Shaopeng LIU, Guohui TIAN, Yongcheng CUI, Xuyang Shao, 2022. A deep Q-learning network based active object detection model with a novel training algorithm for service robots. *Frontiers of Information Technology & Electronic Engineering*, 23(11):1673-1683. https://doi.org/10.1631/FITEE.2200109

A deep Q-learning network based active object detection model with a novel training algorithm for service robots

Key words: Active object detection; Deep Q-learning network; Training method; Service robots

Corresponding author: Guohui TIAN
E-mail: g.h.tian@sdu.edu.cn
ORCID: https://orcid.org/0000-0001-8332-3064

Motivation

1. Active object detection (AOD) is important for service robots to complete tasks in the family environment, and leads robots to approach the target object by taking appropriate moving actions.

2. Most of the current AOD methods are based on reinforcement learning with low training efficiency and testing accuracy.

Main idea

1. Deep Q-learning network is effective for robot active object detection.

2. The Q-learning network (DQN) model is designed to fit the Q-values of various actions, and includes state space, feature extraction, and a multilayer perceptron (MLP).

3. A novel training algorithm based on memory is designed for the proposed DQN model to improve training efficiency and testing accuracy.

4. A method of generating the end state is presented to judge when to stop the AOD task during the training process.

Method

1. The DQN-based model is designed, including the state space, feature extraction, and an MLP.

2. A novel training algorithm based on memory (TAM) is designed to speed up the exploration with high efficiency during the training and to increase the testing accuracy. The proposed training algorithm can generate more training data with positive rewards and prevent the model from repeatedly learning the data with negative rewards.

3. The end state is used to judge whether the robot has arrived at the end point, and the generation method of the end state is presented.

Method (Cont'd)

