Active flow control using machine learning: A brief review

Feng Ren[1], Haibao Hu[2], Hui Tang[1*]

[1]Research Center for Fluid-Structure Interactions, Department of Mechanical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China

[2]School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, Shaanxi, China

[*] E-mail: h.tang@polyu.edu.hk

Abstract

Nowadays the rapidly developing artificial intelligence has become a key solution for problems of diverse disciplines, especially those involving big data. Successes in these areas also attract researchers from the community of fluid mechanics, especially in the field of active flow control (AFC). This article surveys recent successful applications of machine learning in AFC, highlights general ideas, and aims at offering a basic outline for those who are interested in this specific topic. In this short review, we focus on two methodologies, i.e., genetic programming (GP) and deep reinforcement learning (DRL), both having been proven effective, efficient, and robust in certain AFC problems, and outline some future prospects that might shed some light for relevant studies.

Key words: Active flow control; machine learning; genetic programming; deep reinforcement learning.

1. Background

AFC has been hot in fluid mechanics, in which a fluid system is purposely altered by actuators through exerting a small amount of energy input. Compared to passive controls that usually involve geometrical changes, AFC is adaptive and hence can realize more effective control in a much wider operation range. Depending on whether the signals from the system output are fed back to regulate the actuator(s), AFC can be either open-loop or closed-loop. Different from open-loop control, closed-loop control can adjust actuation using feedback signals from sensors, hence realizing automatic operations in a much wider range.

Due to the nonlinearity of fluid dynamics, it is challenging to obtain an effective, efficient, and sufficiently robust control strategy for multi-input-multi-output flow control problems. In its broadest sense, machine learning can be classified into many fields, including regression, classification/clustering, supervised learning, unsupervised learning, ensemble learning, deep learning, reinforcement learning, etc. Note that overlapping often exists in these subclasses. For example, deep neural networks are often adopted in reinforcement learning algorithms, which is then called deep reinforcement learning (DRL). Benefit from advanced algorithms, powerful hardware devices, and massive data, the booming machine learning has witnessed intensive successes in diverse fields. For example, nowadays state-of-the-art machine learning technologies have developed comparable or even superior capabilities over human beings in identifying images, processing natural languages, playing games, manipulating robotic body, etc.

Successes of machine learning have also continuously attracted attentions from the community of fluid mechanics. For some specific areas of fluid mechanics, researchers can find some common features from other disciplines. For example, Cai et al.[1] applied the convolutional neural network (CNN) as a global estimator for particle image velocimetry (PIV), which is trained using a synthetic dataset and evaluated using both artificial and laboratory PIV images, showing superior performance than traditional methods in terms of computational efficiency. Xiao et al.[2] conducted a series of studies for turbulent modeling using machine learning. Huang et al.[3] designed a deep learning model that successfully predicts 3D flame evolution based on historic 2D projections obtained from time-resolved volumetric tomography. Wang et al.[4] developed a reduced-order model for unsteady combustion flows using artificial neural networks (ANNs), showing good efficiency and robustness. Wu et al.[5] used the generative adversarial network (GAN), combining the CNNs, to establish a mapping from a parameterized supercritical airfoil to its corresponding transonic flow field. The flow field is then used to assess aerodynamic

performance and perform optimization. Results show that the machine learning-based surrogate model is superior in terms of accuracy and efficiency over most existing surrogate models.

Note that, for most applications mentioned above, machine learning is used in a supervised manner, which means that the model works mainly as an estimator and does not influence the flow system. Moreover, to train the model, the ground truth data need to be continuously fed to the model. Differently, the machine learning-based AFC is usually conducted in a semi-supervised way, that is, the performance is evaluated through a prescribed cost function other than continuous data input. An ideal evolution/training process would constantly improve the control performance and finally achieve convergence. In subsequent sections, two main methodologies used in the machine learning-based AFC will be discussed, i.e., the GP and the DRL, together with relevant successful applications in the past few years.

