2022 4th International Conference on Communications, Information System and Computer Engineering (CISCE)

Research of target detection and distance measurement technology based on YOLOv5 and depth camera

Dian Yang[1]
Institute of Medical Support, Institute of Systems Engineering
Academy of Military Sciences
Tianjin 300161, China
1205688778@qq.com

Chen Su
Institute of Medical Support, Institute of Systems Engineering
Academy of Military Sciences
Tianjin 300161, China

Hang Wu
Institute of Medical Support, Institute of Systems Engineering
Academy of Military Sciences
Tianjin 300161, China

XinXi Xu
Institute of Medical Support, Institute of Systems Engineering
Academy of Military Sciences
Tianjin 300161, China

XiuGuo Zhao*
Institute of Medical Support, Institute of Systems Engineering
Academy of Military Sciences
Tianjin 300161, China
* Corresponding author: zhaoxg2011@126.com

Abstract—**Target detection and distance measurement technology based on YOLOv5 and depth camera is proposed for positioning the shelter in the field shelter hospitals deployment process by the shelter-transporting automated guided vehicle (AGV). This paper studies the principle of YOLOv5 target detection and the depth camera imaging and ranging model. The square capsule data set and the engineering implementation code of target detection and ranging technology are formed. The scenario test of AGV's positioning to the square capsule is completed. The test results show that the target detection and distance measurement technology proposed in this paper can be better applied to the actual engineering of AGVs for the positioning of the shelter, and the average accuracy of target detection is more than 95%. The distance measurement error is less than 0.1 m. The combination of target detection technology and depth camera can be effectively applied to the engineering practice, improving the field shelter hospitals' rapid deployment capability and intelligence.**

Keywords: target detection; depth camera; YOLOv5; AGV

I. Introduction

The field shelter hospitals are critical equipment to support field medical treatment and disaster medical rescue. It plays a significant role in modern warfare and non-war military operations. The deployment speed of the field shelter hospitals is decisive for the rapid generation of its treatment capability, and the faster its deployment speed, the more it can from the treatment capability in a short time [1~4]. The efficiency and positioning accuracy of the shelter transferring has a decisive impact on the deployment speed of the field shelter hospitals. A shelter-transporting automated guided vehicle (AGV) is an intelligent transfer tool that can ensure faster and more accurate transfer with brilliant navigation capability while saving human and material resources [5]. Zhang et al. proposed a sharp binocular stereo vision ranging based on YOLOv5, which solved the problems of traditional methods to calibrate the camera and the high time complexity of the stereo matching algorithm. However, the binocular camera is influenced by the environment, and the size of the binocular baseline limits the ranging distance [6]. The monocular visual depth estimation algorithm proposed by Fu et al. has a significant error due to the defects of the monocular camera [7]. The automatic guided vehicle for field medical shelter transfer studied by Qu et al. is less intelligent and relies only on GPS to obtain positioning information, lacking general reliability [8].

For this reason, there is an urgent need to study a reliable and stable technology for the positioning of the shelter. In this paper, by studying the principles of target detection and distance acquisition, propose a shelter positioning technology based on the combination of YOLOv5 and depth camera. The research in this paper is of great significance for the treatment of casualties on battlefields and natural disasters and infectious patients in public health events. It is essential for improving the capability of health service protection.

II. Materials and Methods

A. YOLOv5

Object detection is a computer vision task that distinguishes objects in an image or video from other regions of no interest, locates objects, and identifies object types. As deep learning is developing faster and faster, deep learning-based target detection with higher detection efficiency has developed comprehensively. Many scholars have conducted in-depth research and obtained some deep learning target, detection models.

2022 4th International Conference on Communications, Information System and Computer Engineering (CISCE) | 978-1-6654-9848-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/CISCE55963.2022.9851025

Authorized licensed use limited to: Universidad Nacional de Colombia (UNAL). Downloaded on February 10,2023 at 03:44:45 UTC from IEEE Xplore. Restrictions apply.

