Determining the Possibility of Transfer Learning in Deep Reinforcement Learning Using Grad-CAM (Student Abstract)

Ho-Taek Joo, Kyung-Joong Kim
School of Integrated Technology, Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea
{hotaek87, kjkim}@gist.ac.kr

Abstract

Humans are usually good at guessing whether the two games are similar to each other and easily estimate how much time to master new games based on the similarity. Although Deep Reinforcement Learning (DRL) has been successful in various domains, it takes much training time to get a successful controller for a single game. Therefore, there has been much demand for the use of transfer learning to speed up reinforcement learning across multiple tasks. If we can automatically determine the possibility of transfer learning in DRL domain before training, it could efficiently transfer knowledge across multiple games. In this work, we propose a simple testing method, Determining the Possibility of Transfer Learning (DPTL), to determine the transferability of models based on Grad-CAM visualization of the CNN layer from the source model. Experimental results on Atari games show that the transferability measure is successfully suggesting the possibility of transfer learning.

Introduction

Recently, Grad-CAM (Selvaraju et al., 2017) has been proposed to visualize the internal responses of CNN layers for classification tasks. In our previous work, we proposed to use the Grad-CAM tool for reinforcement learning to understand the agent's action selection and attention mechanism (Joo and Kim 2019). During the analysis, we found that the visual layer mostly used to detect objects on the game screen and indicated the possibility of the general use of the layer for multiple games.

In this work, we propose a simple testing method using some screenshots of other games (e.g., 1000 images) to determine whether the game is a good target of transfer learning from the source. It compares the outputs of Grad-CAM and actual images using SSIM (Wang et al. 2004) and reports the measure of transferability. Our testing on four Atari games using DemonAttack source model demonstrated our testing scheme works well to measure the transferability.

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

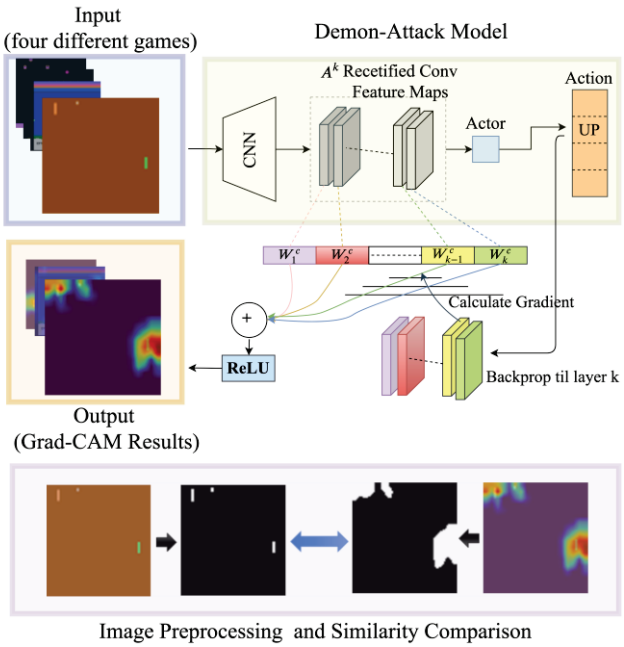

Figure 1: Visualization process of the CNN layer for the Demon-Attack Model with inputs from other game images

Although there are some works on the use of transfer learning in reinforcement learning (Taylor and Stone 2009; Barreto et al. 2017), there are little works on the transferability measure.

Our Model

Our model is shown in Figure 1. First, we train the model specialized DemonAttack in Atari 2600 games. This model is used as the source model for TL. After that, we collect the screen images like the 'Input' in Figure 1. These images are input into the DemonAttack model, while a visualization method called Grad-CAM is applied. Grad-CAM results indicate that the one of the CNN layer's role is to detect objects, and it can surprisingly work well when some games' images are input.

The following process is a way to evaluate how well to detect objects in other games' images. By comparing the area of the objects in the game images with the Grad-CAM results that the objects are detected, we measure the transferability. To verify transferability, we conduct various TL experiments.