Dual-constraint denoising network

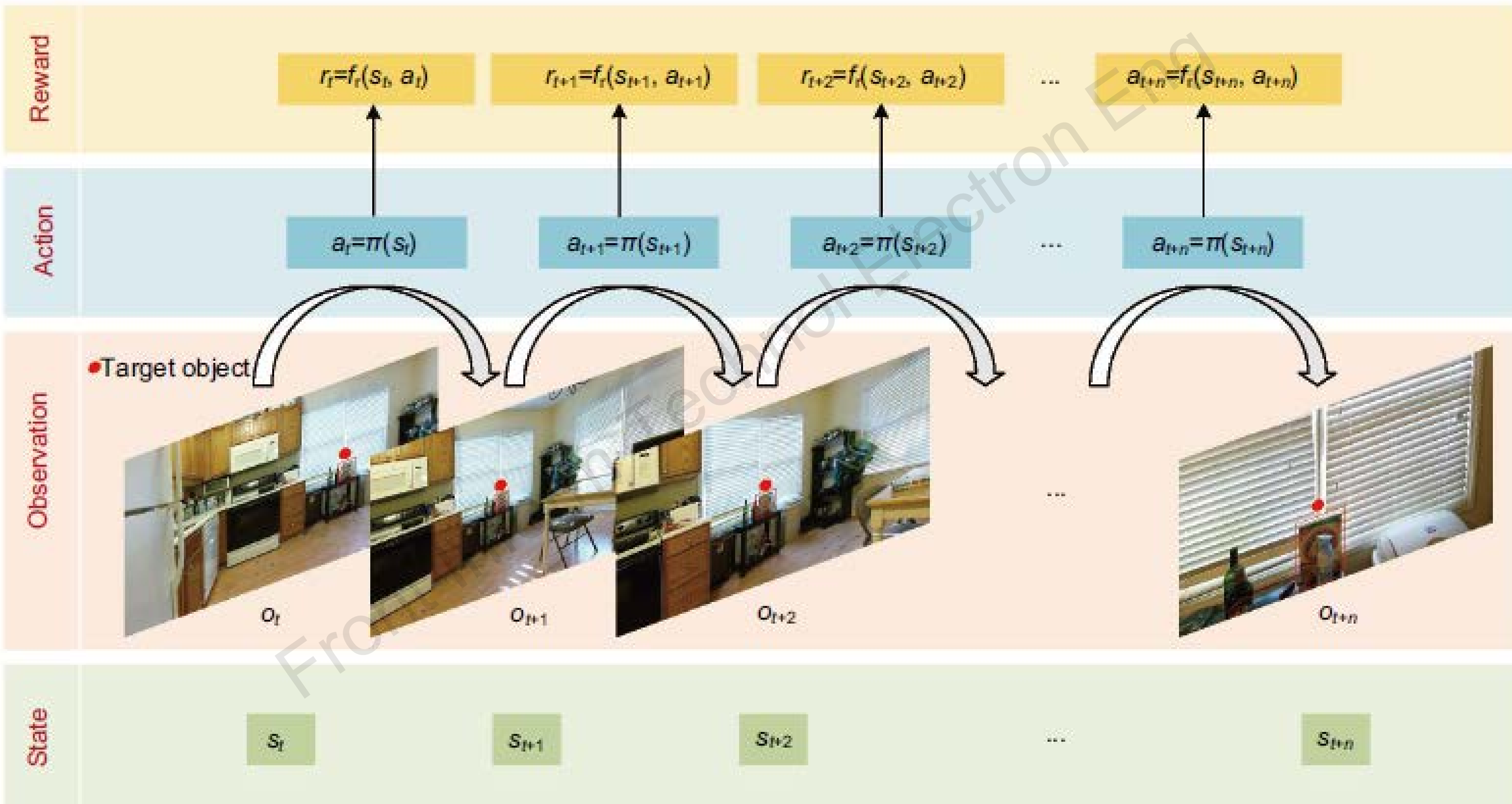

Fig. 1 Workflow of active object detection (AOD) expressed in four parts: state, observation, action, and reward

Method (Cont'd)

A novel training algorithm based on memory (TAM)

Algorithm 1 TAM for robot AOD

Require: initialized training data $\mathcal{D}$
Ensure: well-trained DQN model
Define a storage $\mathbb{S}$ and initialize the evaluation DQN $Q_e(s,a|\theta_e)$ and the target DQN $Q_t(s,a|\theta_t)$ with random parameters θ_e and θ_t
$N_s = 0$ /* N_s is the number of steps */
for episode=1:M **do**
  $N_s = N_s + 1$
  Initialize a scene image I_i with a target object o_{id} as the start state s_t from $\mathcal{D}$
  Obtain s_e /* s_e is the end state of this episode */
  done=False
  continue_search = False
  action_list = $\{a_1, a_2, a_3, a_4, a_5, a_6\}$
  while not done **do**
    if $N_s > 200$ **then**
      Select a random action a_t with probability ϵ; otherwise, $a_t = \max Q_e(s_t, a|\theta_e)$
    else
      Select a random action a_t
    end if
    if continue_search and action_list are not null **then**
      $a_t = a_s$ /* sample an action from action_list */
    end if
    $s_{t+1} = T(s_t, a_t)$
    if s_{t+1} is null **then**
      continue_search = True
      $s_{t+1} = s_t,\ r_t = -1$
    else
      if $s_{t+1} = s_e$ **then**
        done = True, $r_t = 1$
      else
        continue_search = True, $r_t = f_r(s_t, a_t)$
      end if
    end if
    Store (s_t, a_t, r_t, s_{t+1}) in $\mathbb{S}$
    $s_t = s_{t+1}$
    if $N_s > 2000$ **then**
      Sample batch size B_s from $\mathbb{S}$ randomly
      Use B_s to train $Q_e(s,a|\theta_e)$ and update θ_e by the RMSProp optimizer
    end if
    if continue_search **then**
      Remove a_t in action_list
    end if
    if $N_s \% 2000 == 0$ **then**
      $\theta_t = \theta_e$
    end if
  end while
end for
Return $Q_t(s,a|\theta_t)$

Method (Cont'd)

