Foreign Object Detection Algorithm Based on Multi-scale Convolutional Network

Jiexin Zheng 1, Ye Chen 2, Hong Zhang 2, Dan Liu 2,*

1 School of Electronic and Control Engineering, Chang’an University, Xi’an 710064, China
2 College of transportation engineering, Chang’an University, Xi’an 710064, China

*Corresponding author e-mail: liudan88r@chd.edu.cn

Abstract. In order to detect foreign matter invading the track and prevent the intrusion of foreign matter from causing railway safety accidents, the detection algorithm of foreign matter intrusion on the track is studied. Aiming at the problem that the rail transit scene is complex and the obstacle scale changes in the acquired image information, this research proposes a multi-scale target detection algorithm based on the YOLO (you only look once) algorithm. First, an adaptive feature fusion module is designed to make the feature maps used for detection have strong semantic information at various scales; then, a new loss function is designed to alleviate the problem of uneven sample distribution and optimize the training process. Experiments show that the algorithm has obvious advantages in multi-scale detection, which not only improves the accuracy of target detection, especially the accuracy of small targets, but also does not significantly increase the inference time and the amount of parameters, and has high real-time performance.

Keywords: Target detection, multi-scale, rail transit, intrusion detection.

1. Introduction

By the end of 2019, my country's railway operating mileage has reached 139,000 kilometers, and the road network density has reached 145.5 kilometers/10,000 square kilometers. Railway transportation has become one of the indispensable means of transportation for people's daily life and cargo transportation. With the continuous increase of the terrain covered by the railway, a complex train operating environment has been formed, and the continuous increase of train operating speed has made the braking distance larger and larger. When foreign matter in the track invades the railway boundary, it is difficult for the train to brake, which threatens driving safety. Through the analysis of the causes of railway accidents, it can be known that in railway accidents, the impact of foreign body intrusion on trains and people is the primary cause except for man-made and train failures [1,2]. The railway foreign body intrusion limit refers to the color steel tiles, dangerous trees, overhead lines, tower poles, light garbage (kites, plastic sheets, etc.), illegally crossing pedestrians, vehicles or unidentified trees within a certain range on both sides of the high-speed and normal-speed railways. The maintenance personnel who withdrew in time and the stranded mechanical equipment invaded the railway boundary [3], thus causing varying degrees of harm to the railway traffic.
At present, the protection and monitoring methods adopted for the intrusion of railway foreign bodies mainly include: installing protective nets in special sections along the railway to isolate the invading foreign bodies, and at the same time adopting the manual inspection methods of patrol workers and driver's observation to realize the monitoring of intrusive foreign bodies [4, 5]. But in recent years, limit violation accidents still occur frequently. A large amount of accident data shows that the traditional railway intrusion foreign body detection method composed of protective nets, patrolmen, drivers, etc., is difficult to respond to emergencies under the long-term high-speed operation of trains, and has gradually failed to meet the needs of high-speed railway development. Foreign body intrusion detection has become one of the key areas of concern for railway transportation safety protection in various countries.

2. Related work

Aiming at the problem of railway foreign body intrusion detection, according to the actual situation of railway sites in various countries, domestic and foreign railway foreign body intrusion detection systems based on different scenarios have been developed.

Japanese research scholar Okazaki and others developed a 60Gh millimeter-wave spread-spectrum radar system based on the characteristics of high accuracy and strong anti-interference ability of radar sensors in response to the frequent intrusion of private cars and trucks at railway crossings. Turkish scholar Mucahit Karaduman [6] and others used image processing technology to achieve laser detection of railway obstacles. Juan Jesús García [7], a research scholar at the University of Alcala in Spain, and others proposed an FPGA-based multi-sensor railway safety assurance system. The system consists of a transmitter and a receiver, which are located on opposite sides of the railway track. By using infrared (IR) and ultrasonic (US) sensors establish optical and acoustic links between them. French scholar Houssam Salmane et al. [8] proposed a railway safety detection system based on video analysis. The system monitors real-time violation behaviors caused by pedestrians, vehicles and unattended objects entering the railway boundary at level crossings.

