

Logo recognition of vehicles based on deep convolutional generative adversarial networks

Huan Ma¹, Yunfei Han²

¹School of Information Management, Xinjiang University of Finance and Economics, Urumqi, 830012, China

²The Xinjiang Technical Institute of Physics and Chemistry Chinese Academy of Sciences, Urumqi, 830000, China

²Corresponding author

E-mail: ¹23069082@qq.com, ²hanyf@ms.xjb.ac.cn

Received 5 December 2023; accepted 21 April 2024; published online 23 May 2024
DOI <https://doi.org/10.21595/jme.2024.23849>



Copyright © 2024 Huan Ma, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. Vehicle logo recognition plays a critical role in enhancing the efficiency of intelligent transportation systems by enabling accurate vehicle identification and tracking. Despite advancements in image recognition technologies, accurately detecting and classifying vehicle logos in diverse and dynamically changing environments remains a significant challenge. This research introduces an innovative approach utilizing a Deep Convolutional Generative Adversarial Network (DCGAN) framework, tailored specifically for the complex task of vehicle logo recognition. Unlike traditional methods, which heavily rely on manual feature extraction and pre-defined image processing techniques, our method employs a novel DCGAN architecture. This architecture automatically learns the distinctive features of vehicle logos directly from data, enabling more robust and accurate recognition across various conditions. Furthermore, we propose a refined training strategy for both the generator and discriminator components of our DCGAN, optimized through extensive experimentation, to enhance the model's ability to generate high-fidelity vehicle logo images for improved training efficacy. The technical core of our approach lies in the strategic integration of transfer learning techniques. These techniques significantly boost classification accuracy by leveraging pre-learned features from vast image datasets, thereby addressing the challenge of limited labeled data in the vehicle logo domain. Our experimental results demonstrate a substantial improvement in logo detection and classification accuracy, achieving an Intersection over Union (IoU) ratio of 42.67 % and a classification accuracy of 99.78 %, which markedly surpasses the performance of existing methods. This research not only advances the field of vehicle logo recognition but also contributes to the broader domain of measurement science and technology, offering a technically sound and logically coherent solution to a complex problem.

Keywords: vehicle logo detection, vehicle logo recognition, generative adversarial network, convolutional neural network.

1. Introduction

In the realm of measurement science and technology, precise recognition of vehicle logos plays a pivotal role in enhancing the capabilities of intelligent transportation systems. This recognition task, grounded in the principles of measurement science, involves accurately identifying the visual emblems that represent various automotive brands as a crucial means of vehicle identification. The information provided by vehicle logos serves various important roles in intelligent transportation. In intelligent transport systems, the identification of vehicle logos enables vehicle classification and management, tracking, and monitoring of specific vehicles, as well as supporting law enforcement and inspections. It can also aid in detecting illegal modifications and counterfeit vehicles, enhancing road safety, and providing a basis for traffic control and scheduling. While significant advancements have been made in vehicle logo recognition, a gap remains in effectively recognizing logos under varied lighting conditions and perspectives, a

challenge not fully addressed by existing solutions.

Despite the considerable progress in vehicle logo recognition technologies, a significant research gap remains in addressing the complexity of vehicle logos, which often combine text, images, and other elements, without a fixed position. This complexity is further exacerbated in real-world scenarios by varying lighting conditions, perspectives, and partial occlusions, areas that current methodologies struggle to effectively address. In high-resolution vehicle images, the logo area is a small target with limited structural and textural data. The imaging quality is vulnerable to degradation due to factors such as intense illumination, reflective surfaces, or casting of shadows, leading to potential issues of blur, distortion, and visual occlusion. Vehicle logo recognition today typically involves localization and classification steps. As a fundamental feature of vehicles, the vehicle logo provides crucial evidence for vehicle matching and identification. This has immense potential for application in intelligent transportation, criminal investigation, urban computing, social computing, and other fields.

Detecting and locating vehicle logos can prove to be a challenging task. Consequently, previous research has depended on the knowledge of the vehicle registration plate's position to roughly ascertain the location of the vehicle logo. Subsequently, diverse image processing techniques such as morphology processing, edge detection, and candidate region screening are employed to pinpoint the vehicle logo accurately. Machine learning methods are widely used in the classification of vehicle logos. Machine learning classifiers, including Adaboost, SVM, and multi-layer neural networks, are applied to categorise vehicle logos based on the gap between the sample features. Additionally, image features such as edge features, local gradient histogram features, HOG (Histogram of Oriented Gradient), LBP (Local Binary Pattern), or SIFT (Scale-invariant Feature Transform) are extracted for the same purpose.

Although deep learning approaches, including deep convolutional neural networks [1-6], recurrent neural networks [7], and generative adversarial networks [8, 9], residual neural networks [9], have shown promise in vehicle logo recognition, there is a notable research gap in the adaptability of these models to the dynamic conditions of real-world scenarios, such as variations in lighting, angle, and logo occlusion [10-14]. Compared with traditional manual features, the image features automatically obtained by convolutional neural networks are more similar to human biological vision features. Its features can effectively solve various common affine transformations in images, such as scaling, translation, rotation. Moreover, deep learning-based image recognition is considerably more accurate and efficient [15], and deep learning technology offers clear advantages in vehicle logo recognition.

