POLITECNICO DI TORINO
Repository ISTITUZIONALE

Leveraging Computational Geometry for Data Augmentation in Medical Flow Fields Classification

Original
Leveraging Computational Geometry for Data Augmentation in Medical Flow Fields Classification / Margheritti, R., Schillaci, A., Pipolo, C., Quadrio, M., Boracchi, G.. - ELETTRONICO. - 2582:(2025), pp. 109-122. (Engineering Applications of Neural Networks 26th International Conference, EANN 2025 Limassol (CY) June 26–29, 2025) [10.1007/978-3-031-96199-1_9].

Availability:
This version is available at: 11583/3001510 since: 2025-07-03T08:19:45Z

Publisher:
Springer

Published
DOI:10.1007/978-3-031-96199-1_9

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/978-3-031-96199-1_9

(Article begins on next page)

10 October 2026

Leveraging Computational Geometry for Data Augmentation in Medical Flow Fields Classification

Riccardo Margheritti[1,2], Andrea Schillaci[1], Carlotta Pipolo[3], Maurizio Quadrio[1], and Giacomo Boracchi[1]

[1] Politecnico di Milano, 20133 Milano MI, ITA
[2] Politecnico di Torino, 10129 Torino TO, ITA
[3] Unita' di Otorinolaringoiatria ASST Santi Paolo e Carlo, 20133 Milano MI, ITA

Abstract. The application of Machine Learning (ML) to Computational Fluid Dynamics (CFD) has gained significant attention due to its potential in speeding up simulations and approximating numerical solutions of physical equations. Still, the dominant role of ML, which is to infer expert labels that cannot be calculated from explicit equations, has received much less attention in CFD. A major challenge in this direction is the scarcity of large, annotated datasets required to train robust ML models for flow field classification. In this work, we address the problem of training a ML model to classify CFD flow fields, inferring pathologies affecting the human upper airways. We propose a novel data augmentation method to address this limitation which involves an automated pipeline to extract CFD-ready surfaces from CT scans and a computational geometry technique to generate synthetic training samples. By defining deformation functions for specific pathologies on a reference surface and mapping these to healthy anatomical surfaces, we create a large and diverse training set with minimal expert supervision. This method allows for the generation of a dataset with high anatomical variability and well-defined labels, enhancing the model's ability to generalize to unseen geometries. We demonstrate that a Neural Network (NN) can accurately classify two common nasal pathologies, septal deviation and turbinate hypertrophy, achieving strong performance on real pathological patient data despite being trained solely on synthetic samples.

Keywords: Data Augmentation, Computational Fluid Dynamics, Machine Learning, Functional Mapping, Nasal Pathology Classification

1 Introduction

The use of Machine Learning (ML) in Computational Fluid Dynamics (CFD) has rapidly increased due to improved computational power and the potential of deep learning in fluid mechanics [1]. ML techniques are widely used to estimate complex non-linear input-output relationships in motion equations, enhancing or even replacing traditional CFD simulations. As a result, much work has focused

on improving turbulence modeling [2], flow field reconstruction using Convolutional Neural Networks (CNNs) [3], and reducing computational costs with physics-informed deep learning [4]. However, the primary use of ML, which is to infer labels that cannot be computed from equations (but rather defined by expert observations), has been poorly investigated in CFD. This is largely due to the difficulty in collecting CFD data and the vast amount of information contained within a single CFD flow field, which makes it impractical to feed directly into a ML model. As a result, training a Neural Network (NN) with CFD input becomes particularly challenging.

In this work, we address the classification of nasal pathologies, specifically, *septal deviation* and *turbinate hypertrophy*, by leveraging CFD simulations of airflow in the human upper airways. Our approach is motivated by the work of Schillaci et al. [5], which demonstrated the potential of combining ML and CFD for classifying parametric deformation in simplified 3D shapes of the nasal cavities. However, the simplified geometries used in [5] fail to capture the anatomical complexity and variability of human airways, limiting their applicability in clinical practice. To overcome this limitation, we introduce a novel data augmentation method that generates a realistic and labeled training set of CFD data based on real anatomical surfaces extracted from CT scans.

Data augmentation methods [6] are widely recognized as a standard approach to address data scarcity. These techniques expand the training set by generating additional data, typically through either data warping or oversampling. However, traditional augmentation methods used in image classification [7] cannot be directly applied to CFD data, as they might violate physical laws and boundary conditions. Our method integrates computational geometry tools such as functional correspondence [8,9], to transfer deformation functions defined by ENT experts to healthy anatomical surfaces, simulating various pathological conditions with different severity levels and locations. This approach allows us to produce a large dataset with minimal expert supervision while preserving anatomical realism and ensuring well-defined pathology labels.

