Design and Implementation of Safety Helmet Detection System Based on YOLOv5

Yaqi Guan
Jianghan University
Artificial Intelligence Institute
Wuhan, China
13147191997@163.com

Wenqiang Li
Jianghan University
Artificial Intelligence Institute
Wuhan, China
552453124@qq.com

Tianyu Hu
Jianghan University
Artificial Intelligence Institute
Wuhan, China
457139049@qq.com

Qun Hou
Jianghan University
Artificial Intelligence Institute
Wuhan, China
houqun@jhun.edu.cn

2021 2nd Asia Conference on Computers and Communications (ACCC) | 978-1-6654-0743-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ACCC54619.2021.00018

***Abstract*—In order to reduce safety accidents caused by non-standard wearing of helmets, deep learning target detection technology is applied to construction safety detection scenarios, and a helmet detection algorithm based on YOLO v5 is proposed, which can realize real-time detection of helmet wearing. The deep learning part uses the K-means algorithm to cluster the dimensions of the target frame, and Yolov5s.pt is used for deep learning training. During training, the size of the input image is changed to increase the adaptability of the model, and the hyperparameters and optimizer are adjusted to be the best after improvement. The detection model has an accuracy rate of 90%, and the detection speed has reached 37.8fps, which meets the requirements of real-time detection of helmets. Through the combination of this model and hardware such as cameras, a real-time detection of whether a person wears a helmet is designed and implemented. The system realizes the three functions of picture detection, video detection and real-time monitoring.**

Keywords*—*deep learning, YOLOv5, helmet detection, image processing.

I. Introduction

At present, the country is developing and constructing, and the number of construction sites has increased. In order to prevent the dangers caused by the workers not wearing safety helmets, it is necessary to raise the importance of safety awareness of the construction personnel [1], and carry out necessary supervision and testing of the wearing of safety helmets. Nowadays, most of them use deep learning methods to train models, and they have been widely used in the related problems of target detection and recognition [2]. Liu Xiaohui and others use Hu moment feature vectors to realize the detection of helmets through support vector machine models[3]; Li Qirui studied the detection method of helmet wearing based on human body recognition, using HOG features to locate the head and then calculating the color attributes of the helmet, combining the two to generate the wearing detection result [4].

In order to make the trained model meet the speed and accuracy requirements of real-time detection, this paper proposes a helmet wearing detection method using an improved YOLOV5 model. This method mainly performs input size, initial candidate frame adjustment and loss function on the original YOLOV5 model The improvement of YOLOV5 makes the YOLOV5 model more suitable for the recognition of wearing helmets, and combines the trained model with the camera and other hardware to design a complete real-time detection system.

II. Deep Learning Based on Object Detection

Networks trained faster than on models like ResNe t,Faster-RCNN[5-7] have YOLOv3[8], but a disadvantage with lower accuracy. The main reason for the fast speed is mixing the detection and classification datasets and expanding the training set through a joint training algorithm.

The recognition model of YOLOv5 can ensure high accuracy at the speed of real-time detection. According to the network depth size and feature graph width, YOLOv5s is adopted as the use model. The network structure of YOLOv5 consists of four parts, namely Input Sector,Backbone Sector,Neck Sector,Prediction Sector.As shown in Fig.1.

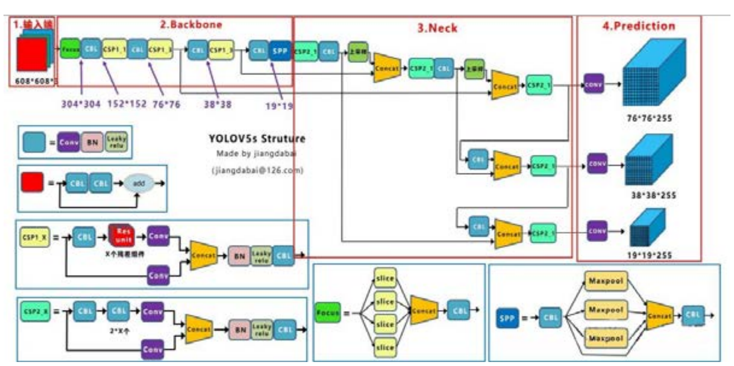

Fig. 1. The YOLOV5 network structure.

