# Aerodynamic Coefficients Prediction via Cross-Attention Fusion and Physical-Informed Training

**Yueqing Wang**[*1,2], **Peng Zhang**[*1,2], **Yushuang Liu**[3,1], **Jianing Zhao**[4,2], **Jie Lin**[1,2,5], **Yi Chen**[3,1†]

[1]State Key Laboratory of Aerodynamics, Sichuan, China,
[2]Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Sichuan, China,
[3]Sichuan Tianfu Fluid Big Data Research Center, Chengdu, China,
[4]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China,
[5]College of Computer Science and Technology, National University of Defense Technology, Changsha, China,
{yqwang2013, zhaojianingx, lukeyichen}@163.com, {zhangpeng14f, linjie15}@nudt.edu.cn, 2738549743@qq.com

## Abstract

Aerodynamic coefficient prediction is pivotal in aircraft and vehicles' design, performance evaluation, and motion control. Integrating artificial neural networks into aerodynamic coefficient prediction offers a promising alternative to traditional numerical methods burdened by extensive computations and high costs. Nevertheless, this data-driven approach faces several critical challenges, which limit its further performance enhancement: i) The current research lacks a profound understanding of the complex interplay between the shape of an object and its aerodynamic characteristics. ii) The scarcity of high-quality aerodynamic data poses a significant barrier. The models trained on limited datasets lack generalization ability, struggling to accurately predict and adapt to diverse aerodynamic performance under new shapes or conditions. To overcome these challenges, we introduce an innovative framework that employs cross-attention to capture the intimate interplay between shape and flow conditions and allows for the direct utilization of pre-trained models on general shape datasets to mitigate the scarcity of aerodynamic data. Furthermore, to bolster the inference capabilities of this data-driven approach, we integrate physical information constraints into the model, leveraging them as guiding principles to enhance the model's predictive power under unknown conditions. Experimental validation demonstrates that our proposed method performs excellently in multiple aerodynamic prediction tasks. This achievement brings a new technological breakthrough to the field of aerodynamic prediction and provides robust support for the design optimization of complex systems such as aircraft and vehicles.

## Introduction

Accurately predicting the aerodynamic coefficients of complex-shaped objects under varying flow conditions has long been a critical challenge for scientific research and engineering applications (Secco and Mattos 2017; Peng, Zhang, and Desmarais 2021). Traditional approaches rely heavily on costly experimental testing or intricate numerical simulations, which are expensive and time-consuming and struggle to encompass the full spectrum of possible shapes and flow factors (Wang, Kou, and Zhang 2022). With the rapid advancements in artificial intelligence, particularly the remarkable performance of deep learning in data processing and pattern recognition (Zeng et al. 2023), data-driven aerodynamic prediction models have gradually emerged as a research hotspot (Yetkin, Abuhanieh, and Yigit 2024). Nevertheless, this data-driven approach faces several critical challenges, including a lack of profound understanding of the complex interplay between shape and aerodynamic characteristics, and the scarcity of high-quality aerodynamic data, which hinders the development of effective and generalized models.

In recent years, point clouds as a powerful tool for representing three-dimensional data have revolutionized the fields of object recognition and shape analysis (Chen et al. 2024; Huang et al. 2024; Wu et al. 2023). Their abundance of geometric detail and adaptability render point clouds an exceptional choice for accurately portraying object shapes, which is vital for aerodynamic coefficient predictions. Building upon existing point cloud processing architectures provides a solid foundation for devising more efficient and effective models. However, the intricacies of aerodynamic predictions demand a holistic approach that encompasses not just the geometric features but also the integration of complex and dynamic flow conditions (Xiong et al. 2022). Modeling the intricate interplay between shape and these variable flow factors remains a significant challenge, necessitating a deep understanding of the physical phenomena involved and innovative methodologies capable of capturing these multifaceted relationships. As such, the quest for a comprehensive model that seamlessly integrates shape and flow conditions represents an open and exciting area of research.

Moreover, purely data-driven methods can be influenced by data biases or noise when confronted with complex aerodynamic phenomena (Jiahao, Wenbo, and Weiwei 2023). This can lead to limitations in their generalization capabilities. Incorporating physical information constraints into the model design is a viable approach to address this and improve prediction accuracy and reliability. By integrating physical laws and equations, the model receives additional supervisory signals to guide it toward learning feature rep-

[*]These authors contributed equally.
[†]Corresponding Author.

resentations and predictions more consistent with physical principles. This, in turn, can strengthen the model's generalization ability and improve its interoperability across various flow conditions.

In this work, we propose a cross-attention-based transformer model. The initial intention is to fully leverage the existing research accumulations in deep learning while facilitating adaptability to different scenarios of aerodynamic coefficient prediction. Structurally, the model separates the extraction and fusion of geometric and aerodynamic features, making it convenient to migrate state-of-the-art (SOTA) models pre-trained on general datasets in different domains into this model, thereby alleviating the impact of insufficient aerodynamic data. Additionally, in terms of training algorithms, a plug-and-play physical constraint method is introduced to incorporate relevant physical theorems or formulas tailored to aerodynamic applications, thereby enhancing the model's extrapolation capability and generalization. The main contributions in this work include:

- By structurally separating the processes of extracting and fusing geometric and aerodynamic features, the model facilitates the integration of deep learning models pre-trained on general datasets from different domains.
- Proposing a cross-attention mechanism within the transformer model enables effective interactions between geometric and aerodynamic representations. This helps the model capture complex relationships between the two types of features, which is crucial for accurate aerodynamic coefficient prediction.
- Introducing a plug-and-play physical constraint method allows the incorporation of relevant physical theorems or formulas tailored specifically to aerodynamic applications. This approach enhances the model's extrapolation capability and generalization, ensuring the predictions align with the underlying physics.