We apply our method to train a binary classifier to identify 2 nasal pathologies: turbinate hypertrophy and septal deviation. The training set is entirely generated with our augmentation method, starting from the CT scans of 7 healthy individuals. Overall, 315 surfaces are produced using deformation functions corresponding to different severity levels and locations of pathologies. We test a classifier exclusively trained on our synthetic samples both on synthetic samples and on a set of 10 CFD simulations from real patients, and we show that the classifier can successfully generalize and correctly identify pathologies.

2 Related Work

The application of machine learning (ML) in fluid mechanics has grown significantly over the past decade, with a surge of interest in integrating ML with Computational Fluid Dynamics (CFD) [1,10]. Traditionally, ML models are used to approximate or accelerate fluid simulations by modeling complex non-linear

relationships in fluid dynamics, often bypassing traditional numerical solvers. A primary example is the use of physics-informed neural networks to replace the Navier-Stokes equations with ML models, significantly speeding up simulations [4,11]. Another research direction involves enhancing turbulence models through ML techniques. Ling et al. [12] applied deep learning to improve the Spalart-Allmaras turbulence model [13], while Fukami et al. [3] used convolutional neural networks (CNNs) for super-resolution tasks, reconstructing high-fidelity flow fields from lower-resolution simulations. These works primarily focus on improving the accuracy and efficiency of CFD simulations by predicting traditional fluid dynamics quantities. In contrast, our approach aims to infer high-level labels, such as medical diagnoses, from CFD flow fields, a less explored area in ML-CFD research. The work by Schillaci et al. [5] is one of the few attempts to combine ML with CFD for classification tasks. However, their study was limited to simplified geometries and toy examples, whereas our method targets real-world clinical scenarios by generating realistic, labeled datasets through data augmentation.

Data augmentation in CFD poses unique challenges due to the strict physical laws and boundary conditions governing flow fields. Traditional augmentation methods, such as geometric transformations, often fail to produce physically consistent results. Wu et al. [14] addressed this issue using Generative Adversarial Networks (GANs) to create synthetic flow fields, while Abucide-Armas et al. [15] leveraged the similarity principle of fluid dynamics to maintain physically consistent flow behavior across different scales by matching the Reynolds number. These approaches, however, primarily focus on augmenting the flow field in simplified scenarios, which limits their applicability to complex medical contexts where the effects of anatomical pathologies on internal flows must be accurately captured. Our method differs by applying data augmentation directly to the anatomical geometries rather than the flow fields. By deforming healthy anatomical surfaces according to expert-defined pathology models, we generate a large and diverse training set with clear pathology labels. This strategy ensures that the augmented data remains physically realistic and allows the ML model to learn meaningful associations between anatomical deformations and their impact on fluid dynamics, improving the classification of nasal pathologies in clinical settings.

3 Problem Formulation

We address the problem of training a classifier K capable of identifying nasal pathologies by analyzing flow fields obtained from CFD simulations. The classifier K operates as follows:

$$K : F \to Y, \quad Y = \{y_j\}_{j=1,\dots,M}, \tag{1}$$

where F represents the space of CFD-simulated flow fields, and Y is a set of predefined pathology labels, with $M = 2$ corresponding to the labels for *turbinate hypertrophy* and *septal deviation*.

Each CFD simulation F_i is computed from a surface $S_i \subset \mathbb{R}^3$ extracted from a CT scan, and consists of a matrix stacking in each column the 3D coordinates and the fluid dynamics quantities of a cell:

$$F_i[:, j] = [x_j, y_j, z_j, p_j, u_j, v_j, w_j]^T \in \mathbb{R}^7, \quad j = 1, \ldots, n_i, \tag{2}$$

where (x_j, y_j, z_j) are the spatial coordinates, (u_j, v_j, w_j) are the components of the velocity vector, p_j is the pressure, and n_i is the number of cells of the simulation domain Ω. In our application, n_i can easily range in the order of $n_i \approx 10^7$.

The training set for the model K consists of labeled pairs:

$$\{(F_i, y_i)\}_{i=1,\ldots,N}, \tag{3}$$

where $y_i \in Y$ is the ground truth label associated with each flow field F_i. Our goal is to train K to accurately predict y_i based on features extracted from F_i, enabling robust and automatic classification of nasal pathologies.

4 Method

Our goal is to develop a method that can efficiently generate a labeled and realistic training set of annotated CFD flow fields, minimizing the need for expert supervision of numerous anatomies, which is both costly and time-intensive. Our solution is illustrated in Figure 1.