R. B. G. thanks Wuhan Municipal Education Bureau and Jianghan University

978-1-6654-0743-4/21/$31.00 ©2021 IEEE
DOI 10.1109/ACCC54619.2021.00018

III. IMPROVED YOLOV5 MODEL

Although YOLOV5 has the advantages of fast detection rate and high detection accuracy, in order to make the trained model better meet the needs of real-time detection and achieve better detection effect, we need to optimize the original YOLOv5 and constantly adjust the super parameters to get a better model.

A. Identification Process of Safety Helmet Wearing

When making helmet wear identification, first use the target detection weight with YOLOv5 to locate the face image with face identification, then use the helmet wearing recognition module to crop the detected face area, classify the image based on the helmet and no helmet, and finally get the image identification results of whether to wear the helmet or not. The helmet wear identification flow chart is shown in Fig.2.

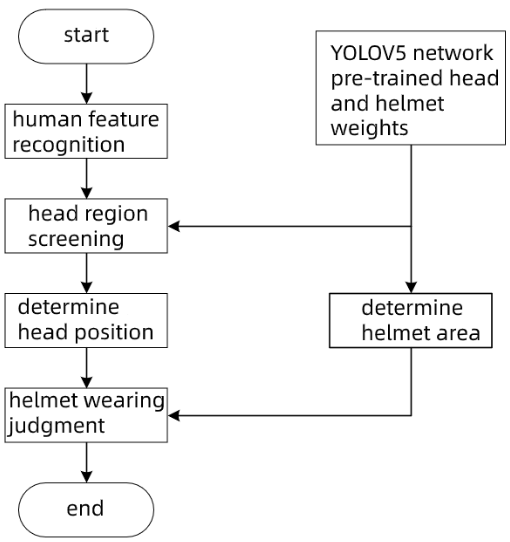

Fig. 2. Flow chart of safety helmet detection and wear identification.

B. Improvement of the Anchor Box

Both the YOLOv5 training stage and the prediction stage use the anchor box, training stage to find a pre-labeled anchor box, loss between this anchor box and ground truth at the feature map layer after a series of convolution and pooling, and the main purpose of training is to train the model parameters to fit the real border with anchor box. The prediction phase first generates multiple anchor box, in the image and then predicts the category and offset of these anchor box based on the trained model parameters, thus obtaining the predicted boundary frame.

The value of the anchor box directly affects the speed of the target recognition and the accuracy of the target box position, so it needs to adjust the optimal anchor box. according to its own dataset This paper selects the K-Means clustering algorithm to analyze the width of the labeled target box in the training set. According to the characteristics of the YOLOV5 model network structure, Locate the wide and high dimensions of the 9 cluster centers as the values of the anchor parameter in the network profile, Best Anchors based on the open-source hard hat dataset of Safety-Helmet-Wearing-dataset clustering [[14.74, 27.64], [23.48, 46.04], [28.88, 130.0], [39.33, 148.07], [52.62, 186.18], [62.33, 279.11], [85.19, 237.87], [88.0, 360.89], [145.33, 514.67]].

C. Selection of the Loss Function

Using GIOU_Loss as the loss function of Bounding box can overcome the problem that IOU cannot directly optimize parts without overlap, as well as retain an optimal box and suppress those redundant ones during the post-processing of target detection. The calculation expression of GIOU is shown in (1), give M,N two detection boxes, X the smallest closed shape, but X contains the ratio of M,N, calculation X does not contain M,N to x, and finally minus the IOU of M,N.

$$GIOU = IOU - \frac{|(M \cup N)/X|}{|X|} \tag{1}$$

GIOU solves the problem that IOU two boxes do not coincide, no matter how far the M,N two boxes are, but the larger the non-coincide values of the two boxes and the more GIOU tends to-1. GIOU is obtained by IOU minus a value so avoiding when Loss is equal to 0 when two boxes do not intersect. GIOU has directable properties, and when IOU=0, the expression is shown in (2):

$$GIOU = -1 + \frac{M \cup N}{X} \tag{2}$$

D. Hyperparameters and Optimizer Selection

The hyperparameters of the YOLOV5 network used in the test are 18, as shown in Table 1.