## Related Works

### Aerodynamic Coefficient Prediction

In the context of aerodynamic coefficient prediction tasks, traditional numerical methods are computationally intensive, limiting rapid design iterations (Karali et al. 2020). (Rajkumar and Bardina 2002; Thirumalainambi and Bardina 2003) have thus shifted towards data-driven Artificial Neural Networks (ANNs) for more efficient and accurate forecasts. Convolutional Neural Networks (CNNs) have also shown promise in extracting features from flow field data, enabling more accurate aerodynamic predictions (Chen et al. 2020a). (Chen et al. 2020b) have harnessed this power, demonstrating CNNs' versatility in tackling aerodynamic challenges. Recurrent Neural Networks (RNNs) have been utilized to capture temporal dependencies in aerodynamic behaviors to predict dynamic aerodynamic responses over time (Ignatyev and Khrabrov 2015). However, these implementations tested on limited datasets, focusing on simple geometries under controlled conditions, which does not align with the reality of diverse shapes and complex flow conditions encountered in real-world applications.

Moreover, the cornerstone of aerodynamic coefficient prediction lies in seamlessly blending geometric features with inflow conditions. Multimodal data fusion, a technique that fuses information from multiple sources, significantly boosts the predictive capacity of data-driven models in this domain. Traditional fusion methods in aerodynamic fields have relied on straightforward feature concatenation or MLPs for combining data streams. In recent studies, such as (Zhao, Zeng, and Shao 2023), where the airfoil shape remains constant, the focus shifts to using inflow conditions as inputs to predict aerodynamic coefficients under varying conditions. Conversely, (Chen et al. 2020b) multiplies airfoil images by Mach number (Ma) and Angle of Attack (AoA), leveraging CNNs to extract modified features for prediction. The DRNN model (Lee et al. 2023) represents an advancement, using residual MLPs to capture geometric and inflow features separately, then concatenating them. The prediction head further refines predictions through a residual MLP. (Xiong et al. 2022) takes a novel approach, utilizing PointNet (Qi et al. 2017) to extract features from point cloud representations while directly concatenating inflow conditions with these features for prediction by an MLP.

### Optimization Models Guided by Incorporating Physical Information

In recent years, the effective integration of physical knowledge into data-driven models has emerged as a research focus within the rapid development of deep learning. Among these frameworks, physics-informed neural networks (PINNs) (Raissi, Perdikaris, and Karniadakis 2019; Cai et al. 2021) have stood out as an innovative approach that directly incorporates physical laws as constraints during neural network training, significantly enhancing the accuracy and physical consistency of models in solving complex physical problems. PINNs achieve this by minimizing a composite loss function that includes both a physics-based loss term and a data-based loss term, showcasing their powerful predictive capabilities across multiple physical domains such as fluid dynamics and solid mechanics.

However, facing challenges like data scarcity and limited model generalization ability in practical applications, the standalone PINN approach often falls short of optimal performance. Consequently, researchers have begun exploring strategies to combine PINNs with other technologies to enhance model performance and applicability further (Li et al. 2024). Among these, transfer learning offers new avenues for PINNs by leveraging knowledge learned from source tasks to assist in the teaching of target tasks (Chen et al. 2021; Liu et al. 2023). Specifically, integrating PINNs with transfer learning aims to accelerate the training process of new PINN models and improve their generalization capabilities in new scenarios by transferring partially pre-trained parameters or feature representations from similar physical problems or related datasets. This approach can effectively mitigate data scarcity issues, leverage existing physical knowledge and model experience, and reduce the time and resources required for training models from scratch.