First, in collaboration with ENT experts, we define a *reference* surface $R \subset \mathbb{R}^3$ representing a healthy nose, along with a set of M *surface deformation functions* $\{(\delta_j, y_j)\}_{j=1,\ldots,M}$, as shown in the top left corner of Figure 1. Each deformation δ_j modifies R to create a variant R^*_j, which is affected by the pathology y_j. Deformation functions are defined by ENT experts, who manually modify the healthy nose R to obtain a surface that is compliant with the nasal cavities of a patient affected by the pathology y_j, as in an inverse *virtual surgery* procedure. We then use R and $\{(\delta_j, y_j)\}_{j=1,\ldots,M}$, together with a computational geometry technique, to generate a large training set of annotated anatomies starting from a small set $\{\tilde{S}_i\}_{i=1,\ldots,N}$ of N surfaces extracted from the CT scans of *healthy* individuals. In particular, the key idea of our method is that each surface from a healthy patient can be considered a deformed variant of R, thus we adopt functional maps [16], an advanced shape registration technique, to establish a point-to-point mapping between R and each $\tilde{S}_i$. This mapping is used to automatically apply deformation δ_j to S_i (the cleaned version of $\{\tilde{S}_i\}$, manually removed of some spurious geometry), resulting in a new surface $S^*_{i,j}$, which is a variant of S_i affected by the pathology y_j. By repeating this procedure for all the surfaces of healthy patients and all the pathologies, we obtain a set of surfaces $\{S^*_{i,j}\}_{i=1,\ldots,N;j=1,\ldots,M}$ with their corresponding labels. We use these surfaces to run $N \times M$ CFD simulations $\{F_{i,j}\}_{i=1,\ldots,N;j=1,\ldots,M}$ which form the training set of our classifier.

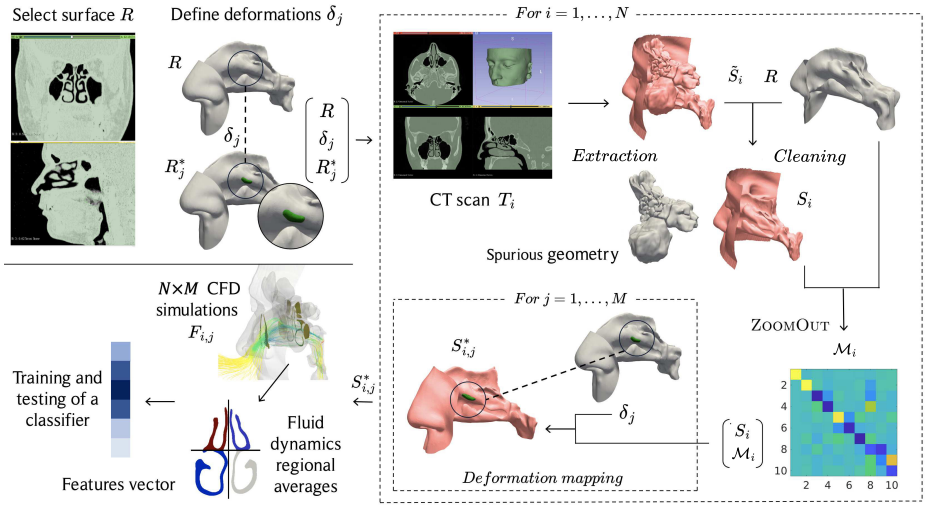

Fig. 1: A schematic representation of our method. The top-left corner illustrates the selection of the reference nose R and the definition of deformation functions $\{\delta_j\}_{j=1,\ldots,M}$, which transform R into pathological variants R_j^*. On the right, we show the generation of synthetic pathological surfaces $S_{i,j}^*$. Surfaces $\tilde{S}_i$ are extracted from the N healthy CT scans T_i by applying a threshold. Experts manually remove spurious parts (e.g., paranasal sinuses) from $\tilde{S}_i$, producing the cleaned surface S_i. The bottom-right corner displays the mapping $\mathcal{M}_i$, enabling the transfer of δ_j from R to S_i, resulting in $S_{i,j}^*$. Finally, CFD simulations on $S_{i,j}^*$ are used to train and test the classifier, as shown in the bottom-left corner.

Our method allows us to automatically produce a large set of surfaces $\{S_{i,j}^*\}$, each associated with a specific pathology. These surfaces are then used to perform CFD simulations, generating a wide and labeled training set for ML models. The anatomical variability of the healthy subjects, combined with the flexibility of applying different deformation intensities and combinations, significantly increases the size and diversity of the dataset.The effectiveness of our method is evaluated by training a classifier on the synthetic CFD data and testing it on real patient anatomies.

