https://doi.org/10.31256/Ss9Aa2K

Robotic Action-state Evaluation via Siamese Neural Network

Xiang Chang
Aberystwyth University
Aberystwyth, United Kingdom
xic9@aber.ac.uk

Fei Chao
Aberystwyth University
Aberystwyth, United Kingdom
fec10@aber.ac.uk

***Abstract*—Robotic imitation learning methods assist robots to operate in evolving and unconstrained environments. However, current robotic state representation imitation learning methods still must involve human experts to provide sparse rewards that indicate whether robots successfully complete tasks. However, enabling robots to make the action-state evaluation autonomously still remains a challenge, especially for multi-stage complex tasks. Therefore, in this work, we propose a novel Siamese neural network-based robotic action state evaluation system in an imitation learning system, so as to replace human experts in a multi-stage imitation learning process and improve the learning efficiency. One target learning footage is divided into several stages; for each stage, two Siamese network frameworks are created to assess the robotic action-states in terms of both movement and environment changes.**

***Index Terms*—Robotic action state evaluation, Siamese neural network, Imitation learning, Few-shot learning.**

I. Introduction

Autonomous robots that can assist humans in situations of daily life have been a long-standing vision of robotics. The first step toward this goal is to create robots that can learn new tasks autonomously or through simple demonstrations, such as video clips, in unconstrained environments.

Many Recent studies have enabled robots to learn complex multi-stage tasks by extracting video demonstrations into step-by-step instruction images in a time-stepped way [1], [2]. For example, a video demonstration of moving a cup is divided into four moving stages (Catch, pick up, move, and put down) with a length-fixed time step, based on the end frame of each stage as the introduction image. By using state representation learning methods, robots can then encode those introduction images, with extra neural networks, into latent states rather than into image pixels. This way of breaking down complex tasks into step-by-step progressive learning makes it possible for the robot to learn complex tasks autonomously through video demonstrations.

However, this type of method also leads to another problem: stage classifiers used in those studies thus focus only on the low-dimensional representation of the state, but cannot accurately determine whether the action has actually been completed at each stage. Therefore, This method causes error accumulations within each stage and even affects subsequential learning. These existing studies must introduce human experts for the stage completion status assessments. In contrast, in this paper, by modeling human judgment processes, we design two Siamese network frameworks to replace human experts in the robotic action states evaluation process, so as to make the robot imitation learning process autonomous and efficient.

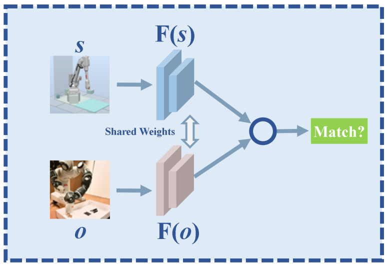

Fig. 1. Assessing Action Changes: A Siamese neural network framework comparing observations (o) and introductions (s)

Siamese neural networks have been used to compare two images and determine if the two images contain the same object [3]. The Siamese neural networks are twin neural networks sharing identical weights. The feature vectors corresponding to the two images are calculated by passing them through two identical convolutional networks. These vectors are used to perform a binary classification indicating if the two domains are similar or not [4], [5]. In other words, with the help of a suitable framework, the Siamese networks are able to display changes in the action and environment as vectors in the latent space, thus enabling a comprehensive judgments of each stage.

II. Approach

We designed two Siamese neural network-based frameworks to replace human experts in the robotic action states evaluation process: One framework focuses on action changes, and the other handles environment changes. The two frameworks work together to provide a comprehensive assessment of the robot's completion of each stage from the view of the action and environment changes, which can provide a more diverse range of rewards.

When human experts judge whether a stage object has been completed, they generally focus on two domains: **(1):**Whether the expected objectives are achieved (e.g., the cup has been placed in the target position; switch has been successfully

Fig. 2. Assessing Environment Changes: A Siamese neural network framework focus on the changes in the environment between different stages.

turned on) **(2):**whether the active state at the end of the stage is consistent with the instruction. (to facilitate the start of the next stage)

Only when both objectives have been met, i.e., only when the target action has been completed and had the desired impact on the environment, we can conclude that the robot has completed this stage. To this end, we design the two frameworks based on the different characteristics of movements and environmental changes:

A. Assessing Action Changes

We propose a two-channel convolutional neural network, where one channel computes feature vectors for introductions, while the other channel computes feature vectors for observations (see Figure 1).

During training, the distance between the feature vectors corresponding to the same stage images is minimized while that between different stage images is maximized. At the end of each stage, we compare the observation image with the introduction image by this framework to determine whether the action at that stage was successfully completed.

If an observation feature is matched to the same stage introduction feature, it means that the robot did the same action with the introduction.

B. Assessing Environmental Changes

We believe that after completing a stage of action, the real environment has the same opportunity as the introduction image. As shown in Figure 2, the second framework focuses on the changes in the environment between two adjacent stages, used as observations or introduction images for different stages, and measures these changes in terms of the vector distance in the latent space.

We used three dual-channel Siamese neural networks (see Figure 2). The first network focuses on changes in the environment between two adjacent stages' introduction images. The change in environment is reflected by the vector distance. The second network is concerned with environment changes between two adjacent stages' observations. The last Siamese neural network judges the vector distance between these two situations. If the vector distance is observed to be the same on both occasions, the robot can be considered to have the same change to the environment and successfully achieved the task goal, as instructed.

III. Experimental Results

To verify the performance of our action-state evaluation framework at the end of each task stage, we designed a robotic task in the simulated environment and recorded it as a video. This video demonstration was divided into four task stages with a length-fixed time step, the last shot of each stage serving as introduction images to the learning process. At the end of each stage, the robotic action-state was captured as the observation image with the same camera angle. The introduction image and observation image at the same stage were compared to assess the state of the robot at this time point to determine whether the robot has completed this task stage.

We extracted the introduction images from the video several times using different time spans to enrich the dataset. A total of 300 different training samples were accumulated. Each sample contained four introduction images and four observation images. After training and comparison, our proposed Siamese network-based system achieved an action state determination success rate up to 83%, such a result proved that our system can preliminarily replace human experts.

IV. Conclusions

In contrast to the sparse rewards by human experts, our system can theoretically give four different kinds of rewards in terms of action and environmental change. In future work, we will further improve the performance of the system and explore the potential applications of the diverse rewards introduced in the learning process.

References

[1] L. Smith, N. Dhawan, M. Zhang, P. Abbeel, and S. Levine, 'AVID: Learning Multi-Stage Tasks via Pixel-Level Translation of Human Videos', presented at the Robotics: Science and Systems 2020, Jul. 2020.

[2] M. Zhang, S. Vikram, L. Smith, P. Abbeel, M. Johnson, and S. Levine, 'SOLAR: Deep Structured Representations for Model-Based Reinforcement Learning', in Proceedings of the 36th International Conference on Machine Learning, 2019, pp. 7444–7453.

[3] S. G. Venkatesh and B. Amrutur, 'One-Shot Object Localization Using Learnt Visual Cues via Siamese Networks', in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019, pp. 6700–6705.

[4] X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, 'Siamese Neural Network Based Few-Shot Learning for Anomaly Detection in Industrial Cyber-Physical Systems', IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5790–5798, 2021.

[5] D. Chicco, 'Siamese Neural Networks: An Overview', in Artificial Neural Networks, H. Cartwright, Ed. New York, NY: Springer US, 2021, pp. 73–94.