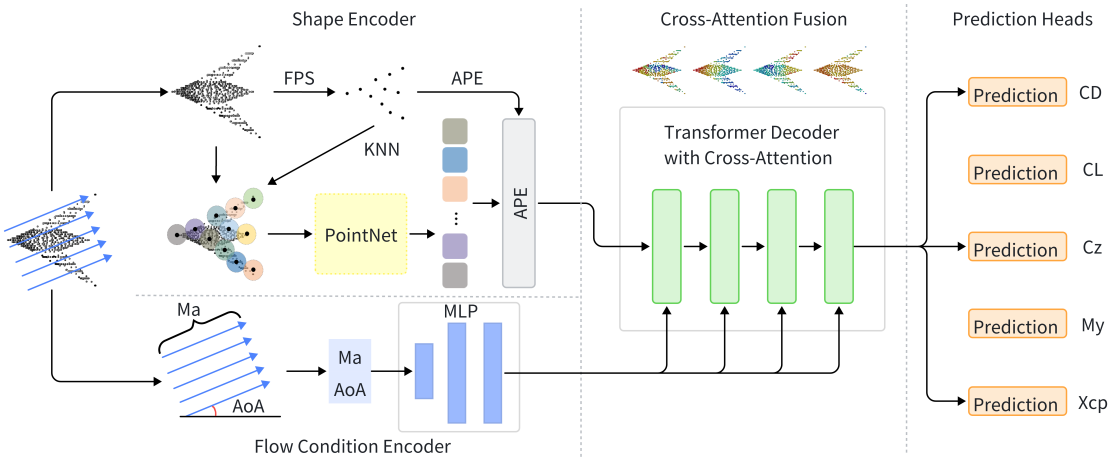


Figure 1: Modular and Scalable Deep-Learning Architecture for Predicting Aerodynamic Coefficients.

# Methods

## Overall Architecture

Firstly, we propose a modular and scalable deep-learning network architecture specifically designed to predict aerodynamic coefficients, ensuring flexibility and adaptability to the ever-evolving landscape of ANNs across various domains. The core philosophy behind this design is to facilitate the seamless integration of future advancements in ANN research. Maintaining a modular structure allows individual components to be easily updated with the latest SOTA models emerging, without requiring significant restructuring of the overall framework. As depicted in Figure 1, this architecture incorporates multiple modules, including shape feature encoder, flow condition encoder, cross-attention fusion, and prediction heads, collectively contributing to achieving high-precision and robust predictions.

**Shape Encoder** Expanding on the theoretical underpinnings, grid-based data can often be prohibitively large for processing, especially when dealing with complex 3D shapes like aircraft. Downsampling the point cloud data significantly reduces data size while preserving critical shape features. This reduction is crucial since the geometric shape of an aircraft directly influences its aerodynamic performance, which is paramount in designing and analyzing such structures.

Hence, the need arises for a sophisticated feature extraction network capable of comprehensively capturing the nuances within the point cloud data. PointGPT's architecture (Chen et al. 2024), adapted for this purpose, offers a robust framework by incorporating point cloud processing techniques tailored to handle sparsity and disorderliness. The three-stage process—partitioning, sorting, and embedding—not only addresses these inherent challenges but also ensures that the extracted features accurately represent the geometric characteristics of the aircraft.

The Farthest Point Sampling (FPS) and K-nearest neighbors (KNN) algorithms used in partitioning effectively condense the point cloud into meaningful point blocks, each capturing local shape information. The sorting stage, leveraging Morton encoding, establishes an order among these blocks, facilitating the subsequent feature extraction process. Finally, PointNet, renowned for its ability to handle unordered point sets, is employed for embedding with Absolute Position Encoding (APE), transforming the sorted point blocks into a high-dimensional feature space that captures rich geometric information.

This comprehensive approach ensures that the feature extraction network is efficient and capable of extracting features highly relevant to aerodynamic analysis. By leveraging these techniques, researchers and engineers can gain deeper insights into the shape-performance relationship, ultimately leading to the design and optimization of more efficient and effective aircraft structures.

**Flow Condition Encoder** The flight attitude of an aircraft significantly impacts its overall aerodynamic forces, such as velocity (commonly represented by Ma), and AoA. To fully consider the influence of flow conditions on aerodynamic coefficients in this technical solution, an MLP is employed to encode the flow conditions, forcibly mapping them into a high-dimensional space to obtain their encodings in this space. To facilitate subsequent feature fusion with the aircraft shape encoding, the dimensions of the flow condition encoding must be consistent with those of the shape encoding, as expressed in the following equation:

$$M_{emb} = MLP([Ma, AoA]), M_{emb} \in R^D \qquad (1)$$

Here, $M_{emb}$ represents the encoded flow conditions, and the MLP takes the $Ma$ and $AoA$ as inputs to produce an encoding vector that lies in the D-dimensional space ($R^D$).

**Cross-Attention Fusion** After obtaining the shape features $S_{emb}$ and flow condition features $M_{emb}$, the model performs feature fusion through a Transformer's decoder module based on cross-attention. This module enables interactions and dependency modeling between shape features and

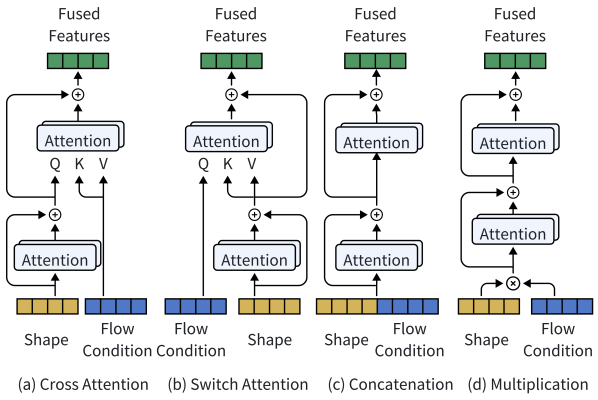


Figure 2: Comparison of Four Different Fusion Strategies

flow condition features, as shown in the following equation:

$$E_{fuse} = Fusion(S_{emb}, M_{emb}), E_{fuse} \in R^D \quad (2)$$

The shape features, augmented with positional encodings $Pos_{emb}$ for the central point coordinates, are fed into a self-attention module to facilitate a more comprehensive learning of global feature representations below:

$$SS_{emb} = S_{emb} + Pos_{emb}, SS_{emb} \in R^{n \times D} \quad (3)$$

$$S_{att} = Softmax(\frac{SS_{emb} SS_{emb}^T}{\sqrt{N}}) SS_{emb}, S_{att} \in R^{n \times D} \quad (4)$$

