Research on vehicle multi-classification object detection algorithm based on Ultralytics/YOLOv5 improvement

Yu Deng

Dalian Maritime University, 1 Linghai Road, Dalian, China

13508332846@163.com

Abstract. This paper investigates an improved vehicle multi-classification object detection algorithm based on Ultralytics/YOLOv5. By combining the YOLOv5 model with a custom dataset, training and testing were conducted on Bangladeshi road vehicle images. The main contributions of this paper include optimizing the hyperparameter settings of the model to adapt to the vehicle multi-classification problem and improving detection performance by adjusting input sizes and the number of categories. The training set contains 2704 images, the validation set contains 300 images, covering 21 classes of vehicles. Experimental results show that the improved algorithm performs well in terms of mAP, Precision, and Recall values, with an average mAP of 0.41 and Precision of 0.669 on the test set. The research results of this paper demonstrate that the improved YOLOv5 model exhibits high accuracy and robustness in vehicle detection tasks in complex traffic environments, providing strong support for intelligent traffic monitoring and autonomous driving.

Keywords: Vehicle Multi-Classification, Object Detection, Ultralytics/YOLOv5, Deep Learning, Real-Time Detection.

1 Introduction

Region-based object detection is an important issue in the field of computer vision, playing a crucial role in many applications such as intelligent surveillance, autonomous driving, medical image analysis, etc. In recent years, with the development of deep learning technology, object detection algorithms based on deep neural networks have made significant progress, among which region-based detection techniques are the most representative. Region-based object detection algorithms use region proposal generators (such as RPN - Region Proposal Network) to generate a series of candidate regions that may contain objects, and then classify and locate these candidate regions. This approach can significantly improve the accuracy and robustness of object detection. In recent years, deep learning-based object detection algorithms have evolved into two main technical routes: Anchor-based methods (two-stage, one-stage) and Anchor-free methods. The two-stage object detection algorithms in Anchor-Based mainly include RCNN, SPPNet, Fast RCNN, Faster RCNN, FPN, Cascade RCNN; one-stage object detection algorithms include YOLO v1, SSD, YOLO v2, RetinaNet, YOLO v3, YOLO v4, YOLO V5. Anchor-Free object detection algorithms include CornerNet, CenterNet, FSAF, FCOS, SAPD, and other mainstream algorithms.

CONF-CIAP 2025, January 17-23, Eskisehir, Turkey

DOI 10.4108/eai.17-1-2025.2355328

In practical applications such as intelligent surveillance, autonomous driving, vehicle traffic detection, smart industry, agricultural automation, etc., accurate detection and positioning of targets are crucial for the stability and reliability of the system. Moreover, region-based object detection algorithms have strong adaptability and can handle object detection problems in complex scenes. For example, in urban traffic monitoring, vehicles may appear in different sizes, angles, and degrees of occlusion. Region-based object detection algorithms can effectively cope with these challenges and improve the detection performance of the system. In maritime ship target detection, region-based object detection can monitor port shipping traffic, obtain ship deployment and dynamic information, and has important research value for effectively understanding the deployment status of maritime ships [1].

This paper will be divided into six main parts. The second part of the article will review classical literature, discussing region-based object detection algorithms and previous research applications. The third part of the article will introduce the algorithm used in this paper, as well as elaborating on the overall experimental process and specific model details (such as hyperparameters) from two dimensions: macro and micro. The fourth part of the article will present the results of experiments through figures and tables. The fifth part of the article will conduct discussions, analyzing specific experimental results, as well as the similarities and differences between this study and similar studies. Finally, the sixth part of the article will summarize, discussing the advantages, limitations, and future research directions of this study.

2 Literature Review

Object detection, as one of the core issues in the field of computer vision, plays a crucial role in various areas such as intelligent surveillance, autonomous driving, and medical image analysis. With the rapid development of deep learning technology, object detection algorithms based on deep neural networks have made significant progress.