2. Active flow control using genetic programming
2.1 Genetic programming method

The GP is a symbolic regression method in machine learning. Initially proposed by Koza[6], this concept was inspired by the genetic algorithm (GA)[7]. Although they share many common features, such as replication, crossover, mutation, etc., the major difference is that the GP generates symbolic expressions using the locator/identifier separation protocol (LISP) language, while the GA only produces optimized values. The symbolic regression enables the GP to derive model-free controllers, where each symbolic expression becomes an explicit control law.
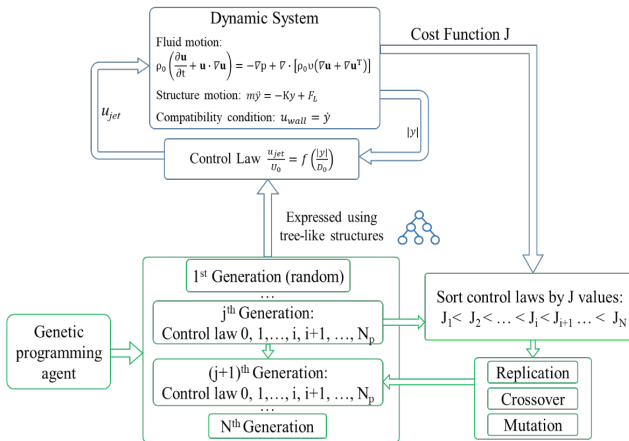

Fig. 1 Schematic of the GP evolution framework for the AFC

Taking the vortex-induced vibration (VIV) suppression as an example[8], Fig. 1 presents a schematic diagram of the GP evolution framework for AFC. Every GP generated control law will be assessed by a fluid-structure-interactions (FSI) simulation module, which couples the fluid flow, structure motion, and the actuation. The outcome of the simulation module, i.e.,

the cost function J that measures the control performance and is also weighted by energy consumption of actuators, will then be sent back to the GP selection module. After the assessment is done for all control laws in one generation, they will be ranked according to their J values. The best few control laws generating smallest J values are chosen to produce the candidate control laws of the next generation using the GP.

For the GP framework, it is assumed that given sufficiently large population size and after sufficiently long generations, the algorithm can converge to globally optimal individuals. Initially, the population members are given random tree-like expressions within a certain range of depth. During this evolution three principal genetic operations are employed, i.e., replication, crossover and mutation. First, the tournaments can replicate themselves to participate in competitions in the next generation. Second, they also get a chance to breed, so that their children can inherit the genes from two excellent parents and could possibly perform better than their parents. This is a process called crossover (as depicted in Fig. 2(a)). Last, the tournaments may also experience mutations, in which a part of the tree-like expression structure happens to be replaced by another randomly produced expression. Three typical mutation modes, i.e., subtree mutation, hoist mutation and point mutation, are adopted in this study, as depicted in Fig. 2(b-d).
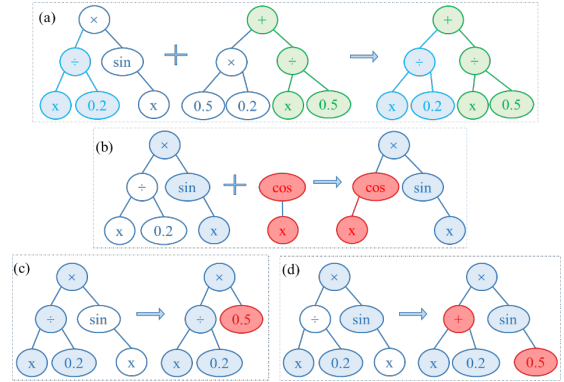

Fig. 2 Examples of main modes in generating new populations: (a) crossover (b) subtree mutation (c) hoist mutation (d) point mutation

