aerospace

MDPI

Article

CNN-LSTM Assisted Multi-Objective Aerodynamic Optimization Method for Low-Reynolds-Number Micro-UAV Airfoils

Jinzhao Peng [1], Enying Li [1,*] and Hu Wang [2,*]

[1] School of Mechanical and Intelligent Manufacturing, Central South University of Forestry and Technology, Changsha 410004, China; pengjinle123@163.com
[2] Shenzhen Automotive Research Institute, Beijing Institute of Technology, Shenzhen 518000, China
* Correspondence: enyasteven@hotmail.com (E.L.); wanghu@szari.ac.cn (H.W.)

Abstract

The optimization of low-Reynolds-number airfoils for micro unmanned aerial vehicles (UAVs) is challenging due to strong geometric nonlinearities, tight endurance requirements, and the need to maintain performance across multiple operating conditions. Classical surrogate-assisted optimization (SAO) methods combined with genetic algorithms become increasingly expensive and less reliable when class–shape transformation (CST)-based geometries are coupled with several flight conditions. Although deep learning surrogates have higher expressive power, their use in this context is often limited by insufficient local feature extraction, weak adaptation to changes in operating conditions, and a lack of robustness analysis. In this study, we construct a task-specific convolutional neural network–long short-term memory (CNN–LSTM) surrogate that jointly predicts the power factor, lift, and drag coefficients at three representative operating conditions (cruise, forward flight, and maneuver) for the same CST-parameterized airfoil and integrate it into an Non-dominated Sorting Genetic Algorithm II (NSGA-II)-based three-objective optimization framework. The CNN encoder captures local geometric sensitivities, while the LSTM aggregates dependencies across operating conditions, forming a compact encoder–aggregator tailored to low-Re micro-UAV design. Trained on a computational fluid dynamics (CFD) dataset from a validated SD7032-based pipeline, the proposed surrogate achieves substantially lower prediction errors than several fully connected and single-condition baselines and maintains more favorable error distributions on CST-family parameter-range extrapolation samples ($\pm$40%, geometry-valid) under the same CFD setup, while being about three orders of magnitude faster than conventional CFD during inference. When embedded in NSGA-II under thickness and pitching-moment constraints, the surrogate enables efficient exploration of the design space and yields an optimized airfoil that simultaneously improves power factor, reduces drag, and increases lift compared with the baseline SD7032. This work therefore contributes a three-condition surrogate–optimizer workflow and physically interpretable low-Re micro-UAV design insights, rather than introducing a new generic learning or optimization algorithm.

Keywords: low-Reynolds-number; airfoil optimization; deep learning; neural networks

Check for updates

Academic Editor: Bosko Rasuo

Received: 27 November 2025
Revised: 31 December 2025
Accepted: 8 January 2026
Published: 11 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Aerodynamic optimization is a cornerstone of wing design and involves inherent complexities that require careful treatment [1,2]. With the emergence of new aviation appli-

cations such as unmanned aerial vehicles (UAVs), the focus of aerodynamic design research has progressively shifted toward addressing the specific challenges of small-scale aircraft [3]. In UAV applications, the relationship between performance requirements and geometric configuration is highly nonlinear, owing to their distinct operational environments and mission profiles [4,5]. Recent advances have shown that systematic aerodynamic optimization can effectively coordinate multiple design variables, objective functions, and constraints, thereby significantly improving the aerodynamic performance of UAV airfoils [6–8].

This shift in application focus has introduced new aerodynamic challenges, particularly for micro unmanned aerial vehicle (micro-UAV) development. Current aerodynamic optimization of micro-UAVs primarily concentrates on refining airfoil shapes under low-Reynolds-number conditions (Reynolds number on the order of 10^4 to 10^5) [9,10]. Although decades of conventional aircraft research have led to substantial progress in high-Reynolds-number airfoil optimization, the aerodynamic phenomena governing micro-UAV operation are fundamentally different and therefore require dedicated investigation [11,12]. The rapid expansion of micro-UAV applications has further increased the demand for dedicated low-Re airfoil studies. Unlike full-scale aircraft, where optimization typically targets maximum lift-to-drag ratios, micro-UAV airfoil design often prioritizes power factor (*PI*) enhancement, which is a critical parameter determining endurance and mission effectiveness in small-scale aerial systems [13]. Representative examples include Chen et al. [14], who optimized power factor using an improved Hicks–Henne parameterization, and Li et al. [15], who combined computational fluid dynamics (CFD) with wind-tunnel measurements to support low-Re optimization and highlighted the practical relevance of power factor for micro-UAV endurance. Abbasi et al. [16] incorporated turbulence-related uncertainties into genetic-algorithm-based optimization to balance lift, drag, and stability considerations. Jiang et al. [17] improved computational efficiency via an enhanced particle swarm optimization (PSO) strategy with selective regeneration, enabling simultaneous treatment of power-factor and pitching-moment requirements with fewer iterations.

For practical applications, surrogate-assisted optimization (SAO) methods are among the most popular tools. Tim et al. [18] optimized low-Reynolds-number airfoils under multiple Reynolds numbers and angles of attack using Kriging surrogate models. Rao et al. [19] developed a multi-level hierarchical Kriging (MHK) model, integrating multi-resolution data to form variable-fidelity surrogates. Wauters et al. [20] employed multi-fidelity collaborative Kriging models that combine coarse-grid, low-accuracy data with high-accuracy unsteady Reynolds-averaged Navier–Stokes (URANS) simulations, achieving Pareto optimization of cascade aerodynamic performance under controllable computational costs. In Mars's extremely low-density regime ($Re < 10^5$), Park et al. [21] coupled the γ–Re_θ transition model, Kriging, and a multi-objective genetic algorithm (MOGA) to improve efficiency and reduce power. However, as the design space becomes higher-dimensional and the response surfaces more nonlinear—especially when class–shape transformation (CST) geometry is coupled with multiple operating conditions—conventional surrogates struggle to maintain accuracy and cross-condition consistency [22,23], which is critical for autonomous micro-UAV applications. Nevertheless, when high-dimensional CST designs are optimized against several mission-representative conditions simultaneously, building a surrogate that remains reliable for multi-condition decision-making without training fully separate models remains nontrivial. These limitations have motivated increasing interest in deep-learning-based surrogates that can learn high-dimensional nonlinear mappings more flexibly.

In parallel, deep learning has progressed rapidly in fluid mechanics [24,25] and aerodynamic optimization [26,27], offering stronger nonlinear feature learning than traditional methods [28] and enabling surrogate replacements for expensive CFD evaluations.

Bakar et al. [29] built a Convolutional Neural Network (CNN) surrogate paired with Non-dominated Sorting Genetic Algorithm (NSGA-II) for efficient multi-objective design. Du et al. [30] developed a rapid interactive design framework combining Generative Adversarial Networks (GANs), Multi-Layer Perceptrons (MLPs), and Recurrent Neural Networks (RNNs), which can quickly predict airfoil aerodynamic characteristics and optimal aerodynamic shapes. Li et al. [31] proposed a deep-learning-driven, geometry-aware parameterization with physical filtering to handle transition-induced discontinuities and support robust multi-point optimization at low-Re. Wu et al. [32] proposed an aerodynamic optimization method combining deep neural networks and gradient optimization, greatly improving the efficiency and accuracy of aerodynamic performance optimization by predicting aerodynamic force gradients. Vinothkumar et al. [33] trained MLPs on finite-volume CFD–generated flow data across angles of attack. Sekar et al. [34] used deep CNNs with inviscid panel-method–derived pressure distributions (single-AoA setting, >1300 airfoils) and demonstrated strong generalization to unseen cases. Hasegawa et al. [7] further combined a CNN autoencoder with an long short-term memory (LSTM), achieving compact flow-field representations while maintaining statistical consistency and enabling generalization to unseen Reynolds-number regimes. Overall, these studies highlight the promise of deep surrogates; yet for micro-UAV airfoil optimization, the practical question is how to incorporate multiple operating conditions into a single design decision context without relying on fully separate condition-wise predictors. Here, "tri-condition surrogate" emphasizes this coupled optimization use case, rather than merely being a conventional multi-output CNN.

Low-Re micro-UAV airfoil design is inherently a multi-condition problem, where the same geometry must remain competitive across representative mission states. This coupling challenges conventional SAO, as independently trained surrogates or condition-wise regressors may compromise cross-condition consistency and inflate the CFD budget when high-dimensional CST variables are optimized under multiple operating conditions. To address this gap, we formulate a joint tri-condition surrogate that couples a compact geometry encoder with an operating-condition aggregator and integrate it into a constrained three-objective optimization workflow for micro-UAV airfoils.

Contributions and originality. While this study builds on established convolutional neural network–long short-term memory (CNN–LSTM) and NSGA-II algorithms, we aim to contribute a practical and reproducible surrogate–optimizer workflow for low-Re micro-UAV airfoils through task-specific formulation, coupling, and evaluation:

(1) Tri-condition, nine-output surrogate formulation. We cast the prediction task as a joint mapping from CST-based geometry to {C_D(drag coefficient), C_L lift coefficient, PI} at three representative operating conditions—cruise, forward flight, and maneuver—for the same airfoil. A compact CNN encoder for geometry, combined with an LSTM aggregator over operating conditions, improves cross-condition consistency and extrapolation robustness compared with Single-CNN, Multi-CNN, and LSTM-only baselines.

(2) Dedicated low-Re CFD–CST pipeline tailored to micro-UAV airfoils. We construct a reusable low-Re dataset and workflow by linking validated CFD settings (mesh adequacy and turbulence-model scope) with Latin-hypercube sampling of SD7032-based CST geometries and further connect the optimization outcomes to interpretable aerodynamic and geometric evidence (e.g., thickness/camber distributions, pressure coefficient (C_P) curves, flow fields, and polar curves). This pipeline provides consistent labels for all geometries and operating conditions and is documented with mesh adequacy, turbulence-model scope, and coefficient-level validation.

(3) Constrained three-objective optimization with robustness analysis. We embed the CNN–LSTM surrogate into an NSGA-II framework that simultaneously maximizes the cruise-condition power factor PI_1, minimizes forward-flight drag C_{D2}, and maximizes maneuver-condition lift C_{L3}, under thickness, internal-area, and pitching-moment constraints. Beyond a single optimization run, we quantify optimization stability through repeated runs, Pareto-front overlays, and statistics of the objective distributions.

Overall, this work emphasizes a problem-oriented and empirically validated workflow and low-Re design insights, rather than proposing a new generic learning or evolutionary algorithm.

2. Methodology

2.1. Numerical Computation Method

The flow-field data used to build the deep learning dataset were generated using CFD simulations. The integral form of the two-dimensional Reynolds-Averaged Navier–Stokes (RANS) equations are:

$$\frac{\partial}{\partial t}\int_{\Omega} W \mathrm{d}\Omega + \oint_{\partial\Omega} (F_C - F_V)\mathrm{d}S = 0 \tag{1}$$

where W represents conservative variables; F_c and F_v denote inviscid flux and viscous flux, respectively, in the form of:

$$W = \begin{bmatrix} \rho & \rho u & \rho v & \rho E \end{bmatrix}^{\mathrm{T}} \tag{2}$$

$$F_{\mathrm{C}} = \begin{bmatrix} \rho V \\ \rho u V + n_x p \\ \rho v V + n_y p \\ \rho H V + p \end{bmatrix}, \quad F_{\mathrm{V}} = \begin{bmatrix} 0 \\ n_x \tau_{xx} + n_y \tau_{xy} \\ n_x \tau_{yx} + n_y \tau_{yy} \\ n_x \Theta_x + n_y \Theta_y \end{bmatrix} \tag{3}$$

where ρ represents density, u and v are velocity components in the x and y directions, E denotes total energy per unit mass, V is the absolute velocity of the fluid, n_x and n_y are the unit normal vector components of the control volume surface, H is the total enthalpy per unit mass, τ_{ij} is the viscous stress tensor, and Θ_i represents the combined terms of viscous stress work and fluid heat conduction.

The governing fluid equations were numerically solved using the finite volume method, with inviscid fluxes computed via the Roe scheme and viscous fluxes discretized using a central-difference method. The Spalart–Allmaras (SA) model was adopted for turbulence simulation. As a fully turbulent RANS closure, SA is used here without an explicit transition model. For low-Re airfoils, laminar–turbulent transition and laminar separation bubbles can noticeably affect profile drag C_D and thus the derived power factor $PI = C_L^{3/2}/C_D$. Accordingly, the CFD coefficients used as learning labels—and the surrogate's C_D/PI predictions and optimization outcomes—should be interpreted as model-consistent results within this fixed SA–RANS setup, rather than as transition-resolved "ground truth" in regimes dominated by transition or near-stall separation. A no-slip boundary condition was applied on the airfoil surface, while non-reflecting boundary conditions were imposed at the far field.

Model-form sensitivity and scope. In addition to boundary-condition specification and mesh-convergence checks, we briefly assess the evidence on turbulence/transition-model sensitivity at low Reynolds numbers in Appendix A.1.2 ("Literature Selection and Harmonization"; see Table A2), based on Ref. [35]. A consistent observation is that the drag coefficient C_D (and hence PI) is generally more sensitive to transition modeling than the lift coefficient C_L, particularly as the angle of attack α approaches stall. Accordingly, we adopt the SA model to ensure label consistency for large-sample data generation, and we explicitly

restrict the scope of the reported conclusions to the present SA–RANS configuration and operating envelope.

2.2. CST Parameterization

Parameterization is a key step in airfoil optimization. In this study, we use the CST method [36] to parameterize the airfoil geometry. The CST method is a widely used aerodynamic shape-parameterization technique that reduces the number of design variables while ensuring a smooth airfoil surface, and it has been successfully applied in many aerodynamic optimization problems. In CST, the airfoil geometry is approximated by polynomial functions, with the general expression given by:

$$\begin{cases} \xi_{upper}(\psi) = C(\psi) \cdot s_v(\psi) + \psi \cdot \xi_{teu} \\ \xi_{lower}(\psi) = C(\psi) \cdot S_1(\psi) + \psi \cdot \xi_{tel} \end{cases} \tag{4}$$

The class function $C(\psi)$ and shape function $S(\psi)$ are represented by the following formulas:

$$C(\psi) = (\psi)^{N_1}[1-\psi]^{N_2} \tag{5}$$

$$S_{upper}(\psi) = \sum_{i=0}^{n} A_{upper}(i) \frac{n!}{i!(n-i)!} \psi^i (1-\psi)^{n-i} \tag{6}$$

$$S_{lower}(\psi) = \sum_{j=0}^{m} A_{lower}(j) \frac{m!}{j!(n-j)!} \psi^j (1-\psi)^{m-j} \tag{7}$$

where S_{upper}, S_{lower} represent the upper and lower surfaces of the airfoil, respectively; $C(\psi)$ determines the geometric type of the airfoil, with N_1 set to 0.5 in the first term to characterize the airfoil's round nose, and N_2 set to 1 in the second term to characterize the airfoil's sharp tail; $\psi \cdot \xi_t$ controls the thickness of the trailing edge; $S_{upper}(\psi)$, $S_{lower}(\psi)$ shape functions are decomposed using nth-order Bernstein polynomials. The first term of the shape function defines the leading edge radius, the last term defines the trailing edge angle and thickness, and the middle terms are used to modify the outer contour to obtain a smoother aerodynamic shape; $A_{upper}(i)$ and $A_{lower}(j)$ are coefficients of the shape function and serve as the airfoil shape parameters to be optimized. A fifth-order CST is adopted, with a total of 12 optimization variables.