There are already many excellent object detection algorithms applied to the detection of natural scenes. For example, Zhao and Yang proposed a lightweight YOLOv5 algorithm based on multi-scale pyramid and multi-scale attention, targeting problems such as small target sizes and complex backgrounds in remote sensing vehicle detection [2]. By reducing the number of downsampling layers and redesigning the multi-scale pyramid network, the detection capability of small objects and feature fusion ability were improved; introducing an improved multi-scale attention module enhanced perception and reduced model parameters; using the K-means++ clustering algorithm to design anchor box scales and aspect ratios suitable for targets. Compared to YOLOv5s on a self-built dataset, the algorithm achieved higher detection accuracy while reducing parameters and model size.

Lv and Jia proposed an improved object detection model GP-YOLOv5n, based on YOLOv5n and incorporating fusion attention into the depth separable neck network [3]. The model adopted Alpha-IoU in the bounding box regression loss function, improving the accuracy of bounding box positioning for objects. For wildlife detection, a lightweight giant panda detection model was designed. Experimental results showed that in complex environments, the model significantly improved the detection accuracy and speed for giant pandas.

Li et al. proposed an ultra-lightweight object detection network for detecting ships in aerial images. The model size of this algorithm is only 1.64MB, achieving a speed of 25fps on a mobile platform with 0.75T computing power while maintaining high accuracy [4]. By employing an extremely lightweight CNN backbone network, fused detection head design, and advanced data preprocessing methods, the network's robustness and capability to capture detailed features of targets were enhanced, thereby improving detection accuracy. The adoption of the tiled region of interest detection strategy provided a wider range of choices for the network's practical applications. The fusion of these strategies achieved a balance between accuracy and speed, and the algorithm's advancedness and practicality were verified on embedded platforms.

Chen et al. proposed a mural damaged area detection method called U^2-DUANet, using a nested U-shaped structure network [5]. The method introduced a Depth-supervised Aggregation (DUA) module to more effectively integrate detailed information from side outputs. Additionally, the Pixel-level Context and Channel Attention (PCCA) module were utilized to capture important features more accurately. Moreover, the use of Self-adaptive Normalization (SN) instead of traditional Batch Normalization (BN) enhanced the flexibility and generalization capability of the network. Finally, a novel attention module, PCCA, was proposed, which combined spatial and channel information of images to more accurately capture important features.

Yuan et al. proposed a drone object detection method based on region-adaptive thresholding [6]. Firstly, target suspected regions were identified using local threshold segmentation, then a clustering method was used to segment suspected regions into the YOLOv5 detection network to avoid loss of target features during global detection due to image compression. Finally, to address the problem of low confidence in small targets, an adaptive threshold based on target size was used to enhance the detection rate of small drones. This method improved the detection rate of small targets, avoided ROI repeated extraction through the region attribution division based on DBSCAN clustering, and effectively reduced the false alarm rate through the filtering method based on adaptive thresholding according to target size.

Chen et al. proposed a drug recommendation algorithm based on dialog structure and graph attention network to address the problem of existing algorithms' inability to reflect patients' real-time health needs [7]. Firstly, a correlation-aware structural graph was constructed by combining grey relational analysis with graph attention networks to better capture the intrinsic correlations between nodes. Then, each dialogue utterance was represented as a node of the graph attention network, and two types of relational structures were designed to learn the adjacency relationships between nodes. Next, dialogue hierarchical encoders and disease encoders were designed to learn dialogue structure representation and patient health problems. Finally, the two feature representations were fused into an MLP layer to achieve real-time prediction and recommendation of drugs. This method fully perceives the correlation of dialogue nodes, learns dialogue structure representation, and combines disease representation for drug prediction, improving the effectiveness of the recommendation algorithm.

Zhang and Yang proposed an improved facial expression recognition method based on Local Binary Pattern (LBP) and attention mechanism in the New Visual Geometry Group Network (NEW-VGG), aiming to improve training speed and recognition performance [8]. By integrating the LBP algorithm and the NEW-VGG model, the training speed of the model was accelerated. The VGG-16 network was improved to create the NEW-VGG model, and through ablation experiments, the effectiveness of global average pooling layer and attention mechanism in improving the speed and accuracy of the model was verified.

