



Aerodynamic optimization of high-lift devices using a 2D-to-3D optimization method based on deep reinforcement learning and transfer learning

Jiahua Dai^a, Peiqing Liu^{a,*}, Qiulin Qu^a, Ling Li^a, Tongzhi Niu^b

^a School of Aeronautic Science and Engineering, Beihang University, Beijing, 100191, China

^b State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China

ARTICLE INFO

Article history:

Received 2 August 2021

Received in revised form 21 December 2021

Accepted 8 January 2022

Available online 14 January 2022

Communicated by Z. Weiwei

Keywords:

High-lift device

Transfer learning

Deep reinforcement learning

Three-dimensional optimization

ABSTRACT

The computational fluid dynamics is the main method to evaluate the aerodynamic performance for the optimization of high-lift devices. Currently, the direct three-dimensional (3D) optimization requires significant computational resources. Additionally, the commonly used heuristic algorithms can not extract the experience of two-dimensional (2D) optimization to accelerate the 3D optimization process. In order to resolve these issues, a novel 2D-to-3D optimization method based on the coupling of Deep Reinforcement Learning (DRL) and Transfer Learning (TL) is proposed to conduct the aerodynamic optimization of the 3D high-lift devices and tested on the NASA Trap Wing model. The 2D optimization is first carried out, and then its neural networks in DRL and the optimal configuration are extracted by TL to turn into the 3D optimization. Compared with the direct 3D optimization, the proposed 2D-to-3D optimization method can result in improved aerodynamic performance for the same computational cost, or can save 51%–81% of the computational cost to obtain a similar performance.

© 2022 Elsevier Masson SAS. All rights reserved.

1. Introduction

High-lift devices can effectively improve both the performance and safety of aircrafts during take-off and landing phases [1,2]. Due to the complex viscous flow, the wind tunnel experiment and computational fluid dynamics (CFD) based on the Navier-Stokes equations have been frequently applied for accurate prediction of the aerodynamic performance of high-lift devices. Thus the CFD is commonly used to predict the aerodynamics in the optimization of high-lift devices. Because of considerable the computational cost, in the traditional design procedure, the multi-element airfoil is firstly optimized using the heuristic algorithms, such as the Tabu Search (TS) [3] and the Genetic Algorithm (GA) [4–6]. It is then extended to a three-dimensional (3D) layout based on the lifting surface theory [7–11] in which the wing aerodynamic performance can be predicted using one section [9–11] or multiple sections [7,8]. For a typical three-element airfoil with slat and flap, its aerodynamic performance is mainly characterized by six parameters, including Gap, Overlap (O/L) and deflection angle θ_{flap}

as shown in Fig. 1, which are widely chosen as the optimization variables. Minervino et al. [12] and Iannelli and Quagliarella [9] compared the wing section of an optimal 3D high-lift device and the airfoil from a pure 2D optimization for the same spanwise position, and found that the optimal values of the six parameters are completely different. These differences due to the complex 3D viscous flow effects [13], including the local effective angle of attack (AoA) [14,15], the interference between wing and fuselage [16], and so on.

In order to improve the traditional design method, Kanazaki et al. [14] conducted the 2.5-dimensional (2.5D) optimization using GA instead of the 2D approach and found that the spanwise flow will cause an earlier separation which should be considered in the design process. Sutcliffe et al. [17] compared the optimal flap positions of different dimensional models, and it was concluded that the optimal Gap of the 2.5D model is larger than that of the 2D. Bourgault-Côté et al. [18] derived a crossflow equation and proposed a novel 2.5D method. Compared with the 2D RANS method, the result of the new method is closer to the actual 3D flow with a little increasing in computational cost. However, in the 2.5D method, aspect ratio, twist angle and other 3D factors can not be considered.

* Corresponding author.

E-mail address: lpq@buaa.edu.cn (P. Liu).

Nomenclature

θ_{flap}	deflection angle of flap	θ	parameter of online actor neural network
χ	quarter chord sweep angle	ω	parameter of online critic neural network
B	span	V	state value function
M_∞	incoming Mach number	$Loss_{actor}$	loss function of online actor network
Re	Reynolds number	td_target	time difference target
$\theta_{spoiler}$	deflection angle of spoilers	td_error	time difference error
P_{in}	inboard axis point of flap rotation axis	r_t	reward obtained at the step t
P_{out}	outboard axis point of flap rotation axis	θ'	parameter of target actor neural network
S_{flap}	displacement of flap along rotation axis	ω'	parameter of target critic neural network
Cl	lift coefficient	γ	future reward discount
K	lift-to-drag ratio	τ	soft update coefficient of target networks
f	target value	Cl_{2D}	lift coefficient of an airfoil
t	a step in a episode	Cd_{2D}	drag coefficient of an airfoil
s_t	the state of the step t	$Cl_{2D,n}$	lift coefficient of a normalized airfoil
a_t	the action of the step t	$Cd_{2D,n}$	drag coefficient of a normalized airfoil
π	strategy of agent	\bar{f}	averaged target value in an episode
Q	action value function		

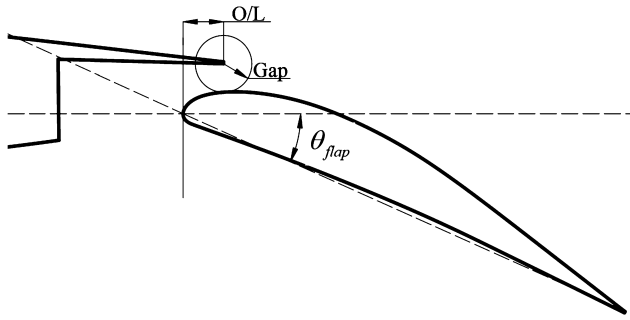


Fig. 1. Parametric variables for the flap of a multi-element airfoil.

In the optimization of 3D configuration, the traditional optimization algorithm (such as GA [19], Particle Swarm Optimization (PSO) [20], and TS [21]) involves a significant CFD computational cost during the initial stage of optimization process. This is due to the random search within a large domain of variables to get enough samples and approach the global optimal. It has been frequently considered to combine these optimization algorithms with a surrogate model which fits the relationship between optimization targets and variables based on the samples to reduce the amount of cost [14,22,23]. However, the surrogate model has a limited effect because it only uses the 3D samples and ignores the experience of 2D optimization. A coupling of Deep Reinforcement Learning (DRL) and Transfer Learning (TL) provides an alternative method to shrink the search domain and obtain the global optimal more efficiently.

