends on the model's predictions. Whether all are recognized or not, the desired classes must be identified as accurately as possible. Precision is the percentage of positive classes in the sample that give a positive result. Sensitivity is the ratio of all detected positive cases to all positive cases. The F1 score is the harmonic mean of memory and accuracy. If the precision or recall is low, the harmonic mean trends towards smaller values, so the harmonic mean gets closer and closer to that value, and the metric gets narrower. Accuracy is a metric used to evaluate classification models.

To compare the recognition performance of the proposed method with other outstanding deep learning models such as AlexNet, VGG16, ResNet18, and EfficientNet-v2 differently, we present them in Figure 7 with specificity, precision, sensitivity, f1-score, and accuracy.

As illustrated in Figure 7, the proposed model outperforms other deep learning models in all five evaluation metrics, showing obvious comprehensive advantages. Compared with the traditional ZFNet, the proposed model effectively compensates for the lack of feature extraction ability due to the smaller number of layers. Compared with ResNet18, the proposed model achieves higher accuracy with fewer network layers, indicating that the proposed model is more efficient in structural design.

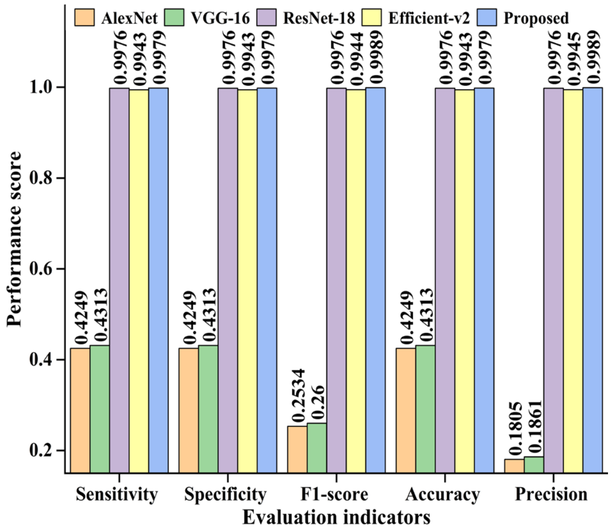

Figure 7. Comparison of evaluation indicators for different models

Although EfficientNet-v2 already performs well in terms of training efficiency, the proposed model not only achieves higher accuracy but also reduces the training time significantly. These results demonstrate that the introduction of the dual-attention mechanism can enhance the feature extraction capability of the model. Meanwhile, the partial convolution of the residual unit is used to replace the traditional convolution, which reduces the redundant computation, effectively reduces the number of parameters and training time of the model, and improves the overall computing efficiency.

An ablation experiment is a technique for expressing how different modules perform differently overall when combined, determining which factors are important to model performance and which have less of an effect, and directing model optimization or the definition of new structures to enhance the model's overall evaluation performance. This paper conducts ablation experiments using specificity, precision, sensitivity, f1-score, and accuracy to confirm the validity and feasibility of the proposed method, which consists of an attention mechanism, a residual unit, and an improved ZFNet. The outcomes of the ablation tests conducted using various techniques are exhibited in Figure 8.

Some important observations can be made from the content of Figure 8. The proposed method, Improved ZFNet + Residual unit + CBAM, is higher than the other combination methods in the five evaluation indexes and is 2.77% higher than the lowest, Improved ZFNet + CBAM, in terms of the f1-score. The recognition performance of adding residual units is significantly higher than that of adding attentional mechanisms, indicating that residual units have higher performance than attentional mechanisms in license plate recognition. Overall, the proposed three combinations of ablation methods do not conflict with each other and can effectively improve the recognition performance of license plates.