2.2 Numerical study

Using a high-fidelity solver, Ren et al.[8] numerically realize the first GP-based AFC, targeting at suppressing the VIV of a circular cylinder that is connected with a spring in the transverse direction, as shown in Fig. 3(a). The chosen working condition for this mass-spring system lies in the lock-in region (with a mass ratio of 2 and a reduced velocity of 5), i.e., the frequency of flow-driven vibration matches the system's resonance frequency and hence large-amplitude vibration occurs[9], making the VIV difficult to suppress.
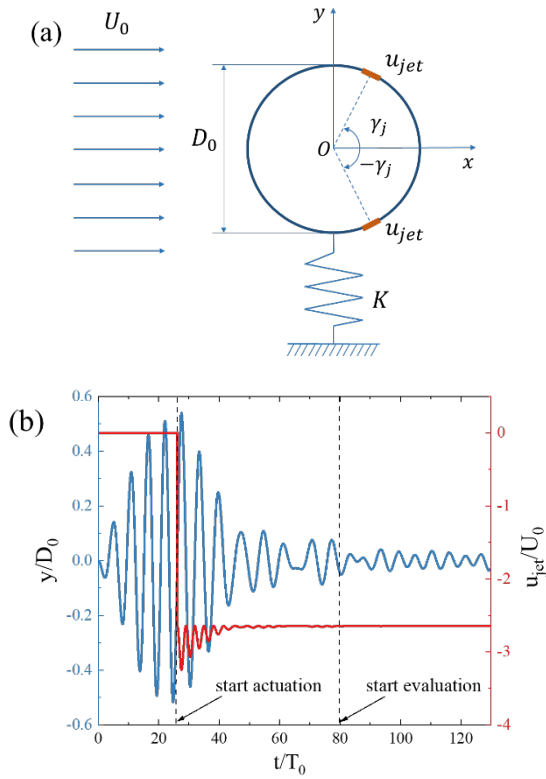
Fig. 3 (a) Schematics of the vortex-induced vibration system (b) Evolution of the actuation strength and the cylinder's transverse location in the case controlled by the best genetic programming control law

In this dynamic system, the instantaneous transverse displacement of the cylinder functions, y, as the control input, and the blowing/suction velocity issued from a jet pair, i.e., u_{jet}, functions as the control output. The control law that establishes their inter-relationship is obtained using the GP, through an evolution of 25 generations and overall 1,250 individual cases. The evolution converges after around 5 generations, generating an unexpected optimal control law that suggests a suction-type actuation with the actuation strength nonlinearly increasing with the cylinder's transverse displacement. Based on this control law, the performance, evaluated by the root-mean-square value of transverse displacement and the energy consumption, is improved by 21.4% than the best open-loop controls. Considering the Reynolds numbers ranging from 100 to 400, the GP control maintains superior performance over the open-loop control, and good robustness than the linear proportional control.

2.3 Experimental studies

Compared with time-consuming numerical simulations, a well-designed experiment can realize machine learning based AFC with relative ease, since the evaluation of each control law usually takes only seconds. Thus we see almost all GP-based AFC were realized in experiments. Gautier et al.[10] first applied the GP to control the recirculation area of a backward-facing step. The optimal control law, converged after 12 generations, could reduce the recirculation area by 80%. Following this work, Debien et al.[11] used the GP control to mitigate the separation and early reattachment of a turbulent boundary layer from a sharp-edge ramp. In turbulent mixing layer manipulation experiments, Parezanovic et al.[12] showed that the GP control can find the same velocity signals essential for optimal control as those observed in reduced-order model based feedback controls. Furthermore, under the lock-on condition, they also demonstrated that the robustness of the GP control significantly outperforms open-loop controls when varying the freestream velocities[13]. Using another variant model of GP, i.e., the linear GP, Li et al.[14] conducted close-loop controls on a car model and achieved a 22% drag reduction.

Targeting at enhancing turbulent jet mixing, Wu et al.[15] designed a jet system, which can generate a group of minijets individually operating with periodic velocities. With two hot-wires providing feedback signals, the ON/OFF mode of each minijet is determined by the GP agent. Based on this AFC system, the researchers not only achieved a large performance gain, but also observed new wake patterns formed during the control[16].