DRL makes a computational agent based on neural networks to learn the strategy of making decisions by trial and error, and finally obtain the maximum benefit. This method is superior to the heuristic algorithms implemented in solving some non-deterministic polynomial-time hardness problems [24–26], and can even be used to optimize the algorithm itself [27]. It has been gradually applied in the aerospace. Lampton [28] used Q-learning to optimize a morphing airfoil under constrained thickness and camber conditions. Based on DRL, a single Unmanned Aerial Vehicle (UAV) can be controlled to smoothly carry the load to different target positions [29]. Multiple UAVs can also be commanded to perform reconnaissance and strike tasks accordingly [30]. TL focuses on storing knowledge gained while solving one problem and applying it to a different but related problem by pre-training the neural network [31]. It has been widely applied in speech recog-

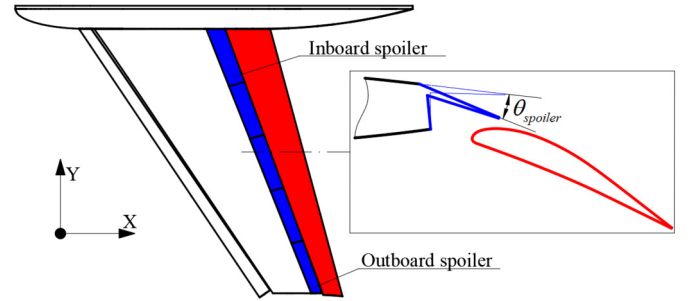


Fig. 2. Schematic representation of the spoilers. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

nition [32] and image recognition [33]. In the optimization of 3D high-lift devices, TL can be used to reduce the computational cost of the high-order method based on the results obtained from the low-order method. Yan et al. [34] coupled DRL and TL to optimize the shape of a missile, where the panel method (low-order method) is used to pre-train the neural network, to reduce the CFD computational cost and enhance the solution accuracy.

In this paper, a coupled method of DRL and TL is used to conduct the aerodynamics optimization of the 3D trailing-edge high-lift devices based on the NASA Trap Wing model, and explore the feasibility of this coupled method in the design of complex high-lift devices.

2. Optimization problem

The NASA Trap Wing model with the quarter chord sweep angle of $\chi = 29.97^\circ$ and the span of $B = 2184.4$ mm from the 1st AIAA CFD High Lift Prediction Workshop [35] is chosen as the basic model. The objective is to optimize the trailing-edge high-lift configuration under the take-off conditions of $M_\infty = 0.2$, $AoA = 8^\circ$, and $Re = 4.3 \times 10^6$. For the trailing-edge devices, the coupled deflection of flap and spoiler (like that adopted in B787 and A350XWB) is regarded as a promising configuration [36,37] to obtain an improved aerodynamic performance. Therefore, five spoilers are designed for the Trap Wing model to conduct the optimization of the configuration as shown in Fig. 2. The origin of the Cartesian coordinate system is located on the model nose. The ratio of spoiler length to local wing section chord length decreases linearly from 15% to 11%, as moving from the wing root to the tip in the spanwise direction. They can deflect the same angle $\theta_{spoiler}$

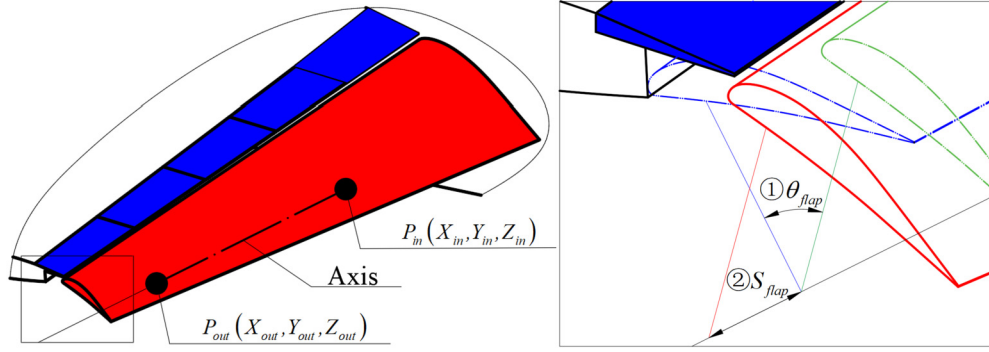


Fig. 3. Parameters describing the spatial position of a flap.

Table 1
Ranges of optimization variables.

Variable	Initial value	Range
$\theta_{spoiler}$	0°	$0^\circ \leq \theta_{spoiler} \leq 6^\circ$
X_{in}	1253.070 mm	$1200 \text{ mm} \leq X_{in} \leq 1263 \text{ mm}$
Z_{in}	-238.130 mm	$-244 \text{ mm} \leq Z_{in} \leq -188 \text{ mm}$
X_{out}	1665.792 mm	$1600 \text{ mm} \leq X_{out} \leq 1675 \text{ mm}$
Z_{out}	-157.871 mm	$-164 \text{ mm} \leq Z_{out} \leq -107 \text{ mm}$
θ_{flap}	26.178°	$22^\circ \leq \theta_{flap} \leq 28^\circ$
S_{flap}	-34.052 mm	$-50 \text{ mm} \leq S_{flap} \leq -25 \text{ mm}$

around their respective rotational axes, which is selected as one of the optimization variables.

According to the principles of kinematics, the complex motion of a 3D flap can be decomposed into rotation around a certain axis and translation along that axis. As illustrated in Fig. 3, the rotational axis passing through the points $P_{in}(X_{in}, Y_{in}, Z_{in})$ and $P_{out}(X_{out}, Y_{out}, Z_{out})$ is defined by four independent variables, since two points are chosen on the planes of 1/4 and 3/4 spanwise positions, i.e., $Y_{in} = -546.1 \text{ mm}$ and $Y_{out} = -1638.3 \text{ mm}$. Origination from the stowed position (the blue dashed curves), the flap first rotates by θ_{flap} to approach the green dashed curves, and then translates by S_{flap} to the deployed position (red solid curves).

There are seven independent variables in the optimization of the coupled deflection of flap and spoilers. Their ranges are mainly determined by the following considerations, which are summarized in Table 1 as well.

a) For general commercial aircrafts, the spoiler deflection angle of is less than 10° and the flap deflection angle is larger than 20° in the take-off configuration [37,38].

b) Numerous related studies [6,17,39] have shown that the optimal Gap value of the take-off configuration is about 2% of the local chord length. Here the limited boundary of the rotational axis cause a Gap range from the smallest value (corresponding to interference of the flap and spoilers) to 4%.

c) The translation range of the flap mainly depends on the interference or clearance between flap and fuselage.

According to the Federal Airworthiness Regulations (FAR) Part 25 [40] and relevant studies [37,41], the take-off process of the transport aircrafts can be divided into three stages, the ground taxiing stage, the first climbing stage and the second climbing stage. In the first two stages, a large lift coefficient Cl is needed to reduce the taxiing and climbing distances; In the second climbing stage, a large lift-to-drag ratio K is required to reach a higher climbing rate. Here, the target value f of optimization combining the lift coefficient and the lift-to-drag ratio can be expressed as Eq. (1).

$$f = 0.7 \times \frac{Cl}{1.64} + 1.3 \times \frac{K}{7.19} \quad (1)$$

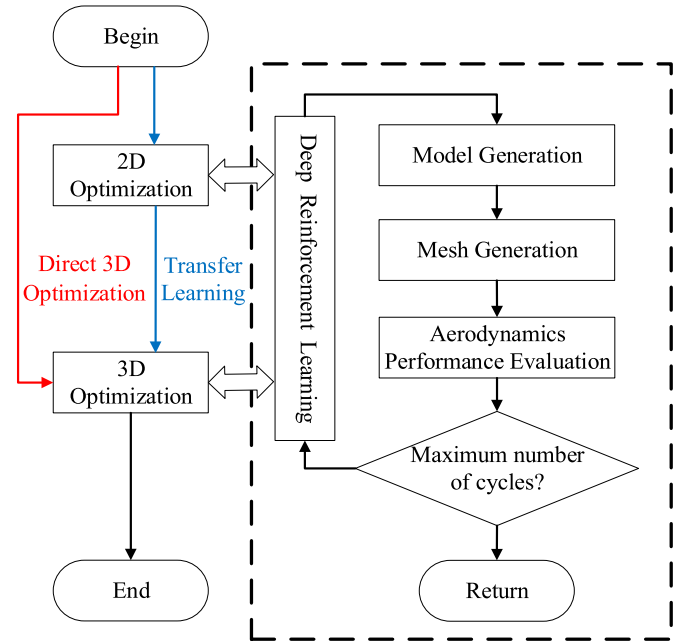


Fig. 4. Optimization process.

