International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 2, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n2.06
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Fuel Tank Position Localization Based on Machine Vision

Changjiang Song

Arizona Collage of Technology at Hebei University of Technology, Tianjin, China
715362264@qq.com

ABSTRACT

Refueling robots, as an important part of intelligent service systems, greatly enhance the safety and efficiency of the refueling process. This study focuses on the machine vision module in refueling robots, specifically exploring the application of high-resolution cameras and LiDAR in data acquisition. We used Convolutional Neural Networks (CNNs) such as ResNet and MobileNet for feature extraction, which ensured high-precision recognition and classification in a variety of environments. Meanwhile, the target detection module uses YOLOv4 and MobileNet3 with fast and accurate target localization capabilities to effectively identify and calibrate the location of fueling ports. In addition, we introduced Extended Kalman Filter (EKF) and Bayesian Filter algorithms for data fusion and state estimation, which improves the robustness and reliability of the system. By combining these advanced vision techniques and algorithms, the refueling robot realizes efficient and accurate automatic refueling operation. This research provides theoretical and technical support for the further development of intelligent refueling robots so that they can still operate stably in complex environments.

KEYWORDS

Machine Vision; ResNet; Convolutional Neural Networks; Kalman Filter (EKF); Bayesian Filtering Algorithm; Refueling Robot; Robustness

1. INTRODUCTION

1.1. Background

As an important part of the transportation system, gas stations need to serve a large number of vehicles every day. However, the traditional way of refueling relies mainly on manual operation, which has many problems. And as the population decreases labor is even more scarce, in the case of rising labor costs, gas stations need to hire a large number of employees to ensure the quality of service, increasing operating costs. Moreover, there are some safety hazards associated with manual refueling. There are certain risks associated with manual operation, such as improper operation that may lead to oil leakage, fire or explosion and other safety accidents. Long waiting time during peak hours and poor user experience under different weather conditions, especially in bad weather.

Many users are reluctant to get out of their vehicles to refuel, especially in cold or hot weather. In recent years robotics technology has made significant progress, including breakthroughs in artificial intelligence, machine vision, and robot control systems, promoting the development and application of refueling robots Among the rapid development of machine vision, machine vision systems can help refueling robots detect the location of the vehicle's refueling ports to ensure accurate docking. This includes identifying the vehicle model, fueling port location, and the shape and size of the fueling port, but also real-time monitoring of the surrounding environment, identifying obstacles and other vehicles, to ensure that the refueling robot avoids collisions during movement and operates safely. Through high-precision image processing technology, machine vision can accurately locate the refueling port and guide the robot arm to complete the refueling operation. This includes a series of actions such as recognizing the fueling port cover, opening the cover, and inserting the fuel gun. In my opinion this technology is necessary that for example, in the COVID-19 epidemic prompted an increase in the need to reduce human contact, refueling robots can effectively reduce the risk of exposure to infection; robotic refueling can reduce the volatilization of oil, reduce environmental pollution, in line with modern environmental standards; labor shortages, and other social issues.

1.2. Significance

Using a refueling robot in a bad environment, or in a remote area and being able to work 24/7 improves the operating time and efficiency of the gas station, thus increasing revenue. All that is required is regular checking by technicians. Secondly, automated systems can accurately control refueling volumes, reducing human error and fuel waste and increasing economic efficiency. In terms of employee health, this technology reduces the amount of time gas station employees are exposed to harmful gases and hazardous environments, improving job safety and health. In terms of technology: the research and development and application of refueling robots promote the development of robotics, machine vision, artificial intelligence and other related technologies, which involves a number of disciplines such as mechanical engineering, computer science, electronic engineering, etc., and promotes the integration and innovation of interdisciplinary technologies.

