



## OPEN ACCESS

## EDITED BY

Xingwen Zheng,  
The University of Tokyo, Japan

## REVIEWED BY

Amar M. Kamat,  
University of Groningen, Netherlands  
Zhiqiang Ma,  
City University of Hong Kong, Hong  
Kong SAR, China

## \*CORRESPONDENCE

Qian Xue,  
✉ qian.Xue@rit.edu  
Xudong Zheng,  
✉ xxzeme@rit.edu

## †PRESENT ADDRESS

Geng Liu,  
Department of Physics and Engineering,  
University of Scranton, Scranton, PA,  
United States

†These authors have contributed equally  
to this work

RECEIVED 30 May 2023

ACCEPTED 17 July 2023

PUBLISHED 03 August 2023

## CITATION

Bodaghi D, Wang Y, Liu G, Liu D, Xue Q  
and Zheng X (2023), Deciphering the  
connection between upstream  
obstacles, wake structures, and root  
signals in seal whisker array sensing  
using interpretable neural networks.  
*Front. Robot. AI* 10:1231715.  
doi: 10.3389/frobt.2023.1231715

## COPYRIGHT

© 2023 Bodaghi, Wang, Liu, Liu, Xue and  
Zheng. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original author(s)  
and the copyright owner(s) are credited  
and that the original publication in this  
journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Deciphering the connection between upstream obstacles, wake structures, and root signals in seal whisker array sensing using interpretable neural networks

Dariush Bodaghi<sup>1†</sup>, Yuxing Wang<sup>2†</sup>, Geng Liu<sup>3†</sup>, Dongfang Liu<sup>2</sup>,  
Qian Xue<sup>1,4\*</sup> and Xudong Zheng<sup>1,4\*</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Maine, Orono, ME, United States, <sup>2</sup>Department of Computer Engineering, Rochester Institute of Technology, Rochester, NY, United States, <sup>3</sup>Department of Engineering, King's College, Wilkes-Barre, PA, United States, <sup>4</sup>Department of Mechanical Engineering, Rochester Institute of Technology, Rochester, NY, United States

This study presents a novel method that combines a computational fluid-structure interaction model with an interpretable deep-learning model to explore the fundamental mechanisms of seal whisker sensing. By establishing connections between crucial signal patterns, flow characteristics, and attributes of upstream obstacles, the method has the potential to enhance our understanding of the intricate sensing mechanisms. The effectiveness of the method is demonstrated through its accurate prediction of the location and orientation of a circular plate placed in front of seal whisker arrays. The model also generates temporal and spatial importance values of the signals, enabling the identification of significant temporal-spatial signal patterns crucial for the network's predictions. These signal patterns are further correlated with flow structures, allowing for the identification of important flow features relevant for accurate prediction. The study provides insights into seal whiskers' perception of complex underwater environments, inspiring advancements in underwater sensing technologies.

## KEYWORDS

bioinspired flow sensing, wake identification, seal whisker, interpretable machine learning, fluid-structure interaction

## 1 Introduction

Given the limitations of the current underwater navigation systems, there has been a growing interest in advancing underwater sensing capabilities. The increasing scientific, military, and commercial demands, along with the exploration of the uncharted underwater world, have been the driving force behind these developments. Underwater robots employ various navigation systems, including inertial, geophysical, and acoustic systems, to estimate their position, velocity, and acceleration (Stutters et al., 2008; Chutia et al., 2017). However, these systems rely heavily on seafloor maps to prevent collisions and have limited capabilities in surveilling, navigating, and maneuvering safely in unpredictable

deep waters, while avoiding obstacles and hazardous environments (Kinsey et al., 2006; Leonard and Bahr, 2016).

To address these challenges, different types of sensors are utilized on underwater robots (Chutia et al., 2017; Elshalakani et al., 2020). Camera-based sensors, relying on vision, are commonly used for monitoring the nearby environment and object detection. However, their effectiveness is much lower in dark deep waters, necessitating the use of artificial light sources. This further reduces the sensor's range and reveals its location, posing a significant disadvantage, particularly in military applications (Yang et al., 2006; Jiang et al. 2019; Jiang et al. 2020). In environments with poor water clarity, the images captured by these sensors suffer from low quality and cloudiness due to light dissipation (scattering) over long distances (Kröger, 2008; Lee et al., 2012). To overcome the limitations of vision-based sensors, sonar sensors are commonly employed to estimate object locations by transmitting acoustic waves (Akyildiz et al., 2004). However, similar to vision sensors, sonar sensors have drawbacks such as high-power consumption and the risk of revealing the sensor's location (Akyildiz et al., 2004; Elshalakani et al., 2020). Furthermore, the use of sonar sensors can have detrimental effects on various underwater species by altering their natural habitats (Schrope, 2002).

In recent years, there has been significant research on exploring the sensing abilities of phocid seals to address challenges in underwater sensing. These seals possess remarkable whisker arrays that enable them to detect and distinguish hydrodynamic disturbances caused by potential prey. Studies have shown that seals can detect disturbances with speeds as low as 245  $\mu\text{m/s}$  and frequencies ranging from 10 to 100 Hz (Dehnhardt et al., 1998). Behavior studies have demonstrated the seals' capability to discern the shape and size of upstream objects, track the travel direction of vortex rings, detect miniature submarines, and even follow other seals, by only using their whisker arrays (Dehnhardt et al., 2001; Müller and Kuc, 2007; Wieskotten et al., 2010; 2011). Notably, seals can track objects at considerable distances and detect weak vortices from past movements (Müller and Kuc, 2007). These findings highlight the extraordinary sensitivity, accuracy, and intelligence of seal whisker arrays in detecting hydrodynamic cues. The remarkable sensing abilities of seal whiskers have sparked great interest and inspiration for applications in underwater navigation, object tracking, and object detection. Several biomimetic sensors has already been designed, fabricated and tested in the laboratories (Beam et al., 2013; Kottapalli et al., 2015; Eberhardt et al., 2016; Sayegh et al., 2022; Wang et al., 2022; Zheng et al., 2022; Liu et al., 2023).

Previous studies on the fundamental mechanism of seal whisker sensing have mostly focused on the geometric effects, revealing that the whisker's unique undulated shape can suppress self-induced vibrations by disrupting the Kármán vortex street in the wake (Hanke et al., 2010; Hans et al., 2013; Beem, 2015; Kottapalli et al., 2015; Morrison et al., 2016). Our recent study (Liu et al., 2019b) further elucidated that the suppression of the vibration is primarily due to the 180° phase shift of the axes of the elliptical cross-sections of seal whisker, which generates stable three-dimensional hair-pin vortices in the wake. Additionally, Beem and Triantafyllou, (2015) found that seal whisker geometry is able to generate large-amplitude “slaloming” motions in vortical flows by extracting energy from passing vortices, thereby enhancing wake capturing sensitivity.

While these findings confirm the high sensitivity of seal whisker sensing, the fundamental mechanism of wake identification and tracking remains largely unknown. Seal whiskers are arranged in a stereotyped grid pattern known as the vibrissa array, and the resulting bend moments at the root of the whiskers collectively form sensory inputs. Recent studies have suggested that seals may correlate temporal-spatial patterns of whisker array signals with surrounding flow patterns for intelligent perception. Various flow features, such as highest velocities, velocity gradients, and wake spatial extension, have been proposed as potential cues for object perception (Wieskotten et al., 2011). Artificial whisker array experiments have also indicated that cross-correlation of bending signals can be used to detect vortices (Glick et al., 2021). Our previous computational studies have shown that whisker array signals can reflect the strength, timing, and trajectories of upstream vortex-induced jets (Liu et al., 2021). Machine-learning models have been used to infer the position of upstream objects based on whisker tip displacements (Elshalakani et al., 2020). However, many questions remain unanswered, including the correlation between mechanical whisker signals and external flow disturbances, the specific flow features seals can sense, and how they interpret flow features to perceive the environment. Further research is needed to unravel these mysteries and advance our understanding of seal whisker sensing.

In this study, we aim to develop a novel method that combines a computational fluid-structure interaction (FSI) model with an interpretable deep-learning model to investigate the fundamental mechanisms underlying seal whisker sensing. The novelty of the method lies in its ability to establish connections between crucial signal patterns, flow characteristics, and attributes of upstream obstacles, thereby enhancing our understanding of the intricate sensing mechanisms. The effectiveness of the method is demonstrated through its ability to accurately predict the location and orientation of a circular plate positioned in front of seal whisker arrays. To achieve this, we placed a circular plate within a free stream, located upstream of the seal's head. Two whisker arrays were integrated onto the seal's head, replicating the realistic location, structure, geometry, and orientation observed in reported data. A one-way FSI model was employed to compute the plate-induced wake and subsequent dynamics and signals of the whiskers, represented by the root bending moment. By systematically varying the plate's location and orientation, we generated diverse wake characteristics and whisker signals.

We further developed and trained an interpretable neural network model to learn and accurately predict the location and orientation of the upstream plate based on the whisker signals. Importantly, the model produces temporal and spatial importance values, allowing for the identification of crucial temporal-spatial signal patterns contributing to the network's predictions. We then correlated these patterns with flow structures, enabling the identification of important flow features for prediction. The outcomes of this study provide novel insights into seal whiskers' ability of perceiving and interpreting complex underwater environments, which have the potential to advance the development of underwater sensing technologies.

The rest of the paper is organized as follows: **Section 2** introduces the computational model including the governing equations and numerical method, whisker array model, and

parametric simulations setup. Section 3 introduces the interpretable deep learning network, including the network architecture, input preparation and training configuration. Section 4 presents the results of parametric simulations and network prediction and interpretation as well as the connection between signals and flow features. Section 4 presents the summary of the results and general conclusion.

## 2 Computational model

### 2.1 Governing equations and numerical method

The flow field is modeled using unsteady incompressible Navier-Stokes equations:

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

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{\text{Re}} \frac{\partial^2 u_i}{\partial x_j \partial x_j}$$

where  $u_i$  represent the velocity components in the three directions and  $p$  is pressure.  $\text{Re}$  is the Reynolds number, defined as  $\frac{U_\infty}{\nu}$ , where  $U_\infty$  is the incoming flow speed,  $D$  is the characteristic length, and  $\nu$  is the kinetic viscosity of water at 20°C. The presence of the whiskers in the flow field is ignored in this study.

The current study employs an in-house immersed boundary method based incompressible flow solver, as described by Mittal et al. (2008). The solver discretizes the spatial terms using a cell-centered collocated arrangement of the primitive variables  $u_i$  and  $p$ . To integrate the equations in time, a three sub-steps fractional step method based on Van Kan (1986) is employed. In the first sub-step, an intermediate velocity field,  $u^*$ , is determined by solving a modified momentum equation without the pressure term. The convective terms in this equation are discretized using a second-order Adam-Bashforth scheme, while the diffusion terms are discretized using an implicit Crank-Nicholson scheme to remove the viscous stability constraints. In the second sub-step, the pressure corrections are computed by solving a pressure Poisson equation, and in the final step, the pressure field and velocity field are updated using the obtained pressure corrections. To treat the immersed boundaries, a multidirectional ghost-cell methodology is employed, resulting in both local and global second-order accuracy. The solver used in this study has been successfully applied to various biological applications, including human phonation, insect flight, fish swimming, and seal whiskers inspired flow sensing (Geng et al., 2017; Liu et al., 2018; Liu et al., 2019b; Bodaghi et al., 2021a; Bodaghi et al., 2021b; Liu et al., 2021). Details of the solver can be found in Mittal et al. (2008).

