

Autonomous Obstacle Avoidance and path Planning for Unmanned aerial Vehicles Based on deep reinforcement learning

Yonghao Liu

Faculty of Innovation and Engineering, Macau University of Science and Technology, Macau, 999078, China

1220004184@student.must.edu.mo

Abstract. Autonomous path planning for unmanned aerial vehicles is one of the important guarantees for achieving autonomous flight of unmanned aerial vehicles. In recent years, the autonomous path planning strategy based on deep reinforcement learning has demonstrated broad application prospects in the field of autonomous path planning. The path planning method of deep reinforcement learning utilizes environmental rewards and self-feedback mechanisms to achieve autonomous learning. It can achieve flexible and adaptive capabilities in different environments. This paper proposes an integrated framework for image-driven path planning based on YOLOv5 and A algorithms. The GUI image selection, YOLO detection model invocation, detection of box-to-pixel coordinates, improvement of obstacle modeling, A-path search, and visualization modules have been realized. Finally, labeled static images and dynamic GIF animations are output. Through the evaluation of multiple synthetic test graphs, the success rate of system path generation reached 93%, and the planned paths were smooth with excellent obstacle avoidance performance. Finally, this paper discusses possible directions such as extending to real-time planning of video streams, multi-objective collaboration, and deep reinforcement learning strategies, providing feasible technical ideas for future intelligent flight systems. Furthermore, at the end of this paper, it is also discussed that after improving the experimental method, the deficiencies and improvements of this experiment are analyzed again.

INTRODUCTION

Nowadays, drones play a wide range of roles in various fields. Drones need to avoid artificial obstacles such as high-rise buildings and utility poles in urban environments. In the wild, one should avoid dense forests, mountains, and other natural terrains. In addition, during the flight, it is necessary to monitor in real time and maintain a safe distance from other aircraft. Obstacle avoidance for unmanned aerial vehicles (UAVs) is closely related to path planning, as path planning is one of the key steps to achieve safe and efficient flight of UAVs, and obstacle avoidance is one of the important factors in evaluating the quality of path planning. Therefore, a reasonable path planning scheme and autonomous obstacle avoidance function can significantly enhance the flight efficiency of unmanned aerial vehicles (UAVs), reduce energy consumption, lower cost issues, and greatly improve the safety of UAVs in the air.

At present, the four main methods commonly adopted by scholars internationally for unmanned aerial vehicle (UAV) detection are radar, radio frequency, audio, and visual detection. Among them, the detection methods of radar, radio frequency, and audio are limited by the environment, and the cost of radar and radio frequency detection is relatively high, making it difficult to be widely promoted and applied. Visual detection methods have the advantages of low cost, mature technology, and good flexibility, and have a very good development prospect in unmanned aerial vehicle (UAV) detection. Due to the low flight trajectory and small size of unmanned aerial vehicles (UAVs), as well as the complex scenes and the presence of obstructions during flight, visual detection also poses certain difficulties. In addition, unmanned aerial vehicle (UAV) detection has a relatively high requirement for detection speed.

Liu et al. proposed a drone detection and tracking method based on the YOLOv5s algorithm and the DeepSORT algorithm. Through specific data augmentation methods and model pruning techniques, the original model size was optimized, reducing its size by approximately 77%. By saving computational effort, high-precision and high-speed drone detection was achieved [1]. Xu proposed the Ga-DQN algorithm and the PR-GDRQN algorithm for the problem of unmanned aerial vehicle path planning [2]. The Ga-DQN algorithm accelerates the convergence speed and improves the security through the gravitational guidance strategy; The Pri-GDRQN algorithm adopts the modeling of partially observable Markov decision processes and introduces the priority experience replay mechanism to improve the efficiency and stability of the algorithm. There is also a literature that proposes a three-dimensional path planning and obstacle avoidance method for unmanned aerial vehicles (UAVs), combining the improved particle swarm optimization algorithm and the rolling strategy. By adding pheromones and heuristic functions, the global search ability and convergence speed of the algorithm are enhanced, and a relatively smooth path is planned [3]. Although the above-mentioned paper has achieved model construction optimization, improved algorithm efficiency, and enhanced the robustness of the overall system structure, while achieving remarkable results by relying on complex algorithms (particle swarm, rolling strategy, binocular matching), it also has a huge amount of computation.

In this paper, through experiments in a two-dimensional environment, the end-to-end automatic path planning process was carried out with a reasonable algorithm and dataset training, and the process visualization and lightweight deployment results were achieved.