This article presents a deep-learning approach to identifying vehicle logos. The approach initially identifies the vehicle's logo by leveraging prior information about its license plate position before utilizing a selective search candidate region algorithm to propose various regions. In contrast to the sliding window method, this technique significantly reduces the data volume, resulting in a faster vehicle logo detection that can recognize multi-scale vehicle logos. Next, the generator G and discriminator D in the adversarial generation network DCGAN [16] are trained on the vehicle logos dataset, and Nash equilibrium is achieved through the zero-sum game process of G and D . This gives the discriminator D the ability to judge whether an image is a vehicle logo image, and the highest confidence is obtained as the vehicle logo candidate region. Finally, the learned vehicle logo features in discriminator D are used for transfer learning, and the pre-trained D network is fine-tuned by connecting three fully connected network layers for vehicle logo classification tasks.

The contributions are listed as follows: 1) We propose a method to locate the vehicle logo roughly based on prior knowledge of the license plate position; 2) We use generative adversarial neural networks to generate a large number of random samples and enrich the diversity features of vehicle logos; 3) We use parameter transfer learning to transfer the discriminator parameters in the network and fine tune them to construct a vehicle logo classifier, enhancing the performance of classification; 4) On the test dataset, our proposed method achieved an average IoU result of 42.67% for vehicle logo detection and an average accuracy of 99.78% for vehicle logo

classification.

2. Related work

2.1. Logo recognition based on artificial features

Until now, many scholars have conducted research on vehicle logo recognition based on artificial features. Sotheeswaran and Ramanan [17] employed a coarse-to-fine strategy in their study on vehicle logo recognition. They first detected the vehicle region, followed by the grille region and logo region. By focusing on local features of structural features, they achieved a vehicle standard accuracy of 86.3 % through the analysis of 25 unique elliptical vehicle logos, each containing 10 images. Geng et al. [18] utilized the local stability of SIFT features and classified the extracted 218-bit SIFT features of vehicle logos using BP neural networks. This approach led to recognition accuracies of over 90 % for both simple and complex vehicle logos. Li and Yu [19] adopted a coarse extraction method for locating vehicle logo regions based on the license plate. They then used edge detection, horizontal projection, and vertical projection techniques to fine-tune the logo position. CSLBP and LBP features were extracted, and SVM and BP neural networks were used for classification. The final result was obtained through weighted voting fusion, achieving an accuracy of 96 %. Hoang V. T. [20] proposed a hybrid approach for vehicle logo recognition, combining feature selection methods with Histograms of Oriented Gradient descriptors. Irrelevant information was reduced through feature selection, leading to improved efficiency. The method was evaluated on the VLR-40 database, demonstrating its effectiveness. Qu et al. [21] presented a joint feature extraction and recognition method for vehicle logos. They utilized joint HOG-LBP features and support vector machine classification to achieve a high recognition rate. Zhang [22] introduced an incremental learning algorithm based on a twin support vector machine for vehicle logo recognition. The algorithm incorporated feature detection and affine transformation for the precise positioning of vehicle logos. HOG features were then extracted, and TSVM incremental learning was performed using inverse matrix operations. Experimental results on the vehicle logo dataset revealed a recognition rate of 91.77 % for the proposed algorithm. As we known, combining traditional manual features with classifiers can enable the classification of vehicle logos. Nevertheless, manual features pose challenges in their selection, stability, and computational requirements, leading to several shortcomings in real-world applications.

2.2. Logo recognition based on deep learning

With the development of deep learning theory and practice, vehicle logo recognition based on deep convolutional neural networks has also attracted many scholars' attention. Huang et al. [23], Bianco et al. [2], and Soon et al. [3] have explored the integration of deep convolutional neural networks, such as Faster-RCNN, VGG-16, and ResNet-50, in vehicle logo recognition, demonstrating the robustness and efficiency of deep learning approaches in handling complex image data and enhancing recognition accuracy. In their experiments, a better mean average precision result of 94.33 % is achieved, which shows that the methods have good robustness. Ke and Du [24] proposed three augmentation strategies for vehicle logo data: the cross-sliding segmentation method, the small frame method, and the Gaussian Distribution Segmentation method. The F1 value of our method in the YOLO framework is 0.7765, and the precision is greatly improved to 0.9295. In the Faster R-CNN framework, the F1 value of our method is 0.7799, which is also better than before. The results of experiments show that the above optimization methods can better represent the features of the vehicle logos than the traditional method. Zhang et al. [25] proposed a new approach called a multi-scale vehicle logo detector, which is based on SSD. This method obtains better results than the current detection methods by setting the parameters of the preset boxes, changing the pre-training strategy, and adjusting the

network structure. Experiments show that the proposed approach is better for multi-scale vehicle logo detection. Yang et al. [26] made further analysis work toward vehicle logo recognition and detection in real-world situations. To begin with, they proposed a new multi-class VLD dataset, called VLD-45 (Vehicle Logo Dataset), which contains 45000 images and 50359 objects from 45 categories respectively. Then, they used 6 existing classifiers and 6 detectors based on deep learning to evaluate their dataset and established new baseline performance. Song et al. [27] proposed a vehicle logo detection network called YOLO-T. It integrates multiple receptive fields and establishes a multi-scale detection structure suitable for VLD tasks. In addition, they design an effective pre-training strategy to improve the detection accuracy of YOLO-T. The experimental results show that the proposed method achieves high detection accuracy and outperforms the existing vehicle logo detection methods. Compared with traditional manual features, deep learning methods can effectively solve various common affine transformations in images, such as scaling, translation, rotation, and so on. Vehicle logo recognition based on deep learning is more accurate and efficient, and deep learning technology has obvious advantages in vehicle logo recognition.

