Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 9-10, 2024

Group invariant neural networks-based deep reinforcement learning for optimal flow control in nuclear reactor systems

Joongoo Jeon [a,b*], Su Yeong Jo [a], Minseo Lee [b], Shilaj Baral [a], Sangam Khanal [a], Ricardo Vinuesa [c], Jean Rabault [d]
[a]*Jeonbuk National University, Graduate School of Integrated Energy-AI, 54896 Jeonju-si, South Korea*
[b]*Jeonbuk National University, Department of Quantum System Engineering , 54896 Jeonju-si, South Korea*
[c]*KTH Royal Institute of Technology, FLOW, Engineering Mechanics, SE-100 44 Stockholm, Sweden*
[d]*Independent Researcher, Oslo, Norway*
[*]*Corresponding author: jgjeon41@jbnu.ac.kr*

Keywords : flow control; reinforcement learning; machine learning; Rayleigh-Bènard convection; Nusselt number

1. Introduction

Currently, countries around the world are in a race to develop cutting-edge design concepts for small modular reactor. For example, Choi et al. investigated the feasibility of a helium bubbling system for natural circulation in molten salt fast reactor [1]. A common strategy among them involves enhancing flow control system efficiency and safety through the application of artificial intelligence techniques [2, 3]. Reinforcement learning (RL), capable of training an agent to find optimal actions in a dynamic environment for maximizing cumulative rewards, emerges as a promising technique in flow controls [4].

Nevertheless, the applicability of state-of-the-art machine learning methods within flow control remains unexplored [4]. The objective of this study is to address this gap by developing a novel deep reinforcement learning (DRL) method focused on maximizing learning efficiency in flow control. Notably, efficient reinforcement learning is particularly crucial to nuclear system applications, where the cost of episode exploration is prohibitive, such as helium bubbling system [1].

We developed a group invariant deep reinforcement learning framework to reduce state representation complexity by exploiting symmetries. In other words, the original flow field can be represented in symmetry-reduced subspace. Even under conditions where the number of episodes is insufficient for general method, this framework can find an optimal policy by leveraging a combination of multi-agent and group invariant approaches. To evaluate the performance of this study, Nu control of two-dimensional Rayleigh-Bénard convection (RBC) was selected as a case study. In nuclear reactor severe accidents, RBC occurs during molten core-concrete interactions. The performance of our DRL framework was compared with the framework in previous study [5].

2. Methods and Results

2.1 Reinforcement learning

Reinforcement learning is a branch of machine learning where an agent interacts with an environment to learn optimal behaviors through trial and error. As shown in **Fig. 1**, The agent is the entity responsible for making decisions within the environment, and it takes actions that influence the state of the environment. The environment represents the external system in which the agent operates, and its state encapsulates the current situation.

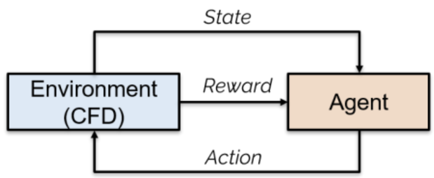

Fig. 1. Example of reinforcement learning framework with computational fluid dynamics (CFD).

Actions are the decisions or moves made by the agent to transition from one state to another. The agent receives feedback in the form of rewards or penalties based on the actions taken, providing a signal to guide its learning. The policy is the strategy or set of rules the agent employs to determine its actions in different states. The goal of reinforcement learning is for the agent to learn a policy that maximizes the cumulative or instantaneous reward over time, leading to optimal decision-making in complex and dynamic environments.

However, Reinforcement learning faces challenges related to computational cost and the potential for the emergence of suboptimal or bad policies. The computational cost arises from the need for agents to interact with the environment, learn from experiences, and update their policies iteratively. In complex and large state or action spaces, training a reinforcement learning model can demand significant computational resources and time, making it impractical for certain applications. Additionally, when we increase the learning rate to reduce computation time, there is a problem of convergence to a bad policy [6].

2.2 Group invariant neural networks

The flow control system in a nuclear reactor also has a complex and large state dimension. For practical reinforcement learning applications, we need to minimize state representation complexity according to characteristics of each flow control system. If the original flow field can be represented in symmetry-reduced subspace, we can explore the optimal policy more efficiently [7]. For example, in 2D Rayleigh-Bénard convection flow, the same action is required in a symmetric state. Thus, a symmetry-invariant network architecture is required as shown in **Fig. 2**.