The output of the self-attention module, along with the flow condition encoding, is then input into a cross-attention module, as shown in Figure 2(a). Through the cross-attention mechanism, the model learns which shape features are more sensitive to specific flow conditions, enabling more accurate predictions of aerodynamic performance, as illustrated by:

$$SC_{att} = Softmax(\frac{S_{att} M_{emb}}{\sqrt{N}}) M_{emb}, SC_{att} \in R^{n \times D} \quad (5)$$

We explored several fusion strategies as in Figure 2: switching attention, concatenation, and multiplication. Switching attention means adjusting the cross-attention mechanism in Transformer-like architectures to separately process shape and flow condition, considering their interactions during attention calculation. Specifically, we change the cross-attention part in Equation 5 to:

$$SC_{att} = Softmax(\frac{M_{emb} S_{att}}{\sqrt{N}}) S_{att}. \quad (6)$$

The fused features $SC_{att}$ then pass through a FeedForward Network (FFN), and the output serves as the input for the subsequent block. Notably, the flow condition encoding input to the cross-attention module of each encoder remains consistent, effectively reinforcing the influence of flow conditions on the fused features.

**Prediction Head** The fused feature vector then enters the prediction module. A multi-head predictor design approach is adopted for the prediction module to accommodate the uniqueness of different aerodynamic coefficients. Each prediction head comprises an independent MLP, specifically trained and optimized for different aerodynamic coefficients. This design enables the model to make accurate predictions for each parameter while simultaneously processing multiple aerodynamic coefficients, thereby avoiding interference among them:

$$\hat{y}_i = MLP_i(E_{fuse}), \hat{y}_i \in R \quad (7)$$

Where $\hat{y}_i$ denotes the predicted value for the $i-th$ aerodynamic coefficient, $E_{fuse}$ represents the fused feature vector, and $MLP_i$ is the MLP designed specifically for the $i-th$ aerodynamic coefficient.

The structural design of this multi-head predictor fully considers the differences in feature representation and data distribution among different aerodynamic coefficients, thereby enhancing the model's prediction accuracy and generalization capability for each parameter.

## Learning Algorithm

The shape encoder component adopts a GPT-like training strategy. The encoding module first generates a point cloud feature sequence, which is then fed into an extractor comprising multiple Transformer layers to learn and capture the latent representations of the point cloud data. Simultaneously, a generator predicts subsequent point patches in an autoregressive manner, assisting the extractor in deeper learning. After the pre-training phase, the generator is discarded, and the extractor, with the dual masking strategy removed, utilizes the learned latent shape features for aerodynamic coefficient prediction.

**Physical-Informed Training** To broaden the applicability of the model, especially in scenarios requiring data extrapolation, we integrate the concepts of PINNs by incorporating physical laws and equations relevant to the prediction target into the model design. Rather than directly relying on the model to predict all independent variables within the physical formula, this approach cleverly selects key predicted parameters such as velocity and pressure, utilizes these physical parameters to perform further calculations through predefined physical equations, and then directly assesses the error between these processed data and actual measurements or experimental data. This process indirectly ensures the consistency between the model predictions and physical laws.

In the specific application of aerodynamics, our goal is to accurately predict critical performance indicators such as the lift coefficient ($C_L$), drag coefficient ($C_D$), X-axial force coefficient ($C_x$), etc. for objects with specific geometric shapes. To achieve this, we leverage fundamental aerodynamic and mechanical principles, integrating these formulas:

$$C_x^{cal} = cos(AoA) \times C_D - sin(AoA) \times C_L, \quad (8)$$

Also, the side force coefficient $C_z^{cal}$ is calculated as:

$$C_z^{cal} = sin(AoA) \times C_D + cos(AoA) \times C_L, \quad (9)$$

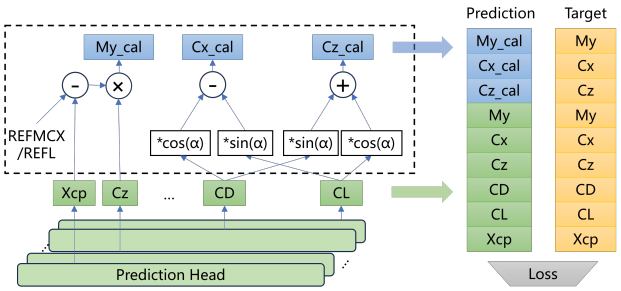


Figure 3: Integration of Physical Equations.

Furthermore, the calculated Y-axial pitch moment coefficient $M^y_{cal}$ is derived as:

$$M_y^{cal} = MCX_{ref} - Xcp \times L_{ref} \times C_z / L_{ref}, \quad (10)$$

where $MCX_{ref}$, $Xcp$, $L_{ref}$, and $C_z$ represent the reference moment center, pressure center position, reference length, and normal force coefficient, respectively.