In recent years, Chinese researchers have done a lot of research work on the detection of foreign objects intruding on railway tracks and have made remarkable achievements. Chen Genzhong [9] proposed a track intrusion detection system based on monocular vision, which uses Hough transform to extract the rail edges in the detection result of Canny operator. The detection of intrusions is realized based on the background difference method, and the intrusions are classified through the multi-frame classification method to determine whether the intrusions are trains. Sun Yongli [10] uses image processing technology to realize automatic detection of railway invading foreign objects. Li Dandan [11] proposed an intelligent recognition method for track foreign body intrusion. This method is based on the rail position to establish a foreign body detection window in the dangerous area of the railway track. The detection of intrusions is realized by a combination of background difference and inter-frame difference method, and finally passed The image correlation coefficient and the area ratio between the foreign object target and the image realize the intelligent identification of foreign objects. In order to prevent the running train from being mistakenly judged as an intruder, the increased or decreased area in the foreground of the continuous sequence of images is used to complete the intelligent identification of the train.

Through in-depth analysis of intrusion detection algorithms in the above research results at home and abroad, there are still problems that need to be improved and resolved. For the detection of intrusions, the complexity of the background environment of railway operation and a single traditional target detection algorithm cannot accurately and completely detect the intrusions, which may easily cause false and missed detections of the intrusions. After the intrusion is detected, it is often only to track the intrusion, but it is rarely judged whether the intrusion has violated the limits.
3. Method
First, YOLO’s [12] backbone network performs multiple convolution-pooling operations on the input image, extracts features, and generates a feature map; then, the feature map is divided into \( s \times s \) grids, and each grid is responsible for detecting the center point. Enter the target in the grid area, classify the target confidence vector in the grid by the features in the grid, and generate \( B \) bounding boxes fixedly; finally, the \( s \times s \times B \) boundaries obtained from the whole image. The box is filtered by the bounding box confidence.

3.1. Network framework
This paper designs a multi-scale target detection algorithm, and its network structure is shown in Figure 1.

![Fig.1 Multi-scale detection algorithm network structure diagram](image)

The DBL (DarkNet convolution batch normalization leaky relu) unit is composed of convolution, batch normalization and activation functions; the residual body is composed of zero padding, a DBL unit and three residual units; the residual unit is stacked by two DBL units, and add identity mapping.

The backbone network of the multi-scale detection algorithm is composed of multiple residual bodies and has strong feature extraction capabilities. The feature pyramid has four layers. The first layer is obtained from the last layer of the backbone network through DBL unit and convolution, and the second, third, and fourth layers are obtained from the penultimate layer of the backbone network, and the upper layer of the feature pyramid passes through the DBL unit and convolution. After upsampling, the channels are spliced, and then obtained through DBL unit and convolution.

3.2. Loss function
The classification loss function used by the YOLO detector is the cross-entropy loss function \( L_c \), which is expressed as follows.

\[
L_c = -\log (p_t)
\]

In the formula: \( p_t \) is the classification probability, which is taken as 1 when the classification is correct, and 0 when the classification is wrong, minus the classification confidence.

In the cross-entropy loss function, the weight of the negative sample and the positive sample are the same, but the difference in the number of the two makes the weight of the negative sample much larger than the positive sample. For easy-to-classify samples, the number is too large, so that the hard-to-classify samples with small \( p_t \) and weight is lower. Aiming at the above-mentioned sample distribution imbalance problem, this research improves the cross-entropy loss function, and the improved cross-entropy loss function \( L_f \) is as in formula.

\[
L_f = \alpha_t (1 - p_t)^\gamma \log (p_t)
\]
In the formula, $a$ is the weighting factor, which is 0.3.

4. Experiment

4.1. Training

According to railway safety monitoring requirements, the classification and recognition rate of intrusions is as high as possible, and the training sample set is obtained by shooting the intrusions at the railway crossing and the training base. When selecting training samples, the positive sample set is set to images of intrusive foreign objects (pedestrians, cars), the number of which is 100 (50 for pedestrians and cars), and the selection of negative sample sets needs to highlight the difference with the positive samples. Representative backgrounds and trains that do not contain intrusive foreign objects, the number of negative samples is 200 (including 50 trains). Due to the different sizes of the intruders in the detection results, it is necessary to normalize the image size of all sample sets to 20×20 pixels.

4.2. Analysis

In order to verify the effectiveness of the algorithm in this paper, the average value method is used to calculate the average value of the 10 recognition results. When the Adaboost algorithm is used to realize the classification and identification of intruders, the positive samples in the test sample include 25 pedestrians and 25 cars; the negative samples are 50 (20 trains). YOLO v3 [13, 14] is used to realize the intrusion detection, and the intrusion detection is realized based on the algorithm of this paper.

| Methods   | Sample size | Mean accuracy | Mean error rate |
|-----------|-------------|---------------|-----------------|
| Adaboost  | 100         | 77.86%        | 22.14%          |
| YOLO v3   | 100         | 80.23%        | 19.77%          |
| Ours      | 100         | 86.97%        | 13.03%          |

It can be found from Table 1 that our algorithm has a better accuracy rate than Adaboost and YOLO v3. Here are a few examples of the method in this article and the detection results of YOLO v3.

