

# An Intelligent Equipment Inspection System based on Robotic Arm

Zewei Ye<sup>1,2,3</sup>, Liwei Zeng<sup>1,2,3,a,\*</sup>, Qingzhang Wang<sup>1,2,3</sup>, Zhicheng Wang<sup>1,2,3</sup>, Jianhua Cui<sup>1,2,3</sup>, Yanning Cao<sup>4</sup>, Lingling Shen<sup>1,2,3,b,\*</sup>, Xiaojun Qian<sup>1,2,3</sup>

<sup>1</sup> School of Computer and Electronic Information, Nanjing Normal University, Nanjing, China

<sup>2</sup> School of Artificial Intelligence, Nanjing Normal University, Nanjing, China

<sup>3</sup> Artificial Intelligence Research Institute, Nanjing Normal University, Nanjing, China

<sup>4</sup> Honors College, Nanjing Normal University, Nanjing, China

a,\* 1426912622@qq.com, b,\* llshen509@163.com

\* Corresponding author: Liwei Zeng and Lingling Shen

**Abstract:** Intelligent robots have replaced manual labor to complete a lot of complicated work, especially in equipment inspection has advantages. The purpose of this paper is to design an intelligent robotic inspection system. The system is based on ROS and adopts YOLO algorithm, RRT\* algorithm, Agile\_grasp function package, and MoveIt! framework to realize the basic target recognition, position recognition, robotic arm navigation planning and other functions. The human-computer interaction interface has been designed, and the simulation test in Gazebo and the real machine test in real environment have also been completed. The system still has a large space for improvement, and further research will be conducted in the future in terms of response speed and accuracy.

**Keywords:** Robotic Arms; YOLO; RRT\*; QT5; RVIZ; GAZEBO.

## 1. Introduction

Nowadays, robotic arms have been widely used in production and life, such as manufacturing, intelligent service areas. The most well-known is the Tesla Shanghai factory of high-efficiency industrial robots and gradually into people's homes cleaning robots, medical robots and other service robots.

In addition, the field of intelligent management. Intelligent robots have replaced manual labor to complete a lot of complicated work, especially in equipment inspection has advantages. Current general robots lack the ability to handle emergencies such as power leaks, fires, equipment downtime, etc. on time. However, the failure of critical equipment within a facility can lead to production stoppages and increased costs, so the ability to quickly and accurately inspect equipment and deal with problems on time is essential. For example, short-term disablement of electronic information facilities such as cloud computing and cloud storage servers may cause huge economic losses and social impacts. Therefore, the intelligent inspection system for equipment based on robotic arm has become an important research direction to solve such problems.

In this paper, an intelligent system using a robotic arm for equipment inspection is designed, which can effectively solve such problems. YOLO algorithm and RRT\* algorithm are used in this system with the ROS environment based on an in-depth study of the critical equipment environment.

## 2. Related Technologies

### 2.1. Deep Learning based Target Detection Algorithm

YOLO (You Only Look Once) algorithm is an end-to-end target detection algorithm widely used in computer vision.

Compared with traditional target detection algorithms, YOLO algorithm has faster processing speed and higher accuracy.

The core idea of the YOLO algorithm is to transform the target detection task into a regression problem by simultaneously predicting the location and category of a target through a single neural network. The YOLO algorithm divides the input image into fixed-size grids and predicts a set of bounding boxes on each grid. Each bounding box contains the location and size of the target and predicts the category of the target that appears in each bounding box.

The v3 version used the Darknet-53 network with 53 convolutional layers, and the subsequent versions of the convolutional layers have been added, bringing higher accuracy while placing higher demands on the hardware. So based on the available hardware, the v3 version was chosen for this paper.

The advantages of YOLO are speed and efficiency, which are very suitable for real-time application scenarios such as robotic arm navigation. However, the current version of YOLOv3 has poor detection performance for small and overlapping targets [1].