4.1 Reference Surface Selection and Deformation Functions

A critical component of our augmentation method is the selection of a high-quality *reference surface* $R \subset \mathbb{R}^3$ representing a healthy nasal anatomy. ENT experts from *ASST Santi Paolo e Carlo* in Milan reviewed a CT scan database to identify a well-defined, healthy scan with high resolution. The surface of R, as well as the surfaces $\{\tilde{S}_i\}$, are extracted from CT scans $\{T_i\}$ using a segmentation process that isolates the internal airways [17]. The segmentation process can be

(a) Paranasal sinuses highlighted in green.

(b) Surface without paranasal sinuses.

Fig. 2: The reference surface before and after the removal of paranasal sinuses, optimizing it for CFD simulations.

challenging, particularly in small regions where nasal mucus may accumulate and increase the local density, leading to irregularities and disconnections in the surface. These small errors need to be evaluated and corrected by experts. The extracted surfaces are further refined by manually removing unnecessary regions [18], namely the paranasal sinuses, which contribute minimally to airflow dynamics and unnecessarily increase the computational cost of CFD simulations (Figure 2). Despite these manual interventions being very time-consuming, it is important to stress that this is limited to just 7 healthy patients out of the 315 synthetic samples produced by our augmentation technique. The remaining synthetic pathological samples are automatically generated from these cleaned surfaces, avoiding further segmentation and manual refinement.

To generate synthetic pathological surfaces, we define a set of *surface deformation functions* $\{(\delta_j, y_j)\}_{j=1,\dots,M}$ over the reference surface R. Each deformation function $\delta_j : \mathbb{R}^3 \to \mathbb{R}^3$ is manually designed by ENT surgeons to simulate specific nasal pathologies, such as *septal deviation* and *turbinate hypertrophy*. A *septal deviation* involves bending the nasal septum, potentially obstructing airflow, while *turbinate hypertrophy* results in enlarged tissue along the nasal walls, leading to nasal obstruction. Figure 3 illustrates how these deformations can impact the reference geometry R. The deformation functions map each point (x, y, z) of the reference surface to a displacement vector $(\Delta x, \Delta y, \Delta z)$, allowing us to transform the healthy geometry R into a pathological variant R_j^*. By varying the location and severity of these deformations, we generate multiple pathological surfaces with minimal expert intervention. This approach not only provides a robust and diverse dataset but also ensures that each generated surface has a clear and well-defined pathology label. The ability to combine and control the intensity of these deformations further enhances the diversity of the training set, contributing to better generalization of the classification model.

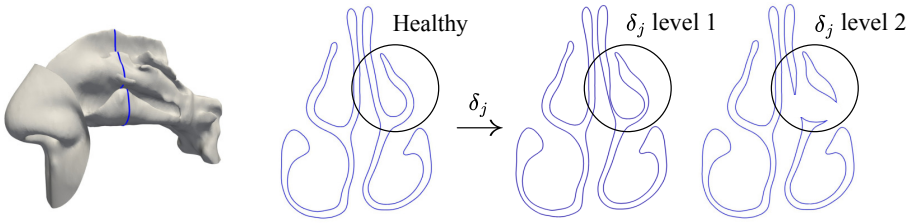

Fig. 3: Left: section of the reference surface used to visualize the pathology. Center: healthy section. Right: simulated turbinate hypertrophy with two levels of severity as defined by deformation functions.

4.2 Mapping Deformations from R to S_i

Our data augmentation method transfers deformation functions $\{\delta_j\}_{j=1,...,M}$, defined by ENT surgeons on the reference surface R, to healthy anatomical surfaces S_i of different patients. Each deformation function δ_j simulates a specific pathology y_j, transforming R into a pathological variant R_j^*. Instead of redefining deformations for every surface, we use *functional mapping* to automatically apply δ_j to any healthy cleaned surface S_i, significantly reducing the need for manual intervention. The process involves estimating a point-to-point correspondence $\mathcal{M}_i : \mathbb{R}^3 \to \mathbb{R}^3$ between R and S_i such that $S_i = \mathcal{M}_i(R)$, that allows us to map each deformation function directly to the target surface. This mapping is achieved using ZoomOut [19], which aims to preserve the anatomical consistency of the deformations despite geometric differences between R and S_i. The deformed surfaces $S_{i,j}^*$ are computed as:

$$S_{i,j}^* = S_i + \delta_j(\mathcal{M}_i^{-1}(S_i)), \quad \forall i = 1, ..., N;\ j = 1, ..., M. \tag{4}$$

This approach generates a set of $N \times M$ annotated surfaces for CFD simulations. Additionally, we can control the severity of pathologies by scaling deformations or combining multiple functions as follows, further enhancing the diversity of the training set:

$$S_{i,(j_1,j_2)}^* = S_i + \alpha\delta_{j_1}(\mathcal{M}_i^{-1}(S_i)) + \beta\delta_{j_2}(\mathcal{M}_i^{-1}(S_i)). \tag{5}$$

This flexibility in generating synthetic data contributes to robust model training, enabling accurate classification of nasal pathologies.

