Helmet wear detection based on neural network algorithm

Ruiyun Cao, Hongxiang Li, Banglie Yang, Ao Feng, Jie Yang, Jiong Mu

Key Laboratory of Agricultural information engineering of Sichuan Province, Sichuan Agriculture University, College of Information Engineering Sichuan Agricultural University, Ya'an, China

*Corresponding author: 378558989@qq.com

Abstract. Wearing safety helmet correctly is an effective means to reduce the accident rate and ensure the safety of construction personnel. But in the complex work site, the helmet detection task will face a variety of challenges. This paper uses YOLOV4's deep neural network architecture and makes fine-tuning and improvement according to the characteristics and difficulties of this task, proposes a new helmet detection system, and achieves 95.1% accuracy. The application can timely and accurately detect whether the personnel wear the helmet correctly, which is of great significance for ensuring construction safety in practical work.

1. Introduction

A large number of safety accidents each year are related to helmet wearing. Construction workers are considered to be one of the most dangerous occupations, and are vulnerable to falling objects, knocks, and flying objects on construction sites, factories, power plants, and transportation. The safety of construction site has been widely concerned by the society. Many high-risk industries have high rates of injury and death from slips, falls, collisions and collisions with falling objects. To protect workers from danger, helmets are important personal protective equipment in many workplaces, including construction, chemical, manufacturing, and mining. Although every worker is required to wear a helmet, there will always be cases where the worker does not wear it or does not wear it correctly. Video surveillance systems have been deployed at most sites to monitor workers' behavior and ensure work safety. However, considering the long time of detection through video system is a difficult task for human beings. It turns out that as video surveillance hours increase, so do human errors. With the advent of the era of intelligent manufacturing, automation of safety management is as important as automation of manufacturing process. Safety helmet, as a basic protection measure for construction workers, is of great significance to protect the life safety of construction workers.

2. Related Work

In the real construction environment, there are often difficult problems such as complex scenes, severe occlusion, light changes, and small targets, which make helmet detection a very challenging problem. In recent years, helmet detection has attracted the attention of many researchers. Traditional helmet detection methods are mainly based on manual characteristics. Jie Li et al. [1] proposed an image processing and machine learning detection method. ViBe background modeling algorithm was used for segmentation, then HOG feature was extracted, support vector machine (SVM) was used for classification on the basis of feature extraction, and finally helmet detection was carried out through color
feature. Ratta Poom Waranusast [2] put forward a method used in the detection of motorcycle helmets, they first use a hybrid model of tectonic background image, background deduction, and then the image binarization processing, connected component labeling and direction detection, and then classified using KNN(K-NearestNeighbor) will extract the object, and then extract the people head pixel information, again using KNN based on roundness, average each quadrant tonal identify wear a helmet and don't wear a helmet. These detection methods have some problems, such as complex steps, weak model expression ability and weak generalization ability. Although they perform well in distinguishing between wearing and not wearing a helmet, they lack semantic information, which makes it difficult to distinguish between correctly wearing a helmet and incorrectly wearing a helmet, and it is difficult to achieve satisfactory results in practical applications. With the continuous development of deep learning in the field of computer vision, deep learning has become the mainstream method in the field of image. Deep learning model has made great achievements in image classification, image segmentation, face recognition, target detection and other fields. In recent years, many excellent algorithms have emerged in the field of target detection, including algorithms based on candidate regions and classified, such as Faster R-CNN[3], Mask R-CNN [4] and regression target detection network, such as SSD[5]. Experiments show that [6] features extracted by CNN (convolutional neural network) are more effective than those designed by human.

3. Proposed an Approach for Safety Helmet Recognition

In the actual helmet wear detection, the detection speed is often as important as the accuracy. Based on the dual requirements of speed and accuracy of helmet wear detection, in this paper, we used Yolov4 [7] and adjusted and extended to develop a new helmet wear detection system to meet the requirements of this task.

In the construction of the actual work, often there will be a large difference of dark and light, our data set for photometric distortions, by adjusting the brightness, contrast, saturation and image noise to better simulate the scenario in real work. In addition, the input image is scaled, clipped and flipped to enhance the robustness and generalization of the model. These methods are commonly used in object detection image enhancement and usually retain the pixel distribution information in the original image area. Therefore, when dealing with image occlusion, target overlap and so on(Fig. 1), still need to use Random Erase[8], Grid Mask [9] and other methods, they are similar to the concept of Dropout in the feature extraction, selection of random or uniform in the image area, randomly will originally no pixel filling zero or replace the original pixel values to zero.

