ISSN: 2664-9640

Mission Assignment of Unknown Threat Multi-UAV Based on MH-MADQN

Zejun Gao, Fei Cao, Hangfei Wang, Yangyang Sun

Rocket Army Engineering University, Xi'an, Shaanxi Province, China

Abstract

Before the drone performs its mission, all known situations in the battlefield environment, that is, the target position and a series of threat areas such as radar and obstacles in the battlefield have been detected in advance, and will no longer occur during the mission of the drone cluster Change. In this case, the multi-UAV task allocation problem model is relatively simple, and it is also the most mature research currently. It is only necessary to use the improved Q-learning algorithm to plan a reasonable route for the multi-UAV before the task is executed. Assign reasonable target and other mission information to multiple drones. However, considering the real battlefield environment, usually before the mission starts, only the location information of the target is known, and it is impossible to fully detect or only detect threat information such as radar and obstacles in a part of the combat area. At this time, it is necessary for the drone to perform detection while performing the mission, and re-planning and task assignment during the mission.

Keywords

Multi—uav, qlearning , task allocation.

1. Introduction

The relevant settings of multi-UAV task allocation in the unknown threat environment are based on Chapter 2. The difference is that due to the introduction of the unknown threat environment, a two-dimensional raster map is used for modeling as shown in Figure 1. Traditional multi-UAV task assignment and path planning need to transform the model into a search optimization problem through constraint conditions, and the model needs to be retrained when the given threat area changes. In this chapter, UAVs use reinforcement learning algorithms for self-learning. Even if the location of the threat area is unknown or changed, the model trained through reinforcement learning does not need to be retrained.

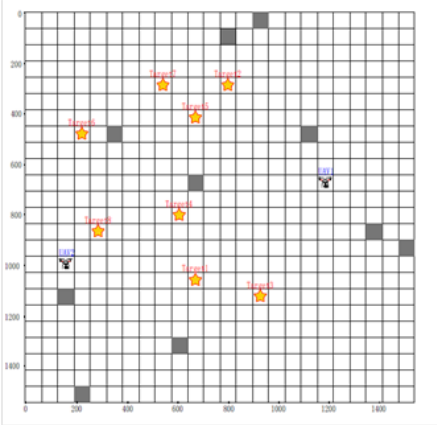

Figure 1 Two-dimensional raster model

ISSN: 2664-9640

2. Mission Assignment Method of Unknown Threat Multi-UAV Based on MH-MADQN

In machine learning methods, deep neural networks can perceive complex external environments, and reinforcement learning is very suitable for complex decision-making problems. Therefore, Mnih et al. [] combined the two and achieved good results in solving perceptual decision-making problems in complex environments. . This paper aims at the problem of multi-UAV task assignment and path planning for unknown threat areas, based on the perception ability of deep learning in the DQN model, combined with the decision-making ability of reinforcement learning, in simple terms, is through the deep neural network framework, the reinforcement learning algorithm Fitting the Q value in, can enable the UAV to perceive and avoid the threat area autonomously and improve the survivability of the battlefield.

2.1. Multi-UAV Task Assignment Environment Model in Unknown Threat Area

The environment model consists of three parts: a global map (3*24*24), a local perception map (3*7*7), and a unit direction vector (1*4) from the current position to the target position, as shown in Figure 2.

(1) Global map (Gm): It contains the location of obstacles, the current location of the drone, the current location of other drones, the current target location of the drone, and the target location of other drones.

(2) Local map (Lm): The definition is similar to the global map, but it limits the drone's perception range. With the current position of the drone as the center, 3 grids of information can be sensed up and down, left and right, and the size of the final input (Dimension) is 3*7*7.

