Fire Recognition and Path Planning for Fire Fighting Robots Based on Machine Vision Slam Technology

Lequan Li

School of Automation, Nanjing Institute of Technology, Nanjing ,211167, China
x00203211215@njit.edu.cn

Abstract. Firefighting robots, as an important part of firefighting, make great contributions to the life safety and economic security of the society. Among the firefighting robots, the firefighting robots utilizing machine vision slam technology are the most promising. This paper describes the main algorithms of today's firefighting robots based on machine vision slam technology. The fire recognition technology mainly utilizes the You Only Look Once(YOLO) algorithm, with YOLOv5 and YOLOv8 as the main algorithms. The main algorithms for path planning techniques are A-star(A*) algorithm and Ant Colony Optimization respectively. For these main algorithms, the article evaluates their advantages and disadvantages, and proposes directions that can be improved. Then, the article introduces two feasible solutions based on transformer model for visual domain and visual Simultaneous Localization and Mapping(SLAM) domain respectively, which have been verified with good results in other domains. Finally, this paper is summarized to evaluate the algorithms and algorithm optimization that have been studied.

Keywords: Machine vision，Machine learning，SLAM, Fire robot.

1 Introduction

With the improvement of science and technology as well as economic development, intelligence enters people's lives. Along with the concepts of Internet of Things, smart city, intelligent firefighting robots also become a safety core of smart city and Internet of Everything [1]. Firefighting robot is a kind of intelligent robot that can recognize fire and extinguish fire autonomously, which plays a great role in firefighting and rescue. The use of machine vision SLAM technology can help firefighting robots to realize the recognition of fires in unknown environments, as well as planning the best path for search and rescue tasks.

The main algorithms of today s firefighting robots based on machine vision slam technology. The fire recognition technology mainly utilizes the YOLO algorithm, with YOLOv5 and YOLOv8 as the main algorithms. The main algorithms for path planning techniques are A* algorithm and Ant Colony Optimization respectively. In addition to the above major algorithms, Benjamin Ronald van Manen proposed a new thermally synchronized localization and mapping (SLAM) algorithm called

© The Author(s) 2024
Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_13

FirebotSLAM [2]. Among the various types of algorithms applied to fire recognition are visual algorithms such as Convolutional Neural Network(CNN) and Efficient Detection. Christyan Cruz Ulloa compared the models of deep learning such as Region-based Convolutional Neural Network(R-CNN),Single Shot MultiBox Detector(SSD),YOLO,EfficientDet and found that the mean Average Precision(MAP) value of YOLO series of algorithms is significantly greater than the other deep learning models [3][4]. This is one of the reasons why YOLO series algorithms have become the main algorithms applied to fire recognition.

This article introduces the optimization and improvement of each algorithm and evaluates the advantages and disadvantages of their improvement. Then this paper introduce two feasible solutions based on transformer model for vision and slam domains respectively, which have been verified in other domains with good results. Finally, this article is summarized.

2 Optimization Improvement of the Algorithm

2.1 A Subsection Sample

The K-means algorithm is difficult to give the number of clustering centers K and initial centers in advance, and Qing optimized the selection of initial points for this problem [5].

Ziyang Zhang proposed a lightweight ship fire detection algorithm based on YOLOv8n called "Ship-Fire Net", which can reduce the computational complexity by introducing GhostnetV2-C2F and spatial channel reconstruction convolution (SCConv) [6].Saydirasulov Norkobil Saydirasulovich uses YOLOv8, introduces WIoUv3 as a bounding box regression loss, improves the model's localization accuracy through a reasonable gradient allocation strategy, replaces the traditional convolution process in the middle neck layer, reduces the model parameters, accelerates the convergence speed, and introduces the BiFormer Attention Mechanism to focus on the features of forest fire smoke, while suppressing the interference of background information [7].

In fire recognition, the YOLO family serves as the most commonly used visual model. The authors above have made improvements to the YOLO series, but it is still the framework of the YOLO series. the YOLO series, although its adaptability, real-time performance, and accuracy compared to other visual models. However, the YOLO series is not accurate in detecting smaller targets and is prone to misdetection and omission. It can also cause overlapping detection or missed detection for dense targets. Even the newer YOLOv8 is unsatisfactory in terms of small target detection, although accuracy has improved, it has increased the algorithm and sacrificed speed.

2.2 Ant Colony Optimization

For the complex fire environment, Yang Song established a raster map combining two-dimensional map and terrain, increased the ants to expand the direction, and combined with the robot's own factors to optimize the heuristic function, to improve the efficiency of its firefighting robot [8].

