Mechanics of Advanced and Smart Materials Journal

http://masm.araku.ac.ir

ISSN: 2783-4220

Spatiotemporal Wake Learning via Three-Dimensional Convolution for Instantaneous Aerodynamic Load Estimation on an Elastically Mounted Square Cylinder

Ahmadreza Rabiee [a], Amir Hossein Rabiee [b]*, Mostafa Esmaeili [c]

[a] Department of Computer Engineering, Faculty of Engineering, Arak University, Arak 38181-75846, Iran
[b] School of Mechanical Engineering, Arak University of Technology, Arak 38181-41167, Iran
[c] Department of Mechanical Engineering, Faculty of Engineering, Kharazmi University, Tehran 15719-14911, Iran

Original Article

Use your device to scan and read the article online

Citation: Rabiee A, Rabiee A H, Esmaeili M. Spatiotemporal Wake Learning via Three-Dimensional Convolution for Instantaneous Aerodynamic Load Estimation on an Elastically Mounted Square Cylinder. Mechanics of Advanced and Smart Materials. 2026;5(4):638-659.

10.22034/masm.2026.742758

KEYWORDS

Flow-induced vibration,
Three-dimensional,
Convolutional neural network,
Lift coefficient prediction,
Fluid-structure interaction,
Square cylinder.

ABSTRACT

This study presents a unified three-dimensional convolutional neural network (3D CNN) for the direct prediction of the instantaneous lift coefficient in turbulent flow-induced vibration of a two-degree-of-freedom elastically mounted square cylinder. Physics-consistent data were generated using a validated two-way coupled URANS-FSI model with the k-ω SST turbulence closure. After excluding the initial transient response, 1017 synchronized wake-field snapshots and lift-coefficient values were collected over 101.7 s. Consecutive flow images were processed as a single spatio-temporal volume, allowing the Conv3D kernels to learn vortex morphology, shear-layer evolution, and short-term wake dynamics without recurrent units or reduced-order preprocessing. A chronological train-validation-test division was adopted to preserve temporal causality. Sensitivity analysis identified a sequence length of T = 18 and a batch size of 14 as the optimal configuration. The final network achieved R^2_{train} = 0.932 and R^2_{test} = 0.919, with approximate MAE = 0.19 and MSE = 0.06. Time-history comparisons showed accurate recovery of oscillation amplitude, extrema, zero crossings, and phase over the unseen test interval. These results demonstrate that direct three-dimensional convolution can provide an effective image-to-force surrogate for turbulent fluid-structure interaction.

Extended Abstract

1. Introduction

Flow-induced vibration (FIV) arises when unsteady fluid loading interacts with the inertia, stiffness, and damping of a compliant structure. Among its principal forms, vortex-induced vibration (VIV) is distinguished by the synchronization of vortex shedding with structural motion, which can produce large-amplitude oscillations and sustained cyclic loading [1,2]. For elastically mounted bodies, the response is governed by a nonlinear feedback loop: wake asymmetry generates fluctuating lift, the resulting motion alters separation and vortex formation, and the modified wake changes the subsequent aerodynamic force [3,4]. Accurate estimation of the instantaneous lift coefficient, C_L, is therefore important for fatigue assessment, vibration control, and energy-

* Corresponding author.
E-mail: rabiee@arakut.ac.ir
DOI: 10.22034/masm.2026.742758
Received: September 09, 2026; Received in revised form: October 07, 2026; Accepted: October 11, 2026.
© Author

harvesting design. Recent reviews also show that FIV is increasingly treated not only as a structural hazard but also as a mechanism for extracting low-power renewable energy [5].

The difficulty of predicting C_L becomes greater for sharp-edged bodies and turbulent operating conditions. Unlike a circular cylinder, a square cylinder has nominally fixed separation points at its leading corners, while the separated shear layers, corner vortices, and near-wake recirculation remain highly sensitive to structural motion and flow conditions. Recent investigations by Rabiee, Esmaeili, and co-workers have demonstrated that incidence angle, thermal buoyancy, surface condition, cylinder arrangement, and wake interference can substantially modify vibration amplitudes, synchronization ranges, and energy conversion [6-9]. Wind-tunnel studies have further confirmed that the response of interacting stationary and vibrating cylinders is strongly configuration-dependent [10], while rectangular plate–cylinder arrangements can intensify wake-induced oscillations and improve harvested power [11]. Machine-learning analysis of experimental FIV data has also shown that nonlinear output variables such as voltage and power can be predicted from coupled geometric and flow parameters [12]. These developments reinforce the need for efficient models that directly connect the evolving wake to the instantaneous force response.

High-fidelity computational fluid dynamics and two-way fluid–structure interaction solvers can resolve this coupling, but their repeated use in design exploration, uncertainty analysis, optimization, and real-time estimation remains computationally demanding. Machine learning has consequently become an important complementary tool in fluid mechanics [13]. Current research includes data-driven turbulence modeling, reduced-order representation, surrogate simulation, flow reconstruction, and accelerated numerical integration [14]. Hybrid learning strategies have even produced stable, high-speed approximations of governing-equation solvers for turbulent flows [15]. Nevertheless, model usefulness depends on whether the architecture can preserve the spatial organization and temporal development of coherent structures rather than merely fitting global statistical correlations.

Convolutional neural network (CNN) is particularly suitable for image-like flow data because their shared kernels identify local gradients, separated shear layers, vortex cores, and multiscale spatial patterns. CNN-based surrogates have predicted aerodynamic velocity and pressure fields for previously unseen conditions and geometries [16,17], reconstructed under-resolved turbulent fields [18], and produced compact nonlinear descriptions of wake dynamics [19]. Deep networks have also predicted the unsteady cylinder wake while incorporating conservation-related information [20]. Physics-informed neural network (PINN), provide another route by embedding differential equations into the training objective [21], and have been applied to two-degree-of-freedom cylinder VIV [22], and turbulent VIV reconstruction [23]. These methods confirm the potential of learning-based fluid–structure interaction (FSI) modeling, but they often require equation residuals, reduced variables, or separate stages for spatial and temporal processing.

Temporal prediction in VIV has commonly relied on recurrent networks. Long Short-Term Memory (LSTM) based reduced-order models can represent nonlinear wake evolution and structural response over successive time steps [24], while neural-network-based FSI solvers can replace part of the conventional flow solution and reproduce lock-in behavior at reduced computational cost [25]. More recently, Rabiee et al. directly predicted turbulent VIV lift and drag coefficients for a freely vibrating circular cylinder using an LSTM architecture [26]. Although recurrent networks are effective for sequential data, they typically compress each frame into a feature vector before processing temporal dependence through memory units. This separation increases architectural complexity and may weaken the direct association between local vortex evolution and the aerodynamic force generated at the following instant.

Three-dimensional convolutional neural network (3D CNN) offers a different formulation. A Conv3D kernel moves simultaneously through time and the two spatial directions, allowing local wake morphology and its short-term evolution to be encoded within the same operation. Previous studies have shown that 3D CNNs can forecast high-dimensional fluid states [27], while convolutional models can infer turbulent fields from indirect measurements [28]. Physics-constrained autoregressive networks [29], and deep convolutional generative models [30], further demonstrate that temporally ordered flow fields can be learned without relying exclusively on recurrent units. However, direct force regression from complete wake-image sequences using a fully 3D convolutional architecture remains comparatively underexplored, particularly for turbulent VIV of a two-degree-of-freedom square cylinder.

Accordingly, the present study develops a unified end-to-end 3D CNN for predicting the instantaneous lift coefficient from consecutive wake-field images generated by a validated two-way coupled FSI simulation based on the unsteady Reynolds-averaged Navier–Stokes (URANS) equations. The principal novelty is the direct and joint extraction of temporal and spatial information from the complete input volume, without an LSTM, recurrent cell, or separately trained reduced-order representation. The model is applied to turbulent VIV of an elastically mounted square cylinder, for which fixed corner separation and strong wake asymmetry create a demanding force-prediction problem. Sequence length and batch size are selected through sensitivity analysis, and the final network is assessed using a chronological training–validation–test division to preserve temporal causality. The proposed formulation therefore establishes a direct image-to-force mapping that relates short-horizon vortex development to C_L within a single three-dimensional convolutional framework.

The conceptual distinction of the proposed architecture lies in how spatial and temporal information is integrated. In conventional CNN–LSTM formulations, a CNN typically converts each flow snapshot into a spatial feature representation, which is subsequently processed by recurrent units to describe temporal evolution. In contrast, a 3D convolutional kernel operates jointly across two spatial dimensions and a finite temporal window, learning local patterns of wake evolution directly. Although recurrent architectures can represent longer temporal dependencies, the present formulation targets short-term flow evolution associated with the instantaneous lift response.

The approach also differs from reduced-order learning pipelines that first project the flow onto a separately constructed low-dimensional basis and then learn the dynamics of the resulting coefficients. Here, feature extraction and force prediction are trained jointly from image sequences without a separate modal decomposition. Moreover, using CFD-generated training data does not make the model physics-informed in the sense of explicitly enforcing governing equations: the loss minimizes prediction error without equation residuals or physical constraints.

The expected advantage is the joint representation of local wake morphology and its short-term evolution within a single convolutional feature extractor. The contribution is the application and evaluation of this formulation for the present turbulent FSI problem; superiority in accuracy or computational efficiency over CNN–LSTM, reduced-order, or physics-informed alternatives is not claimed without a controlled comparative assessment.

2. Methodology

This study investigates a fully 3D CNN for the direct prediction of the instantaneous lift coefficient in the turbulent, vortex-induced vibration of a two-degree-of-freedom, elastically mounted square cylinder. The proposed network convolves jointly across the temporal and spatial axes of the input sequence: each 3D kernel slides simultaneously along time, height, and width, so that spatial patterns, such as vortex shape and shear-layer curvature, and their short-term temporal evolution are learned within a single convolutional operation. This joint spatio-temporal formulation allows the network to directly associate the evolving wake structure with the instantaneous aerodynamic response, within a single, unified, end-to-end architecture. Physics-consistent training data are obtained from a validated, two-way coupled numerical simulation of the (FSI) problem, as described in Section 2.1

2.1 Physics-Based Framework and Dataset Generation

The physical problem considered is the two-dimensional, two-way coupled FSI response of a square cylinder elastically mounted with two degrees of freedom (in-line and cross-flow), immersed in a turbulent, incompressible flow. The surrounding fluid is governed by the URANS equations,

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial(\rho u_i u_j)}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \rho\overline{u_i' u_j'}\right] \tag{2}$$

with the Reynolds stress closed via the Boussinesq approximation. Turbulence is modeled using the k-ω Shear Stress Transport (SST) model [31], selected for its established capability in resolving flow separation, vortex shedding, and near-wall behavior around bluff bodies. The Reynolds number is defined as Re = ρ U D /μ, with U

the free-stream velocity and D the cylinder side length; the computational domain extends 20 D upstream, 50 D downstream, and 40 D in height, as shown in Fig. 1 together with the mesh topology and boundary conditions.

At the inlet, a uniform flow is prescribed with $(u_x = U_\infty)$ and $(u_y = 0)$. At the outlet, zero streamwise gradients are imposed for both velocity components, expressed as $(\partial u_x/\partial x = 0)$ and $(\partial u_y/\partial x = 0)$. The upper and lower boundaries satisfy slip and no-penetration conditions, defined by $(\partial u_x/\partial y = 0)$ and $(u_y = 0)$. On the moving cylinder surface, the no-slip condition requires the fluid velocity to match the cylinder velocity, such that $(u_x = \dot{x})$ and $(u_y = \dot{y})$.