Each whisker is modeled as elastic structure and dynamics of whisker is governed by the Navier equation:

$$\frac{\partial^2 d_i}{\partial t^2} = \frac{1}{\rho_w} \frac{\partial \sigma_{ij}}{\partial x_j} + f_i$$

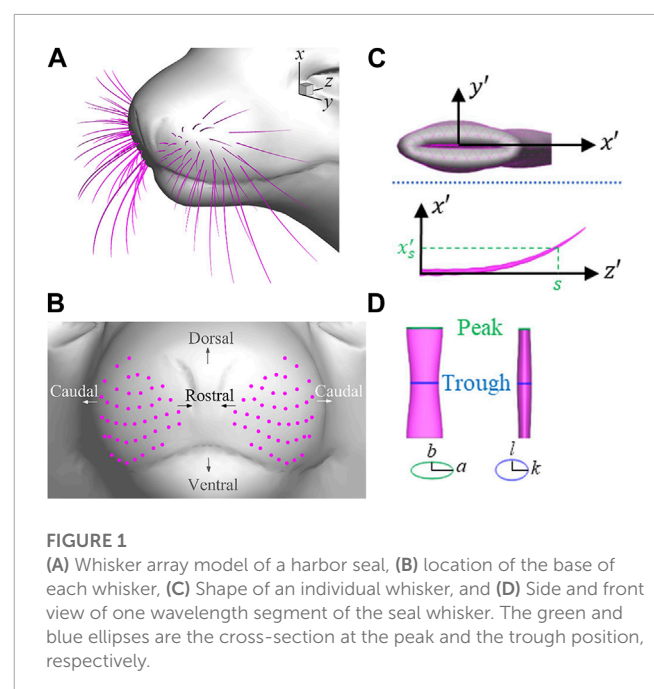
where  $d_i$ ,  $\rho_w$ ,  $\sigma_{ij}$  and  $f_i$  are displacement vector, whisker density, stress tensor and body force vector, respectively. An in-house finite element method code is utilized to solve the equation for

$d_i$  based on the hydrodynamic loads obtained from the flow solver. This solver was fully validated by Liu et al. (2019a) and was successfully employed for different biological simulations, including fish swimming (Liu et al., 2018) and seal whiskers (Liu et al., 2021).

Upon solving the dynamic equation at each time step, the hydrodynamic loadings along each whisker are estimated based on the local flow speeds using a simplified drag-lift model. To achieve this, each individual whisker is discretized into a series of whisker segments, each one spanning one segment. The hydrodynamic force acting on each segment is computed using  $(F_D, F_L) = (C_D, C_L) \cdot 0.5\rho U^2 A_{ref}$ , where  $F_D$  and  $F_L$  are the drag and lateral forces along the primary and minor axis of the cross section, respectively.  $\rho$  is flow density,  $U$  is local flow speed and  $A_{ref}$  is referenced cross section area.  $C_D$  and  $C_L$  are the drag and lateral force coefficients of the whisker segment and depend on local Reynolds number and angle of attack of the segment. We obtained these coefficients by using a tabulated function from our previous numerical parametric study. The resulting hydrodynamic loadings are then applied the centroid points of each segment, together driving the motion of the entire whisker. The details of hydrodynamic load calculation can be found in Liu et al. (2021).

### 2.2 Whisker arrays model

The seal whisker array model, depicted in Figure 1A, is adopted from our previous study (Liu et al., 2021), which closely replicates a real harbor seal's whisker array. The model comprises of two arrays, each containing 44 individual whiskers, placed on the muzzle of a realistic harbor seal head model. The length, curvature, tapering, location, and orientation of each whisker are taken from the reported data of realistic harbor seal whisker arrays (Murphy, 2013; Graff, 2016) and are illustrated in Figures 1B, C. The undulated geometry of each whisker is constructed using a simplified elliptical cylinder



**FIGURE 1** (A) Whisker array model of a harbor seal, (B) location of the base of each whisker, (C) Shape of an individual whisker, and (D) Side and front view of one wavelength segment of the seal whisker. The green and blue ellipses are the cross-section at the peak and the trough position, respectively.

model, superimposed with two sinusoidal undulations with a  $180^\circ$  phase shift of the axes of the elliptical cross-sections (Figure 1D) (Rinehart et al., 2017). The material properties of the whiskers are also determined using experimental data (Hans et al., 2014). For more detailed information, please refer to Liu et al. (2021).

## 2.3 Simulation setup and parametric space

Figure 2 illustrates the simulation setup. A circular plate with a diameter of 40 mm is placed 30 cm in front of the nasal tip of a seal head. This diameter is determined based on previous studies (Wieskotten et al., 2011; Liu et al., 2021) for generating typical-size vortices of prey. The seal head and the plate are immersed into a  $900\text{ mm} \times 1000\text{ mm} \times 950\text{ mm}$  flow domain. A uniform flow speed of 20 cm/s from left to right is applied as far field conditions on left, top, bottom, front and back boundaries of the flow domain and the outflow condition is implemented on the right boundary of flow domain with zero velocity gradient and zero pressure gradient. To reduce the computational time, we increased the kinematic viscosity of fluid to  $4 \times 10^{-5}\text{ m}^2/\text{s}$ , leading to a reduced Reynolds number of  $Re_D = 200$ , defined using the diameter of the circular disk. To ensure grid independence, two different grids were tested, with resolutions of  $128 \times 128 \times 256$  and  $256 \times 256 \times 256$  grid cells in the  $x$ ,  $y$ , and  $z$  directions, respectively. The highest grid density is concentrated near the circular disk, its wake region as well as the whisker arrays, with minimum grid cell sizes of  $3.2\text{ mm} \times 4.4\text{ mm} \times 2.6\text{ mm}$  and  $1.6\text{ mm} \times 2.2\text{ mm} \times 1.3\text{ mm}$  for the coarse and fine grids, respectively. The grid independence study shows that using the coarse grid results in less than 0.5% difference in the drag force experienced by the circular disk and 1% difference in the vortex shedding period compared to the fine grid. Considering the significant reduction of computational cost, the  $128 \times 128 \times 256$  coarse grid is chosen for the study, as depicted in Figure 2B. The flow simulations are executed on the XSEDE Expanse cluster (AMD EPYC 7742 type CPU, clock speed: 2.25 GHz, and flop speed: 4608 GFlop/s)

using 16 processors. The simulations are run for 5,000 time steps to capture three steady vortex shedding cycles, resulting in an average computational cost of 224 CPU hours per case for the flow simulation.

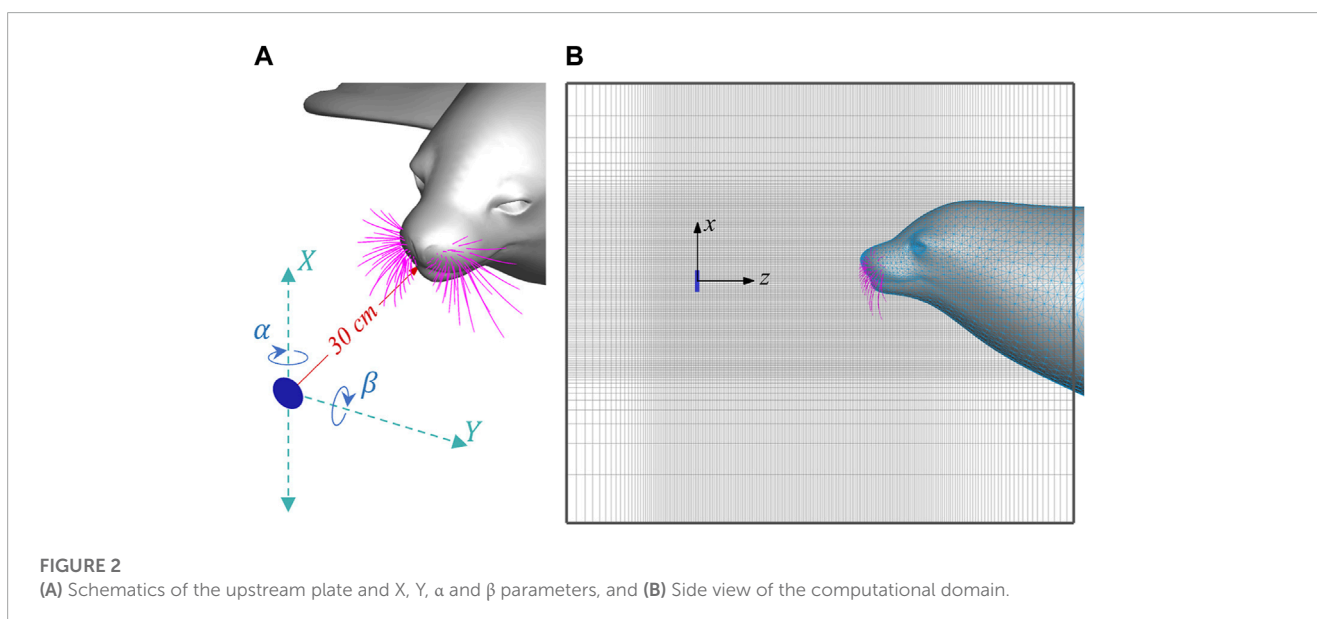
The whiskers are discretized using ten-node tetrahedral cells, resulting in a range of 9,700–34,100 cells depending on the length of the whiskers. The whiskers are fixed at the base, preventing any movement or rotation. The solid simulations are performed on the same XSEDE Expanse cluster, with one processor allocated for each whisker. Thus, for each parametric case that includes 88 whiskers, the average computational cost for the solid simulation is 184 CPU hours.

A group of parametric simulations were conducted by systematically varying the location ( $X$ ,  $Y$ ) and orientation ( $\alpha$ ,  $\beta$ ) of the circular plate, as defined in Figure 2A. The ranges of these parameters were determined to ensure the wakes would intersect at least one of the whiskers. The selected parameters are:  $X$  from  $-120\text{ mm}$  to  $120\text{ mm}$  with an interval of 60 mm,  $Y$  from 0 to 200 mm with an interval of 50 mm,  $\alpha$  from  $-45^\circ$  to  $45^\circ$  with an interval of  $22.5^\circ$ , and  $\beta$  from  $-45^\circ$  to  $45^\circ$  with an interval of  $22.5^\circ$ . This results in a total of 625 simulation cases.

## 3 Interpretable deep learning network

### 3.1 Network architecture

Vision Transformer (ViT) (Vaswani et al., 2017; Dosovitskiy et al., 2020) has demonstrated effective structural interpretation of spatial correlations between pixels in images by dividing raw images into multiple image patches and leveraging comprehensive imaging information. Flexibility of ViT to handle various input sequences allows for the utilization of video information in addition to image information. In this study, we propose a network architecture called Video Transformer (VT) as a straightforward extension of ViT, designed specifically for



video processing. ViT initially splits the video into spatio-temporal tokens and applies a self-attention model using spatial and temporal blocks to extract features based on spatial and temporal changes of moments. These features are then optimized by passing through transformer blocks to determine the target position and orientation. In this section, we first introduce ViT, followed by a discussion on tubelet embedding for video processing (Arnab et al., 2021), and finally, we present the self-attention model for video processing (Arnab et al., 2021).