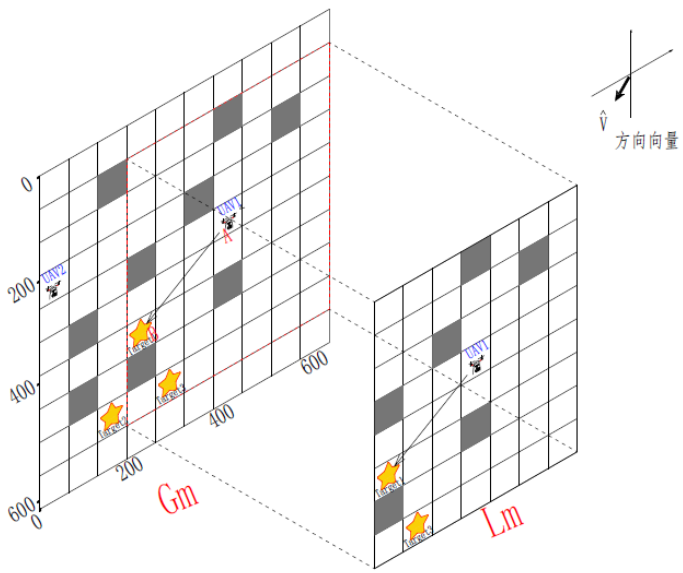

Figure 2 Environmental simulation map

2.2. MH-MADQN network structure design

Convolutional Neural Network (CNN) has become a mainstream method in the field of computer vision. Generally speaking, the deeper the network, the higher the nonlinearity and accuracy. However, the gradient of network training is backward propagation. Therefore, the larger the number of network layers, the smaller the gradient received by the previous layer of

ISSN: 2664-9640

the network or even the training stops. In 2015, He[4,5,6] proposed a residual network (ResNet), which can make information transfer between the deep and shallow layers. In 2017, Huang borrowed the idea of residual network and established DenseNet to improve the back propagation of the gradient, making the network easier to train. As shown in Figure 3, this chapter uses an L-layer convolutional network (L=4), each layer of the network has a non-linear transformation, and the output of the layer is. Then the output of the first layer is

$$x_l = H([x_0, x_1, ..., x_{l-1}])$$

Among $[x_0, x_1, ..., x_{l-1}]$ them is the output feature map generated by layer dense blocks. This connection can make the model more compact, and the information can be transmitted to the deeper network, enhance the features between the connected layers, and prevent the gradient from disappearing.

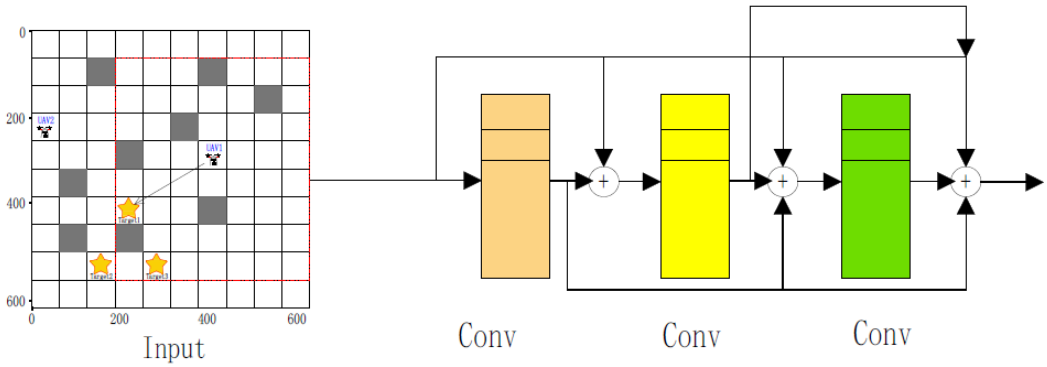

Figure 3 Dense Network structure

Figure 4 shows the MH-MADQN algorithm network structure. The network structure includes a Densenet network that perceives a global map. The network parameter settings are shown in Table 3.1. The network is composed of a preprocessing layer, a dense module, and a fully connected layer. , Its input is the global map (3*24*24). The first layer is the convolutional layer. The ReLU function is used as the activation function. The size of the convolution kernel is 3*3, the step size is 1, and more features are extracted. The output is, corresponding to the 9 action values of the drone.