In Eq. (2), u_i and u_j are the mean velocity components along the spatial coordinates x_i and x_j, t is time, ρ is the fluid density, p is the mean pressure, and μ is the dynamic viscosity. The quantities u'_i and u'_j denote velocity fluctuations; $-\rho\,\overline{u'_i u'_j}$ represents a Reynolds-stress component, with the overbar indicating averaging. The indices i and j refer to the two spatial directions, and repeated indices imply summation.

The flow simulations were performed using ANSYS Fluent 6.3.26 and the finite volume method. The computational mesh contained 69,050 cells. The minimum and maximum cell-edge lengths in the initial mesh were 0.005602D and 1.167281D, respectively, where D is the side length of the square cross section.

The cylinder is modeled as a rigid body mounted on a linear spring–damper system with two structural degrees of freedom:

The cylinder has two translational degrees of freedom, in the streamwise and transverse directions. Rotation about its longitudinal axis is constrained, so the square cross section retains a fixed orientation. The support stiffness and damping are equal in both directions: $k_x = k_y = k$ and $c_x = c_y = c$. In the dataset-generation simulation, the mass ratio is $m^* = 2.4$ and the mass–damping parameter is $m^*\zeta = 0.013$. The natural frequency and damping ratio are defined as $f_n = \sqrt{(k/m)}/(2\pi)$ and $\zeta = c/(2\sqrt{(km)})$, respectively, with equal natural frequencies in the two directions. Here, m, k, and c are the structural mass, stiffness, and damping per unit span.

For the square cross section, the mass ratio is defined as $m^* = m/(\rho D^2)$, where m is the structural mass per unit span, ρ is the fluid density, and D is the side length [33]. The reduced velocity is $U_r = U_\infty/(f_n D)$, where U_∞ is the inlet velocity and f_n is the structural natural frequency.

$$\begin{aligned} m\ddot{x} + c\dot{x} + kx &= F_D(t), \\ m\ddot{y} + c\dot{y} + ky &= F_L(t) \end{aligned} \tag{3}$$

where F_D and F_L are the instantaneous drag and lift forces obtained from surface integration of the pressure and viscous stresses, giving the force coefficients

$$\begin{aligned} C_L &= \frac{2F_L}{\rho U^2 D}, \\ C_D &= \frac{2F_D}{\rho U^2 D} \end{aligned} \tag{4}$$

A strongly coupled, two-way FSI algorithm is used to advance the solution, as illustrated in Fig. 2: at each time step, the URANS equations are solved to obtain the instantaneous flow field and hydrodynamic forces, which are transferred to the structural solver; the resulting displacement and velocity update the cylinder position, and the computational mesh is deformed accordingly before the next flow solution. The numerical framework was validated against reference VIV data from the literature prior to dataset generation.

The numerical FSI framework was assessed using a spring-supported cylinder with two translational degrees of freedom, a mass ratio of 2.7325, and zero structural damping in turbulent flow. Table 1 compares the dimensionless transverse amplitude Y/D with the numerical data of Zhu et al. [32] at 11 reduced velocities from 4 to 14. Both datasets reach their maximum at $U_r = 12$. The peak amplitudes are 0.524242 and 0.578788 for the present and reference results, respectively, corresponding to a relative difference of approximately 9.4%. Differences are larger at some velocities, particularly on the post-peak branch; the comparison therefore supports reproduction of the overall response trend and peak location.

The final dataset was generated at a fixed inlet velocity in two-dimensional, incompressible turbulent flow, using the URANS equations with the k–ω SST model and two-way coupling to the structural motion. The structural parameters specified above apply to dataset generation and are distinct from the validation conditions. This

operating condition was selected to provide evolving wake structures and a corresponding time-varying lift response for evaluating joint spatial–temporal feature extraction under a fixed physical configuration.

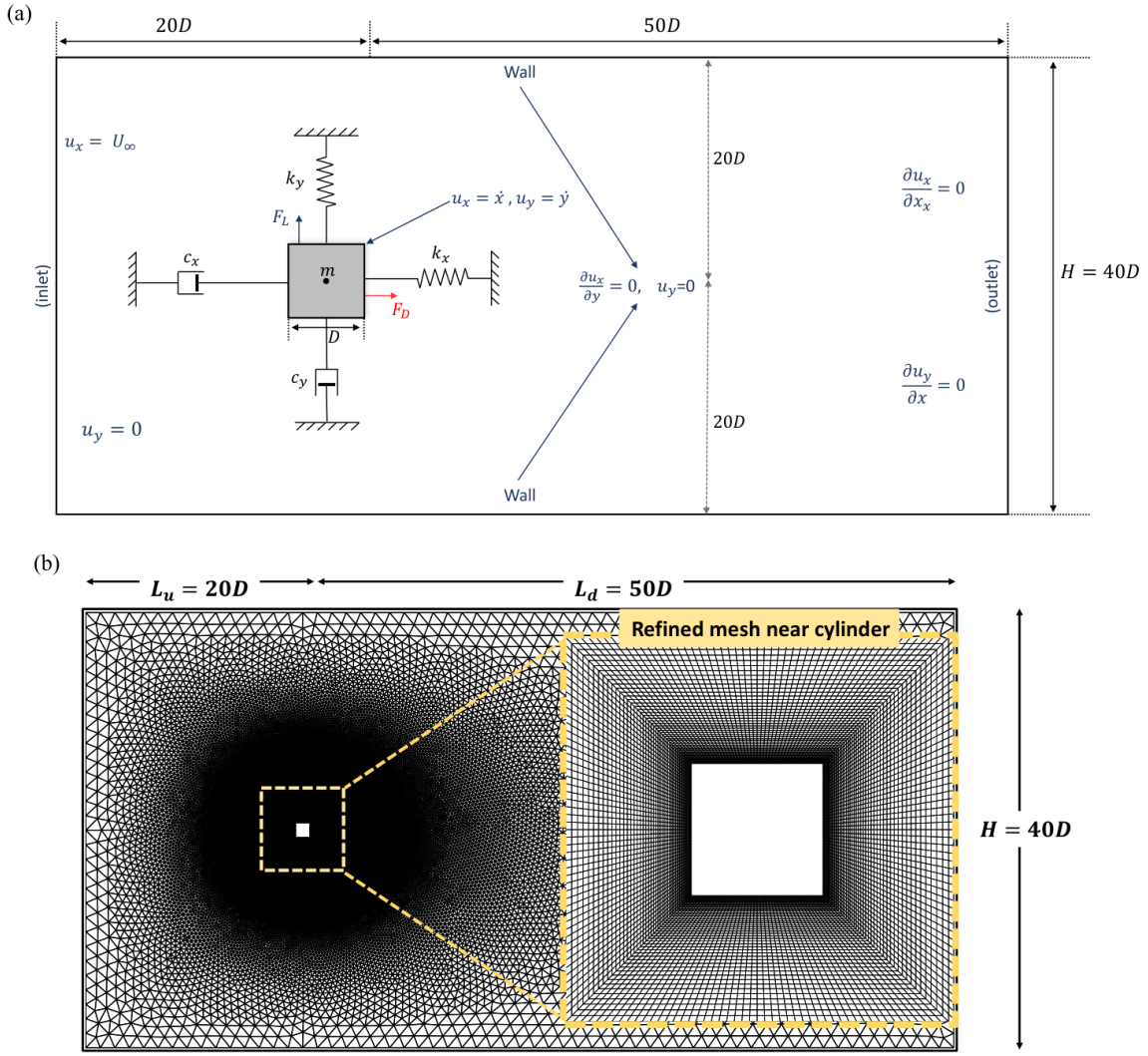


Figure 1. Computational domain, mesh, and boundary conditions

Table 1. Comparison of transverse response amplitudes at reduced velocities 4–14 for the validation case (m* = 2.7325, zero structural damping, two translational degrees of freedom).

U_r	Zhu et al. [32] Y/D	Present Y/D	Present − reference
4	0.013636	0.015152	0.001516
5	0.013636	0.022727	0.009091
6	0.031818	0.051515	0.019697
7	0.074242	0.133333	0.059091
8	0.157576	0.204545	0.046969
9	0.256061	0.293939	0.037878
10	0.501515	0.380303	-0.121212
11	0.546970	0.498485	-0.048485
12	0.578788	0.524242	-0.054546
13	0.106061	0.212121	0.106060
14	0.081818	0.201515	0.119697

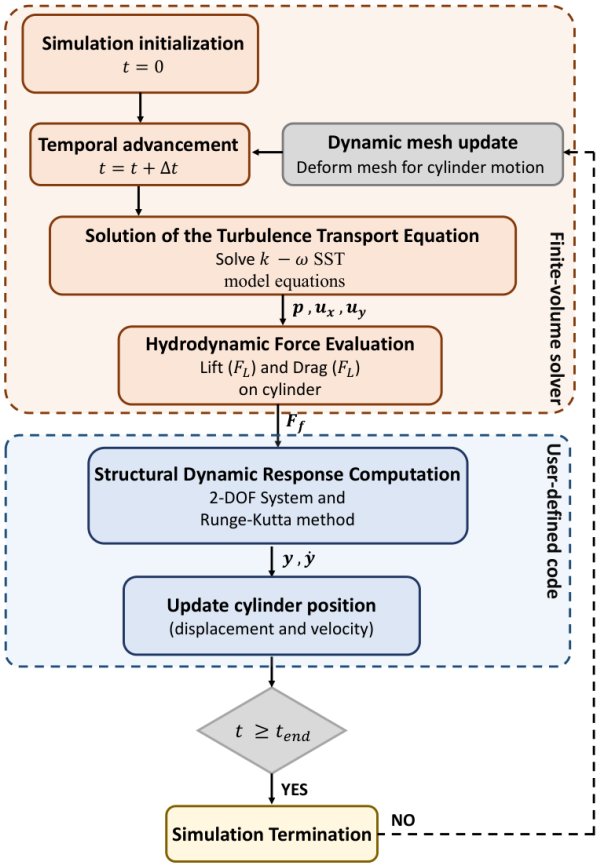


Figure 2. Two-way FSI coupling algorithm flowchart

Following the validated FSI simulation, a time-accurate run was performed at a fixed inflow velocity within the turbulent VIV regime. The initial period (t < 10 s), during which the structural response had not yet built up and the flow field exhibited only limited variation, was excluded from the dataset. A total of 1017 flow-field snapshots were then uniformly sampled over the physical time interval t = 10–111.7 s (a duration of 101.7 s), together with the synchronized instantaneous lift coefficient at each corresponding time index. The flow-field images were geometrically aligned, cropped to remove inactive boundary regions, and resized to a common spatial resolution; pixel intensities were normalized. The lift-coefficient signal was filtered using a fourth-order Butterworth low-pass filter with a 0.9 Hz cutoff, implemented through forward-backward filtfilt filtering with odd-symmetric endpoint padding. Comparison of the raw and filtered Hann-windowed spectra revealed an unchanged principal peak at 0.454 Hz, retaining 99.64% of its spectral amplitude. The displacement spectrum peaked at 0.464 Hz, supporting force–motion frequency consistency. These results indicate preservation of the dominant oscillatory component, although localized endpoint effects remain possible; the comparison is not an independent measurement of vortex-shedding frequency and the C_L signal was smoothed with a low-pass filter to suppress high-frequency numerical noise without introducing phase distortion. Fig. 3 shows the resulting time history of the lift coefficient together with representative instantaneous flow-field snapshots, illustrating the dynamic diversity of the generated dataset.

