Journal of Electronic Research and Application, 2024, Volume 8, Issue 4
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Tactile Paving Recognition Method Based on Improved YOLOv7

Chengyi Shi[1], Qun Zhao[1]*, Nan Fang[1], Yufeng Mao[2]

[1]School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China
[2]NR Electric Co., Ltd., Nanjing 211102, China

Corresponding author: Qun Zhao, qunzhaoqz@outlook.com

Abstract: Tactile paving is a professional road facility to ensure the safe travel of people with visual impairment. However, there are many problems with tactile paving travel in practice. For one, some tactile paving is seriously damaged, and the other is the accumulation of obstacles. How to help visually impaired people recognize and locate obstacles in tactile paving is a problem worth studying. In this paper, image recognition technology is used to recognize the tactile paving pictures with obstacles, and an attention mechanism is used to optimize samples to improve recognition accuracy.

Keywords: YOLOv7; Tactile paving recognition; Attention mechanism

Online publication: August 12, 2024

1. Technology background

In China, there are approximately 17 million people with visual impairments, accounting for 1% of the total population. With the development of modern society, most people with visual impairments mainly rely on blind canes, assistive devices, etc. to carry out their daily lives and improve their quality of life. To enable the visually impaired to walk safely outside, this paper conducts relevant research on the current tactile paving facilities and tactile paving obstacle recognition.

In the field of computer vision, object recognition is an important issue [1]. Some methods transfer objects to visually impaired individuals through one-dimensional data drawing, while others solve the problem of interaction between visually impaired individuals and the surrounding environment through the real value local dissimilarity mapping method in the measurement of inter-frame dissimilarity in images. However, in practical applications, object detection may vary in different scenarios [2]. Therefore, based on the differences in the regional distribution of different tactile paving, this paper proposes a tactile paving recognition method based on You Only Look Once (YOLOv7). At the same time, based on the attention mechanism, a top-down information screening method is used to screen out a large amount of irrelevant information, and computer resources are allocated to more meaningful work to improve the ability of neural networks to process information [3].

2. Related works

To assist visually impaired individuals in controlled environments such as hospitals and homes, Kim *et al.* proposed a new assistive device that uses data beyond one dimension to depict objects for blind individuals [4]. Additionally, the output-only program proposed by Hossein and Alireza can effectively assist blind people in extracting seismic input motion and identifying systems with severe permutation patterns [5]. To achieve an assistive device that allows blind people to adapt quickly, Alsayaydeh *et al.* built automated hardware using Raspberry Pi, aiming to enable visually impaired individuals to immediately detect objects or people in front of them and inform them of what is in front of them using audio [6]. Compared to the above-mentioned blind assistance methods, the smart home model proposed by Tayyaba *et al.* for safe and stable movement of blind people can enable them to complete daily activities without wearing heavy equipment [7]. This model uses fuzzy logic for simulation, and the output of Internet of Things (IoT) devices composed of sensors and Bluetooth serves as the input of the fuzzy controller. The output communicates through IoT devices to assist blind people or users in safe movement.

For the object detection model, this article chooses YOLOv7 as the basis. Before this, Qi and Gao proposed an improved YOLOv7 object detection model [8]. Combining the idea of feature separation and merging improves the Multi-Path Convolution (MPConv) module in the YOLOv7 network model. In addition, Wang *et al*. proposed an automatic dataset expansion method for transmission line defect detection based on the improved YOLOv7 algorithm, which can use a small number of images to establish a transmission line defect dataset [9]. Cao *et al.* also proposed an improved night pedestrian detection algorithm for YOLOv7 [10]. In the improved algorithm, the YOLOv7 tiny network is used as the baseline to meet both accuracy and high detection speed. Mao also proposed an automated road defect detection algorithm based on the YOLOv7 distillation model, which combines the advantages of high object detection accuracy and fast speed [11]. Zhang *et al.* proposed a three-dimensional detection method for autonomous driving platforms based on roadside monocular cameras [12].