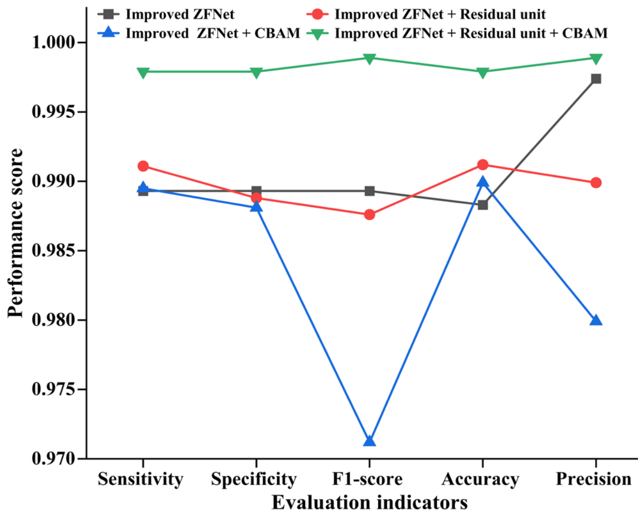

Figure 8. Comparison of assessment indicators for different models

A confusion matrix is used to clearly present the specific classification situation of each method for each class, such as the difference between the model's good and poor classification of a certain class. The quantity on the main diagonal usually represents the number of correctly classified items, while the quantity on top represents the number of incorrectly classified items in a certain category. In this paper, the confusion matrix is used to show the comparison of classification results of different models for license plate recognition in Figure 9.

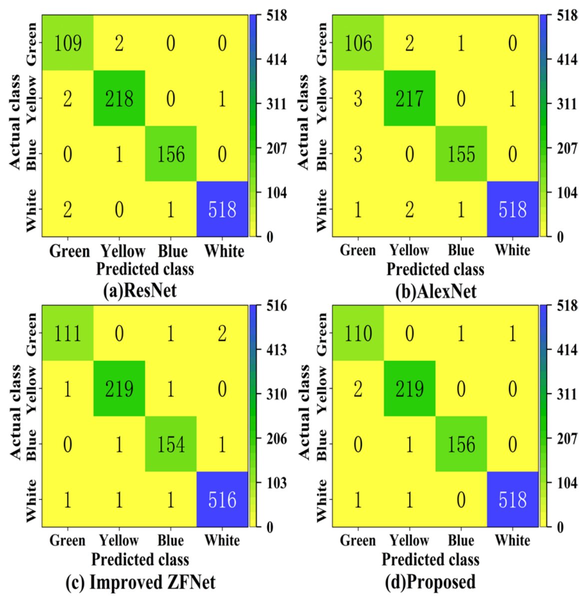

Figure 9. Classification results of the four methods

Based on Figure 9, it can be inferred that the suggested method has 99.2% correct overall recognition classification, ResNet has 99.0% overall recognition correct classification accuracy, AlexNet has 98.3% overall recognition correct classification, and Improved ZFNet has 99.0% overall recognition correct classification. The suggested model performs better than the others and is 0.9% better than the lowest AlexNet. All things considered, the suggested approach for license plate identification has a high recognition classification accuracy.

We randomly selected 4 images from the test set using the proposed method for testing, and the results are shown in Figure 10. All predictions were correct, and the recognition probability was above 85%.

(a)Test blue card (b)Test green card (c)Test yellow card (d)Test white card

Figure 10. The predicted results on our own test dataset

The violin plot can provide a more intuitive representation of data such as concentration, dispersion density, distribution status, etc., which helps guide model optimization and improvement. In this paper, the violin is used to show the probability distribution graph of the proposed method for correct license plate recognition in Figure 11.

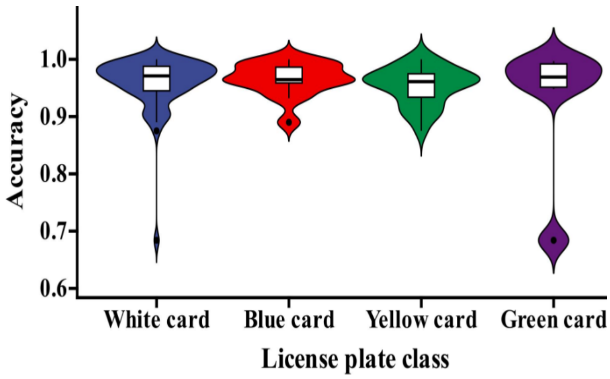

Figure 11. The distribution of violin probability for correct license plate recognition

The suggested model in this paper has a probability of accurately identifying the license plate of over 65%, as shown in Figure 11. When compared to the White card, which has a probability distribution between 0.65 and 0.7, the Green card has the greatest median recognition probability of 99.7%. In this interval, the green card has no meaning, meaning that its overall probability distribution is marginally higher than the white card's. Overall, the model presented in this work performs well in terms of the distribution of license plate recognition probabilities, with no notable disparity.