The network input consists of sequences of color contour images of the vorticity field. Each image contains three red, green, and blue (RGB) channels that encode the color representation of the same vorticity field, rather than separate physical fields. The chronological division and within-subset window construction are described in Section 2.3.

2.2 Problem Formulation and Network Architecture

The prediction of the instantaneous lift coefficient is formulated as a supervised spatio-temporal regression problem:

$$\mathcal{F}: I_{t-T+1}, I_{t-T+2}, \ldots, I_t \rightarrow C_L(t+1) \quad (5)$$

where I_t denotes the flow-field image at time t, and T is the length of the input temporal window. The entire input sequence is treated as a single four-dimensional tensor of shape (T, H, W, C) and processed directly by a stack of 3D convolutional layers, so that spatial and temporal features are learned jointly from the input volume. The sequence length T was treated as a tunable hyperparameter and selected, together with the batch size, through the sensitivity analysis described in the following section. Based on this analysis, a sequence length of T = 18 was adopted, so that the network input tensor has dimensions 18 × 117 × 293 × 3. Based on this problem formulation, a dedicated 3D CNN was designed to learn the spatio-temporal relationship between wake evolution and the instantaneous lift coefficient.

The proposed architecture, shown in Fig. 4, consists of three functional stages: an input stage, a 3D convolutional feature extractor, and a fully connected regression head.

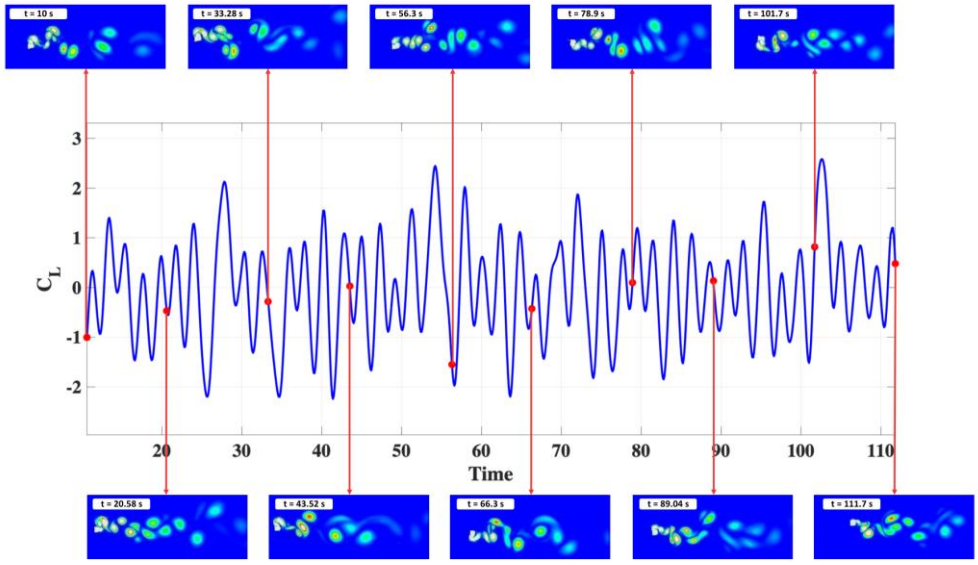


Figure 3. Reconstructed time history of C_L with representative flow-field snapshots

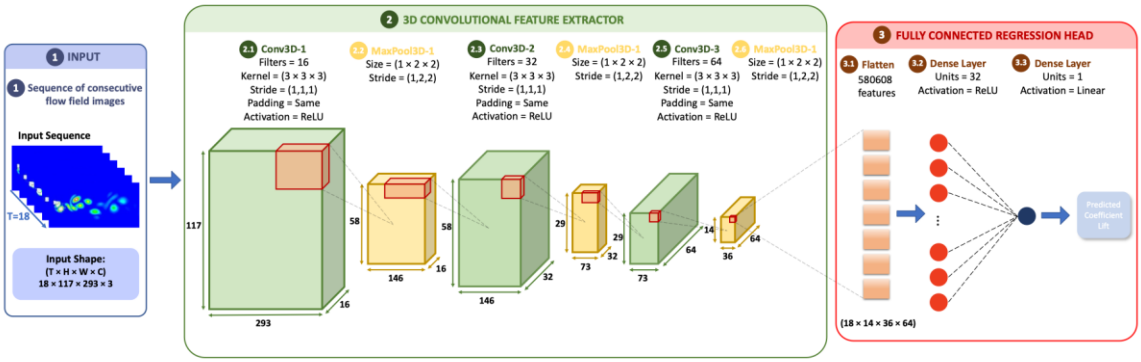


Figure 4. Three-Dimensional CNN Architecture

The 3D convolution operation generalizes standard 2D convolution by sliding a kernel simultaneously along the temporal and the two spatial axes:

$$y_{t,i,j}^{(l,f)} = \phi\left(b^{(l,f)} + \sum_{c}\sum_{\tau=1}^{3}\sum_{p=1}^{3}\sum_{q=1}^{3} W_{\tau,p,q}^{(l,f,c)}, x_{t+\tau,,i+p,,j+q}^{(l-1,c)}\right) \quad (6)$$

where $W^{l,f,c}$ denotes the 3 × 3 × 3 kernel connecting input channel c to output feature map f at layer l, and φ(·) is the ReLU activation function.

The feature extractor comprises three sequential Conv3D-MaxPool3D blocks. All convolutions use 3 × 3 × 3 kernels, unit stride, "same" padding, and ReLU activation, and are each followed by 3D max-pooling with a (1,2,2) window and stride, which down samples only the spatial dimensions while leaving the temporal axis unchanged. The first block applies 16 filters (117 × 293 → 58 × 146), the second block applies 32 filters (→ 29 × 73), and the third block applies 64 filters (→ 14 × 36); the temporal dimension remains at T = 18 throughout, since pooling is never applied along the time axis.

The resulting feature volume, of size 18 × 14 × 36 × 64 (580,608 elements), is flattened into a single feature vector and passed to the regression head, consisting of a dense layer with 32 units and ReLU activation, followed by a single linear output unit that produces the predicted lift coefficient, $C_L(t + 1)$. Neither dropout nor batch normalization is applied within the convolutional feature extractor. After defining the network architecture, the training strategy, hyperparameter selection procedure, and evaluation methodology are described in the next section.

The input sequence length and the local temporal receptive field describe different scales. T = 18 is the number of frames supplied to the network, while each feature after the three convolutional layers has a nominal temporal receptive field of seven frames, calculated as $1 + 3(3 - 1) = 7$. With "same" padding and no temporal pooling, all 18 temporal positions are retained, although boundary features include padded values. The resulting feature volume is flattened and passed to the dense layer, which combines features across all temporal positions. Thus, the final prediction can use information from the entire 18-frame sequence, even though each individual convolutional feature captures a local seven-frame context. The physical temporal span of a kernel depends on the sampling interval (Δt); for a three-frame kernel, this span is $(2\Delta t)$. Nevertheless, the kernel dimensions are a design choice for local feature extraction and need not scale automatically with sampling frequency. The $(3 \times 3 \times 3)$ kernels are retained as compact local operators, with broader temporal context provided by the stacked convolutions and the Flatten–Dense head. For a different sampling interval, the total input duration, $\big((T - 1)\Delta t\big)$, can be adjusted through (T), although this does not preserve the local physical receptive field. The suitability of the retained kernel dimensions at other sampling rates should therefore be assessed empirically.

The model has a unified end-to-end architecture comprising three convolutional–pooling blocks followed by a regression head. Its architectural simplicity lies in integrating spatial–temporal feature extraction and lift prediction without recurrent units or a separately constructed reduced-order representation. However, connecting the 580,608-element flattened feature volume to 32 dense units introduces approximately 18.58 million weights in that connection alone. Thus, a simple processing pipeline should not be interpreted as a small parameter count or low computational cost. Computational savings relative to recurrent or multi-stage alternatives have not been established through a controlled benchmark.

2.3 Training and Evaluation

The 1017 synchronized samples were divided into contiguous chronological subsets: approximately 70% for training, 15% for validation, and 15% for testing. The earliest samples formed the training subset, followed by validation and then testing, without random shuffling. The split was performed before constructing temporal windows. Each input sequence and its corresponding next-step target were confined to the same subset; windows crossing subset boundaries were excluded. These percentages refer to the original snapshots, before window construction. This procedure prevents cross-subset sharing of snapshots through overlapping windows while preserving temporal order; it does not imply complete statistical independence between segments of the same simulation. Specifically, the chronological partitions contain 711 training snapshots (indices 1–711), 152 validation snapshots (indices 712–863), and 154 test snapshots (indices 864–1017). For (T=18) input frames, a next-step target, and a sliding stride of one frame, a subset containing $\big(N_j\big)$ snapshots yields $\big(N_j - T\big)$ windows. The final configuration therefore contains 693 training, 134 validation, and 136 test windows, totalling 963 windows. Each window and its corresponding target are confined to their own subset, and windows crossing partition boundaries are excluded. The test block follows both the training and validation blocks chronologically.

The network was trained by minimizing the mean squared error between the predicted and ground-truth lift coefficients:

In Eq. (7), N is the number of samples used to evaluate the loss, i is the sample index, $C_{L,actual}^{(i)}$ is the lift coefficient obtained from the numerical simulation, and $C_{L,pred}^{(i)}$ is the network prediction for the same sample.

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(C_{L,actual}^{(i)} - C_{L,pred}^{(i)})^2 \quad (7)$$

using the Adam optimizer with its default learning rate of 0.001. Unlike the sequence-length and batch-size hyperparameters, the learning rate was not subjected to a separate sensitivity sweep, as this default value already yielded the best observed performance. Training was carried out for a maximum of 100 epochs, with early stopping monitored on the validation loss to prevent overfitting; the network weights corresponding to the best validation performance were retained. The dataset was not shuffled during training, and sliding temporal windows were used to construct the input–output pairs specified in Eq. (5), so that no future information was accessible during training.

The sequence length T and batch size were treated as tunable hyperparameters and determined through a sequential, rather than joint, sensitivity analysis. This ordering reflects a fundamental distinction between the two parameters: T determines the dimensionality of the input tensor and, consequently, the resulting network architecture, since the size of the flattened feature vector scales directly with T, whereas the batch size only affects the optimization dynamics of an already-fixed architecture. Accordingly, T was varied first, from 2 to 20 frames, using a fixed baseline batch size of 2, in order to establish the temporal window that provides the network with sufficient information on the vortex-shedding dynamics; this step identified T = 18 as optimal. The batch size was then varied from 2 to 20 with T fixed at 18, identifying a batch size of 14 as optimal for this architecture. The resulting configuration, T = 18 and a batch size of 14, was adopted for the final model reported here; the complete sensitivity analysis supporting this choice is presented in Section 3. The predictive performance of the trained model was subsequently evaluated using several standard regression metrics, as described below.