### 3.1.1 ViT

Initially, ViT processes a 2D image,  $I \in \mathbb{R}^{H \times W \times C}$ , where  $(H, W)$  represents the resolution of the original image, and  $C$  is the number of channels. The standard ViT extracts  $M$  non-overlapping image patches,  $PATCH = \{patch_i\}_{i=1}^M \in \mathbb{R}^{M \times (h \times w \times c)}$ , where  $(h, w)$  represents the resolution of each image patch, and  $c$  is the number of patch channels. The value of  $M$  is determined by  $M = (H/h) \times (W/w)$ . These image patches are then flattened into 1D tokens,  $Q = \{q_i\}_{i=1}^M \in \mathbb{R}^{M \times g}$ , where  $g = (h \times w \times c)$ . These flattened tokens serve as the effective input sequence tokens for ViT. A constant latent vector size  $G$  is defined in ViT across all network layers, projecting the flattened 1D token  $g$  into a  $G$ -dimensional vector using a trainable linear projection. The output of this projection is considered as token embeddings. Similar to BERT's token (Devlin et al., 2018), an optional learnable classification token  $q_{class}$  is prepended to the sequence of embedded patches. Additionally, standard learnable 1D position embeddings  $E'_{pos}$  are added to the token embeddings to preserve positional information, which aids in learning the spatial relationship between whiskers. Finally, the sequence of token embeddings and position embeddings is passed through a transformer encoder (Vaswani et al., 2017), consisting of  $L$  transformer blocks. Each block  $\ell$  comprises

alternating layers of layer normalization (LN), Multi-Headed Self-Attention (MSA), and Multi-Layer Perceptron (MLP), with a residual connection after MSA and MLP (Baevski and Auli, 2018; Wang et al., 2019). Therefore, the transformer encoder can be expressed as:

$$Z_0 = [q_{class}; q_1 E'; q_2 E'; \dots; q_M E'] + E'_{pos},$$

$$E' \in \mathbb{R}^{g \times G}, E'_{pos} \in \mathbb{R}^{(M+1) \times G},$$

$$\dot{Z}_\ell = MSA(LN(Z_\ell)) + Z_\ell, \quad \ell = 0 \dots L-1,$$

$$Z_{\ell+1} = MLP(LN(\dot{Z}_\ell)) + \dot{Z}_\ell, \quad \ell = 0 \dots L-1,$$

where MLP consists of two layers with a GELU non-linearity (Hendrycks and Gimpel, 2016). Finally, an MLP is employed to classify the output of  $L$  transformer blocks, which can be expressed as:

$$target = MLP(LN(Z_L))$$

### 3.1.2 Tubelet embedding

Given the flexibility of ViT to accommodate any sequence of input tokens  $Z_0 \in \mathbb{R}^{M \times G}$ , we can leverage this property to convert image information into video information. For a video  $V \in \mathbb{R}^{T' \times H \times W \times C}$ , where  $T'$  represents the temporal dimension, compared to image information  $I \in \mathbb{R}^{H \times W \times C}$ , the video data introduces an additional temporal dimension. Our goal is to map the video data to a sequence of tokens  $Q^v \in \mathbb{R}^{n_t \times n_h \times n_w \times g}$ . As depicted in "embed to token" in Figure 3, the process of tokenizing video information involves extracting non-overlapping spatio-temporal tubes from the input. These tubes are then projected to a sequence of tokens  $Q^v \in \mathbb{R}^{n_t \times n_h \times n_w \times g}$ , extending the embedding approach of ViT to a 3D setting.

In this context, we define a tubelet with dimensions  $t' \times h \times w$ , where tokens are derived from the temporal dimension  $t'$  and the

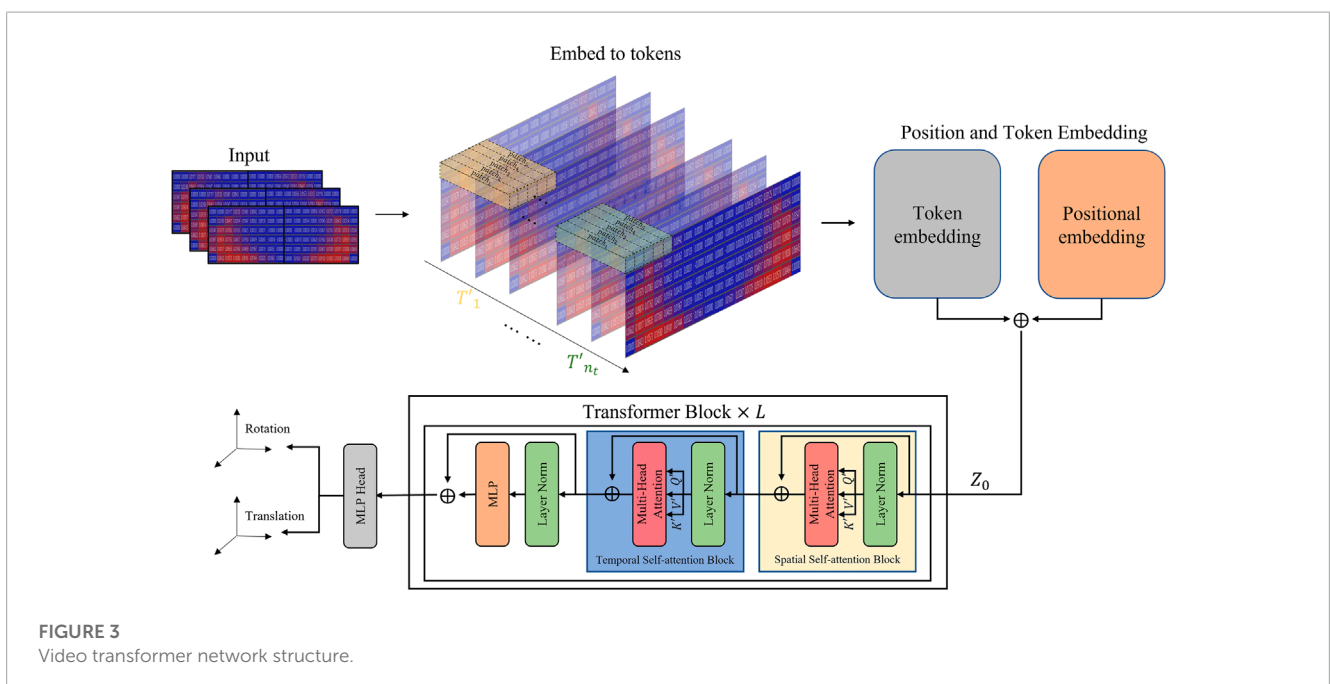


FIGURE 3 Video transformer network structure.

spatial dimensions  $h$  and  $w$ . This allows us to obtain  $n_{t'} = \left\lfloor \frac{T'}{t'} \right\rfloor$ ,  $n_h = \left\lfloor \frac{H}{h} \right\rfloor$ , and  $n_w = \left\lfloor \frac{W}{w} \right\rfloor$  frames. It is worth noting that smaller tubelets in any dimension result in an increase in computational cost during the training process, as more tokens are generated. Through tubelet embedding, the spatio-temporal information from video data can be directly fused into a transformer encoder, eliminating the need for the transformer to fuse temporal information from different frames.

### 3.1.3 Factorized self-attention model

The factorized self-attention model (Ho et al., 2019; Weissenborn et al., 2019; Bertasius et al., 2021) described in the transformer block of Figure 3 processes the spatio-temporal tokens  $Z_0$  obtained through the tubelet embedding method. Within each transformer block, the spatial self-attention block is initially applied to learn the spatial relationship among all tokens with the same temporal index, denoted as  $T_{n_i}^t$ . Subsequently, the temporal self-attention block is employed to learn the temporal relationship among all tokens from the same spatial index, denoted as  $\{patch_i\}_{i=1}^M$ . This enables effective learning of spatiotemporal interactions within each transformer block.

Specifically, the spatial self-attention block calculates the spatial relationship by reshaping the tokens  $Z_\ell \in \mathbb{R}^{n_{t'} \times n_h \times n_w \times g}$  to  $Z_\ell^s \in \mathbb{R}^{n_{t'} \times (n_h \cdot n_w \cdot g)}$ . Then, the output of the spatial self-attention block  $\dot{Z}_\ell^s \in \mathbb{R}^{n_{t'} \times n_h \times n_w \times g}$  is reshaped to  $Z_\ell^t \in \mathbb{R}^{(n_h \cdot n_w) \times (n_{t'} \cdot g)}$  and used as the input for the temporal self-attention block to calculate the temporal relationship. The factorized self-attention model can be summarized as follows:

$$\dot{Z}_\ell^s = MSA\left(LN\left(Z_\ell^s\right)\right) + Z_\ell^s, \quad \ell = 0 \dots 1,$$

$$\dot{Z}_\ell^t = MSA\left(LN\left(Z_\ell^t\right)\right) + Z_\ell^t, \quad \ell = 0 \dots 1,$$

$$Z_{\ell+1} = MLP\left(LN\left(\dot{Z}_\ell^t\right)\right) + \dot{Z}_\ell^t, \quad \ell = 0 \dots 1,$$

Unlike previous approaches, this model does not utilize a classification token since it can handle token reshaping across spatial and temporal dimensions without introducing ambiguities.

## 3.2 Network input preparation

The primary input signal for the network is derived from the time history of the bending moment at the base of each whisker. The total bending moment is decomposed in two components:  $M_{x'}$  and  $M_{y'}$ , where  $x'$  and  $y'$  denotes body-fixed coordinates of each whisker, as depicted in Figure 1C. Based on the direction of  $x'$  and  $y'$ ,  $M_{x'}$  represents the bending moments resulting from the lift force while  $M_{y'}$  represents the bending moments resulting from the drag force. Please note that an additional simulation of the freestream passing the seal head without an upstream plate was conducted and the bending moments obtained from this simulation were then subtracted from the bending moments observed in the other cases where wakes were present. The process is to eliminate the influence of the freestream and solely examine the effect of the wakes.

Tactile vibrissal systems in general possess two distinct types of mechanoreceptors in their whiskers, namely Merkel

cells (SAI mechanoreceptors) and Pacinian corpuscles (PC mechanoreceptors), specialized in detecting static and dynamic changes (Zimmerman et al., 2014). Merkel cells are adept at sensing static stimuli, such as sustained pressure, while Pacinian corpuscles excel at perceiving dynamic changes, including vibrations and rapid movements. To emulate the underlying nerve system, we further decompose the bending moments into two components: time-averaged (DC) components, denoted as  $M_{DC}$ , and oscillatory components (AC), denoted as  $M_{AC}$ . This yields a four-component vector:  $(M_{DC,x'}; M_{AC,x'}; M_{DC,y'}; M_{AC,y'})$ , obtained from each whisker for the network input.

The subsequent step involves a data transformation process aimed at consolidating the inputs from individual whiskers to generate a signal map, as depicted in Figure 4. Initially, the base locations of each whisker were extracted from the model and then topologically mapped onto a  $9 \times 22$  two-dimensional rectangular grid based on their relative positions. Subsequently, each grid point was treated as a pixel and associated with the bending moment vector of the corresponding whisker. Each moment component was considered a distinct channel, with the intensity of each channel determined by the magnitude of its respective moment component. As a result, a “representative sensing image” comprising four channels was created. In cases where a pixel lacked correspondence to any whisker, a NULL (0) value was assigned. Given that bending moments exhibit temporal variations, the resulting data ultimately takes the form of a video file, reflecting the dynamic nature of the sensory input.

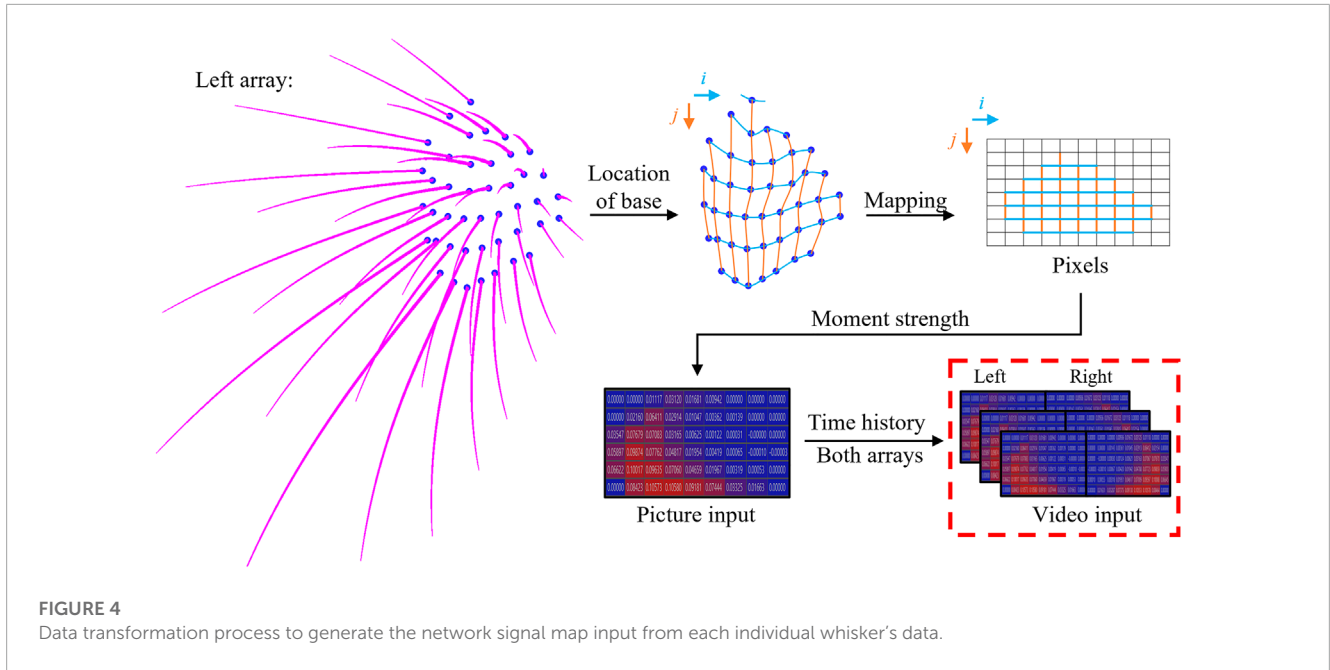
## 3.3 Training configuration

The network training process involves regression training for 1,000 epochs using the Adam optimizer. A batch size of 25 and an initial learning rate of 0.001 are used, along with a multi-step learning rate scheduler. The input samples are divided into train and validation groups using the 5-fold method, with 500 samples for training and 125 samples for validation. Each video sample has a size of  $80 \times 9 \times 22 \times 4$ , representing the temporal dimension (80, corresponding to the last 1,500 flow time steps to only consider the steady vortices), the whisker array dimensions ( $9 \times 22$ ), and the moment channels (4). To capture the relationships between each whisker, the model parameters are initialized as follows:  $t' = 8, h = 1$  and  $w = 1$ , with  $L = 4$  and  $G = 96$ .