4.3 CFD Simulations

We run numerical simulations to obtain the flow fields inside $\tilde{S}_{i,j}$ using OpenFOAM, an open-source CFD toolbox. The simulations employ the *Large Eddy*

Simulation (LES) technique under the assumption of incompressible flow, capturing the complex turbulent structures within the nasal cavities. The computational domain is defined by enclosing the nostrils with a spherical surface. This is a standard practice in CFD to simulate the boundary condition of an ideally infinite external ambient. The simulations replicate a steady-state inspiration at a flow rate of 16 l/min, ensuring consistent physiological conditions. Approximately 12 million mesh cells are used per simulation, leveraging parallel computing with 96 cores on the Galileo supercomputing system. We maintain the same boundary conditions across all synthetic and real samples to ensure consistency and comparability of the results among simulations. The simulations produce detailed flow fields containing spatial coordinates, velocity components, and pressure values, which are subsequently processed for feature extraction.

5 Experiments

To demonstrate the effectiveness of our approach, we train a classifier using CFD flow fields $\{F_{i,j}\}$ computed from the set of surfaces $\{S^*_{i,j}\}$. In light of the study conducted in [5] on simplified nose shapes, we extract six transversal sections (right and left sides) from each surface $S^*_{i,j}$. Regional averages of the velocity $\boldsymbol{U}$ are then computed to build the feature vector fed to the classifier. The classification performance is assessed in two settings: *i*) we evaluate the classifier on synthetic samples using a Leave One Patient Out Cross Validation (LOPOCV); *ii*) we test the model on CFD data from a set $\{D_i\}_{i=1,\dots,10}$ of real, *pathological*, and *never-seen-before* patients. Each patient D_i is extracted from a CT scan and undergoes the same surface extraction and cleaning (Section 4.1), and CFD simulation procedures as healthy samples. We aim to demonstrate that a classifier trained solely on manually deformed data can effectively identify pathologies in real and anatomically diverse cases.

5.1 Training Set Generation

We generate the training set by applying our method to the CT scans of 7 healthy individuals and then perform CFD simulations for each surface $S^*_{i,j}$. These CT scans are provided by *ASST Santi Paolo e Carlo* and have been diagnosed as *healthy* by ENT experts. Before processing, all scans were verified for high quality and resolution to ensure accurate surface extraction and CFD simulations.

For both septal deviation and turbinate hypertrophy, ENT experts manually defined deformation functions $\{(\delta_j, y_j)\}$ at different locations and with two levels of severity, as described in Section 4.1. This process yielded six variations of septal deviation and nine of turbinate hypertrophy. Additionally, combining deformations of the same pathology at different locations and severities (as shown in 5) resulted in 12 combinations of septal deviations and 18 combinations of turbinate hypertrophy. The initial set of 7 individuals was thus expanded to a total of 315 pathological samples $S^*_{i,j}$ with clearly defined labels. CFD simulations

Table 1: Summary of the datasets used in our experiments, distinguishing between hypertrophies (Hyp) and septal deviations (Dev), as well as the severity level of pathologies.

			Dataset Name			
			$S1$	$S2$	S	D
Samples	Hyp	Lev 1	35	-	189	5
		Lev 2	-	28		
	Dev	Lev 1	21	-	126	5
		Lev 2	-	21		
	Tot	Lev 1	56	-	315	10
		Lev 2	-	49		

were then performed as outlined in Section 4.3. Healthy samples were excluded since the focus of this study is on classifying pathological cases.

To clarify the structure of our dataset and simplify the notation, we denote as $S1$ and $S2$ the subsets of $S^*_{i,j}$ containing surfaces affected by a single pathology of severity level 1 and 2, respectively, and we name S the complete set that extends $S1$ and $S2$ with combined deformations. We also denote D as the set $\{D_i\}_{i=1,\dots,10}$ of real, pathological, and never-seen-before patients. These sets are used to validate the classifier performance on real anatomical data. Table 1 summarizes all datasets used in our experiments, categorizing samples by pathology type and severity level.