The weight coefficients for Cl and K are 1.3 and 0.7 respectively because the time of the third stage is longer, and Cl is included into K as well. Additionally, the scale coefficients 1.64 and 7.19 correspond to Cl and K of the baseline configuration.

3. Optimization approach

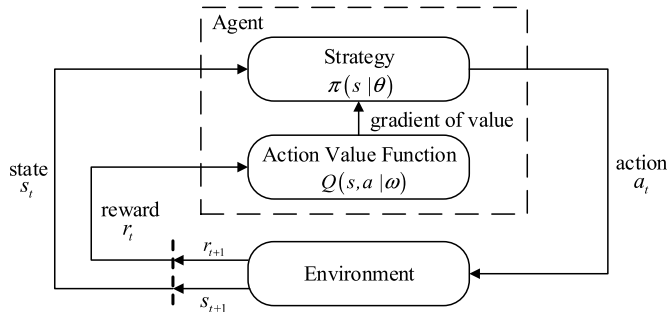
In this paper, a novel 2D-to-3D optimization method based on the coupling of DRL and TL is proposed to solve the above problem, as depicted by the blue arrows in Fig. 4. The 2D optimization is first carried out, and its experience is then extracted by TL to turn into the 3D optimization. Both of the 2D and 3D optimizations implement the same algorithm based on DRL and involve the same workflow including model generation, mesh generation and aerodynamics performance evaluation, as demonstrated in the dashed block in Fig. 4. The only difference between them is the method of aerodynamics performance evaluation. In the 2D optimization, the performance is approximately evaluated at two typical 2D sections of the wing as presented in Subsection 3.3 while in the 3D optimization accurately evaluates the performance based on the complete wing-body configuration.

In order to assess the accuracy and efficiency of the proposed 2D-to-3D optimization method, a direct 3D optimization based on DRL is also conducted and the results are compared, as illustrated by the red arrow in Fig. 4.

Table 2

The optimization pseudo code based on DDPG.

The optimization pseudo code based on DDPG
Define the objective function of f
Set the soft update parameter $\tau = 0.01$
Randomly initialize the online actor network $\pi(s \theta)$ and the online critic network $V(s \omega)$
Initialize the target actor network π' and the target critic network V' with $\theta' = \theta$ and $\omega' = \omega$
Initialize a replay buffer
for episode = 0 to 99
Set the initial state s_0
for $t = 0$ to 7
Set the total training times $M = 8 \times \text{episode} + t$
Set the action variance, $\sigma^2 = \text{Max}(0.15 - 0.00125 \times M, 0.025)$
Select action $a_t = \mathcal{N}(\mu, \sigma^2)$, $\mu = \pi(s_t \theta)$
Set the next state $s_{t+1} = a_t$
Calculate the reward:
If the model interferes or the Gap is too small, $r_t = -10$
Else if there is a negative volume of the fluid mesh, reselect a_t
Else $r_t = 1000(f_{t+1} - f_t)$
Store the experience (s_t, s_{t+1}, r_t) in the replay buffer
Set the learning rates of the neural networks $\alpha = \text{Max}(0.01 - 0.00016 \times M, 0.002)$
Nine experiences (s_i, s_{i+1}, r_i) are selected randomly from the replay buffer:
Calculate $td_target = r_i + \gamma V'(s_{i+1} \omega') = r_i + \gamma V'(\pi'(s_i \theta') \omega')$
Calculate $td_error = td_target - V(s_t \omega)$
Update the online critic network $V(s \omega)$ with:
$Loss_{critic} = \frac{1}{9} \sum_i (td_target - V(s_i \omega))^2$
Update the online actor network $\pi(s \theta)$ with:
$Loss_{actor} = \frac{1}{9} \sum_i td_error \times (s_{i+1} - \pi(s_i \theta))^2$
Update the target networks:
$\theta' \leftarrow \tau \theta + (1 - \tau) \theta'$
$\omega' \leftarrow \tau \omega + (1 - \tau) \omega'$
End for
End for

**Fig. 5.** Process diagram of the A-C algorithm.

3.1. Deep reinforcement learning

The process of the Actor-Critic (A-C) algorithm, commonly used in DRL based on Markov Decision Process (MDP), can be expressed in Fig. 5 from a step t to the next step $t + 1$ where the agent first makes actions a_t only according to the current state s_t , and then adjusts strategy $\pi(s|\theta)$ (described by an artificial neural network named actor) according to the gradient of value provided by action value function $Q(s,a|\omega)$ (described by a neural network named critic) which is influenced by the reward r from environmental feedback. Here θ and ω are the parameters of the neural networks. After several steps the agent will return to the origin,

which is called an episode. The mechanism of returning is similar to the “heating process” in simulated annealing algorithm [42] which is helpful to avoid the agent falling into local optimization. In order to improve the stability and convergence speed of the A-C algorithm, the Depth Deterministic Policy Gradient (DDPG) algorithm [43] was proposed by adding a replay buffer, a twin actor network (online actor network and target actor network) and a twin critic network (online critic network and target critic network).

In this paper, the aerodynamic optimization problem is conducted using an optimization algorithm based on DDPG. The aerodynamic performance of the current configuration can be regarded as the state s_t , and the change of optimization variables depends on the agent can be regarded as the action a_t . The problem meets the requirements of MDP. The pseudo code of the algorithm is shown in Table 2. The optimization process includes 100 episodes and each episode includes 8 steps. In order to solve the above aerodynamic problem appropriately, four modifications are conducted to improve the DDPG algorithm proposed by Lillicrap [43].

Firstly, instead of the action value function $Q(s,a|\omega)$, the state value function $V(s|\omega)$ is considered to evaluate the quality of the agent's strategy at each step. This is because the configuration performance is more concerned than the improvement direction of this configuration. Accordingly, the action a_t output from the actor neural network is directly regarded as the next state s_{t+1} . There-

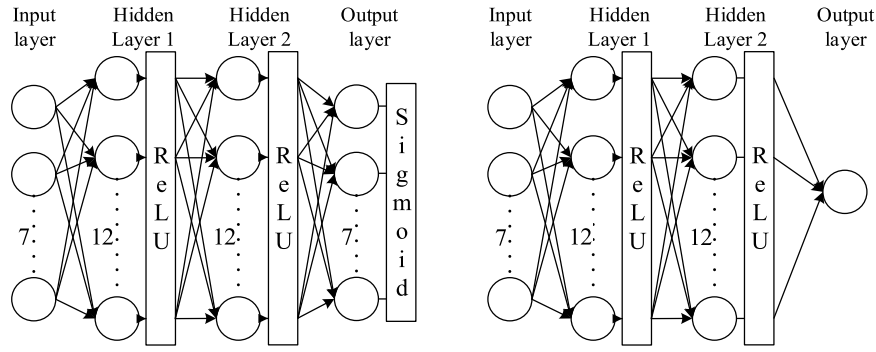


Fig. 6. Structures of the actor (left diagram) and the critic (right diagram) networks.