The implementation of the method utilizes PyTorch 1.12.1 and Scorch 0.11.0 (Tietz et al., 2017) frameworks. The training process is performed using an NVIDIA GeForce RTX 2080 Ti GPU with 11GB memory.

## 4 Results and discussion

This section begins with a comprehensive examination of the influence of the plate's location and orientation on both the flow field and whisker signals. Subsequently, the outcomes of the interpretable deep learning network are presented, including accuracy measurements and importance value maps of the signals. Lastly, the correlation between the importance value maps, whisker signals, and flow field is analyzed, while also investigating the



**TABLE 1** Definition of the exemplary cases.

Case	$X$	$Y$	$\alpha$	$\beta$
$X_{60}$	60 mm	0	0	0
$Y_{50}$	0	50 mm	0	0
$\alpha_{22.5}$	0	0	22.5°	0
$\beta_{22.5}$	0	0	0	22.5°
$\alpha_{45}\beta_{22.5}$	0	0	45°	22.5°

underlying mechanisms. To facilitate the discussion, five exemplary cases are selected for a detailed investigation, which are summarized in Table 1.

### 4.1 Characteristics of the flow field

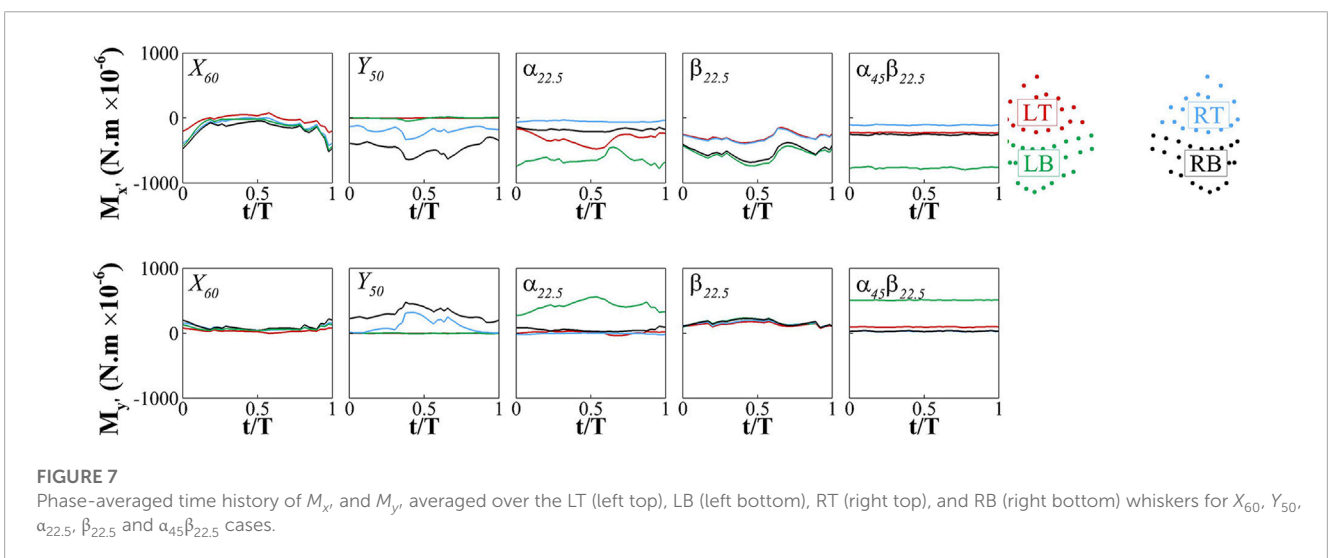
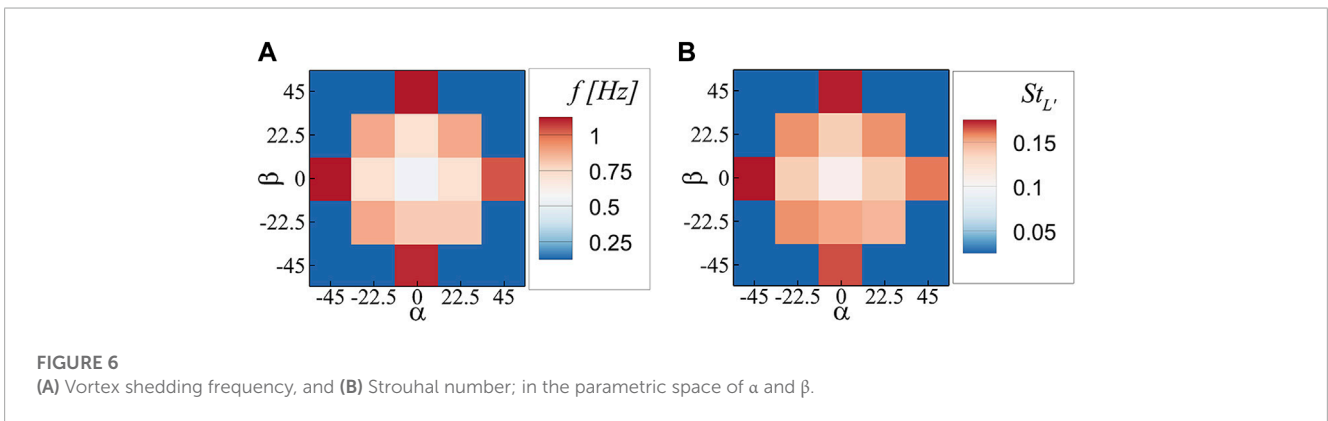
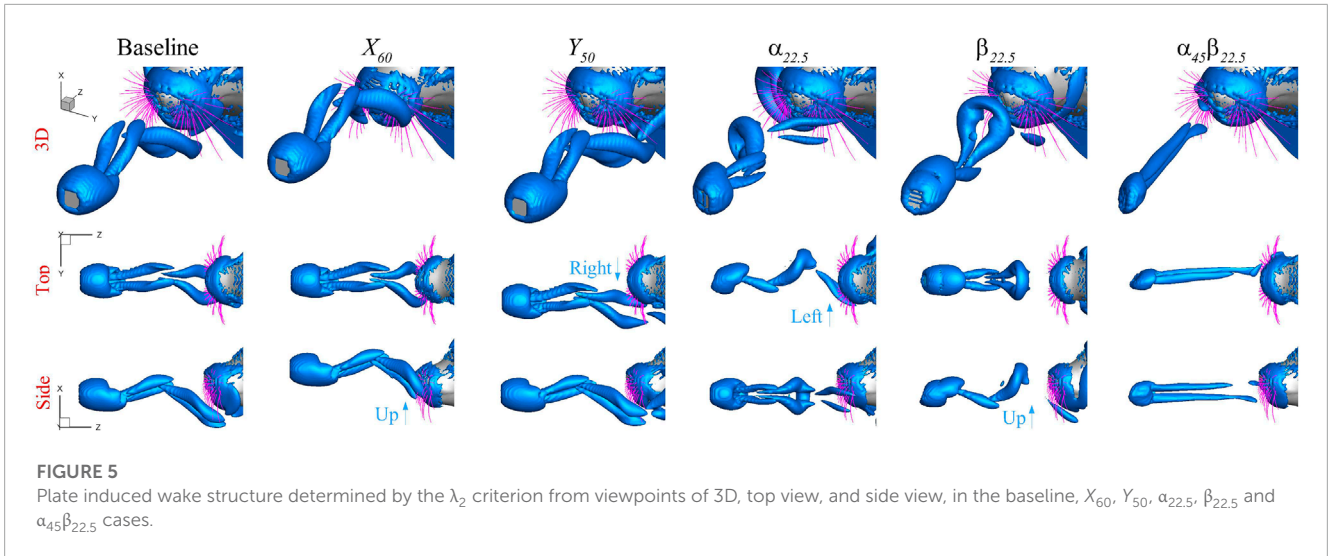
Figure 5 shows the wake structure induced by the plate in the exemplary cases, determined by the  $\lambda_2$  criterion, viewed from different viewpoints: 3D, top view, and side view. To aid the comparison, the baseline case is added in the figure. Except for  $\alpha_{45}\beta_{22.5}$  case, the wake displays two distinct vortex rings that alternate between the upper and lower sides ( $X_{60}$ ,  $Y_{50}$ , and  $\beta_{22.5}$  cases) or the left and right sides ( $\alpha_{22.5}$  case), with a phase difference of 180°. In the baseline case, where the plate is directly in front of the seal's nose, the wake interacts with all whiskers. As the plate moves upward in the  $X_{60}$  case and rightward in the  $Y_{50}$  case, the wake translates accordingly without changing the pattern. In the  $\alpha_{22.5}$  case, where the plate tilts to the left, the wake inclines in the same direction. Additionally, due to plate rotation, the vortex ring forming on the left side is stronger than the one on the right side. In the  $\beta_{22.5}$  case, where the plate tilts upward, the wake also inclines upward accordingly, and the rotation generates a stronger vortex ring on the

upper side than the lower side. In contrast, the  $\alpha_{45}\beta_{22.5}$  case does not exhibit vortex shedding but instead displays two streamwise vortex legs attached to the plate's sides.

The vortex shedding frequency also changes with the orientation of the plate. To illustrate the relationship, Figure 6A presents the vortex shedding frequency in the parametric space of  $\alpha$  and  $\beta$ . When vortex shedding is absent, a zero value is assigned. The plot reveals that the vortex shedding frequency increases symmetrically along both the  $\alpha$  and  $\beta$  axes. When one orientation angle is at its maximum while the other is non-zero, no vortex shedding occurs. To account for the changes in both the vortex shedding frequency and the frontal area of the plate, the Strouhal number was calculated for each case using the formula  $St_{L'} = fL'/U_{\infty}$ , where  $f$  represents the vortex shedding frequency,  $U_{\infty}$  denotes the inflow velocity, and  $L'$  is the hydraulic diameter of the frontal area of the plate. Figure 6B demonstrates the change in the Strouhal number with orientation, mirroring the behavior of the vortex shedding frequency. The maximum increase in the Strouhal number is 80%. A similar observation was reported by Chen and Fang (1996), where a decrease in the angle of attack of an inclined plate (comparable to increasing  $\alpha$  and/or  $\beta$  in our study) resulted in an increase in the Strouhal number of its vortex shedding.