The generation method of end state

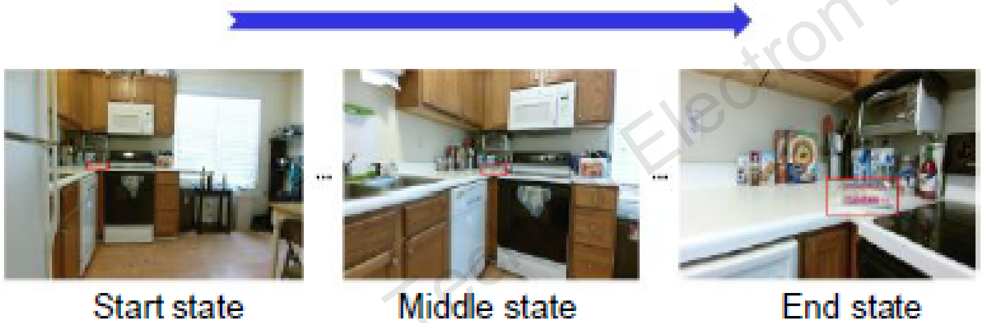

Fig. 4 An example of the end state generation based on the GMES, where the bounding box represents the target object

Major results

Table 3 Results of different DQN-based methods

Method	SR			AS		
	Home_001_2	Home_005_2	Average	Home_001_2	Home_005_2	Average
DQN	0.1984	0.7229	0.4607	26.1421	21.6625	23.9023
DDQN	0.3669	0.8705	0.6187	28.5379	18.2388	23.3884
D3QN	0.1294	0.8223	0.4759	28.1935	19.7875	23.9905
DQN_dueling	0.2559	0.9494	0.6027	23.0592	18.0036	20.5314
Our method	0.6020	0.9849	0.7935	24.1997	17.8532	21.0265

Table 5 Results of the DQN model based on different object detection accuracy (ODA) values

ODA	SR			AS		
	Home_001_2	Home_005_2	Average	Home_001_2	Home_005_2	Average
0.7	0.47	0.69	0.58	22.94	13.88	18.41
0.8	0.54	0.80	0.67	22.80	13.86	18.33
0.9	0.61	0.88	0.75	22.81	13.80	18.31
1.0	0.67	1.00	0.84	22.79	13.88	18.34

Table 6 Results of the DQN model with different state spaces

State space	SR			AS		
	Home_001_2	Home_005_2	Average	Home_001_2	Home_005_2	Average
rgb+box	0.6747	0.9970	0.8359	22.7858	13.8761	18.3310
depth+box	0.4730	0.9398	0.7064	26.4076	19.5962	23.0019
rgb+depth+box	0.5815	0.8102	0.6959	26.5360	19.4572	22.9966

Major results (Cont'd)