1.3. Current Status At Home and Abroad

Major domestic players such as SINOPEC, Jingdong and Hangzhou Yule are also actively promoting the development and application of refueling robots. Domestic research often combines its own technological advantages to launch refueling robot systems with a high degree of intelligence, and pilot applications have been launched in a number of cities. Currently, the domestic system mostly adopts technologies such as AI image recognition, robotic arm control and automatic settlement, with favorable feedback from users, showing high market potential. Sinopec is actively exploring the application of intelligent refueling systems, and many of its gas stations have been equipped with self-developed intelligent refueling machines. Hangzhou Uno Information Technology Co., Ltd. focuses on the research and development of intelligent equipment for gas stations, and its intelligent refueling robot has been launched in Zhejiang, Jiangsu and other places for pilot applications. The feedback from the pilot users is very positive, saying that the system has greatly improved the convenience and safety of refueling.

Overseas in the refueling robot technology has accumulated rich experience, a number of companies such as Fuelmatics, RoboPump, etc. in the technical realization and market promotion have made some progress. Based on machine vision, AI and automatic control technology, these systems have been able to realize accurate refueling operations and have been put into use on a large scale, helping to improve the service capacity and operational efficiency of gas stations. The Rotterdams Tank Terminal in the Netherlands has adopted a fully automated refueling robot system provided by TAE Vision Robotics, and Fuelmatics Systems, a U.S.-based company that focuses on the development of automated refueling systems, has developed robots that can complete the entire refueling process automatically.

1.4. Basic Introduction to Refueling Robots

The system aims to automate vehicle refueling operations using machine vision technology to reduce manual intervention and improve the efficiency and safety of the refueling process. The system consists of the following main components: cameras and sensors are used to capture images of the vehicle and detect the fuel tank position. The computer vision algorithm is used to recognize the position and status of the vehicle and the fuel tank cap. The control system safety system is used to perform refueling operations, including opening the fuel tank lid, inserting the fuel gun and refueling, and coordinating the work of each part to ensure that the refueling process runs smoothly coordinating the work of each part to ensure that the refueling process runs smoothly.

2. VISUAL ORIENTATION STUDIES

2.1. Overall Architecture and Program Design

In refueling robot applications, real-time positional localization is critical for autonomous robot navigation and refueling operations. Based on the overall framework provided previously, we can further refine and optimize it to fit the specific needs of refueling robots. The following is the improved framework for the visual localization system in the refueling robot paper. The first step is to acquire image data of the refueling station environment using a high-resolution camera and a depth camera. The addition of LiDAR can be considered to provide more accurate depth information. For the preprocessing of the data, the images are denoised, corrected and cropped to accommodate the changing light and other disturbing factors of the gas station environment. Next, for the feature extraction module, features in the images are extracted using pre-trained ResNet or MobileNet models, which are fine-tuned in conjunction with the task. Efficient feature description methods such as ORB are used to ensure real-time operation despite limited computational resources. For the target detection the detection algorithms are lightweight and efficient target detection algorithms such as YOLOv4 or SSD are used to detect the location and category of gasoline ports, parking spaces and obstacles. For the localization fusion module is utilizing Kalman filtering, fusing the data from the camera and LIDAR, and using Extended Kalman Filter (EKF) for state estimation to improve the stability and accuracy of localization. Finally, the output is performed to display the localization results in real time on the user interface, including the fueling interface and the parking space location information to help the operator monitor. As shown in the figure below.