Hu et al. proposed an ACE-YOLO adaptive local image detection algorithm based on deep learning for fast and high-precision detection of apple defects [9]. This algorithm reduces the detection area through deep learning, utilizes channel attention mechanism to concentrate computing resources on local detection range, employs image enhancement algorithm to improve detail clarity, and adds a small target detection layer to enhance detection accuracy. This method effectively solves the difficulty of identifying small defects in apple detection, amplifies defect details through local image processing techniques, and avoids environmental interference. By improving the network structure and introducing channel attention mechanism, a balance between detection accuracy and speed was achieved.

Wu et al. addressed the issue of feature blurring in small target detection in aerial image object detection by proposing an improved YOLO_v5x algorithm [10]. By adding the SPD module and small target detection head, the loss of fine-grained information was reduced, effectively improving the efficiency of small target detection; the introduction of the CA attention mechanism and a new loss function significantly improved the positioning accuracy of small targets. Experimental results based on the Visdron dataset validated the effectiveness of this method.

Hua et al. proposed an improved target detection network for the inefficient detection of targets in vehicle-mounted infrared images [11]. Firstly, by adding a dynamic detection head based on attention mechanism, the network focused more on foreground targets, enhancing the expression ability of the detection head. Secondly, the MPDIOU was used to replace the CIOU bounding box loss function during training, improving the model's positioning accuracy and efficiency. Finally, the lightweight network FasterNet was added to the C3 module at the end of the neck network to further improve the real-time performance of the model.

Wang et al. proposed a controller's sleeping behavior recognition method based on a dual-stream adaptive graph convolutional network [12]. This method designs dual-stream networks to process first-order and second-order information of the controller's skeleton separately, achieving full extraction of skeleton data; by adaptively learning the skeleton's topological connectivity matrix, functional connection relationships between different joints of the controller are explored; meanwhile, a spatio-temporal channel attention mechanism is introduced into the convolutional layer to enhance the model's ability to extract important information in time, space, and channel directions for recognizing the controller's sleeping behavior.

Zou et al. proposed a spatio-temporal attention graph convolutional network model for dynamic traffic flow prediction, and based on the attention mechanism, established a spatio-temporal attention graph convolutional network model, considering the application of temporal and spatial correlations, dynamic spatial correlations, and external features [13]. The spatio-temporal attention graph convolutional network model for dynamic traffic flow prediction can effectively extract spatio-temporal features in the traffic network. Compared to other baseline models, the ATST-GCN model shows more stable prediction results for medium to long-term (40, 60 minutes) traffic flow.

Zhou et al. proposed a multi-branch joint network structure based on attention mechanism to improve the accuracy of pedestrian re-identification [14]. This method, based on the Full-scale Network (OSNet), introduced an attention mechanism module to refine the semantic representation of features. Meanwhile, a multi-branch joint network was employed to organically combine local and global features, enhancing the correlation between features. Batch feature erasure was utilized for data augmentation to reduce the influence of occlusion, and a multi-loss joint function was adopted to reinforce the model's supervised training. This approach embedded attention mechanism in the backbone network to obtain richer global features, while utilizing batch feature erasure and improved horizontal segmentation method in the branch network to increase the model's robustness and focus more on extracting local fine-grained features.

Xie and Zhang published a review article on graph convolutional neural networks [15]. The article categorized graph convolutional neural networks into two major categories: graph-based methods and spatial-based methods. It summarized the research progress of graph convolutional neural networks, analyzing the development of graph convolutional neural networks from graph-based convolution and spatial-based convolution respectively. Graph-based convolution, rooted in mathematical expression and oriented towards reducing computational complexity, laid the theoretical foundation of graph convolutional neural networks. Spatial-based convolution, based on aggregation functions and update functions, combined theories such as sampling, attention mechanisms, and pooling operations, forming diverse graph convolutional neural networks, which have been widely applied.

Chai et al. proposed an unsupervised document graph embedding learning and classification model GGCN-DDC [16]. GGCN-DDC combines the scalability of TextING text representation and the advantages of DAEGC deep graph clustering. By improving the convolutional layer and proposing a new reconstruction matrix loss function, this model can better extract text features and implicitly learn the relationship between unlinked documents, thereby achieving better classification and clustering effects when dealing with unsupervised document corpora.