3. Active flow control using deep reinforcement learning
3.1 Deep reinforcement learning

The DRL is applied in an interactive manner between the environment and the agent, as depicted in Fig. 4. The main elements are the agent, the environment, and the interactions between them. From the perspective of machine learning, the agent receives full or partial information from the environment, i.e., the states s_t, as input, and generates control signals as output, i.e., the actions a_t. The actions will then alter the states of the environment, and also the control performance in the next step. The performance is evaluated using a specified value, i.e., the reward r_t. Then the reward will be transferred to the agent. Historical series of states, actions and rewards will form a chain and be used for optimization purpose when updating the control policy. From the perspective of AFC, the agent plays a key role in building up connections between the input and output of the active control, and the formation of the closed loop.
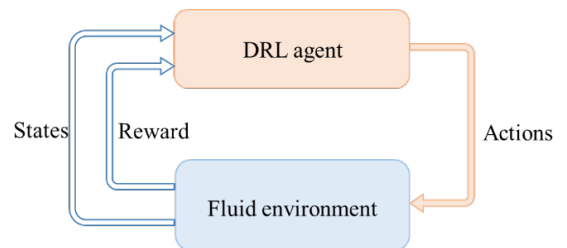

Fig. 4 Schematic diagram of the control loop

Herein we take the state-of-the-art proximal policy optimization (PPO) algorithm as an example. This model has been proposed individually by Heess et al. [17] and Schulman et al. [18] and described in detail therein, so we would briefly introduce the algorithm architecture while focusing on physical problems.

In one episode, the agent runs the policy for T times and collects a sequence of states-actions-reward, i.e.,

$$\tau = (s_1, a_1, r_1), (s_2, a_2, r_2), \ldots, (s_t, a_t, r_t), \ldots, (s_T, a_T, r_T) \quad (1.)$$

The policy, π_Θ, is modeled by ANNs, which can be determined by the weights Θ. As shown in FIG 5, the PPO uses two sets of ANNs: an actor network whose input layer is the states and the output layer is the actions, and a critic network whose input layer is the states and the output layer is an approximation of discounted reward. Here the discounted reward is defined as

$$R(t) = \sum_{t'>t} \gamma^{t'-t} r_{t'} \quad (2.)$$

where γ is the discount factor usually close to 1. Generally, appropriate choices of network width (i.e., the number of neurons in each hidden layer) can improve the training performance, while bad choices will lead to over-fitting or under-fitting problems.
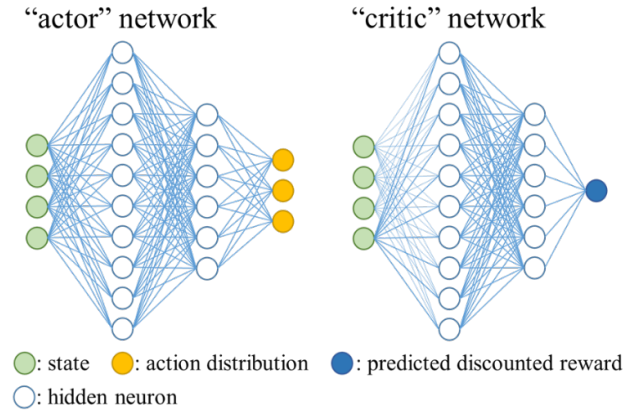

"actor" network   "critic" network

⬤: state   🟠: action distribution   🔵: predicted discounted reward
◯: hidden neuron

Fig. 5 Architectures of the artificial intelligence networks

In order to update the policy π_Θ, the objective function of each set of ANN is essential. First, with the output of the critic network and the long-term discounted reward, one can estimate the advantage as

$$\hat{A}_t = \sum_{t'>t} \gamma^{t'-t} r_{t'} - V_\Theta(s_t) \quad (3.)$$

The objective of the critic network is to minimize the discrepancy between the predicted and real values of the discounted reward, so the following objective function can be applied, i.e.,

$$J_{CRITIC}(\Theta) = -\sum_{t=1}^{T} \hat{A}_t^2 \quad (4.)$$

Next, in order to update the actor network, we follow the work of Schulman et al. [18], where the following clipped surrogate objective is used

$$J_{ACTOR}(\Theta) =$$
$$\hat{E}_t \left[\min\left(\frac{\pi_\Theta(a_t|s_t)}{\pi_{old}(a_t|s_t)} \hat{A}_t, clip\left(\frac{\pi_\Theta(a_t|s_t)}{\pi_{old}(a_t|s_t)}, 1-\varepsilon, 1+\varepsilon \right) \hat{A}_t \right) \right]$$
(5.)

where ε is a hyper-parameter set to be 0.2 as in Schulman et al.[18] The clip term inside the above equation means the ratio between the new policy and the old policy, i.e., π_Θ/π_{old}, is constrained to an interval $[1 - \varepsilon, 1 + \varepsilon]$.