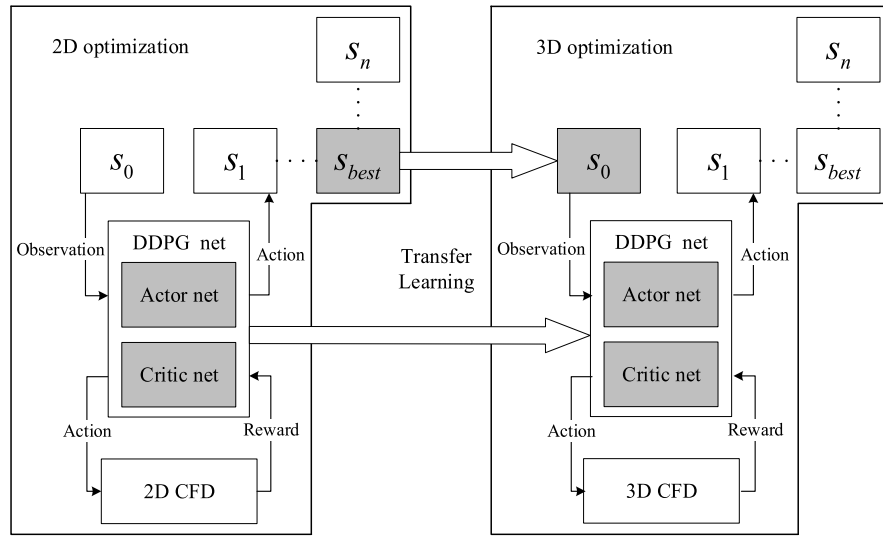


Fig. 7. Transfer learning from 2D optimization to 3D optimization.

fore, the actor network consists of an input layer and an output layer both including 7 cells according to the number of the optimization variables. A Sigmoid activation function is applied after the output layer in order to make the whole network non-linear. This leads to an output in the range of 0 to 1, that normalizes the state s before being used into the actor network. The frameworks of the actor and the critic networks are depicted in Fig. 6.

Secondly, a loss function is proposed for the online actor network ($Loss_{actor}$) which is multiplied by a time difference error td_error and a mean square error. The sign of td_error is positive if the actor network makes decisions that are beneficial to performance improvement, otherwise it is negative. Furthermore the mean square error has a positive value to indicate the extent of performance changes. Through the backward propagation, the actor network learns how to adjust the optimization variables to get a better configuration with higher performance.

Thirdly, the noise and learning rate are set to decrease linearly. In the initial stage of optimization, the large noise can make the agent more likely to explore the unknown domain and therefore avoid falling into the local extremums. Moreover, the large learning rate can enhance the convergence speed of the algorithm in the initial stage. Reduction of the noise and learning rate at the later stages can improve the algorithm stability.

Finally, the reward function r_t is designed as Eq. (2) when all operations (model and mesh generation and numerical simulation) are successfully accomplished in a step.

$$r_t = 1000 (f_{t+1} - f_t), \quad (2)$$

where the coefficient of 1000 is used to amplify the reward value to a suitable range. Two types of operation failures are considered.

- If the model generation fails (the Gap is too small or negative), the configuration is unsatisfactory and the agent should avoid taking this action. In such situation, the reward function will be set to $r_t = -10$ and the neural network will be updated.
- If there is any failure in the mesh generation or the numerical simulation, the network training process will be skipped and the next state will be reselected because the failure is not caused by the poor aerodynamic performance.

3.2. Transfer learning

TL includes the transfers of both the initial state and the initial neural network, as shown in the grey boxes of Fig. 7. The best result in the 2D optimization is selected as the starting point of the 3D optimization. Different from the random initialization of the neural network in the direct 3D optimization, the online networks obtained by the 2D optimization are used as the initial networks of the 3D optimization. TL not only transfers the high-performance samples of 2D, but also transfers the trend of each sample. In other words, not only the high-performance samples accelerate the

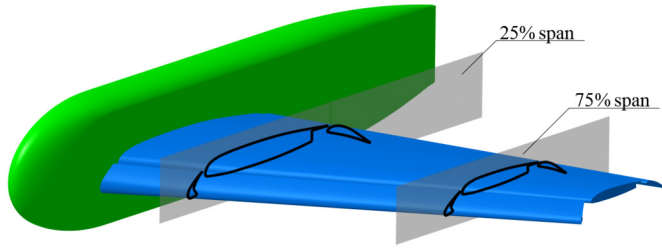


Fig. 8. Typical wing sections in 2D optimization.

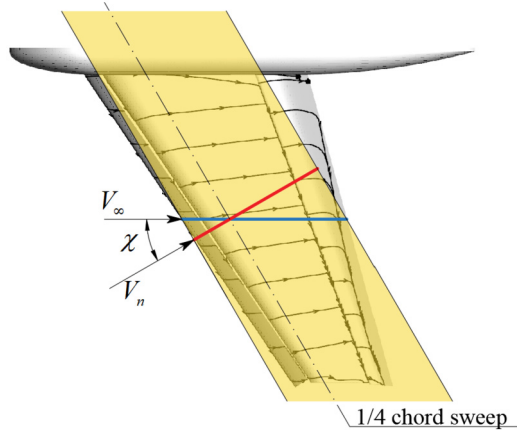


Fig. 9. Normalized airfoil based on the theory of infinite swept wing.

3D optimization, but also the low-performance samples guide the agent to move towards the optimal direction.

3.3. Aerodynamics evaluation model of 2D optimization

In the 2D optimization, the lift coefficient and the lift-to-drag ratio of the wing-body configuration are approximately evaluated by the CFD results of the inboard (25% span) and the outboard (75% span) sections of the wing, as presented in Fig. 8 and Eq. (3). In general, the approximated aerodynamic performance has the same trend as the accurate performance of the wing-body configuration, thus the 2D optimization can be applied in pre-training of the agent to solve the optimization problem preliminarily.

$$Cl = (Cl_{2D}^{in} + Cl_{2D}^{out}) / 2, \quad K = (K_{2D}^{in} + K_{2D}^{out}) / 2 \quad (3)$$

In the traditional theory of an infinite swept wing [11,18], the incoming airflow can be divided into the chordwise branch and the spanwise branches originating from the leading edge. The chordwise branch mainly contributes to the aerodynamic forces. Thus, the aerodynamic coefficients in Eq. (3) can be calculated from the aerodynamic coefficients $Cl_{2D,n}$ and $Cd_{2D,n}$ of the normalized airfoil and the quarter-chord sweep angle χ , as shown in Eq. (4) and Eq. (5).

$$Cl_{2D} = Cl_{2D,n} \cos^2 \chi \quad (4)$$

$$Cd_{2D} = Cd_{2D,n} \cos^3 \chi \quad (5)$$

The CFD results obtained for the present finite swept wing (as shown in Fig. 9) indicates that the limiting streamlines on the wing surface are more likely to be perpendicular to the quarter-chord line. Therefore, the normalized airfoils in the 2D optimization are considered based on the quarter-chord sweep angle χ .

Table 3

Summary of the grids.