### 2.2. RRT\* Path Planning Algorithm

Another important part of the system is the motion planning of the robotic arm. Steven M. LaValle proposed the RRT algorithm, which generates paths using random sampling for navigation. However, the quality of the paths generated by the RRT algorithm cannot meet the actual needs in the complex and unstructured server setting environment [2].

In order to improve the performance of the RRT algorithm, this paper adopts a more advanced solution, namely the RRT\* algorithm. The RRT\* algorithm is an improved version of the RRT, which can obtain better path planning results by globally optimizing the tree structure during the search process[3]. On the one hand, the RRT\* algorithm performs

some heuristic planning before generating the tree structure in a randomized way, and at the same time calculates the cost function of the paths, and the algorithm prioritizes the direction of expansion with less cost when the nodes are expanded, in order to reduce the cost of the paths. On the other hand, the algorithm structure of RRT\* is more complex, but it is able to converge to the optimal solution in an asymptotic sense. It employs incremental expansion, i.e., it expands one step at a time, so the algorithm can cover the entire space and find a better path over a longer period [4].

### 2.3. Grabbing Algorithm Agile\_grasp

Agile\_grasp is an object grasping pose detection algorithm based on deep learning and geometric computation. The algorithm obtains the 3D pose information of an object by processing the object point cloud data. The algorithm is mainly applied in robot autonomous grasping tasks.

Specifically, the agile\_grasp algorithm first predicts a number of candidate grasping poses using a neural network, then scores each pose using a scoring function based on geometry and contact stability, and selects the pose with the highest score as the final grasping pose. The algorithm is implemented using the TensorFlow framework, which employs Convolutional Neural Network (CNN) and Fully Connected Neural Network (FCN) as neural network models.

## 3. System Design

### 3.1. Overall System Design

The system aims to realize the inspection function of key electronic and electrical energy equipment, including timely detection, handling of simple faults, and feedback of fault causes. By combining RGB camera and depth camera, utilizing the target recognition module and robotic arm module, it realizes the recognition of cabinet and cabinet handles, the acquisition of position information, as well as the functions of robotic arm's navigation, motion solving, motion obstacle avoidance, and end-effector control through MoveIt! The overall architecture is based on the ROS operating system, with the help of RVIZ 3D visualization platform to monitor the robotic arm, and the MoveIt! tool for integrated development and control of the robotic arm.

Table 1. Items required by the system

Object Name	Type	Parameter
Six-axis Collaborative Robots	Lebai LM3	Payload:3kg Working radius:638mm
Electric Gripper	Lebai LMG-90	NONE
Kinect Camera	NONE	NONE
IPC	DELL G15	NONE