3. Research methods

The YOLOv7 algorithm consists of three parts: input, backbone network, and prediction head. The backbone network adopts an efficient aggregation network structure, which allows the network to learn more features by controlling the shortest and longest gradient paths. The environment of tactile paving is complex, and some are seriously damaged or have obstacles piled up, which sometimes blocks the target tactile paving [13]. After the original image is normalized in size, the effective part of the target tactile paving will be smaller. Because the shallow features of the convolutional neural network have more local information, it is easy to cause the key information of the deep feature map to lose the tactile paving part after multiple convolution operations, thus causing false recognition. Given the above problems, this paper improves the YOLOv7 model. To improve the model's ability to extract tactile paving features in different situations, a convolutional attention module is added to the backbone network, to improve the network's attention to the target area of interest in the input image.

3.1. Multi-scale feature fusion

When facing the tactile paving and ordinary road with a small color difference, the original network structure cannot extract the local features of small occlusion well and fuse them, which results in the loss of local information of the tactile paving and the false recognition of tactile paving. YOLOv7 extracts key features from the 80 × 80, 40 × 40, and 20 × 20 feature maps in the backbone network, and performs feature fusion. To better distinguish the tactile paving from the ordinary road and improve the accuracy of classification and recognition, a 160 × 160 feature scale is added to the original three scales. This method outputs the extracted features on the original 160 × 160 feature map of the backbone network and fuses the scale features with the original three scale features.

3.2. Convolutional Block Attention Module (CBAM) attention mechanism

The road condition of tactile paving is complex. Due to the influence of tactile paving damage, obstacle accumulation, and other factors, many interference messages appear in the picture. The attention mechanism is a dynamic selection process for important information input into an image, which is achieved through feature adaptive weights, focusing on important features of the image and suppressing unnecessary regional responses. Therefore, this article introduces the attention mechanism CBAM into YOLOv7.

The insertion position of the attention mechanism will also affect the performance of the network. In this paper, CBAM is embedded into the 11^{th}, 24^{th}, 37^{th}, and 51^{st} layers of the backbone network respectively, and the feature maps obtained from each layer of the network are used as the input feature maps. After the CBAM module, the weight of tactile paving features is added, and the feature maps are output, to further improve the ability to extract key features of the tactile paving under different feature scales, such as 160 × 160, 80 × 80, 40 × 40 and 20 × 20, which are represented by L-1 to L-4 respectively.