The target detection algorithms based on deep learning models can be divided into two main categories: two-stage detection algorithms and one-stage detection algorithms. One-stage detection algorithms are most representative of the YOLO series. In June 2020, the Ultralytics team proposed the YOLOv5 model, introducing a new Pytorch training and deployment framework, enabling custom models to be used in the detection process [9]. The YOLOv5 algorithm has four network structures, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, which differ in the width and depth of the CSP structure in the Backbone and Neck parts, i.e., the number of convolutions and the number of residual blocks is different. The YOLOv5 simple network structure is shown in Fig. 1, and its network model is divided into four parts, namely Input, Backbone, Neck, and Prediction.

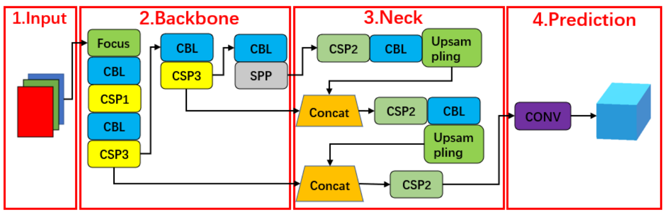


Figure 1. YOLOv5 network structure diagram

B. Depth Camera

Implementing target detection and ranging technology is inseparable from measuring feature objects by cameras to obtain color and depth information of things simultaneously. This paper uses an RGBD depth camera as a target detection tool. Compared with the traditional monocular and binocular cameras, the most important feature and function of depth cameras are that they can actively acquire the depth information of images. In this paper, the Intel REALSENSE D455 structured light camera is used to obtain the depth information of the target, as shown in Fig. 2. The left and right imager and IR projector measure the metrology information. The camera is connected to the IPC through the USB interface.

Figure 2. RealSense D455 structured light camera

The RGBD model includes a pinhole camera and a depth information calculation model. The schematic diagram of the pinhole camera model is shown in Fig. 3, and the depth information calculation model is shown in Fig. 4.

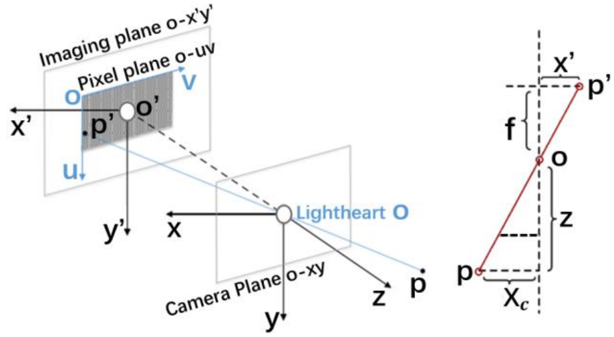


Figure 3. Pinhole camera model

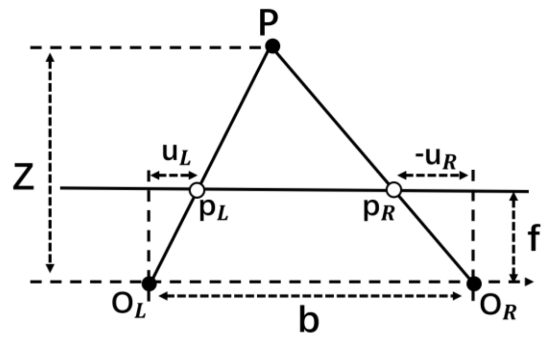


Figure 4. Depth information calculation model

C. Programming and physical

The development system of the target detection and ranging program is the Ubuntu system, and the development language is Python. Download the target detection model in GitHub YOLOv5. Download the D455 working program from the RealSense website. The two programs are fused, and the RGB images in the D455 working program are acquired as the target detection input of YOLOv5. After completing the target detection, the distance from the camera to the shelter is received by the D455 depth reading program, which achieves the positioning of the shelter. Its workflow is as follows:

- AGV movement.
- Depth camera work.
- Detecting the square module in the video stream.
- Getting depth distance information.
- Performing shelter positioning.