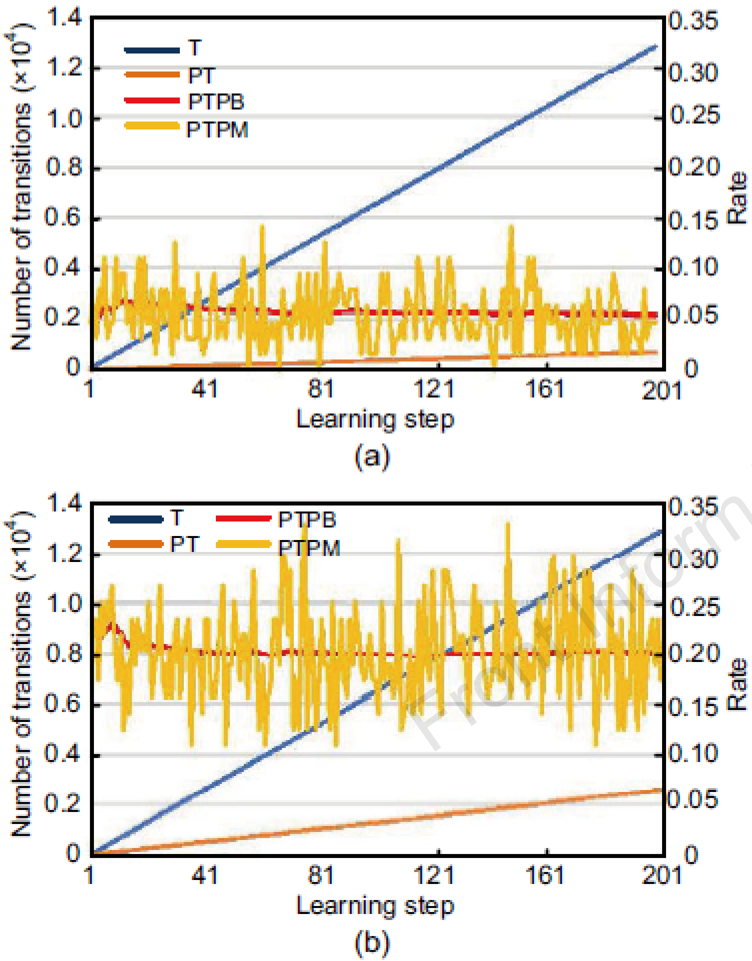

Fig. 7 Curves of T, PT, PTPB, and PTPM based on RTA (a) and TAM (b) during 201 learning steps

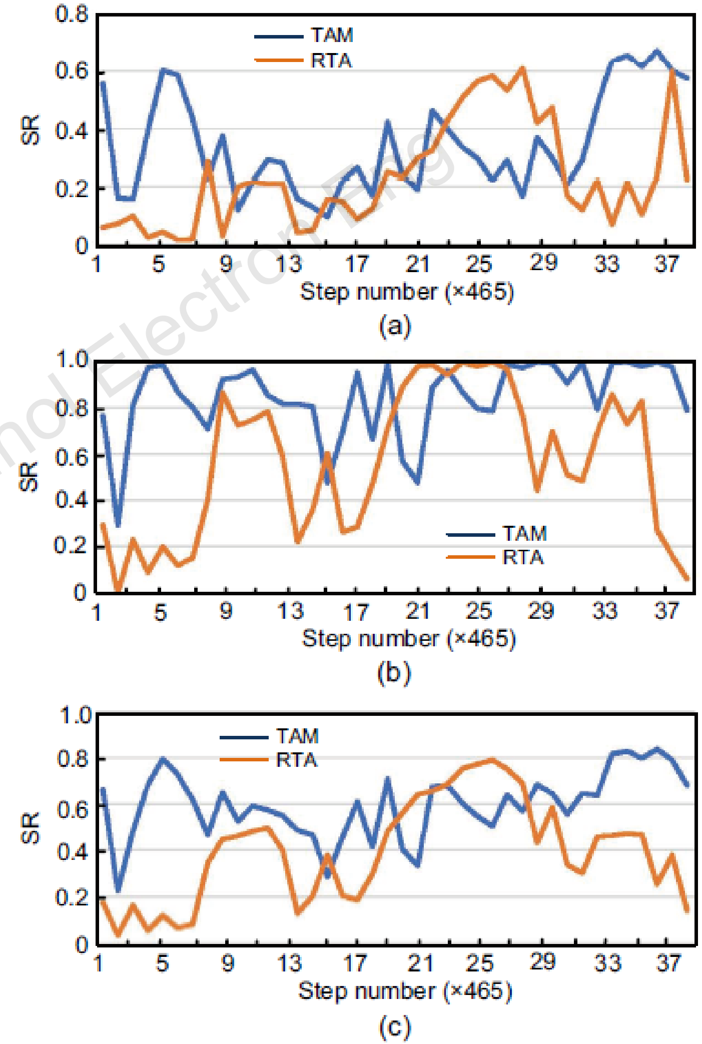

Fig. 8 Test curves of SR based on RTA and TAM in Home_001_2 (a), Home_005_2 (b), and the average SR (c)

Conclusions

1. A DQN-based AOD model with a novel training algorithm is proposed to increase the training efficiency and improve the performance of the DQN model.

2. A training algorithm based on memory is designed to avoid repetitive learning of the data with negative rewards and increase the amount of data with positive rewards in the training process.

Guohui TIAN is a professor at Shandong University. He received his PhD degree from Northeastern University in 1997. From 1999 to 2001, he was a postdoctor with the School of Mechanical Engineering, Shandong University. From 2003 to 2005, he was a visiting professor with the Graduate School of Engineering, Tokyo University, Tokyo, Japan. His current research interests include service robot, intelligent space, cloud robotics, and brain-inspired intelligent robotics.