As shown in Figure 3, the calculated additional aerodynamic coefficients $M_y^{cal}$, $C_x^{cal}$, $C_z^{cal}$ are then concatenated with the directly predicted aerodynamic coefficients:

$$\hat{y} = Concat(\hat{y}, M_y^{cal}, C_x^{cal}, C_z^{cal}), \quad (11)$$

to compute an overall loss with target ones.

It is noteworthy that when introducing physical equations, special attention must be paid to the sources of the various parameters in these equations. These parameters must be information already contained within the sample datasets, serving either as input features for the model or as prediction targets. This design ensures that the model has actual data to support its learning process, enabling it to acquire useful information relevant to the task at hand.

Furthermore, we aspire for the model to utilize this physical information in a "moderate" manner. By "moderate", we mean that we hope the model can benefit from these physical laws to assist in its primary task without becoming overly reliant on them to the point of neglecting other important information or features in the data. In practical applications, data often contains rich and complex information that may not fully align with simplified physical models but is crucial for improving prediction accuracy.

## Experiments

To comprehensively demonstrate the capabilities of the proposed rapid aerodynamic prediction model, we conduct tests on 3D point cloud datasets from both automotive and aircraft domains. For the automotive dataset, we adopted a publicly available dataset to facilitate direct comparisons with other methods, ensuring reproducibility and fairness in our experiments. In contrast, for the aircraft dataset, we took a more unique and in-depth approach. Recognizing the diversity and complexity of aircraft shapes, we generated a series of custom aircraft geometries and obtained rich aerodynamic coefficients through high-fidelity computational fluid dynamics (CFD) simulations. These data empower us to conduct ablation studies to elucidate the effects of key techniques such as cross-attention mechanisms, and physical constraints.

### Datasets

**DrivAerNet++** DrivAerNet++ comprises 8000 diverse car designs modeled with high-fidelity CFD simulations (Elrefaie et al. 2024), and is integrated into NVIDIA Modulus. The dataset includes diverse car configurations. Each entry in the dataset features detailed 3D meshes, parametric models, aerodynamic coefficients, and extensive flow and surface field data, along with segmented parts for car classification and point cloud data.

**AircraftData** The construction of AircraftData began with a thorough capture of aerodynamic characteristics, encompassing 1000 unique aircraft geometries tested across a wide range of AoA from $-7$ to 13 and various Ma. Specifically, for each geometry, we calculated key aerodynamic coefficients under 96 different flow conditions, including but not limited to the $C_x$, $C_z$, $M_y$, $Xcp$, $C_D$, and $C_L$.

We split the dataset into the training&testing set, and three distinct validation sets focused on shape generalization $V_s$, condition generalization $V_c$, and both shape and condition generalization $V_{s+c}$. In subsequent experiments, unless specifically stated, the three validation datasets will be merged into a single dataset, denoted as $V_{all}$, for use. The training and testing set comprises shapes 1 to 900, encompassing samples under flow conditions with a refined range of the middle $80\%$ of the values for Ma and AoA. The validation sets are further categorized into three types: the shape validation set utilizes shapes 901 to 1000, totaling 4524 samples under the same flow conditions; the condition validation set employs shapes 1 to 900 but under extreme flow conditions,i.e. the extreme $10\%$ of the data for both Ma and AoA, containing 2425 samples; finally, the shape and condition generalization validation set also utilizes shapes 900 to 1000 but under these extreme flow conditions, with 291 samples. This split methodology comprehensively evaluates the model's generalization capabilities under various conditions.

### Metrics

To facilitate a consistent comparison with the prediction task in the automotive aerodynamic dataset, we evaluate the performance of our models using Mean Squared Error ($MSE$), Mean Absolute Error ($MAE$), Mean Absolute Error ($MAE$) and Coefficient of Determination ($R^2$ Score). In the ablation study utilizing aircraft aerodynamic datasets, we primarily utilize Mean Relative Error ($MRE$) as a comprehensive evaluation metric and $MRE_n$ represents inverse standard normalization after prediction.

### Implementations

We adopt a unified training setup for automotive and aeronautical aerodynamic prediction models. The work is carried out on Nvidia A100 GPU featuring 40GB of memory capacity, leveraging the PyTorch 1.10.1 framework. We utilize AdamW as the optimizer and employ the Cosine Learning Rate strategy for learning rate scheduling, where the learning rate gradually increases from a minimum of $10^{-7}$ to a maximum of $10^{-4}$, repeating every 60 epochs. To mitigate overfitting, we set the weight decay coefficient to $5 \times 10^{-4}$.