	Number of Cells	Y plus	Near-wall growth rate
Extra-coarse	3,419,090	0.88	1.32
Coarse	7,385,091	0.66	1.22
Medium	13,826,416	0.55	1.15
Fine	28,149,935	0.44	1.08

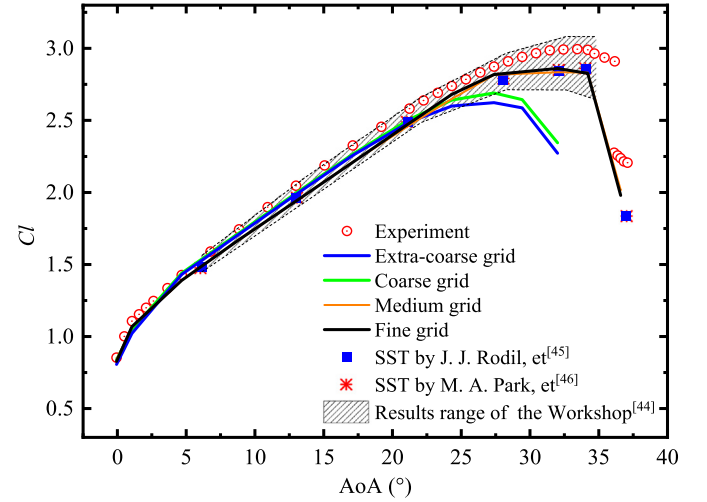


Fig. 10. Comparison of the lift curves obtained by the four grid with different studies.

3.4. Mesh and CFD validation

The commercial software ANSYS ICEM CFD is used to generate the structure grid. A cuboid computational domain is considered for the 3D CFD simulation. The cube dimensions are considered based on the Mean Aerodynamic Chord (MAC) of the wing-body model. More specifically, the downstream side of this cube is 100 times greater than the MAC and the others are 50 times larger than the MAC. The no-slip boundary conditions are set at the walls and a symmetry boundary condition is applied to the central symmetry plane of the domain. The opening boundary condition with specified speed is considered for the other boundaries. Four grids are generated to conduct the grid independence study, which are termed as the extra-coarse, coarse, medium and fine, as presented in Table 3.

The steady RANS equations with the SST $k - \omega$ turbulence model are solved using the commercial software ANSYS CFX based on the finite volume method. A second order upwind "High Resolution" advection scheme is used for the mass, momentum, energy and turbulence equations.

The lift curves obtained for different grids are compared with the results reported by other studies [44–46], as shown in Fig. 10, where the gray shadow region is the envelope of the results from the 1st AIAA CFD High Lift Prediction Workshop. For the range of small to middle AoA, the lift curves predicted by the four sets of grids are completely located inside the margins of the Workshop results and match the experimental data well. For the stall region, only the curves predicted by the medium and fine grids are located in the mentioned margin, while the predicted maximum lift and the stall AoA are smaller than those of the experimental data. The studies [45,46] have shown that the SST $k - \omega$ model result in early prediction of the flow separation, that in turn decreases both the lift and the stall AoA.

In this paper, the take-off case with AoA of 8° is studied. Thus, the coarse grid is used in the optimization process to reduce the

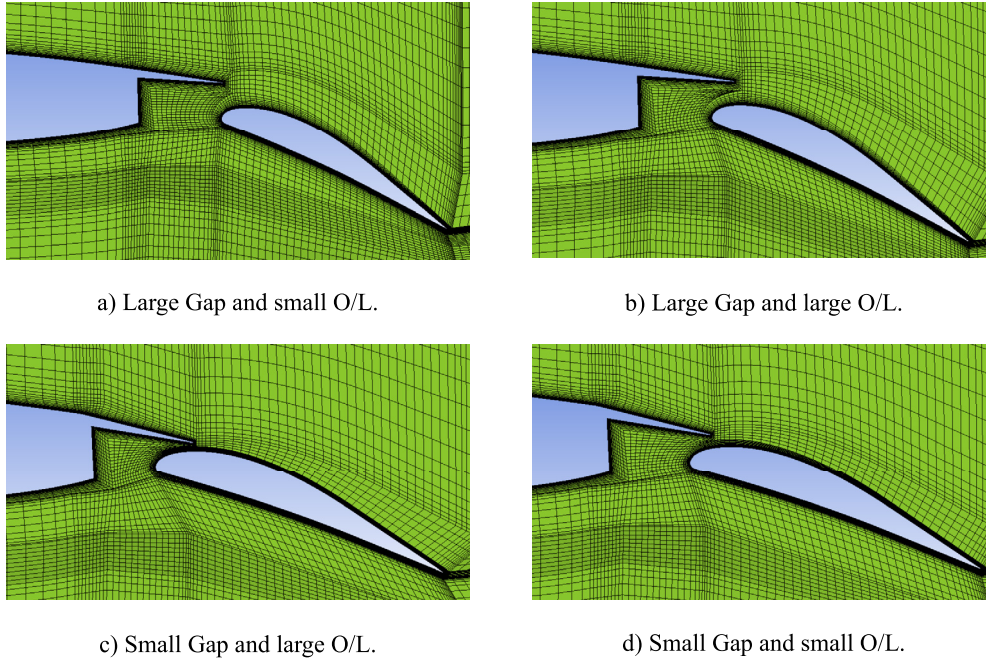


Fig. 11. Mesh layouts for typical amounts of Gap and O/L.

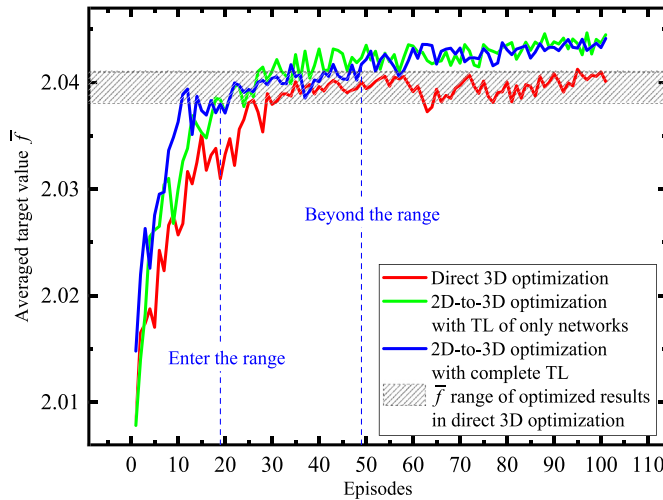


Fig. 12. Comparison of various optimization method with and without TL.

computational cost, and the fine grid is employed for the final assessment of aerodynamic performance.

The 3D mesh quality of the wing-body is critical to the optimization process. A block-topology is generated for the baseline configuration, and the ICM macro [47] is utilized to construct the adaptive mesh. The mesh layouts for typical amounts of Gap and O/L are shown in Fig. 11.

4. Results and discussion

In this section, the optimization performance, computational cost and optimal results of the 2D-to-3D optimization method based on the coupling of DRL and TL are studied in detail.

First, the optimization performance is analyzed. Fig. 12 presents variation of averaged target value \bar{f} (calculated based on 8 steps in each episode) with episode for various optimization methods. Compared with the direct 3D optimization, the 2D-to-3D method leads to an obviously greater target value \bar{f} that is closer to the

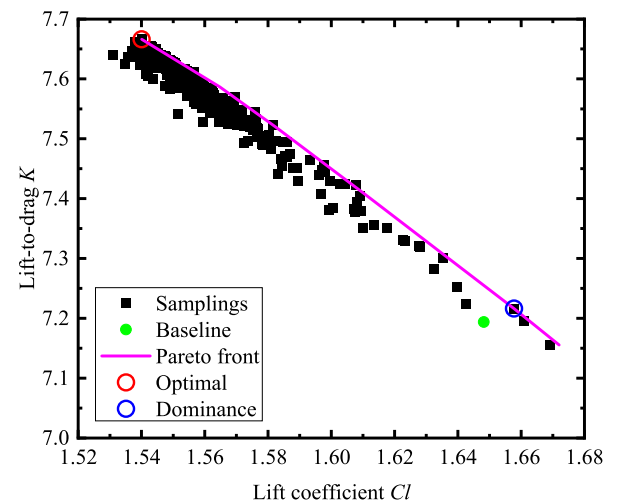


Fig. 13. Variation of the lift-to-drag ratio with the lift coefficient based on the optimization results.

global optimal solution. Furthermore, a rapid convergence rate can be observed in the stage of initial random search.