### 4.2 Whisker signals

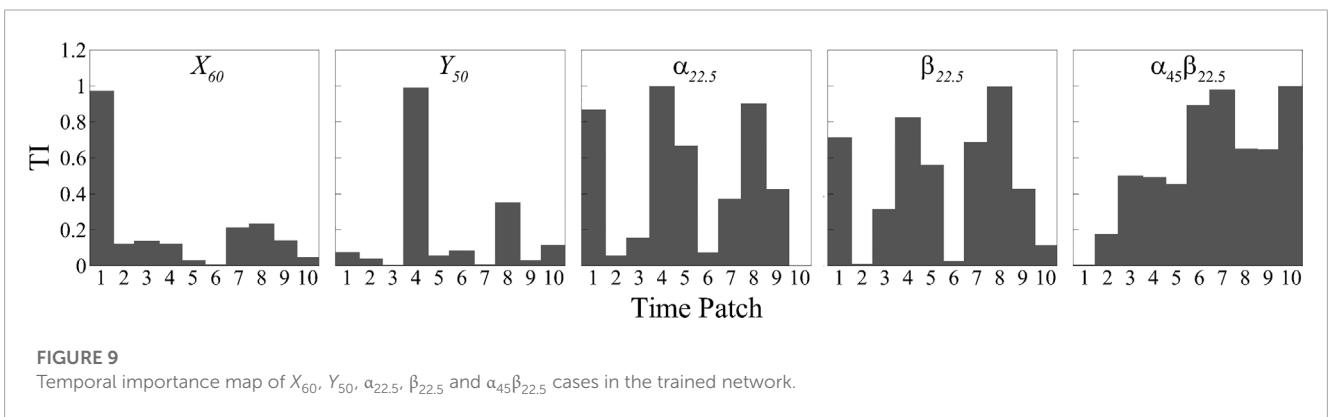
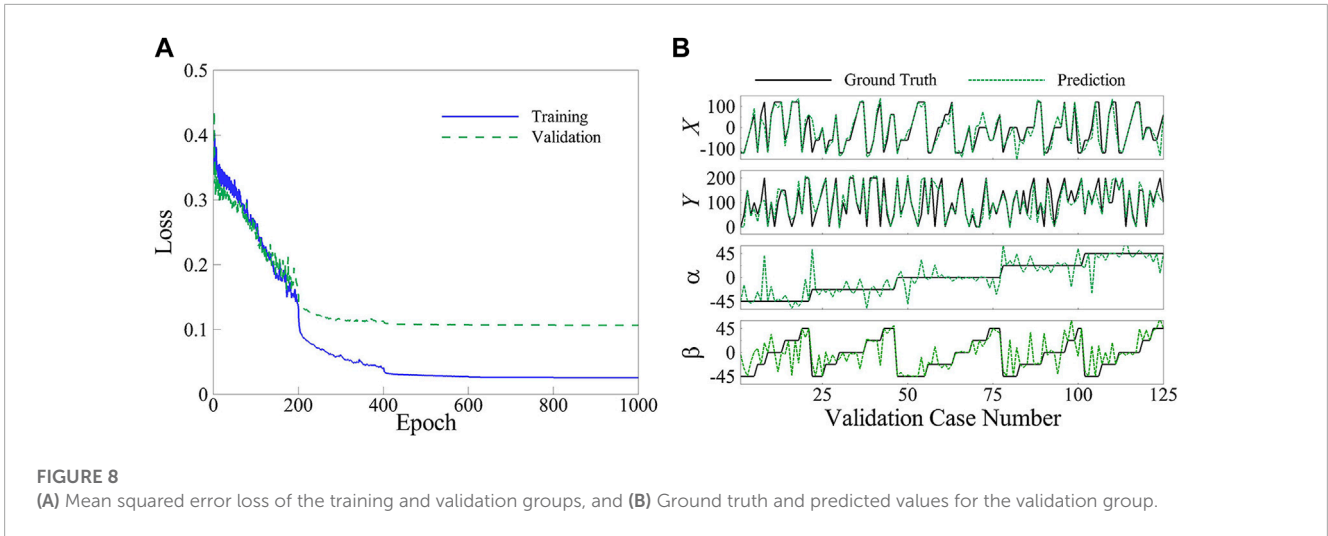
We partition the whisker arrays into four regions: LT (left top), LB (left bottom), RT (right top), and RB (right bottom) whiskers, and calculated the average signals,  $M_{x'}$  and  $M_{y'}$ , across the whiskers within each region. Figure 7 presents the phase-averaged time history of  $M_{x'}$  and  $M_{y'}$  in each region in each exemplary case. Note that  $M_{x'}$  is mostly negative values in all the cases, while  $M_{y'}$  exhibits both positive and negative values. The behaviour of the signals aligns with the temporal and spatial dynamics of the vortices in each case, (e.g. Frequency). In the  $X_{60}$  case, the signals



are evenly distributed across the regions since the vortices impact the seal head at the center. In the  $Y_{50}$  case, where the wake translates rightward, both  $M_{x_i}$  and  $M_{y_i}$  are stronger on the right whiskers. The right bottom whiskers experience stronger signals compared to the right top whiskers because of the impact location of the

vortices. In the  $\alpha_{22.5}$  case, the left whiskers experience stronger signals than the right whiskers, consistent with the plate rotation that generates stronger vortices on the left side. Additionally, the left bottom whiskers experience stronger moments than the left top whiskers. In the  $\beta_{22.5}$  case, interestingly, despite the upper vortex





ring being stronger than the bottom one, the bottom whiskers experience stronger  $M_{x'}$  than the top whiskers. This is attributed to the upward inclination of the wake, causing the upper vortex ring moving away from the whiskers. Also interestingly, different from  $M_{x'}$ ,  $M_{y'}$  shows no variation across the regions in this case. In the  $\alpha_{45}\beta_{22.5}$  cases, both  $M_{x'}$  and  $M_{y'}$  are almost constant with time while the left bottom whiskers have stronger moments compared to the other whiskers due to the stronger wake in this region.

Overall, the signals at the root of the whisker arrays exhibit temporal and spatial patterns that align with the dynamics of the vortices. There is a clear correlation between the arrival of the vortices and the temporal variation in the signals. The regions where the vortices directly impact the whiskers tend to exhibit stronger signal amplitudes. The distinct vortex dynamics in each case result in different temporal and spatial patterns of the signals, providing an explanation for the discernible differences among the cases.

### 4.3 Network results

The neural network was trained successfully to predict the location and orientation of the upstream plate. Figure 8A illustrates the best mean squared error (MSE) losses of the training and

validation groups over 1,000 epochs, indicating the convergence of the network. The final MSE loss for the training and validation corresponding to the best fold is 0.0255 and 0.1064, respectively. The decision to limit the epochs to 1,000 was based on the observation that the drop in MSE loss for both groups during the last 100 epochs is less than 0.1%. The accuracy of the trained network corresponding to the best fold is depicted in Figure 8B through a comparison of the ground truth and predicted locations and orientations of the upstream plate in the validation cases. The accuracy metric is quantified using the following formula:

$$Accuracy = \left( 1 - \frac{\sum \left( \frac{|y_{GT} - y_{Pr}|}{Range} \right)}{N} \right) \times 100$$

where  $y_{GT}$  represents the ground truth value,  $y_{Pr}$  denotes the predicted value,  $Range$  represents the range of changes of  $y_{GT}$ , and  $N$  is the number of validation cases. The obtained 5-fold-average accuracy of the parameters  $X$ ,  $Y$ ,  $\alpha$ , and  $\beta$  is  $91.4 \pm 1.1\%$ ,  $91.0 \pm 1.1\%$ ,  $84.8 \pm 2.6\%$ , and  $83.3 \pm 0.6\%$ , respectively, resulting in an average accuracy of  $87.6 \pm 0.7\%$  for the network.

An essential capability of the interpretable network is its ability to generate temporal and spatial importance of the signals within the trained network, allowing for the identification of significant temporal-spatial signal patterns that contribute to the network's

predictions. In our study, as outlined in Section 3.3, we utilized a temporal dimension of 80 and a time interval of  $t = 8$ , resulting in 10 distinct time patches for the prediction. In the  $X_{60}$  and  $Y_{50}$  cases, these 10 time patches correspond to two consecutive cycles, while in the  $\alpha_{22.5}$  and  $\beta_{22.5}$  cases, they correspond to three consecutive cycles.

In Figure 9, we present the temporal importance (TI) analysis of the signals in each representative case. Except for the  $\alpha_{45}\beta_{22.5}$  case, the results highlight that the temporal importance is not uniformly distributed across all time instances. Instead, certain time points exhibit significantly higher importance values, while others demonstrate relatively lower importance values. This disparity suggests that specific time instances play a crucial role in the prediction process, while others have a lesser impact. Furthermore, it is noteworthy that the temporal importance patterns differ across all the cases. Each case exhibits a unique distribution of importance values across the time patches. In the  $\alpha_{45}\beta_{22.5}$  case where vortex shedding is absent, the temporal importance is nearly uniform across the time patches within a cycle. It is because the flow field remains nearly unchanged with time.

The variation in temporal importance patterns highlights the importance of considering the temporal dynamics and their specific correlations with the underlying flow structures when interpreting the prediction process. To illustrate this relationship, we present the vortex structures determined by the  $\lambda_2$  criterion and contours of flow velocity magnitude at two planes intersecting the whisker arrays at the first five individual time patches in each representative case in Figure 10. Note that the flow velocity in the case of freestream passing the seal head without an upstream plate was subtracted from the flow velocity in other cases where wakes were present to focus on the disturbance induced by the wake. To facilitate analysis, the temporal importance map is also included in each case. These five patches cover either one or more than one cycle of the wake. The  $\alpha_{45}\beta_{22.5}$  case is not included as the temporal importance and flow structure remains unchanged in this case.

In the  $X_{60}$  case, time patch 1 exhibits significantly higher importance compared to the other time patches. This observation corresponds to the vortex dynamics, where at time patch 1, the lower vortex ring actively interacts with the whiskers. In contrast, at the other time patches, either the top vortex ring arrives in the whisker area but lacks strong interaction due to its distance from the whiskers, or the bottom vortex ring has already moved past the whiskers. The flow velocity contour reveals that the flow disturbance is strongest at time patch 1 around the whisker compared to other time patches.