Next, after completing the program design, the depth camera D455 is installed on the AGV. Then, the target detection and ranging program is written to the AGV IPC for engineering applications. The installation schematic is shown in Fig. 5 below.

Authorized licensed use limited to: Universidad Nacional de Colombia (UNAL). Downloaded on February 10,2023 at 03:44:45 UTC from IEEE Xplore. Restrictions apply.

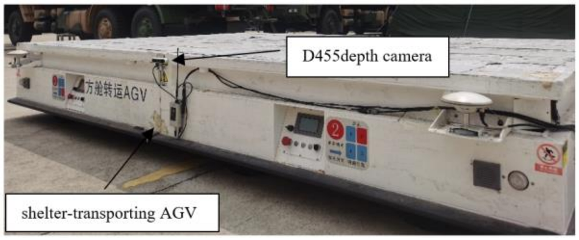


Figure 5. installation schematic

D. Model Training

Before conducting the model training, 1000 images are created for the shelter dataset, and then all the images in the dataset are labeled by Labelimg. The experimental environment is the Ubuntu18.04 operating system based on the Pytorch framework. CPU: Intel Core I9-10900K, GPU: NVIDIA RTX 3090, 24GB. Training parameter settings for ten pre-training models of YOLOv5 are shown in Table 1.

TABLE I. EXPERIMENTAL TRAINING PARAMETERS

Parameter	Size
Epochs	1000
Batch-size	32
Learning rate	0.01
Decay	0.0005

Among the YOLOv5 series algorithms, YOLOv5l has both detection performance and speed, and YOLOv5l can be better apply it to practical engineering. Therefore, in this paper, YOLOv5l is selected as the target detection algorithm, and it is trained using the shelter data. The trend of each parameter change during the training process is shown in Fig. 6, from which YOLOv5l can conclude that the curve fit state is good, and the parameters tend to be stable. So, the model obtained by using YOLOv5l training data is available.

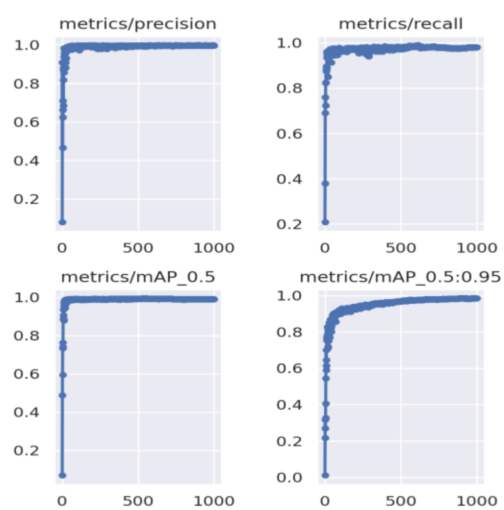


Figure 6. The trend of parameter changes during training

E. Methods

In order to evaluate the accuracy of the test results, this paper applies a tape measure to mark the critical locations on the ground and use them as the actual distance values. The marking process is shown in Fig. 7.

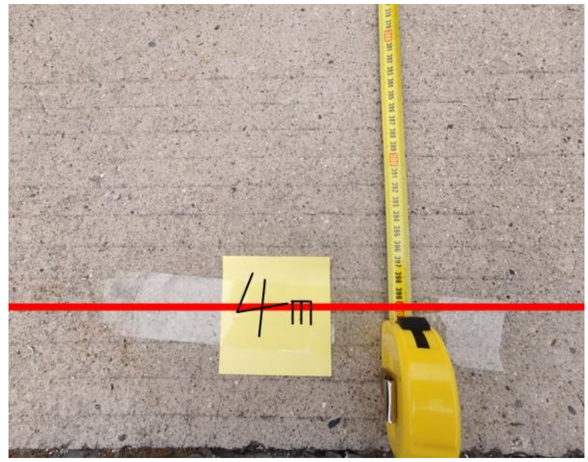
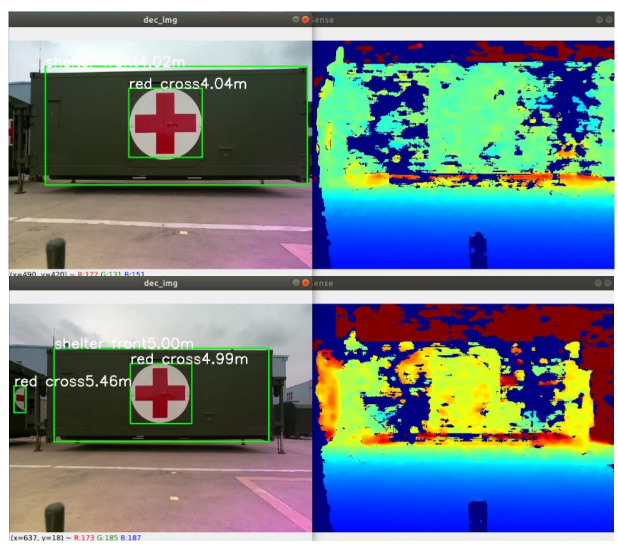