| Methods | | $MSE\downarrow$ ($\times 10^{-5}$) | $MAE\downarrow$ ($\times 10^{-3}$) | $MaxAE\downarrow$ ($\times 10^{-3}$) | $R^2\uparrow$ | $MRE\downarrow$ ($\times 10^{-2}$) |
|---|---|---|---|---|---|---|
| PointNet (Charles et al. 2017) | | 14.9 | 9.60 | 12.45 | 0.643 | - |
| GCNN (Kipf and Welling 2017) | | 17.1 | 10.43 | 15.03 | 0.596 | - |
| RegDGCNN (Elrefaie, Dai, and Ahmed 2024) | | 14.2 | 9.31 | 12.79 | 0.641 | - |
| Shape only | 64c32n | 9.031 | 7.306 | 10.341 | 0.935 | 2.627 |
| | 100c64n | 5.293 | 5.654 | 7.316 | **0.962** | 2.021 |
| Shape+LWH | 64c32n | 8.096 | 6.825 | 9.318 | 0.941 | 2.453 |
| | 100c64n | **5.275** | **5.629** | **7.243** | **0.962** | **2.016** |

Table 1: Comparative analysis of deep learning models for aerodynamic drag prediction.

| Fusion Strategies | $MAE(\times 10^{-3})$ | $MSE(\times 10^{-3})$ | $MRE(\times 10^{-2})$ | $MRE_n(\times 10^{-2})$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $Total$ | $C_X$ | $C_Z$ | $M_y$ | $C_D$ | $C_L$ | $Xcp$ |
| Cross Attention | **6.728** | 13.304 | **3.661** | **7.254** | **7.629** | **5.569** | **14.523** | **3.855** | **5.579** | 6.404 |
| Switch Attention | 6.996 | 12.409 | 3.807 | 7.653 | 7.712 | 5.698 | 15.976 | 3.920 | 5.736 | 6.914 |
| Concatenation | 7.185 | **11.876** | 3.910 | 7.973 | 8.386 | 5.820 | 17.879 | 4.238 | 5.882 | 5.679 |
| Multiplication | 7.746 | 13.256 | 4.215 | 8.753 | 8.623 | 6.242 | 19.600 | 4.263 | 6.282 | **4.263** |

Table 2: Comparison between different shape-condition fusion methods configuration.

Furthermore, to smoothly initiate the training process, we include a warm-up period of 60 epochs and continually adjust and optimize the model performance over the total 300 epochs. Additionally, we use MRE Loss to complete the network training, ensuring that our models minimize the average absolute difference between predicted and actual aerodynamic values.

In the shape encoder and cross-attention fusion, we adopt the same network structure as PointGPT, allowing for direct fine-tuning of the pre-trained network. For other components, to reduce the impact of specialized network architectures and to highlight the cross-attention fusion and physical-informed training approaches presented in this work, we utilize simple MLP structures. Specifically, in the flow condition encoder section, we employ an MLP with dimensions $\{f_{in}, 64, 128, 512\}$, where $f_{in}$ represents the flow condition dimension. Each prediction head utilizes an MLP with dimensions $\{256, 256, 1\}$.

## Standard Benchmarks

To benchmark against other aerodynamic coefficient prediction models, we initially conducted experiments on automotive aerodynamic prediction dataset. Given that all vehicles in this dataset travel at one uniform speed, i.e., one single flow condition, the proposed method's full potential leveraging cross-attention for information fusion could not be fully realized. Consequently, we incorporated simplified vehicle information as flow conditions into our model.

Table 1 presents the predictive outcomes of various methods on the automotive aerodynamic prediction dataset. The terminology "Shape only" denotes using solely point cloud data of vehicles as input, whereas "Shape+LWH" signifies incorporating vehicle length, width, and height as flow features into the flow condition encoding. The notations "c" and "n" represent the number of center points and neighbors utilized during point cloud feature extraction.

Our approach outperforms all others across all metrics, highlighting its efficacy. By incorporating vehicle dimensions as flow condition inputs, we enhanced input diversity, aiding the model in capturing finer shape characteristics. Furthermore, as the number of sampled points in the point cloud increases, the model achieves more precise shape representations, yielding superior results.

## Ablation Studies

To comprehensively evaluate the capabilities of our proposed model and training algorithm in more practical scenarios, we conducted a series of ablation studies on a self-generated dataset for aircraft aerodynamic prediction. These studies focused on investigating the impact of different fusion strategies for shape and flow conditions and the effects of varying training strategies.

**Fusion Strategies** Accurate fusion of aircraft shape features and flow conditions is crucial for aerodynamic performance prediction. We explored several fusion strategies to assess their influence on prediction accuracy: concatenation, multiplication, and switching the cross-attention.

The experimental results in Table 2 indicate that the modified cross-attention strategy generally performs best. Element-wise multiplication of shape features and flow conditions at the feature level assumes that feature products reveal nonlinear relationships. Concatenating the shape features and flow conditions is a straightforward approach. While simple, it may fail to capture complex interactions between the two. This demonstrates that a well-designed information interaction mechanism enables the model to utilize shape and flow conditions, enhancing prediction accuracy more effectively.