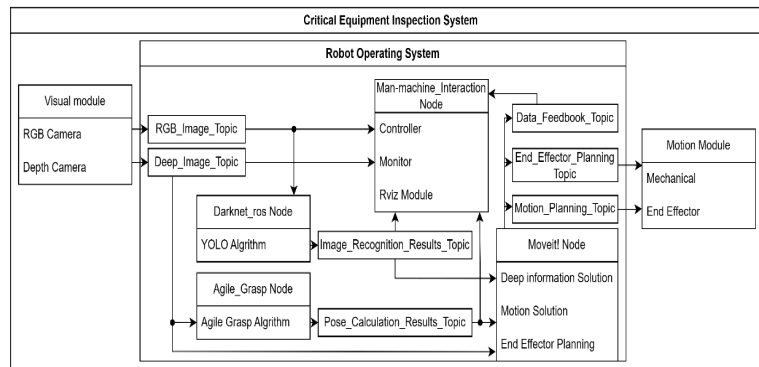


Fig 1. Functional structure diagram of key equipment inspection system

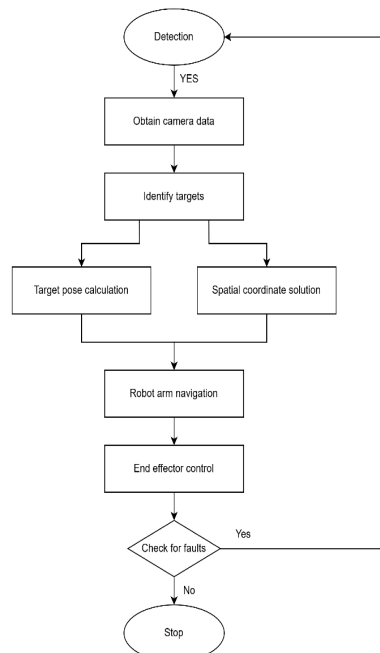


Fig 2. System Flow Chart

### 3.2. Human-machine Interface

This paper uses QT5 to complete the human-computer interaction interface, which can realize the functions of user login, robotic arm monitoring and robotic arm control. After entering the account password and successfully logging in, click Connect to connect to the local ROS Master and operate the robotic arm. First of all, this paper sets up a monitor containing camera images, YOLO recognition results, and a real-time position model of the robotic arm. At the same time, it is possible to plan by manipulating the position model. Then there is a manual control window, which contains the robotic arm joint adjustment, end position control and end actuator control. Finally, some fixed behavior settings are set, and the self-starting operation of the robotic arm is also realized by it.

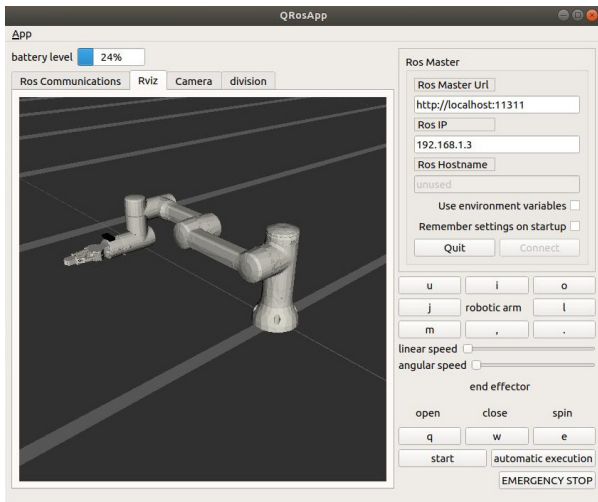


Fig 3. Control Interface

### 3.3. Target Identification

It is important to note that the YOLO algorithm has a problem - it is not very accurate in recognizing small targets. Since the topmost feature map is extracted in the native YOLO algorithm, this means that details are ignored in the constant filtering operation, and thus small targets cannot be extracted in the final feature map. To solve this problem, the following scheme is proposed in this paper.

Adjust the parameters of the Darknet model to adapt it to the scene. For small target detection, higher resolution images can be input, and it should be noted that this will undoubtedly significantly increase the computational difficulty. Multi-scale feature extraction can also be used. Using a multi-scale feature pyramid, detection is performed on different levels of the feature map to ensure that small targets can be captured. For overlapping target detection optimization, the IoU threshold can be adjusted. IoU, which refers to the intersection and concurrency ratio, is used as a measure of how much two bounding boxes overlap. A lower intersection and merge ratio threshold preserves overlapping information, but may produce redundant results[5].

Introducing new mechanisms - Attention Mechanisms, including Channel Attention, Spatial Attention, and Hybrid Attention. Attention mechanisms have better results for small target detection. The idea is to weight the input and then output, want the network to pay attention to the region to give a greater weight, and do not want the region side to give a smaller weight. Channel attention i.e. weighting the input image channels and then output. Spatial attention that is for

the input image region weighted and then output, the representative model is the spatial transformation neural network, but also a kind of neural network algorithm, can be a variety of data in the space conversion, and automatically get the important region. And the hybrid attention mechanism is the combination of two attention mechanisms[6].

