



POLITECNICO DI MILANO
Dipartimento di Scienze e Tecnologie Aerospaziali
Doctoral Programme in Aerospace Engineering

Analysis of complex systems
with CFD and Machine Learning

Doctoral Dissertation of:
Andrea Schillaci

Supervisors:
Prof. Maurizio Quadrio
Prof. Giacomo Boracchi

Academic year 2022/2023 - Cycle XXXIV

Моей любимой Юлии

*"Beware of long explanations.
They are the most dangerous enemies of spontaneity."*

Mikhail Bulgakov. The Master and Margarita.

Abstract

Computational Fluid Dynamics (CFD) provides detailed quantitative information in a wide range of applied fluid mechanics problems. However, sometimes CFD alone cannot predict high-level functional properties that do not descend directly from the equations of fluid motion.

One such instance is the diagnosis of nasal breathing difficulties (NBD). While NBD affect a large fraction of the world population, to date the CFD-computed airflow in the nose is not enough to provide a satisfactory diagnosis. This Thesis introduces a data-driven framework for extracting a medical diagnostic output from a CFD solution. A data-driven approach requires a suitable training set. A robust and reliable CFD procedure to compute the flow solution must be set up; meaningful features must be extracted from the solution and fed to a machine-learning model. Large samples of annotated data are not available to train the model; in this work, a novel procedure is described to generate a large and consistent dataset containing arbitrary combinations of a set of well-defined pathologies.

By using few informative features extracted from the CFD solution, a neural network is shown to be able to successfully classify nasal pathologies. Results are encouraging, and support the intuition that using CFD as a powerful feature extraction tool makes the ML problem more tractable, thanks to the non-linear Navier–Stokes equations, which act as a filter on the geometry.

Keyword: Computational Fluid Dynamics, Machine Learning, flow in the human nose

Preface

This thesis deals with the application of Computational Fluid Dynamics and Machine Learning to nasal breathing difficulties. A first Part (Chapter 1) is an overview that introduces the application, the key concepts, the methodology and the main results. A second Part reproduces seven published articles and manuscripts. The papers are adjusted to comply with a uniform thesis format for consistency, while their content remains true to the original counterparts.

Chapter 2. A. Schillaci & M. Quadrio, 2022. *Importance of the numerical schemes in the CFD of the human nose.* Journal of Biomechanics **138**, 111100.

Chapter 3. M. Macellari, A. Schillaci, U. Tanzini, M. Trimarchi & M. Quadrio. *An adjoint-based approach for the surgical correction of nasal septal deviations.* Manuscript.

Chapter 4. A. Schillaci, K. Hasegawa, C. Pipolo, G. Boracchi & M. Quadrio. *Comparing flow-based and anatomy-based features in the data-driven study of nasal pathologies.* Flow. In revision.

Chapter 5. A. Schillaci, M. Quadrio, C. Pipolo, M. Restelli & G. Boracchi. 2021. *Infering Functional Properties from Fluid Dynamics Features.* 2020 25th International Conference on Pattern Recognition (ICPR), pp. 4091-4098.

Chapter 6. R. Margheritti, A. Schillaci, M. Quadrio & G. Boracchi. *Data Augmentation for Training Neural Networks to Classify Flow Fields: a Medical case study.* Manuscript.

Chapter 7. A. Schillaci, C. Pipolo, G. Boracchi & M. Quadrio. *Enhancing Machine Learning with Computational Fluid Dynamics.* Manuscript.

Chapter 8. C. Pipolo, A. M. Bulfamante, A. Schillaci, J. Banchetti, L. Castellani, A. M. Saibene, G. Felisati & M. Quadrio, 2021. *Through the back door: expiratory accumulation of SARS-CoV-2 in the olfactory mucosa as mechanism for CNS penetration.* International Journal of Medical Sciences, **18(10)**, 2102.

December 2023, Milan
Andrea Schillaci

Other publications:

The following paper is not included in the thesis:

S. Wang, P. Rohlfs, M. Börnhorst, A. Schillaci, H. Marschall, O. Deutschmann, & M. Wörner. 2022. *Bubble Cutting by Cylinder–Elimination of Wettability Effects by a Separating Liquid Film*. Chemie Ingenieur Technik, **94(3)**, 385-392.

Conferences:

A Schillaci, M. Quadrio, C. Pipolo, M. Restelli & G. Boracchi. *Inferring Functional Properties from Fluid Dynamics Features*. 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021.

A. Schillaci, G. Boracchi, C. Pipolo, & M. Quadrio. *A CFD-augmented machine-learning approach for the classification of nasal pathologies*. In 14th European Fluid Mechanics Conference (EFMC14), Athens, Greece, 2022.