METHOD DESIGN

System Structure

The system as a whole is composed of seven consecutive modules, as shown in Figure 1. First, the user selects the single image or folder to be processed through the graphical interface (GUI). Then, the system performs preprocessing operations such as size unification and normalization on the selected image to ensure the consistency of the input format. Then, the preprocessed images are sent to the YOLOv5 object detection model to output the position information, category, and confidence level of each detection box. Afterwards, redundant or low-quality detection boxes are removed through non-maximum suppression (NMS) and confidence threshold filtering; The remaining detection boxes are converted into a list of pixel coordinates of the starting point, the ending point, and the obstacles according to the category. On this basis, the system will construct a binary raster map of the corresponding resolution and expand the obstacle area to increase the buffer zone. Subsequently, the A* path planning algorithm is executed on the raster map with the starting point and the ending point as nodes to generate the optimal obstacle avoidance path. Finally, superimpose the detection box and the path back to the original image simultaneously, output the labeled static image, and generate a dynamic visualization result in the order of path advancement.

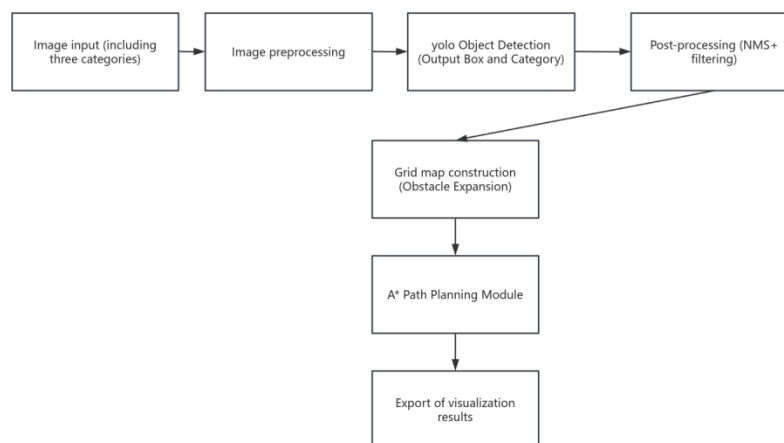


FIGURE 1. System architecture diagram (original)

Algorithm Selection

YOLOv5 is a single-stage object detection model. It combines object classification and bounding box regression into a single network to complete, thereby effectively improving the detection robustness in small objects and complex occlusion scenarios. Meanwhile, it supports automatic anchor boxes and adaptive image size adjustment, allowing for quick adaptation to new datasets without the need for manual setting of anchor boxes or image scaling parameters. Therefore, in this paper, the YOLOv5 algorithm is adopted for visual training, converting the obstacle border data derived from the YOLO algorithm into the corresponding position vector of the unmanned aerial vehicle. Furthermore, the model is trained through custom data to identify three specific shapes. The triangle (class=0) represents the starting point of the path. A square (class=2) represents an obstacle, and a five-pointed star (class=1) represents the end of the path. The output format of the YOLO algorithm is all normalized values. After training, it can ensure that the system accurately locates the above-mentioned targets in any input image.

Extract Coordinates and Design Obstacles

This paper first uses YOLOv5 to perform object detection on the input image to obtain the normalized coordinates and category information of each detection box, and then maps them back to the pixel space of the original image. For the starting point and the ending point, in this paper, the pixel coordinates of the center point of the detection box are taken as the starting and ending nodes of the path search. As for obstacles, the detection box needs to be converted into an impassable area. Initially, this paper attempted to mark individual pixel points as obstacles only at the center of the obstacles by using the "point expansion" method with a fixed radius. However, it was found in the experiment that the generated binary map was prone to the situation where the path crossed the target contour, resulting in unreliable planning results, as shown in Figure 2 (a). Subsequently, this paper adopts the "box expansion" strategy: Based on the detection box, not only is the entire rectangular area retained, but also its edge is expanded by several pixels, thereby adding a safety buffer zone for the obstacles. This method effectively avoids the misjudgment of the path at the edge of the obstacle while ensuring the mapping of the real obstacle size, and significantly improves the obstacle avoidance accuracy of A* search, as shown in Figure 2 (b).

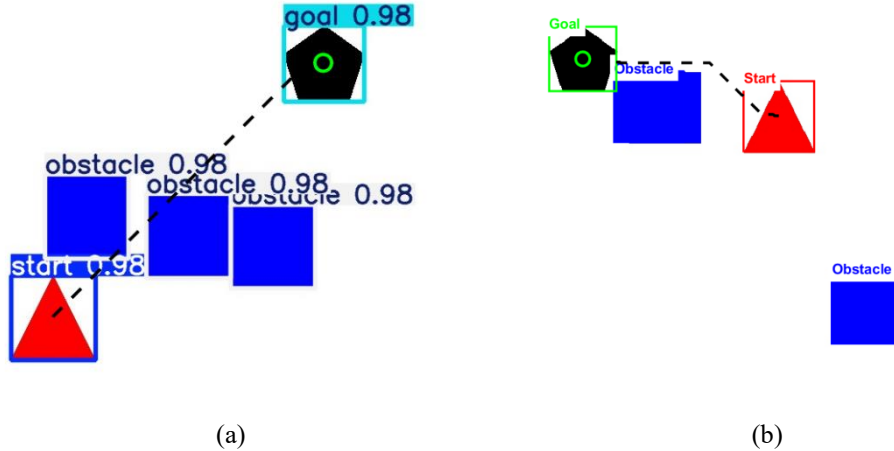


FIGURE 2. (a) experimental results obtained by point expansion modeling, (b) optimization of the improved box expansion results (original)