### 3.4. Robotic Arm Motion Planning

MoveIt! is one of the most versatile robotic arm planning and control packages in ROS, providing a powerful set of tools for motion planning, motion control, collision detection, visualization and simulation for robotic systems. In terms of motion planning, Moveit! provides a flexible motion planning framework that can support a wide range of robotic arm and actuator types, including support for multi-degree-of-freedom, series and parallel robotic arm planning. It contains a variety of motion planning algorithms, including the RRT and RRT\* algorithms that need to be used in this paper. The visualization and simulation tools of Moveit! make it possible to visualize the state of the robotic arm and the environment for easy path planning and debugging[7].

### 3.5. End-effector Control

When the robotic arm moves to the right position, it will be operated by the end-effector.

The robotic arm is controlled using the MoveIt! function package, which can be utilized to complete the MoveIt! robot framework to achieve the recognition of the target positional attitude using `moveit_simple_grasps`. Although the function package provides easy to use and convenient ports to realize the operation of end-effector, and it can set rich end-effector movements and postures. However, there are shortcomings in some functions, such as the accuracy of unstructured object position recognition is not high enough. In this regard, the mentioned `agile_grasp` algorithm can be used.

First, `agile_grasp` implements an agile grasping method based on depth images. Its main process is to receive the input RGB-D image, followed by preprocessing the depth image, and then by using a segmentation algorithm, the objects in the depth image are segmented from the background for the subsequent grasping bit position calculation. On the basis of object segmentation, `agile_grasp` extracts candidates for grasping poses from the segmented image using a special image feature extraction method such as depth-channel based voting method. The grasping poses are filtered and scored to select the best grasping pose [8].

## 4. Test

### 4.1. Simulation and Analysis

Before the real machine test, this paper conducts a simulation experiment, the simulation platform used in this experiment is the joint simulation platform of RVIZ and GAZEBO in ROS, which is aimed at verifying whether the co-operation between the robotic arm and the working scenario can be carried out smoothly.

RVIZ (ROS Visualization) is a visualization tool in ROS, which is mainly used for presenting sensory information, state data, and simulation results of a robot system. It provides a rich set of functions, including 3D scene visualization, motion planning and control of robotic arms, and visualization of sensor data.

And GAZEBO is a widely used open source robotics

simulation platform that creates highly realistic physical simulation environments, including a wide range of sensors, objects and robot models, and provides a physics engine for simulating real-world physical motion.

Firstly, this paper carried out the model building and configuration based on the URDF (Unified Robot Description Format) files of the robotic arm and the scene, including the size, weight, material and other parameters of the robotic arm and other devices, as well as the establishment of the environment model. After completing the configuration, RVIZ in ROS was used to carry out the visualization test of some functions, such as the forward and reverse kinematics calculation of the robotic arm, collision detection, and so on.