Figure 1 demonstrates the effectiveness of the CST parameterization method in fitting the SD7032 airfoil, showing favorable fitting results. Table 1 summarizes the fitting error data under different Bernstein polynomial order (BPO) orders. As shown in Table 1, Root Mean Square Error (RMSE) of the upper and lower airfoils exhibit significant reduction as BPO increases from 3 to 5. At BPO = 5, the RMSE values reach 1.436×10^{-3} for the upper airfoil and 1.564×10^{-3} for the lower airfoil. Furthermore, Figure 2 displays the absolute error distributions of both airfoil surfaces, revealing Maximum Absolute Errors (MAEs) below 0.006 for the upper airfoil and 0.0075 for the lower airfoil at BPO = 5. Based on these geometric-fit indicators, BPO = 5 is identified as the most accurate candidate order; its final adoption is confirmed by the aerodynamic invariance analysis presented below (Figure 3).

Table 1. CST fitting errors for SD7032 airfoil of different Bernstein polynomial orders.

BPO	CST Parameters	RMSE of CST Fitting of Upper Airfoil	RMSE of CST Fitting of Lower Airfoil
3	4	0.00407167	0.00241547
4	5	0.00245478	0.00208179
5	6	0.00143571	0.00156391

Abbreviations: CST, class-shape transformation; BPO, Bernstein polynomial order; RMSE, root mean square error.

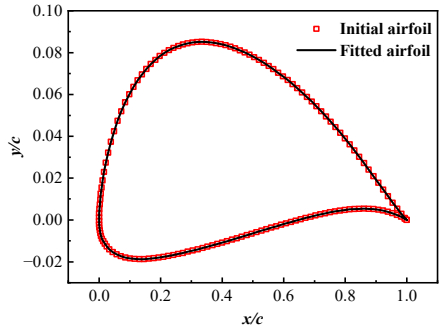


Figure 1. Comparison between original airfoil and fitted airfoil configuration.

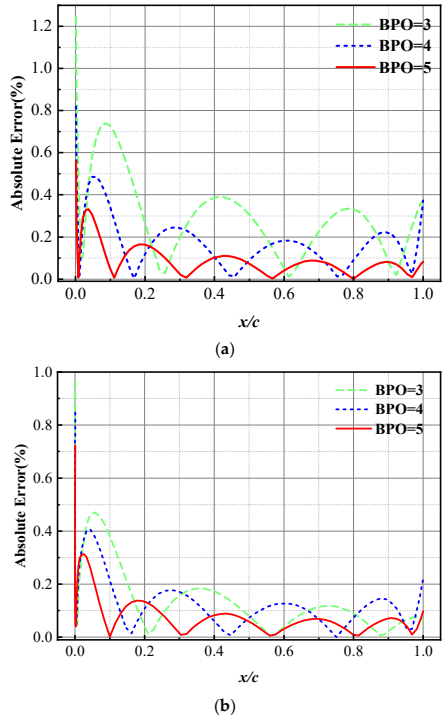


Figure 2. Fitting error analysis of SD7032 airfoil. (**a**) Upper airfoil; (**b**) Lower airfoil.

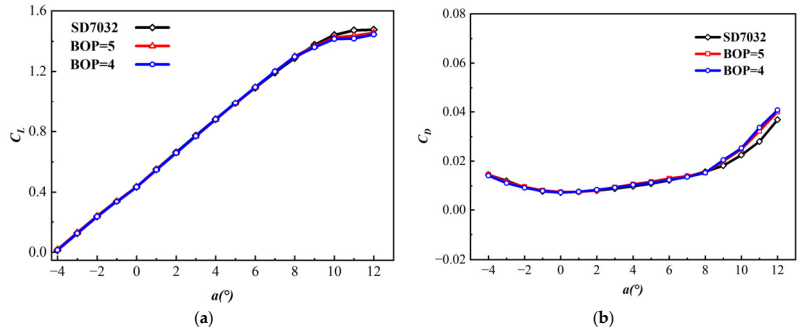


Figure 3. Original (black) vs. CST reconstructions (red: BPO = 5, blue: BPO = 4). (**a**) Lift coefficient (C_L); (**b**) Drag coefficient (C_D).

Order selection and aerodynamic invariance check. To complement the geometric-fit indicators, we verify that the CST reconstruction preserves coefficient-level behavior of the original SD7032. Figure 3 overlays C_L-α and C_D-α for the original shape and its CST reconstructions with BPO = 4 and BPO = 5 over $\alpha \in [-4^\circ, 12^\circ]$. In our operational envelope $\alpha \in [2^\circ, 8^\circ]$, the BPO = 5 curves are visually indistinguishable from the original and remain well within the acceptance bands ($| \Delta C_L | \leq 2\%$, $| \Delta C_D | \leq 5\%$ for $\alpha \leq 10^\circ$). Minor

high-α deviations near stall lie outside this envelope and therefore do not affect subsequent analyses; BPO = 4 shows slightly larger divergence only in that region. Taken together with the preceding geometric-fit results, these overlays corroborate and finalize the choice of BPO = 5 as the minimal order that preserves coefficient-level invariance for the remainder of this study.

Scope of applicability. The surrogate and optimization in this paper are defined on a fixed CST design space with BPO = 5 (12 variables); changing the CST order changes the input dimension and the underlying design space, and therefore would require re-parameterization, re-labeling, and re-training rather than direct transfer. Moreover, the reported robustness/generalization is confined to within-family CST parameter interpolation/extrapolation under the same CFD setup (including the $\pm 40\%$ parameter-range stress test in Section 3.2.3 with geometry-valid screening), and we do not claim validated applicability to non-CST parameterizations or markedly different low-Re airfoil families (e.g., highly cambered/thin shapes) without rebuilding the dataset.

2.3. CFD Validation

To verify the reliability of the numerical simulation method used in this study, numerical validation was performed for the NACA0012 airfoil based on Ladson's low-speed wind-tunnel experimental data [37]. This benchmark case is used primarily for numerical verification (solver implementation/discretization and boundary-condition treatment) and should not be interpreted as a direct validation of low-Re micro-UAV aerodynamics. The computational grid is shown in Figure 4, with the airfoil chord length c set to 1 m. The computational domain extends 25 c in the upper, lower, and upstream directions, and 25 c in the downstream direction. A C–H-type two-dimensional structured grid was employed.

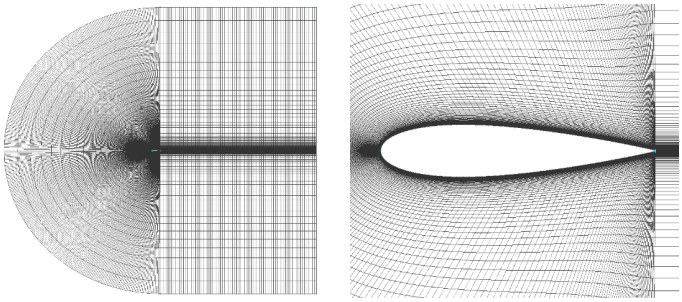


Figure 4. Computational mesh. (**a**) Overall mesh; (**b**) Local mesh.

Several angles of attack were examined under freestream conditions with a Mach number $Ma = 0.30$ and a Reynolds number $Re = 6 \times 10^6$. Figure 5 compares the numerical results with the reference data. The lift coefficient increases approximately linearly with angle of attack, and the CFD predictions show good agreement with the reference results. In contrast, the drag coefficient exhibits a nonlinear trend, with a relatively large error at an angle of attack of 15°, possibly due to stall. The errors in lift and drag coefficients are less than 3.3% and 11%, respectively, which are acceptable for this type of CFD simulation. Therefore, the present CFD setup is numerically verified and provides consistent aerodynamic-coefficient labels for training and evaluating the deep learning models under a fixed CFD configuration. However, because low-Re airfoil aerodynamics is highly sensitive to transition and laminar separation, the agreement obtained here at $Re = 6 \times 10^6$ should not be directly extrapolated to low-Re regimes—particularly for C_D and derived metrics such as PI. Accordingly, the low-Re CFD labels and surrogate/optimization results in this work are interpreted as model-consistent within the CFD setup used throughout the study (see Appendix A.1 for related scope and limitations).

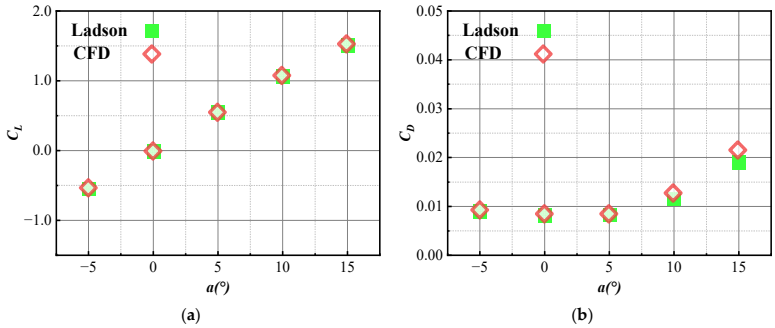


Figure 5. Comparison between numerical results and wind tunnel experiments. (**a**) Lift coefficient; (**b**) Drag coefficient.Abbreviations: CFD, computational fluid dynamics.

To quantify the agreement with the Ladson-corrected NACA0012 data on the same set of angles of attack, we report RMSE, the Coefficient of Determination (R^2) (see Equation (12)), and a range-normalized error nRMSE for both lift and drag. Let $\{y_i^{\mathrm{CFD}}, y_i^{\mathrm{EXP}}\}_{i=1}^{N}$ denote paired coefficients (C_L or C_D) over N angles of attack. We compute

$$\mathrm{RMSE}(y) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i^{\mathrm{CFD}} - y_i^{\mathrm{EXP}}\right)^2}, \tag{8}$$

And the range-normalized error

$$\mathrm{nRMSE}(y) = \frac{\mathrm{RMSE}(y)}{max_i y_i^{\mathrm{EXP}} - mim_i y_i^{\mathrm{EXP}}} \times 100\% \tag{9}$$

The metrics for the grid-selected case indicate near-unity agreement for C_L and acceptable error for C_D (Table 2), consistent with the visual overlays.

Table 2. NACA0012 Validation: Agreement Metrics (RMSE, R^2, nRMSE).

Quantity	RMSE (%)	R^2 (%)	nRMSE (%)
C_L	0.7624	99.811	0.53
C_D	0.0512	91.065	5.59

Abbreviations: RMSE, root mean square error; R^2, coefficient of determination; nRMSE, normalized root mean square error.

To ensure the accuracy and reliability of the computational results, grid independence was assessed through numerical simulations of the NACA0012 airfoil flow field. Three computational grids with varying densities were utilized: coarse (8000 meshes), medium (12,800 meshes), and fine (20,800 meshes). These simulations were conducted under the conditions of $Ma = 0.15$, $Re = 6 \times 10^6$, and $\alpha = 10.18°$, with the first layer grid height of 1.5×10^{-5} and $y^+ < 1$, as detailed in the study. Figure 6 compares the computed and experimental surface pressure coefficients for the NACA0012 airfoil under the three mesh densities. Here, c denotes the airfoil chord length and C_p is the surface pressure coefficient. As shown in Figure 6, the computational results agree well with the experimental data.

Table 3 presents a detailed comparison between computational results on the coarse, medium, and fine grids and the corresponding experimental values. The analysis shows that the lift coefficients computed on all three grids exhibit only small discrepancies relative to the experimental data, consistent with previous studies [38,39]. However, for the drag coefficient, the results from the medium and fine grids are closer to the experimental values, whereas the coarse grid exhibits larger deviations. Considering both computational

efficiency and accuracy, the medium-grid parameters were selected for generating the airfoil mesh in this study.

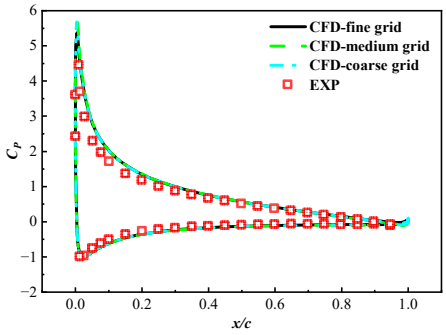


Figure 6. Surface pressure coefficient distribution. Abbreviations: EXP, experiment.

Table 3. Comparison of numerical computation results with experimental values under different grid sizes.

Item	C_L	ΔC_L/%	C_D	ΔC_D/%
EXP	1.081	/	0.0117	/
Coarse grid	1.099	1.66	0.0137	17.09
Medium grid	1.089	0.74	0.0124	5.98
Fine grid	1.083	0.185	0.0119	1.71
very-fine	1.084	0.28	0.0121	3.42

Mesh adequacy (engineering criteria). In addition to the coarse and medium grids, we assessed a finer 42,678-cell mesh on the same C–H domain and kept the first-layer height and growth ratio identical to maintain comparable y^+. The two-finest-grid difference is small—ΔC_L = 0.092% (1.083→1.084) and ΔC_D = 1.68% (0.0119→0.0121). Against the wind-tunnel reference (C_L = 1.081, C_D = 0.0117), the 20.8 k mesh yields 0.185% (C_L) and 1.71% (C_D) errors, and the 42.7 k mesh yields 0.28% and 3.42%, respectively. Representative Cp overlays show negligible sensitivity of surface-pressure distributions within the medium–fine range. It is also worth noting that prior published work has reproduced NACA0012 experimental C_P using a ~9.3 k-cell two-dimensional structured mesh, indicating that O(10^4) cells are adequate when the far-field extents and near-wall spacing are properly set [32]. On this basis we adopt 20.8 k as a cost-effective, mesh-adequate resolution for dataset generation, with 42.7 k serving as an additional confirmation. Given the observed mesh adequacy and agreement with reference data, this configuration is applied consistently to generate labels for all geometries and operating conditions in our dataset. This choice is intended to ensure numerical consistency and mesh adequacy for label generation, rather than to claim experimental-level validation for low-Re transition-dominated flows.

2.4. CNN-LSTM Architecture

2.4.1. Convolutional Neural Network (CNN)

CNN is a neural-network architecture characterized by local connectivity and weight sharing. As illustrated in Figure 7, CNNs typically stack convolution and pooling operations to progressively extract compact latent features, and then use a lightweight regression head to map features to target quantities. In this study, the CNN is employed as a geometry encoder: it takes the CST-parameterized design-variable vector as input and learns local correlations among neighboring CST coefficients along the airfoil profile, producing a compact geometry feature representation for downstream prediction. With this geometry encoding, the subsequent sequence module aggregates how the same geometry responds under multiple operating conditions (described next).