TABLE I. OVERPARAMETER CONFIGURATION

Hyperparameter	Value
Learning rate	0.01
Batch size	8
Epochs	50
Giou loss gain	0.05
Cls loss gain	0.50
Cls BCELoss positive_weight	1.00
Obj loss gain	1.00
Obj BCELoss positive_weight	1.00
Iou training threshold	0.20
Anchor-multiple threshold	4.00
Focal loss gamma	0.00
Image HSV-Hue augmentation	5.00
Image HSV-Saturation augmentation	0.70
Image HSV-Value augmentation	0.40
Image rotation	0.00
Image translation	0.00
Image scale	0.50
Image shear	0.00

The optimizer uses a SGD optimizer, where the momentum coefficient momentum and the weight decay coefficients are 0.937 and 0.0005, respectively.

IV. Analysis of the Training Results

A. Subjective Analysis

The test results of the test diagram in the experiment are shown in Fig.3.

It can be seen from the detection results that the algorithm extracts the overall features of the image comprehensively, even if the target is in a different environment or whether the target in the picture can be detected. Moreover, the bounding boxes generated by each grid obtains the continuous iteration and screening of the detection box with weighted nonmaximum suppression. Under the condition of certain occlusion, poor image clarity and more recognition objects, the algorithm can also accurately identify whether a safety helmet is worn.

Fig. 3. Model test results.

B. Model Evaluation Indicators

The model evaluation indicators used in this paper mainly include: accuracy, recall rate, average precision mean, and harmonic mean. The higher the accuracy and recall, the better the helmet recognition, but the negative correlation. Average precision mean and harmonic mean are quantitative indicators considering both accuracy and recall, and the greater their value, the better the helmet wear recognition effect.

C. Evaluation of the Experimental Results

The YOLOV5 model for helmet wear identification evaluates the model after 50 epoch training, and the results are shown in Fig.4.

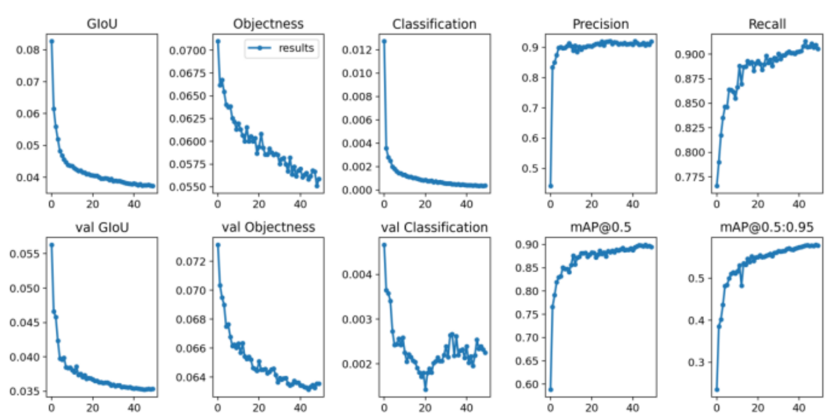

Fig. 4. Model evaluation results.

Fig.4 shows the evaluation results of the YOLOv5 model for helmet wear identification. The model achieves a convergence state after 50 epoch,. Its accuracy and recall improvement are very stable during model training. After the model reaches saturation, the accuracy (Precision) can be stable above 90%; the recall (Recall) can be stable near 90%, and the average accuracy mean and harmonic mean can remain at a high level. The mAP index performs the objective reaction algorithm, obtaining the head, safety helmet and overall mAP. As shown in Table 2.

TABLE II. SPOINTSTABLE

Classification	Precision.	Recall.	mAP.
All	90.7	91.1	90.1
Head	89.8	89.9	88.1
Safety helmet	91.5	92.2	91.7

After training, the mAP on the test set reached 90.1%, a high score, indicating that the YOLOv5 algorithm performed satisfactorily on helmet wearing detection. At the same time, the model is fast in practical application. When GTX 2070 graphics card is tested, its recognition speed can reach 37FPS, and basically meets the speed requirements of real-time detection. Due to the small volume of YOLOv5, it is easy to be applied in some industrial computers, resulting in an automated supervision that workers in production areas wear safety helmets.

V. Overall Framework Design and System Implementation

The software development environment is developed by Windows10, using Anaconda3 and Visual Studio Code. The functions realized by the software are: (1) helmet detection of local media files; (2) real-time detection of access camera.