Model performance was quantified using three complementary metrics: the mean squared error defined in Eq. (7), the mean absolute error,

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|C_{L,actual}(i) - C_{L,pred}(i)\right| \quad (8)$$

and the coefficient of determination,

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(C_{L,actual}(i) - C_{L,pred}(i)\right)^2}{\sum_{i=1}^{N}\left(C_{L,actual}(i) - \overline{C_{L,actual}}\right)^2} \quad (9)$$

R^2 is regarded as the primary indicator, as it quantifies the fraction of the physical variance in the lift signal captured by the model, and is used throughout this study both to select hyperparameters and to evaluate overall predictive performance.

3. Results and Discussion

Figure 5 presents the sequential sensitivity analysis used to determine the temporal sequence length and batch size of the 3D CNN. Panels (a) and (b) show the variations of MAE, MSE, and R^2 with the sequence length T, whereas panels (c) and (d) show the corresponding changes with batch size after fixing the selected temporal window. The results demonstrate that the amount of recent wake information supplied to the network has a strong effect on the prediction of C_L. At T = 2, the model produces approximately MAE = 0.80 and MSE = 0.98, while R^2 is close to −0.40. This negative value indicates that a two-frame input does not provide sufficient information for the network to establish a useful relationship between wake evolution and the instantaneous lift coefficient. Increasing T progressively decreases both errors and increases R^2. At approximately T = 6, R^2 approaches zero, after which the model begins to explain an increasingly meaningful fraction of the signal variance. This improvement is consistent with the seven-frame local temporal receptive field produced by the three successive convolutions, calculated as $(r_t = 1 + 3(3-1) = 7)$. For $(T < 7)$, part of each local receptive field is occupied by padding, whereas for $(T \geq 7)$, interior features can incorporate seven actual flow-field frames. Providing this complete local temporal context enables a richer representation of wake evolution and supports the learning of its relationship with the lift coefficient. The improvement beyond $(T = 6)$ therefore supports the importance of supplying sufficient temporal information to utilize the network's local feature-extraction capacity. The Flatten–Dense head further integrates features across the entire input sequence, with $(T = 18)$ providing the best performance among the tested sequence lengths. The best sequence-length result is obtained at T = 18, where MAE and MSE fall to approximately 0.40 and 0.24, respectively, and R^2 reaches about 0.65. Extending the window to T = 20 causes a small deterioration, indicating that the additional frames increase the input dimensionality without

adding equally useful dynamical information. The batch-size analysis shows a second optimum. Increasing the batch size from 2 to 14 reduces MAE from about 0.40 to 0.19 and MSE from about 0.24 to 0.06, while R^2 rises from approximately 0.65 to 0.92. Larger batches gradually reduce the accuracy, with R^2 decreasing to about 0.88 at a batch size of 20. Therefore, T = 18 and a batch size of 14 were selected for the final network. This combination provides enough temporal coverage to describe short-horizon vortex development while maintaining effective gradient updates during `optimization.

Short input sequences provide limited information about the evolution of vortex shedding and wake asymmetry associated with the lift response. Increasing T supplies additional temporal context, helping the network distinguish flow structures at different stages of their development. The improvement up to T = 18 is consistent with this interpretation. Extending the sequence further may introduce largely redundant information while increasing the flattened feature dimension and the number of dense-layer parameters, making learning more demanding for the available dataset.

Batch size influences optimization rather than the physical temporal context. The improvement up to a batch size of 14 may reflect more stable gradient estimates. Larger batches reduce the number of parameter updates per epoch and may alter convergence and generalization, potentially explaining the slight deterioration observed. These interpretations have not been independently verified; the selected values are specific to the present dataset and training configuration.

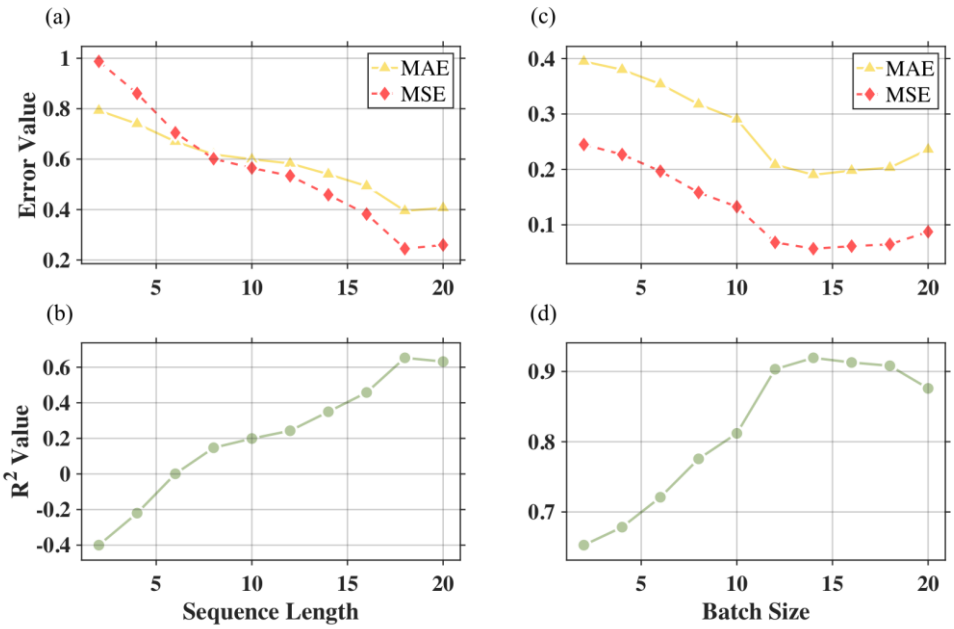


Figure 5. Sensitivity analysis of the 3D CNN hyperparameters: variations of (a) MAE and MSE and (b) R^2 with input sequence length T; variations of (c) MAE and MSE and (d) R^2 with batch size after fixing T = 18.

Figure 6 shows the regression behavior of a relatively low-accuracy network by plotting the predicted C_L values against the corresponding numerical values for the training and test subsets. The model achieves R^2_{train} = 0.690 and R^2_{test} = 0.653, meaning that it explains 69.0% of the training variance and 65.3% of the test variance. The point cloud follows the general one-to-one trend, confirming that the network learns a meaningful relationship between the wake images and the aerodynamic response. However, the broad distribution around the diagonal reveals considerable pointwise error. Predictions near moderate values of C_L are generally closer to the reference line, whereas the deviations increase near large positive and negative amplitudes. The small generalization gap, $R^2_{train} - R^2_{test}$ = 0.037, suggests that the limited accuracy is not mainly caused by severe overfitting. Instead, this configuration appears to provide an incomplete representation of the nonlinear spatio-temporal mapping between vortex evolution and lift generation. The similar train and test regression trends indicate that the model learns a repeatable, but insufficiently precise, approximation of the underlying dynamics.

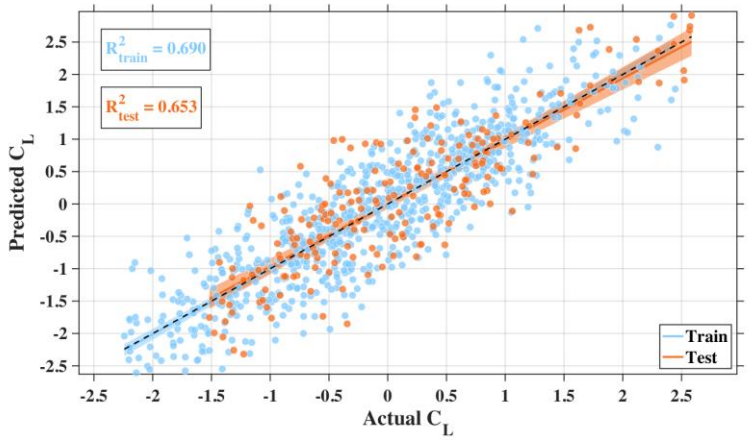


Figure 6. Regression plots comparing the numerical and predicted instantaneous lift coefficients, C_L, for the training and test datasets using the representative lower-accuracy 3D CNN configuration.

Figure 7 compares the actual and predicted histories of C_L for the same low-accuracy configuration. Panel (a) presents the training samples, and panel (b) presents the chronologically later test samples. In the training interval, the predictions reproduce the broad oscillatory structure and many changes of sign, but they fluctuate substantially around the actual curve. Several local maxima and minima are overestimated or underestimated, and isolated predictions depart markedly from the target signal. The test results show a similar pattern. The network identifies the dominant positive peak near the middle of the test segment and captures the approximate positions of several negative minima and later oscillations, but many individual estimates remain displaced from the numerical values. The largest deviations tend to occur near sharp extrema and rapidly changing portions of the signal. This behavior indicates that the network recognizes the principal vortex-shedding cycle but does not resolve the local amplitude modulation with adequate precision. Nevertheless, the predictions do not exhibit a persistent positive or negative bias, and the oscillatory structure is not lost. The principal weakness is therefore local amplitude error rather than complete failure to detect the force dynamics. Figures 6 and 7 together show that a model can recover the overall trend of C_L while still producing substantial instantaneous errors.

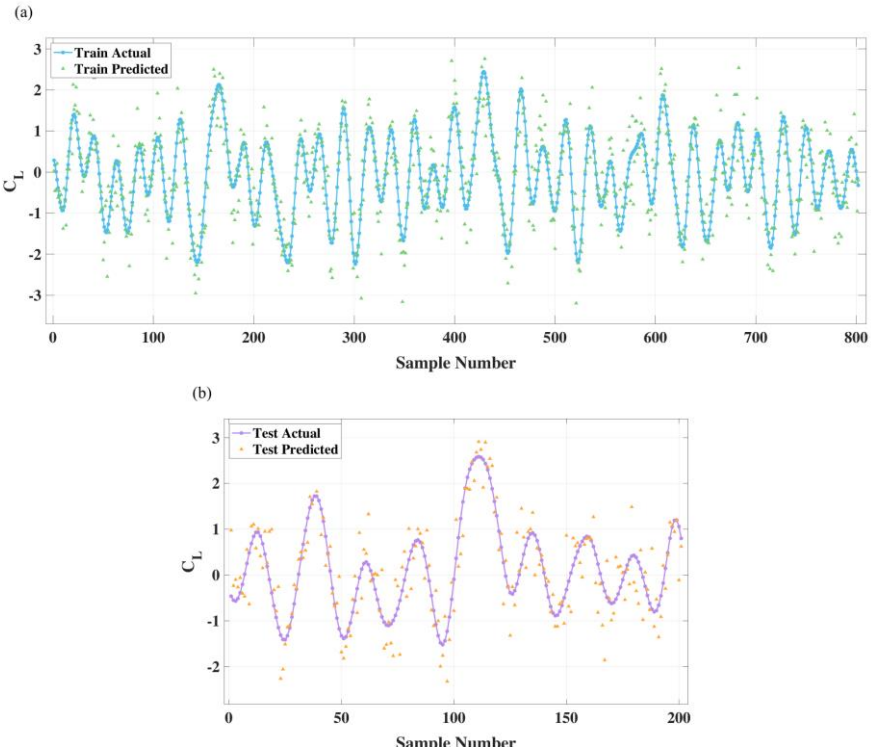


Figure 7. Time-history comparison of the numerical and predicted lift coefficients, C_L, obtained using the representative lower-accuracy configuration: (a) training dataset and (b) test dataset.