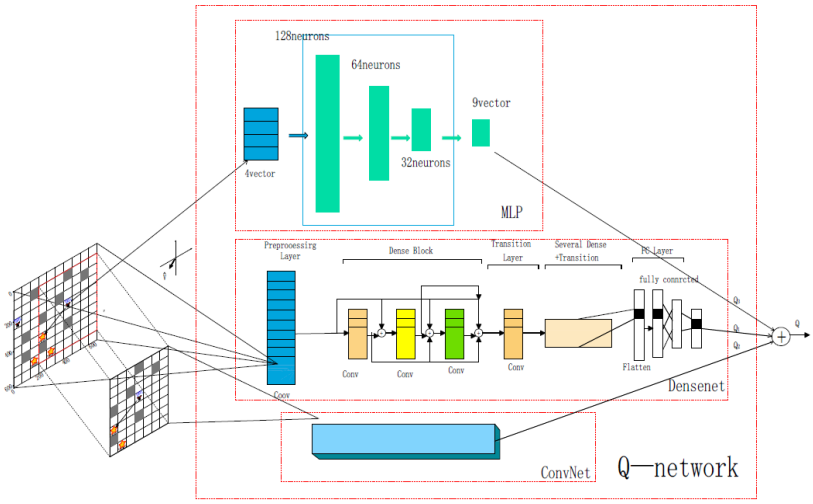

Figure 4 MH-MADQN algorithm network structure\

A convolutional network that perceives a local map, its input is a local map (3*7*7), and its output is, corresponding to the value of 9 actions.And a multi-layer perceptron (MLP) that senses the direction of the target position. Its network parameter settings are shown in Table 3.2. Its input is a four-dimensional direction vector R, and its output is O3, corresponding to the

ISSN: 2664-9640

Q values of 9 actions. The three networks together form the Q network of the MH-MADQN algorithm. Then the output of the MH-MADQN algorithm Q network is Q= O1+ O2+ O3.

3. Simulation Experiment and Analysis of Unknown Threat Multi-UAV Task Assignment

The initial simulation environment includes randomly distributed threat areas, the initial position of the drone, and the position of the target as shown in Figure 5.