When updating the policy, the conventional stochastic gradient descent optimizer, or the currently popular moment-based optimizers can be adopted, e.g., the adaptive moment estimation (adam) optimizer. To deal with continuous control, the actor network does not directly output the action values, but generates probability distributions, e.g., the Gaussian distribution, so that actions in the following steps can be sampled from these distributions.

3.2 Drag reduction of a blunt body

Rabault et al.[19, 20] applied the DRL to train a two-layer ANNs, targeting at reducing the drag exerting on a circular cylinder while avoiding large lift fluctuations. In this work, the latest policy-based DRL algorithm, i.e., the PPO, is adopted, which is superior in continuous control. In the flow system, two jets located at angular position of 90º and 270º of the cylinder surface are implemented as actuators. The two jets are operated in the one-suction-one-blowing mode, thus ensuring a zero-net mass flux rate into the flow field. 151 velocity probes, located both in the surrounding and in the wake of the cylinder, are used to obtain partial information of the flow system. The Reynolds number is fixed in the laminar flow regime, i.e., Re = 100. Fluid dynamics is simulated using an open-source solver based on the finite element scheme.

The DRL agent interacts with the flow via receiving information from the probes and determining the jet strength. The control performance is then evaluated by considering both the drag and lift using a function, i.e., $r_t = -C_d - 0.2|C_l|$, which is then fed back to the DRL agent. Under the DRL framework, the policy is determined by two sets of ANNs and updated using the stochastic descent method, so as to achieve higher objective values defined by the PPO. Through the training process, it is shown that the DRL agent can achieve a high drag reduction rate after hundreds of episodes.

A follow-up study was reported by Ren et al.[21], where the above DRL control was pushed to the turbulent flow regime. A drag reduction rate of around 30% has been demonstrated for cases at Re = 1000, proving that the DRL is also effective in the turbulent flow regime.

3.3 Maneuvering of biological body

As a typical cross-disciplinary field, maneuvering of biological body has been a hot topic. A thorough study in

this field can deepen our knowledge in nature and help design useful robotic systems that mimic real creatures. With a set of sensors for perceiving surrounding environments and a well-designed reward for evaluating the performance, the DRL agent can play a core role similar as brains or neural systems. In this field, impactful researches includes bird flying[22], fish swimming[23], and gravitaxis of microswimmers[24], just to name a few.

Migratory birds usually adopt an energy-saving strategy when encountering rising atmospheric currents. Reddy et al.[22] adopts the DRL to explore this phenomena, where the atmospheric boundary layer is obtained from a 3D numerical solver by directly solving the Navier-Stokes equation and a thermal convective equation, the glider is modeled using a group of kinematic equations and is assumed not to affect the background flow field. In this study, the state-action-reward-state-action (SARSA) algorithm is used, which models the decision process as a Markov chain, and the optimal policy is found via estimating the Q function. Learned policy reveals different strategies in exploiting the turbulent thermals and avoiding risky situations in regimes of both moderate and strong turbulent flow. This study is believed to technically help autonomous gliders in extending their flying range.

Verma et al.[23] investigate schooling fish, where the follower can harness and exploit energy in the vortex wakes of its companion. This study was realized in a numerical framework, where the flow environment is directly simulated by solving the Navier-Stokes equation, the body undulations are determined by a group of curvature functions whose parameters are determined by the DRL. Results show a surprising fact that smart swimmers will place themselves at off-vortex-center positions. In this study, the deep Q network (DQN) algorithm is adopted to train the swimming policy, and the recurrent neural network is also used so that the agent can "remember" useful historical information to better learn the policy.