Moreover, Figure 4 shows how cross-attention focuses shift under diverse flow conditions. Specifically, the first row reveals a distinct shift in attention between the nose and fuselage as the AoA varies. The second row underscores a concentration of attention on the wings under varying Ma.

| Methods | $MRE(\times 10^{-2})$ | $MRE_n(\times 10^{-2})$ | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Total | $C_L$ | $C_Z$ | $M_y$ | $C_D$ | $C_L$ | $Xcp$ |
| Base | 4.793 | 6.527 | 7.760 | 4.765 | 13.328 | 3.125 | 4.714 | 5.504 |
| Shape Feature Reuse | 4.760 | 6.075 | 7.328 | 4.426 | 12.510 | 2.940 | 4.422 | 4.856 |
| Fine-tuning | 4.582 | 5.778 | 7.085 | **4.293** | 11.850 | 2.825 | **4.236** | **4.408** |
| Physical-Informed | 4.676 | 5.919 | 7.146 | 4.489 | 12.127 | 2.847 | 4.424 | 4.512 |
| Physical-Informed Fine-tuning | **4.610** | **5.755** | **6.874** | 4.512 | **11.537** | **2.765** | 4.447 | 4.423 |

Table 3: Comparison between different training strategies.

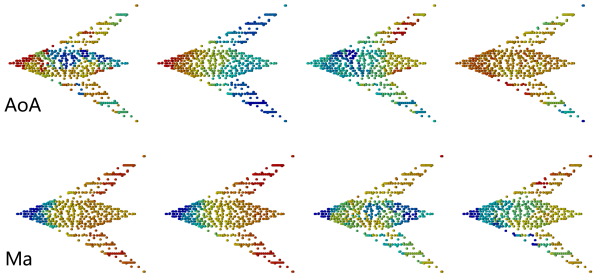


Figure 4: The cross-attention focuses on different positions of the aircraft under different incoming flow conditions.

This capability of our methods to dynamically attend to various exterior features contingent upon flow conditions underscores their potential for achieving superior prediction outcomes.

**Training Strategies** To explore the effects of different training strategies on model performance, we designed the following training methods and compared them under identical experimental settings. The experimental outcomes in Table 3 reveal that the physical-informed fine-tuning achieves the best performance in most cases. This underscores that by incorporating physical constraints into a fine-tuned pre-trained model, we can further enhance prediction accuracy and generalization capabilities.

**Generalization** To comprehensively evaluate the generalization ability of the proposed model in practical applications, i.e., its predictive performance on unseen data, we have designed three distinct strategies for dataset splitting to individually test the model's performance under random conditions, with unknown specific shapes, and with unknown specific flow conditions. The following outlines the detailed experimental design and implementation process.

Figure 5 displays the $MRE_n(\times 10^{-2})$ for various generalization types, each evaluated across differing amounts of training data. As the size of the training set diminishes, the performance of the models deteriorates, evidenced by a rise in the $MRE_n$ values. A comparative analysis among the three generalization types $V_s$, $V_c$ and $V_{s+c}$ in contrast to the $V_{all}$ baseline reveals that incorporating generic shape data and leveraging a pre-trained shape feature extraction network enables the model to maintain relatively consistent performance across varying training data volumes. This finding

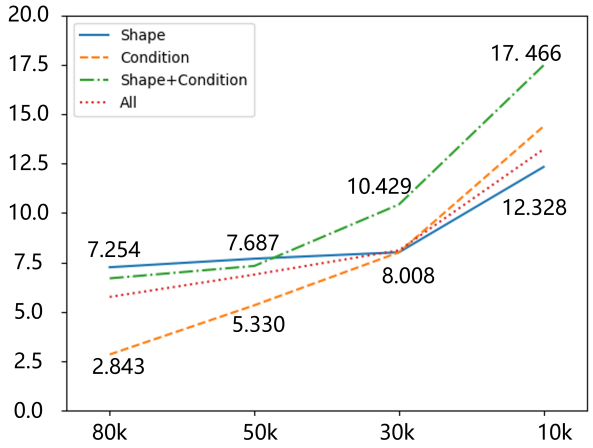


Figure 5: $MRE_n(\times 10^{-2})$ of different generalization types based on varying amounts of training data.

underscores the fundamental premise of our work: harnessing the "free lunch" offered by deep learning to bolster the effectiveness of aerodynamic-related tasks.

## Conclusion

In conclusion, this paper proposes an innovative cross-attention-based transformer model that combines the strengths of deep learning advancements with the specific demands of aerodynamic coefficient prediction. By segregating the extraction and fusion of geometric and aerodynamic features in a strategic manner, the model establishes a versatile framework that adeptly integrates state-of-the-art techniques from diverse disciplines, thereby effectively overcoming the limitation of limited aerodynamic data availability. The modular physical constraint method during training is a significant contribution, as it significantly boosts the model's ability to extrapolate and generalize, thereby ensuring its practical relevance and applicability in real-world aerodynamic analysis scenarios. The results demonstrate the model's efficacy in dynamically focusing attention on pertinent features under varying flow conditions, leading to improved prediction accuracy and performance. In summary, this work presents a promising approach that advances the field of aerodynamic coefficient prediction and paves the way for future research endeavors aimed at enhancing accuracy, robustness, and practicality.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 12472300 and the Natural Science Foundation of Sichuan Province under Grant 2024NSFSC0446.

## References

Cai, S.; Mao, Z.; Wang, Z.; Yin, M.; and Karniadakis, G. E. 2021. Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12): 1727–1738.

Charles, R. Q.; Su, H.; Kaichun, M.; and Guibas, L. J. 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 77–85.