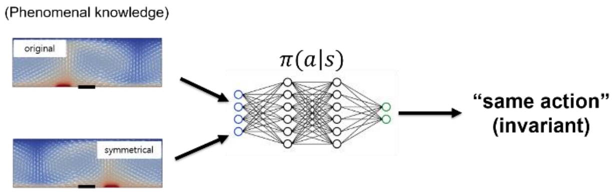

Fig. 2. 2D RBC: symmetrical invariant policy problem

Recently, Lafarge et al. developed group equivariant convolutional networks for image classification purposes. Their roto-translation equivariant convolutional layer is an architecture whose output is equivariant depending on the translation and rotation of the image [8]. However, in our reinforcement learning framework, an architecture where actions are invariant (not equivariant) depending on the state can improve learning performance.

For this reason, we developed the group invariant convolutional neural networks (GI-CNNs) as shown in **Fig. 3.** In our novel network, the combination of direct lifting layer and shifted kernels produces completely invariant output under symmetrical state conditions. As shown in **Fig. 4**, the group invariant convolutional layers were built in the front of the fully connected layer.

2.3 Shenfun with the Spectral Galerkin method

As described in *Section 2.1*, the environment of reinforcement learning represents the external system in which the agent operates. To investigate the optimal policy of flow control, we can employ CFD as the environment. In this study, Shenfun, high performance computing platform for solving partial differential equations (PDEs) by the spectral Galerkin method, was used as the CFD solver. Because this platform was developed by Python language, it is compatible with Tensorforce (a Python-based reinforcement learning framework).

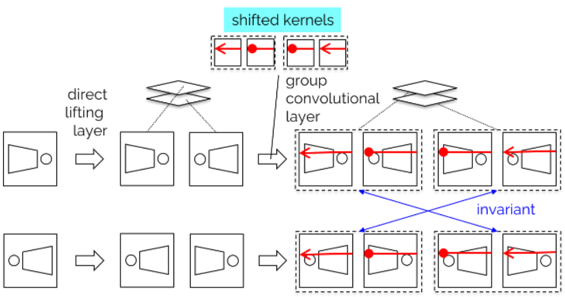

Fig. 3. Group invariant convolutional networks for symmetry-reduced representation.

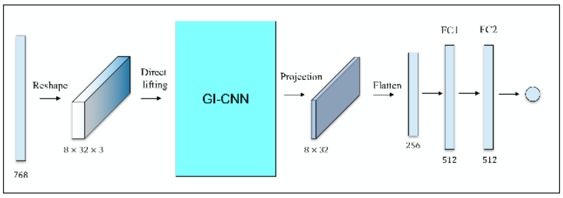

Fig. 4. The overall framework of deep neural networks (agent) in our reinforcement learning.

2.4 RBC Simulation

Fig. 5 illustrates the schematic of the RBC simulation in this study. The flow occurs between a lower hot wall with an average temperature of $T_H = 2$ and an upper wall kept cooler at a constant temperature of $T_C = 1$. While T_H remains consistent in what we term the 'baseline' scenario (i.e., without any control applied), it varies across spatial positions in what we refer to as the 'controlled' scenario, as elaborated below. Both walls are separated by a distance of $2H$, and the no-slip boundary condition applies to both. The lateral ends of the domain, with a normalized width of $L = 2\pi H$, are subjected to periodic boundary conditions.

Throughout the study, a Prandtl number of $Pr = 0.71$ is utilized, corresponding to air's Prandtl number. Ra represents the ratio of time scales related to thermal transport through diffusion and convection. This indicates that higher Ra values render flows more susceptible to instabilities due to buoyancy-driven convection. The Ra employed in this study is $Ra = 10^4$, surpassing the critical Ra_c when no control is exerted on the flow. It means that effective controlling convection flow was required to minimize/maximize heat transfer.

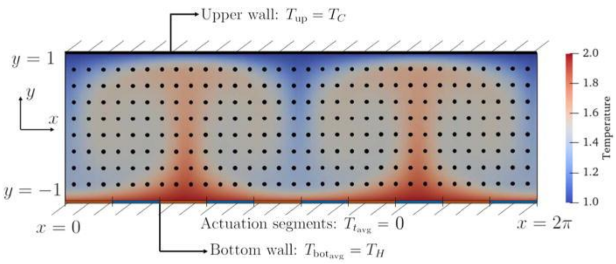

Fig. 5. Schematic representation of 2D RBC simulation

2.5 Results

In this study, the performance of our novel architecture was evaluated on the Nu control of 2D RBC. This is a case study that minimizes Nu by controlling the temperature of the lower wall divided into 10 sections. A uniform lower temperature initial condition is given to the agent and no correct answer is given to agent as to the optimal policy.