A. Overall System Framework

Using image processing library functions such as opencv and numpy, the input media files are read frame by frame, then convert each frame picture into a matrix format and input to the neural network of YOLOv5 for processing. Finally, we output a vector that can specify the image information classification. The main process of the detection program is shown in Fig. 5.

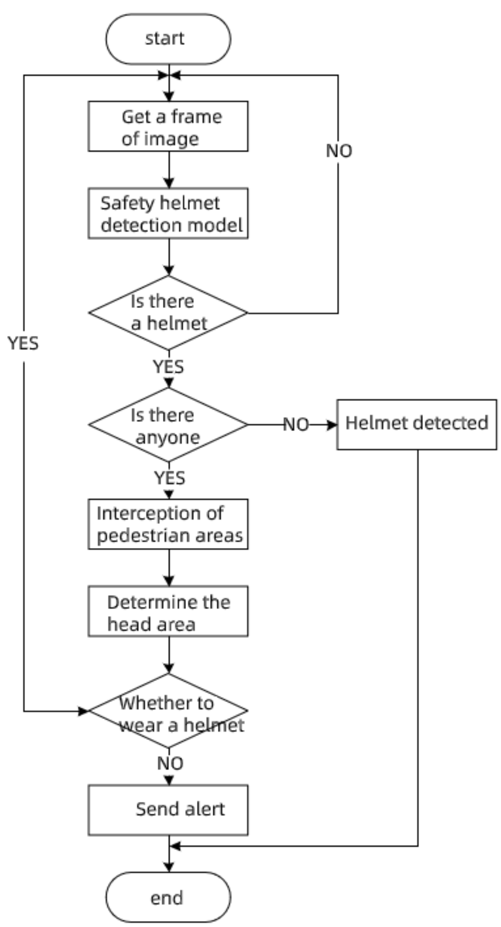

Fig. 5. Safety helmet detection process.

The image detection module writes the overall framework using Tensorboard, and the overall architecture of GUI, software using PYQT and Tkinter is as shown in Fig.6.

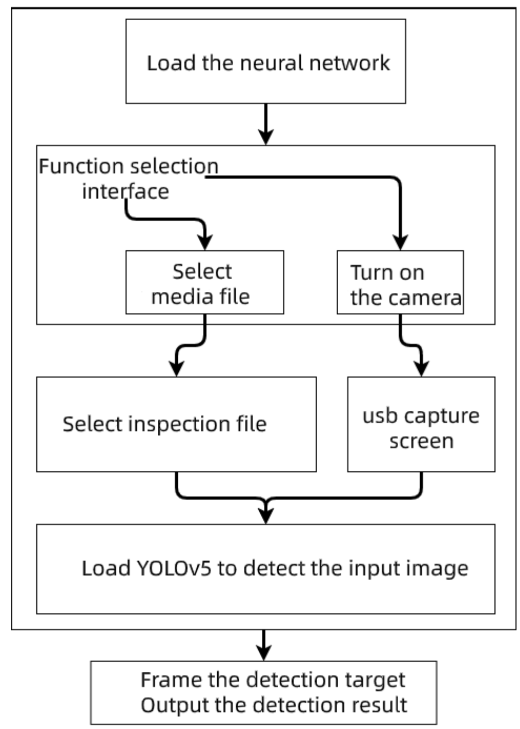

Fig. 6. Overall framework diagram of the software.

B. System Function Implementation

After opening the software, first load the YOLOv5's model and weight files, and enter the main interface. It is divided into media file detection and real-time camera detection. The user selects the required detection method through this interface. The main interface of the system is shown in Fig.7.

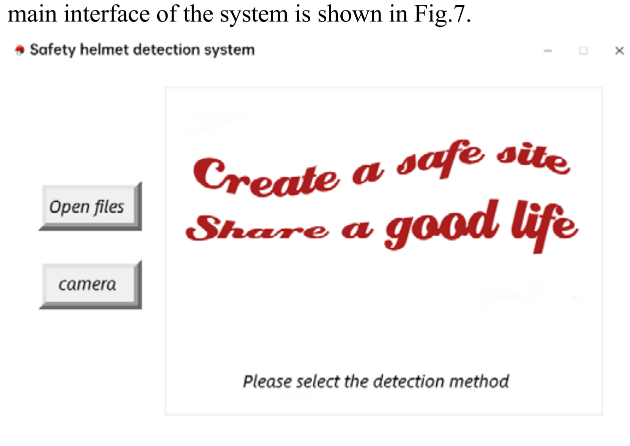

Fig. 7. Software main interface.