Another very interesting study conducted by Colabrese et al[24] is focused on smart gravitactic particles. In this study, the particles are endowed with the ability of obtaining surrounding vorticity information to decide which direction to go in the next step. If uncontrolled, the particles will most likely to be trapped by the underlying vortices. Through trainings using the DRL, the particles can finally develop a good ability to swim to a high altitude, even in a perturbed environment. In this study, the Q learning algorithm is used, whose idea is similar as the SARSA.

It is essential to point out that in all the above studies, the actions are discretized other than continuous, indicating that the number of actions used is limited.

4. Challenges and future prospects

Although having attracted extensive attentions, the machine learning-based AFC still faces huge obstacles before it becomes really feasible. These obstacles include:

(1) Time and hardware costs for the evolution/learning process are huge. For numerical studies, since the whole evolution/training process would involve hundreds or even thousands of cases, it is essential to reduce the time cost for each case, via fast algorithms as well as parallel computing devices. Differently, for experimental studies, machine learning-based AFC requires low uncertainties in the measurements, thus the precision, sampling rate, and time delay characteristics of sensory system will become key issues.

(2) Due to the randomness in machine learning algorithms, the results are sometimes difficult to reproduce, which also brings uncertainties to the analysis. Moreover, choices of machine learning algorithms, neural network structures (only for the DRL), and hyper-parameters still rely on experience.

In spite of these challenges, we would also outline some future prospects:

(1) There are many potential flow control problems that can get solutions from machine learning. In addition to traditional problems like flow past a blunt body, more complicated systems involve FSI, convective flow, chemical reactive flow, etc. For most systems, when the flow becomes turbulent, the strong nonlinearity and high dimensionality bring challenges to the control. As a core issue in fluid mechanics, turbulence control deserves more efforts. Since the DRL has virtually no limit on the nonlinearity and dimensionality of control problems, it might be a good direction to explore.

(2) Latest progress in machine learning also brings us some novel, promising concepts and tools for AFC. For example, the GAN has superior capability in data argumentation and imitation learning, which can imitate most features of a given control strategy and even improve them. A new idea then emerges by combining the GAN with the reinforcement learning, i.e., generative adversarial imitation learning (GAIL)[25], which has recently been applied in complicated control tasks such as autopilot.

5. Summary

In this paper, we briefly reviewed the studies of machine learning-based AFC in the recent years, mainly focusing on the GP or DRL enabled controls. Through most aforementioned studies, machine learning-based AFC has been proven a success, as evaluated from different aspects such as effectiveness, efficiency, and robustness. These proof-of-concept studies have offered solid evidence that machine learning can be a powerful alternative for the AFC, especially when conventional control methods fail in complex flow problems involving strong nonlinearity and high dimensionality. Having said that, however, it should be noted that this field is still undeveloped, with very limited studies being reported. Thus more efforts should be put in this nice area.

References

[1] S.Z. Cai, S.C. Zhou, C. Xu, Q. Gao, Dense motion estimation of particle images via a convolutional neural network [J], Experiments in Fluids, 2019 **60**, 73.

[2] K. Duraisamy, G. Iaccarino, H. Xiao, Turbulence Modeling in the Age of Data [J], Annual Review of Fluid Mechanics, 2019 **51**, 357-377.

[3] J.Q. Huang, H.C. Liu, W.W. Cai, Online in situ prediction of 3-D flame evolution from its history 2-D projections via deep learning [J], Journal of Fluid Mechanics, 2019 **875**.

[4] Q. Wang, J.S. Hesthaven, D. Ray, Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem [J], Journal of Computational Physics, 2019 **384** 289-307.

[5] H. Wu, X. Liu, W. An, S. Chen, H. Lyu, A deep learning approach for efficiently and accurately evaluating the flow field of supercritical airfoils [J], Computers & Fluids, 2020 **198** 104393.

[6] J.R. Koza, Genetic programming: On the programming of computers by means of natural selection [M], MIT Press, 1992.