Yang et al. proposed a preference-aware denoising graph convolutional network social recommendation model PD-GCN [17]. PD-GCN is a social recommendation model that utilizes a preference-aware denoising graph convolutional network. By unsupervised learning, users are assigned to interaction subgraphs and social subgraphs to alleviate the over-smoothing problem and improve the model's robustness to noise. Meanwhile, by identifying and removing noise nodes through denoising strategies, the complex interactions and social relationships between users and items are effectively modeled.

Tang et al. proposed a knowledge graph convolutional network recommendation model KGCN-SHCN based on structural holes and common neighbors [18]. KGCN-SHCN achieves learning resource recommendation by computing the sampled neighborhood of central entities and utilizing the message aggregation mechanism of the KGCN model, enriching learning resources with auxiliary information from the knowledge graph, stimulating learners' interests, and improving the efficiency of recommendation by improving the sampling method and combining the message aggregation idea of graph convolutional networks.

Wang et al. proposed a low-light target detection algorithm based on image adaptive enhancement [19]. By designing an adaptive enhancement network and jointly optimizing it with the YOLOv5 target detection network end-to-end, the enhancement effect is more conducive to the target detection task. Meanwhile, by combining channel attention and pixel attention, a feature enhancement module is designed to improve target detection accuracy.

Luo et al. proposed a commodity news event extraction model SAT-GCN-DPT based on self-attention mechanism and average pooling graph convolutional network, aiming to solve the problems of weak correlation between trigger words and entity vectors and insufficient accuracy in parameter role extraction in commodity news event extraction [20]. This model combines self-attention mechanism, average pooling graph convolutional network, and dependency parsing tree. By using the ComBERT pre-trained model for data preprocessing, it enhances the correlation between trigger words and entity vectors and improves the accuracy of role segmentation using the average pooling function.

3 Methodology

Vehicle images contain various rich road information, with a huge amount of data, and a wide variety of vehicle types on the road, making it difficult to distinguish in detail. At the same time, different types of vehicles vary in size, posing great challenges for multi-class vehicle object detection. If using second-stage object detection algorithms like Faster RCNN, although the accuracy is high, the inference speed is slow, making it unsuitable for real-time detection applications. Additionally, it is complex, with a complicated training and deployment process involving candidate region generation and multi-stage processing, making it difficult to apply in resource-constrained environments. Algorithms like RetinaNet use Focal Loss to improve detection accuracy for small targets and imbalanced data, but require large computational resources and high hardware configurations. Although EfficientDet strikes a balance between performance and speed, it requires high-performance hardware to fully leverage its advantages. SSD (Single Shot MultiBox Detector), while faster, is not as accurate as YOLOv5, especially in handling small targets and complex backgrounds. Compared to YOLOv5, YOLOv3 lags behind in both accuracy and speed, and its code and community support are not as active as YOLOv5. Although R-FCN (Region-based Fully Convolutional Networks) is faster, it lacks flexibility in handling high-resolution images and complex scenes. Although CenterNet is simple and efficient, it does not perform as well as YOLOv5 in handling small targets and heavily occluded scenes. Therefore, proposing a road vehicle multi-class object detection method based on the YOLOv5 algorithm has the advantages of high real-time performance, high accuracy, simple and efficient processes, and lightweight models, enabling vehicle classification and detection in complex road environments.

Firstly, collect a dataset of images containing various types of vehicles and accurately annotate the vehicles in the images, including bounding boxes and category information. The vehicle image dataset should include images from various angles, lighting conditions, and environments. Next, perform operations such as scaling, rotating, and cropping on the training data to enhance the model’s generalization ability. Then, configure the network structure and training parameters of YOLOv5 according to the requirements of the multi-class vehicle object detection task. Subsequently, train the YOLOv5 model using the annotated vehicle dataset to learn the multi-class features of vehicles. After model training, optimize the model using the YOLOv5 loss function, including bounding box regression loss, confidence loss, and classification loss. Evaluate the model’s performance on the validation set using metrics such as mAP, Precision, Recall to measure detection accuracy. Adjust hyperparameters such as learning rate and batch size based on the evaluation results to optimize the performance of the vehicle multi-class object detection model. Finally, test the model on an independent test set to ensure its generalization ability, and deploy the trained model to an actual vehicle detection system. Utilize the deployed model to perform real-time vehicle detection on vehicle images or video streams, and perform post-processing operations such as NMS to remove redundant detection boxes and improve detection quality.