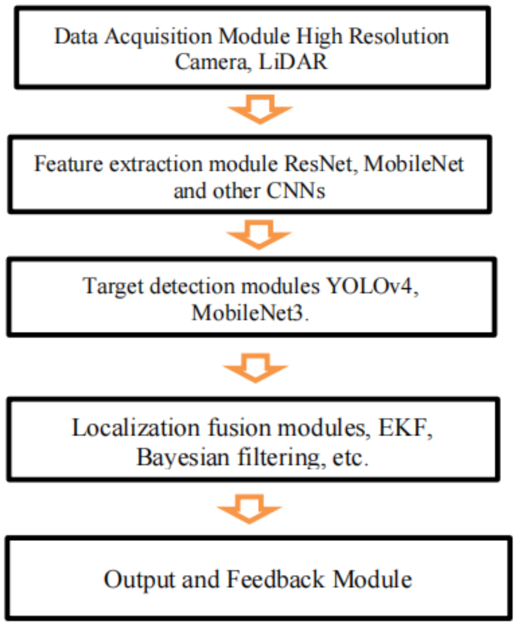

Figure 1. A mind map of tank positioning

2.2. Camera Selection

For camera selection CMOS (Complementary Metal Oxide Semiconductor) and CCD (Charge Coupled Device) are two common types of sensors used in industrial cameras. CCD sensors typically have higher image quality and lower noise for high-precision imaging, such as in scientific, medical, and professional photography, but are slower and consume more power. CMOS sensors, on the other hand, have faster readout speeds, lower power consumption, and lower manufacturing costs for applications requiring fast response and low power consumption, such as consumer electronics, automotive driving aids, and security monitoring. CMOS sensors, on the other hand, offer faster readout speeds, lower power consumption and lower manufacturing costs, making them suitable for applications requiring fast response and low power consumption, such as consumer electronics, automotive assisted driving and security surveillance. Although CMOS is slightly inferior to CCDs in terms of image quality and noise control, its technological advances have made it increasingly popular in many areas. For camera selection CMOS (Complementary Metal Oxide Semiconductor) and CCD (Charge-Coupled Device) are two common types of industrial camera sensors. CCD sensors typically have higher image quality and lower noise for high-precision imaging such as scientific, medical and professional photography, but are slower and consume more power. CMOS sensors, on the other hand, have faster readout speeds, lower power consumption, and lower manufacturing costs for consumer applications. CMOS sensors, on the other hand, offer faster readout speeds, lower power consumption and lower manufacturing costs, making them suitable for applications requiring fast response and low power consumption, such as consumer electronics, automotive assisted driving and security surveillance. Although CMOS is slightly inferior to CCDs in terms of image quality and noise control, its technological advances have made it increasingly popular in many areas.

The lens of an industrial camera is equivalent to a convex lens, and the imaging principle of the lens is shown in Figure.

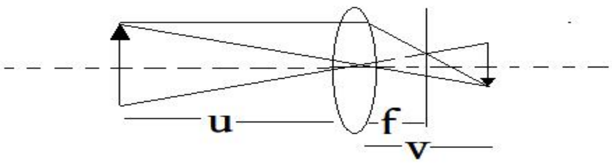

Figure 2. the imaging principle of the lens

2.3. Analysis of Algorithms

In refueling robots, the use of pre-trained ResNet or MobileNet models can significantly improve the performance and accuracy of visual localization systems. The following is a detailed explanation of these two models and their application in refueling robots, as well as some necessary formulas. ResNet (Residual Network) ResNet solves the problem of gradient vanishing in deep neural networks by introducing the Residual Block, which allows the network to extract features more deeply and accurately. The residual block formula used in it, the deep separable convolution formula.

$$DWConv(x)\ \backslash ast\ PWConv(x) \tag{1}$$

$$y = F(x, \{W_i\}) + x \tag{2}$$

In the refueling robot, both models are used for the following tasks: environment awareness to identify refueling interfaces, vehicle positions and obstacles in the refueling station. Accurately detecting and localizing refueling interfaces so that the robot can accurately perform refueling operations. Extract key features in the environment to support subsequent attitude estimation and path planning. Acquire image data of the refueling station environment using a camera and depth sensor. Data preprocessing denoises, corrects, and crops the images to accommodate lighting variations and other distractions at the refueling station.

The experimental data and results in the article [1] show that I found that the results of the study indicate that Convolutional Neural Network (CNN) models using YOLO subclasses are currently the best choice for solving the problem of image object detection in the eigenvalue extraction module of refueling robots. The analysis of existing YOLO subclass CNN models shows that it is necessary to replace some of the layers with different combinations of Inception ResNet modules in order to increase the computational speed and maintain the accuracy of image detection in refueling robots. By replacing multiple layers in the original tiny YOLO model with a sequence of two Inception ResNet modules, we developed a new CNN model, tiny-YOLO-Inception-ResNet2.