3. Methods

3.1. Overview

The framework of our methods can be divided into 3 parts: proposal, recognition, and classification. Before applying our proposed methodology, we undertook a comprehensive data collection and preprocessing phase. The cameras used in real environments have a resolution of 720P or higher, support a shooting speed of 30 frames per second, work well in temperatures ranging from $-10\text{ }^{\circ}\text{C}$ to $+50\text{ }^{\circ}\text{C}$, and are capable of automatically adjusting the exposure time according to the light conditions to keep the image brightness appropriate. The cameras are usually deployed in front of the left or right side of the pass port. We collected vehicle logo images from various sources, ensuring a diverse representation of lighting conditions, angles, and resolutions. Each image was then preprocessed to normalize the size, adjust the contrast, and remove noise. Specifically, image preprocessing involved resizing all images to a uniform dimension of 224×224 pixels, applying grayscale conversion to enhance contrast, and employing Gaussian blur for noise reduction. Then, we apply appropriate knowledge rules to preliminarily filter the vehicle logo regions. In the recognition part, we train a DCGAN on the vehicle logo dataset, where G generates rich vehicle logo samples, and D learns to recognize vehicle logo images to identify the most likely vehicle logo region from numerous candidate regions. Finally, in the classification part, we transfer the knowledge of vehicle logo features learned in D to perform the final classification. Fig. 1 demonstrates the framework of our method.

Before feeding the vehicle logo images into our DCGAN model for training, a series of preprocessing steps were undertaken to ensure the data quality and consistency. This preprocessing involved resizing images to a uniform dimension of 224×224 pixels to match the input size of our model, converting images to grayscale to reduce computational complexity while retaining essential features, and applying a Gaussian blur filter to reduce image noise and improve the model's focus on relevant features. The standardized preprocessing pipeline is crucial for the effective training of deep learning models and contributes significantly to the robustness of the vehicle logo recognition system.

Our DCGAN model architecture for vehicle logo recognition is composed of a generator (G) that synthesizes new vehicle logo images from a noise distribution and a discriminator (D) that evaluates the authenticity of both real and generated images. The generator utilizes deconvolutional layers to progressively upscale the input noise to the size of a vehicle logo image, incorporating batch normalization and ReLU activation to stabilize the training process. Conversely, the discriminator employs convolutional layers to downscale the input images to a single probability score, using LeakyReLU activation to avoid the vanishing gradient problem. We optimized the model using the Adam optimizer, with a learning rate of 0.0002 for both G and

D, and a beta parameter of 0.5. Training was conducted in batches of 128 images over 200 epochs, ensuring the model's capability to generate high-quality vehicle logo images and accurately classify them.