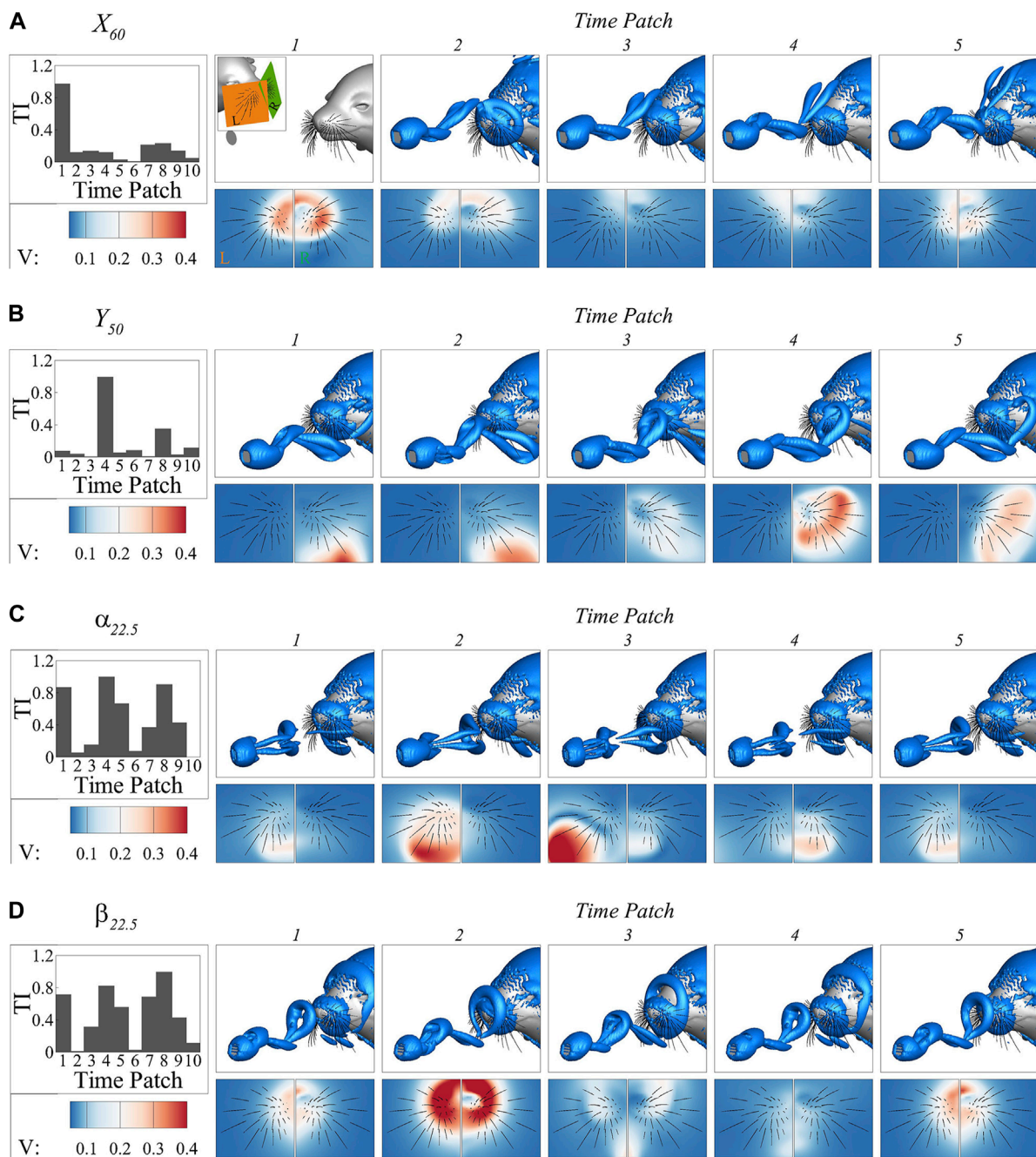
In the  $Y_{50}$  case, time patch 4 demonstrates significantly higher importance compared to other time patches. This time patch corresponds to the instances when the top vortex ring interacts with the right whiskers. At the other time patches, either the top vortex ring has not yet reached the whiskers or has already moved past them. In this case, the bottom vortex ring has minimal effect due to its distance from the whiskers resulting from the translation of the upstream plate. The flow velocity contour also shows that the vortex ring interacts with the right whisker array and the disturbance is the strongest on most whiskers at time patch 4.

In the case of  $\alpha_{22.5}$  time patch 4 exhibits the highest importance, while time patches 1 and 5 are slightly less important, and time patches 2 and 3 have minimal importance. Interestingly, upon

analyzing the vortex dynamics, it is observed that time patches 2 and 3 capture the interactions between the left whiskers and the stronger vortex ring from the left side. Conversely, time patch 4 captures the interactions between the right whiskers and the weaker vortex ring shed from the right edge. Time patches 1 and 5 represent the transitional stages, where the weaker vortex ring is about to leave and the stronger vortex ring is about to arrive. The flow velocity contour further confirms that the flow disturbance is strongest at time patches 2 and 3 due to the presence of the stronger vortex, whereas it is weaker at time patch 4 attributed to the weaker vortex ring. One possible explanation for the higher importance of the weaker vortex in this case is that the stronger vortex ring shed from the left edge aligns towards the center of the nose, similar to the vortices observed in other cases. In contrast, the weaker vortex ring is oriented perpendicular to it and only interacts with the right whisker arrays. This unique interaction provides additional crucial orientation information, making the weaker vortex more influential in the prediction process.

In the case of  $\beta_{22.5}$ , time patches 1, 4, and 5 exhibit greater significance compared to time patches 2 and 3. Through vortex dynamics analysis, it is observed that time patches 1 and 5 capture the arrival of the top vortex at the whiskers, while time patch 4 captures the interactions between the bottom vortex and the whiskers. Interestingly, time patch 2 captures the strong interactions between the top vortex ring and the whiskers, but its contribution is nearly negligible. This observation is further supported by the flow velocity contour. The flow disturbance during time patches 1, 2, and 5 is primarily associated with the top vortex, whereas during time patch 4, it is linked to the bottom vortex. Time patch 3 presents a scenario where disturbances from both vortex rings are present. Notably, similar to the  $\alpha_{22.5}$  case, the two vortex rings differ in strength, with the top ring being stronger than the bottom one, but the weaker ring plays the most important role in prediction. It is possible for the same reason that the weaker ring possesses a unique orientation and only interacts with the bottom whiskers. However, unlike the  $\alpha_{22.5}$  case, the significance of the other ring is increased in this case. It may be because the orientation information provided by the bottom ring is insufficient for accurate prediction, thereby necessitating the incorporation of additional information.

The current interpretable deep learning network also provides the importance of each whisker through spatial importance (SI) maps at individual time patch. Figure 11 illustrates the locations of the whiskers with the highest importance values during each of the first five time patches in each case. These whisker locations generally exhibit a strong correlation with areas of strong flow disturbance, as depicted in Figure 10. Furthermore, the spatial distribution of these important whiskers provides insights into the location and size of the vortex rings in the wake. For instance, in the  $X_{60}$  case, the important whiskers reflect the vortex ring impacting the center of the seal head. In the  $Y_{50}$  case, the important whiskers correlate with strong flow disturbance on the right side, and their extension reflects the extension of the vortex ring. In the  $\alpha_{22.5}$  case, the important whiskers capture both vortex rings, with one located on the left side and the other on the right side. Also, they reveal that the extension of the left vortex ring is larger than the right one. Similarly, in the  $\beta_{22.5}$  case, the important whiskers capture two vortex rings. In this scenario, they

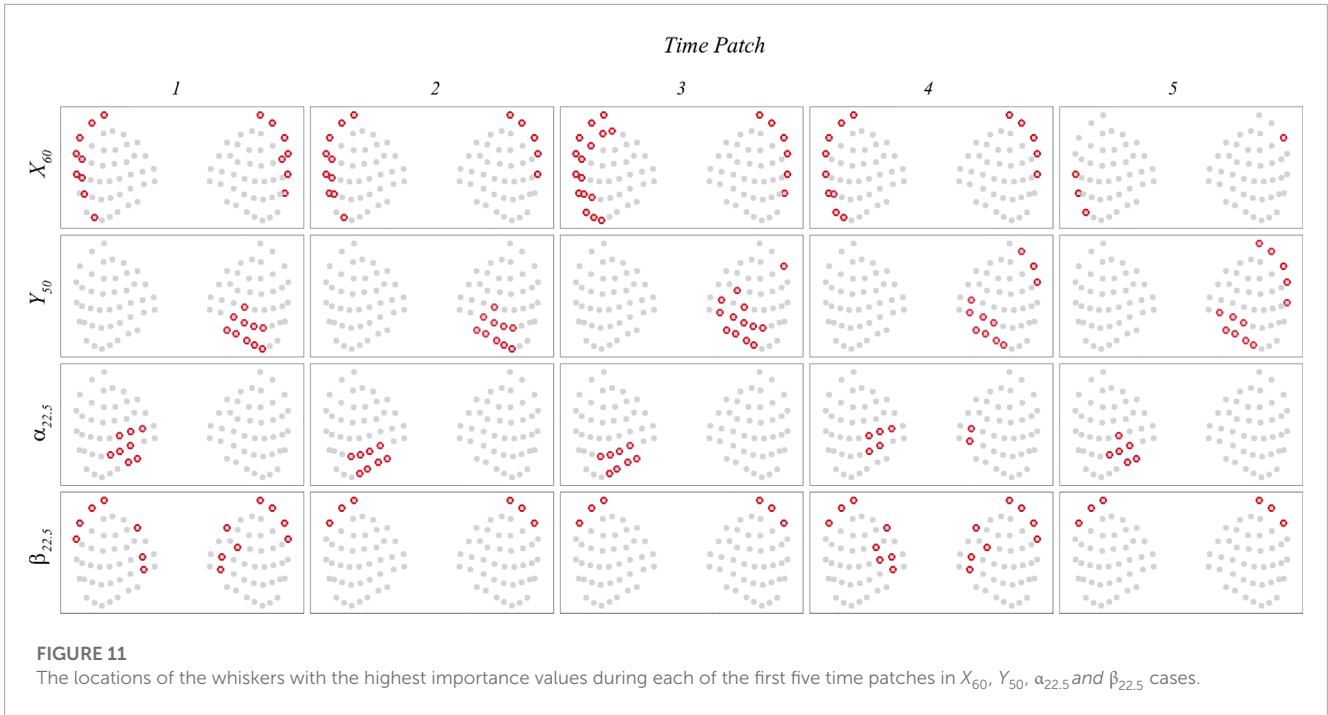


**FIGURE 10** The vortex structures determined by the  $\lambda_2$  criterion and contours of flow velocity magnitude at two planes intersecting the whisker arrays at the first five individual time patches of (A)  $X_{60}$ , (B)  $Y_{50}$ , (C)  $\alpha_{22.5}$ , and (D)  $\beta_{22.5}$  cases. Note that the flow velocity in the case of freestream passing the seal head without an upstream plate was subtracted from the flow velocity in other cases where wakes were present to focus on the disturbance induced by the wake.

capture that the higher vortex ring is larger and situated near the center, while the lower vortex ring is smaller and positioned closer to the bottom.

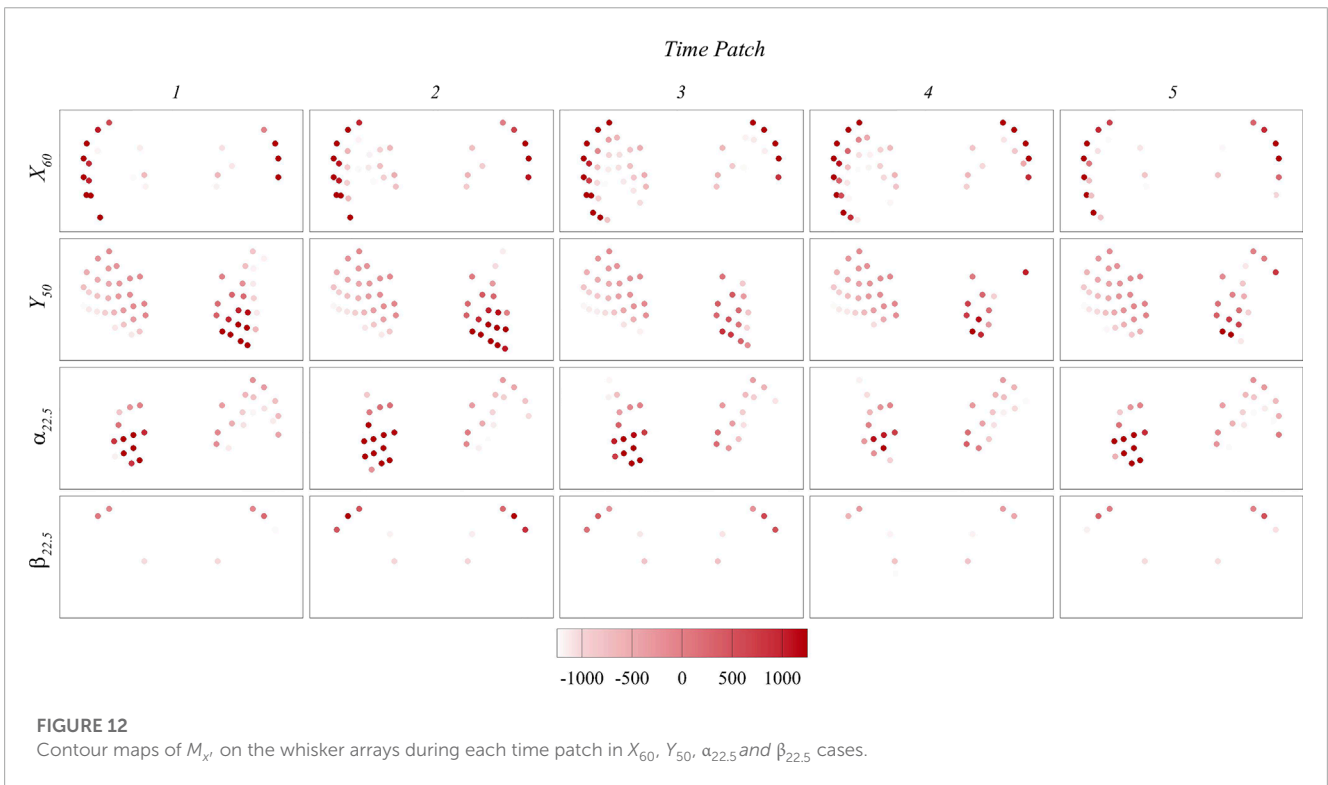
To gain a deeper understanding of the signals used for prediction, Figures 12, 13 present the maps of  $M_{x'}$  and  $M_{y'}$  on the whisker arrays during each time patch for each case. Note that  $M_{x'}$  represents the moment generated by the lift force, while