First, the relationship between the detection time of an image based on the number of computational units (CUs) in each layer of the CNN model was investigated. The results show that an increase in the number of such units in the CU improves the performance almost linearly. In addition, the effect on the accuracy of object detection in an image based on the floating point bit depth used in the CUs (16-bit or 32-bit) was investigated. It was found that 32-bit floating-point computation in the CU is required to maintain the accuracy of object detection in images in a refueling robot. It was also shown that the power consumption of the CU did not exceed 12 W in all the experiments. The results in the article [2] showed that increasing the number of Convolutional Neural Network (CNN) layers does not improve the detection accuracy. The researchers proposed a new architecture using ResNet modules in which the data from the layer inputs are also passed through two layers to the inputs of the next layer. This design allows the use of fewer CNN layers while improving the accuracy of object detection These results show that the application of this CU to the eigenvalue extraction module of a refueling robot has great promise for application.

In the feature extraction module, we can use the network model structure of YoLov4 and MobileNet by comparing the depth separable convolution and traditional convolutional neural networks found to be different in the number of parameters. The model combines the advantages of the YoLov4 network with the MobileNetv3 network to optimize its performance. [3] YoLov4 performs well in target detection tasks with high accuracy, can accurately recognize multiple targets, and can maintain stable detection performance under different lighting, viewing angle and background conditions. However, YoLov4's high accuracy and high real-time performance depend on high computational resources, which requires high hardware, complex model structure, and more time and resources for training and tuning.

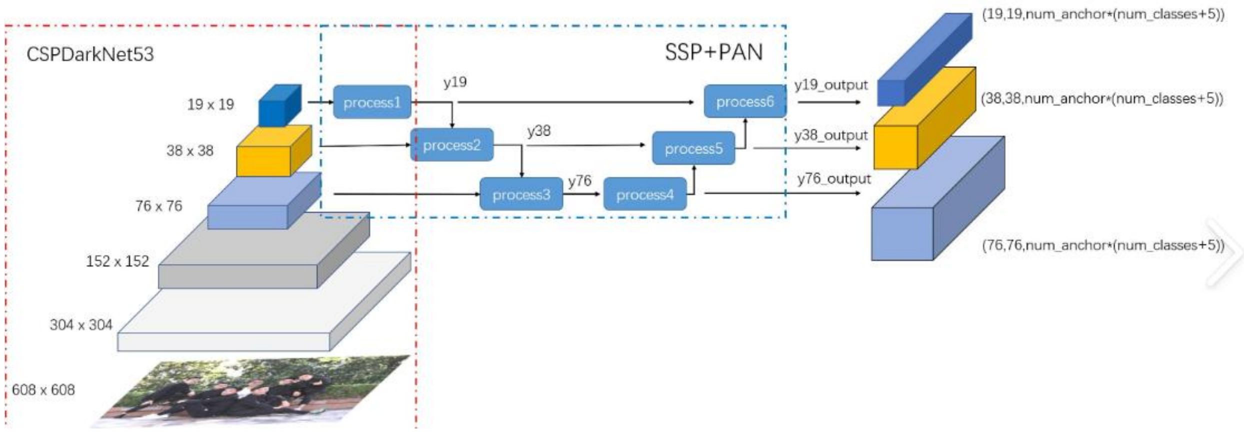

Figure 3. Analysis of the network structure of YoLov4

Total loss function:

$$textLoss = \mathrm{Loss}_{cls} + \mathrm{Loss}_{box} + \mathrm{Loss}_{obj} \tag{3}$$

Categorized losses:

$$textLoss_{cls} = \sum_{i=1}^{N} -[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)) \tag{4}$$

Localization loss:

$$textLoss_{box} = \sum_{i=1}^{N} \left((x_i - \widehat{x_i})^2 + (y_i - \widehat{y_i})^2 + (w_i - \widehat{w_i})^2 + \left(h_i - \widehat{h_i}\right)^2\right) \tag{5}$$

Loss of confidence:

$$textLoss_{obj} = \sum_{i=1}^{N} -[c_i \log(p_i) + (1 - c_i)\log(1 - p_i)] \quad (6)$$

MobileNet uses deep separable convolution to dramatically reduce the number of parameters and the amount of computation, which is ideal for resource-constrained devices, MobileNet runs fast on mobile devices and is suitable for real-time applications, and MobileNet is easy to deploy on embedded and mobile devices due to its lightweight modeling.