Figure 7. Schematic diagram of the marking process

III. RESULTS & DISCUSSION

In practical engineering applications, the target detection technology steadily identifies the shelter, and then the active vision ranging technology obtains information on the distance between the AGV and the shelter. The integration of the two technologies improves the ability of AGV to transfer the shelter.

The experimental results are shown in Fig. 8. The left is the RGB pixel information acquired by the depth camera, and the right is the depth D information received by the depth camera. Different color areas in the depth map represent different distances.



Authorized licensed use limited to: Universidad Nacional de Colombia (UNAL). Downloaded on February 10,2023 at 03:44:45 UTC from IEEE Xplore. Restrictions apply.

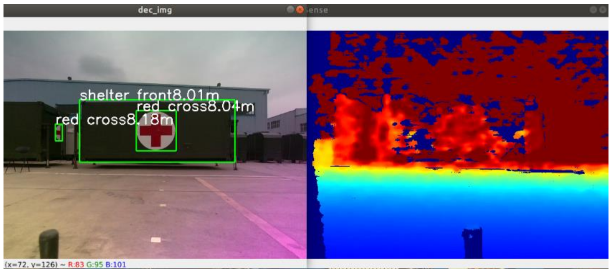


Figure 8. Experiment to obtain information results

The AGV moves from a distance of 10 m to a length of 4 m. The AGV performs target detection and ranging procedure and acquires image and distance information through a laptop. Data are collected ten times at 1m intervals, and the distance error is shown in Fig. 9 below, the maximum error variation curve is shown in Fig. 10.