The basic algorithm framework Ultralytics/YOLOv5 used in this paper is an advanced object detection model developed based on Jocher Glenn's work, which exhibits good robustness and generalization ability in the problem of multi-class vehicle object detection [21]. For the selection of hyperparameters, the Model Architecture adopts YOLOv5n, which can better adapt to the computational resources and accuracy requirements of the dataset in this problem. At the same time, the model needs to crop the Input Size to dimensions such as 640*640 or 1280*1280. It is worth noting that the number of categories in the yaml file needs to be preset, and the number of categories in this dataset is 21, indicating that there are 21 different vehicle models that need to be classified and recognized.

4 Results

The training data for the multi-class vehicle model in this paper is selected from the road vehicle image dataset on the Kaggle website, which consists of Bangladeshi road vehicle images labeled with YOLOv5 tags. The dataset contains a total of 3004 images, with the image size of the test and validation datasets approximately 640 pixels by 640 pixels, and the labels are predefined. The training set of the dataset comprises 2704 images, and the validation set consists of 300 images, containing data for 21 categories of vehicles. There are 2568 vehicle target instances in the validation dataset.

Based on the default hyperparameters of Ultralytics/YOLOv5, optimizations and improvements were made in this paper. The pre-training weights chosen for training are from the YOLOv5n model, which, although not significantly superior in numerical performance, is more suitable for the multi-class vehicle classification problem studied in this paper, thus improving mAP. Additionally, the epoch is set to 300 to achieve optimal training results. The experiments were conducted using an NVIDIA GeForce RTX 3080 Laptop GPU, so the batch_size was set to 8, the number of workers was set to 4, and patience was set to 100 to stop training when the model reaches its optimum.

The evaluation metrics used in this paper are mAP, Precision, and Recall values, as shown in Table 1 with the experimental results.

Table 1. Experimental Results.

Class	Images	Instances	Precision	Recall	mAP50	mAP50-95
all	300	2568	0.669	0.358	0.41	0.244
bicycle	300	32	0.47	0.375	0.366	0.176
bus	300	425	0.799	0.567	0.661	0.392

Table 1. (continued).

Car	300	842	0.782	0.691	0.741	0.466
Minibus	300	2	0.645	0.5	0.495	0.446
Minivan	300	110	0.494	0.436	0.397	0.274
motorbike	300	335	0.619	0.505	0.51	0.173
pickup	300	142	0.525	0.218	0.299	0.163
policecar	300	1	1	0	0	0
rickshaw	300	192	0.717	0.708	0.694	0.397
scooter	300	1	1	0	0.00229	0.000457
suv	300	60	0.26	0.264	0.175	0.113
taxi	300	19	1	0	0.509	0.299
three wheelers	300	252	0.813	0.583	0.681	0.429
truck	300	84	0.546	0.607	0.609	0.375
van	300	62	0.267	0.161	0.205	0.127
wheelbarrow	300	9	0.761	0.111	0.213	0.0714

The experimental results indicate that the improved multi-class vehicle detection algorithm based on Ultralytics/YOLOv5 in this paper achieves high accuracy and recall rates on most vehicle types. The algorithm can accurately detect various vehicle types such as cars, buses, motorbikes, rickshaws, etc., demonstrating good generalization and robustness.

The comparison between the actual anchor boxes and the algorithm-detected anchor boxes is illustrated in Figure 1.

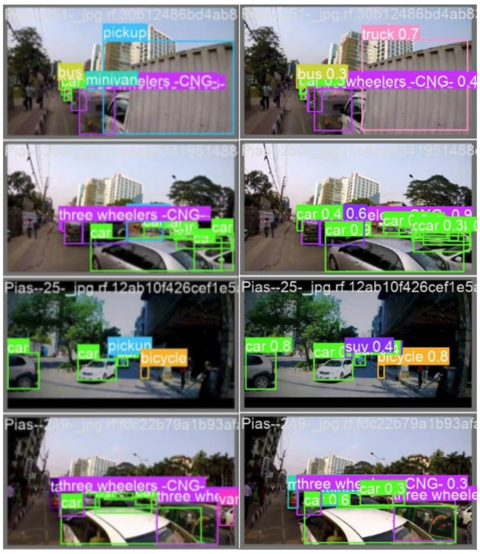

Fig. 1. Anchor Box Comparison (Image on the left side: Real Anchor Boxes; Image on the right side: Ultralytics/YOLOv5).