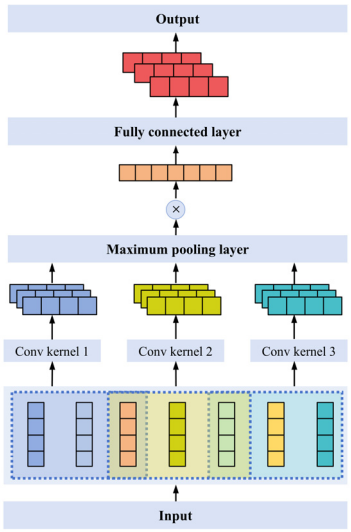


Figure 7. Convolutional neural network (CNN) Architecture.

2.4.2. Long Short-Term Memory (LSTM)

LSTM is a gated recurrent network that aggregates information along an ordered sequence through a controlled memory pathway (Figure 8). In this work, it is not used for temporal forecasting; instead, the sequence index corresponds to operating conditions for the same geometry, so that the model can learn inter-condition dependencies under a unified prediction objective.

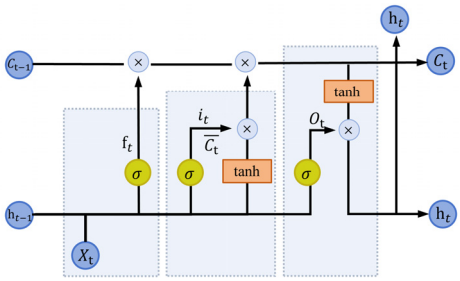


Figure 8. Long short-term memory (LSTM) Model Schematic Diagram.

For clarity of Figure 8, X_t denotes the input at step t, h_{t-1} and c_{t-1} are the previous hidden and cell states, and $\sigma(\cdot)$, $\tanh(\cdot)$, and $\otimes$ denote the sigmoid, hyperbolic tangent, and element-wise product, respectively.

In the proposed hybrid surrogate, the LSTM therefore acts as a compact condition-aggregator, enabling feature-level fusion of geometry and operating conditions before regressing the aerodynamic coefficients across all conditions.

2.4.3. CNN-LSTM Model

To establish an efficient nonlinear mapping between low-Reynolds-number airfoil geometries and their aerodynamic coefficients, we adopt a hybrid deep neural architecture that couples convolutional neural networks (CNNs) with a lightweight LSTM aggregator. The CNN-LSTM paradigm was originally introduced for video recognition to aggregate frame-wise representations [40]. In this study, however, the LSTM is repurposed to capture inter-condition dependencies—i.e., the relationships among the three operating conditions (cruise, forward flight, and maneuver; combinations of *Re* and α) for the same geometry—rather than explicit temporal evolution. Similar CNN-LSTM variants have been used in fluid mechanics and sequential modeling, demonstrating strong generalization and engineering adaptability (e.g., Hasegawa et al. [7]).

Within this framework, we tailor the input organization and network hierarchy to the aerodynamic prediction task with CST-based shapes and multi-condition parameters. Specifically, the CST design variables are first encoded by the CNN into a shared geometry representation. The operating conditions are then organized as a short ordered sequence of three steps, following a fixed mission-representative order (cruise → forward flight → maneuver) as specified by the three design points in Section 4.2 (Design Requirements for Multiple Operating Conditions). Each step is described by a scaled/normalized operating-condition tuple <*Re*, *α*> under the same preprocessing convention used throughout the dataset and model training. At each step, the condition tuple is fused with the shared geometry representation at the feature level, and the LSTM aggregates the sequence to model cross-condition coupling in a unified manner. This design enables an efficient mapping from design variables to the targets C_D, C_L, and PI across the three operating conditions and produces the nine coefficients $[C_D, C_L, PI] \times 3$ in a single forward pass.

The proposed network supports end-to-end training and can be seamlessly integrated into surrogate-assisted multi-objective optimization to reduce the reliance on expensive CFD evaluations. Architectural and implementation details are provided in Section 3.1.1.

2.5. Optimization Framework

The aerodynamic optimization framework developed in this study is illustrated in Figure 9. The entire airfoil aerodynamic optimization process includes:

1. Parameterizing the airfoil;
2. Inputting design variables into the Deep Predictive Neural Network (pre-trained and frozen; only queried to evaluate objectives) to predict the airfoil's aerodynamic coefficients;
3. Updating design variables using the optimization algorithm;
4. Checking the optimization termination criteria.

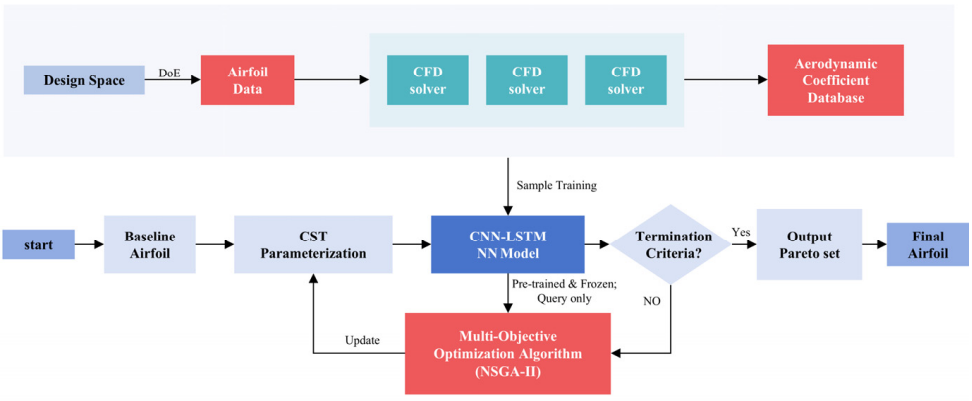


Figure 9. Optimization architecture based on deep learning.Abbreviations: DoE, design of experiments; NSGA-II, non-dominated sorting genetic algorithm II.

If the criteria are met, the Pareto-optimal set of geometries is output and the final design is selected; if not, the process returns to steps 2 and 3 until completion. This section optimizes the airfoil geometry (CST parameters); the surrogate is never updated during optimization and is only queried for objective evaluation.

3. CNN–LSTM Surrogate Modeling and Validation

3.1. CNN-LSTM Modeling

In low-Reynolds-number airfoil aerodynamic prediction, a complex nonlinear coupling exists between local geometric sensitivities (e.g., leading-edge curvature and trailing-edge thickness) and the dynamic responses under multiple flight conditions (e.g., varying

Reynolds numbers and angles of attack). While traditional fully connected neural networks (FCNNs) possess global nonlinear fitting capabilities, they struggle to capture the localized spatial features embedded in airfoil geometries. Conversely, standalone LSTM models, which are effective for sequential data, are ill-suited for structured geometric inputs that lack explicit temporal dependencies.

To address these challenges, this study employs a hybrid CNN-LSTM architecture designed to simultaneously fulfill the dual requirements of geometric feature extraction and flight-condition-aware dynamic modeling. The CNN module is used to encode CST-based geometry into a compact feature representation that captures local correlations among design variables associated with geometric variations (e.g., near the leading and trailing edges), while the LSTM module aggregates inter-condition dependencies across varying flight conditions. These two components work in synergy to construct a coupled surrogate model for multi-condition aerodynamic-coefficient prediction under low-Reynolds-number regimes. The detailed architecture of the CNN-LSTM model is illustrated in Figure 10.

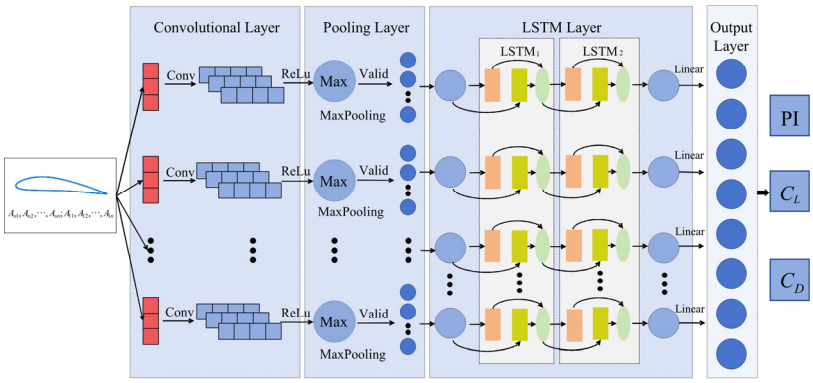


Figure 10. Architecture of the CNN-LSTM model, a hybrid neural network that effectively combines the strengths of CNNs for feature extraction and LSTM networks for capturing long-term dependencies, making it ideal for tasks involving multi-condition responses for the same geometry. Abbreviations: *PI*, power index; ReLU, rectified linear unit.

3.1.1. Model Architecture

(1) Local Sensitivity Modeling with CNN

We employ a 1D convolutional encoder as a compact feature extractor for CST-parameterized design variables. In this study, the input to the CNN is a 12-dimensional vector of airfoil design variables parameterized by the CST method. To clarify the meaning of "locality" under this representation, the CST variables (Bernstein coefficients) are arranged in a fixed, geometry-informed order: coefficients on each surface follow the chordwise direction from leading edge to trailing edge, and the upper- and lower-surface groups are concatenated consistently. Under this convention, neighboring coefficients correspond to adjacent Bernstein basis contributions with strongly overlapping chordwise influence, so short-range 1D convolutions capture coupled sensitivities among neighboring geometric control parameters. More generally, one-dimensional convolution (Conv1D) is also used as a parameter-efficient alternative to fully connected layers, leveraging weight sharing and a limited receptive field to reduce parameters and improve generalization for low-dimensional inputs.

A series of one-dimensional convolutional layers (kernel size = 3) is applied to capture local correlations associated with geometric sensitivities across the airfoil profile: with small kernels capturing short-range correlations among neighboring CST coefficients. As the encoder depth increases, the effective receptive field expands, enabling the network to represent broader interactions across coefficient groups. We therefore describe the CNN as

a hierarchical feature encoder, without assigning specific geometric semantics to particular depths in the absence of dedicated attribution evidence.

The CNN consists of three convolutional layers with increasing channel dimensions: 32, 64, and 128, respectively. Each convolutional layer is followed by a max-pooling operation to progressively condense and refine high-dimensional feature representations. The final output is a 128-dimensional feature vector that serves as input to the subsequent LSTM module. This hierarchical feature encoding facilitates the modeling of nonlinear relationships among geometric parameters in a compact and efficient manner.

(2) Dynamic Response Modeling with LSTM

LSTM resolves long-term dependencies through gating (input/forget/output) but here is not used for temporal prediction; instead, it aggregates dependencies across operating conditions for the same geometry. The CNN first encodes the CST variables into a 128-dimensional geometry vector. We define an ordered three-step condition sequence with a fixed order: cruise ($t = 1$) $\rightarrow$ forward flight ($t = 2$) $\rightarrow$ maneuver ($t = 3$), consistent with the three design points listed in Section 4.2 . At each step t, the operating condition is encoded as a scaled tuple $\left\langle \widetilde{Re}_t, \widetilde{\alpha}_t \right\rangle$. Because the operating conditions are predefined (Section 4.2), we apply linear scaling using the corresponding ranges, i.e., $Re_{\min}$ and $Re_{\max}$ are taken from the minimum and maximum Reynolds numbers across the three design points, and $\alpha_{\min}$ and $\alpha_{\max}$ are taken analogously. We then concatenate the shared geometry feature vector with the condition tuple to form the step input $x_t = \left[g; \widetilde{Re}_t; \widetilde{\alpha}_t \right]$, and feed $\{x_1, x_2, x_3\}$ into a two-layer LSTM.

The decoder then outputs nine coefficients $[C_D, C_L, PI] \times 3$ conditions in one pass. This feature-level fusion targets multi-condition consistency under coefficient-level supervision. Since LSTMs are order-aware, the condition order must be used consistently during training and inference. In our workflow, the fixed order in Section 4.2 is used throughout, and the optimization loop interprets the outputs according to the same indexing (e.g., PI_1, C_{D2}, C_{L3}); changing the order without retraining would change the semantics of the output indices and may introduce a mismatch. Quantitative performance comparisons under the fixed evaluation protocol are reported in Section 3.2.

(3) Engineering Advantages of the CNN-LSTM Architecture

The integration of CNN and LSTM leverages the complementary strengths of spatial feature extraction and sequential modeling, yielding a hybrid surrogate model tailored to the challenges of low-Reynolds-number airfoil prediction. Specifically, the CNN module efficiently identifies localized geometric sensitivities, reducing reliance on redundant inputs, while the LSTM module aggregates cross-condition dependencies (sequence-over-conditions) across flight regimes rather than temporal dynamics.

This synergistic combination aims to balance predictive accuracy with computational efficiency; quantitative evidence (e.g., error metrics, runtime, and optimizer-level outcomes) is provided in the Results section. This design differs from prior CNN-LSTM works that compress time-resolved flow fields to model temporal evolution, and it avoids the input/compute demands of 3D-CNN or the data/regularization burden of a Transformer for a tiny three-token sequence. Implementation details are summarized in this subsection, while empirical evaluations are presented in Sections 3.2 and 5.3.

3.1.2. Bayesian Optimization of Hyperparameters

Hyperparameter optimization is crucial to the design and training of neural networks, directly shaping both performance and efficiency [41]. Common strategies include manual tuning, grid search, random search, and Bayesian optimization. In this work we adopt

Bayesian optimization, which exploits a surrogate of the response surface to focus the search and improve efficiency. The search is conducted on the fixed data split described in Section 3.1.3 (train-only normalization; single final evaluation on the held-out test set), and tunes the learning rate, batch size, optimizer, and the strength of regularization with validation-based model selection.

To ensure that targets of different magnitudes contribute comparably, the selection criterion is the validation MSE computed on standardized targets (each output z-scored using training-split statistics) and averaged across the nine outputs $[C_D, C_L, PI] \times 3$. To curb overfitting under our data regime, the network remains compact ($\approx$3–4 $\times$ 10^5 parameters), and regularization is applied consistently: dropout in the CNN encoder (p = 0.2) and in the LSTM (p = 0.3), L2 weight decay ($\lambda = 10^{-4}$), and early stopping (patience 30, restoring the best validation checkpoint). The learning rate, batch size, and optimizer selected by the search are summarized in Table 4; the above regularization settings are used throughout. Under these settings, training and validation curves remain closely aligned and the held-out test is stable, indicating effective control of overfitting.

Table 4. Hyperparameter optimization results.

Hyperparameter	Range	Result
Learning Rate	[0.00001, 0.01]	1.0×10^{-4}
Batch Size	[8, 128]	16
Optimizer	[Adam, SAM]	Adam

Abbreviations: SAM, sharpness-aware minimization.