Li Erchao optimized the effective vertices of the obstacles based on two-way Ant Colony Optimization, divided into two batches of ants set off searching at the target point and the starting point respectively to determine whether there are the same vertices or not, and repeatedly conducted the search to finally form the path [9].

Qiao Jia optimized the formula of pheromone increment in the Ant Colony Optimization to improve the concentration of the optimal solution and optimized the formula:

$$\Delta \tau_{\mathrm{ijk}} = \frac{\lambda \times (N - M) + K}{\min(\{length(g)\})} \quad (1)$$

The optimized formula can be used in non-two dimensional environments with uneven terrain can efficiently perform fire rescue and avoid higher fire areas [10].

Ant Colony Optimization is an optimization algorithm that imitates the behavior of ants searching for food in nature, with strong global search ability, can find the global optimal solution during the search process, and has good robustness and adaptability. However, the Ant Colony Optimization is a global search through a large number of calculations, which is computationally intensive, slow in the search process, and requires many iterations to obtain a better solution. It is easy to be influenced by the initial solution and the algorithm falls into the optimal solution. In practice, Ant Colony Optimization(ACO) can be used in combination with other optimization algorithms to improve its accuracy and performance.

2.3 A* Algorithm

Sen Li used global A* algorithm with local D-star(D*) algorithm to plan the path, through the global A* algorithm, global screening of static data, with the D* algorithm to deal with unexpected events [11].

The traditional A* algorithm is good at searching the static environment, but the fire environment is complex and variable.DU Chuan-sheng improves on the traditional A* algorithm for its problems, using two-way search to improve efficiency, optimizing the heuristic function, adding dynamic weights w(n) to h(n),adding the normalizing factor function to improve the adaptability of complex environments[12].

In order to improve the efficiency of the firefighting robot, Wang Wanfu improved the original valuation function on top of the original valuation function. The improved formula is:

$$f(n) = \mathrm{g(n)} + q_g(n) + s_h(n) h^*(n) \quad (2)$$

In this improved formula, qg(n) is the penalty function,which represents the cost of turning from the starting point to the current node n, and is proportional to the number of turns; sh(n) is the reward factor $0 \leq sh(n) \leq 1$. Its optimized A* algorithm can combine with real-time sensing data to dynamically adjust the path planning, so that firefighting robots can reach the target location quickly and safely. It can also perform multiple tasks simultaneously, such as searching for survivors, firefighting and surveying [13].

The A* algorithm, as a heuristic algorithm, is able to find the shortest path from the start state to the goal state.The A* algorithm can efficiently prioritize the search for nodes that are close to the goal state and reduce the search time, but its algorithm has obvious drawbacks. Since the valuation function of the A* algorithm needs to estimate the goal state, if the valuation function is not accurate enough, it will cause the algorithm to search for subpaths that are not the best path. Moreover, the A* algorithm is only suitable for offline search of the global goal state, for real-time dynamic environments, such as dynamic path planning for robots, then the A* algorithm is not applicable. So, for fire scenarios with complex and changing environments, the A* algorithm needs to be used in combination with other optimization algorithms and dynamic algorithms.

3 Feasible Solution Based on Transformer Model

Transformer model is a deep learning model for processing sequence data, which is widely used in language processing and machine vision. Two feasible schemes are given in this article, which are in the field of machine vision and in the field of Visual SLAM(VSLAM) algorithm.

3.1 Machine Vision Direction

Cong proposed a new module that combines transformer and convolution module called Unified Weighting Transformer Network(UWT-Net).This module is intended to be used to solve the problem of color distortion and contrast degradation in images captured by robots[14].

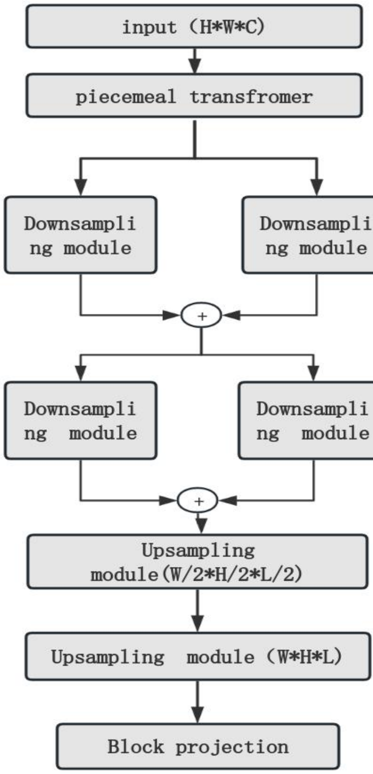

Fig. 1. The network uses an encoder and decoder architecture[14].