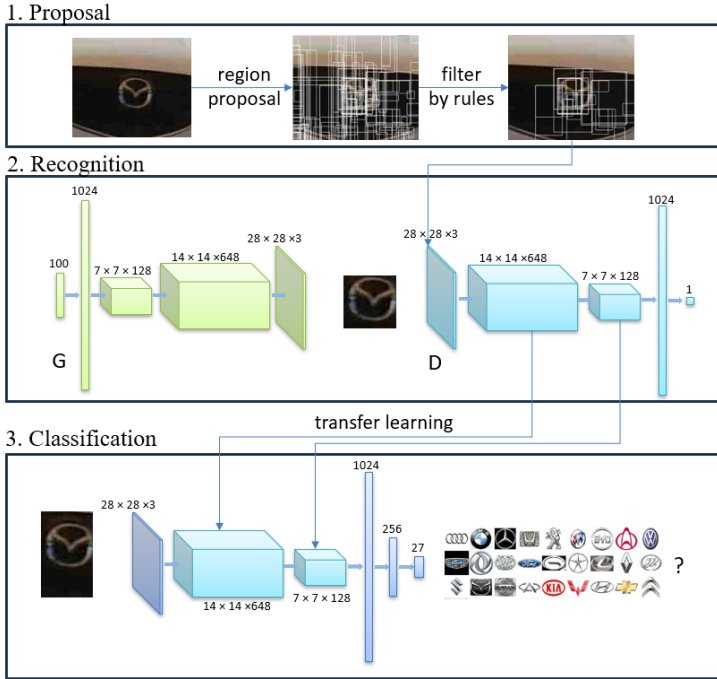


Fig. 1. The framework of our methods

3.2. Region proposals for vehicle logo

The framework of our methods can be divided into 3 parts: proposal, recognition, and classification. In the proposal part, we use prior knowledge of the license plate position to roughly determine the vehicle logo location and generate multiple possible regions via a region proposal algorithm. Based on prior knowledge of license plate position, coarse localization of the license plate area is performed. Assuming the height and width of the license plate are H , W , and the upper left coordinate position is x , y , then the coarse localization of the vehicle logo area is $[y - 3H : y, x : x + W]$. The selective search candidate region algorithm is used to recommend candidate regions for vehicle logos, and multiple candidate regions that may be vehicle logos are recommended. The selective search [28] candidate region algorithm first obtains n pre-segmented regions on the image based on the image segmentation algorithm in reference, and then calculates the similarity of each adjacent region to obtain a similarity matrix. In the selective search algorithm, an initial similarity matrix is constructed to record the similarity scores between every pair of adjacent regions. This matrix is pivotal for the algorithm's iterative process, where, in each iteration, the two regions exhibiting the highest similarity according to the matrix are merged. Consequently, the total count of regions decreases by one ($n - 1$), and the matrix is updated to reflect the new similarities between the merged region and its neighbors. Repeat this process until finally merging into one region. In terms of similarity measurement, the selective search algorithm adopts a weighted sum of multiple metrics, which integrates four metrics: color similarity, texture similarity, size similarity, and shape compatibility measurement of regions.

In terms of color similarity calculation, this algorithm uses 8 different color space transformations, fully considering the color similarity of different scenes and lighting conditions. Obtain histograms on three color channels of the image, with 25 histogram statistical features on

each color channel. Each region can generate a total of 75-dimensional feature vectors, denoted as $\{c_i^1, \dots, c_i^n\}$. The color similarity calculation between regions is shown in Eq. (1):

$$s_{colour}(r_i, r_j) = \sum_{k=1}^n \min(c_i^k, c_j^k). \quad (1)$$

In terms of texture similarity calculation, SIFT-Like feature is used to calculate Gaussian differentiation for 8 different directions of each color channel when variance is 1, obtain 10 histogram statistical features in each direction, generate a feature vector of 240 dimensions, denoted as $\{t_i^1, \dots, t_i^n\}$, and the texture similarity between regions is shown in Eq. (2):

$$s_{texture}(r_i, r_j) = \sum_{k=1}^n \min(t_i^k, t_j^k). \quad (2)$$

In order to ensure that regions do not engulf surrounding areas during merging, priority should be given to merging smaller regions, giving them greater weight, and ensuring multi-scale merging of images at each position. The calculation formula for size similarity is shown in Eq. (3):

$$s_{size}(r_i, r_j) = 1 - \frac{size(r_i) + size(r_j)}{size(im)}. \quad (3)$$

The inclusion relationship between two regions can help with more reasonable regions merging. The calculation is shown in Eq. (4), where BB_{ij} represents the bounding box of two regions. If two regions have inclusions, they are considered to be included. If the distance between the two regions is large, merging should not be considered:

$$s_{fill}(r_i, r_j) = 1 - \frac{size(BB_{ij}) - size(r_i) - size(r_j)}{size(im)}. \quad (4)$$

Taking into account the above four similarity measures, the comprehensive similarity calculation is shown in Eq. (5), where $a_i \in \{0,1\}$:

$$s(r_i, r_j) = a_1 s_{colour}(r_i, r_j) + a_2 s_{texture}(r_i, r_j) + a_3 s_{size}(r_i, r_j) + a_4 s_{fill}(r_i, r_j). \quad (5)$$

3.3. Identification of vehicle logo regions based on DCGAN

DCGAN, as demonstrated in vehicle logo recognition systems [5], is a very successful attempt to combine CNN and GAN for generating realistic images [29] and improving recognition tasks [9, 30, 31]. In addition, CNN is introduced into G and D, and the structure of CNN is also adjusted accordingly to improve the quality of samples and the speed of convergence. In constructing our DCGAN model, the generator G employs upsampling techniques to efficiently generate higher-resolution images from low-dimensional latent spaces. This choice avoids the checkerboard artifacts commonly associated with deconvolution operations. Batch normalization is used in both G and D networks to make data distribution more centralized, making model learning more stable and efficient. The fully connected network layers are deleted, and the convolutional layers are directly used to connect the input and output of generator G and discriminator D, making G and D become the full convolutional neural network. In the generator G of our DCGAN model, the Rectified Linear Unit (ReLU) is employed as the activation function for all layers except the last, facilitating a faster and more efficient training process by allowing only positive values to pass and blocking negative values. The hyperbolic tangent function (tanh)

is utilized in the final layer to ensure that the generated images' pixel values fall within the range of $[-1, 1]$, aligning with the preprocessing normalization of input images. Conversely, the discriminator D adopts the Leaky Rectified Linear Unit (LeakyReLU) as its activation function. LeakyReLU allows a small, non-zero gradient when the unit is not active and the input is negative, preventing the dying ReLU problem and ensuring a more robust learning of discriminative features.

In the DCGAN network, the generator G is responsible for the reverse generation of realistic images from simple features. The generator uses convolutional networks and upsampling to generate realistic images layer by layer from randomly simple input data. Fig. 2 shows the network structure of the generator. On the other hand, discriminator D in the DCGAN network is responsible for discriminating generated images. It extracts features such as colors, points, lines and faces from images through convolutional layers to understand an image semantically. By doing this, the discriminator can determine whether the distribution of the generated image is similar to that of the vehicle logo image dataset. Fig. 3 illustrates the network structure of the discriminator, which has also been fine-tuned here.