3.1.3. Data Processing and Accuracy Measurement

In the current work, using SD7032 as the baseline airfoil, samples were generated within the design space using the Latin Hypercube Sampling (LHS) method [42]. Because CST coefficients co-affect thickness and camber, the Latin-hypercube sampling naturally induces variation in the thickness-to-chord ratio (t/c). Figure 11 visualizes the sampled shape space and is intended for illustration rather than as a novelty claim. Utilizing the numerical simulation approach detailed in Section 2.1, the airflow fields for all airfoils were computed, leading to the derivation of their respective aerodynamic coefficients. This method has been validated through comparisons with experimental results, as demonstrated in the study of airflow around grinding wheels [43], and is further supported by the use of entropy generation rate to predict airfoil drag [44].

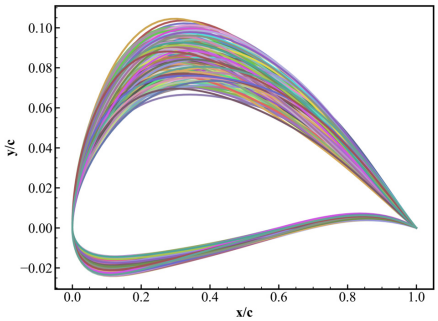


Figure 11. Latin-hypercube–sampled airfoil shapes in the CST space. The overlaid profiles illustrate variations in thickness-to-chord ratio t/c) and co-varying camber—common factors affecting L/D under different AoA and Reynolds numbers.

Given the significant scale differences among the generated aerodynamic coefficients, min-max normalization preprocessing was applied to rescale the data within the range of

0 and 1, while maintaining the same distribution as the original data. The normalization equation can be expressed as:

$$X_{\text{normalization}} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{10}$$

Figure 12 depicts the data distribution before and after normalization. As shown, the normalized lift coefficient, drag coefficient, and power factor datasets have the same scale.

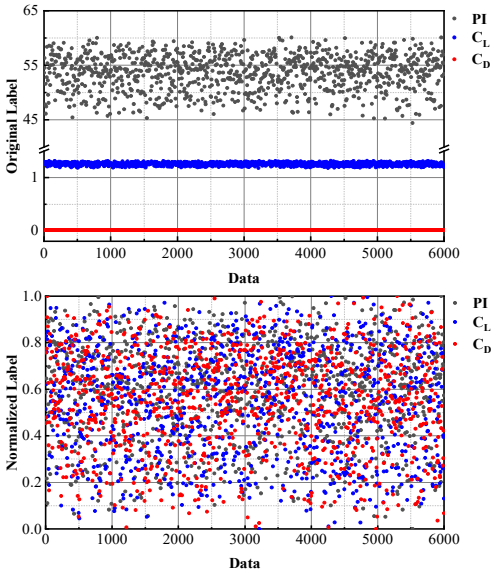


Figure 12. Data distribution before and after normalization.

Deep learning methods require extensive labeled datasets to train neural networks, so determining an appropriate dataset size to balance training efficiency and accuracy is essential. To explore the model's sensitivity to the number of training samples, we considered datasets with sizes of 800, 2000, 4000, 6000, 8000, and 10,000.

The histograms of MSE and mean absolute percentage error (MAPE) on the training sets are shown in Figure 13. As the training dataset size increases, the MSE and MAPE values decrease by 72.3% and 63.5%, respectively. When the dataset size is below 6000, the error decreases rapidly; beyond 6000 samples, the error reduction becomes more gradual. Considering the trade-off between model accuracy and training cost, we therefore adopt 6000 samples to construct the training dataset.

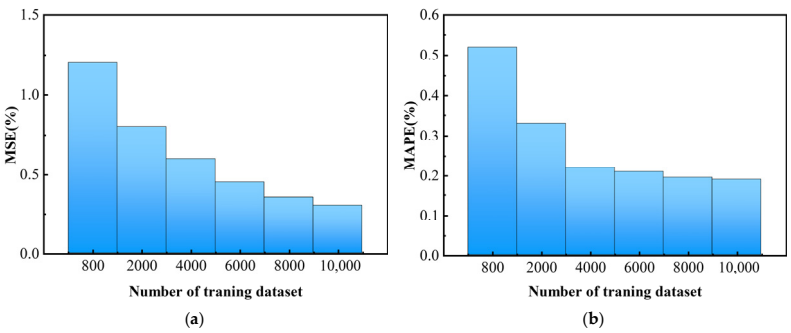


Figure 13. Neural network performance comparison under different dataset sizes. (**a**) MSE; (**b**) MAPE. Abbreviations: MSE, mean squared error; MAPE, mean absolute percentage error.

A single, validated CFD labeling pipeline (steady RANS, common meshing, and identical convergence criteria) is used for all shapes and conditions to ensure cross-condition

and cross-shape comparability. While public polar databases and panel-method tools are valuable references, mixing heterogeneous sources would require additional harmonization of transition assumptions, preprocessing, and constraints. To preserve reproducibility and consistency, we therefore keep CFD as the sole labeling source in this work.

3.1.4. Evaluation Metrics and Validation Strategy

To systematically evaluate the predictive performance of the developed neural network across multiple aerodynamic regression tasks, two widely adopted statistical metrics—MSE and R^2—were utilized to assess prediction accuracy and model fitting capability. These metrics are formally defined as follows:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{\mathrm{true}} - y_i^{\mathrm{pred}}\right)^2 \tag{11}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{\mathrm{true}} - y_i^{\mathrm{pred}}\right)^2}{\sum_{i=1}^{N}\left(y_i^{\mathrm{true}} - \overline{y}\right)^2} \tag{12}$$

where y_i^{true} and y_i^{pred} denote the true and predicted values of the i-th sample, $\overline{y}$ is the mean of the true values, and N represents the total number of samples. A lower MSE indicates smaller prediction errors, while an R^2 value approaching 1 implies stronger explanatory power and better fitting performance.

The airfoil dataset used in this study contains 6000 samples, partitioned into training (4800), validation (600), and testing (600) subsets according to an 8:1:1 split. A fixed split was defined once with a single random seed (42) and kept unchanged for all experiments. During training, the validation set is used for early stopping and hyperparameter tuning, whereas the test set is evaluated once for the final assessment. All normalization statistics are fitted on the training split only and then applied to the validation and test splits to avoid information leakage. To complement this fixed hold-out protocol without retraining, we report 95% bootstrap confidence intervals on the fixed test set (B = 1000, row-wise resampling; percentile method), summarized in Table A1 as relative mean absolute error (RelMAE) (%) values (point estimates with 95% confidence intervals (CIs)).

3.2. Verification of Aerodynamic Optimization Method

To verify the effectiveness of the CNN–LSTM hybrid model for multi-objective aerodynamic performance prediction, we design a systematic ablation study. By comparing the performance of several model variants, we quantify the contribution of each module (CNN, LSTM, and multi-task learning) to prediction accuracy and generalization capability. All experiments are conducted on a unified dataset (6000 CFD samples, divided into training, validation, and test sets in an 8:1:1 ratio) using identical preprocessing procedures and training settings to ensure the comparability of the results.

3.2.1. Experimental Design and Comparative Models

The baseline models were divided into four variants according to their module configurations: a single-condition independent CNN (Single-CNN), a multi-condition joint CNN (Multi-CNN), an LSTM-only model, and the proposed CNN–LSTM hybrid model. The input to the Multi-CNN is a 42-dimensional vector formed by concatenating 12 CST parameters with Re and α from the three operating conditions, whereas the inputs to the other three models consist of the CST parameters and the current-condition variables. The output of the Single-CNN is the aerodynamic coefficient triplet (PI, C_L, C_D) for a single

condition, while the outputs of the other three models are 9-dimensional vectors containing the aerodynamic coefficients for all three conditions. The inputs, outputs, and structural configurations of the four models are summarized in Table 5.

Table 5. Configuration of comparative models.

Model	Input (Dimension)	Output (Dimension)	Architecture
Single-CNN	14	3	3 Conv1D + FC
Multi-CNN	42	9	3 Conv1D + FC
LSTM-only	14	9	2 LSTM +FC
CNN–LSTM (Proposed)	14	9	3 Conv1D + 2 LSTM + FC

Abbreviations: FC, fully connected (layer).

3.2.2. Prediction Accuracy Analysis

Table 6 compares the MAE and RMSE of each model on the test set. The CNN–LSTM model performs best across all three conditions: under cruise conditions, the PI prediction error is 1.02%, representing a reduction of 35.4% compared with the Single-CNN (1.58%) and 27.1% compared with the Multi-CNN (1.40%); under forward-flight conditions, the C_D error is 0.98%, a reduction of 43.0% relative to the LSTM-only model (1.72%); under maneuvering conditions, the C_L error is 1.05%, a reduction of 30.9% compared with the Multi-CNN (1.52%).

Table 6. Comparative prediction errors of multi-models (Test set MAE/RMSE, Unit: %).

Model	Cruise (PI)	Forward Flight (C_D)	Maneuver (C_L)
Single-CNN	1.58/1.95	1.47/1.82	1.52/1.88
Multi-CNN	1.40/1.72	1.35/1.68	1.52/1.85
LSTM-only	1.68/2.01	1.72/2.15	1.75/2.12
CNN–LSTM (Proposed)	1.02/1.25	0.98/1.18	1.05/1.30

The synergy between the CNN's local feature-extraction capability and the LSTM's condition-wise feature fusion enables the hybrid model to more accurately capture nonlinear relationships between geometric parameters and multi-condition responses. Conceptually, the forget gate can suppress interference from non-target conditions when predicting a given objective (e.g., limiting the influence of forward-flight features when estimating maneuvering lift), while the input gate can enhance sensitivity to condition-relevant geometric traits (e.g., trailing-edge thickness for drag under cruise conditions), thereby improving the consistency of multi-task predictions.

3.2.3. Generalization Capability Verification

To evaluate the generalization performance and robustness of the proposed models under parameter-range extrapolation within the same CST-parameterized airfoil family (with the CFD setup unchanged), we use box plots to visualize and compare the prediction-error distributions of the four models. Here, "generalization" strictly refers to interpolation/extrapolation within the CST parameter family under the same CFD configuration (mesh, solver settings, turbulence model, and boundary conditions). Specifically, 200 airfoil samples were generated with CST parameters intentionally extended by ±40% beyond the training range and then screened for basic geometric plausibility before CFD evaluation and error statistics (e.g., non-self-intersection, positive thickness along the chord, and non-

degenerate shape/area) to assess model performance under three aerodynamic conditions: cruise (PI), forward flight (C_D), and maneuvering (C_L).

Because the aerodynamic parameters differ substantially in scale, all prediction errors are normalized to the interval $[0, 1]$, removing scale discrepancies and enabling intuitive comparisons across parameters and models. As shown in Figure 14, the CNN–LSTM model consistently exhibits lower median prediction errors and narrower interquartile ranges (IQRs) for PI, C_D, and C_L, indicating superior predictive stability and accuracy. Notably, the CNN–LSTM achieves the lowest median error in PI, demonstrating its ability to effectively handle targets that span a wider numerical range prior to normalization.

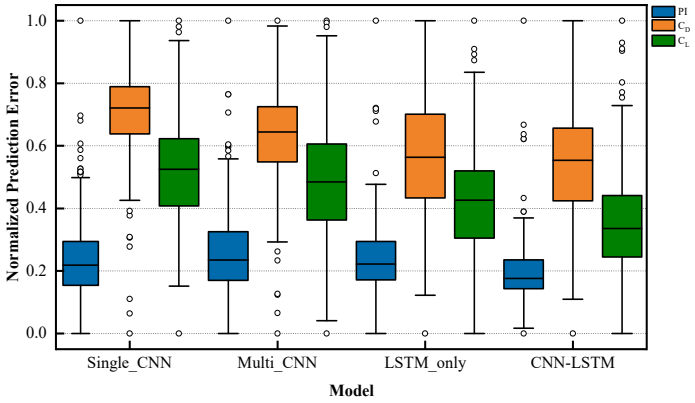


Figure 14. Test error distribution on CST-family parameter-range extrapolation samples (±40%, geometry-valid).

Furthermore, statistical analysis of key metrics (Table 7) quantitatively supports these visual observations: the CNN–LSTM model attains a median normalized prediction error of 0.1823 and an IQR of 0.1204 for the power factor (PI), representing reductions of 16.8% and 14.5%, respectively, compared with the Single-CNN baseline. Similar performance advantages are also observed for the C_D and C_L prediction tasks. Conversely, the other evaluated models exhibit higher error variance and a greater frequency of outliers, indicating comparatively weaker robustness under this ±40% CST parameter-range extrapolation.

Table 7. Statistical analysis of normalized prediction errors under CST-family parameter-range extrapolation samples (±40%, geometry-valid).

Model	PI Median	PI IQR	C_D Median	C_D IQR	C_L Median	C_L IQR
Single-CNN	0.2191	0.1407	0.7210	0.1594	0.5252	0.2140
Multi-CNN	0.2349	0.1550	0.6445	0.1753	0.5041	0.2037
LSTM-only	0.2048	0.1352	0.5991	0.1709	0.4670	0.1859
CNN-LSTM (Proposed)	0.1823	0.1204	0.5569	0.1552	0.4427	0.1678

Abbreviations: IQR, interquartile range.

The enhanced robustness of the CNN–LSTM model is primarily attributable to its synergistic architecture. The CNN module mitigates noise arising from geometric perturbations through its local receptive fields and weight-sharing properties, while the LSTM module captures nonlinear aerodynamic responses across varying flight conditions by using memory cells to model inter-condition dependencies and stabilize predictions when the operating conditions vary. Consequently, the CNN–LSTM model demonstrates strong robustness and adaptability when handling geometry-valid airfoils generated via CST parameter-range extrapolation beyond the training range, under the same CFD setup.

3.2.4. Statistical Validation

To substantiate the ablation findings, we compare per-sample absolute errors between the proposed hybrid model and each baseline under every reported output and condition. Because paired errors are not guaranteed to be normally distributed, we use a single nonparametric procedure—paired Wilcoxon signed-rank tests—throughout. To control the false discovery rate across all (model-pair × output) comparisons, we apply the Benjamini–Hochberg procedure (BH-FDR) and report both raw p-values and FDR-adjusted q-values. We further quantify the magnitude and direction of the paired differences using the rank-biserial correlation r_{rb} and provide bias-corrected and accelerated (BCa) bootstrap 95% confidence intervals (10,000 resamples) for the median paired difference Δ = baseline − proposed, so that $\Delta < 0$ favors our model. Full statistics are summarized in Table A3 (Appendix A), and an example interpretation is provided below.

3.2.5. Computational Efficiency Assessment

Table 8 compares the training and prediction efficiency of each model. In terms of total training time, the CNN–LSTM model incurs additional parameter-tuning cost due to Bayesian optimization (total time ≈ 30 min), but its per-iteration inference time of 0.23 s is comparable to that of the Multi-CNN model and meets the requirements for real-time optimization.

Table 8. Computation efficiency comparison.