We collected the literature on license plate recognition for the last four years. The results are interpreted in Table 3.

As shown in Table 3, the recognition accuracy of the method proposed in this paper is 2.14% higher than that of the improved CRNN model constructed by Pan et al. Compared with the recognition accuracy of the better-performing MRNet network, the method proposed in this paper is slightly lower by 0.88%. Overall, the deep learning model constructed in this chapter has a slightly higher recognition accuracy than the type recognition, which can be satisfied for license plate recognition.

Table 3. Comparison between the proposed method and the recent literature

Ref., Year	Methods	Dataset size	Number of classes	Accuracy
Pan et al. [30], (2022)	Improved CRNN	6000	4	96.78%
Ke et al. [31], (2023)	MRNet	18000	8	99.80%
Vizcarra et al. [32], (2024）	OCR	1000	3	97.80%
Yang et al. [33], (2024)	SvRetina-LPD	3249	3	97.90%
Proposed, (2024)	Improved ZFNet	13275	4	98.92%

5 Conclusions and Outlook

In this work, we simplify the classic ZFNet to improve the feature extraction efficiency. Channel and spatial attention mechanisms are introduced to improve the richness of feature extraction for license plate character positions. The recognition speed and accuracy of the model are optimized by designing the residual network module. The proposed method achieves 98.92% average recognition accuracy and high FPS. Although this paper uses deep learning algorithms to achieve a high-precision license plate recognition task, there are still some deficiencies that need further research and improve. Possible future research is as follows.

(1) We currently only collect four common types of license plates, and their types and quantities are not comprehensive. In order to increase the accuracy and universality of the model, the future plan is to expand the data sample size to cover more license plate types from different countries and regions, with special attention to license plate data under different lighting, angles, and weather conditions, in order to improve the stability and adaptability of the system in various environments.

(2) The degree of damage, stains, and other elements that influence license plate recognition have not yet been examined by the existing system, which primarily concentrates on identifying the type and category of license plates. Deep learning technology will be used in the future to evaluate the readability and recognition difficulties of license plates, repair and improve license plates, and boost the system's capacity to identify damaged license plates.

(3) The proposed license plate recognition system can accurately recognize license plate numbers under common conditions, meeting the needs of application scenarios such as vehicle management and parking lots. However, the current system only supports files in static image format and does not support video files or real-time monitoring. In order to improve user experience, a system that supports video stream analysis will be developed in the future, allowing users to obtain license plate information through real-time monitoring and achieve automated intelligent recognition and storage.

(4) At present, license plate recognition systems are only deployed on personal computers and are not ported to other systems. In the future, we will develop specialized mobile applications that enable the system to perform license plate recognition on mobile devices. In addition, by integrating mobile phone cameras and augmented reality technology, the application scenarios of the system can be further expanded.

Acknowledgements

This work is supported by Henan Provincial Higher Education Teaching Reform Research and Practice Project under Grant (No. 2024SJGLX0475).

References

[1] Y. Gao, S. Mu, G. Xu, Toward Unified End-to-End License Plate Detection and Recognition for Variable Resolution Requirements, *IEEE Transactions on intelligent transportation systems*, Vol. 25, No. 9, pp. 10689-10701, September, 2024.

[2] C.-M. Pun, W.-Y. Ho, An Edge-Based Macao License Plate Recognition System, *International Journal of Computational Intelligence Systems*, Vol. 4, No. 2, pp. 244-254, April, 2011.

[3] H. R. Johnson, J. E. Scofield, B. Kostic, The effect of color on license plate recall, *Applied Ergonomics*, Vol. 118, Article No. 104279, July, 2024.

[4] S. Azizi, E. Asadi, S. Howard, B. W. Muir, R. O'Shea, A. Bab-Hadiashar, Autonomous Hyperspectral Characterization Stati: Robot Aided Measuring of Polymer Degradation, *IEEE Transactions on Automation Science and Engineering*, pp. 1-12, June, 2024.

[5] R. Islam, M. Islam, K. Talukder, An efficient ROI detection algorithm for Bangla text extraction and recognition from natural scene images, *Journal of King Saud University-computer and information sciences*, Vol. 34, No. 8, pp. 6150-6164, September, 2022.