Figure 8 presents the regression plot for the network with intermediate accuracy. The coefficients increase to R^2_{train} = 0.789 and R^2_{test} = 0.756. Compared with Figure 6, the samples are more concentrated around the one-to-

one line and the fitted training and test trends more closely follow the ideal relationship. The improvement in test performance from 0.653 to 0.756 reduces the unexplained fraction of the test variance from $1 - 0.653 = 0.347$ to $1 - 0.756 = 0.244$. The improvement occurs for both positive and negative C_L values, showing that it is not limited to one direction of the transverse force. Moderate-amplitude samples display the strongest concentration around the diagonal, although some scatter remains near the largest positive and negative values. The generalization gap is $R^2_{train} - R^2_{test} = 0.033$, which is small and slightly lower than that of the previous configuration. Consequently, the improvement is retained on unseen data rather than being restricted to the training subset. The model therefore captures a larger portion of the nonlinear relationship between wake morphology and lift, although some local features associated with instantaneous amplitude remain unresolved.

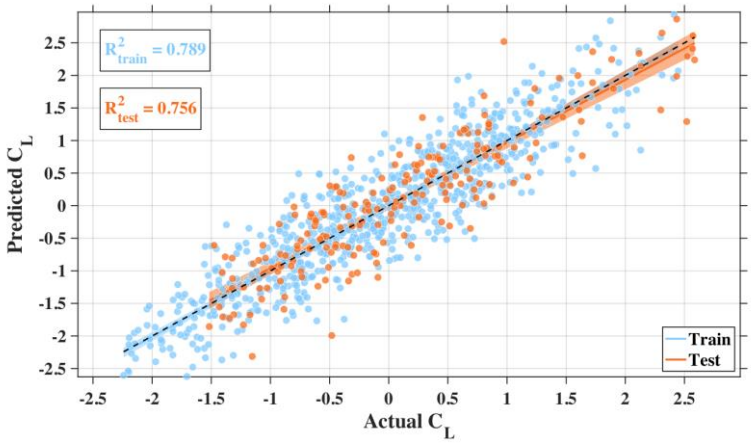


Figure 8. Regression plots comparing the numerical and predicted instantaneous lift coefficients, C_L, for the training and test datasets using the representative intermediate-accuracy 3D CNN configuration.

Figure 9 shows the corresponding training and test histories for the intermediate-accuracy network. In panel (a), the predicted training values follow the main oscillatory pattern more consistently than those in Figure 7. Most major peaks and troughs occur at the correct sample positions, and the spread of the predictions around the numerical curve is reduced. The improvement is especially visible in intervals containing repeated moderate-amplitude oscillations, where the predicted values remain close to the target over several consecutive samples. Some isolated errors persist near abrupt extrema and intervals with rapid amplitude changes. Panel (b) shows that the model also preserves the overall waveform throughout the test interval. The large positive response near sample 110, the preceding negative trough, and the subsequent lower-amplitude oscillations are all identified. Compared with the low-accuracy configuration, fewer predictions lie far from the reference curve, and the rising and falling portions of the signal are tracked more closely. However, local overprediction and underprediction remain visible around several extrema, explaining why R^2_{test} performance is not controlled by only a few outliers. It reflects a generally improved representation of the signal with moderate residual errors distributed across the sample sequence.

Figure 10 displays the regression results for the high-accuracy network obtained after hyperparameter selection. The model reaches $R^2_{train} = 0.932$ and $R^2_{test} = 0.919$, corresponding to 93.2% and 91.9% of the variance explained in the training and test signals. The samples form a narrow distribution around the ideal diagonal over nearly the complete range of C_L. The fitted trends for the two subsets are also close to each other and to the one-to-one relation, while the regression bands are noticeably narrower than those of the lower-performing networks. The generalization gap decreases to $R^2_{train} - R^2_{test} = 0.013$. This small difference indicates that the final network does not obtain its accuracy by simply memorizing the training samples. Instead, it learns spatio-temporal wake features that remain informative during the later test interval. The residual deviations are mainly isolated and are not accompanied by a persistent offset, a systematic slope error, or an obvious asymmetry between positive and negative values. The network therefore estimates both the sign and magnitude of C_L with high fidelity. Since the input is processed using 3D kernels rather than a separate recurrent module, this result confirms that joint convolution across time and space can extract the short-term wake evolution required for accurate lift prediction. The selected sequence length T = 18 allows the network to distinguish between visually similar wake states whose vortices are evolving in different temporal directions.

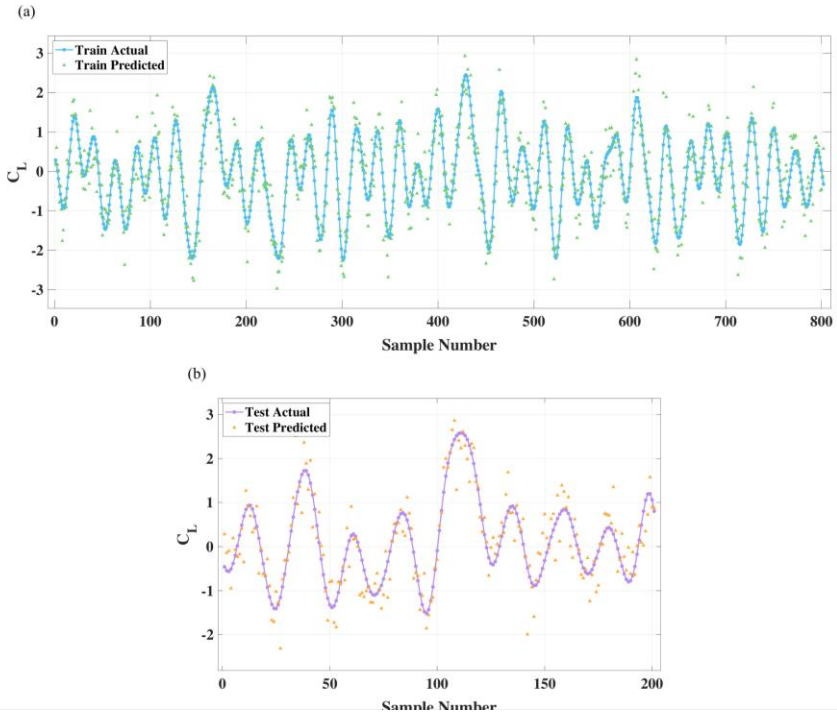


Figure 9. Time-history comparison of the numerical and predicted lift coefficients, C_L, obtained using the representative intermediate-accuracy configuration: (a) training dataset and (b) test dataset.

Figure 11 compares the actual and predicted C_L histories for the high-accuracy model. Panel (a) shows that the training predictions closely follow the numerical signal over the entire sample range. The model reproduces the alternating positive and negative oscillations, changes in amplitude, local maxima, local minima, and most zero crossings. The agreement is maintained during intervals with relatively large amplitudes, showing that the network is not accurate only near the mean of the signal. Panel (b) demonstrates similarly strong performance for the test data. The large positive peak near the center of the interval, the preceding negative response, and the subsequent lower-amplitude cycles are all reproduced with close temporal alignment. No progressive phase drift is visible as the sample number increases. This is important because even a small phase error can cause substantial pointwise disagreement in an oscillatory signal despite an approximately correct frequency. A limited number of isolated overpredictions and underpredictions remain, particularly near some extrema, but these errors are local and do not alter the overall waveform. The time-domain agreement therefore provides a direct explanation for $R^2_{test} = 0.919$. The high coefficient is supported not only by a correct overall distribution of C_L but also by accurate tracking of its instantaneous evolution.

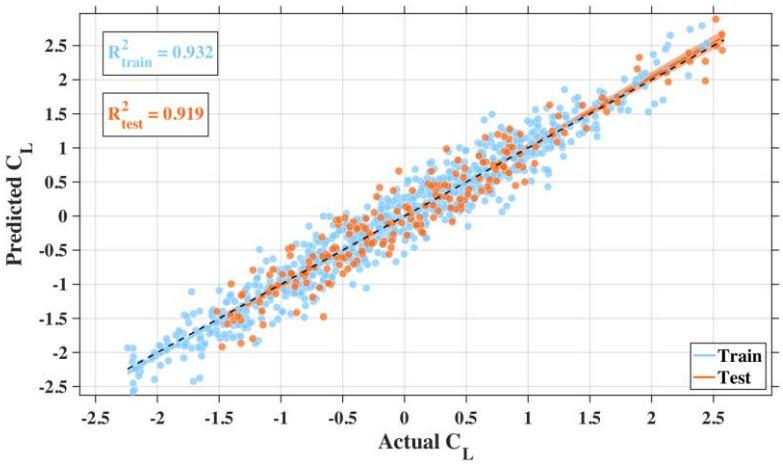


Figure 10. Regression plots comparing the numerical and predicted instantaneous lift coefficients, C_L, for the training and test datasets using the optimized high-accuracy 3D CNN configuration.

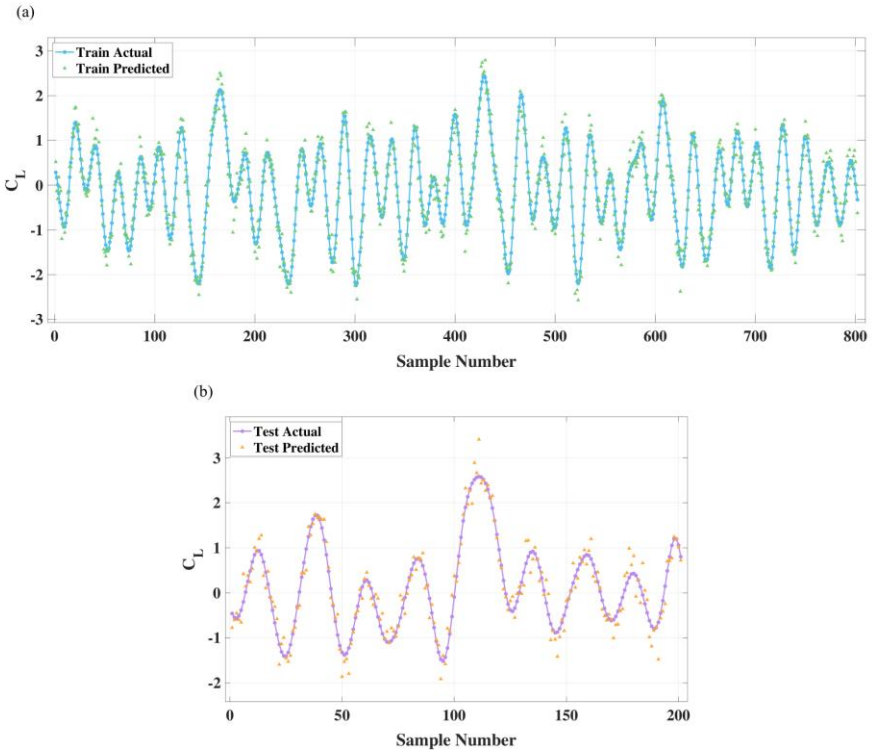


Figure 11. Time-history comparison of the numerical and predicted lift coefficients, C_L, obtained using the optimized high-accuracy 3D CNN configuration: (a) training dataset and (b) test dataset.