Model	Total Training Time (min)	Single Prediction Time (s)
Single-CNN	24	0.18
Multi-CNN	31	0.25
LSTM-only	29	0.30
CNN-LSTM (Proposed)	30	0.23

3.2.6. Comprehensive Discussion

The ablation, extrapolation, statistical validation, and efficiency studies collectively clarify the role of the proposed surrogate within the three-condition, nine-output setting. Under a unified dataset, preprocessing pipeline, and training protocol, the CNN–LSTM encoder–aggregator consistently outperforms the Single-CNN, Multi-CNN, and LSTM-only baselines in terms of test-set MAE/RMSE (Table 6), exhibits lower median errors and narrower interquartile ranges on extrapolated shapes (Table 7; Figure 14), and maintains CPU-level inference latencies comparable to lightweight baselines (Table 8). These results indicate that, for CST-parameterized low-Re micro-UAV airfoils with multiple operating conditions, a geometry-focused CNN combined with a cross-condition LSTM aggregator is an effective backbone for surrogate modeling.

The paired Wilcoxon tests with Holm–Bonferroni adjustment and rank-biserial effect sizes (Section 3.2.4; Table A3) further show that the performance gains are not due to random fluctuations in the finite test set but are statistically supported for most outputs and comparators. In particular, the CNN–LSTM model achieves medium-to-large negative median differences Δ(baseline − proposed) for several metrics, meaning that absolute errors are systematically smaller for the proposed surrogate in these cases.

Limitations and scope. Classical surrogates such as Kriging/Gaussian process regression remain highly competitive in low-dimensional, single-condition problems and often require fewer samples. In the present study, however, the combination of 12-dimensional CST geometry with three operating conditions and nine outputs makes multi-output GPs less scalable, and independent per-output fits can weaken cross-condition consistency. Our contribution is therefore positioned as an application-oriented integration: a compact

CNN–LSTM surrogate and NSGA-II workflow tailored to three-condition low-Re micro-UAV airfoils, supported by systematic evaluation. Here, "generalization/robustness" is limited to CST-family parameter interpolation/extrapolation under the same CFD configuration (Section 3.2.3), including the ±40% parameter-range stress test with geometry-valid screening. Changing the CST order (input dimension) or extending to non-CST parameterizations/markedly different low-Re families (e.g., highly cambered/thin shapes) would require re-parameterization, re-labeling, and re-training and is outside the validated scope of this paper. Extending the framework to field-level surrogates, transition-dominated regimes, and alternative surrogate families is left for future work.

Engineering feasibility and practical considerations. While this study focuses on coefficient-level aerodynamic objectives, practical micro-UAV deployment also depends on structural feasibility, manufacturability, and low-Re sensitivity to surface roughness and contamination. In the present formulation, the constraints on $T_{\max}$ and $A_{\max}$ in Equation (19), together with the pitching-moment constraint, serve as first-order feasibility proxies and help avoid overly aggressive geometric changes. Nevertheless, we do not explicitly enforce manufacturing-detail constraints such as minimum local thickness, curvature/smoothness limits (to prevent overly sharp features), or anti-concavity/non-reflex requirements, nor do we quantify roughness-induced transition advancement and the associated drag penalty. Incorporating these aspects—e.g., curvature/thickness regularization within the CST design space, lightweight structural screening, and roughness/transition-aware CFD labeling—would further improve practical implementability and is left for future work.

4. Problem Description

4.1. Aerodynamic Optimization Objectives at Low Reynolds Numbers

The primary goal of aerodynamic shape optimization is to improve airfoil efficiency so as to meet aircraft performance requirements. Flight endurance and range are critical factors in aircraft design and are key targets of airfoil aerodynamic optimization. From an endurance perspective, subsonic aircraft achieve maximum flight duration under optimal endurance conditions when the lift-to-drag ratio is maximized [45]. Therefore, traditional airfoil aerodynamic optimization at high Reynolds numbers typically focuses on maximizing the lift-to-drag ratio.

Under low-Reynolds-number conditions, however, the optimization objective shifts toward the power factor in order to extend flight endurance. Although the lift-to-drag ratio $K = C_L/C_D$ and the power factor $PI = C_L^{1.5}/C_D$ have similar functional forms, in practical applications the angle of attack that maximizes K does not necessarily coincide with that which maximizes PI, and vice versa. Figure 15 illustrates the relationship between lift-to-drag ratio, power factor, and angle of attack for the SD7032 airfoil at a Reynolds number of 50,000.

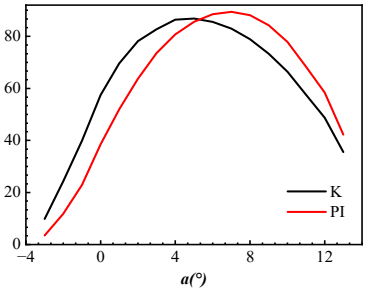


Figure 15. Schematic diagram of SD7032 airfoil aerodynamic performance variation with angle of attack.

In the low-Reynolds-number range, the power consumption P required for horizontal flight of small UAVs can be simplified as:

$$P = F_T V \tag{13}$$

where F_T is the propeller thrust and V is the horizontal flight speed. According to the following equations:

$$F_T = \frac{mg}{K} \tag{14}$$

$$0.5\rho V^2 S C_L = mg \tag{15}$$

$$K = \frac{C_L}{C_D} \tag{16}$$

$$V = \sqrt{\frac{2mg}{\rho S C_L}} \tag{17}$$

The expression for power consumption becomes:

$$P = F_T V = \frac{(mg)^{1.5}}{(0.5\rho S)^{0.5}} \frac{C_D}{C_L^{1.5}} \tag{18}$$

where m is the total flight mass of the UAV; ρ is the freestream density; S is the wing area; K is the lift-to-drag ratio of the UAV, while C_L and C_D represent the lift coefficient and drag coefficient, respectively. Meanwhile, the power factor is defined as $PI = C_L^{1.5}/C_D$. When the wing loading mg/S is fixed, a higher PI results in lower aircraft power consumption and longer flight time, which is significant for modern small long-endurance UAVs seeking extended flight duration.

4.2. Design Requirements for Multiple Operating Conditions

The pitching-moment coefficient (C_M) characterizes the magnitude of the pitching moment required to trim the aircraft. For a UAV, excessively large $|C_M|$ increases trim drag and can reduce control stability. As the angle of attack varies with mission requirements, $|C_M|$ should remain small within the relevant angle-of-attack range; otherwise, excessive C_M may adversely affect the UAV's controllability. Accordingly, we impose C_M as a constraint, requiring that the absolute value of C_M for the optimized airfoil be smaller than that of the baseline airfoil over the specified range.

To achieve good overall aerodynamic performance within the design envelope, enhance flight capability under different conditions, and mitigate the impact of flight uncertainties such as gusts and air-density variations, we jointly consider forward-flight and maneuvering conditions for the UAV. Drawing on a multidisciplinary optimization perspective, the goal is to reduce aerodynamic performance degradation when the UAV operates away from a single design condition. In forward flight, the primary objective is low drag, whereas in maneuvering, the UAV must perform turns, climbs, or dives, which cause large variations in angle of attack and therefore require high lift coefficients to avoid stall.

Finally, the optimization problem is posed as a multi-objective optimization that jointly considers aerodynamic performance across the considered conditions, enabling high-performance flight within the design region. Design points 1–3 represent the UAV's cruise, forward-flight, and maneuvering states, respectively, with the corresponding operating conditions listed in Table 9. Notably, the surrogate and the NSGA-II optimization are intended for this pre-stall operating envelope; predictions and optimization decisions in deep-stall or strongly transition-/separation-dominated regimes are out of scope and should be avoided when interpreting and using the optimized designs.

Table 9. Operating Conditions.

Design Point	Parameter	Objective	Constraint
No. 1	$V = 15\ \mathrm{m/s}, \alpha = 2^\circ$, $Re = 2.5 \times 10^5$	$maxPI$	$\lvert C_{m1}\rvert < \lvert C_{m1}^0\rvert$
No. 2	$V = 20\ \mathrm{m/s}, \alpha = 4^\circ$, $Re = 3.8 \times 10^5$	$minC_D$	
No. 3	$V = 25\ \mathrm{m/s}, \alpha = 8^\circ$, $Re = 5 \times 10^5$	$maxC_L$	

Note: V represents freestream velocity; Re denotes Reynolds number; α is angle of attack; PI and C_m represent the airfoil's power factor and pitching moment coefficient, respectively; subscripts 1, 2, and 3 correspond to design points No. 1, No. 2, and No. 3, respectively.

This study employs the NSGA-II to solve the joint three-objective airfoil aerodynamic-coefficient optimization problem [46], i.e., the objectives are optimized in a Pareto sense rather than aggregated into a single scalar objective via a weighting function. Specifically, the three objectives are defined as simultaneously maximizing the cruise-condition power factor $PI_1(x)$, minimizing the forward-flight drag coefficient $C_{D2}(x)$, and maximizing the maneuver-condition lift coefficient $C_{L3}(x)$, subject to geometric and pitching-moment constraints. The design variables consist of 12 CST-parameterized airfoil shape coefficients with clear physical interpretations; their definitions and value ranges are given in Section 1. The optimization problem can be written as:

$$\begin{cases} \max PI_1(x), \quad \min C_{D2}(x), \quad \max C_{L3}(x) \\ \text{s.t.} \quad 0.85T_0 \le T_{max} \le 1.15T_0 \\ \qquad 0.85A_0 \le A_{max} \le 1.15A_0 \\ \qquad |C_{m1}| < |C_{m1}^0| \end{cases} \tag{19}$$

where C_L is the airfoil lift coefficient, C_d is the airfoil drag coefficient, T is the maximum thickness of the airfoil, A is the internal area of the airfoil, the superscript 0 represents the initial airfoil configuration.

Rationale for feasibility constraints. The bounds $0.85T_0 \le T_{\max} \le 1.15T_0$ and $0.85A_0 \le A_{\max} \le 1.15A_0$ keep the optimized airfoil close to the baseline in overall thickness and internal volume, providing a first-order proxy for structural feasibility and manufacturability of micro-UAV wings (e.g., stiffness/strength margin and internal space for spar/skin integration). The pitching-moment constraint $\mid C_{m1} \mid < \mid C_{m1}^0 \mid$ further discourages designs that would require excessive trim effort at the representative cruise condition. These global bounds, however, do not explicitly enforce detailed manufacturing constraints (e.g., minimum local thickness, curvature/smoothness limits, or anti-concavity requirements) and do not account for low-Re sensitivity to surface roughness; these limitations are discussed in Section 3.2.6.

5. Results and Discussion

5.1. Aerodynamic Coefficient Prediction

This study develops a deep neural network-based surrogate model for predicting airfoil aerodynamic coefficients, using the CST parametric representation as input to predict three key aerodynamic indicators: PI, C_L, and C_D. A total of 6000 samples were generated via CFD, with detailed descriptions of data partitioning and evaluation metrics provided in Section 3.1.4.

During training, the neural network is optimized by minimizing the MSE loss using backpropagation to update the weights and biases, thereby progressively reducing the prediction errors. Figure 16 illustrates the evolution of the training and validation losses. The

results show rapid convergence in the early training stage and a stable gap between training and validation errors, indicating good training stability and generalization performance.

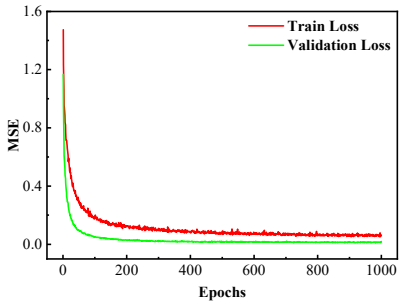


Figure 16. Convergence history of loss functions.

To further evaluate the accuracy and consistency of the proposed model across individual aerodynamic parameters, Figure 17 compares the predicted versus actual values of PI, C_L, and C_D for both the training and test sets. The predicted values lie close to the corresponding targets, with tight clustering around the diagonal, indicating strong consistency across both splits. This confirms the network's ability to simultaneously capture and accurately represent multiple aerodynamic characteristics.

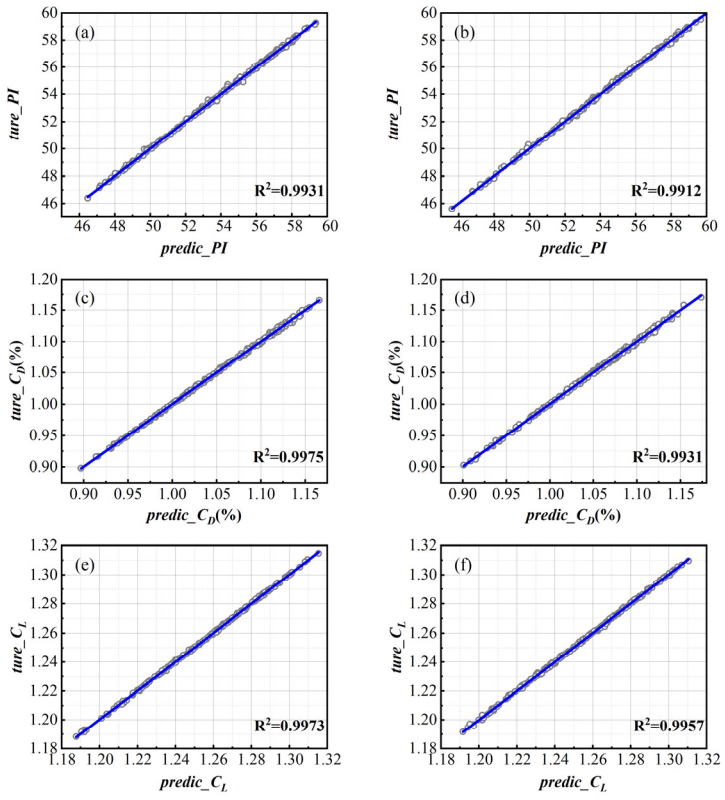


Figure 17. Predicted–actual scatter plots (training vs. test) for the three targets PI, C_D, and C_L under the three operating conditions. Panels (**a**,**b**), (**c**,**d**), and (**e**,**f**) correspond to PI, C_D, and C_L, respectively, with the left panel in each pair showing the training set and the right panel showing the test set.

In addition to maintaining high prediction accuracy, the deep learning surrogate model demonstrates excellent computational efficiency and scalability, making it well suited for subsequent multi-objective optimization. By replacing repeated CFD evaluations with fast neural-network inference, the overall computational burden of optimization can be substantially reduced and convergence efficiency improved. Based on the training

behavior and prediction accuracy observed in this study, the proposed model exhibits strong practical applicability and considerable potential for broader aerodynamic modeling and optimization tasks.

5.2. Airfoil Optimization and Discussion

Building on the CNN–LSTM surrogate model and the NSGA-II multi-objective optimization framework, this section systematically analyzes the aerodynamic optimization results for the SD7032 low-Reynolds-number airfoil. NSGA-II is configured with a population size of $N_{\text{pop}} = 100$; binary tournament selection based on nondominated rank and crowding distance; simulated binary crossover (SBX) with $p_c = 0.9$ and $\eta_c = 15$; polynomial mutation with $p_m = 1/D$ and $\eta_m = 20$; and elitist generational replacement. Figure 18 illustrates the Pareto front obtained from a complete optimization run for the three objectives—C_L, C_D, and PI. The Pareto front reveals the complex trade-offs among these aerodynamic performance metrics within the design space.