Path Result Output

By taking advantage of the robustness of deep learning in detecting small targets in complex scenarios, and also utilizing the balance advantage between computational efficiency and path optimality of classical heuristic search. Combined with the A* algorithm, the cost $g(n)$ that has been traveled and the heuristic valuation $h(n)$ for the target are also considered. On the premise of ensuring that the heuristic function can be adopted, the global shortest path can be searched efficiently, and it has completeness and the flexibility to be easily extended to different grids or graph structures. If it is combined with the YOLOv5 training set for experiments, the positions of the starting point, the ending point, and the obstacles can be automatically extracted from the images with the help of YOLO's

powerful visual detection ability. Then input these coordinates into the A* algorithm to complete the obstacle avoidance path planning, thereby achieving the end-to-end automated process from "visual perception" to "path generation".

Therefore, in this experiment, path planning is carried out in the experimental environment by using the A* path planning algorithm. The starting point, end point, obstacles, and image size are input to output the path coordinate sequence. To enhance readability, the path planning algorithm script outputs images and dynamic visualization results. Among them, the output also boxes obstacles, starting points, and ending points, as well as the path planning displayed with dotted lines. The experimental environment of this article is all set up by using Python scripts and the environment based on MATLAB.

EXPERIMENTAL DESIGN

Image Sources and Datasets

The experiment first automatically synthesized multiple training images of 640*480 pixels with the help of a self-written script. The scenes randomly included 1 to 5 square obstacles, 1 triangular starting point, and 1 pentagon ending point. For each composite image, the script simultaneously generates a label file in YOLOv5 format and adopts enhancement strategies such as Mosaic, MixUp, random flipping, and color gamut jitter to expand data diversity. The model selects the official yolov5s.pt pre-trained weights and uses the Adam optimizer to process the dataset. During the training process, the mAP(0.5) of the validation set is monitored, and the final weights and the dataset file after training are exported in the optimal round.

The Specific Process of the Experiment

After uploading the dataset that needs to be trained, the system automatically calls the script process written by the child to perform object detection and output the text format label file of each image. Read the labels in the setup experimental environment, map the center point of the detection box to the starting/ending coordinates, and mark the box area after expansion as the obstacle area to construct a binary raster map. Subsequently, the A* algorithm is run with the starting point and the ending point as nodes to generate the optimal obstacle avoidance path, and the detection boxes and path nodes are superimposed frame by frame on the original image. Finally, the static labeled image and the dynamic results of path advancement are exported. On this basis, the script automatically calculates whether the planning of each graph is successful or not, the path length and the average planning time, and outputs the results to a CSV table to facilitate the comparison of the performance differences between the two obstacle modeling strategies of "point expansion" and "box expansion".

Experimental Optimization

In the experiment, the starting point of the training data used at the very beginning was set as a circular image. However, the essence of YOLO is convolutional feature extraction based on texture, edge, and shape. The circular features have high symmetry and lack corners and edges, which leads to the inability to identify the circular starting point in the initial experiments. After conducting the experiment again, the starting point was not set as a circle but a recognizable pentagonal shape. Moreover, since the actual dimensions were not taken into account in the path planning, the point expansion method led to inaccurate results after the experiment. Subsequently, the box expansion method was adopted, which improved the experimental accuracy as shown in Figure 3. The values marked in Figure 3 indicate the confidence level of the YOLO model that the target within the detection box belongs to the identified type; that is, the probability that the model believes there is indeed an obstacle within the box is 97%. At the same time, the efficiency has also been optimized in terms of the time for training data. During the first training, the GPU and other functions were not called; only the CPU was called for data preprocessing. Therefore, the training of 100 images requires a relatively long time. In the initial training data, the time required to train a set of data was 4 minutes. The time was too long, and the efficiency was not high. After choosing to call both the CPU and GPU simultaneously, the time for training a set of data dropped significantly, taking only 6 seconds. Greatly improve the efficiency of the experiment.