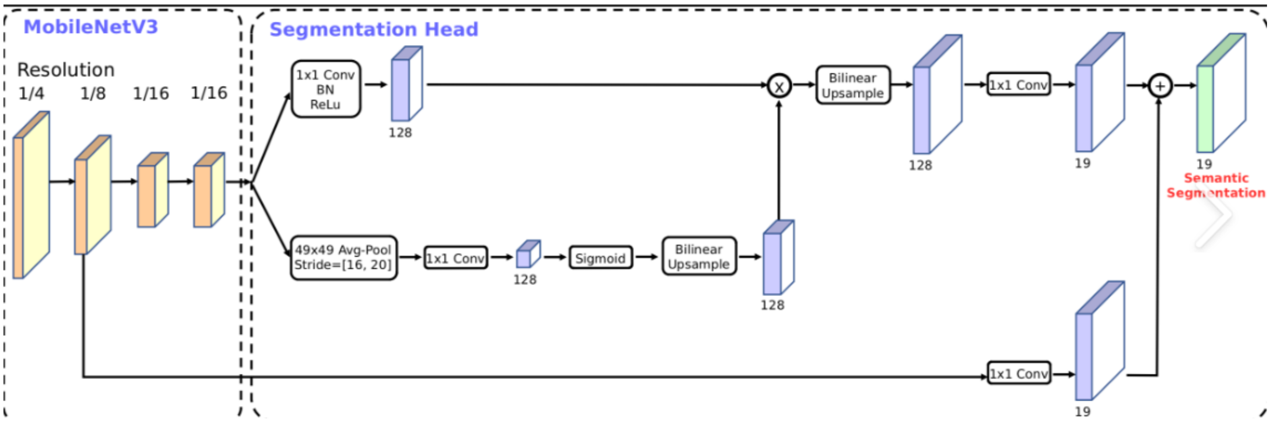

Figure 4. Analysis of the network structure of MobileNet

Depthwise Convolution:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m-1,j+n-1,k} \cdot K_{m,n,k} \quad (7)$$

Pointwise Convolution:

$$Z_{i,j,l} = \sum_{k} Y_{i,j,k} \cdot W_{k,l} \quad (8)$$

loss function:

$$Loss_{cls} = -\sum_{i=1}^{N} y_i \log(p_i) \quad (9)$$

The refueling robot has sufficient computational resources and has high requirements for detection accuracy and robustness, YoLov4 is a more suitable choice. MobileNet is a viable alternative in resource-constrained situations [4].

2.4. Positioning Fusion Module

In the localization module refueling robots need to be equipped with a variety of sensors, such as GPS, IMU (Inertial Measurement Unit), LIDAR and cameras. These sensors provide position information, velocity information, and environmental characteristics, respectively. EKF is able to effectively fuse the data from these sensors to improve the accuracy and reliability of localization. First of all, EKF is capable of fusing data from multiple sensors, such as GPS, IMU (Inertial Measurement Unit), LIDAR, and camera, etc. These sensors provide position, velocity, and environmental characteristics. These sensors provide position, velocity, and environmental feature information, respectively, and by fusing these data, EKF can more accurately estimate the robot's current position and orientation. Second, the EKF is designed to handle nonlinear systems [5].

Motion models and sensor observation models of refueling robots are usually nonlinear. EKF is able to deal with nonlinear problems efficiently by calculating and linearizing the Jacobi matrices of these nonlinear systems at each time step. In the dynamic prediction and update process, EKF consists of two main steps: prediction and update. In the prediction step, the motion model is utilized to predict the state at the next moment based on the current state and control inputs. In the updating step, the predicted state is corrected by combining the observed data from the sensors. Through continuous prediction and updating, EKF is able to track the position and attitude of the robot in real time. In addition, the EKF not only provides state estimation (e.g., robot position and attitude), but also estimates the state uncertainty. This is particularly important for the navigation and decision-making of refueling robots in dynamic and complex environments, and can help the system react more robustly when encountering uncertainty and noise. Through these mechanisms, EKF significantly improves the localization accuracy and robustness of the refueling robot. Even in the absence of sensor data or in the presence of interference, EKF provides reliable position information, ensuring that the refueling robot can accurately locate the tank opening and complete the automatic refueling operation.