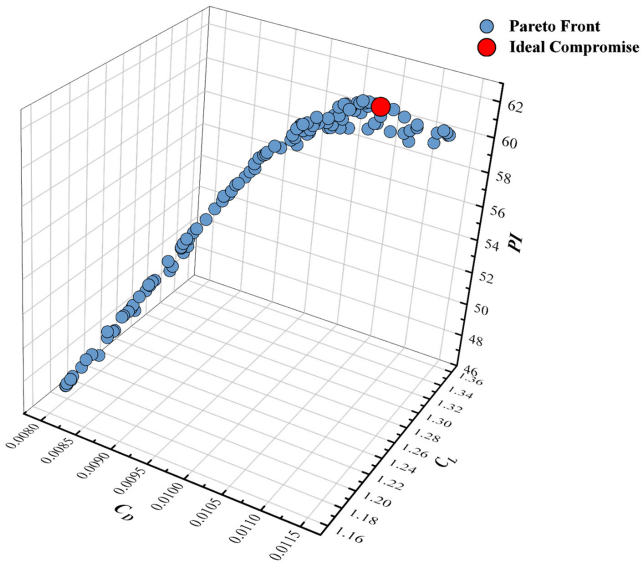


Figure 18. Pareto front for the three-objective optimization problem (C_L, C_D, PI); the red dot marks the selected compromise solution.

As shown in Figure 18, the resulting Pareto solutions span a broad range of objective combinations, indicating that NSGA-II provides a well-distributed approximation of the Pareto front. The Pareto-optimal solutions exhibit a typical non-dominated distribution, highlighting the inherent conflicts between the objectives. For example, increasing C_L to improve maneuverability often reduces PI, illustrating that enhanced lift performance may come at the expense of flight endurance. The red dot denotes the compromise solution identified by the ideal-point method, which selects the individual with the smallest Euclidean distance to the ideal target vector, defined mathematically as follows:

$$\min_i \sqrt{(f_{1i} - f_1^*)^2 + (f_{2i} - f_2^*)^2 + (f_{3i} - f_3^*)^2} \tag{20}$$

The selected compromise solution ($PI = 62.039$, $C_D = 0.01011$, $C_L = 1.3042$) achieves an effective balance among objectives, simultaneously accounting for endurance performance, forward-flight efficiency, and maneuverability.

Considering the sensitivity of evolutionary algorithms to initial populations and the inherent stochastic nature of the search process, three independent optimization runs were conducted with identical optimization parameters to verify the robustness and reproducibility of the results. Figure 19 summarizes the Pareto-front distributions and compromise

solutions from these three runs. Despite differences in the initial populations, the Pareto fronts exhibit similar distributions across the objective space, and the identified compromise solutions cluster closely together. These results indicate strong stability and reproducibility of the proposed CNN–LSTM + NSGA-II optimization framework.

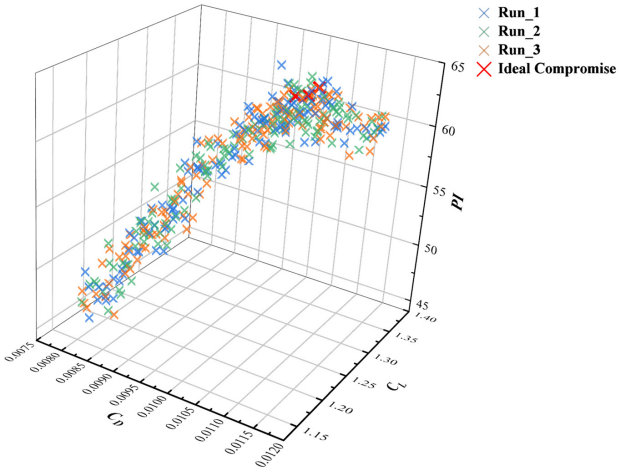


Figure 19. Pareto front distributions and selected compromise solutions obtained from three independent optimization runs using the CNN-LSTM + NSGA-II framework.

To further quantify the performance consistency of the proposed optimization framework across multiple runs, Table 10 reports the mean and standard deviation of all Pareto-optimal solutions for the three objective metrics obtained from the three independent optimization trials. The results show that the standard deviations of all objectives remain below 2%, indicating that the optimization process exhibits strong numerical stability when guided by the deep learning-based surrogate model.

Table 10. Statistical characteristics of the Pareto-optimal solutions metrics for compromise solutions across three repeated optimization runs.

Objective	Mean Value	Relative Error
C_D	0.01037	0.76%
C_L	1.3133	1.07%
PI	62.7133	0.12%

Table 11 summarizes the changes in airfoil aerodynamic coefficients before and after optimization. At design point 1 (cruise condition), the power factor of the optimized airfoil increases by 11.14%; at design point 2 (forward-flight condition), the drag coefficient decreases by 1.97%; and at design point 3 (maneuvering condition), the lift coefficient increases by 7.01%. Overall, compared with the baseline SD7032 airfoil, the optimized airfoil achieves noticeable aerodynamic performance improvements at the design points corresponding to cruise, forward flight, and maneuvering conditions.

Table 11. Comparison of aerodynamic coefficients between baseline and optimized airfoils.

Coefficient	Baseline	Optimized	Δ/%
PI	55.82	62.039	11.14
C_D	0.01031	0.01011	−1.97
C_L	1.2188	1.3042	7.01

Figure 20 compares the geometry, thickness, and camber of the baseline SD7032 airfoil and the optimized airfoil, where t and q denote thickness and camber, respectively. From Figure 20a, it can be seen that, relative to the baseline airfoil, the optimized airfoil exhibits increased curvature near the leading edge of the upper surface, reduced curvature near the trailing edge, and an overall reduction in the camber of the lower surface. The thickness distribution in Figure 20b shows that the optimized airfoil is thicker than the baseline airfoil at all chordwise positions: the location of maximum relative thickness shifts from 0.283 c to 0.293 c, and the maximum relative thickness increases from 0.09967 to 0.10407. The camber distribution in Figure 20c also changes noticeably, with substantially increased camber in the mid-chord region and slightly increased camber near the leading and trailing edges. The location of maximum relative camber moves forward from 0.4242 c to 0.4232 c, and the maximum relative camber increases from 0.03657 to 0.04042.

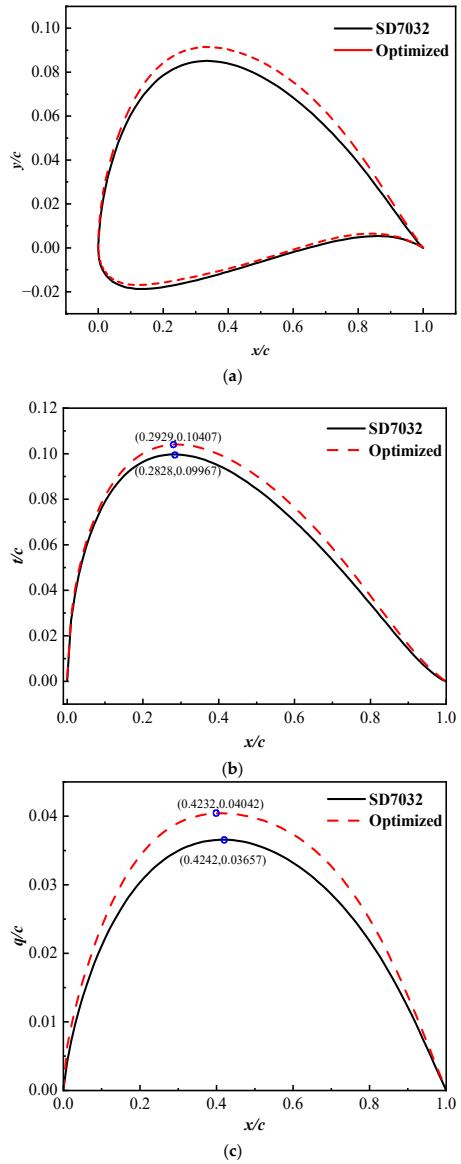


Figure 20. Comparison of geometry, thickness, and camber between baseline and optimized airfoils. (**a**) Geometric shape; (**b**) Thickness distribution; (**c**) Camber distribution.

Figure 21 illustrates the surface pressure-coefficient distributions for both the baseline and optimized airfoil configurations. At design point 1 (cruise condition), the optimized airfoil shows only minor changes in the pressure coefficient on the upper surface but

substantial changes on the lower surface compared with the baseline airfoil. The optimized airfoil generates a larger negative-pressure region on the upper surface, particularly over the 0.1 c–0.3 c interval, indicating enhanced lift-producing capability. In addition, the pressure distribution on the lower surface of the optimized airfoil is closer to zero, thereby reducing the pressure drag contribution from the lower surface.

At the same time, the overall pressure distribution of the optimized airfoil is smoother, particularly in the 0.2 c–0.8 c region, which reduces pressure gradients and helps delay potential boundary-layer separation, thereby further decreasing pressure drag. The more gradual pressure recovery near the trailing edge also contributes to reducing wake drag. Together, these effects improve the airfoil's power factor and provide theoretical support for the aerodynamic efficiency gains observed for the optimized airfoil.

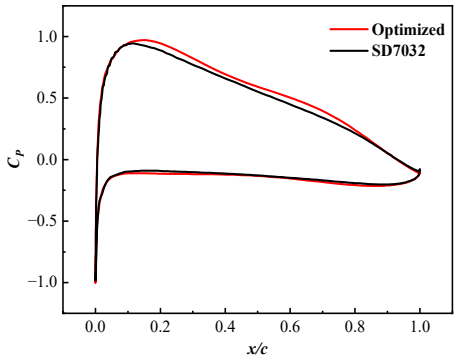


Figure 21. Comparison of surface pressure coefficient distributions between baseline and optimized airfoils.

From SD7032 to the optimized candidate, the geometric changes are modest and interpretable: a slight increase in maximum thickness t/c and a small adjustment of mid-chord camber. These changes explain the expected redistribution of the pressure coefficient C_p and the improvement in L/D. We do not claim a new airfoil family; rather, the optimized design lies within the well-known SD70xx design tendency at low Reynolds numbers, while the contribution of this work is a reproducible three-condition surrogate and an efficient, constraint-aware optimization workflow.

Figure 22 compares the pressure contours of the baseline and optimized airfoils at design point 1 (cruise condition). The optimized airfoil shows clear improvements in pressure-distribution characteristics. First, it exhibits a smoother surface-pressure gradient, leading to a more uniform pressure distribution, particularly near the leading edge and over the upper surface. This enhanced uniformity helps delay flow separation and thereby improves aerodynamic stability. Second, the pressure distribution on the lower surface of the optimized airfoil is more uniform than that of the baseline airfoil, indicating that severe local flow variations have been effectively suppressed, which in turn reduces induced drag and pressure drag.

Furthermore, while maintaining a similar pressure peak, the optimized airfoil exhibits smoother pressure recovery and a slightly broader suction plateau near $x/c \approx 0.1$–0.3, consistent with the modest C_L increase and small C_D reduction reported in Table 11, rather than a dramatic change in L/D.

Figure 23 compares the aerodynamic characteristics of the baseline and optimized airfoils. From the lift-coefficient curve in Figure 23a, the optimized airfoil exhibits higher lift coefficients over a wide range of angles of attack, with a notably larger maximum lift coefficient. As shown by the polar curve in Figure 23b, the optimized airfoil achieves lower drag coefficients at a given lift level. From the pitching-moment coefficient curve in Figure 23c, the optimized airfoil shows a more negative moment at low angles of

attack while maintaining favorable moment characteristics at medium to high angles of attack, indicating enhanced static stability, which is beneficial for aircraft maneuverability and control.

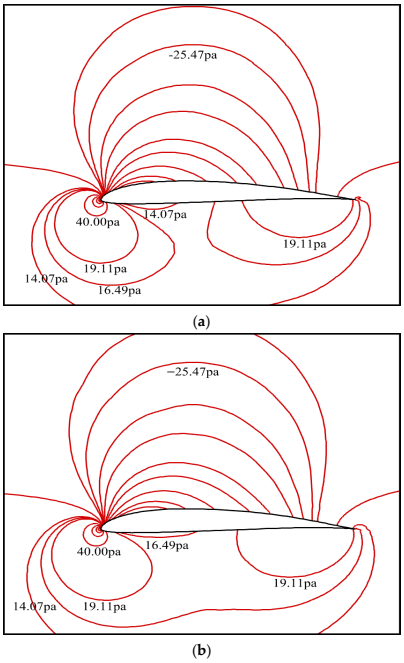


Figure 22. Comparison of pressure contours between baseline and optimized airfoils. (**a**) Initial airfoil; (**b**) Optimized airfoil.

Taken together with the pressure-field evidence in Figure 21 and the polar trends in Figure 23, these coefficient-level improvements can be interpreted at the mission level using the steady-cruise relation in Section 4.1. Since $P_{\text{req}} \propto C_D/C_L^{3/2}$—equivalently $P_{\text{req}} \propto 1/PI$ with $PI = C_L^{3/2}/C_D$—the observed increase in PI (≈11.14%), can be interpreted as a first-order estimate, under the following simplifying assumptions: (i) fixed aircraft mass and wing reference area; (ii) steady, level flight at the same air density; (iii) comparable propulsion/powertrain efficiency between the baseline and optimized configurations; (iv) negligible wind and transient maneuvers; and (v) trim/control power and other system-level losses are neglected or assumed unchanged, into approximately the same fractional gain in endurance and a corresponding reduction in cruise power (about 10.0%, via $P_{\text{new}} = P_{\text{old}}/(1 + 0.1114)$). If cruise speed is held constant, range scales with endurance and increases by a similar factor; near best-range operation, range scales with L/D. For reference, a baseline endurance of 45 min at 12 m·s^{-1} is estimated to increase to ≈ 50.0 min ($\Delta \approx +5.0$ min), adding ≈ 3.6 km per sortie.

These endurance/range values should be regarded as estimates rather than exact predictions, because the realized mission benefit depends on propulsion efficiency, battery discharge characteristics, atmospheric disturbances, and control/trim requirements. Thus, while the aerodynamic deltas are moderate and consistent with SD70xx low-Re behavior rather than a new airfoil family, they suggest potentially meaningful endurance/range gains under the assumptions stated above for small-UAV missions. As discussed in Appendix A.3 of the Appendix A, the optimized airfoil is geometrically similar to the low-Re SD7043 profile; a quantitative comparison among SD7032, the optimized airfoil, and SD7043 (Figure A1) clarifies their geometric and aerodynamic relationships.