[6] J. Shen, Y. Xia, H. Ding, W. Cabrel, Smart Parking Locks Based on Extended UNET-GWO-SVM Algorithm, *Sensors*, Vol. 23, No. 20, Article No. 8572, October, 2023.

[7] Y. Wen, Y. Lu, J. Yan, Z. Zhou, K. M. von Deneen, P. Shi, An Algorithm for License Plate Recognition Applied to Intelligent Transportation System, *IEEE Transactions on Intelligent Transportation Systems*,Vol. 12, No. 3, pp. 830-845, September, 2011.

[8] A. Zafar, M. Aamir, N. Nawi, A. Arshad, S. Riaz, A.Alruban, A. K. Dutta, S. Almotairi, A Comparison of Pooling Methods for Convolutional Neural Networks, *Applied Sciences-Basel*, Vol. 12, No. 17, Article No. 8643, September, 2022.

[9] R. Bhandari, S. Singla, P. Sharma, S. S. Kang, AINIS: An Intelligent Network Intrusion System, *International Journal of Performability Engineering*, Vol. 20, No. 1, pp. 24-31, January, 2024.

[10] M. Chawla, M. Pareek, Hybridizing Intelligence: A Comparative Study of Machine Learning Algorithm and ANN-PSO Deep Learning Model for Software Effort Estimation, *International Journal of Performability Engineering*, Vol. 20, No. 11, pp. 668-675, November, 2024.

[11] Z. Rao, D. Yang, N. Chen, J. Liu, License plate recognition system in unconstrained scenes via a new image correction scheme and improved CRNN, *Expert Systems with Applications*, Vol. 243, Article No. 122878, June, 2024.

[12] C. Lin, C. Chuang, H. Lin, Edge-AI-Based Real-Time Automated License Plate Recognition System, *Applied Sciences-Basel*, Vol. 12, No. 3, Article No. 1445, February, 2022.

[13] D. Habeeb, F. Noman, A. Alkahtani, Y. Alsariera, G. Alkawsi, Y. Fazea, A. M. Al-jubari, Deep-Learning-Based Approach for Iraqi and Malaysian Vehicle License Plate Recognition, *Computational Intelligence and Neuroscience*, Vol. 2021, Article No. 3971834, November, 2021.

[14] C. Tung, C. Wang, B. Peng, A Deep Learning Model of Dual-Stage License Plate Recognition Applicable to the Data Processing Industry, *Mathematical Problems in Engineering*, Vol. 2021, Article No. 3723715, November, 2021.

[15] T. Saidani, Y. Touati, A vehicle plate recognition system based on deep learning algorithms, *Multimedia Tools and Applications*, Vol. 80, No. 30, pp. 36237-36248, December, 2021.

[16] T. Shahwar, F. Mallek, A. Rehman, M. T. Sadiq, H. Hamam, Classification of Pneumonia via a Hybrid ZFNet-Quantum Neural Network Using a Chest X-ray Dataset, *Current Medical Imaging*, Vol. 20, Article No. e15734056317489, August, 2024.

[17] G. Corneto, F. Silva, D. Pereira, L. Almeida, A. Artero, J. Papa, V. H. C. de Albuquerque, H. M. Sapia, A New Method for Automatic Vehicle License Plate Detection, *IEEE Latin America Transactions*, Vol. 15, No. 1, pp. 75-80, January, 2017.

[18] W. Zhu, M. Xie, J.-F. Xie, A decision tree algorithm for license plate recognition based on bagging, *2012 International Conference on Wavelet Active Media Technology and Information Processing (ICWAMTIP)*, Chengdu, China, 2012, pp. 136-139.

[19] K. Hwan, L. Jun-Hong, License Plate Detection and Recognition Algorithm using Deep Learning, *Journal of IKEEE*, Vol. 23, No. 2, pp. 642-651, August, 2023.

[20] E. Huallpa, A. Macalupu, J. Luque, J. Sánchez-Garces, Determinación del mejor algoritmo de detección de matrículas en ambientes controlados y no controlados, *Risti - Revista Ibérica De Sistemas Tecnologias De Informação*, Vol. 49, pp. 83-99, March, 2023.

[21] R. Islam, M. Islam, K. Talukder, An efficient method for extraction and recognition of bangla characters from vehicle license plates, *Multimedia Tools and Applications*, Vol. 79, No. 27-28, pp. 20107-20132, July, 2020.