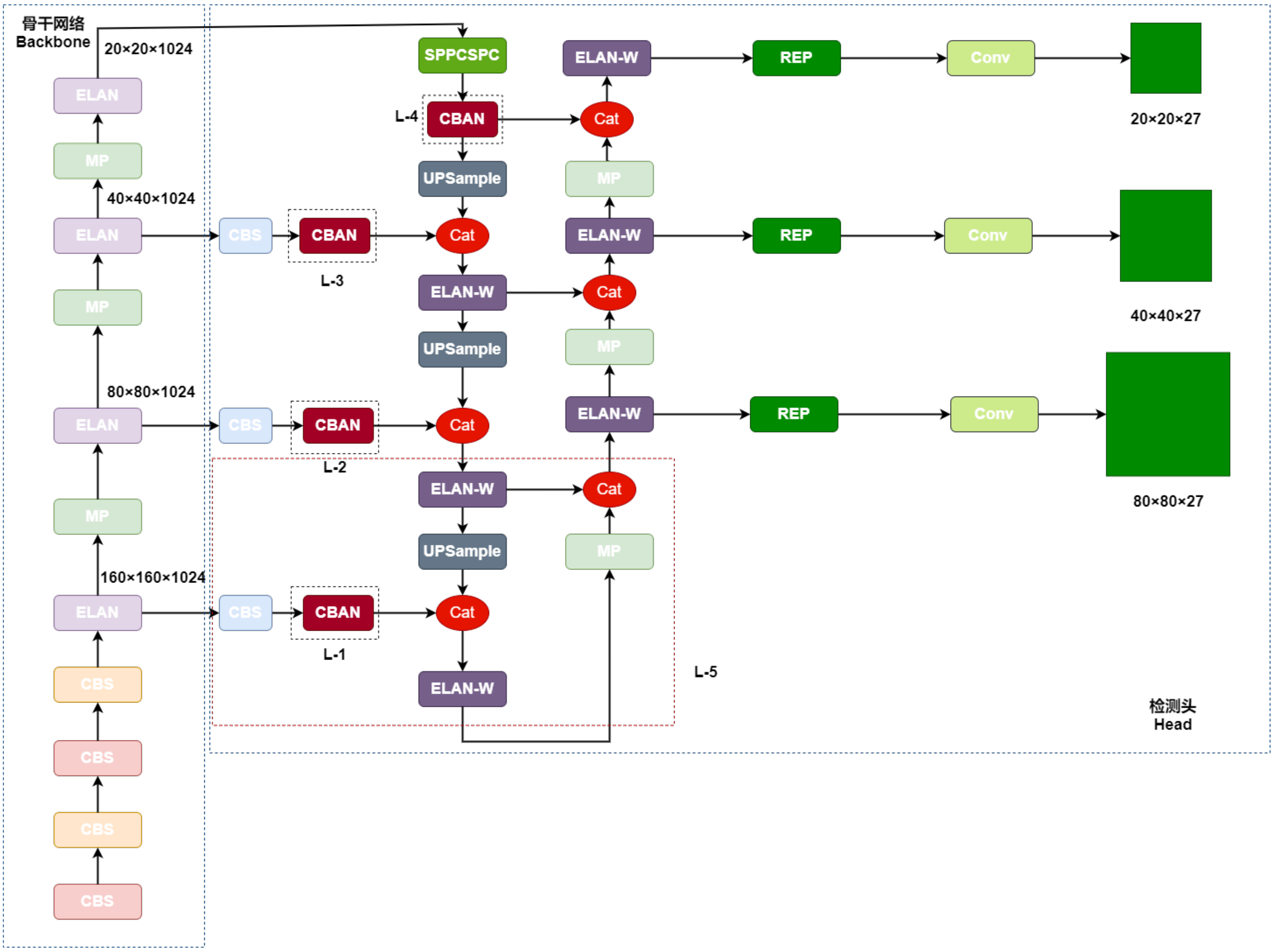

Figure 1. Improved YOLOv7 network model.
Note: L-1 to L-4 are the final addition positions of the CBAM module, and L-5 is the added 160 × 160 scale feature

4. Experiment

To evaluate the performance of the improved YOLOv7 algorithm for tactile paving recognition, this experiment conducted training and testing on the tactile paving dataset. The image data set collected in this paper comes from AI STUDIO and live scenes. The data set includes complete tactile paving, damaged tactile paving, and partially blocked tactile paving. As shown in Figure 2, use the LabelImg annotation tool to annotate. The format of the annotation information file is a text file, which contains information such as the category and bounding box of the target tactile paving. The total number of image sample data used in the experiment is 8,697. Among them, there are 6,221 training sets, 1,230 validation sets, and 1,246 test sets.

(a) Complete tactile paving (b) Damaged tactile paving (c) Covering tactile paving

Figure 2. Schematic diagram of tactile paving under different conditions

Due to limitations in experimental equipment, the input image size was scaled to 640 × 640 pixels, and the optimizer used Mini-Batch Gradient Descent (MBGD) with a learning rate of 1×10^{-2}, momentum of 0.9, and weight attenuation of 5×10^{-4}. The learning rate was adjusted using the MultiStepLR method. The batch size is 12, with a training duration of 200 epochs, alternating between 10 training epochs and 1 testing epochs.

The experiment uses the average recognition accuracy of all categories with mean average precision (mAP) at 0.5. The Intersection over Union (IoU) threshold is set to be greater than 0.5 as the evaluation indicator, and the parameters used include accuracy, recall, etc. [14].

The experiment compared the proposed algorithm with mainstream object detection methods YOLOv5, YOLOv7, and Region-based Convolutional Neural Network (RCNN). As shown in Table 1, the original YOLOv7 algorithm uses a Path Aggregation Feature Pyramid Network (PAFPN) for efficient fusion of features at different levels, but it cannot effectively handle the problem of target boundaries and angle periodicity. Therefore, compared to the original YOLOv7 model, the detection accuracy of our algorithm has been improved by 4.1 percentage points [15]. Among them, RCNN can accurately predict by adding rotation equivariant networks to extract equivariant features, but the detection accuracy is still 5.9% lower than the algorithm proposed in this paper. The sampling fusion network and loss function designed by YOLOv5 is good for small and messy target detection, but the detection effect for tactile paving is not as good as the algorithm in this paper. Experimental results show that the improved algorithm proposed in this paper has better detection performance on tactile paving data sets.