The DCGAN is trained to generate vehicle logo images, and the discriminator D is used to determine whether the image is a vehicle logo image for all candidate regions. The candidate region with the highest score is set as the default vehicle logo region, and this region image is the vehicle logo image to be classified.

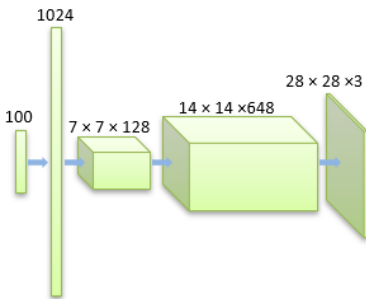


Fig. 2. Structure of fine-tuned generator

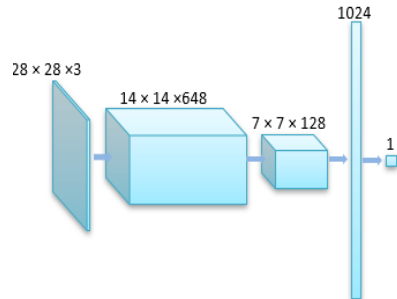


Fig. 3. Structure of fine-tuned discriminator

3.4. Logo classification based on discriminator parameter transfer

We propose to optimize the vehicle logo classification by transferring the discriminator D, which has already learned the features of the vehicle logo during DCGAN training and can support the discrimination of vehicle logo images. To this end, we remove the final fully connected layer from the discriminator D and connect three fully connected layers with 1024, 256 and 27 neurons, respectively. The first two layers are activated by ReLU, while the last layer employs the softmax activation function. This model is referred to as D3F. The network structure of D3F is depicted in Fig. 4. Finally, we divide the vehicle logos into 27 categories using parameter transfer [32, 33].

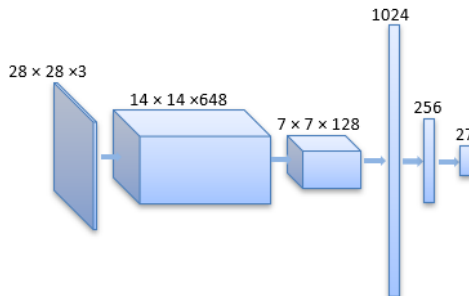


Fig. 4. Structure of D3F

4. Results and discussion

License plate location detection can be implemented using the open-source license plate recognition tool HyperLPR (<https://github.com/zeusees/HyperLPR>). The tool is based on a cascade classifier for multi-scale detection, where the license plate region is considered to be detected when it satisfies all of the classifiers in the cascade. If any classifier is not satisfied, the region is not considered a license plate region. However, the recognition of license plates in compressed images can be affected by the compression of the images, which in turn affects the detection effect of license plates. Based on the testing of 6368 samples of the vehicle dataset, 4408 license plate regions were obtained, resulting in a license plate detection rate of 69.22 %. The coarse localization area of the vehicle logo obtained through HyperLPR detection of license plate position can be observed in Fig. 5.



Fig. 5. Coarse logo region

Selective search candidate region algorithm produces a large number of candidate regions which are labeled with white rectangular boxes, as shown in Fig. 6. To reduce computational load, rule-based filtering is performed based on prior knowledge. Firstly, the length and width of the candidate area are limited in proportion to the relative size of the vehicle logo and the entire candidate area. The length and width of the candidate area should not be greater than half of the length and width of the entire candidate area. Secondly, the aspect ratio of the candidate areas for the vehicle logo should not exceed 3. Finally, candidate areas with less than 400 pixels and height or width less than 10 are filtered out. Fig. 7 shows the result of rule-based filtering on the vehicle logo candidate regions.



Fig. 6. Logo regions proposal generated by Selective Search

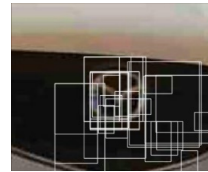


Fig. 7. Logo regions proposal filtered by rules

Manual labeling was performed based on candidate regions of vehicle logo to obtain 27 types of vehicle logos, as shown in Table 1. The number of samples for each type of vehicle logo varies from about 20 to 40. The standard legends corresponding to each vehicle logo are shown in Fig. 8, arranged in the above order from top left to bottom right.

Table 1. Logos

Name	Name	Name	Name	Name	Name	Name	Name	Name
Audi	BMW	Benz	Honda	Peugeot	Buick	BYD	Chang'an	Volkswagen
Geely Emperor Hao	Dongfeng	Toyota	Ford	GAC Trumpchi	Jianghuai	Lexus	Reynolds	Lifan
Suzuki	Mazda	Nissan	Chery	KIA	Wuling	Hyundai	Chevrolet	Citroën

In our experiments, we utilized a dataset comprising 27 different vehicle logo classes, with a total of 1200 samples for each class in the training set and 400 samples for each class in the

validation and testing sets. Using classification algorithms in machine learning can achieve good results in the research of vehicle logo classification. However, many works often suffer from relatively missing annotation data. Because machine learning classification algorithms require a large amount of annotated data, expanding data through image reduction, scaling, and rotation using data augmentation can effectively improve classification results. To expand data, the cutting range was set to -0.1 to 0.1 , and the scaling range to -0.1 to 0.1 . The rotation angle range was set to -5 to 5 , and the length and width to 64 . For pixels outside the boundary, the nearest pixel was used for interpolation. Finally, a balanced dataset containing 1200 samples for each type of vehicle logo in the training set and 400 samples for each type of vehicle logo in the validation and testing sets was obtained.