Taken together, Figures 5–11 show a systematic progression from an incomplete approximation of the lift dynamics to a highly accurate 3D CNN predictor. The test coefficient increases from $R^2_{test} = 0.653$ for the lower-accuracy configuration to $R^2_{test} = 0.756$ for the intermediate network and finally to $R^2_{test} = 0.919$ for the optimized model. The total increase is $0.919 - 0.653 = 0.266$, while the training–test gap decreases from 0.037 to 0.013. Thus, the gain in accuracy is accompanied by more consistent performance across the temporal subsets. The sensitivity analysis indicates that both insufficient temporal context and unsuitable optimization settings limit the ability of the network to associate wake structures with C_L. Short sequences do not adequately describe recent vortex development, whereas excessively large batches slightly weaken predictive performance. With T = 18 and a batch size of 14, the final architecture captures the dominant coupling between the evolving wake and the transverse aerodynamic force. The remaining errors are concentrated mainly in isolated extrema or rapidly varying samples rather than in persistent bias or phase displacement. These results support the use of direct 3D convolution as an end-to-end approach for learning coupled spatial and short-term temporal information from flow-field image sequences and converting that information into an accurate instantaneous prediction of C_L.

The small training–test R^2 difference indicates consistent predictive performance on a later, held-out portion of the same numerical trajectory under fixed operating conditions. It should therefore be interpreted as generalization within the simulated condition, rather than evidence of transfer to other Reynolds numbers, structural parameters, or geometries. Applicability to limited changes that preserve the response regime remains a possibility requiring independent evaluation. Substantial changes in structural dynamics or response regime require new simulations, corresponding training data, and retraining and reassessment of the model. Transfer to experimental wake images also remains unverified, since measurement noise and differences in image acquisition and flow representation may affect prediction accuracy.

4. Conclusion

This study developed an end-to-end 3D CNN for predicting the instantaneous lift coefficient directly from temporally ordered wake-field images in turbulent VIV of a two-degree-of-freedom square cylinder. The training dataset was obtained from a validated two-way coupled URANS–FSI simulation. Validation reproduced the characteristic lock-in peak at $U_r = 12$, with a maximum cross-flow amplitude of $Y / D = 0.524$, compared with the reference value of $Y / D = 0.579$, corresponding to a deviation of approximately 9.4%. Following removal of the initial transient, 1017 synchronized flow-field and C_L samples were collected over a physical duration of 101.7 s.

The sensitivity analysis showed that temporal context is essential for reliable force prediction. Increasing the input sequence from T = 2 to T = 18 improved R^2 from approximately −0.40 to 0.65 and reduced MAE and MSE from about 0.80 and 0.98 to 0.40 and 0.24, respectively. With T fixed at 18, increasing the batch size to 14 further improved the performance to approximately MAE = 0.19, MSE = 0.06, and R^2 = 0.92. Larger temporal windows or batch sizes produced slight deterioration, confirming the existence of an architecture-specific optimum.

For the final configuration, the model obtained R^2_{train} = 0.932 and R^2_{test} = 0.919. The training–test gap was only $R^2_{train} - R^2_{test}$ = 0.013 indicating consistent performance on the held-out temporal segment. Relative to the lower-accuracy configuration, the test coefficient increased from R^2_{test} = 0.653 to R^2_{test} = 0.919, corresponding to an absolute improvement of 0.266. The optimized network reproduced the principal amplitudes, extrema, zero crossings, and phase of the lift history, with the remaining errors concentrated mainly near isolated high-gradient samples. The results confirm that joint three-dimensional convolution across time and space can capture the short-horizon wake dynamics governing C_L without LSTM units or separate temporal feature processing.

مکانیک مواد پیشرفته و هوشمند
شاپا: ۴۲۲۰-۲۷۸۳
مجله علمی مکانیک مواد پیشرفته و هوشمند

یادگیری فضازمانی دنباله با استفاده از کانولوشن سه‌بعدی برای برآورد بار آیرودینامیکی لحظه‌ای استوانه مربعی با تکیه‌گاه الاستیک

احمدرضا ربیعی[الف]، امیرحسین ربیعی[ب*]، مصطفی اسماعیلی[ج]

[الف] کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه اراک، اراک، ایران، ahmad.reza.rabiee65243@gmail.com
[ب] دانشیار، گروه طراحی کاربردی، دانشکده مهندسی مکانیک، دانشگاه صنعتی اراک، اراک، ایران، rabiee@arakut.ac.ir
[ج] دانشیار، گروه مهندسی مکانیک، دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران، m.esmaeili@khu.ac.ir

واژگان کلیدی

ارتعاش ناشی از جریان، شبکه عصبی کانولوشنی سه‌بعدی، پیش‌بینی ضریب برآ، اندرکنش سیال–سازه، استوانه مربعی.

تاریخ دریافت: ۱۴۰۵/۰۶/۱۸
تاریخ بازنگری: ۱۴۰۵/۰۷/۱۵
تاریخ پذیرش: ۱۴۰۵/۰۷/۱۹

چکیده

در این پژوهش، یک شبکه عصبی کانولوشنی سه‌بعدی فشرده برای پیش‌بینی مستقیم ضریب برآی لحظه‌ای در ارتعاشات ناشی از جریان آشفته یک استوانه مربعی با تکیه‌گاه الاستیک و دو درجه آزادی ارائه شده است. داده‌های سازگار با فیزیک از یک مدل اعتبارسنجی‌شده اندرکنش سیال–سازه دوطرفه، مبتنی بر معادلات ناپایای میانگین‌گیری‌شده ناویر–استوکس و مدل آشفتگی k-ω SST، به دست آمد. پس از حذف پاسخ گذرا، ۱۰۱۷ تصویر هم‌زمان از میدان دنباله و ضریب برآ طی ۱۰۱/۷ ثانیه گردآوری شد. تصاویر متوالی به‌صورت یک حجم فضازمانی واحد به شبکه داده شدند تا ریخت گردابه‌ها، تحول لایه‌های برشی و پویایی کوتاه‌مدت دنباله بدون استفاده از واحدهای بازگشتی یاد گرفته شود. تقسیم داده‌ها به‌ترتیب زمانی انجام گرفت. تحلیل حساسیت، طول توالی ۱۸ و اندازه دسته ۱۴ را به‌عنوان تنظیم بهینه مشخص کرد. مدل نهایی به R^2_{train} = ۰/۹۳۲ و R^2_{test} = ۰/۹۱۹، با MAE حدود ۰/۱۹ و MSE حدود ۰/۰۶، دست یافت و دامنه، نقاط اوج و فرود، عبور از صفر و فاز سیگنال آزمون را با دقت مناسب بازسازی کرد.

۱- مقدمه

ارتعاش ناشی از جریان زمانی رخ می‌دهد که بارگذاری ناپایای سیال با جرم، سختی و میرایی سازه انعطاف‌پذیر برهم‌کنش کند. یکی از مهم‌ترین صورت‌های آن، ارتعاش ناشی از گردابه است که در محدوده همگامی رهایش گردابه و حرکت سازه می‌تواند به نوسان‌های بزرگ و بارگذاری چرخه‌ای پایدار منجر شود [۱–۴]. در اجسام با تکیه‌گاه الاستیک، عدم تقارن دنباله نیروی برآی نوسانی ایجاد می‌کند؛ حرکت حاصل، نقاط جدایش و شکل‌گیری گردابه‌ها را تغییر می‌دهد و دنباله اصلاح‌شده نیز بار آیرودینامیکی بعدی را تعیین می‌کند. ازاین‌رو، برآورد لحظه‌ای ضریب برآ برای ارزیابی خستگی، کنترل ارتعاش و طراحی سامانه‌های برداشت انرژی اهمیت مستقیم دارد [۵–۱۲].

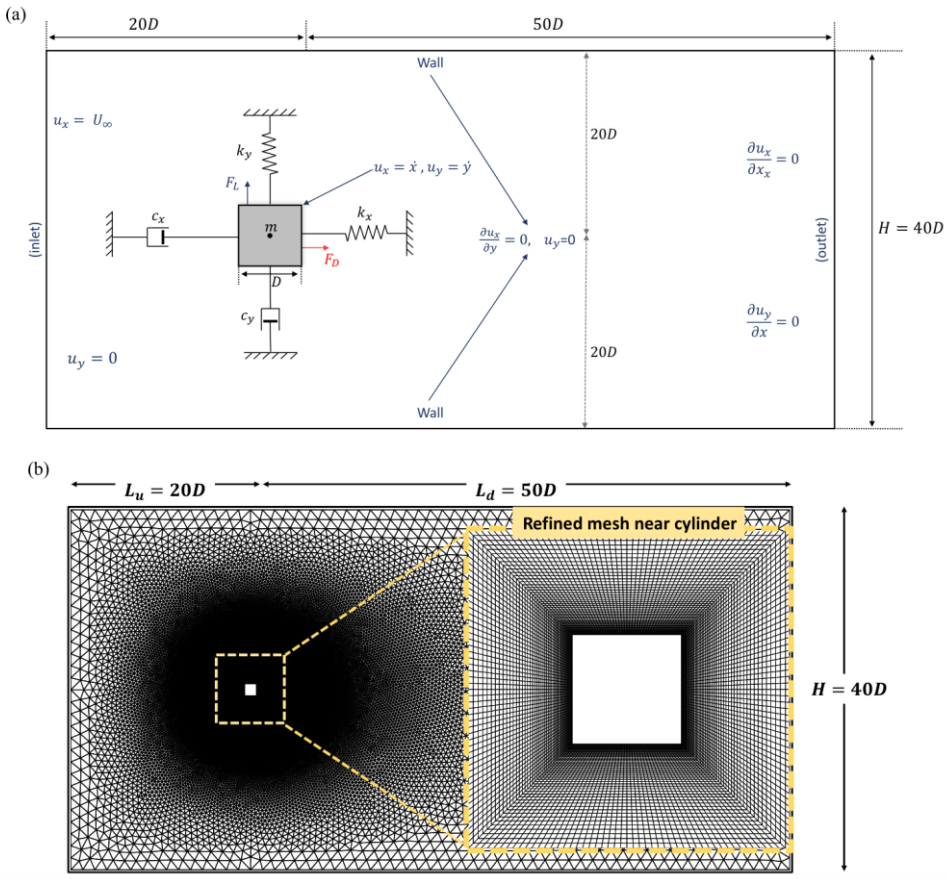
حل‌های دقیق دینامیک سیالات محاسباتی و اندرکنش سیال–سازه این چرخه بازخورد را بازنمایی می‌کنند، اما تکرار آن‌ها در بهینه‌سازی، تحلیل عدم‌قطعیت و تخمین برخط پرهزینه است. شبکه‌های کانولوشنی برای داده‌های تصویری جریان مناسب‌اند، زیرا گرادیان‌های موضعی، لایه‌های برشی، هسته گردابه و الگوهای چندمقیاسی را استخراج می‌کنند [۱۳–۲۰]. نوآوری این مطالعه، نگاشت مستقیم توالی کامل تصاویر دنباله به ضریب برآ با کانولوشن سه‌بعدی است؛ در این روش، اطلاعات مکانی و زمانی

* نویسنده مسئول
آدرس پست الکترونیک: rabiee@arakut.ac.ir

در یک عملیات واحد آموخته می‌شوند و نیازی به شبکه حافظه بلند–کوتاه‌مدت یا مرحله جداگانه کاهش مرتبه نیست [۲۱–۳۰].

