SAC+GAARA:
Generate Augmented And Reconstructable latent states with Ae

Jizheng Chen, Junru Gong

Shanghai Jiao Tong University

Objectives

Combine SAC+AE with contrastive learning. Use new model SAC+GAARA to solve DMControl tasks.

- cheetah-run
- finger-spin
- walker-walk

Introduction

We revisit the concept of adding an autoencoder to model-free RL approaches focused on *off_policy* algorithms and process of data augmentation. By runnning them as baselines we confirm that apart from using a decoder to minimize the pixel reconstruction loss, using contrastive learning and data augmentation to further enhance the encoder's capability is vital. Also, in order to dynamically balance contrastive learning and reconstruction, we proposed an annealing way on the data augmentation loss, and ablation study on following parts verifies its accesibility. We name the work GAARA, indicating the process of generating augmented and reconstructable latent states with Ae.(The name GAARA also pays tribute to the beloved charactor Gaara in Cartoon Naruto Shippuden.)

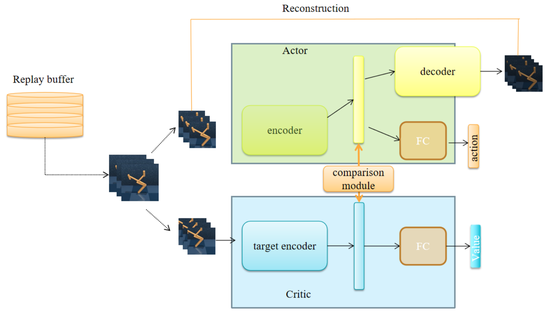

Figure 1:Frame of GAARA

Group Information

Jizheng Chen

Junru Gong

Figure 2:portrait

Information about Jizheng Chen:

- Major in Computer Science
- Responsible for the idea GAARA and code for released version, write proposed method, technical details and conclusions parts in final report

Information about Junru Gong:

- Major in Soft Engineering
- Responsible for experiments and analyse,write related work, background and experiment Results in final report

Results

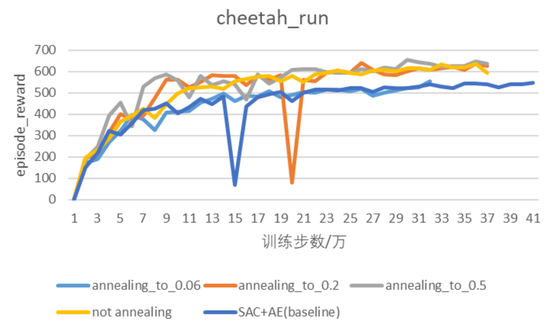

Figure 3:Experiments results

In cheetah run task, SAC+GAARA shows amazing performance.Our model converges faster and gain better reward than SAC+AE baseline.
In walker walk task, SAC+GAARA learns better than baseline in 360k steps. In finger spin task, SAC+GAARA also reach the standard of baseline. As a result, SAC+GAARA is talented and powerful on current experimental environment.

Important Result

In cheetah run task, SAC+GAARA shows amazing performance.Our model converges faster and gain better reward than SAC+AE baseline.

Mathematical Section

The reconstruction loss can be formulated as:

$$J(RAE) = \mathbb{E}_{\mathbf{o}_t \sim \mathcal{D}}[\log p_\theta(\mathbf{o}_t|\mathbf{z}_t) + \lambda_{\mathbf{z}}||\mathbf{z}_t||^2 + \lambda_\theta||\theta||^2]$$

The comparison loss can be formulated as:

$$loss_{com} = \alpha_{ann} \times log \frac{exp(q^T W k_+)}{exp(q^T W k_+) + \sum_{i=0}^{K-1} exp(q^T W k_i)}$$

Methods

This section describes the structure of our SAC+GAARA method, within which the SAC part isn't changed much from [2], where an actor, a critic, and a critic target network is maintained. The whole framework of SAC+GARRA is demostrated in Figure 1. We are inspired by CURL:[1].

Conclusion

From the process of improving and experimenting with the original model, we can see that combining the method of contrastive learning with the AE architecture can indeed improve the speed of training convergence and the model effect. And in this process, the two students in our group did not simply pursue the reproduction of the paper, but actively explored and sought ways to improve. In the process of reading the original paper and code over and over again, I not only deepened our understanding and comprehension of the original algorithm, but also experienced the fun of exploration, and was able to propose innovative methods and skills to further improve the effect of the model. The final results also confirm our hard work and efforts, proving that GAARA's method really has a lot of usefulness and research space. In the future, if there is time, we will consider further improving the network architecture, enhancing the expression ability of the network, and further improving the effect of the model.

References

[1] A. Srinivas, M. Laskin, and P. Abbeel. Curl: Contrastive unsupervised representations for reinforcement learning. 2020.

[2] D. Yarats, A. Zhang, I. Kostrikov, B. Amos, J. Pineau, and R. Fergus. Improving sample efficiency in model-free reinforcement learning from images. 2019.

Contact Information

- Web: github/Otsuts/SAC-GAARA
- Email: humihuadechengzhi@sjtu.edu.cn
- Phone: 15104367355