Fig. 8. Logo examples

The DCGAN generative adversarial network will first be trained to generate vehicle logo images. Before training, the image pixel values are normalized to the range of $[-1, 1]$, and the size is uniformly scaled to $28 \times 28 \times 3$. For the training of both the generator G and discriminator D in our DCGAN model, we employ stochastic gradient descent (SGD) with mini-batches of size 128 . Importantly, we set the momentum parameter to 0.9 , a decision based on empirical evidence suggesting that a higher momentum helps in accelerating the SGD convergence towards the optimal solution by integrating past gradients' direction into the current update. This technique effectively reduces oscillations during training, leading to a smoother and faster convergence process. The learning rate is 0.0005 , and the logarithmic loss function is used for the loss function. During the training process, it begins with training the discriminator D . After marking the real data samples in the batch as 1 and marking the generated samples based on random input as 0 , the discriminator D was trained to be able to accurately identify the real and fake samples. This allows D to be trained and updated, after which D is marked as untrained and the generator G is then trained to generate samples which match the distribution of the real dataset as much as possible. After 100 epochs with 296 batches per epoch of training, the discriminator D is used to distinguish the candidate regions of the vehicle logo filtered by the rules, and the most suitable vehicle logo is obtained after discriminant processing. Moreover, in our study, we manually annotated a dataset comprising 141 vehicle logo samples for the detection experiments. The performance of our proposed method was evaluated using the Intersection over Union (IoU) metric, a common measure for the accuracy of object detection models. The IoU calculates the percentage overlap between the detected logo area and the ground truth area. For our experiments, the IoU achieved was 42.67% . This result indicates a significant improvement over the baseline method, which utilizes image morphology transformation for vehicle logo detection, achieving an IoU of 33.42% . Thus, our method's performance was 9.25 percentage points higher.

After obtaining the vehicle logo area, the convolutional neural network based on parameter transfer is used to extract the vehicle logo features, and then the full connection layers are used for classification. In order to better identify vehicle logos in different environments, the data is further enhanced based on larger morphological changes, thereby improving the robustness of the model. To enhance the data, the cutting range was set to -0.2 to 0.2 , and the scaling range to -0.2 to 0.2 . The rotation angle range was set to -10 to 10 , and the length and width to 28 . During the training process, validation is conducted after each epoch, and a test is conducted after the training is completed.

Based on the trained discriminator model for parameter transfer, a comparative experiment was designed to verify that the parameter transfer can indeed optimize the classification model. In the comparative experiment, the D3F model without parameter transfer was used to classify the vehicle logo. The training process is shown in Fig. 9, it can be seen that the models were actively learning and updating parameters before 20 epochs. After 20 epochs, the classification accuracy reached a saturation state. The D3F model with parameter transfer has great initial results in the early training stage and is more efficient in approaching the optimal results than the model without parameter transfer.

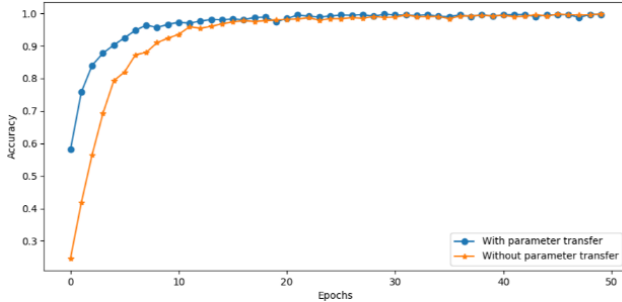


Fig. 9. Logo classification learning process comparison

Finally, the D3F model with parameter transfer was used to predict each vehicle logo, with a mean classification accuracy of 99.78 %. The classification results are shown in Table 2.

To ensure the robustness and reliability of our classification results presented in Table 2, we employed a comprehensive statistical analysis approach. This involved computing the mean accuracy across all logo classes to gauge the overall performance of our model. Furthermore, we performed variance analysis to assess the consistency of our model's performance across different vehicle logos. The statistical significance of our results was evaluated using a one-way ANOVA test, with a significance level set at $p < 0.05$, ensuring that the observed differences in accuracy among various vehicle logos were not due to random chance.