In the convex optimization method for the solution process, Carmi et al [6] proposed a compressed perception algorithm for solving the sensor selection combination problem. In target tracking optimization, the standard extended Kalman filter obtains the suboptimal Kalman gain matrix by linearizing the nonlinear measurement model, aiming at minimizing the trace of the error covariance matrix. I found that the ADMM (Alternating Direction Method of Multipliers) method [7] was used to solve this optimization problem.

The nonlinear system equations are discretized and then Equation:

$$[x_{k+1} = x_k + v_k \cos(\theta_k)\, \Delta t] \tag{10}$$

$$[y_{k+1} = y_k + v_k \sin(\theta_k)\, \Delta t] \tag{11}$$

$$[\theta_{k+1} = \theta_k + \omega_k \Delta t] \tag{12}$$

2.5. Emulation Processing

Measurement model: $[z_k = h(x_k) + v_k]$

Defining the measurement model: (z_k)

Observation noise: (v_k)

In the EKF initialization we set the initial state to be $\boldsymbol{x_0}$, The initial covariance matrix p_0 is obtained by entering the correlation coefficients:

$$\hat{x}k|k-1 = f\hat{x}(k-1|k-1, u_{k-1}) \tag{13}$$

In handling sample data analysis and model fitting:

$$S_m = S_{\emptyset f}\sqrt{\frac{N}{N\sum_{i=1}^{k} fi^2 - \left(\sum_{i=1}^{k} f_i\right)^2}} \tag{14}$$

We utilize the Extended Kalman Filter (EKF) simulation for a refueling robot. The environment setup includes the robot's initial position, velocity, angular velocity, and noise settings. The code will be run in 20 seconds and a plot of the true trajectory and the EKF estimated trajectory will be generated, plotting the true trajectory (green solid line) and the EKF estimated trajectory (blue dashed line) to compare the difference between the two. The following is shown in the figure 5.

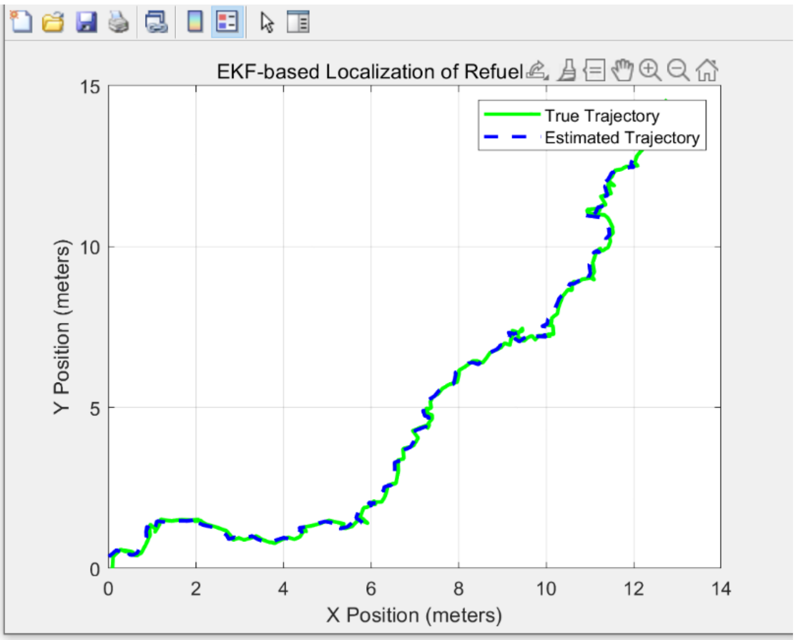

Figure 5. Comparison between actual and estimated trajectories

Then we can find that the extended Kalman filter performs well in trajectory estimation, with the real trajectory (green solid line) and estimated trajectory (blue dashed line) being generally close, especially providing high-precision position information in refueling robot positioning. Although there may be slight deviations at turning points or curve segments, the overall error is small, effectively reducing noise interference and making the estimated trajectory smooth and close to the real trajectory.

