Joint event Euromech Colloquium on Data-Driven Fluid Dynamics/2nd ERCOFTAC Workshop on Machine Learning for Fluid Dynamics, 2-4 April 2025, London, UK

# Combining Physics-Informed Clustering and Deep Learning to Identify Pathologies and Defects from CFD Data

**Riccardo Margheritti**[1], Onofrio Semeraro[2], Maurizio Quadrio[1], Giacomo Boracchi[1]

[1] *Politecnico di Milano,* [2] *LISN-CNRS, Universite' Paris-Saclay*

*Corresponding author: Riccardo Margheritti (riccardo.margheritti@mail.polimi.it)*

Machine learning (ML) is widely used in Computational Fluid Dynamics (CFD) to approximate complex input-output relationships, accelerating or replacing numerical simulations. However, classifying flow fields by inferring labels that are not directly computable from equations, remains almost unexplored [1]. Existing methods [2] rely on expert-driven features from specific flow sub-regions, limiting generalization. We propose a novel approach combining physics-informed clustering with modern deep learning for a comprehensive flow field classification.

Inspired by Callaham et al. [3], we apply Bayesian Gaussian Mixture Models (BGMM) to cluster flow fields based on RANS equation-derived features, ensuring clusters reflect underlying physics. Clusters characterized by geometric properties and statistical features form a point cloud fed to a PointNet-based architecture, which captures spatial relationships and provides permutation invariance.

We test our method in two scenarios: NACA airfoils with defects (e.g., bumps, cavities) and real human upper airways showing pathologies. Approximately 4000 simulations for airfoils and airway datasets validate the effectiveness of our framework, accurately identifying airfoil defects and pathologies while reducing the computational costs of processing such high-dimensional data. By leveraging all CFD data and incorporating spatial and physical relationships, our approach offers a scalable, holistic solution for analyzing complex aerodynamic and biomedical flows.

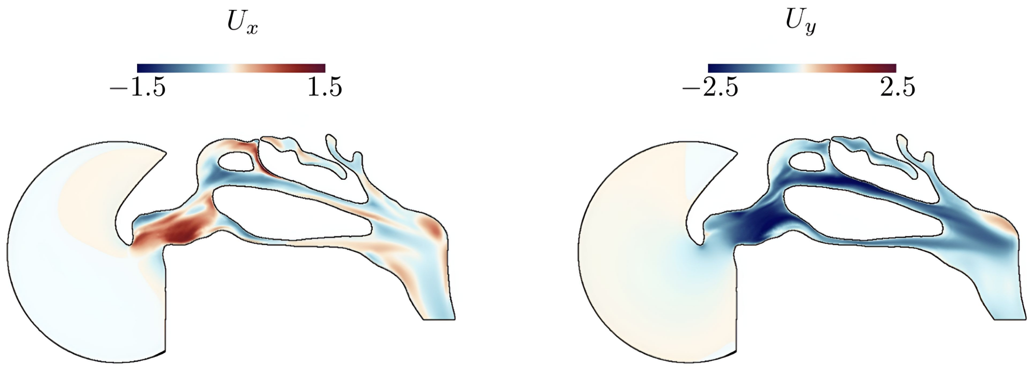


Figure 1: The $x$ and $y$ components of the velocity in a sagittal section obtained from a RANS simulation in real human upper airways.

## References

[1] Ettore Saetta and Renato Tognaccini. Identification of flow field regions by Machine Learning. In *AIAA SCITECH Forum*. American Institute of Aeronautics and Astronautics, 2022.

[2] Andrea Schillaci, Maurizio Quadrio, Carlotta Pipolo, Marcello Restelli, and Giacomo Boracchi. Inferring Functional Properties from Fluid Dynamics Features. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 4091–4098, 2021.

[3] Jared L. Callaham, James V. Koch, Bingni W. Brunton, J. Nathan Kutz, and Steven L. Brunton. Learning dominant physical processes with data-driven balance models. *Nature Communications*, 12(1):1016, 2021.