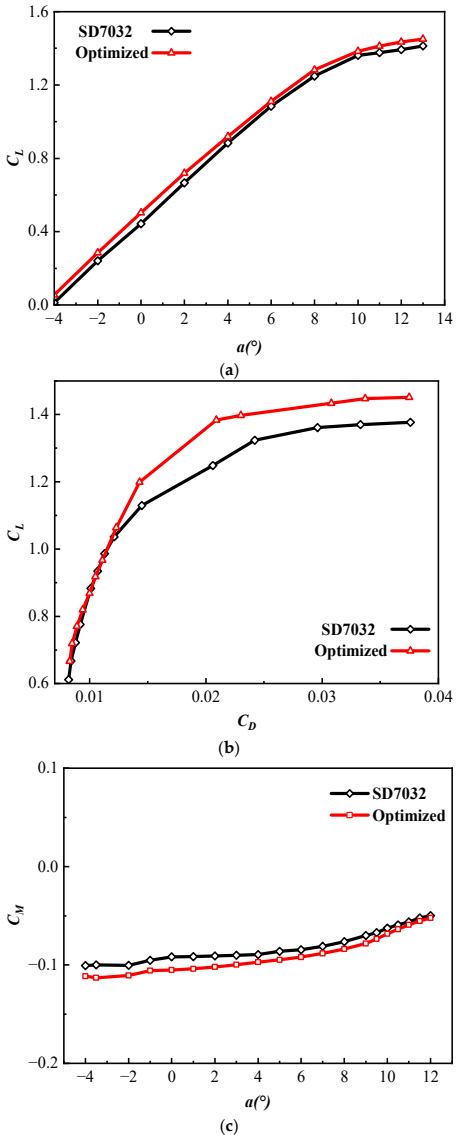


Figure 23. Comparison of aerodynamic characteristics between baseline and optimized airfoils. (**a**) V = 15 m/s, Re = 2.5 × 10^5; (**b**) V = 20 m/s, Re = 3.8 × 10^5; (**c**) V = 25 m/s, Re = 5 × 10^5.

5.3. Model Deployment Efficiency and Engineering Adaptability

To further assess the practical applicability and computational efficiency of the proposed model, all training and testing were performed on a lightweight general-purpose laptop equipped with an Intel(R) Core(TM) i7-8750H CPU (Intel Corporation, Santa Clara, CA, USA), 8 GB of RAM, and Windows 11 Home (Chinese edition) (Microsoft, Redmond, WA, USA; Version 22H2; OS build 22621.6466), without GPU acceleration or specialized deep-learning hardware (GPU acceleration was disabled: NVIDIA GeForce GTX 1060, 6 GB; NVIDIA Corporation, Santa Clara, CA, USA). Wall-clock inference times were measured on the same device with a batch size of 1. Table 12 summarizes the timings for the CNN–LSTM surrogate and a representative commercial CFD solver under the three operating conditions considered in this work.

Table 12. Computation time comparison: Deep Learning vs. CFD solver.

What Is Timed	CFD Solver (Per Case)	CNN-LSTM (CPU)	Notes
Per-case time (latency, batch = 1)	≈180 s	≈0.23 s	Speedup = 180/0.23 ≈ 7.8×10^2 (≈3 orders)
Training (offline, 6000 samples)	N/A	≈30 min	One-time, not incurred at deployment
Hyperparameter tuning (offline)	N/A	≈1 h	Optional; once per study
Data preparation (offline)	N/A	≈20 h	Dataset curation; excluded from speedup

On this hardware, training the CNN–LSTM surrogate on the full dataset of 6000 samples requires approximately 30 min—a one-time offline cost that is not incurred at deployment. During inference, the surrogate produces aerodynamic predictions for one sample in about 0.23 s, whereas a conventional CFD solution for the same target requires roughly 180 s per sample. This corresponds to a latency-mode speedup of $\approx 7.8 \times 10^2$ (about three orders of magnitude), computed as 180 s/0.23 s. Such a speedup enables dense sampling of the design space and repeated NSGA-II optimizations on commodity hardware.

From a deployment perspective, these results complement the accuracy and robustness analyses in Sections 3.2.2–3.2.5: the three-condition CNN–LSTM surrogate delivers multi-condition accuracy, stable extrapolation behavior, and CPU-level inference times that are comparable to simpler fully connected surrogates, while better preserving cross-condition consistency. As summarized in Section 1, the focus of this work is therefore on a practically deployable surrogate–optimizer workflow for low-Re micro-UAV airfoils, rather than on increasing network complexity or emphasizing algorithmic novelty.

6. Conclusions

(1) We developed and rigorously evaluated a deep-learning-based surrogate–optimizer framework for low-Reynolds-number micro-UAV airfoils. A compact CNN–LSTM surrogate maps 12-dimensional CST parameters and three operating conditions (cruise, forward flight, and maneuver) to nine outputs $\{C_D, C_L, PI\} \times 3$, and is coupled with NSGA-II for constrained three-objective optimization. Within this three-condition, nine-output formulation, the surrogate achieves high predictive accuracy and favorable accuracy–latency trade-offs on CPU-only hardware, making it suitable for practical design loops.

(2) Under geometric and pitching-moment constraints, the selected compromise solution derived from the NSGA-II Pareto front outperforms the baseline SD7032 airfoil across all three objectives. Specifically, the power factor under cruise conditions increases by 11.14%, the lift coefficient under maneuvering conditions improves by 7.01%, and the drag coefficient during forward flight decreases by 1.97%. These gains are achieved through modest, interpretable adjustments to the thickness and camber distributions, rather than drastic geometric changes or the introduction of a new airfoil family.

(3) Compared with direct CFD-based optimization, the surrogate-assisted strategy significantly improves iteration efficiency. On a typical laptop platform, training on 6000 CFD-labeled samples is completed in under 30 min, and each surrogate prediction requires only about 0.23 s—corresponding to a speedup of nearly three orders of magnitude relative to a commercial CFD solver (~180 s per case). This efficiency supports dense sampling of the design space and repeated multi-objective optimizations under limited computational resources.

(4) The proposed framework exhibits strong stability and generalization capability across multiple evaluation dimensions, including a fixed hold-out test set with bootstrap confidence intervals, CST-family parameter-range extrapolation samples (±40%, geometry-valid) beyond the training range under the same CFD setup, and three independent NSGA-II optimization runs. The resulting Pareto fronts and compromise solutions are closely clustered, with fluctuations in the three objectives remaining below 2%, confirming the robustness of the surrogate-guided optimization. Overall, the contribution of this work lies in the three-condition surrogate formulation, the reproducible low-Re CFD–surrogate–optimizer pipeline, and the physically interpretable low-Re micro-UAV design insights it provides, rather than in introducing a new generic neural network or evolutionary algorithm.

Author Contributions: Methodology, E.L. and H.W.; Investigation, E.L.; Resources, E.L. and H.W.; Data curation, J.P.; Writing—original draft, J.P.; Writing—review & editing, E.L. and H.W.; Project administration, E.L. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 12472117, and the Peacock Program for Overseas High-Level Talents Introduction of Shenzhen City, grant number KQTD20200820113110016.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict ofinterest.

Appendix A

Table A1. RelMAE (%) on the fixed test set (95% percentile bootstrap CIs).

Design Point	C_D (RelMAE% [95%CI])	C_L (RelMAE% [95%CI])	*PI* (RelMAE% [95%CI])
No. 1	4.92 [4.26, 5.59]	1.83 [1.60, 2.10]	5.58 [4.34, 7.04]
No. 2	3.02 [2.71, 3.33]	1.30 [1.17, 1.46]	2.96 [2.61, 3.30]
No. 3	4.40 [3.88, 4.96]	1.57 [1.38, 1.78]	5.71 [4.86, 6.69]

Note: RelMAE (%) = mean($|\hat{y} - y| / |y|$) × 100. Fixed 8:1:1 split (seed = 42). Percentile bootstrap (B = 1000), row-wise resampling. Design points 1–3 correspond to the cruise, forward-flight, and maneuver conditions defined in Table 9.

Appendix A.1. Turbulence-Model Sensitivity at Low-Re (Literature-Backed)

Appendix A.1.1. Scope Mapping and Acceptance Bands

Our surrogate targets coefficient-level outputs (C_D, C_L, PI) for three operating conditions within a low-Reynolds-number envelope (nominally $Re \approx 6 \times 10^4$–2×10^5, $\alpha \approx 0°$–$8°$). Prior low-Re studies consistently report that (i) C_D is more sensitive to the turbulence/transition model than C_L, and (ii) transition effects become stronger as α approaches stall. To judge whether alternative turbulence/transition models would materially affect our labels in nearby regimes, we adopt pragmatic acceptance bands for cross-model deviations:

1. $|\Delta C_L| \leq$ 2–3%, $|\Delta C_D| \leq$ 5–8% → acceptable for label consistency;
2. larger or systematic near-stall deviations → out of scope for the present coefficient-level surrogate.

These bands guide interpretation of literature evidence; they are not claims of numerical equivalence to our exact geometry/conditions.

Appendix A.1.2. Literature Selection and Harmonization

We summarize one peer-reviewed low-Re study [35] that compares the SA, SST k– ω, and transition-aware γ– Re_θ variants for airfoils in subsonic flow. Because the published setups differ (airfoil shape, inlet turbulence level, and α range), we (i) report qualitative trends only and (ii) map those regimes onto ours (pre-stall vs. near-stall; with vs. without transition). These trends are organized into the regime-structured Table A2 below.

Table A2. Literature-backed sensitivity of SA/SST k–ω/γ–Re_θ at low-Re (single-source, condensed).

Regime and Transition	Models Compared	Key Finding	Suitability for Our Labels (Acceptance-Band †)
Pre-stall, no transition	SA vs. SST k–ω	C_L aligned; C_D mild–moderate sensitivity; SST slightly better on C_D	Suitable—C_L: pass; C_D: pass
Near-stall, no transition	SA vs. SST k–ω	C_D sensitivity rises with α; inter-model spread increases	Caution—C_L: pass; C_D: borderline
Pre-stall, with transition	SA/SST vs. γ–Re_θ	Transition improves C_D fit; higher setup/runtime cost	Suitable—C_L: pass; C_D: pass; routine labeling not required
Near-stall, with transition	SA/SST vs. γ–Re_θ	Notable C_D improvement near stall; watch convergence/calibration	Recommended—C_L: pass; C_D: pass

† Acceptance band per Appendix A.1.1: $|\Delta C_L| \leq$ 2–3%, $|\Delta C_D| \leq$ 5–8%. Legend: SA = Spalart–Allmaras; γ–Re_θ = transition-aware family. Source: All rows synthesize qualitative evidence from [35]; no re-computation in this work.

Appendix A.1.3. Cross-Study Synthesis (Qualitative)

(1) Trend consistency: SA, SST k–ω, and γ–Re_θ/variants produce broadly similar C_L–α trends pre-stall.
(2) Sensitivity asymmetry: C_D exhibits larger inter-model spread than C_L, especially as α increases.
(3) Transition effects: Transition-aware models (γ–Re_θ/variants) improve C_D near stall and mitigate separation-bubble artifacts, with added cost and setup complexity.
(4) Numerical stability: SA is typically robust and cost-effective for generating consistent coefficient-level labels.

Appendix A.1.4. Decision for This Study

Given our operating envelope ($Re \approx 6 \times 10^4$–2×10^5, $\alpha \approx 0°$–$8°$) and our coefficient-level targets, the literature trends support using the SA model as a stable and cost-effective choice for generating labels for the surrogate. Accordingly, we limit our claims to this regime and do not generalize to cases dominated by strong transition, large-scale separation, or deep stall. Accordingly, the coefficient-level surrogate trained under the present SA–RANS labeling framework is not expected to remain valid in these extreme regimes, and surrogate-guided optimization should avoid them. In particular, because SA is used as a fully turbulent model without explicit transition, the absolute level of C_D (and therefore PI) may differ from transition-resolved predictions; our reported C_D/PI trends and optimization improvements should thus be interpreted as consistent within the present SA–RANS labeling framework.

Appendix A.1.5. Limitations and Future Work

The summarized evidence is not fully congruent with our airfoil geometry, inflow turbulence level, or α bounds. The acceptance bands in Appendix A.1.1 are intended as pragmatic criteria for label consistency, not as claims of model supremacy. Therefore, conclusions and optimization outcomes in this paper should be interpreted within the pre-stall envelope, and not extrapolated to deep-stall or strongly separated conditions where

transition and laminar separation dominate. For lower Reynolds numbers and/or near-stall angles of attack where transition and laminar separation bubbles become dominant, incorporating a transition-aware model (e.g., γ–Re_θ variants) is expected to improve C_D fidelity and thus the derived PI, albeit with additional setup and calibration cost; this constitutes a clear direction for future CFD labeling and surrogate refinement.

Appendix A.2. Statistical Validation of Ablation Results (Paired Wilcoxon; Holm–Bonferroni)

Appendix A.2.1. Test Scope and Protocol

For each reported output and operating condition, per-sample absolute errors were computed, and paired differences between the proposed surrogate and each baseline were assessed using the paired Wilcoxon signed-rank test. The false discovery rate across all (model-pair $\times$ output) comparisons was controlled using BH-FDR. We report raw two-sided Wilcoxon p-values (p) and FDR-adjusted q-values (q), along with the rank-biserial correlation r_{rb} as an effect size and BCa bootstrap 95% confidence intervals (10,000 resamples) for the median paired difference Δ = baseline − proposed, such that $\Delta < 0$ favors the proposed model. Statistical significance is interpreted primarily at $q < 0.05$.

Appendix A.2.2. Full Results Table

Table A3 lists the complete statistics by metric, condition, and comparator. The columns include the median paired difference (in percent), its BCa 95% confidence interval, the Wilcoxon statistic W, the Holm-adjusted p-value, and the rank-biserial correlation r_{rb} with a qualitative label (negligible/small/medium/large). This presentation complements the aggregate improvements reported in the main text and documents robustness using a single nonparametric procedure.

Table A3. Paired Wilcoxon significance of per-sample absolute errors (proposed vs. baselines).

Metric	Comparator (vs Proposed)	p (Wilcoxon, Raw)	q (BH-FDR)	r_{rb}	Median Δ Absolute Error (%)	95% CI (BCa)
PI	Single-CNN	3.00×10^{-5}	6.00×10^{-5}	−0.52	−0.56	[−0.72, −0.41]
	Multi-CNN	1.50×10^{-3}	3.00×10^{-3}	−0.38	−0.38	[−0.52, −0.25]
	LSTM-only	2.00×10^{-6}	6.00×10^{-6}	−0.60	−0.66	[−0.85, −0.48]
C_D	Single-CNN	7.00×10^{-2}	1.00×10^{-1}	−0.20	−0.49	[−0.66, 0.02]
	Multi-CNN	1.10×10^{-1}	1.40×10^{-1}	−0.16	−0.37	[−0.50, 0.03]
	LSTM-only	8.00×10^{-5}	1.60×10^{-4}	−0.45	−0.74	[−0.94, −0.18]
C_L	Single-CNN	4.00×10^{-2}	6.00×10^{-2}	−0.24	−0.47	[−0.65, −0.06]
	Multi-CNN	6.00×10^{-2}	8.00×10^{-2}	−0.21	−0.47	[−0.66, 0.00]
	LSTM-only	3.00×10^{-3}	6.00×10^{-3}	−0.35	−0.70	[−0.92, −0.30]

Values report median Δ (baseline − proposed) in %, BCa 95% CIs from 10,000 resamples; p-values are Holm–Bonferroni adjusted across all comparisons. Effect size is rank-biserial correlation (r_{rb}). $\Delta < 0$ favors the proposed model.