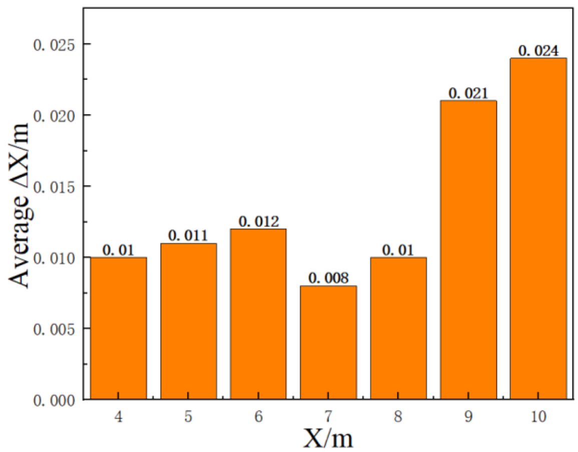


Figure 9. Distance measurement average error graph

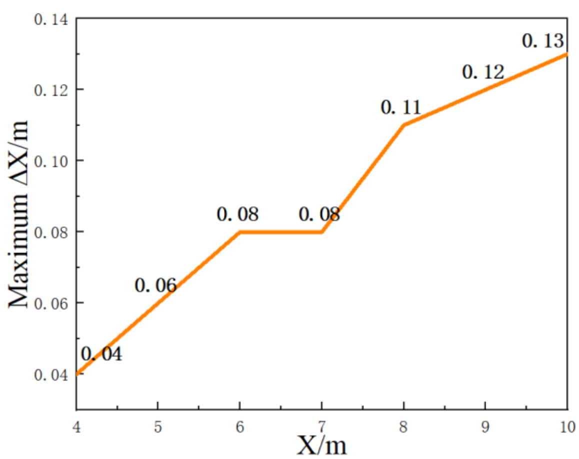


Figure 10. the maximum error variation curve

The average error of the target detection and ranging program is less than 0.025m at a distance of 4~10m from the shelter, and the maximum ranging error increases with the distance but remains around 1%. In general, the error between the range data and the actual data in the test is small, which indicates that the reliability and stability of the target detection and range program are reasonable. However, the error gradually increases at longer distances. The reason is that the depth camera ranging by structured light is not accurate at long distances.

IV. Conclusions

Target detection and distance measurement technology based on the combination of YOLOv5 and depth camera is proposed for the positioning problem of the AGV opponent cabin. In this paper, the target detection and ranging principle is analyzed. The target detection and distance measurement program is designed to realize the practical engineering combination of the D455 depth camera and AGV. An excellent model is obtained by training YOLOv5l with an original square pod dataset. The experimental results show that the range values obtained by the target detection and distance measurement procedure have an error of about 1% from the actual values. The target detection and distance measurement is performed under normal environmental conditions. Its performance and reliability under dense obstacles and rainy and exposed environmental conditions still need to be further verified, and in-depth follow-up studies are needed. The target detection and distance measurement technology can effectively improve the intelligence of AGV. It is imperative to enhance the non-war and wartime defense and security capabilities.

References

[1] Zhou Ying, Wang Lingling, Xu Yayin, Ding Lieyun, Tang Zhouping, “Intelligent fangcang shelter hospital systems for major public health emergencies: The Case of the Optics Valley Fangcang Shelter Hospital,” Journal of Management in Engineering,2022,38(1).

[2] “Emergency biosafety management practice in laboratory of shelter hospital,” Current medical science,2020,40(5).

[3] Simiao Chen, Zongjiu Zhang, Juntao Yang, Jian Wang, Xiaohui Zhai, Till Bärnighausen et al, “Fangcang shelter hospitals: a novel concept for responding to public health emergencies,” The Lancet,2020,395(10232).

[4] Song Bai, Bao-guo Yu, Yong-Zhong Zhang, Hui Ding, Zhou-wei Wu, Bin Fan et al, “Challenges of treating adenovirus infection: application of a deployable rapid-assembly shelter hospital,” Disaster Medicine and Public Health Preparedness,2017,12(1).

[5] D. Morikawa, T. Nakatani, T. Hirogaki, E. Aoyama, "Simultaneous planning method for number and allocation of AGVs in an AGV control system under uncertain transportation conditions," 2020 20th International Conference on Control, Automation and Systems (ICCAS), 2020, pp. 41-46.

[6] Zhongnan Zhang, Wei Huo, Ming Lian, Lei Yang, "Research on Fast Binocular Stereo Vision Ranging Based on YOLOv5," Journal of Qingdao University (Engineering Technology Edition), 2021,(02):20-27.(In Chinese)

[7] Q. Fu, J. Chen, W. Yang and S. Zheng, "Nearshore ship detection on SAR image based on YOLOv5," 2021 2nd China International SAR Symposium (CISS), 2021, pp. 1-4.

[8] Pengtao Qu, Hang Wu, Xinxi Xu, Chen Su, Sheng Gao, Xiuguo Zhao, "Research and design of automatic guided vehicle for field medical cabin transport, " Military Medicine: 1-7. (In Chinese)

[9] F. Majeed, F. Z. Khan, M. J. Iqbal and M. Nazir, "Real-time surveillance system based on facial recognition using YOLOv5," 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), 2021, pp. 1-6.

Authorized licensed use limited to: Universidad Nacional de Colombia (UNAL). Downloaded on February 10,2023 at 03:44:45 UTC from IEEE Xplore. Restrictions apply.