5.2 Feature Extraction

Due to its large size, CFD flow fields cannot be directly fed to a NN. Therefore, we perform feature extraction to reduce the dimensionality of the input. Following the approach of [5], we compute *regional averages* of the velocity magnitude $|\boldsymbol{U}|$ over six transversal sections of each surface $S^*_{i,j}$, including the olfactory region boundaries and four equally spaced sections along the nasal cavity (see Figure 4).

To account for varying mesh densities and to ensure a consistent representation of flow characteristics, regional averages are computed as area-weighted means:

$$\overline{|\boldsymbol{U}|}_n = \frac{\sum_i |\boldsymbol{U}|_i A_i}{\sum_i A_i}, \tag{6}$$

where A_i is the area of cell i, and the index i spans all cells in region $\mathcal{R}_n$. This approach ensures that larger cells, which cover more significant portions of the anatomical region, contribute proportionally to the final feature values.

Regional averages are calculated separately for the left and right halves of each section, therefore, each CFD simulation F_i is reduced to a vector of 12 regional averages, providing a compact and informative input for the NN.

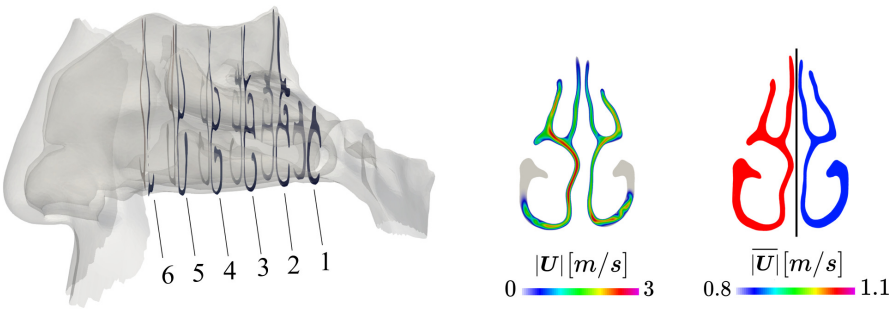

Fig. 4: Sections used for computing feature vectors. The left-hand side shows the positions of the 6 sections we considered. On the right-hand side, the module of the velocity is averaged on the left and right half-sections.

5.3 Model Training

We train a model to classify feature vectors composed of regional averages of fluid dynamics quantities. The classifier K takes as input a vector of 12 features, namely, regional averages of $|\boldsymbol{U}|$ described in Section 5.2.

The classifier is implemented as a multilayer perceptron (MLP) with a standard architecture comprising four hidden layers of 60, 40, 20, and 10 neurons, respectively, totaling 4261 trainable parameters. The output layer consists of a neuron with sigmoid activation where 0 indicates hypertrophy and 1 indicates septal deviation. Other neurons use the ReLU activation function.

Each input feature is standardized to have zero mean and unit variance on the training set. The model is trained using binary cross-entropy as the loss function for up to 1000 epochs, with early stopping based on validation loss. During training, 85% of the samples are used for training and 15% for validation.

5.4 Results

We evaluate our classifier using two experimental settings: *(i)* Leave One Patient Out Cross-Validation (LOPOCV) on synthetic data generated through our augmentation method, and *(ii)* testing on a set D of 10 real pathological patients, with expert-confirmed diagnoses of either hypertrophy (5 patients) or septal deviation (5 patients).

Leave One Patient Out Cross Validation (LOPOCV) The LOPOCV approach involves iteratively excluding all synthetic samples derived from a single healthy individual during training and using them as the test set, eventually, we average the inference performance over all the patients. We rely on the LOPOCV to evaluate the model's ability to generalize to unseen synthetic samples produced by our augmentation method. As shown in Figure 5, the classifier K achieves 94% accuracy on hypertrophies and 84% on septal deviations,

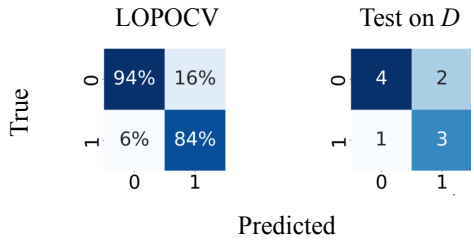

Fig. 5: Confusion matrices for the classifier K. Left: LOPOCV on synthetic data. Right: Test on real patient data. Labels 0 and 1 correspond to hypertrophy and septal deviation, respectively.