Table 2. Logo classification results

Name	Accuracy (%)	Name	Accuracy (%)
Audi	99.81	Jianghuai	99.91
BMW	99.99	Lexus	99.86
Benz	99.99	Reynolds	99.91
Honda	99.87	Lifan	99.84
Peugeot	99.83	Suzuki	99.54
Buick	99.87	Mazda	99.91
BYD	99.17	Nissan	99.87
Chang'an	98.97	Chery	99.90
Volkswagen	99.91	KIA	99.91
Geely Emperor Hao	99.80	Wuling	99.99
Dongfeng	99.20	Hyundai	99.67
Toyota	99.83	Chevrolet	99.89
Ford	99.84	Citroën	99.91
GAC Trumpchi	99.86	Avg	99.78

5. Conclusions

In this study, we have developed and implemented a deep convolutional generative adversarial network (DCGAN) tailored for the complex task of vehicle logo recognition in natural scenes. The intricacies associated with this task, ranging from the variability in perspective and lighting conditions to the small size and obscured nature of target logos, were systematically addressed

through our two-step approach: logo localization and classification. Initially, leveraging prior knowledge of license plate positions enabled the coarse localization of logo areas, which was then refined through the application of a DCGAN model, culminating in precise region recognition. Our innovative approach yielded a commendable average Intersection over Union (IoU) ratio of 42.67 % and a mean classification accuracy of 99.78 % across 27 distinct vehicle logo classes.

The notable advancements highlighted by these results emphasize the strength of our DCGAN model in surmounting the challenges inherent to vehicle logo recognition. The employment of transfer learning techniques and our model's optimization strategies have significantly bolstered both the robustness and the precision of logo identification, setting new benchmarks for performance in this domain.

Nevertheless, our exploration has illuminated certain limitations that warrant further investigation. The effectiveness of our model is contingent upon the availability of high-quality, well-annotated training data, a requirement that might not always be feasible in practical scenarios. The recognition system can approximate the same detection performance as during the day by increasing lighting equipment in low-light environments at night. However, the performance of our model under extreme conditions, such as rain, snowfall, or when logos are heavily occluded, remains inadequately tested, suggesting areas for future exploration.

Looking ahead, we propose several avenues for advancing the field of vehicle logo recognition. A critical area of future work involves the integration of advanced image enhancement techniques, which could significantly improve model performance under challenging environmental conditions. Additionally, exploring semi-supervised or unsupervised learning methodologies promises to reduce the dependency on large annotated datasets, potentially increasing the adaptability and efficiency of models for real-world application. By pursuing these directions, we aim to further refine vehicle logo recognition technologies, making them more resilient and versatile in the face of varying real-world conditions.

Acknowledgements

This research was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region of China [2020D01B55] and the West Light Foundation of The Chinese Academy of Sciences [2019-XBQNXZ-B-009].

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Huan Ma: data curation, formal analysis, investigation, methodology, software, validation, visualization, and writing – original draft preparation. Yunfei Han: conceptualization, funding acquisition, project administration, resources, and writing – review and editing.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, Vol. 60, No. 6, pp. 84–90, May 2017, <https://doi.org/10.1145/3065386>

- [2] S. Bianco, M. Buzzelli, D. Mazzini, and R. Schettini, "Deep learning for logo recognition," *Neurocomputing*, Vol. 245, pp. 23–30, 2017.
- [3] F. C. Soon, H. Y. Khaw, J. H. Chuah, and J. Kanesan, "Vehicle logo recognition using whitening transformation and deep learning," *Signal, Image and Video Processing*, Vol. 13, No. 1, pp. 111–119, Jul. 2018, <https://doi.org/10.1007/s11760-018-1335-4>
- [4] Y. Huang, R. Wu, Y. Sun, W. Wang, and X. Ding, "Vehicle logo recognition system based on convolutional neural networks with a pretraining strategy," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 4, pp. 1951–1960, Aug. 2015, <https://doi.org/10.1109/tits.2014.2387069>
- [5] J. Zhang, S. Yang, C. Bo, and Z. Zhang, "Vehicle logo detection based on deep convolutional networks," *Computers and Electrical Engineering*, Vol. 90, p. 107004, Mar. 2021, <https://doi.org/10.1016/j.compeleceng.2021.107004>
- [6] Alsheikhy, A., Said, Y., Barr, and M., "Logo recognition with the use of deep convolutional neural networks," *Engineering, Technology and Applied Science Research*, Vol. 10, No. 5, pp. 6191–6194, 2020.
- [7] P. Liu, X. Qiu, and X. Huang, "Recurrent neural network for text classification with multi-task learning," in *International Joint Conference on Artificial Intelligence*, Jan. 2016.
- [8] I. J. Goodfellow et al., "Generative adversarial networks," *arXiv:1406.2661*, Jan. 2014, <https://doi.org/10.48550/arxiv.1406.2661>
- [9] X. Wang, "Vehicle-logo recognition algorithm based on convolutional neural network," *Journal of Information and Computational Science*, Vol. 12, No. 18, pp. 6945–6952, Dec. 2015, <https://doi.org/10.12733/jics20150130>
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, <https://doi.org/10.1109/cvpr.2016.90>
- [11] C. Pany, U. K. Tripathy, and L. Misra, "Application of artificial neural network and autoregressive model in stream flow forecasting," *Journal of Indian Water Works Association*, Vol. 33, No. 1, pp. 61–68, 2001.
- [12] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, Vol. 335, pp. 327–335, 2019.
- [13] D. Sun, P. Chopra, J. Bhola, and R. Neware, "Computer communication network fault detection based on improved neural network algorithm," *Electrica*, Vol. 22, No. 3, pp. 351–357, 2022.
- [14] S. Tang, S. Yuan, and Y. Zhu, "Deep learning-based intelligent fault diagnosis methods toward rotating machinery," *IEEE Access*, Vol. 8, No. 1, pp. 9335–9346, Jan. 2020, <https://doi.org/10.1109/access.2019.2963092>
- [15] Y. Han, T. Jiang, Y. Ma, and C. Xu, "Pretraining convolutional neural networks for image-based vehicle classification," *Advances in Multimedia*, Vol. 2018, pp. 1–10, Oct. 2018, <https://doi.org/10.1155/2018/3138278>
- [16] Radford A. et al., "Unsupervised representation learning with deep convolutional generative adversarial networks," in *International Conference on Learning Representations*, 2016.
- [17] S. Sotheeswaran and A. Ramanan, "A coarse-to-fine strategy for vehicle logo recognition from frontal-view car images," *Pattern Recognition and Image Analysis*, Vol. 28, No. 1, pp. 142–154, Mar. 2018, <https://doi.org/10.1134/s1054661818010170>
- [18] Q. Geng, F. Yu, Y. Wang, H. Zhao, and D. Zhao, "Vehicle logo recognition algorithm based on SIFT," *Journal of Jilin University*, Vol. 56, No. 3, pp. 639–644, 2018.
- [19] Z. Li and M. Yu, "Vehicle logo recognition based on ensemble learning with multiple LBP features," *Computer Engineering and Applications*, Vol. 55, No. 20, pp. 134–138, 2019, <https://doi.org/10.3778/j.issn.1002-8331.1806-0330>
- [20] Hoang and V. T., "Vehicle logo recognition using HOG descriptor and sparsity score," *Telkommika (Telecommunication Computing Electronics and Control)*, Vol. 18, No. 6, 2020, <https://doi.org/10.12928/telkommika.v18i6.16133>
- [21] A. Qu et al., "Research on intelligent recognition method based on vehicle logo area," *Software Engineering*, Vol. 24, No. 10, pp. 37–40, 2021.
- [22] H. Zhang, "Vehicle logo recognition based on HOG feature and TSVM algorithm," *Information Technology*, Vol. 47, No. 2, pp. 185–190, 2023.