In order to explore the contributions of initial state transfer and neural network transfer in the positive benefit of the TL method, a 2D-to-3D optimization is performed that only utilizes TL of neural networks. It begins from the same starting point as that of the direct 3D optimization. Compared with the complete TL, Fig. 12 indicates that the TL of just neural networks can lead to almost the same optimal results and a relatively slower convergence rate. That means the TL of neural networks mainly enhances the optimization performance, while the initial state transfer mainly accelerates the convergence rate in the initial stage.

Second, the computational costs of the 2D-to-3D optimization with complete TL and the direct 3D optimization are compared. The 2D optimization is performed using a PC equipped with a 6-core CPU at 3.00 GHz, while the 3D optimization is carried out by a high-performance cluster equipped with 96 CPU cores at 2.30 GHz. Thus the computational cost of the 2D pre-training in the 2D-

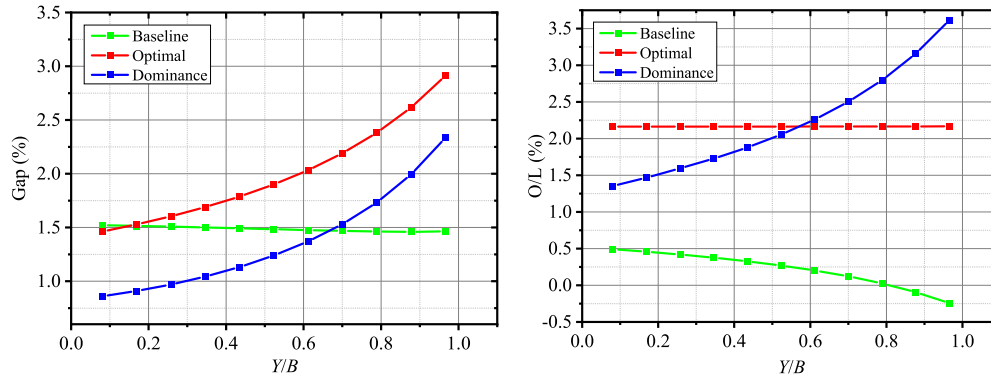


Fig. 14. Spanwise distributions of Gap and O/L for three configurations.

Table 4
Optimization variables and target values of three configurations.

	Baseline	Optimal	Dominance
$\theta_{spoiler}$	0°	0.05°	2.73°
X_{in}	1665.792 mm	1629.066 mm	1628.837 mm
Z_{in}	-157.871 mm	-157.298 mm	-132.443 mm
X_{out}	1253.070 mm	1227.452 mm	1236.707 mm
Z_{out}	-238.130 mm	-238.612 mm	-225.442 mm
θ_{flap}	26.178°	22.000°	24.649°
S_{flap}	-34.052 mm	-40.333 mm	-40.553 mm
Cl	1.648	1.540	1.658
K	7.194	7.666	7.216
f	2.008	2.047	2.016

to-3D optimization can be ignored, and only the cost consumed by the 3D part is compared with direct 3D optimization method. The range of \bar{f} corresponding to the direct 3D optimization results is marked by the gray shadow region in Fig. 12. The averaged target value of the 2D-to-3D optimization enters this range from the 19th episode, and it totally exceeds the range at the 49th episode by an enhanced performance. The results confirm that the 2D-to-3D method can save 51%-81% of the computational cost and still leads to a similar performance as the direct 3D optimization. It should be noted that the separation pattern of the Trap Wing model tested in this paper is simpler than that of an actual wing, and only two sections were used for TL. Complex separation pattern may impair the effect of TL. In the future work, the influence of the number of sections and selection methods for TL is worth studying.

Finally, the Baseline, Optimal (maximum target value f), and Dominance configurations (Cl and K are both larger than the Baseline configuration) of the 2D-to-3D optimization with complete TL are compared and analyzed, as shown in Fig. 13. The contradiction between Cl and K is prominent for the present 3D optimization, such that the agent prefers to achieve the optimal target value by decrease of Cl and increase of K , due to the greater weight of K .

The optimization variables and target values f of the three configurations are presented in Table 4. The spanwise distribution of the Gap and O/L of the three configurations are illustrated in Fig. 14. From the root to the tip, the trends observed for the Gap and O/L in the Optimal and Dominance configurations are completely different with those of the Baseline configuration. More specifically, the Gap of the Baseline configuration almost keeps unchanged as 1.5% of the local chord length, while those of the Optimal and Dominance ones gradually increase (the absolute Gap values tend to keep a constant value). Additionally, the O/L of the Baseline configuration steadily decreases in the spanwise direction, while it keeps a constant value for the Optimal and steadily increase for the Dominance one.

Fig. 15 shows the effect of the Gap and O/L on the flow pattern. A large flow separation region occupies on the upper surface of the flap trailing edge in the Baseline configuration, while the separation region shrinks for the two optimized configurations and only extends from the tip to the middle of the wing. The flow separation control is achieved by suitable coupling of the flap deflection angle, Gap and O/L.

Fig. 16 illustrates variation of the target value f with AoA for the above-mentioned three configurations. In the range of small AoA, the target value f exhibits a great enhancement in the Optimal configuration and a slight improvement in the Dominance one, which indicates satisfactory results due to the optimization process. For the stall AoA range, the Optimal configuration does not show superiority, because the large AoA target is not considered in the optimization process of the present study.

5. Conclusion

In this paper, a novel 2D-to-3D optimization method based on the coupling of DRL and TL is proposed to conduct the aerodynamic optimization of the 3D trailing-edge high-lift devices based on the NASA Trap Wing model. In the 2D pre-training, the aerodynamic performance of the 3D high-lift devices is approximately evaluated by the typical 2D wing sections. In the 3D optimization, the performance is accurately evaluated by the complete wing-body configuration. In the proposed method, the 2D optimization is first carried out, and its experience is subsequently extracted by TL to turn into the 3D optimization; both of them use the same optimization algorithm based on DRL.