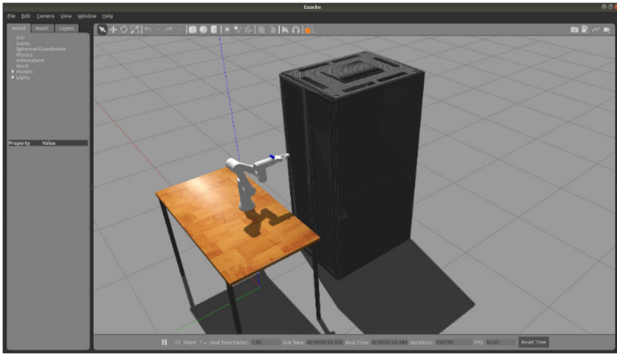


Fig 4. Gazebo Simulation and Robotic Arm Navigation

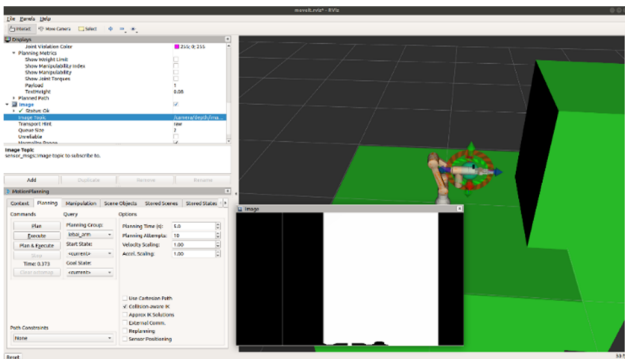


Fig 5. Testing the MoveIt! Planning Function in the RVIZ Interface

## 4.2. Real Machine Testing

In this paper, the motion planning test of a real environment robotic arm is completed under controlled conditions with MoveIt!. In the actual deployment, it is necessary to set up the motion controller in ROS, as well as the corresponding hardware interface of the robotic arm, and the joints of the robotic arm are directly controlled by the ROS motion controller. In addition, it is also necessary to set up the state feeder to provide real-time feedback on the state of the joints for planning purposes. The most important part lies in the configuration of MoveIt! which will directly determine the way the robotic arm is planned.

## 5. Further Work

At present, the system designed in this paper still has many problems. For example, the target recognition algorithm and the position recognition algorithm are too redundant, which can be realized by expanding the dataset and using the YOLO algorithm to achieve target detection and position detection in one step. The end-effector type is less and single function. The structure of end-effector can be improved or end-effector can be added to enrich the function. For example, adding an infrared camera can detect the heat location and temperature

of the server and find the problem faster. For example, add a suction device to realize the operation of dust cleaning and so on. The mechanical arm is fixed, and the flexibility is greatly limited. The arm can be mounted on a chassis or a track to realize spatial mobility.



Fig 6. Real Machine Operation Scenarios

With the further development of the information age, the demand for large-scale server hardware maintenance will become greater and greater. And most of these maintenance behaviors are mechanical, the use of robotic arms to complete simple maintenance work can effectively reduce labor costs. Moreover, the industry has a low degree of development and has great potential for development.

## Acknowledgments

This work was financially supported by Innovation Training Fund of Artificial Intelligence Research Institute, Nanjing Normal University.

## References

- [1] Ye A, Pang B, Jin Y, et al. A YOLO-based neural network with VAE for intelligent garbage detection and classification [C]// 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, 2020:1-7.
- [2] Ramses Reyes, Israel Becerra, Rafael Murrieta-Cid, Seth Hutchinson. Visual-RRT: Integrating IBVS as a steering method in an RRT planner. 2023:09-22.
- [3] Fangbin Wang, Yefei Gao, Zhong Chen, Xue Gong, Darong Zhu, Wanlin Cong. A Path Planning Algorithm of Inspection Robots for Solar Power Plants Based on Improved RRT\*. 2023:11-01.
- [4] WU Y, XIE X J. Robust adaptive control for state - constrained nonlinear systems with input saturation and unknown control direction[J]. IEEE Transactions on systems, man, and cybernetics: systems, 2021, 51(2):1192-1202.
- [5] Gong H, Liu T, Luo T, et al. Based on FCN and DenseNet Framework for the Research of Rice Pest Identification Methods [J]. Agronomy, 2023, 13(2): 410-423.
- [6] Tian L, Cao Y, He B, et al. Image Enhancement Driven by Object Characteristics and Dense Feature Reuse Network for Ship Target Detection in Remote Sensing Imagery[J]. Remote Sensing, 2021, 13(7):1327.
- [7] LIU G L, HE H Y, TIAN G H, et al. Online collision avoidance for human-robot collaborative interaction concerning safety and efficient [C]//2020 International Conference on Advanced Intelligent Mechatronics, 2020.
- [8] Tulbure A A, Tulbure A A, Henrietta E D. A review on modern defect detection models using DCNNs-Deep convolutional neural networks[J]. Journal of Advanced Research, 2021, 35: 33-48.