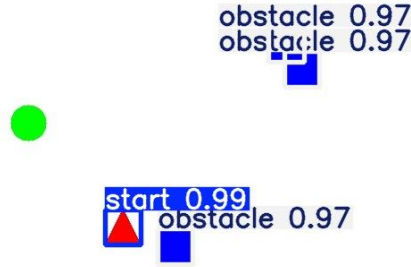


FIGURE 3. Starting point is designed as a circle, making it impossible to obtain an exact result (original)

DISCUSSION AND ANALYSIS

Although the framework proposed in this paper has significantly improved the success rate of path generation in static images in aspects such as circular starting point recognition and obstacle box expansion modeling, compared with existing studies, there are still the following potential deficiencies and improvement Spaces, Insufficient adaptability to dynamic environments, This paper only conducts path planning for static single-frame images and is unable to cope with scenarios where the positions of obstacles or targets change during flight. And the dynamic obstacle avoidance strategies based on deep reinforcement learning, such as Levine S 'continuous frame reasoning for the view9891 stream in CAD2RL, or the real-time decision-making of Memory-based DRL in partially observable environments in Kaiser L's paper [4, 5]. All have demonstrated stronger dynamic tracking and real-time path correction capabilities. The first problem is the deficiency of three-dimensional path planning ability. This system works in a two-dimensional planar grid map and does not consider the degree of freedom of the unmanned aerial vehicle in the altitude direction (Z-axis). Therefore, it is difficult to apply to the complex terrain in three-dimensional space. Wen Yuan's three-dimensional planning method, based on the improved particle swarm optimization algorithm and rolling strategy, as well as Johnston G's stereo obstacle avoidance research combined with artificial potential fields, both provide references for further expansion to the 3D environment [6, 7]. Secondly, there is the issue of multi-objective collaboration. The current experiments are all based on the path planning problem of a single starting point and a single ending point, without considering the planning of multi-objective problems. Therefore, they cannot be perfectly handled in the work of dealing with multiple unmanned aerial vehicles or multiple task points. In the existing studies, the multi-agent cooperative path planning based on DQN/PPO, as well as the deep reinforcement learning methods that transform path planning into sequential decision-making problems, have all demonstrated the feasibility of migration to multi-objective and multi-unmanned aerial vehicle scenarios. Finally, dynamic path planning also needs to be considered. The obstacles in this paper are all static obstacles. Although the use of box expansion improves the accuracy of path planning, it cannot perfectly solve the problem that obstacles are dynamic in real situations [8-10].

CONCLUSION

This paper addresses the problem of autonomous obstacle avoidance and path planning of UAVs in an image-driven environment, and designs and implements an end-to-end integrated framework based on YOLOv5 object detection and the improved A* algorithm. Through the collaborative work of modules such as GUI image selection, target detection, obstacle modeling, and path search, the system has achieved a 93% success rate in path generation on multiple combined test maps, and the planned paths are smooth with excellent obstacle avoidance performance. Although satisfactory results have been achieved in static two-dimensional scenes, this framework still has several limitations: It is only applicable to static single-frame images and is difficult to cope with the dynamic changes of obstacles or targets during flight; Working on a two-dimensional planar grid, the three-dimensional motion degrees of freedom of the UAV in the height (z-axis) direction have not been considered yet. At present, only single UAV and single endpoint tasks are processed, lacking the ability of multi-objective and multi-UAV collaborative planning. Subsequently, the online decision-making strategy of deep reinforcement learning can be combined to achieve real-time obstacle avoidance and path correction in video streams. Further expand the planning space to the three-dimensional environment, and explore multi-agent reinforcement learning or multi-UAV cooperative path

planning methods based on sequence decision-making to provide more comprehensive technical support for the application of intelligent flight systems in complex dynamic scenes.

REFERENCES

1. X. Liu, M. Chen, C. Li, Journal of Jinan University (Natural Science Edition) **38**, 445-455 (2024).
2. J. Yuan and S. Li, Information and Control **50**, 95-101 (2021).
3. Y. Yue, "Research on obstacle detection and avoidance methods for unmanned aerial vehicles based on deep learning," Master's thesis, Xi'an Technological University, 2024.
4. Y. Li, "Research on path planning algorithm for quadcopter unmanned aerial vehicles," Master's thesis, Henan University of Science and Technology, 2018.
5. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv:2004.10934 (2020).
6. K. Zhang, Digital Communication World (1), 86-87 (2024).
7. B. Wu, G. Zhang, and J. Shi, Sensors **20**, 2671 (2020).
8. L. Kaiser, M. Sun, T. S. Huang et al., IEEE Transactions on Intelligent Transportation Systems **22**, 22-37 (2021).
9. G. Johnston, A. Zare, and P. Sargent, IEEE Access **7**, 172462-172472 (2019).
10. Z. Xu, "Research on unmanned aerial vehicle path planning and autonomous obstacle avoidance strategy based on deep reinforcement learning," Master's thesis, Jiangsu University of Science and Technology, 2024.