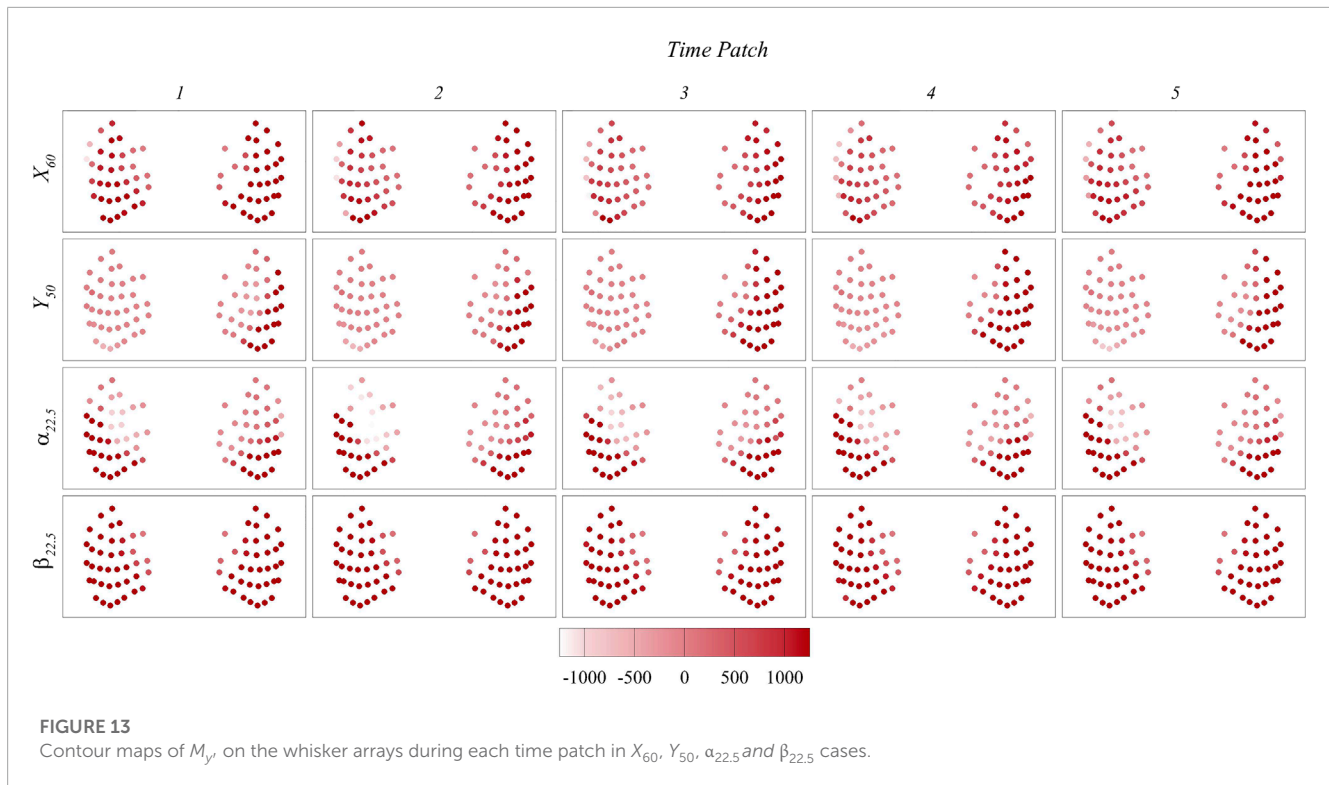
$M_{y'}$  represents the moment generated by the drag force. A strong correlation between the important whiskers and  $M_{x'}$  maps is observed across all cases, indicating the significant role of  $M_{x'}$  in the prediction process. Conversely, we did not observe any notable correlation with the  $M_{y'}$  maps, suggesting that  $M_{y'}$  has minimal impact on the prediction. Interestingly, the direction of  $M_{x'}$  plays a crucial role in the prediction, as the important



whiskers consistently correlate with positive  $M_{x'}'$  values across all cases or, in cases where positive  $M_{x'}'$  values are absent, with less negative  $M_{x'}'$  values. It is important to note that in our algorithm, the calculation of  $M_{x'}'$  involves subtracting the moment induced by the freestream flow. This subtraction is intended to

focus on the moment induced specifically by vortices. Further studies are needed to clarify whether the direction information remains important if the original moment, without subtracting the freestream moment, is used. Nevertheless, the results suggest that the bending moment generated by the lift force plays a crucial role in





the prediction process, and the direction of this moment may also be important.

## 5 Summary and conclusion

In this study, we developed a novel method that combines a computational fluid-structure interaction (FSI) model with interpretable deep-learning models to investigate the fundamental mechanisms of seal whisker sensing. We demonstrate that the method is able to accurately predict the location and orientation of a circular plate upstream of a seal's head by using the bending moment signals at the root of whisker arrays. The algorithm also reveals crucial temporal-spatial signal patterns utilized in the prediction process, enabling establishing correlations between important signal patterns, flow features, and characteristics of upstream obstacles for enhanced understanding.

The FSI model couples an incompressible flow solver and structural dynamics solver in a one-way coupling manner. We generated diverse wake characteristics and corresponding whisker dynamics by placing a circular plate upstream of a seal's head and systematically varying the location and orientation of the plate. Two types of wake patterns were observed. For most of the cases, the wake is characterized by two vortex rings shed from two edges of the plate with a phase difference. The translation of the upstream plate solely affects the position of the wake, without altering its structures. However, when the orientation of the plate is changed, it induces an inclination in the wake's trajectory, a shift in the orientation of the vortices, and an increase of the vortex shedding frequency.

Furthermore, the side farther from the inflow corresponds to the location of the stronger vortex, while the side closer to the inflow is associated with the weaker vortex. In contrast, for the cases where one of the orientation angles is the largest while the other is non-zero, no vortex shedding occurs in the wake. Instead, the wake is characterized by two streamwise vortex legs attached to the side of the plate. The whisker signals are represented by the bending moments at the root of each whisker. The signals exhibit temporal and spatial patterns that align with the dynamics of the vortices. There is a clear correlation between the arrival of the vortices and the temporal variation in the signals. The regions where the vortices directly impact the whiskers tend to exhibit stronger signal amplitudes.

We train the network model to learn the location and orientation of the upstream plate based on the whisker signals. The results show that the model is able to predict the location of the plate with the accuracy of 90%–92.0% and the orientation of the plate with the accuracy between 86% and 89%. The analysis on the temporal importance of the signals reveals that the prediction process primarily relies on the time instances when the wake vortices actively interact with the whiskers. When the vortices are at varying distances from the whiskers, such as when the upstream plate is translated, interactions with the vortices in close proximity to the whiskers have a more significant impact, while interactions with the distant vortex ring have minimal influence. In the context of predicting the orientation of the upstream plate, interactions with all vortices can be important. However, interactions involving the vortices that carry more orientation information tend to hold greater significance, even if these vortices do not cause the strongest disturbance and signals. It suggests that the vortices that provide

crucial orientation cues play a crucial role in the prediction process, even if their influence on the overall flow disturbance and signal strength is not the most pronounced. Spatially, the whiskers in stronger flow disturbance areas are more important in the prediction. The spatial arrangement of the important whiskers not only captures the vortex ring structure in the wake, but also provides valuable information about the location and size of the rings. In the context of the signals used for prediction, our results suggest that the bending moment generated by the lift force plays a crucial role in the prediction process. The direction of the moment may also be important as important whiskers are all associated with the bending moment in one direction.

Taken together, we demonstrate that the developed interpretable neural network is able to accurately predict the location and orientation of the upstream object of seal whisker arrays. More importantly, it allows for the identification of significant temporal-spatial signal patterns that contribute to the network's predictions, which can then be further correlated with flow structures, enabling the identification of the crucial flow features that are sensed for accurate prediction. These insights are crucial for the development of intelligent flow sensors capable of accurately perceiving and interpreting complex underwater environments.

We would like to acknowledge that our study utilized a simplified drag-lift model in our one-way coupling approach to estimate the hydrodynamic loadings along each whisker. This simplified model does not take into account the feedback effect of the whiskers on the surrounding flow. In reality, the presence of the whiskers would influence the surrounding flow, resulting in different drag and lift values that could potentially affect the outcomes of our research. However, we believe that the impact of the whiskers on the surrounding flow would be small due to the small diameter of the whiskers (with a mean diameter of 0.66 mm). This diameter is approximately 1-2 orders of magnitude smaller than the scale of the vortices in the wakes, which range from 20 mm to 120 mm. Considering this scale difference, we believe that the one-way coupling approach still provides valuable insights while significantly reducing the computational cost associated with conducting full FSI simulations. It is important to note that future investigations could explore the implications of the two-way coupling between the whiskers and the surrounding flow for a more comprehensive understanding of the hydrodynamic interactions.

In all the simulation cases, we observed that the frequency of the bending moments consistently matched the shedding frequency of the wake vortices. This behavior is a direct consequence of the one-way coupling approach, where the feedback effect of whisker dynamics on the flow was disregarded. It is worth noting, however, that in our previous research (Liu et al., 2019b), we discovered that the distinctive undulated geometry of seal whiskers can effectively mitigate vortex-induced vibration when exposed to steady flows. This unique characteristic enables the whiskers to better synchronize with the wake frequency. Nonetheless, it is important to acknowledge that further investigations are required to explore the interactions between vortex-induced vibration and wake-induced vibration. Understanding these complex dynamics

will provide valuable insights into the combined effects and enhance our comprehension of the overall system behavior.

Finally, seal whisker arrays display distinct spatial grid patterns combined with variations in whisker length and orientation. These patterns are believed to play crucial functional roles in enhancing the seals' sensing capabilities. For instance, it is believed that longer whiskers located on the caudal side of the array can provide a larger detection area, extending the reach and sensitivity to water movements. Moreover, the variation in whisker length and orientation may allow for high sensitivity in wide frequency range and flow direction. Further studies are needed to fully investigate these effects, particularly by comparing the sensing capabilities between different whisker arrangements. The current findings are limited to a single configuration, and more research is necessary to gain a comprehensive understanding.

## Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## Author contributions

XZ and QX conceived the original study. XZ, QX, DL, and GL designed the study. DB conducted numerical simulation and data analysis. YW implemented and trained the neural network. DB, YW, GL, DL, QX, and XZ collectively wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by NSF under grant number 214421. The computation was supported by XSEDE award TG-CTS180004 and MCH220042. This study also used the HPC systems of the UMS Advanced Computing Group.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Akyildiz, I. F., Pompili, D., and Melodia, T. (2004). Challenges for efficient communication in underwater acoustic sensor networks. *ACM Sigbed Rev.* 1 (2), 3–8. doi:10.1145/1121776.1121779
- Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M., and Schmid, C. (2021). “Vivit: A video vision transformer,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 6836–6846.
- Baevski, A., and Auli, M. (2018). *Adaptive input representations for neural language modeling*. arXiv preprint arXiv:1809.10853.
- Beam, H. R., Hildner, M., and Triantafyllou, M. (2013). Calibration and validation of a harbor seal whisker-inspired flow sensor. *Smart Mat. Struct.* 22, 014012. doi:10.1088/0964-1726/22/1/014012
- Beem, H. R. (2015). *Passive wake detection using seal whisker-inspired sensing*. Woods Hole, MA: Massachusetts Institute of Technology.
- Beem, H. R., and Triantafyllou, M. S. (2015). Wake-induced ‘slaloming’ response explains exquisite sensitivity of seal whisker-like sensors. *J. Fluid Mech.* 783, 306–322. doi:10.1017/jfm.2015.513
- Bertasius, G., Wang, H., and Torresani, L. (2021). July Is space-time attention all you need for video understanding? In *ICML 2*. doi:10.48550/arXiv.2102.05095
- Bodaghi, D., Jiang, W., Xue, Q., and Zheng, X. (2021a). Effect of supraglottal acoustics on fluid–structure interaction during human voice production. *J. Biomechanical Eng.* 143 (4), 041010. doi:10.1115/1.4049497
- Bodaghi, D., Xue, Q., Zheng, X., and Thomson, S. (2021b). Effect of subglottic stenosis on vocal fold vibration and voice production using fluid–structure–acoustics interaction simulation. *Appl. Sci.* 11 (3), 1221. doi:10.3390/app11031221
- Chen, J. M., and Fang, Y. C. (1996). Strouhal numbers of inclined flat plates. *J. wind Eng. industrial aerodynamics* 61 (2–3), 99–112. doi:10.1016/0167-6105(96)00044-x
- Chutia, S., Kakoty, N. M., and Deka, D. (2017). A review of underwater robotics, navigation, sensing techniques and applications. *Proc. Adv. Robotics*, 1–6.
- Dehnhardt, G., Mauck, B., and Bleckmann, H. (1998). Seal whiskers detect water movements. *Nature* 394 (6690), 235–236. doi:10.1038/28303
- Dehnhardt, G., Mauck, B., Hanke, W., and Bleckmann, H. (2001). Hydrodynamic trail-following in harbor seals (*Phoca vitulina*). *Science* 293 (5527), 102–104. doi:10.1126/science.1060514
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). *An image is worth 16x16 words: Transformers for image recognition at scale*. arXiv preprint arXiv:2010.11929.
- Eberhardt, W. C., Wakefield, B. F., Murphy, C. T., Casey, C., Shakhsheer, Y., Calhoun, B. H., et al. (2016). Development of an artificial sensor for hydrodynamic detection inspired by a seal’s whisker array. *Bioinspir. Biomim.* 11, 056011. doi:10.1088/1748-3190/11/5/056011
- Elshalakani, M., Muthuramalingam, M., and Bruecker, C. (2020). A deep-learning model for underwater position sensing of a wake’s source using artificial seal whiskers. *Sensors* 20 (12), 3522. doi:10.3390/s20123522
- Geng, B., Xue, Q., Zheng, X., Liu, G., Ren, Y., and Dong, H. (2017). The effect of wing flexibility on sound generation of flapping wings. *Bioinspiration biomimetics* 13 (1), 016010. doi:10.1088/1748-3190/aa8447
- Glick, R., Muthuramalingam, M., and Brücker, C. (2021). Fluid-structure interaction of flexible whisker-type beams and its implications for flow sensing by pair-wise correlation. *Fluids* 6 (3), 102. doi:10.3390/fluids6030102
- Graff, M. M. (2016). “Mechanisms of whisker based flow sensing.” Doctoral dissertation (Evanston, IL: Northwestern University).
- Hanke, W., Witte, M., Miersch, L., Brede, M., Oeffner, J., Michael, M., et al. (2010). Harbor seal vibrissa morphology suppresses vortex-induced vibrations. *J. Exp. Biol.* 213 (15), 2665–2672. doi:10.1242/jeb.043216
- Hans, H., Miao, J. M., and Triantafyllou, M. S. (2014). Mechanical characteristics of harbor seal (*Phoca vitulina*) vibrissae under different circumstances and their implications on its sensing methodology. *Bioinspiration biomimetics* 9 (3), 036013. doi:10.1088/1748-3182/9/3/036013
- Hans, H., Miao, J., Weymouth, G., and Triantafyllou, M. (2013). “Whisker-like geometries and their force reduction properties,” in 2013 MTS/IEEE OCEANS-Bergen, Bergen, Norway, 10–14 June 2013 (IEEE), 1–7. doi:10.1109/OCEANS-Bergen.2013.6608113
- Hendrycks, D., and Gimpel, K. (2016). *Gaussian error linear units (gelus)*. arXiv preprint arXiv:1606.08415.
- Ho, J., Kalchbrenner, N., Weissenborn, D., and Salimans, T. (2019). *Axial attention in multidimensional transformers*. arXiv preprint arXiv:1912.12180.
- Jiang, Y., Ma, Z., and Zhang, D. (2019). Flow field perception based on the fish lateral line system. *Bioinspiration biomimetics* 14 (4), 041001. doi:10.1088/1748-3190/ab1a8d
- Jiang, Y., Zhao, P., Ma, Z., Shen, D., Liu, G., and Zhang, D. (2020). Enhanced flow sensing with interfacial microstructures. *Biosurface Biotribology* 6 (1), 12–19. doi:10.1049/bsbt.2019.0043
- Kinsey, J. C., Eustice, R. M., and Whitcomb, L. L. (2006). “September. A survey of underwater vehicle navigation: Recent advances and new challenges,” in IFAC conference of manoeuvring and control of marine craft 88. 1–12.
- Kottapalli, A. G. P., Asadnia, M., Miao, J. M., and Triantafyllou, M. S. (2015). “January. Harbor seal whisker inspired flow sensors to reduce vortex-induced vibrations,” in 2015 28th IEEE International Conference on Micro Electro Mechanical Systems (MEMS) (IEEE), 889–892.
- Kröger, R. (2008). “The physics of light in air and water,” in *Sensory Evolution on the Threshold-Adaptations in Secondly Aquatic Vertebrates* (University of California Press), 113–119.
- Lee, D., Kim, G., Kim, D., Myung, H., and Choi, H. T. (2012). Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean. Eng.* 48, 59–68. doi:10.1016/j.oceaneng.2012.04.006
- Leonard, J. J., and Bahr, A. (2016). *Autonomous underwater vehicle navigation*. Springer handbook of ocean engineering, 341–358.
- Liu, G., Jiang, Y., Wu, P., Ma, Z., Chen, H., and Zhang, D. (2023). Artificial whisker sensor with undulated morphology and self-spread piezoresistors for diverse flow analyses. *Soft Robot.* 10 (1), 97–105. doi:10.1089/soro.2021.0166
- Liu, G., Geng, B., Zheng, X., Xue, Q., Dong, H., and Lauder, G. V. (2019a). An image-guided computational approach to inversely determine *in vivo* material properties and model flow-structure interactions of fish fins. *J. Comput. Phys.* 392, 578–593. doi:10.1016/j.jcp.2019.04.062
- Liu, G., Geng, B., Zheng, X., Xue, Q., Wang, J., and Dong, H. (2018). “An integrated high-fidelity approach for modeling flow-structure interaction in biological propulsion and its strong validation,” in 2018 AIAA Aerospace Sciences Meeting, 1543.
- Liu, G., Jiang, W., Zheng, X., and Xue, Q. (2021). Flow-signal correlation in seal whisker array sensing. *Bioinspiration Biomimetics* 17 (1), 016004. doi:10.1088/1748-3190/ac363c
- Liu, G., Xue, Q., and Zheng, X. (2019b). Phase-difference on seal whisker surface induces hairpin vortices in the wake to suppress force oscillation. *Bioinspiration Biomimetics* 14 (6), 066001. doi:10.1088/1748-3190/ab34fe
- Mittal, R., Dong, H., Bozkurt, M., Najjar, F. M., Vargas, A., and Von Loebbecke, A. (2008). A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries. *J. Comput. Phys.* 227 (10), 4825–4852. doi:10.1016/j.jcp.2008.01.028
- Morrison, H. E., Brede, M., Dehnhardt, G., and Leder, A. (2016). Simulating the flow and trail following capabilities of harbour seal vibrissae with the Lattice Boltzmann Method. *J. Comput. Sci.* 17, 394–402. doi:10.1016/j.jocs.2016.04.004
- Müller, R., and Kuc, R. (2007). Biosonar-inspired technology: Goals, challenges and insights. *Bioinspiration biomimetics* 2 (4), S146–S161. doi:10.1088/1748-3182/2/4/s04
- Murphy, C. T. (2013). *Structure and function of pinniped vibrissae*. Doctoral dissertation. St. Petersburg, FL: University of South Florida.
- Rinehart, A., Shyam, V., and Zhang, W. (2017). Characterization of seal whisker morphology: Implications for whisker-inspired flow control applications. *Bioinspiration Biomimetics* 12 (6), 066005. doi:10.1088/1748-3190/aa8885
- Sayegh, M., Daraghma, H., Mekid, S., and Bashmal, S. (2022). Review of recent bio-inspired design and manufacturing of whisker tactile sensors. *Sensors* 22 (7), 2705. doi:10.3390/s22072705
- Schrope, M. (2002). Whale deaths caused by US Navy’s sonar. *Nature* 415 (6868), 106. doi:10.1038/415106a
- Stutters, L., Liu, H., Tiltman, C., and Brown, D. J. (2008). Navigation technologies for autonomous underwater vehicles. *IEEE Trans. Syst. Man, Cybern. Part C Appl. Rev.* 38 (4), 581–589. doi:10.1109/tsmcc.2008.919147
- Tietz, M., Fan, T. J., Nouri, D., and Bossan, B. (2017). *skorch: A scikit-learn compatible neural network library that wraps PyTorch*.
- Van Kan, J. J. M. (1986). A second-order accurate pressure-correction scheme for viscous incompressible flow. *SIAM J. Sci. Stat. Comput.* 7 (3), 870–891. doi:10.1137/0907059
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* 30. doi:10.5555/3295222.3295349
- Wang, Q., Li, B., Xiao, T., Zhu, J., Li, C., Wong, D. F., et al. (2019). *Learning deep transformer models for machine translation*. arXiv preprint arXiv:1906.01787.
- Wang, S., Xu, P., Wang, X., Zheng, J., Liu, X., Liu, J., et al. (2022). Scaling autoregressive video models, 97 (15), 107219. arXiv preprint arXiv:1906.02634.

Wieskotten, S., Dehnhardt, G., Mauck, B., Miersch, L., and Hanke, W. (2010). Hydrodynamic determination of the moving direction of an artificial fin by a harbour seal (*Phoca vitulina*). *J. Exp. Biol.* 213 (13), 2194–2200. doi:10.1242/jeb.041699

Wieskotten, S., Mauck, B., Miersch, L., Dehnhardt, G., and Hanke, W. (2011). Hydrodynamic discrimination of wakes caused by objects of different size or shape in a harbour seal (*Phoca vitulina*). *J. Exp. Biol.* 214 (11), 1922–1930. doi:10.1242/jeb.053926

Yang, Y., Chen, J., Engel, J., Pandya, S., Chen, N., Tucker, C., et al. (2006). Distant touch hydrodynamic imaging with an artificial lateral line. *Proc. Natl. Acad. Sci.* 103 (50), 18891–18895. doi:10.1073/pnas.0609274103

Zheng, X., Kamat, A. M., Cao, M., and Kottapalli, A. G. P. (2021). Creating underwater vision through wavy whiskers: A review of the flow-sensing mechanisms and biomimetic potential of seal whiskers. *J. R. Soc. Interface* 18 (183), 20210629. doi:10.1098/rsif.2021.0629

Zheng, X., Kamat, A. M., Krushynska, A. O., Cao, M., and Kottapalli, A. G. P. (2022). 3D printed graphene piezoresistive microelectromechanical system sensors to explain the ultrasensitive wake tracking of wavy seal whiskers. *Adv. Funct. Mat.* 32, 2207274. doi:10.1002/adfm.202207274

Zimmerman, A., Bai, L., and Ginty, D. D. (2014). The gentle touch receptors of mammalian skin. *Science* 346 (6212), 950–954. doi:10.1126/science.1254229