Chen, G.; Wang, M.; Yang, Y.; Yu, K.; Yuan, L.; and Yue, Y. 2024. Pointgpt: Auto-regressively generative pre-training from point clouds. *Advances in Neural Information Processing Systems*, 36.

Chen, H.; He, L.; Qian, W.; and Wang, S. 2020a. Multiple aerodynamic coefficient prediction of airfoils using a convolutional neural network. *Symmetry*, 12(4): 544.

Chen, H.; He, L.; Qian, W.; and Wang, S. 2020b. Multiple Aerodynamic Coefficient Prediction of Airfoils Using a Convolutional Neural Network. *Symmetry*, 12(4).

Chen, X.; Gong, C.; Wan, Q.; Deng, L.; Wan, Y.; Liu, Y.; Chen, B.; and Liu, J. 2021. Transfer learning for deep neural network-based partial differential equations solving. *Advances in Aerodynamics*, 3: 1–14.

Elrefaie, M.; Dai, A.; and Ahmed, F. 2024. DrivAerNet: A Parametric Car Dataset for Data-Driven Aerodynamic Design and Graph-Based Drag Prediction. *arXiv preprint arXiv:2403.08055*.

Elrefaie, M.; Morar, F.; Dai, A.; and Ahmed, F. 2024. DrivAerNet++: A Large-Scale Multimodal Car Dataset with Computational Fluid Dynamics Simulations and Deep Learning Benchmarks. *arXiv preprint arXiv:2406.09624*.

Huang, Z.; Wen, Y.; Wang, Z.; Ren, J.; and Jia, K. 2024. Surface reconstruction from point clouds: A survey and a benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Ignatyev, D. I.; and Khrabrov, A. N. 2015. Neural network modeling of unsteady aerodynamic characteristics at high angles of attack. *Aerospace Science and Technology*, 41: 106–115.

Jiahao, S.; Wenbo, C.; and Weiwei, Z. 2023. FD-PINN: Frequency domain physics-informed neural network. *Chinese Journal of Theoretical and Applied Mechanics*, 55(5): 1195–1205.

Karali, H.; Demirezen, M. U.; Yukselen, M. A.; and Inalhan, G. 2020. Design of a deep learning based nonlinear aerodynamic surrogate model for UAVs. In *AIAA Scitech 2020 forum*, 1288.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations*.

Lee, D. H.; Lee, D.; Han, S.; Seo, S.; Lee, B. J.; and Ahn, J. 2023. Deep residual neural network for predicting aerodynamic coefficient changes with ablation. *Aerospace Science and Technology*, 136: 108207.

Li, Z.; Zheng, H.; Kovachki, N.; Jin, D.; Chen, H.; Liu, B.; Azizzadenesheli, K.; and Anandkumar, A. 2024. Physics-informed neural operator for learning partial differential equations. *ACM/JMS Journal of Data Science*, 1(3): 1–27.

Liu, Y.; Liu, W.; Yan, X.; Guo, S.; and Zhang, C.-a. 2023. Adaptive transfer learning for PINN. *Journal of Computational Physics*, 490: 112291.

Peng, W.; Zhang, Y.; and Desmarais, M. 2021. Spatial convolution neural network for efficient prediction of aerodynamic coefficients. In *AIAA Scitech 2021 Forum*, 0277.

Qi, C. R.; Su, H.; Mo, K.; and Guibas, L. J. 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 652–660.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378: 686–707.

Rajkumar, T.; and Bardina, J. E. 2002. Prediction of Aerodynamic Coefficients Using Neural Networks for Sparse Data. In *Flairs*, 242–246.

Secco, N. R.; and Mattos, B. S. d. 2017. Artificial neural networks to predict aerodynamic coefficients of transport airplanes. *Aircraft Engineering and Aerospace Technology*, 89(2): 211–230.

Thirumalainambi, R.; and Bardina, J. 2003. Training data requirement for a neural network to predict aerodynamic coefficients. In *Independent component analyses, wavelets, and neural networks*, volume 5102, 92–103. SPIE.

Wang, X.; Kou, J.; and Zhang, W. 2022. Unsteady aerodynamic prediction for iced airfoil based on multi-task learning. *Physics of Fluids*, 34(8).

Wu, C.; Zheng, J.; Pfrommer, J.; and Beyerer, J. 2023. Attention-based point cloud edge sampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5333–5343.

Xiong, F.; Zhang, L.; Hu, X.; and Ren, C. 2022. A point cloud deep neural network metamodel method for aerodynamic prediction. *Chinese Journal of Aeronautics*.

Yetkin, S.; Abuhanieh, S.; and Yigit, S. 2024. Investigation on the abilities of different artificial intelligence methods to predict the aerodynamic coefficients. *Expert Systems with Applications*, 237: 121324.

Zeng, D.; Liu, W.; Chen, W.; Zhou, L.; Zhang, M.; and Qu, H. 2023. Substructure aware graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 11129–11137.

Zhao, J.; Zeng, L.; and Shao, X. 2023. A novel prediction method for unsteady aerodynamic force on three-dimensional folding wing aircraft. *Aerospace Science and Technology*, 137: 108287.