Figure 1 illustrates the comparison between the detection results of the algorithm proposed in this paper and the real anchor boxes. The left image shows the effect of vehicle label annotations, while the right image depicts the detection results of the improved Ultralytics/YOLOv5 algorithm.

5 Discussion

The dataset used in this paper consists of 3004 images, including 2704 training set images and 300 validation set images, with most of the dataset images being 640 pixels*640 pixels. Compared to similar studies in object detection, the scale of this dataset is relatively small. The pre-training weights used in the algorithm of this paper differ from those used in similar studies. Adopting the pre-training weights of YOLOv5n can maximize computational efficiency under limited hardware conditions, but it also results in very high CPU thread usage.

Due to the relatively small sample size involved in this study and the presence of multiple effective anchor boxes within individual samples, coupled with considerations of computational power constraints, a batch size of 8 was set. This helps control sampling freedom and aids in model generalization. Considering the total size of the training set and sample density, the number of epochs was set to 300, while the patience was set to 100 to ensure that the model receives sufficient training.

Results show variations in mAP, Precision, and Recall performance across different vehicle categories. Firstly, some classification label samples have fewer quantities, leading to insufficient training. Secondly, certain vehicle types exhibit obvious features, while others may have less distinct features due to occlusion or other factors. Generally, as the number of categories increases, the success rate of recall for individual categories tends to decrease. In contrast, precision is not affected by this issue, resulting in better numerical performance. Compared to similar multi-classification models, the model in this paper demonstrates better numerical performance in classifying most vehicles with distinct features.

In complex road environments, some vehicles may overlap, leading to poor detection results for the algorithm in this paper. However, under good lighting conditions and minimal vehicle overlap, the proposed method can effectively handle the multi-classification of vehicles.

6 Conclusion

This paper proposes a vehicle multi-classification object detection method based on the Ultralytics/YOLOv5 algorithm for vehicle images under complex road traffic conditions. The method enhances the model's generalization capability by performing operations such as scaling, rotation, and cropping on the training data. Based on the requirements of the vehicle object detection multi-classification task, the network structure and training parameters of YOLOv5 are configured, and the model is optimized using YOLOv5's loss function. Compared to similar studies, the algorithm in this paper, through data augmentation, adjustment of training parameters, and reasonable selection of pre-training weights, can detect vehicles in road environments with a wide variety of categories and achieve high detection accuracy. When detecting most common vehicle types, the method proposed in this paper demonstrates stronger robustness and higher detection accuracy compared to other object detection methods. However, due to limitations of the model itself and the presence of numerous vehicle types in specific complex road environments, the detection effectiveness of rare vehicle types in this paper did not reach the optimal level. In future work, we will seek more urban vehicle image datasets for training and explore research on object detection tasks under low-light and overlapping conditions.

References

[1] Liang, S., Wang, S., Chen, J., & Yu, J. (2023). Research on port ship target detection algorithm based on coastline candidate area. Radio Engineering, 10, 2270-2276.

[2] Zhao, Q., & Yang, Y. (2023). Lightweight remote sensing vehicle small target detection algorithm based on multiple pyramids. Electronic Measurement Technology, 13, 88-94. doi:10.19651/j.cnki.emt.2211674.

[3] Lv, H. T., & Jia, X. L. (2024, June 27). Lightweight giant panda target detection model based on attention mechanism. Laser Journal, 1-7. http://kns.cnki.net/kcms/detail/50.1085.TN.20231207.1051.006.html.

[4] Li, S., Wang, C., & Wang, M. (2023). Extremely lightweight aerial image port ship target detector. Computer Engineering and Design, 12, 3606-3612. doi:10.16208/j.issn1000-7024.2023.12.012.