![Fig. 2 The results of YOLO v3](image)
Through the comparison of Figure 2 and Figure 3, it can be found that the detection accuracy of the method in this paper is higher. YOLO v3 is prone to miss detection for small targets, and the network based on multi-scale convolution in this paper is more effective for small targets.

5. Conclusion
The high-performance target detection algorithm is a powerful guarantee to avoid foreign body intrusion into the railway track, and it is also the most critical core technology to realize the foreign body intrusion detection system based on the railway track environment. Based on the superior characteristics of the deep learning model, this paper designs a detection algorithm that is applied to the track foreign body intrusion detection scene. This article sets a specific scene of the track foreign body intrusion, delimits the detection area, and analyzes the results based on the YOLO_V3 deep learning model. A complete set of orbital foreign body intrusion detection process can realize the functions of comprehensive detection, real-time monitoring, timely warning and early processing.

Acknowledgments
This work was financially supported by xxx fund. This work was supported by College Students' Innovative Entrepreneurial Training Plan Program (Nos.300102210120) and Basic Research Foundation of the Central Universities.

References
[1] Bosquet Brais, Mucientes Manuel, Brea Victor M.. STDnet-ST: Spatio-temporal ConvNet for small object detection[J]. Pattern Recognition, 2021, 116.
[2] Zhao Xiaoli, Chen Zheng, Hwang Jenq Neng, Shang Xiwu. AFLNet: Adversarial focal loss network for RGB-D salient object detection[J]. Signal Processing: Image Communication, 2021, 94 (prepublish).
[3] Quang Tran Ngoc, Lee Seunghyun, Song Byung Cheol. Object Detection Using Improved Bi-Directional Feature Pyramid Network[J]. Electronics, 2021, 10(6).
[4] Liu BiYuan, Chen HuaiXin, Huang Zhou, Liu Xing, Yang YunZhi. ZoomInNet: A Novel Small Object Detector in Drone Images with Cross-Scale Knowledge Distillation[J]. Remote Sensing, 2021, 13(6).
[5] Paul Smriti Nandan, Frueh Carolin. Analytical expressions for orbital perturbations due to Lorentz
force[J]. Acta Astronautica,2021,182.

[6] Zou Cheng,He Bingwei,Zhu Mingzhu,Zhang Liwei,Zhang Jianwei. Encode-decode network with fully connected CRF for dynamic objects detection and static maps reconstruction[J]. Signal Processing: Image Communication,2021(prepublish).

[7] John Vijay,Mita Seiichi. Deep Feature-Level Sensor Fusion Using Skip Connections for Real-Time Object Detection in Autonomous Driving[J]. Electronics,2021,10(4).

[8] Fan Wang,Guohua Peng. Salient object detection via cross diffusion-based compactness on multiple graphs[J]. Multimedia Tools and Applications,2021(prepublish).

[9] Fonseca Erika,Santos Joao F.,Paisana Francisco,DaSilva Luiz A.. Radio Access Technology characterisation through object detection[J]. Computer Communications,2020(prepublish).

[10] Ding Lianghui,Xu Xin,Cao Yuan,Zhai Guangtao,Yang Feng,Qian Liang. Detection and tracking of infrared small target by jointly using SSD and pipeline filter[J]. Digital Signal Processing,2020,110(prepublish).

[11] Gao Ce, Wu Yiquan, Hao Xiaohui. Hierarchical Suppression Based Matched Filter for Hyperspectral Imagery Target Detection[J]. Sensors,2020,21(1).

[12] Frederik E.T. Schöller,Martin K. Plenge-Feidenhans'l,Jonathan D. Stets,Mogens Blanke. Assessing Deep-learning Methods for Object Detection at Sea from LWIR Images[J]. IFAC PapersOnLine, 2019,52(21).

[13] Junho Lee, Hyukjun Jang. Image Classification using Deep Learning Algorithm and 2D Lidar Sensor[J]. Journal of IEEE,2019,23(4).

[14] Xurshid Farhodov,Oh-Heum Kwon,Kwang-Seok Moon,Oh-Jun Kwon,Suk-Hwan Lee,Ki-Ryong Kwon. A New CSR-DCF Tracking Algorithm based on Faster RCNN Detection Model and CSRT Tracker for Drone Data[J]. Journal of Korea Multimedia Society,2019,22(12).