Pathology	*Hypertrophy*			*Septal Deviation*		
Metric	Precision	Recall	F_1-score	Precision	Recall	F_1-score
Score	0.899	0.941	0.919	0.906	0.841	0.872

Table 2: Precision, Recall and F_1-score in the LOPOCV for both classes: hypertrophy and septal deviation.

resulting in an overall accuracy of 89%. Accurate classification performance of the LOPOCV, denotes that the classifier trained on synthetic data successfully learned how deformations δ_j impact CFD data, and that K can properly generalize when the anatomy changes. The confusion matrix indicates that most classification errors occur when predicting septal deviations, suggesting that this pathology may not always produce distinct flow differences. The higher accuracy in identifying hypertrophies (94%) aligns with clinical expectations, as these conditions often cause significant anatomical obstructions and distinct airflow patterns, making them easier to classify. The solid performance of $|\boldsymbol{U}|$ highlights that velocity magnitude is an effective indicator of pathological changes in the nasal airflow.

Table 2 reports the mean values of *precision*, *recall*, and F_1*-score* for both hypertrophy and septal deviation computed during the LOPOCV procedure. Given a binary classification setting, and denoting by:

$$\text{TP} = \text{true positives}, \quad \text{FP} = \text{false positives}, \quad \text{FN} = \text{false negatives},$$

these metrics are computed as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

The results indicate that hypertrophies (class 0) are associated with a higher recall and F_1-score, suggesting that the model is generally able to detect most true cases of this condition, likely due to the more distinct effects they produce on the flow field. Conversely, septal deviations (class 1) exhibit a slightly higher

precision, meaning that when the model predicts a septal deviation, it is more likely to be correct. However, the lower recall for class 1 implies that some true cases are missed, possibly due to the more subtle and variable effects of septal deviation on the airflow field. Overall, these results suggest that while the model is more sensitive to hypertrophies, it achieves higher classification precision for septal deviations.

Test on Real Patients The test on real patients is conducted using a NN having the same architecture as in the previous experiment, but trained on all the synthetic data in S and tested only on data from a set D of real pathological individuals. When applied to D, the classifier trained exclusively on synthetic data correctly classifies 7 out of 10 cases, showing good generalization capabilities. The confusion matrix (Figure 5) reveals consistent results with the LOPOCV, showing that hypertrophies are generally easier to detect, with 4 out of 5 correct classifications, while septal deviations exhibit more variability in prediction accuracy. This behavior underscores the clinical complexity of septal deviations, where mild cases might not induce significant changes in the airflow, leading to potential misclassification. The performance on real pathological patients shows the real potential of our augmentation method: given a large anatomical variability, and considering that we apply our augmentation method to a set of only 7 healthy patients, our classifier performs well on real patients, underscoring the method's ability to generalize to real cases even when trained exclusively on synthetic data.

Overall, these results validate our augmentation method's effectiveness in generating synthetic samples that not only enhance model performance in synthetic scenarios but also ensure reliable generalization to real clinical cases. This demonstrates the potential of combining CFD with ML to support ENT specialists in diagnosing nasal pathologies more accurately.

6 Discussion and Conclusions

In previous work, Schillaci et al. [5] demonstrated the ability of neural networks to classify CFD flow fields, but their tests were restricted to simplified scenarios due to the lack of training data. Our work extends this approach to a more realistic clinical setting, focusing on the classification of nasal pathologies in human upper airways using only CFD data. We present a data augmentation method that generates a large and diverse set of labeled CFD data by applying expert-defined deformations to a limited number of healthy CT scans. This approach significantly reduces the need for extensive manual labeling while maintaining anatomical realism, a critical advantage in medical applications where annotated datasets are often scarce.

Our experiments focused on two key evaluations: *i*) Leave One Patient Out Cross-Validation (LOPOCV) on synthetic data, and *ii*) testing on a set of 10 real patients with confirmed diagnoses of either hypertrophy or septal deviation. Our classifier trained exclusively on synthetic data achieved strong performance,

with up to 89% accuracy in LOPOCV and correctly identifying pathologies in 7/10 real pathological patient cases. The classification performance demonstrates that a model trained on augmented synthetic data can effectively generalize to real clinical scenarios, despite never encountering real pathological samples during training. The consistency of the results across both synthetic and real datasets confirms the effectiveness of our augmentation method in generating high-quality training data. This approach holds the potential to support ENT specialists in diagnosing nasal pathologies through non-invasive CFD analysis, enhancing diagnostic accuracy and treatment planning.

It is important to stress that creating the real patient dataset D required significantly more effort compared to the synthetic dataset of 315 samples, as the former involved the manual interventions described in Section 4.1. This highlights the efficiency of our proposed augmentation method, which automates the generation of synthetic pathological samples, drastically reducing manual intervention while preserving data quality and label reliability.