[22] P. Wang, Z. Wu, S. Zhang, H. Zhang, A GPU-free license plate detection based on fused color-edge and Retina approach, *Multimedia Tools and Applications*, Vol. 83, No. 7, pp. 18649-18666, February, 2024.

[23] F. Sultan, K. Khan, Y. Shah, M. Shahzad, U. Khan, Z. Mahmood, Towards Automatic License Plate Recognition in Challenging Conditions, *Applied Sciences-Basel*, Vol. 13, No. 6, Article No. 3956, March, 2023.

[24] D. Avianto, A. Harjoko, Afiahayati, CNN-Based Classification for Highly Similar Vehicle Model Using Multi-task learning, *Journal of Imaging*, Vol. 8, No. 11, Article No. 293, November, 2022.

[25] R. Castro-Zunti, J. Yepez, S. Ko, License plate segmentation and recognition system using deep learning and OpenVINO, *IET Intelligent Transport Systems*, Vol. 14, No. 2, pp. 119-126, February, 2020.

[26] S. Khokhar, D. Kedia, Integrating YOLOv8 and CSPBottleneck based CNN for enhanced license plate

character recognition, *Journal of Real-time Image Processing*, Vol. 21, No. 5, Article No. 168, October, 2024.
[27] Q. Huang, Z. Cai, T. Lan, A New Approach for Character Recognition of Multi-Style Vehicle License Plates, *IEEE Transactions on Multimedia*, Vol. 23, pp. 3768-3777, 2021.
[28] T. Aqaileh, F. Alkhateeb, Automatic Jordanian License Plate Detection and Recognition System Using Deep Learning Techniques, *Journal of Imaging*, Vol. 9, No. 10, Article No. 201, October, 2023.
[29] L. Wang, C. Cao, B. Zou, J. Ye, J. Zhang, License Plate Recognition via Attention Mechanism, *CMC-Computers Materials & Continua*, Vol. 75, No. 1, pp. 1801-1814, February, 2023.
[30] X. Pan, S. Li, R. Li, N. Sun, A Hybrid Deep Learning Algorithm for the License Plate Detection and Recognition in Vehicle-to-Vehicle Communications, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 12, pp. 23447-23458, December, 2022.
[31] X. Ke, G. Zeng, W. Guo, An Ultra-Fast Automatic License Plate Recognition Approach for Unconstrained Scenarios, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 24, No. 5, pp. 5172-5185, May, 2023.
[32] C. Vizcarra, S. Alhamed, A. Algosaibi, M. Alnaeem, A. Aldalbahi, N. Aljaafari, A. Sawalmeh, M. Nazzal, A. Khreishah, A. Alhumam, M. Anan, Deep learning adversarial attacks and defenses on license plate recognition system, *Cluster Computing-the Journal of Networks, Software Tools and Applications*, Vol. 27, No. 8, pp. 11627-11644, November, 2024.
[33] S. Yang, Z. Zhang, W. Shi, B. Feng, Y. Hua, J. Lei, SvRetina-LPD: A Sliding Vertex-Based RetinaNet for Robust Multi-Oriented License Plate Detection, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 25, No. 11, pp. 16421-16434, November, 2024.

Biographies

Xinhong Wei received the B.S. and M.S. degree from Naval Aeronautical University, Yantai, China, in 2001, 2004 respectively. She is currently an associate professor at Henan University of Urban Construction. Her research interests include computer vision, smart city, and big data analysis.

Hongyun Song, Male, born in April 2007. He is currently pursuing the B.S. degree in internet of things engineering at Henan University of Urban Construction, China. He mainly engaged in the research on deep learning algorithm optimization and license plate detection technology.

Yibu Chang, Female, born in April 2006. She is currently pursuing the B.S. degree in internet of things engineering at Henan University of Urban Construction, China. She mainly engaged in the research on deep learning algorithm optimization and license plate detection technology.

Wenxin Jiang, Male, born in April 2004. He is currently pursuing the B.S. degree in internet of things engineering at Henan University of Urban Construction, China. He mainly engaged in the research on deep learning algorithm optimization and license plate detection technology.

Laixiang Xu received the Ph.D. degree from Hainan University in 2024. He is currently working in Henan University of Urban Construction, China. He published over 10 peer reviewed international journal papers. His research interests mainly include artificial intelligence and image processing.