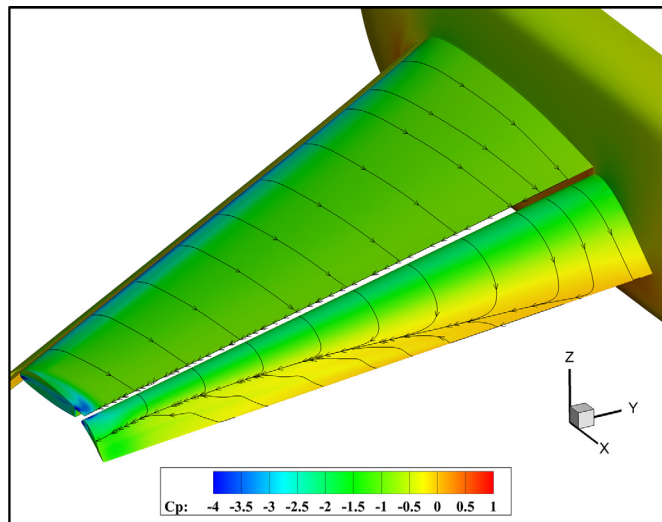
Compared with the direct 3D optimization, the proposed 2D-to-3D optimization can obtain an improved aerodynamic performance for the same computational cost, or can save 51%-81% of the computational cost to achieve the same performance. These benefits are originated from the two transfers introduced by TL: the transfer of the neural networks in DRL mainly improves the optimal configuration performance; the transfer of the optimal configuration in 2D pre-training mainly accelerates the convergence rate.

Declaration of competing interest

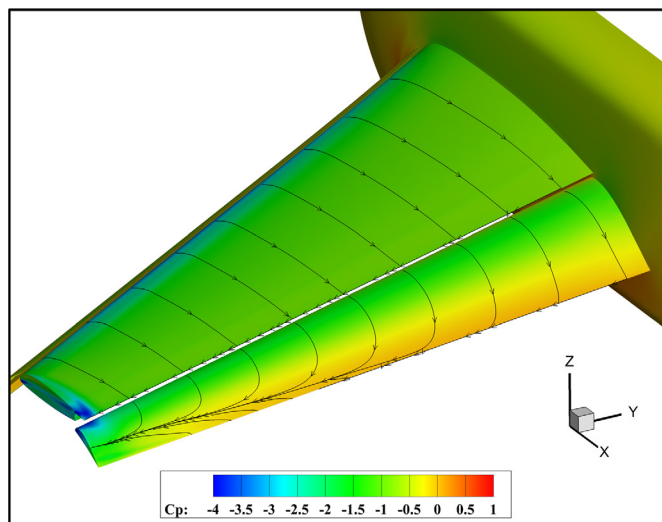
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

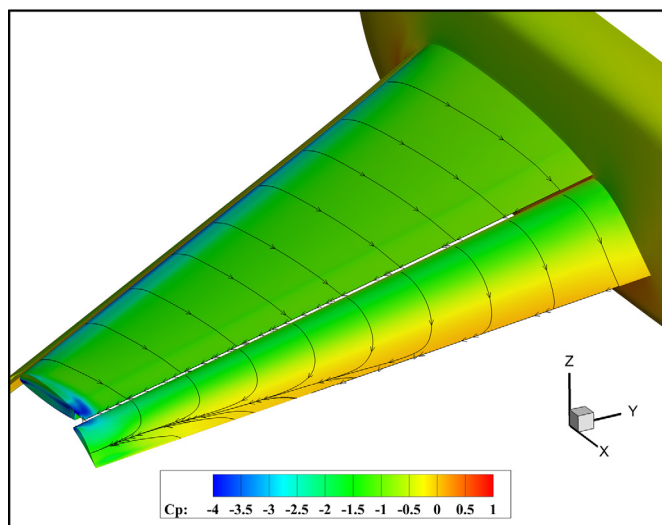
This work was partially supported by the National Natural Science Foundation of China (Grant Nos. 12072014, 12002021).



a) The Baseline configuration.



b) The Optimal configuration.



c) The Dominance configuration.

Fig. 15. Pressure contour and surface streamlines of three configurations.

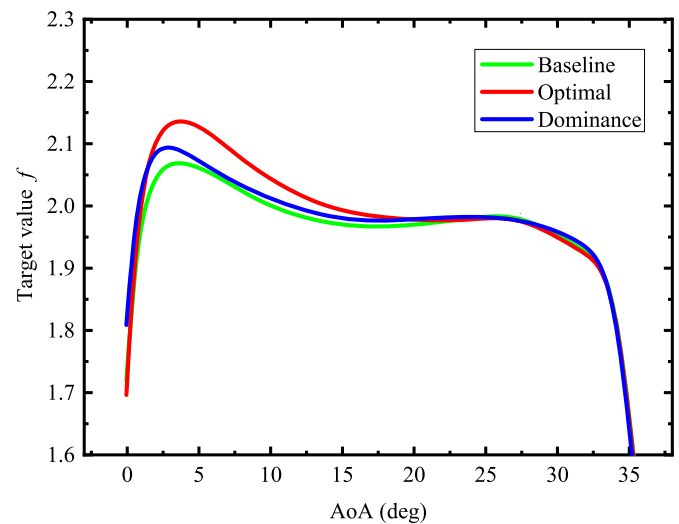


Fig. 16. Variation of the target value with AoA for the three configurations.

References

- [1] D. Reckzeh, Aerodynamic design of the high-lift-wing for a megaliner aircraft, *Aerosp. Sci. Technol.* 7 (2) (2003) 107–119, [https://doi.org/10.1016/S1270-9638\(02\)00002-0](https://doi.org/10.1016/S1270-9638(02)00002-0).
- [2] J. Tao, X.Y. Wang, G. Sun, Stall characteristics analyses and stall lift robustness inverse design for high-lift devices of a wide-body commercial aircraft, *Aerosp. Sci. Technol.* 111 (2021) 106570, <https://doi.org/10.1016/j.ast.2021.106570>.
- [3] G. Trapani, A.M. Savill, T. Kipouros, Multi-objective optimization of 2D high-lift airfoil configurations using tabu search, in: *V Pegasus-AIAA Student Conference*, 2010.
- [4] Y. Tian, J.C. Quan, et al., Mechanism/structure/aerodynamic multidisciplinary optimization of flexible high-lift devices for transport aircraft, *Aerosp. Sci. Technol.* 93 (2019) 104813, <https://doi.org/10.1016/j.ast.2018.09.045>.
- [5] J.F. Wang, Y.Z. Wu, J. Periaux, Combinatorial optimization using genetic algorithms and game theory for high lift configuration in aerodynamics, in: *41st Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, 2003.
- [6] M. Kanazaki, K. Tanaka, et al., Multi-objective aerodynamic exploration of elements' setting for high-lift airfoil using kriging model, *J. Aircr.* 44 (3) (2007) 858–864, <https://doi.org/10.2322/jjass.54.419>.
- [7] J. Weissinger, The lift distribution of swept-back wings, *Technical Report Archive & Image Library*, 1947.
- [8] R.S. Pepper, C.P. Dam, P.A. Gelhausen, Design methodology for high-lift systems on subsonic transport aircraft, in: *Symposium on Multidisciplinary Analysis & Optimization*, 1996.
- [9] P. Iannelli, D. Quagliarella, Multi-objective/multi-point shape and setting high-lift system optimization by means of genetic algorithm and 2D Navier-Stokes equations, in: *Evolutionary and Deterministic Methods for Design, Optimization and Control*, Capua, Italy, 2011.
- [10] J. Wild, J. Brezillon, et al., Advanced design by numerical methods and wind tunnel verification within European high-lift program, *J. Aircr.* 46 (1) (2009) 157–167, <https://doi.org/10.2514/1.37148>.
- [11] J. Wild, J. Brezillon, et al., Realistic high-lift design of transport aircraft by applying numerical optimization, in: *European Conference on Computational Fluid Dynamics*, 2006.
- [12] M. Minervino, P. Iannelli, D. Quagliarella, 3D flap design using Navier-Stokes equations and evolutionary optimization techniques on an industrial platform, in: *Evolutionary and Deterministic Methods for Design, Optimization and Control*, Capua, Italy, 2011.
- [13] C.L. Rumsey, S.X. Ying, Prediction of high lift: review of present CFD capability, *Prog. Aerosp. Sci.* 38 (2002) 145–180, [https://doi.org/10.1016/S0376-0421\(02\)00003-9](https://doi.org/10.1016/S0376-0421(02)00003-9).
- [14] M. Kanazaki, T. Imamura, et al., High-lift wing design in consideration of sweep angle effect using kriging model, in: *46th AIAA Aerospace Sciences Meeting and Exhibit*, 2008.
- [15] D.L. Mathias, K.R. Roth, et al., Navier-Stokes analysis of the flow about a flap edge, *J. Aircr.* 35 (6) (1998) 833–838, <https://doi.org/10.2514/2.2384>.
- [16] C.L. Rumsey, E.M. Lee-Rausch, R.D. Watson, Three-dimensional effects in multi-element high lift computations, *Comput. Fluids* 32 (5) (2003) 631–657, <https://doi.org/10.2514/6.2002-845>.
- [17] M. Sutcliffe, D. Reckzeh, M. Fischer, Hicon aerodynamics - high lift aerodynamic design for the future, in: *25th Congress of the International Council of the Aeronautical Sciences*, Bremen, 2006.