۲- چارچوب فیزیکی و تولید مجموعه‌داده

مسئله فیزیکی شامل جریان آشفته، تراکم‌ناپذیر و دوبعدی پیرامون استوانه مربعی صلبی است که در راستاهای هم‌جهت و عمود بر جریان دو درجه آزادی دارد. جریان با معادلات ناپایای میانگین‌گیری‌شده ناویر–استوکس حل و تنش‌های رینولدز با مدل k-ω SST بسته شدند [۳۱]. دامنه محاسباتی از مرکز استوانه تا مرز ورودی ۲۰D، تا خروجی ۵۰D و در راستای عمودی در مجموع ۴۰D امتداد دارد. حرکت سازه با دستگاه جرم–فنر–میرای خطی در دو راستا محاسبه شد. در هر گام زمانی، حل‌گر سیال نیروهای برآ و پسا را به حل‌گر سازه منتقل می‌کرد؛ جابه‌جایی و سرعت جدید موقعیت استوانه را به‌روزرسانی می‌کرد و شبکه محاسباتی پیش از حل گام بعد تغییرشکل می‌یافت.



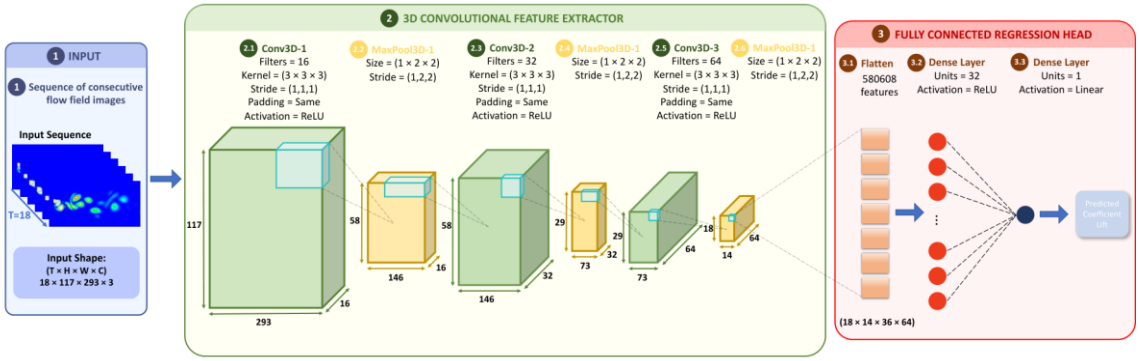
شکل ۱ دامنه محاسباتی، شبکه و شرایط مرزی مسئله اندرکنش سیال–سازه.

پیش از تولید داده، چارچوب عددی با نتایج مرجع یک استوانه دارای دو درجه آزادی، نسبت جرم ۲/۷۳۲۵ و میرایی سازه‌ای صفر اعتبارسنجی شد. پاسخ عرضی در سرعت‌های کاهش‌یافته ۴ تا ۱۴، قله مشخص ناحیه قفل‌شدگی را در سرعت کاهش‌یافته ۱۲ بازتولید کرد. بیشینه دامنه بی‌بعد محاسبه‌شده Y/D = ۰/۵۲۴ بود، درحالی‌که مقدار مرجع Y/D = ۰/۵۷۹ گزارش شده بود؛ بنابراین اختلاف قله حدود ۹/۴ درصد به دست آمد. پس از اعتبارسنجی، شبیه‌سازی اصلی در رژیم ارتعاش ناشی از گردابه اجرا شد. بازه ابتدایی t < ۱۰ s به دلیل رشد ناکامل پاسخ سازه حذف شد و ۱۰۱۷ تصویر میدان جریان، همراه با ضریب برآی هم‌زمان، در فاصله t = ۱۰–۱۱۱/۷ s نمونه‌برداری شد. تصاویر هم‌تراز، برش، تغییراندازه و نرمال شدند و نویز پرفرکانس سیگنال ضریب برآ با فیلتر پایین‌گذر، بدون ایجاد اعوجاج فازی، کاهش یافت. تقسیم آموزش، اعتبارسنجی و آزمون نیز بدون برهم‌زدن ترتیب زمانی انجام شد تا علیت حفظ و از نشت اطلاعات جلوگیری شود.

۳- معماری شبکه و روش آموزش

ورودی شبکه یک تانسور چهاربعدی با ابعاد زمان، ارتفاع، عرض و کانال رنگ است. هر هسته سه‌بعدی هم‌زمان در امتداد زمان و دو راستای مکانی حرکت می‌کند؛ بنابراین شکل گردابه، انحنای لایه برشی و جهت تحول کوتاه‌مدت آن در یک نمایش مشترک رمزگذاری می‌شوند. معماری از سه بلوک پیاپی کانولوشن سه‌بعدی و بیشینه‌گیری تشکیل شده است. تعداد فیلترها به‌ترتیب ۱۶، ۳۲ و ۶۴ و اندازه همه هسته‌ها ۳×۳×۳ است. پنجره بیشینه‌گیری ۱×۲×۲ فقط ابعاد مکانی را کاهش می‌دهد و طول زمانی T=۱۸ ثابت باقی می‌ماند. پس از سه لایه کانولوشن، میدان پذیرش زمانی هر خروجی هفت قاب متوالی را می‌پوشاند. حجم ویژگی نهایی با ابعاد ۱۸×۱۴×۳۶×۶۴، شامل ۵۸۰۶۰۸ مؤلفه، تخت و به یک لایه چگال ۳۲ واحدی با فعال‌ساز ReLU متصل می‌شود؛ یک نورون خطی نیز ضریب برآی گام بعد را تولید می‌کند.

ورودی شبکه شامل توالی تصاویر کانتور رنگی میدان گردابه‌مندی است. هر تصویر با سه کانال قرمز، سبز و آبی (RGB) نمایش داده می‌شود که مقادیر آن‌ها از نگاشت رنگی همان میدان گردابه‌مندی حاصل شده‌اند و معرف میدان‌های فیزیکی مستقل نیستند.

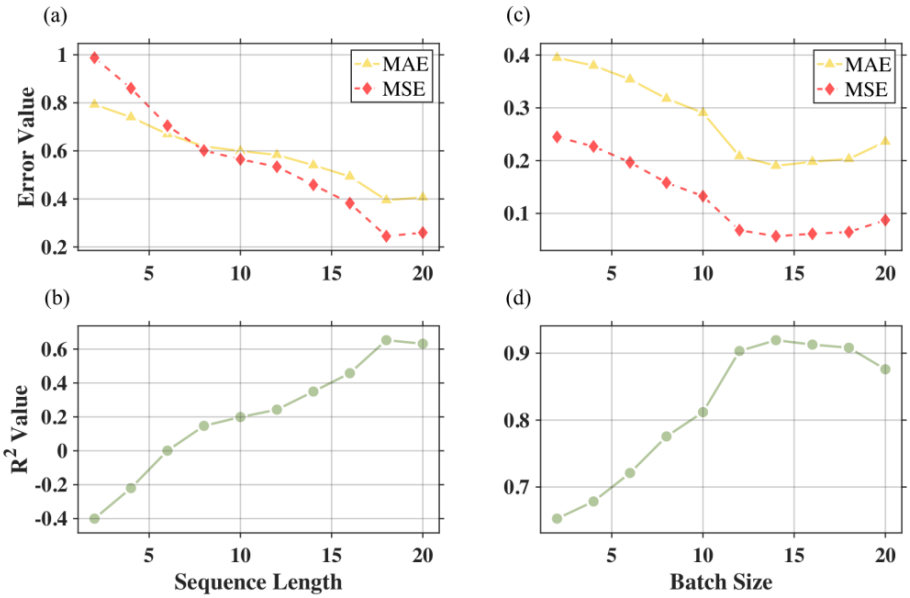


شکل ۲. معماری شبکه عصبی کانولوشنی سه‌بعدی پیشنهادی.

آموزش با کمینه‌سازی میانگین مربعات خطا، بهینه‌ساز Adam و نرخ یادگیری ۰/۰۰۱ انجام شد. حداکثر ۱۰۰ دوره آموزش در نظر گرفته شد و توقف زودهنگام بر پایه خطای اعتبارسنجی از بیش‌برازش جلوگیری کرد. دو ابرپارامتر به‌صورت ترتیبی بررسی شدند: ابتدا طول توالی از ۲ تا ۲۰ قاب با اندازه دسته ثابت ۲ تغییر کرد؛ سپس با تثبیت T=۱۸، اندازه دسته از ۲ تا ۲۰ تغییر یافت. این ترتیب ضروری بود، زیرا طول توالی ابعاد ورودی و اندازه بردار ویژگی را تغییر می‌دهد، درحالی‌که اندازه دسته فقط پویایی بهینه‌سازی معماری ثابت را کنترل می‌کند. عملکرد با MAE، MSE و ضریب تعیین R^2 سنجیده شد.

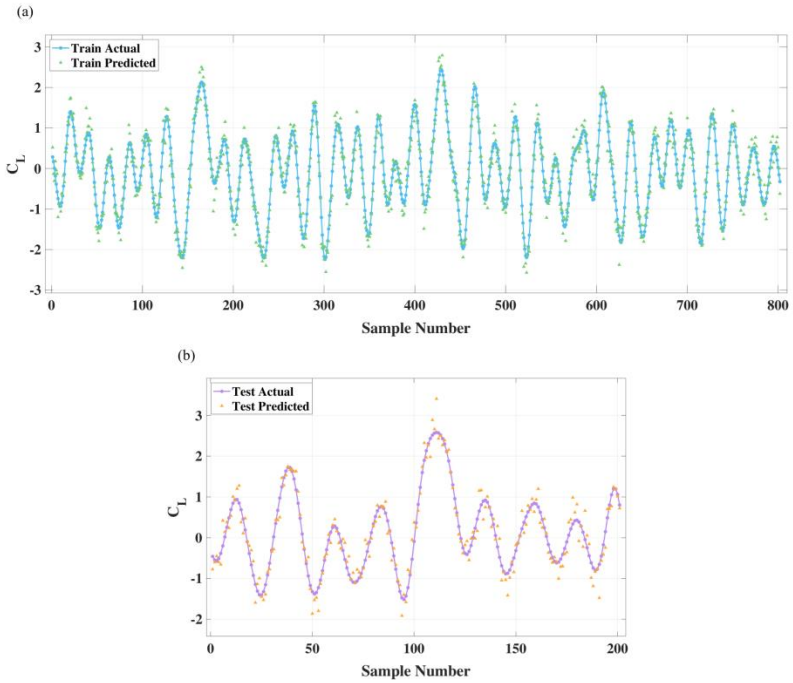
۴- نتایج و بحث

تحلیل حساسیت نشان داد که اطلاعات زمانی ناکافی مانع تشخیص رابطه میان تکامل دنباله و نیروی لحظه‌ای می‌شود. در T = ۲، مقادیر تقریبی MSE = ۰/۹۸، MAE = ۰/۸۰ و R^2 = −۰/۴۰ بود؛ بنابراین دو قاب برای توصیف تحول گردابه کافی نبود. با افزایش طول توالی، هر دو خطا کاهش و ضریب تعیین افزایش یافت. در T = ۱۸، مقادیر MAE و MSE به حدود ۰/۴۰ و ۰/۲۴ رسیدند و R^2 حدود ۰/۶۵ شد. افزایش طول توالی به ۲۰ قاب افت اندکی ایجاد کرد. در مرحله دوم، افزایش اندازه دسته از ۲ به ۱۴، MAE را از حدود ۰/۴۰ به ۰/۱۹ و MSE را از حدود ۰/۲۴ به ۰/۰۶ کاهش داد و R^2 را به حدود ۰/۹۲ رساند؛ دسته‌های بزرگ‌تر دقت را اندکی کاهش دادند. ازاین‌رو، T = ۱۸ و اندازه دسته ۱۴ به‌عنوان تنظیم نهایی انتخاب شدند.