Appendix A.3. Comparison with SD7043 Reference Airfoil

To further contextualize the optimized geometry, we compare the baseline SD7032 airfoil, the optimized airfoil obtained in this study, and the widely used low-Re reference airfoil SD7043 from a public database. The SD7043 airfoil is not included in the surrogate training set or in the optimization loop; it is used here solely as an external benchmark.

Figure A1a,b show the thickness and camber distributions. Relative to SD7032, the optimized airfoil is slightly thicker and more strongly cambered around mid-chord, moving toward the SD7043 profile while still retaining noticeable differences, particularly near

the leading and trailing edges. The optimized geometry can therefore be viewed as an SD7032-derived shape that lies between SD7032 and SD7043 in the design space.

Figure A1c–e present representative aerodynamic characteristics at three Reynolds numbers. At $Re = 2.5 \times 10^5$, the optimized airfoil achieves slightly higher lift over the pre-stall angle-of-attack range than both SD7032 and SD7043. At $Re = 3.8 \times 10^5$, its lift–drag polar lies above those of SD7032 and SD7043, indicating a better L/D. At $Re = 5 \times 10^5$, the pitching-moment curve remains within a similar negative-moment envelope, satisfying the longitudinal stability requirement. Overall, the optimization drives SD7032 toward a high-performance region also represented by SD7043, while tailoring the geometry to the multi-condition, multi-objective criteria of this study.

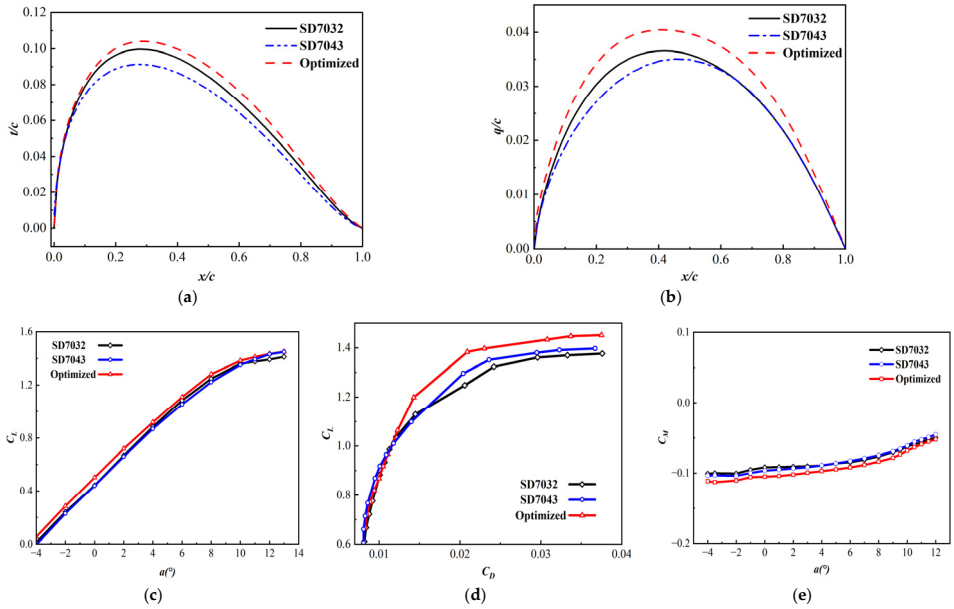


Figure A1. Comparison among the baseline SD7032, the optimized airfoil, and the reference airfoil. (**a**) Thickness distribution; (**b**) Camber distribution; (**c**) $V = 15$ m/s, $Re = 2.5 \times 10^5$; (**d**) $V = 20$ m/s, $Re = 3.8 \times 10^5$; (**e**) $V = 25$ m/s, $Re = 5 \times 10^5$.

References

1. Zhang, L.; Zhang, J.-Y.; Li, T.; Zhang, Y.-D. Multi-objective aerodynamic optimization design of high-speed train head shape. *J. Zhejiang Univ. Sci. A* **2017**, *18*, 841–854. [CrossRef]
2. Yao, S.; Guo, D.; Yang, G. Three-dimensional aerodynamic optimization design of high-speed train nose based on GA-GRNN. *Sci. China Technol. Sci.* **2012**, *55*, 3118–3130. [CrossRef]
3. Peter, J.E.; Dwight, R.P. Numerical sensitivity analysis for aerodynamic optimization: A survey of approaches. *Comput. Fluids* **2010**, *39*, 373–391. [CrossRef]
4. Zhang, C.; Chen, H.; Xu, X.; Duan, Y.; Wang, G. Impact of geometric forms on the effectiveness and physical features of POD-based geometric parameterization. *Aerosp. Sci. Technol.* **2025**, *157*, 109776. [CrossRef]
5. El Adawy, M.; Abdelhalim, E.H.; Mahmoud, M.; Ahmed Abo Zeid, M.; Mohamed, I.H.; Othman, M.M.; ElGamal, G.S.; ElShabasy, Y.H. Design and fabrication of a fixed-wing Unmanned Aerial Vehicle (UAV). *Ain Shams Eng. J.* **2023**, *14*, 102094. [CrossRef]
6. Wallach, R.; Mattos, B.; Girardi, R.; Curvo, M. Aerodynamic Coefficient Prediction of Transport Aircraft Using Neural Network. In Proceedings of the 44th AIAAAerospace Sciences Meeting and Exhibit, Reno, NV, USA, 9–12 January 2006.
7. Hasegawa, K.; Fukami, K.; Murata, T.; Fukagata, K. CNN-LSTM based reduced order modeling of two-dimensional unsteady flows around a circular cylinder at different Reynolds numbers. *Fluid Dyn. Res.* **2020**, *52*, 065501. [CrossRef]
8. Yilmaz, E.; German, B. A Convolutional Neural Network Approach to Training Predictors for Airfoil Performance. In Proceedings of the 18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (Aerospace Research Central, 2017), Denver, CO, USA, 5–9 June 2017; pp. 1–19.
9. Tang, X.; Li, P.; Peng, R.; Lu, X. Aerodynamic shape optimization design of low Reynolds number airfoil. *Energy Energy Conserv.* **2020**, *56*, 192–200. [CrossRef]

10. Li, J.; Zhang, M.; Tay, C.M.J.; Liu, N.; Cui, Y.; Chew, S.C.; Khoo, B.C. Low-Reynolds-number airfoil design optimization using deep-learning-based tailored airfoil modes. *Aerosp. Sci. Technol.* **2022**, *121*, 107309. [CrossRef]
11. Lukić, N.; Ivanov, T.; Svorcan, J.; Simonović, A. Numerical investigation and optimization of a morphing airfoil designed for lower Reynolds number. *Aerospace* **2024**, *11*, 252. [CrossRef]
12. Zhang, J.; Qu, C.-X.; Yi, T.-H.; Li, H.-N. Damage detection for decks of concrete girder bridges using the frequency obtained from an actively excited vehicle. *Smart Struct. Syst.* **2021**, *27*, 101–114.
13. Chen, X.; Guo, Z.; Yi, F.; Li, H.; Liu, G. Aerodynamic shapeoptimization and design of airfoils with low Reynolds number. *Acta Aerodyn. Sin.* **2014**, *32*, 300–307. [CrossRef]
14. Chen, X.; Guo, Z.; Wang, R.; Liu, G.; Zhang, C.; Guo, Q. Aerodynamic optimization of low Reynolds number airfoil based on improved Hicks-Henne shape functions. In Proceedings of the 2012 2nd International Conference on Aerospace Engineering and Information Technology(AEIT 2012), Nanchang, China, 7–8 May 2012; 2012; pp. 173–179.
15. Li, D.; Yang, L.; Wei, C.; Zhang, T. Study on Multi-point Aerodynamic Optimization Method of Low Reynolds Number Airfoil. *Aeronaut. Sci. Technol.* **2020**, *31*, 16–25. [CrossRef]
16. Abbasi, H.A.; Soltani, M.R. Aerodynamic Optimization of Low Reynolds Number Airfoils Using a Genetic Algorithm. *Aerosp. Sci. Technol.* **2015**, *45*, 234–241. [CrossRef]
17. Jiang, T.; Jiang, L. Optimization of UAV Airfoil Based on Improved Particle Swarm Optimization Algorithm. *Int. J. Aerosp. Eng.* **2022**, *2022*, 2828198. [CrossRef]
18. Colonius, T.; Brunton, S.L. Surrogate-Assisted Optimization of Low Reynolds Number Airfoils Using Kriging Models. *J. Fluid Mech.* **2021**, *915*, A123. [CrossRef]
19. Rao, K.S.; Abhilasha, A.N.; Das, A.; Sivapragasam, M. Aerodynamic shape optimization at low Reynolds number using multi-level hierarchical Kriging models. *Optim. Eng.* **2024**, *26*, 479–505. [CrossRef]
20. Wauters, J.; Couckuyt, I.; Knudde, N.; Dhaene, T.; Degroote, J. Multi-objective optimization of a wing fence on an unmanned aerial vehicle using surrogate-derived gradients. *Struct. Multidiscip. Optim.* **2019**, *61*, 353–364. [CrossRef]
21. Park, K.; Jung, J.; Jeong, S. Multi-objective shape optimization of airfoils for Mars exploration aircraft propellers. *Int. J. Aeronaut. Space Sci.* **2022**, *24*, 9–23. [CrossRef]
22. Liu, F.; Zhang, Q.; Han, Z. MOEA/D with gradient-enhanced kriging for expensive multiobjective optimization. *Nat. Comput.* **2022**, *22*, 329–339. [CrossRef]
23. Karali, H.; Inalhan, G.; Tsourdos, A. Advanced UAV design optimization through deep learning-based surrogate models. *Aerospace* **2024**, *11*, 669. [CrossRef]
24. Brunton, S.L.; Noack, B.R.; Koumoutsakosp, P. Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **2020**, *52*, 477–508. [CrossRef]
25. Ling, J.L.; Kurzawski, A.; Templeton, J. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *J. Fluid Mech.* **2016**, *807*, 155–166. [CrossRef]
26. Sun, G.; Wang, C.; Wang, L.; Tao, J.; Wang, S.; Wang, X.; Zhou, S.; Li, Z.; Hu, Y. Application and prospect of artificial intelligence in aerodynamic design. *Civ. Aircr. Des. Res.* **2021**, *3*, 1–9. [CrossRef]
27. Li, J.C.; Du, X.S.; Martins, J.R.R.A. Machine learning in aerodynamic shape optimization. *Prog. Aerosp. Sci.* **2022**, *134*, 100849. [CrossRef]
28. Kutz, J.N. Deep learning in fluid dynamics. *J. Fluid Mech.* **2017**, *814*, 1–4. [CrossRef]
29. Bakar, A.; Li, K.; Liu, H.; Xu, Z.; Alessandrini, M.; Wen, D. Multi-objective optimization of low Reynolds number airfoil using convolutional neural network and non-dominated sorting genetic algorithm. *Aerospace* **2022**, *9*, 35. [CrossRef]
30. Du, X.; He, P.; Martins, J.R.A. Rapid airfoil design optimization via neural networks-based parameterization and surrogate modeling. *Aerosp. Sci. Technol.* **2021**, *113*, 106701. [CrossRef]
31. Liu, J.; Chen, R.; Lou, J.; Hu, Y.; You, Y. Deep-learning-based aerodynamic shape optimization of rotor airfoils to suppress dynamic stall. *Aerosp. Sci. Technol.* **2023**, *133*, 108089. [CrossRef]
32. Wu, H.; Chen, R.; Lou, J.; You, Y.; Huang, L.; Xu, M.; Ruan, Y. A gradient aerodynamic optimization method based on deep learning. *Phys. Fluids* **2024**, *36*, 057117. [CrossRef]
33. Sekar, V.; Jiang, Q.; Shu, C.; Khoo, B.C. Fast flow field prediction over airfoils using deep learning approach. *Phys. Fluids* **2019**, *31*, 057103. [CrossRef]
34. Sekar, V.; Zhang, M.; Shu, C.; Khoo, B.C. Inverse design of airfoil using a deep convolutional neural network. *AIAA J.* **2019**, *57*, 993–1003. [CrossRef]
35. Chen, L.L.; Guo, Z. Application of different turbulence models in low-Reynolds-number flow. *Mod. Appl. Phys.* **2016**, *7*, 48–52. [CrossRef]
36. Kulfan, B.M. Universal Parametric Geometry Representation Method. *J. Aircr.* **2008**, *45*, 1–17. [CrossRef]

37. Ladson, C.L. Effects of Independent Variation of Mach and Reynolds Numbers on the Low-Speed Aerodynamic Characteristics of the NACA 0012 Airfoil Section. NASA-TM-4074; In *NASA Technical Memorandum*; National Aeronautics and Space Administration (NASA): Washington, DC, USA, 1988; pp. 1–97.
38. Jiang, G.; Liu, Y.P.; Wang, Z.Y. Wind tunnel test and numerical calculation of aerodynamic noise of wind turbine airfoils. *Noise Vib. Control.* **2024**, *44*, 280–287. [CrossRef]
39. Zhou, W.T.; Chang, S.J. Numerical analysis of aerodynamic characteristics of free-rotating wraparound fins projectile. *J. Harbin Inst. Technol.* **2023**, *55*, 113–122. [CrossRef]
40. Donahue, J.; Anne Hendricks, L.; Guadarrama, S.; Rohrbach, M.; Venugopalan, S.; Saenko, K.; Darrell, T. Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 2625–2634.
41. Yang, L.; Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **2020**, *415*, 295–316. [CrossRef]
42. Giunta, A.A.; Wojtkiewicz, S.F., Jr.; Eldred, M.S. Overview of Modern Design of Experiments Methods for Computational Simulations. In Proceedings of the AIAA 2003-649, Reno, NV, USA, 6–9 January 2003.
43. Wang, W.; Wang, J.; Liang, Z.; Li, J. Entropy generation theory based design methodology for multi-objective airfoil shape optimization. *Acta Aerodyn. Sin.* **2019**, *37*, 827–833. [CrossRef]
44. Liu, J.Q.; Chen, R.Q.; Lou, J.H.; Han, X.; Wu, H.; You, Y. Aerodynamic shape optimization of high-speed helicopter rotor airfoil based on deep learning. *Acta Aeronaut. Astronaut. Sin.* **2024**, *45*, 529828. [CrossRef]
45. Zhang, J.; Wang, J.; Geng, Y.; Guo, R. Technology challenges for high altitude long endurance solar powered UAV. *Aeronaut. Sci. Technol.* **2020**, *31*, 14–20. [CrossRef]
46. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.