In conclusion, our method establishes a practical and effective connection between CFD data and real clinical scenarios, offering a practical solution for training NNs when labeled data is limited. Future work will focus on expanding the dataset by including a broader range of pathologies to improve the generalization of the model. We also plan to design new features that better capture the fluid dynamic behaviors associated with specific clinical conditions, enhancing the model's predictive power. Additionally, we will explore advanced interpretability techniques to provide deeper insights into which features contribute most significantly to the classification of pathologies, ultimately supporting clinical decision-making with more transparent and explainable models.

Acknowledgements The Authors are grateful to A. Bulfamante for the time spent as ENT expert. Computing time has been provided by the Italian CINECA HPC Center in Bologna under the ISCRA-C grant AICFD-23. This publication is part of the project PNRR-NGEU which has received funding from MUR-DM 351/2022. MQ acknowledges support from the PRIN 2022 project OpenNOSE, cod. 2022BYA5AF, CUP D53D2300343006.

References

1. Steven Brunton, Bernd Noack, and Petros Koumoutsakos. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.
2. Karthik Duraisamy, Gianluca Iaccarino, and Heng Xiao. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, 51(1):357–377, 2019.
3. Kai Fukami, Koji Fukagata, and Kunihiko Taira. Assessment of supervised machine learning methods for fluid flows. *Theoretical and Computational Fluid Dynamics*, 34(4):497–519, 2020.
4. Maziar Raissi, Alireza Yazdani, and George Em Karniadakis. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481):1026–1030, 2020.

5. 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.
6. Connor Shorten and Taghi M. Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):60, 2019.
7. Suorong Yang, Weikang Xiao, Mengcheng Zhang, Suhan Guo, Jian Zhao, and Furao Shen. Image data augmentation for deep learning: A survey, 2022.
8. Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics*, 31(4):30,1–30,11, 2012.
9. Maks Ovsjanikov, Etienne Corman, Michael Bronstein, Emanuele Rodolà, Mirela Ben-Chen, Leonidas Guibas, Frederic Chazal, and Alex Bronstein. Computing and processing correspondences with functional maps. In *ACM SIGGRAPH 2017 Courses*, pages 1–62, 2017.
10. M. P. Brenner, J. D. Eldredge, and J. B. Freund. Perspective on machine learning for advancing fluid mechanics. *Physical Review Fluids*, 4(10):100501, 2019.
11. Pin Wu, Kaikai Pan, Lulu Ji, Siquan Gong, Weibing Feng, Wenyan Yuan, and Christopher Pain. Navier–stokes generative adversarial network: a physics-informed deep learning model for fluid flow generation. *Neural Computing and Applications*, 34(14):11539–11552, 2022.
12. Julia Ling, Andrew Kurzawski, and Jeremy Templeton. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166, 2016.
13. P. Spalart and S. Allmaras. A one-equation turbulence model for aerodynamic flows. In *30th Aerospace Sciences Meeting and Exhibit*, pages 439–445. American Institute of Aeronautics and Astronautics, Reno, NV, 1992.
14. Haizhou Wu, Xuejun Liu, Wei An, and Hongqiang Lyu. A generative deep learning framework for airfoil flow field prediction with sparse data. *Chinese Journal of Aeronautics*, 35(1):470–484, 2022.
15. Alvaro Abucide-Armas, Koldo Portal-Porras, Unai Fernandez-Gamiz, Ekaitz Zulueta, and Adrian Teso-Fz-Betoño. A data augmentation-based technique for deep learning applied to CFD simulations. *Mathematics*, 9(16):1843, 2021.
16. Simone Melzi, Jing Ren, Emanuele Rodolà, Abhishek Sharma, Peter Wonka, and Maks Ovsjanikov. Zoomout: spectral upsampling for efficient shape correspondence. *ACM Transactions on Graphics*, 38(6):155,1–155,14, 2019.
17. Maurizio Quadrio, Carlotta Pipolo, Stefano Corti, Francesco Messina, Chiara Pesci, Alberto M. Saibene, Samuele Zampini, and Giovanni Felisati. Effects of ct resolution and radiodensity threshold on the cfd evaluation of nasal airflow. *Medical & Biological Engineering & Computing*, 54(2):411–419, 2016.
18. Hanhui Jin, Rongrui Fan, M.J. Zeng, and K.F. Cen. Large eddy simulation of inhaled particle deposition within the human upper respiratory tract. *Journal of Aerosol Science*, 38:257–268, 2007.
19. Luca Cosmo, Emanuele Rodolà, Jonathan Masci, Andrea Torsello, and Michael M. Bronstein. Matching deformable objects in clutter. In *Fourth International Conference on 3D Vision (3DV)*, pages 1–10, 2016.