Digital Twins in Engineering & Artificial Intelligence and Computational Methods in Applied Science,
DTE - AICOMAS 2025
February 17-21, 2025, Paris, France

Leveraging Machine Learning for CFD Flow Field Classification

Riccardo Margheritti*, Onofrio Semeraro†, Maurizio Quadrio‡, Giacomo Boracchi‡

*Politecnico di Torino
Corso Duca degli Abruzzi, 24 - 10129 Torino (TO), Italy
Riccardo.margheritti@polito.it
Riccardo.margheritti@mail.polimi.it

† LISN-CNRS, Universite' Paris-Saclay
Rue René Thom (Bât Digitéo), 91190 - Gif-sur-Yvette, France
onofrio.semeraro@universite-paris-saclay.fr

‡ Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
Maurizio.quadrio@polimi.it
Giacomo.boracchi@polimi.it

Key Words: *Computational Fluid Dynamics, Flow Field Classification, Machine Learning, Physics-based Clustering, Deep Learning Architecture*

ABSTRACT

Machine learning (ML) is widely used in Computational Fluid Dynamics (CFD) to estimate sophisticated non-linear input-output relations embedded in the equations of motion, thereby speeding up or sometimes replacing numerical simulations [1]. However, flow field classification, namely, inferring labels that cannot be computed from explicit equations, has been poorly investigated [2]. Given the huge dimensionality of CFD data, existing classifiers for CFD simulations [3] rely on extracting expert-driven features from specific portions of the entire flow field, which are then classified by a traditional ML model. This solution is in contrast with a modern design paradigm for data-driven models and might limit the generalization abilities of the learned classifier.

We present a novel methodology that addresses these limitations by using entirely the CFD data while still extracting essential physical information for classification. Inspired by Callaham et al. [4], our approach relies on a physics-informed clustering using Bayesian Gaussian Mixture Models (BGMM), which extract a manageable number of clusters representing distinct physical phenomena in the flow field. By leveraging the RANS equations as clustering features, we ensure that the clusters are grounded in the underlying physics of the flow. Additionally, we associate each cluster with features based on the statistical properties of the values within the cluster, such as mean values of flow variables and turbulence quantities, as well as geometric properties of the cluster.

Our study focuses on simulations of NACA 4-digit airfoils with various defects, such as bumps and cavities. We conducted approximately 3000 simulations to generate a comprehensive dataset capturing a wide range of aerodynamic behaviors. Since flow fields can be divided into a varying number of clusters in random order, we treat the resulting clusters as points in a point cloud, each characterized by its geometric coordinates and a feature vector. This enables the use of point cloud classification techniques like PointNet++ [5] to analyze the flow behavior, providing permutation invariance with respect to input clusters and capturing spatial relationships between clusters. PointNet++ enables us to efficiently perform a multivariate regression on the simulations, predicting the characteristics of airfoil defects based on the spatial and physical information embedded in the clustered data.

Our experiments demonstrate that the proposed method effectively classifies flow fields corresponding to different airfoil defects, significantly reducing computational requirements without compromising

accuracy. By leveraging all the CFD data and associating features based on the values within each cluster, we capture both the spatial and physical relationships inherent in the data. This comprehensive approach enables a more accurate and holistic understanding of the flow fields compared to methods that rely only on specific features computed from the flow field.

This methodology offers a scalable solution for analyzing complex aerodynamic simulations by integrating physics-informed clustering with point cloud deep learning. Additionally, it enables efficient analysis of high-dimensional CFD data, easing the design and optimization in aerodynamics and potentially extending to other fluid dynamics areas.

REFERENCES

[1] Brunton, S., Noack, B., Koumoutsakos, P.: Machine learning for fluid mechanics. In: 2020 Annual Review of Fluid Mechanics, vol. 52, pp. 477–508.

[2] E. Saetta and R. Tognaccini, “Identification of Flowfield Regions by Machine Learning,” AIAA Journal, vol. 61, no. 4, pp. 1503–1518, Apr. 2023, doi: 10.2514/1.J061907.

[3] Schillaci, A., Quadrio, M., Pipolo, C., Restelli, M., Boracchi, G.: Inferring functional properties from fluid dynamics features. In: 2020 25th International Conference on Pattern Recognition (ICPR), pp. 4091–4098.

[4] Callaham, J. L., Koch, J. V., Brunton, B. W., Kutz, J. N., & Brunton, S. L. (2021). Learning dominant physical processes with data-driven balance models. *Nature Communications*, 12, Article number: 5636.

[5] Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 5099–5108).