![Figure 1. Dense target overlap and object occlusion in the actual scene.](image)
Figure 2. Mosaic data enhancement.

There is another Mosaic data enhancement method (Fig. 2). Four images are randomly cropped and then pieced together into one image. As the pixel of the original image changes, the corresponding detection box will also change. This approach enriches the picture background and when it goes to the Batch Normalization, it also calculates four pictures, which in some ways increases the batch size without adding too much computing cost and still can be trained on a single GPU.

Bochkovskiy et al. [7] proposed that, compared with a classifier, an excellent detector needs to meet the following conditions, as shown in Table 1.

Based on Yolov3 Darknet53, Yolov4 added CSP module to make CSPDarknet53, which well solved the problem of high cost of reasoning calculation. The CSP [10] module first divides the feature mapping of the base layer into two parts, and then merges them through the cross-stage hierarchy, which can reduce the computation while ensuring the accuracy rate.

| Conditions                              | Function                                      |
|-----------------------------------------|-----------------------------------------------|
| Higher network input resolution         | Detect multiple small objects                 |
| More layers (greater receptive field)   | The increase of network input scale should be addressed |
| More parameters (greater network capacity) | Multiple objects of different scales are detected in a single image |

Table 1. Structure Selection of Detector

Figure 3. CSP module structure.

4. Experiment and Result
We have trained the helmet monitoring model based on Yolov4 on the Tesla V100 GPU, and achieved 95.1% accuracy. Compared with many network algorithms with excellent performance, our network model meets the requirements of helmet monitoring task in terms of speed and accuracy, as shown in Table 2. In the actual detection work, our network model can well complete the detection task (Fig. 4).
Table 2. Comparison of Network Model Speed and Accuracy

| Method            | FPS  | mAP@0.5 |
|-------------------|------|---------|
| Yolov4(ours)      | 68   | 95.1%   |
| Yolov3            | 61   | 90.4%   |
| Efficient det-0   | 57.6 | 92.0%   |
| Efficient det-1   | 50.3 | 94.7%   |

Figure 4. Actual Detection Effect.

5. Conclusion
Safety helmet wearing is an important safety issue in actual production, which is of great significance to the personal safety of workers. In this paper, a new helmet wear detection model is developed based on Yolov4, and the accuracy is 95.1% and the speed is 68FPS. It provides an intelligent method to protect the safety of workers, which is sufficient to replace the manual safety helmet wearing detection.

Acknowledgments
Thanks to Sichuan Agricultural University research project "Design and Implementation of Intelligent Helmet Monitoring System Based on Neural Network Algorithm and Sensor" for funding this project.

References
[1] Li J, Liu H, Wang T, Jiang M, Wang S, Li K and Zhao X 2017 Safety helmet wearing detection based on image processing and machine learning 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI) pp 201–5
[2] Waranusast R, Bundon N, Timtong V, Tangnoi C and Pattanathaburt P 2013 Machine vision techniques for motorcycle safety helmet detection International Conference of Image & Vision Computing New Zealand
[3] Ren S, He K, Girshick R and Sun J 2017 Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks IEEE Trans. Pattern Anal. Mach. Intell. 39 1137–49
[4] He K, Georgia G, Piotr D and Ross G 2018 Mask R-CNN IEEE Trans. Pattern Anal. Mach. Intell. 1–1
[5] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y and Berg A C 2016 SSD: Single Shot MultiBox Detector
[6] Wu H and Zhao J 2018 Automated visual helmet identification based on deep convolutional neural networks 13th International Symposium on Process Systems Engineering (PSE 2018) Computer Aided Chemical Engineering vol 44, ed M R Eden, M G Ierapetritou and G P Towler (Elsevier) pp 2299–304
[7] Bochkovskiy A, Wang C-Y and Liao H-Y M 2020 YOLOv4: Optimal Speed and Accuracy of Object Detection
[8] Zhong Z, Zheng L, Kang G, Li S and Yang Y 2017 Random Erasing Data Augmentation
[9] Chen P, Liu S, Zhao H and Jia J 2020 GridMask Data Augmentation
[10] Wang C-Y, Liao H-Y M, Yeh I-H, Wu Y-H, Chen P-Y and Hsieh J-W 2019 CSPNet: A New Backbone that can Enhance Learning Capability of CNN