The Kalman filter is used for sensor fusion, and first data preprocessing is required to obtain data from cameras and LiDAR and perform preprocessing. Afterwards, describe the state vector of the system state. Define the system motion model and measurement model required for the Kalman filter using the equations of motion and observation, and calculate the Jacobian matrix. Many positioning robots use Kalman filtering [8].

I use MATLAB simulation to fuse data from LiDAR and camera to estimate the position of the refueling robot. We will use Extended Kalman Filter (EKF) to process these nonlinear models. Define the state vector as $x = [x, y, \theta]^T$, Set control input: $u = [v, \omega]^T$.

Using Matlab, set the step size (dt) to 0.1s and the number of simulation steps (N) to 50. We obtained the image.

The green line (True Path) shows the true trajectory of the robot's movement, which may take the form of a curve. The blue line (Estimated Path) shows the trajectory estimated by EKF, which should be close to the green line. The red dots (Lidar Measurements) and purple dots (Camera Measurements) are distributed around the true trajectory, but there may be some random errors between each point.

Through these graphical results, the performance of EKF can be intuitively evaluated. If the estimated trajectory is very close and smooth to the real trajectory, it indicates that EKF effectively integrates sensor data and reduces the influence of noise. If the estimated trajectory deviates significantly from the true trajectory, it is necessary to adjust the model parameters to improve the filtering effect. Kalman filtering has a wide range of applications in many fields. Bogdan [9] successfully achieved the localization of robots in office environments by using laser sensors and histogram matching, and employing particle filtering methods. Andrew et al. [10] successfully located the robots inside the building using wireless network sensors and particle filtering methods. By integrating LiDAR and camera data, the Kalman filter can significantly improve the positioning accuracy and stability of refueling robots in complex environments.

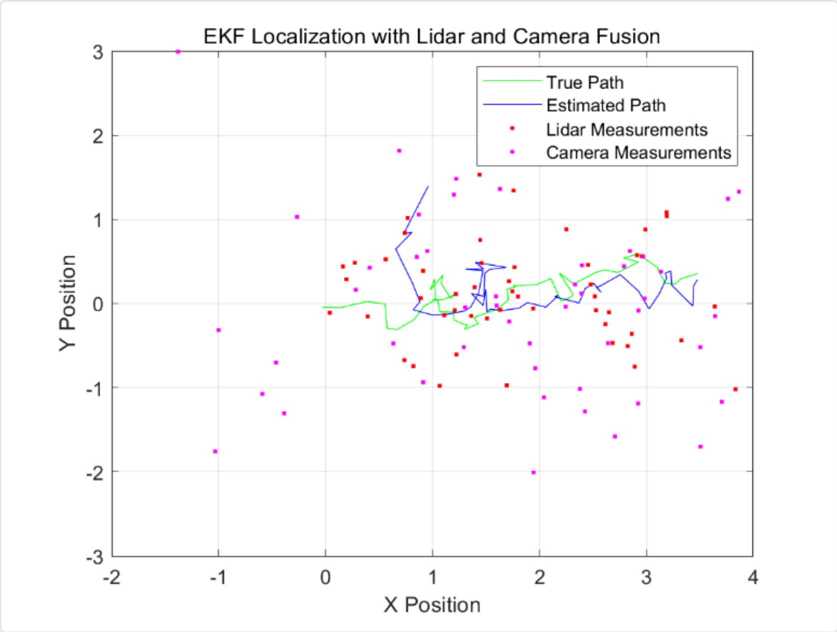

Figure 6. Trajectory comparison

3. CONCLUSION

3.1. Summary of the Paper

This article designs and implements a machine vision based refueling robot system, which integrates YoLov4 and MobileNet for feature extraction, achieving accurate recognition and localization of fuel tank positions. The experimental results of algorithm performance evaluation show that using Extended Kalman Filter (EKF) for trajectory estimation can effectively reduce noise interference and improve positioning accuracy. The comparison chart between the real trajectory and the estimated trajectory shows that the estimated trajectory is highly consistent with the real trajectory and has a small error.