Experiments and Results

Training the Model using A3C We implement DRL model called Asynchronous Advantage Actor-Critic (A3C) (Mnih et al., 2016), and train the DemonAttack model.

Visualization using Grad-CAM We apply Grad-CAM to the DemonAttack model with input from other games' images as follows.

- By inputting four games' images to the DemonAttack model till the last CNN layer, we get the feature map (A_{ij}^{k}).
- From predicted action to last CNN layer, we can calculate the weighted gradient ($\mathbf{w}_k^c$) by back-propagation.
- Calculate the dot product of (A_{ij}^{k}) and ($\mathbf{w}_k^c$), and apply the activation function; ReLU.
- eq. (1) is the whole process for Grad-CAM

$$S^c = L^c_{\text{Grad-CAM}} = ReLU\left(\underbrace{\sum_k}_{\text{class feature weights}} \underbrace{w_k^c}_{} \overbrace{\frac{1}{Z}\sum_i\sum_j}^{\text{global average pooling}} \underbrace{A_{ij}^k}_{\text{feature map}}\right) \quad (1)$$

We experiment with 1,000 game images for each game.

Measurement for Transferability We measure the transferability depending on how well the object is detected. We use a traditional method for detecting objects called SSIM. We first preprocess the images using a locally adaptive threshold, as shown in Figure 1. Then we measure transferability using SSIM.

	Phoenix	Pong	SpaceInvaders	Seaquest
SSIM	**0.62**	**0.46**	**0.32**	**0.17**
Rank	**1**	**2**	**3**	**4**

Table I: Results of the Measurement for Transferability

Transfer Learning We use the fine-tuning that holds the weights of a source neural network except for the last Fully-Connected (FC) layer and the FC layer is replaced by a new random weighted layer. Similarly, our algorithm updates only the parameters of the FC layers while keeping the parameters of the CNN layers. Figure 2 shows the comparison between the original model and transfer learning. 'Reward Mean' on the y-axis is total episode reward/number of the episode. Comparing the ranking of transferability in the previous section with the results of the Figure 2 experiment, the higher the rank, the better the transfer learning.

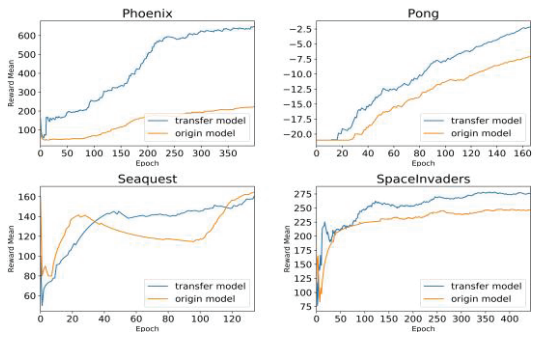

Figure 2: Comparison between origin model and transfer model.

Conclusion

In this work, we propose an approach to determine and evaluate the transferability in DRL. Our experimental results show that CNN layers in the model is mostly trained to detect objects on the Atari game screen. Therefore, if the layer is also working well in other games' screen images, it indicates the possibility of transfer learning is high. Our contribution is to propose an automatic method using Grad-CAM visualization to help transfer learning decisions.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (grant no. 2017R1A2B4002164).

References

Barreto, A.; Dabney, W.; Munos, R.; Hunt, J. J.; Schaul, T.; van Hasselt, H. P.; and Silver, D. 2017. Successor features for transfer in reinforcement learning. In *Advances in neural information processing systems*.

Joo, H. T., and Kim, K. J. 2019. Visualization of Deep Reinforcement Learning using Grad-CAM. In: 2019 *IEEE Conference on Games.*

Mnih, V.; Badia, A. P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Lillicrap, T. P.; Silver, D, and and Kavukcuoglu, K. 2016. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*

Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; and Batra, D. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision*.

Taylor, M. E., and Stone, P. 2009. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*. 10(1):1633–1685.

Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing.*