Table 1. Recognition accuracy of four methods applied to four different tactile paving

Methods	*P* (%)	*R* (%)	mAP at 0.5 (%)
YOLOv5	79.1	82.7	80.7
RCNN	81.5	77.2	79.2
YOLOv7	83.3	84.6	80.2
This article's algorithm	87.4	85.1	87.2

5. Conclusion

This paper proposes a tactile paving recognition method based on improved YOLOv7 to solve the problem of difficulty in recognizing tactile paving and obtaining information in different situations at present. By improving the multi-scale feature fusion network, this method is used to enhance the recognition sensitivity of the model to small local features, introduce the attention mechanism CBAM, improve the network's attention to the target area of interest in the input image, and improve the detection accuracy of the model to tactile paving in various situations. Compared with other typical algorithms, this method can effectively realize the recognition of tactile paving in different situations.

Funding

This work has been partially supported by the Jiangsu Province College Student Innovation Training Program (Project No. 20221127684Y) and the Talent Startup project of Nanjing Institute of Technology (Project No. YKJ202117)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Yang J, Li H, Tian C, et al., 2021, Dynamic Histogram Equalization Algorithm Based on Adaptive Correction. Computer Engineering and Design, 42(5): 1264–1270.

[2] Wu X, Wu L, Wu Z, et al., 2018, A Fast Multi-Target Recognition and Synchronous Ranging Method for Visual Assistance of Blind People. Journal of Fuzhou University: Natural Science Edition, 46(4): 472–480.

[3] Wang L, Zhang L, Li S, Zhou G, 2019, Context Sensitive Question Answering Pairing Method Based on Attention Mechanism. Chinese Journal of Information Science, 33(1): 125–132.

[4] Kim JH, Park JE, Lee JM, 2020, 3-D Space Visualization System Using Ultrasonic Sensors as an Assistive Device for the Blind. IEEE Journal of Translational Engineering in Health and Medicine, 8: 2700305.

[5] Hossein J, Alireza G, 2022, Input Motion Extraction and Closely-Spaced Modes Recognition Through Partial System Identification: An Orthogonality-Aided Blind Source Separation Approach. Measurement, 2022: 200.

[6] Alsayaydeh JAJ, Irianto, Zainon M, et al., 2022, Intelligent Interfaces for Assisting Blind People Using Object Recognition Methods. International Journal of Advanced Computer Science and Applications (IJACSA), 13(5): 734–741.

[7] Tayyaba S, Ashraf MW, Alquthami T, et al., 2020, Fuzzy-Based Approach Using IoT Devices for Smart Home to Assist Blind People for Navigation. Sensors, 20(13): 3674.

[8] Qi L, Gao J, 2023, Based on the Improved Detection of Small Target YOLOv7. Computer Engineering, 49(01): 41–48. https://doi.org/10.19678/j.iSSN.1000-3428.0065942

[9] Wang Y, Shang J, Zhang Y, et al., 2023, Real-Time Transmission Line Defect Detection Method Based on Improved YOLOv7. Southern Power Grid Technology, 2023: 1–8.

[10] Cao Y, Li C, Peng Y, 2022, Night Pedestrian Detection Algorithm Based on Improved YOLOv7. Yangtze River Information and Communication, 35(10): 57–60.

[11] Mao J, 2023, Study on Algorithm of Road Defect Detection Based on YOLOV7 Distillation Model. Modern Manufacturing Technology and Equipment, 59(02): 222–224. https://doi.org/10.16107/j.cnki.mmte.2023.0084

[12] Zhang L, Wang W, Wang B, et al., 2023, A 3D Detection Method for Autonomous Driving Platforms Based on Roadside Monocular Cameras. Information and Computer, 35(2): 201–205
[13] Chen Y, 2023, Research on Small Object Detection Algorithm Based on Improved YOLOv7, thesis, Jiangsu University of Science and Technology.
[14] Xie J, Wang W, Liu T, 2021, Fabric Defect Detection Using Different Backbone Networks Based on YOLO v3 Algorithm. Measurement and Control Technology, 40(3): 61–66.
[15] Cao J, Lv J, Wu X, et al., 2021, A Fall Detection Algorithm Integrating Motion Features and Deep Learning. Computer Applications, 41(2): 583–589.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.