As shown in fig. 1 able to train end-to-end. The encoder is responsible for learning the feature representation and the decoder is responsible for the reconstruction of the image pixels.

Tables 1 and 2 show a comparison table of results from two large publicly available underwater image datasets, the Underwater Image Enhancement Benchmark Dataset (UIEB) and Enhanced Underwater Visual Perception (EUVP-US). The evaluated algorithms include Transmission estimation in underwater single images(UDCP), Image-Based Lighting Algorithm(IBLA), Underwater Light Attenuation Prior(ULAP), green blue enhancement (GB), Underwater Residual Neural Network(UResNet), Underwater Network(UWNet), and WaterNet.In Table 1, UWT-Net has the highest structural similarity index (SSIM) and underwater image quality metric (UIQM), and the second highest peak signal-to-noise ratio (PSNR). The results in Table 2 show that UWT-Net has the highest PSNR and SSIM values of 28.343 and 0.871, respectively, and the UIQM is also high. Overall, UWT-Net has the best experimental results.[14].

Table 1. Quantitative experimental results on the UIEB dataset [14].

Index	UDCP	IBLA	ULAP	GB	UResNet	UWNet	WaterNet	UWT-Net
PSNR	12.368	17.299	17.601	15.078	18.684	18.913	22.694	21.163
SSIM	0.526	0.678	0.692	0.608	0.754	0.784	0.854	0.870
UIQM	1.842	1.783	1.878	2.272	2.772	2.746	2.883	3.084

Table 2. Quantitative experimental results of the EUVP-US dataset [14].

Index	UDCP	IBLA	ULAP	GB	UResNet	UWNet	WaterNet	UWT-Net
PSNR	14.994	22.138	21.897	16.553	27.638	27.727	23.722	28.343
SSIM	0.570	0.721	0.761	0.606	0.839	0.832	0.818	0.871
UIQM	2.114	2.214	2.369	2.583	2.867	2.888	2.889	2.806

3.2 VSLAM Directions

Alexey Dosovitskiy proposed VSLAM algorithm based on Transformer architecture for deep learning model Vision Transformer(ViT). Fig.2 shows the working principle of ViT [15].

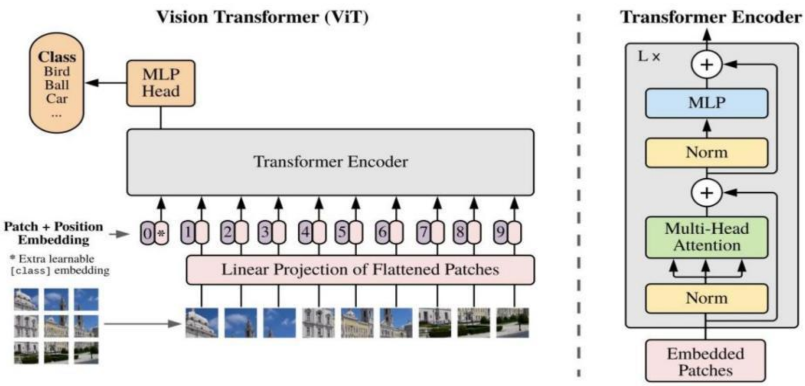

Fig. 2. Demonstrates the direct application of pure transformer to a sequence of image blocks without relying on CNN [15].

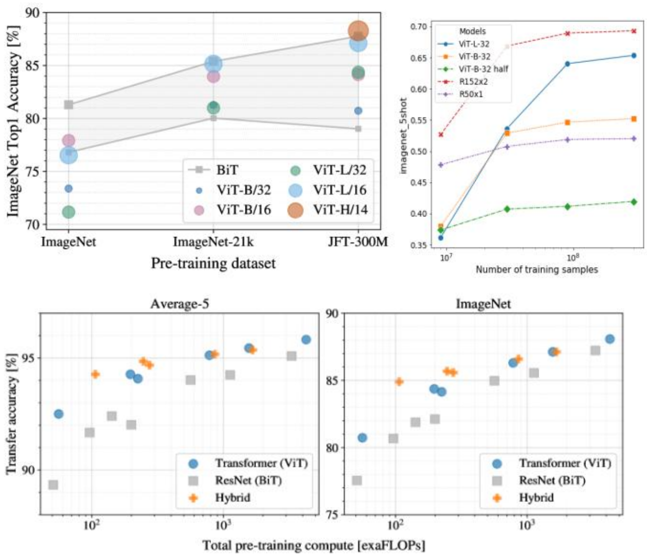

Fig. 3. Performance and pre-training computation for pre-training and different architectures [15].