- [18] S. Bourgault-Côté, S. Ghasemi, et al., Extension of a two-dimensional Navier-Stokes solver for infinite swept flow, *AIAA J.* 55 (2) (2017) 662–667, <https://doi.org/10.2514/1.j055139>.
- [19] J. Holland, *Adaptation in Natural and Artificial Systems: an Introductory Analysis with Application to Biology, Control & Artificial Intelligence*, 1975.
- [20] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: *MHS95 Sixth International Symposium on Micro Machine & Human Science*, IEEE, 2002.
- [21] F. Glover, Future paths for integer programming and links to artificial intelligence, *Comput. Oper. Res.* 13 (5) (1986) 533–549.
- [22] X.S. Du, P. He, J. Martins, Rapid airfoil design optimization via neural network-based parameterization and surrogate modeling, *Aerosp. Sci. Technol.* 113 (2021) 106701, <https://doi.org/10.1016/j.ast.2021.106701>.
- [23] E. Iuliano, Global optimization of benchmark aerodynamic cases using physics-based surrogate models, *Aerosp. Sci. Technol.* 67 (2017) 273–286, <https://doi.org/10.1016/j.ast.2017.04.013>.
- [24] H.Y. Hu, X.D. Zhang, et al., Solving a new 3D bin packing problem with deep reinforcement learning method, *arXiv preprint*, arXiv:1708.05930, 2017.
- [25] I. Bello, H. Pham, et al., Neural combinatorial optimization with reinforcement learning, *arXiv preprint*, arXiv:1611.09940, 2016.
- [26] H.J. Dai, E.B. Khalil, et al., Learning combinatorial optimization algorithms over graphs, *arXiv preprint*, arXiv:1704.01665, 2018.
- [27] K. Li, J. Malik, Learning to optimize, *arXiv preprint*, arXiv:1606.01885, 2016.
- [28] A. Lampton, A. Niksch, J. Valasek, Reinforcement learning of morphing airfoils with aerodynamic and structural effects, *J. Aerosp. Comput. Inf. Commun.* 6 (1) (2015) 30–50, <https://doi.org/10.2514/1.35793>.
- [29] X.X. Li, J.L. Zhang, J.D. Han, Trajectory planning of load transportation with multi-quadrotors based on reinforcement learning algorithm, *Aerosp. Sci. Technol.* 116 (2021) 106887, <https://doi.org/10.1016/j.ast.2021.106887>.
- [30] X.Y. Zhao, Q. Zong, et al., Fast task allocation for heterogeneous unmanned aerial vehicles through reinforcement learning, *Aerosp. Sci. Technol.* 92 (2019) 588–594, <https://doi.org/10.1016/j.ast.2019.06.024>.
- [31] S.J. Pan, Y. Qiang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2010) 1345–1359, <https://doi.org/10.1109/TKDE.2009.191>.
- [32] G.E. Dahl, D. Yu, et al., Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition, *IEEE Trans. Audio Speech Lang. Process.* 20 (1) (2011) 30–42, <https://doi.org/10.1109/TASL.2011.2134090>.
- [33] W.W. Cai, J.Q. Huang, et al., Volumetric reconstruction for combustion diagnostics via transfer learning and semi-supervised learning with limited labels, *Aerosp. Sci. Technol.* 110 (2021) 106487, <https://doi.org/10.1016/j.ast.2020.106487>.
- [34] X.H. Yan, J.H. Zhu, et al., Aerodynamic shape optimization using a novel optimizer based on machine learning techniques, *Aerosp. Sci. Technol.* 86 (2019) 826–835, <https://doi.org/10.1016/j.ast.2019.02.003>.
- [35] J.P. Slotnick, J.A. Hannon, M. Chaffin, Overview of the first AIAA CFD high lift prediction workshop, in: *49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, January 2011, Orlando, Florida, 2011.
- [36] P. Scholz, S.S. Mahmood, et al., Design of active flow control at a drooped spoiler configuration, in: *31st AIAA Applied Aerodynamics Conference*, June 2013, San Diego, 2013.
- [37] W.H. Wang, P.Q. Liu, et al., Numerical study of the aerodynamic characteristics of high-lift droop nose with the deflection of Fowler flap and spoiler, *Aerosp. Sci. Technol.* 48 (2016) 75–85, <https://doi.org/10.1016/j.ast.2015.10.024>.
- [38] X.L. Wang, F.X. Wang, Y.L. Li, Aerodynamic characteristics of high-lift devices with downward deflection of spoiler, *J. Aircr.* 48 (2) (2011) 730–735, <https://doi.org/10.2514/1.C031301>.
- [39] C.P. Dam, S.G. Shaw, et al., Aero-mechanical design of high-lift systems, *Aircr. Eng. Aerosp. Technol.* 71 (5) (1999) 436–443, <https://doi.org/10.1108/00022669910296873>.
- [40] PART 25: Airworthiness Standards: Transport Category Airplanes.
- [41] P.K.C. Rudolph, High-lift systems on commercial subsonic airliners, *NASA-CR-4746*, 1996.
- [42] S. Kirkpatrick, C.D. Gelatt, et al., Optimization by simulated annealing, *Science* 220 (4598) (2015) 671–680.
- [43] T.P. Lillicrap, J.J. Hunt, et al., Continuous control with deep reinforcement learning, *Comput. Sci.* (2015).
- [44] C.L. Rumsey, J.P. Slotnick, et al., Summary of the first AIAA CFD high-lift prediction workshop, *J. Aircr.* 48 (6) (2012) 387–478, <https://doi.org/10.2514/6.2011-939>.
- [45] J.J. Rodio, X. Xiao, et al., NASA trapezoidal-wing simulation using stress- ω , and one- and two-equation turbulence models, *J. Aircr.* 52 (4) (2015) 1189–1200.
- [46] M.A. Park, E.M. Lee-Rausch, C.L. Rumsey, FUN3D and CFL3D computations for the first high lift prediction workshop, in: *49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, January 2011, Orlando, Florida, 2011.
- [47] ANSYS Inc., ANSYS ICEM CFD Programmer's Guide Release 18.2, 2017.