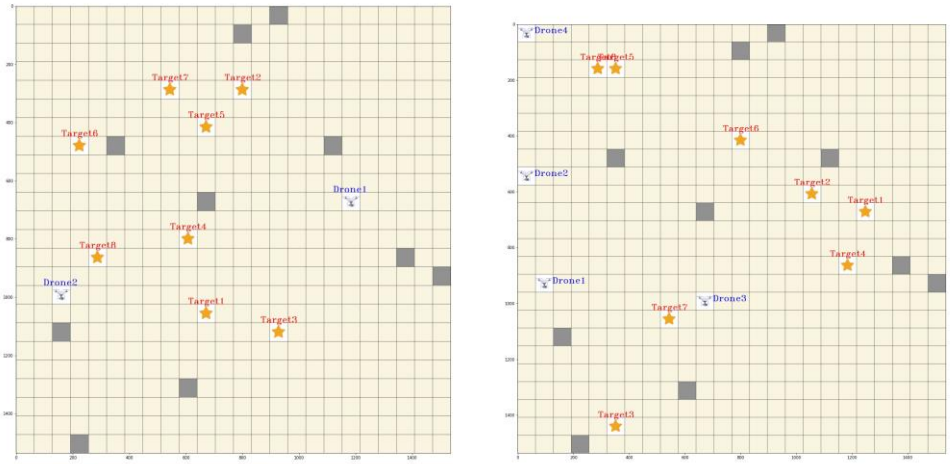

Figure 5 Initialization environment

Task distribution result graph

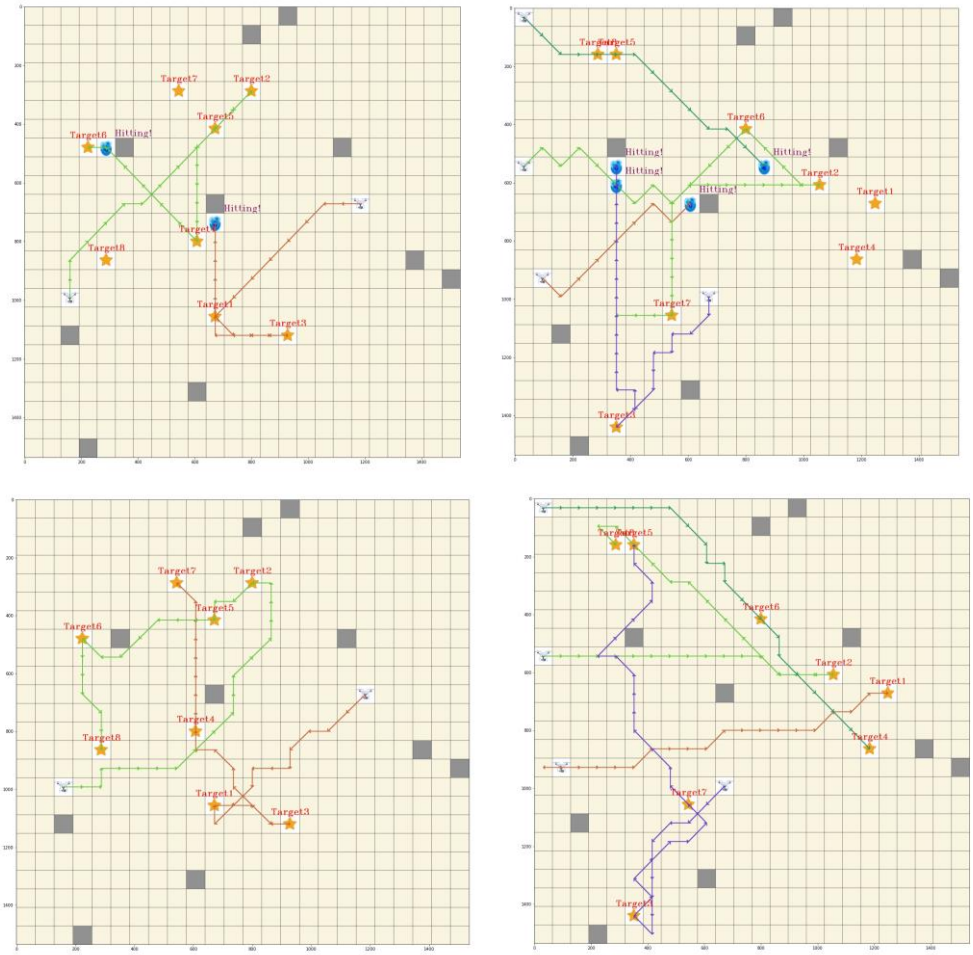

Figure 6 The result of multi-drone task allocation

ISSN: 2664-9640

4. Conclusion

This paper builds a new network structure, which is composed of multi-layer perceptron (MLP), DenseNet, and convolutional neural network (CNN) to form the Q network in the MH-MADQN algorithm, so that the network can not only perceive The static information in the multi-UAV system can also perceive the dynamic information in the environment, which can effectively avoid the unknown threat area, so that the UAV can quickly complete the target task.

References

[1] FENG S, SHU H, XIE B Q. 3D Environment path plan-ning based on improved deep reinforcement learning[J].Computer Applications and Software, 2021, 38(1):250-255.
[2] RODRIGUEZ R, ALEJANDRO S, CARLOS B, et al. A deep reinforcement learning strategy for UAV autono-mous landing on a moving platform[J]. Journal of Intelli-gent & Robotic Systems, 2019, 93(1): 351-366.
[3] ZHANG Y Z, XU J L, YAO K J, et al. Pursuit missions for UAV swarms based on DDPG algorithm[J]. Hang-kongXuebao, 2020, 41(10): 314-326.
[4] LI B, YANG Z P, CHEN D Q, et al. Maneuvering target tracking of UAV based on MN-DDPG and transfer learning[J]. Defence Technology, 2021, 17(2): 457-466.
[5] XIONG L Y, TANG G A, YANG X, et al. Geomorpholo-gy-oriented digital terrain analysis: Progress and perspectives[J]. ActaGeographicaSinica, 2021,76(3): 595-611.
[6] TONG X R, Modeling and realization of real time electronic countermeasure simulation system based on System Vue[J]. Defence Technology, 2020, 16(2): 470-486.
[7] ZHOU W Q, ZHU J H, KUANG M C. An unmanned air combat system based on swarm intelligence [J]. SCIENTIA SINICA Informatiois.2020, 050(003):363 -374 (in Chinese).
[8] Lakas A, Belkacem A N, Hassani S A. An adaptive multi-clustered scheme for Autonomous UAV Swarms [C]. 2020 International Wireless Communications and Mobile Computing (IWCMC), 2020.