On the left side of the media file detection menu is the original file display area of the input image or video file. The right output box outputs the identified media file and marks the corresponding probability at the mark. The comparison interface for the input and outputs is shown in Fig.8.

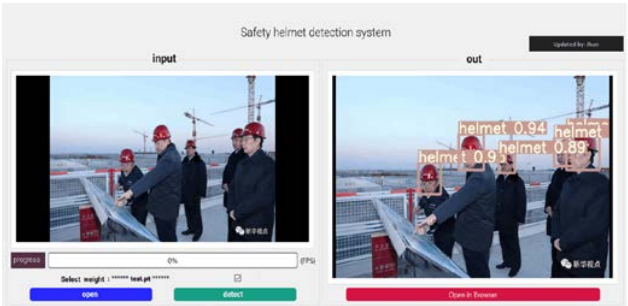

Fig. 8. Target image detection.

You can also turn on the camera for real-time detection. The detection results are marked in real time, and the detected head and helmet frame are marked with credibility. The detection effect is shown in Fig.9.

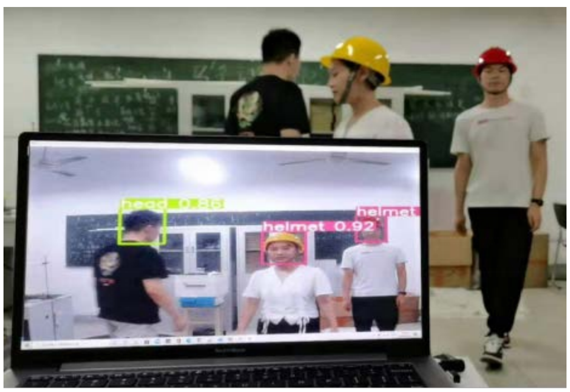

Fig. 9. Real-time camera detection.

This function can be used for real-time detection of the entrance and exit of the site to prevent personnel from entering the site without wearing a helmet and reduce the occurrence of safety accidents. The detected information will be output through the command line, and the software users can monitor the test results in real time.

VI. SUMMARY

This paper presents an optimized YOLOv5 algorithm for a helmet wear detection model. With the YOLO v5 network improved by improving the anchor box, loss function, adjusting the network architecture parameters and optimizer, the detection speed and accuracy are applied to the helmet real-time detection system.This helmet detection system has good stability and reliability, but there are still some aspects that can be continuously improved:

- Continue to improve the accuracy to reduce error detection and leak detection.
- Solve the system delay problem.
- Continue to optimize and expand the system functions.

ACKNOWLEDGMENT

Sponsored by the Wuhan Municipal Education Bureau (project number: 2019012); the open project of Wuhan Research Institute of Jianghan University (project number: IWHS20202006); and the Hubei Provincial Key Discipline Management Science and Engineering Open Project of Jianghan University.

REFERENCES

[1] He Jie. Study on the identification method of human abnormal behavior in video surveillance. [Chongqing University dissertation]. Chongqing: Chongqing University, 2014

[2] Zhang Hui, Wang Kunfeng, Wang Feiyue. Progress and prospect of the application of deep learning in object visual detection.

[3] Liu Xiaohui, Ye Xining. Application of skin color detection and Hu moment in safety helmet identification [J]. Journal of East China University of Technology, 2014,40 (3): 20-25.

[4] Li Chery. Research and Implementation of Hnet Video Detection System Based on Human Identification [D]. Chengdu, University of Electronic Technology, 2017.

[5] Peng Qiuchen, Song Yixu. Object identification and positioning of the based Mask R-CNN. Journal of Tsinghua University (Natural Science Edition), 2019,59 (2): 135-141

[6] Wang Chaoyang, Fan Shaosheng, Liu Zheng, etc. Abnormal state detection of overhead network based on FasterRCNN. Journal of Electricity, 2019,34 (04): 322-329

[7] Tian Yang, Qiu Ling-A road disease detection algorithm based on Fast-RCNN. Municipal technology

[8] Redmon J,Divvala S,Girshick R,et al. You only look once: Unified,real-time object detection[C].. Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas,NV,USA,2016: 779-788.