شکل ۳ تحلیل حساسیت طول توالی و اندازه دسته در مدل سه‌بعدی.

برای نشان‌دادن مسیر بهبود، سه سطح عملکرد مقایسه شد. در پیکربندی کم‌دقت، R^2_{train}=۰/۶۹۰ و R^2_{test}=۰/۶۵۳ بود و فاصله تعمیم ۰/۰۳۷ نشان می‌داد که ضعف اصلی از بازنمایی ناکامل نگاشت غیرخطی، نه صرفاً بیش‌برازش، ناشی می‌شود. پیکربندی میانی این مقادیر را به ۰/۷۸۹ و ۰/۷۵۶ رساند و فاصله تعمیم را به ۰/۰۳۳ کاهش داد. در مدل بهینه، R^2_{train}=۰/۶۹۰ و R^2_{test}=۰/۹۱۹ و فاصله تعمیم فقط ۰/۰۱۳ بود. ابر نقاط رگرسیون در کل دامنه ضریب برآ پیرامون خط یک‌به‌یک متمرکز شد و روندهای آموزش و آزمون شیب‌های نزدیک به هم داشتند. بهبود مطلق ضریب تعیین آزمون نسبت به پیکربندی کم‌دقت ۰/۲۶۶ بود.



شکل ۴ مقایسه تاریخچه زمانی ضریب برآی عددی و پیش‌بینی‌شده برای مدل بهینه.

مقایسه تاریخچه‌های زمانی نتیجه آماری را تأیید کرد. مدل نهایی تناوب علامت، تغییر دامنه، بیشینه‌ها، کمینه‌ها و بیشتر

نقاط عبور از صفر را در داده‌های آموزش و آزمون بازتولید کرد. قله مثبت بزرگ در میانه بازه آزمون، افت منفی پیش از آن و نوسان‌های کم‌دامنه بعدی با هم‌ترازی زمانی مناسب ثبت شدند و با افزایش شماره نمونه، رانش فاز پیشرونده مشاهده نشد. خطاهای باقی‌مانده عمدتاً به چند نقطه با گرادیان زیاد یا قله‌های تیز محدود بودند و سوگیری پایدار، خطای شیب نظام‌مند یا عدم تقارن آشکار میان مقادیر مثبت و منفی دیده نشد. این یافته نشان می‌دهد کانولوشن مشترک در زمان و فضا می‌تواند بدون واحد بازگشتی، ویژگی‌های کوتاه‌مدت دنباله لازم برای برآورد دقیق نیروی برآ را استخراج کند.

۵- نتیجه‌گیری

این مطالعه یک نگاشت مستقیم تصویر به نیرو برای ارتعاش آشفته استوانه مربعی ارائه کرد. مجموعه‌داده از حل معتبر دوطرفه سیال–سازه به دست آمد و تقسیم زمانی داده‌ها، ارزیابی واقع‌بینانه روی بازه دیده‌نشده را ممکن ساخت. نتیجه اصلی آن است که ترکیب ۱۸ قاب متوالی با اندازه دسته ۱۴، توازن مناسبی میان پوشش تحول گردابه و پایداری بهینه‌سازی ایجاد می‌کند. دقت R^2_{test}=۰/۹۱۹ و بازسازی مناسب دامنه و فاز نشان می‌دهد شبکه سه‌بعدی یکپارچه می‌تواند جانشینی مؤثر برای تخمین سریع بار لحظه‌ای باشد. بااین‌حال، خطاهای موضعی نزدیک قله‌های تند و وابستگی مدل به دامنه شرایط شبیه‌سازی‌شده، حدود کاربرد آن را مشخص می‌کنند ارزیابی شبکه به بخش زمانی دیده‌نشده از همان شبیه‌سازی محدود است. تعمیم به شرایط فیزیکی یا داده‌های تجربی دیگر بررسی نشده و تغییر قابل‌توجه دینامیک یا رژیم پاسخ به تولید داده‌های جدید، بازآموزی و ارزیابی مجدد نیاز دارد.

Authorship Contribution Statement

Ahmadreza Rabiee

Biography: Researcher in computer engineering with interests in deep learning, image-based regression, and computational data analysis.

Contribution Statement: Methodology, Software, Data curation, Formal analysis, Visualization, Investigation, Writing – original draft.

Dr. Amir Hossein Rabiee

Biography: Associate Professor of Mechanical Engineering at Arak University of Technology. His research focuses on flow-induced vibration, fluid–structure interaction, flow control, energy harvesting, and data-driven modeling.

Contribution Statement: Conceptualization, Methodology, Validation, Formal analysis, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Dr. Mostafa Esmaeili

Biography: Associate Professor of Mechanical Engineering at Kharazmi University. His research interests include computational fluid dynamics, flow control, energy harvesting, turbulent flow, heat transfer, and renewable energy.

Contribution Statement: Conceptualization, Methodology, Validation, Resources, Supervision, Writing – review & editing.

۶- مراجع

[1] Williamson CH, Govardhan R. Vortex-induced vibrations. Annu Rev Fluid Mech. 2004;36(1):413–55.

[2] Sarpkaya T. A critical review of the intrinsic nature of vortex-induced vibrations. J Fluids Struct. 2004;19(4):389–447.

[3] Jauvtis N, Williamson C. The effect of two degrees of freedom on vortex-induced vibration at low mass and damping. J Fluid Mech. 2004;509:23–62.

[4] Williamson CH, Govardhan R. A brief review of recent results in vortex-induced vibrations. J Wind Eng Ind Aerodyn. 2008;96(6–7):713–35.

[5] Ma X, Zhou S. A review of flow-induced vibration energy harvesters. Energy Convers Manag. 2022;254:115223.

[6] Rabiee AH, Naderi A, Esmaeili M. Enhancing energy harvesting capability through analysis of flow-induced vibrations in double sprung circular cylinders exposed to turbulent flow: influence of incidence angle. Ocean Eng. 2024;309:118598.

[7] Esmaeili M, Fakhri Vayqan H, Rabiee AH. The role of flow angle in mixed convection and vortex-induced vibration of a thermally controlled elastic cylinder. Int J Numer Methods Heat Fluid Flow. 2025;35(2):799–824.

[8] Esmaeili M, Fakhri Vayqan H, Rabiee AH. Impact of temperature-induced buoyancy on the 2DOF-VIV of a heated/cooled cylinder. Arab J Sci Eng. 2025;50(4):2807–22.

[9] Sane AG, Esmaeili M, Rabiee AH. Dual benefits of superhydrophobic surfaces: suppression of flow-induced vibrations and enhancement of heat transfer in cylindrical structures. Ocean Eng. 2026;344:123643.

[10] Rabiee AH, Najafpour A, Rajabi M, Esmaeili M. Wind tunnel experiments on flow-induced vibration and energy harvesting: vibrating and stationary cylinder interactions. Results Eng. 2025:108792.

[11] Najafpour A, Rajabi M, Esmaeili M, Rabiee AH. Augmented wake-induced wind energy harvesting in tandem rectangular plate-cylinder arrangements. Energy. 2025:138769.

[12] Rabiee AH, Esmaeili M, Rajabi M. Assessing voltage and power prediction in vibrating cylinders using machine learning algorithms: insights from wind tunnel experiments. Eng Appl Artif Intell. 2026;163:112848.

[13] Brunton SL, Noack BR, Koumoutsakos P. Machine learning for fluid mechanics. Annu Rev Fluid Mech. 2020;52:477–508. doi:10.1146/annurev-fluid-010719-060214.

[14] Vinuesa R, Brunton SL. Enhancing computational fluid dynamics with machine learning. Nat Comput Sci. 2022;2(6):358–66.

[15] Kochkov D, Smith JA, Alieva A, Wang Q, Brenner MP, Hoyer S. Machine learning–accelerated computational fluid dynamics. Proc Natl Acad Sci U S A. 2021;118(21):e2101784118.

[16] Bhatnagar S, Afshar Y, Pan S, Duraisamy K, Kaushik S. Prediction of aerodynamic flow fields using convolutional neural networks. Comput Mech. 2019;64(2):525–45.

[17] Thuerey N, Weißenow K, Prantl L, Hu X. Deep learning methods for Reynolds-averaged Navier–Stokes simulations of airfoil flows. AIAA J. 2020;58(1):25–36.

[18] Fukami K, Fukagata K, Taira K. Super-resolution reconstruction of turbulent flows with machine learning. J Fluid Mech. 2019;870:106–20.

[19] Murata T, Fukami K, Fukagata K. Nonlinear mode decomposition with convolutional neural networks for fluid dynamics. J Fluid Mech. 2020;882:A13.

[20] Lee S, You D. Data-driven prediction of unsteady flow over a circular cylinder using deep learning. J Fluid Mech. 2019;879:217–54.

[21] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019;378:686–707.

[22] Cheng C, Meng H, Li YZ, Zhang GT. Deep learning based on PINN for solving 2 DOF vortex induced vibration of cylinder. Ocean Eng. 2021;240:109932.

[23] Bai XD, Zhang W. Machine learning for vortex induced vibration in turbulent flow. Comput Fluids. 2022;235:105266.

[24] Nazvanova A, Ong MC, Yin G. A data-driven reduced-order model based on long short-term memory neural network for vortex-induced vibrations of a circular cylinder. Phys Fluids. 2023;35(6).

[25] Bublík O, Heidler V, Pecka A, Vimmr J. Neural-network-based fluid–structure interaction applied to vortex-induced vibration. J Comput Appl Math. 2023;428:115170.

[26] Rabiee AH, Esmaeili M, Shafiei P. LSTM-based prediction of lift and drag coefficients in vortex-induced vibration of circular cylinder in turbulent flow. Neural Comput Appl. 2026;38(10):375.

[27] Lopez-Martin M, Le Clainche S, Carro B. Model-free short-term fluid dynamics estimator with a deep 3D-convolutional neural network. Expert Syst Appl. 2021;177:114924.

[28] Guastoni L, et al. Convolutional-network models to predict wall-bounded turbulence from wall quantities. J Fluid Mech. 2021;928:A27.

[29] Geneva N, Zabaras N. Modeling the dynamics of PDE systems with physics-constrained deep auto-regressive networks. J Comput Phys. 2020;403:109056.

[30] Cheng M, Fang F, Pain CC, Navon I. Data-driven modelling of nonlinear spatio-temporal fluid flows using a deep convolutional generative adversarial network. Comput Methods Appl Mech Eng. 2020;365:113000.

[31] Menter FR. Two-equation eddy-viscosity turbulence models for engineering applications. AIAA J. 1994;32(8):1598–605.

[32] Zhu Y, Chen F, Duan Y, Li Q, Li Y. Effect of side ratio on the two-degrees-of-freedom vortex-induced vibrations of the rectangular cross section model. Phys Fluids. 2024;36:105105.

[33] Manzoor S, Khawar J, Sheikh NA. Vortex-induced vibrations of a square cylinder with damped free-end conditions. 2013.