[7] D.E. Goldberg, J.H. Holland, Genetic algorithms and machine learning [J], Machine learning, 1988 **3** 95-99.

[8] F. Ren, C.L. Wang, H. Tang, Active control of vortex-induced vibration of a circular cylinder using machine learning [J], Phys. Fluids, 2019 **31** 093601.

[9] W.W. Zhang, X.T. Li, Z.Y. Ye, Y.W. Jiang, Mechanism of frequency lock-in in vortex-induced vibrations at low Reynolds numbers [J], Journal of Fluid Mechanics, 2015 **783** 72-102.

[10] N. Gautier, J.L. Aider, T. Duriez, B.R. Noack, M. Segond, M. Abel, Closed-loop separation control using machine learning, Journal of Fluid Mechanics, 2015 **770** 442-457.

[11] A. Debien, K. von Krbek, N. Mazellier, T. Duriez, L. Cordier, B.R. Noack, M.W. Abel, A. Kourta, Closed-loop separation control over a sharp edge ramp using genetic programming [J], Experiments in Fluids, 2016 **57** 40.

[12] V. Parezanovic, J.-C. Laurentie, C. Fourment, J. Delville, J.-P. Bonnet, A. Spohn, T. Duriez, L. Cordier, B.R. Noack, M. Abel, M. Segond, T. Shaqarin, S.L. Brunton, Mixing Layer Manipulation Experiment From Open-Loop Forcing to Closed-Loop Machine Learning Control [J], Flow Turbulence and Combustion, 2015 **94** 155-173.

[13] V. Parezanovic, L. Cordier, A. Spohn, T. Duriez, B.R. Noack, J.P. Bonnet, M. Segond, M. Abel, S.L. Brunton, Frequency selection by feedback control in a turbulent shear flow [J], Journal of Fluid Mechanics, 2016 **797** 247-283.

[14] R.Y. Li, B.R. Noack, L. Cordier, J. Boree, F. Harambat, Drag reduction of a car model by linear genetic programming control [J], Experiments in Fluids, 2017 **58** 103.

[15] Z. Wu, D.W. Fan, Y. Zhou, R.Y. Li, B.R. Noack, Jet mixing optimization using machine learning control [J], Experiments in Fluids, 2018 **59** 131.

[16] D. Fan, Y. Zhou, B.R. Noack, Artificial intelligence control of a turbulent jet [C], 21st Australian Fluid Mechanics Conference, Adelaide, Australia, 2018.

[17] N. Heess, S. Sriram, J. Lemmon, J. Merel, G. Wayne, Y. Tassa, T. Erez, Z. Wang, S. Eslami, M. Riedmiller, Emergence of locomotion behaviours in rich environments [J], arXiv preprint, 2017 1707.02286.

[18] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms [J], arXiv preprint, 2017 1707.06347.

[19] J. Rabault, M. Kuchta, A. Jensen, U. Réglade, N. Cerardi, Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control [J], Journal of Fluid Mechanics, 2019 **865** 281-302.

[20] J. Rabault, A. Kuhnle, Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach [J], Phys. Fluids, 2019 **31** 094105.

[21] F. Ren, H. Tang, Active flow control of flow past a circular cylinder at moderate Reynolds number using deep reinforcement learning [C], HiFiCoMa, Shanghai, China, 2019, 141.

[22] G. Reddy, A. Celani, T.J. Sejnowski, M. Vergassola, Learning to soar in turbulent environments [J], Proceedings of the National Academy of Sciences of the United States of America, 2016 **113** 4877-4884.

[23] S. Verma, G. Novati, P. Koumoutsakos, Efficient collective swimming by harnessing vortices through deep reinforcement learning [J], Proceedings of the National Academy of Sciences of the United States of America, 2018 **115** 5849-5854.

[24] S. Colabrese, K. Gustavsson, A. Celani, L. Biferale, Flow Navigation by Smart Microswimmers via Reinforcement Learning [J], Physical Review Letters, 2017 **118** 158004.

[25] J. Ho, S. Ermon, Generative adversarial imitation learning [J], Advances in neural information processing systems, 2016, 4565-4573.