3.2. Application Prospect

This system has good application prospects, especially in scenarios such as unmanned gas stations and self-service refueling systems, which can improve the automation and efficiency of the refueling process and reduce labor costs.

This system is not only suitable for fuel tank positioning, but can also be applied to other automated tasks that require high-precision positioning and navigation through improved algorithms and hardware, such as autonomous driving, warehousing and logistics robots, etc. It can not only refuel but also add other dangerous chemicals to ensure the safety of workers.

3.3. Future Research Directions

In the future, the feature extraction algorithms of YoLov4 and MobileNet can be further optimized to improve recognition accuracy and speed, while exploring other more advanced machine vision algorithms. Combine with other sensors (e.g. LIDAR, ultrasonic sensors, etc.) to realize multi-sensor data fusion and improve the stability and reliability of the system. Study how to further improve the real-time performance of the system under the premise of guaranteeing the recognition accuracy to meet the requirements of practical applications.

Finally, the output feedback module in the refueling robot is an important part to ensure the safety, accuracy and efficiency of the refueling process. The design of this module can be divided into two parts: hardware and software. The hardware part contains a variety of sensors and measuring devices. First, the flow sensor is used to monitor the refueling volume in real time to ensure the accuracy of the refueling amount; the pressure sensor is installed on the refueling pipeline to detect the pressure in the pipeline in real time to prevent leakage and over-pressure problems; and the temperature sensor is used to monitor the temperature of the fuel to avoid the safety hazards caused by the temperature. The component parts include data acquisition and processing, user interface, control logic and communication module. The Data Acquisition and Processing module is responsible for reading data from each sensor and performing the appropriate calibration and filtering to ensure the accuracy of the data.

REFERENCE

[1] Rega, G., Settimi, V., & Lenci, S. (2020b). Chaos in one-dimensional structural mechanics. Nonlinear Dynamics, 102(2), 785–834. https://doi.org/10.1007/s11071-020-05849-3Rega, G., Settimi, V., & Lenci, S. (2020b). Chaos in one-dimensional structural mechanics. Nonlinear Dynamics, 102(2), 785–834. https://doi.org/10.1007/s11071-020-05849-3Thermocouples: Figliola and Beasley, pp. 342-357.

[2] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Proceedings of the Thirty-First AAAI Conference on Artificial IntelligenceFebruary, pp. 4278–4284, 2017.

[3] Huang, L., Xiang, Z., Yun, J., Sun, Y., Liu, Y., Jiang, D., Ma, H., & Yu, H. (2023). Target detection based on Two-Stream Convolution neural network with Self-Powered Sensors information. IEEE Sensors Journal, 23(18), 20681–20690. https://doi.org/10.1109/jsen.2022.3220341

[4] D.T. Nguyen, T.N. Nguyen, H. Kim, H-J Lee. A High-Throughput and Power-Efficient FPGA Implementa-tion of YOLO CNN for Object Detection. IEEE Transactions on Very Large Scale Integration (VLSI) Systems; vol. 27, pp. 1861–1873, 2019.

[5] MASAZADE E, NIU R, VARSHNEY P K, et al. Energy aware iterative source localization for wireless sensor networks [J]. IEEE Transactions on Signal Processing, 2010, 58(9)4824-4835.

[6] CARMI A. Sensor scheduling via compressed sensing[C]// Proceedings of 13th international Conference on Information Fusion. Edinburgh: IEEE, 2010:1-8.

[7] BOYD S, PARIKH N, CHU E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers [J]. Found. Trends Mach. Learn, 2011,3(1):1-122.

[8] Ziar ko W. Variable Precisio n Roug h Set Mo del [J]. J of Comp uter Sy stems and Science, 1993, 46(1): 39-59.

[9] Bo gdan K. Finding Locat ion Using a Part icle Filter and Histo g ram M atching [A]. P roc of A rtif icial I ntelligence and Sof t Comp uting [C]. Poland: Spring er, 2004: 786-791.

[10] Andr ew H, Sa jid S. An Ex perimental Study of Lo calization Using Wir eless Ether net [A]. Pr oc of the I nt Conf on Field and Serv ice Robotics [C]. Lake Yamanaka, 2003: 201-206.