- [23] Z. Huang, M. Fu, K. Ni, H. Sun, and S. Sun, "Recognition of vehicle-logo based on faster-RCNN," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Singapore, 2018, pp. 75–83, https://doi.org/10.1007/978-981-13-1733-0_10
- [24] X. Ke and P. Du, "Vehicle logo recognition with small sample problem in complex scene based on data augmentation," *Mathematical Problems in Engineering*, Vol. 2020, pp. 1–10, Jul. 2020, <https://doi.org/10.1155/2020/6591873>
- [25] J. Zhang, L. Chen, C. Bo, and S. Yang, "Multi-scale vehicle logo detector," *Mobile Networks and Applications*, Vol. 26, No. 1, pp. 67–76, Feb. 2021, <https://doi.org/10.1007/s11036-020-01722-0>
- [26] S. Yang, C. Bo, J. Zhang, P. Gao, Y. Li, and S. Serikawa, "VLD-45: a big dataset for vehicle logo recognition and detection," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 12, pp. 25567–25573, Dec. 2022, <https://doi.org/10.1109/tits.2021.3062113>
- [27] L. Song, W. Min, L. Zhou, Q. Wang, and H. Zhao, "Vehicle logo recognition using spatial structure correlation and YOLO-T," *Sensors*, Vol. 23, No. 9, p. 4313, 2023.
- [28] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2014, <https://doi.org/10.1109/cvpr.2014.81>
- [29] Yunfei Han, Yi Wang, and Yupeng Ma, "Generative difference image for blind image quality assessment," in *International Conference on Computer Engineering and Artificial Intelligence (ICCEAI)*, 2021.
- [30] S. Mao, M. Ye, X. Li, F. Pang, and J. Zhou, "Rapid vehicle logo region detection based on information theory," *Computers and Electrical Engineering*, Vol. 39, No. 3, pp. 863–872, Apr. 2013, <https://doi.org/10.1016/j.compeleceng.2013.03.004>
- [31] S. Yang et al., "Fast vehicle logo detection in complex scenes," *Optics and Laser Technology*, Vol. 110, pp. 196–201, 2019.
- [32] L. Long and T. Liang, "Multi-distributed speech emotion recognition based on MEL frequency cepstogram and parameter transfer," *Chinese Journal of Electronics*, 2022, <https://doi.org/10.1049/cje.2020.00.080>
- [33] R. O. Ogundokun, R. Maskeliūnas, and R. Damaševičius, "Human posture detection using image augmentation and hyperparameter-optimized transfer learning algorithms," *Applied Sciences*, Vol. 12, No. 19, p. 10156, Oct. 2022, <https://doi.org/10.3390/app121910156>



Huan Ma received Master's degree from Xinjiang University, Urumqi, China, in 2006. Now she works at Xinjiang University of Finance and Economics. Her current research interests include e-commerce, data process and analysis.



Yunfei Han received Ph.D. degree in Xinjiang Technical Institute of Physics and Chemistry Chinese Academy of Sciences, Urumqi, China, in 2019. Now he works at The Xinjiang Technical Institute of Physics and Chemistry Chinese Academy of Sciences. His current research interests include computer vision and data mining.