Alexey Dosovitskiy explores the performance of the Vision Transformer model on different dataset sizes and its advantages over traditional convolutional neural networks. When trained on large-scale datasets, the ViT model outperforms the ResNet model with similar computational cost. For example, ViT outperforms on the JFT-300M dataset than on ImageNet. On smaller datasets, ViT may not perform as well as ResNet, but as the size of the dataset increases, ViT gradually outperforms ResNet.ViT outperforms ResNet computationally in large-scale image recognition tasks, especially in the pre-training phase. These findings suggest that training ViT on large-scale datasets allows it to learn enough patterns to achieve good performance in vision tasks[15].

4 Conclusion

There is still a lot of room for development of today's vision SLAM-based firefighting robots. The mainstream model YOLO algorithm for fire recognition is not accurate enough for small target detection and dense target detection. In the event of a fire, firefighting robots are prone to misjudgment resulting in missing the best time for firefighting and rescue. The A* algorithm and ant colony applied to path planning for firefighting robots also have huge room for improvement.Both A* algorithm and Ant Colony Optimization can be combined with other optimization algorithms to improve their efficiency and accuracy.The A* algorithm can also be combined with other dynamic algorithms to overcome the disadvantage of retrieving only static environments. This article presents two optimization schemes based on the transformer model. These schemes have demonstrated promising results in other fields. After that, the research will continue to follow the development of firefighting robots in fire recognition and path planning. Future research on firefighting robots with visual SLAM technology can incorporate more advanced algorithms or be equipped with more advanced artificial intelligence to facilitate the development of firefighting.

References

1. Talaat, F M. ZainEldin, H.: An improved fire detection approach based on YOLO-v8 for smart cities. Neural Computing and Applications **35**(28), 20939-20954 (2023)
2. Manen, B.R. Sluiter, V. Mersha, A.Y.: FirebotSLAM: Thermal SLAM to Increase Situational Awareness in Smoke-Filled Environments. Sensors **23**, 7611 (2023)
3. Cruz, U, C. Prieto, Sánchez, G. Barrientos, A.: Autonomous thermal vision robotic system for victims recognition in search and rescue missions. Sensors **21**(21), 7346 (2021)
4. Ann, H., Koo, K, Y.: Deep learning based fire risk detection on construction sites. Sensors **23**(22), 9095 (2023)
5. An, Q., Chen, X., Zhang, J.: A robust fire detection model via convolution neural networks for intelligent robot vision sensing. Sensors **22**(8), 2929 (2022)
6. Zhang, Z., Tan, L.: Tiong, R.L.K. Ship-Fire Net: an Improved YOLOv8 Algorithm for Ship Fire Detection. Sensors **24**, 727 (2024)
7. Saydirasulovich, S, N.: Mukhiddinov, M.; Djuraev, O.; Abdusalomov, A.; Cho, Y.-I. An Improved Wildfire Smoke Detection Based on YOLOv8 and UAV Images. Sensors **23**, 8374 (2023).
8. Yang, S.,Hong, T.,Zhu, L.: Improved ant colony algorithm for forest fire mobile robot path planning. Forest Engineering **40**(01), 152-159 (2024)
9. Li, E., Q, Q.: Improved two-way ant colony algorithm for mobile robot path planning. Computer Engineering and Applications **57**(18), 281-288 (2021)
10. Qiao, J.,Shi, Y.,Wang, C.: Three-dimensional path planning for forest fire rescue robot based on improved ant colony algorithm. Fire Science and Technology **42**(07), 989-994 (2023)
11. Li, S.: A fire reconnaissance robot based on SLAM position, thermal imaging technologies, and AR display. Sensors **19**(22), 5036 (2019)
12. Du, C.: Path Planning of Fire Robot Based on Improved Bidirectional A* Algorithm. Computer and Modernization **6**(04), 15-19 (2023)
13. Wang. W.: Forest and grass fire rescue path planning based on improved A* and inner spiral algorithms. Fire Science and Technology **41**(08), 1138-1142 (2022)
14. Cong, X.: Multi-loss fusion underwater image enhancement network based on visual transformer. Journal of Intelligent Science and Technology **4**(04), 522-532 (2022)
15. Dosovitskiy, A., Beyer, L., Kolesnikov, A.: An image is worth 16x16 words: Transformers for image recognition at scale (2020)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.