- ✓ State: 2D temperature and velocity field
- ✓ Action: Wall temperature change (10 segments)
- ✓ Reward: Function by instantaneous Nusselt number

Fig 5. shows the learning curves obtained in both fully connected based RL and GI-CNNs based RL. Reinforcement learning searches for the optimal policy through episodes, and a numerous number of episodes means high computing costs. It was conformed that the GI-CNNs exploiting symmetry-reduced subspace converges to the optimal policy much more efficiently (about 150 episodes faster). However, in several independent learning curves, it was confirmed that the minimum Nu number was lower for FC architecture. It seems that the invariant representations have side effects that can limit the agent's actions. A detailed analysis of this is our future work.

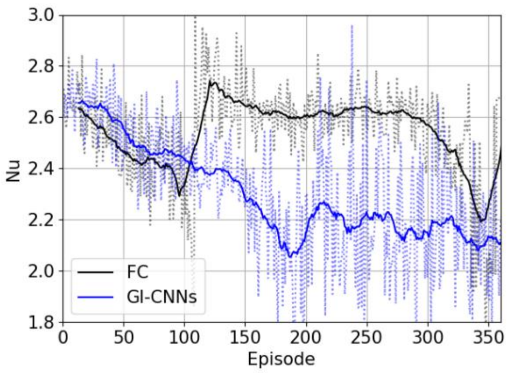

Fig. 5. Comparison of learning curves obtained in both the fully connected (FC) and GI-CNNs.

3. Conclusions

In this study, we proposed a group invariant neural networks to enhance the DRL performance. This architecture can reduce state representation complexity by exploiting symmetries. In Nu control of 2D RBC case study, it was confirmed that this architecture finds the optimal policy 150 episodes faster than previous architecture model. Thus, this work demonstrates the potential of group invariant DRL in the nuclear reactor flow control applications. However, in several independent learning curves, it was confirmed that the minimum Nu number was lower for FC architecture. A detailed analysis of this is our future work.

ACKNOWLEDGEMENT

This work was supported by the "Human Resources Program in Energy Technology" of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), granted financial resource from the Ministry of Trade, Industry and Energy, Republic of Korea (No. 20204010600470) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00278121).

REFERENCES

[1] Choi, W. J., Park, J. H., Lee, J., Im, J., Cho, Y., Kim, Y., & Kim, S. J. (2023). Experimental and numerical assessment of helium bubble lift during natural circulation for passive molten salt fast reactor. Nuclear Engineering and Technology, In Press.

[2] Jeon, J., Lee, J., & Kim, S. J. (2022). Finite volume method network for the acceleration of unsteady computational fluid dynamics: Non-reacting and reacting flows. International Journal of Energy Research, 46(8), 10770-10795.

[3] Jeon, J., Lee, J., Vinuesa, R., & Kim, S. J. (2024). Residual-based physics-informed transfer learning: A hybrid method for accelerating long-term CFD simulations via deep learning. International Journal of Heat and Mass Transfer, 220, 124900.

[4] Garnier, P., Viquerat, J., Rabault, J., Larcher, A., Kuhnle, A., & Hachem, E. (2021). A review on deep reinforcement learning for fluid mechanics. Computers & Fluids, 225, 104973.

[5] Vignon, C., Rabault, J., Vasanth, J., Alcantara-Avila, F., Mortensen, M., & Vinuesa, R. (2023). Effective control of two-dimensional Rayleigh-Benard convection: invariant multi-agent reinforcement learning is all you need. Physics of Fluids, 35, 065146.

[6] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

[7] Belus, V., Rabault, J., Viquerat, J., Che, Z., Hachem, E., & Reglade, U. (2019). Exploiting locality and translational invariance to design effective deep reinforcement learning control of the 1-dimensional unstable falling liquid film. AIP Advances, 9(12).

[8] Lafarge, M. W., Bekkers, E. J., Pluim, J. P., Duits, R., & Veta, M. (2021). Roto-translation equivariant convolutional networks: Application to histopathology image analysis. Medical Image Analysis, 68, 101849.