[5] Chen, A., Yu, Y., Zhao, H. R., et al. (2024, June 27). Detection of damaged areas in ancient murals based on nested U-shaped structure network. Computer Engineering and Applications, 1-11. http://kns.cnki.net/kcms/detail/11.2127.TP.20231218.1537.026.html.

[6] Yuan, J., Lan, Z., & Xiong, P. (2023). Unmanned aerial vehicle target detection method based on region adaptive threshold. Computer and Digital Engineering, 12, 2883-2888+2990.

[7] Chen, J. M., Zhang, W. D., & Tan, R. P. (2024, June 27). Drug recommendation algorithm based on dialogue structure and graph attention network. Journal of Computer Science and Exploration, 1-12. http://kns.cnki.net/kcms/detail/11.5602.TP.20240108.0943.002.html.

[8] Zhang, Z., & Yang, H. (2024). Improved VGG network face expression recognition method based on LBP and attention mechanism. Software Engineering, 01, 23-26+31. doi:10.19644/j.cnki.issn2096-1472.2024.001.006.

[9] Hu, T., Gao, X., Hua, Y., & Cai, L. (2024). Adaptive apple image multi-defect detection based on deep learning. Journal of Shandong University of Technology (Natural Science Edition), 01, 42-47. doi:10.13367/j.cnki.sdgc.2024.01.005.

[10] Wu, H., Zhang, Y., & Hu, P. (2024). Small target detection method for unmanned aerial photography images. Journal of Anhui University of Technology (Natural Science Edition), 01, 65-73.

[11] Hua, C., Mo, S., Chen, Y., Hu, H., & Wu, S. (2024). Analysis of target detection algorithm based on infrared images. Automotive Practical Technology, 02, 59-66. doi:10.16638/j.cnki.1671-7988.2024.002.012.

[12] Wang, C., Wang, Z., & Li, W. (2024). Controller sleep behavior recognition based on dual-stream adaptive graph convolutional network. Journal of Safety and Environment, 02, 596-601. doi:10.13637/j.issn.1009-6094.2023.0240.

[13] Zou, Z. B., Liu, Y. Z., Liao, Z. H., et al. (2024, June 27). Spatio-temporal attention graph convolutional network for dynamic traffic flow prediction. Journal of Shandong University (Engineering Edition), 1-12. http://kns.cnki.net/kcms/detail/37.1391.t.20240221.1017.004.html.

[14] Zhou, H., Zhan, F., Zhou, C., Ren, T., & Luo, L. (2024). Pedestrian re-identification method based on attention mechanism and multi-branch joint. Microelectronics and Computer, 02, 1-10. doi:10.19304/J.ISSN1000-7180.2023.0016.

[15] Xie, J., & Zhang, J. (2024). Review of graph convolutional neural networks. Journal of Shaanxi Normal University (Natural Science Edition), 02, 89-101. doi:10.15983/j.cnki.jsnu.2024003.

[16] Chai, B., Li, Z., Zhao, X., & Wang, R. (2024). Deep document clustering model based on generalized graph convolutional neural network. Journal of Nanjing Normal University (Natural Science Edition), 01, 82-90.

[17] Yang, X. Y., Ma, S., Zhang, Z. L., et al. (2024, June 27). Social recommendation based on preference-aware denoising graph convolutional network. Computer Engineering, 1-10. https://doi.org/10.19678/j.issn.1000-3428.0068748.

[18] Tang, Z., Wu, Y., Li, C., & Tang, Y. Learning resource recommendation based on knowledge graph convolutional network. Computer Engineering. doi:10.19678/j.issn.1000-3428.0068409.

[19] Wang, F., Chen, X., Ren, W., Guan, Y., Han, Z., & Tang, Y. Low-light target detection algorithm based on image adaptive enhancement. Computer Engineering. doi:10.19678/j.issn.1000-3428.0068407.

[20] Luo, Q., Li, H., Wang, Z., Gan, C., & Hu, Z. Event extraction of commodity news based on self-attention mechanism and average pooling graph convolutional network. Journal of Chengdu University of Technology (Natural Science Edition).

[21] Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., Kwon, Y., Michael, K., ... & Mammana, L. (2022). ultralytics/yolov5: v6.2-yolov5 classification models, apple m1, reproducibility, clearml and deci.ai integrations. Zenodo.