M. Quadrio, A. Schillaci & G.Boracchi. *Classification of nasal pathologies: enhancing Machine Learning with Computational Fluid Dynamics*. The 18th OpenFOAM Workshop (OFW18), Genoa, Italy, 2023.

Contents

Abstract	I
1 Overview	1
1.1 The anatomy of the nose	4
1.2 Defining labels: Y	5
1.2.1 Parametric dataset D1	6
1.2.2 Realistic dataset D2	7
1.3 CFD setup and flow results: X	11
1.3.1 The flow main features	13
1.4 Feature extraction procedure: f	15
1.4.1 Expert-driven features	16
1.4.2 Mapped fields	16
1.5 The classifier \mathcal{K}	18
1.6 Results of the data-driven approach	18
1.6.1 Geometric vs Flow features	19
1.6.2 Classification of nasal pathologies	19
1.7 Conclusions	23
2 Importance of the numerical schemes in the CFD of the human nose	25
2.1 Abstract	25
2.2 Introduction	25
2.3 Methods	27
2.4 Results	29
2.4.1 First- vs second-order schemes	31
2.4.2 RANS vs LES	33
2.5 Discussion	33
2.6 Conclusion	36
2.7 Appendix A: The numerical approach	37
2.7.1 Boundary conditions	37
2.7.2 RANS model and procedures	37
2.7.3 LES model and procedures	38
2.8 Appendix B: Geometry and mesh quality	39
2.9 Appendix C: CT vs TrCT	41

3 An adjoint-based approach for the surgical correction of nasal septal deviations	43
3.1 Abstract	43
3.2 Introduction	43
3.3 Adjoint-based optimization	45
3.4 Methods	47
3.4.1 Anatomies and discretization	47
3.4.2 Direct RANS simulations	49
3.4.3 Adjoint solution and surface sensitivity	50
3.5 Results	51
3.5.1 RANS	52
3.5.2 Adjoint field	53
3.5.3 Surface sensitivity	54
3.6 Discussion	57
3.6.1 Validation by virtual surgery	59
3.6.2 The choice of the cost function	60
3.7 Conclusions	60
4 Comparing flow-based and anatomy-based features in the data-driven study of nasal pathologies	62
4.1 Abstract	62
4.2 Impact Statement	63
4.3 Introduction	63
4.4 Methods	65
4.4.1 The anatomies	65
4.4.2 Simulations	68
4.4.3 Functional maps	69
4.4.4 The classifier	70
4.5 Experiments	72
4.5.1 Geometrical features	72
4.5.2 Flow features	74
4.5.3 Performance and discussion	76
4.6 Conclusions and outlook	78
5 Inferring Functional Properties from Fluid Dynamics Features	81
5.1 Abstract	81
5.2 Introduction	81
5.3 Related work	83
5.4 Problem Formulation	84
5.5 Proposed solution	85
5.5.1 Pre-processing	86
5.5.2 Feature Extraction	86
5.5.3 Model training	88
5.6 Experiments	88
5.6.1 Prediction of Geometrical Features of an Airfoil	89

5.6.2	Prediction of pathologies in a simplified human nose	93
5.7	Conclusions	94
6	Data Augmentation Based on Computational Geometry for Neural Network Training in Medical Flow Field Classification	96
6.1	Abstract	96
6.2	Introduction	97
6.3	Related Work	98
6.4	Problem Formulation	99
6.5	Methodology	100
6.5.1	Selection of the reference surface R	101
6.5.2	Definition of Deformation Functions	103
6.5.3	Extraction of the surface \tilde{S}_i from the CT scan	104
6.5.4	Cleaning \tilde{S}_i via registration with R	104
6.5.5	Mapping Deformations from R to S_i	107
6.5.6	CFD Simulations	108
6.6	Experiments	109
6.6.1	Training set generation	109
6.6.2	Features extraction	110
6.6.3	Model training	112
6.6.4	Results	112
6.7	Conclusions	114
7	Enhancing Machine Learning with Computational Fluid Dynamics	116
7.1	Abstract	116
7.2	Introduction	116
7.3	Methods	120
7.3.1	The geometry: discretizing \mathcal{V}	121
7.3.2	The pathologies: defining Y	125
7.3.3	The simulations: computing X	128
7.3.4	Feature extraction: computing f	129
7.3.5	Neural network: defining \mathcal{K}	131
7.3.6	Neural network: dataset partition and training	132
7.4	Results	132
7.4.1	Airfoils	133
7.4.2	Noses	134
7.5	Discussion	136
7.5.1	The regional average of the velocity magnitude	136
7.5.2	Other features based on regional averaging	140
7.5.3	Streamlines	142
7.5.4	Robustness of results	144
7.6	Conclusion	146

8 Through The Back Door: Expiratory Accumulation of SARS-CoV-2 in the Olfactory Mucosa as Mechanism for CNS Penetration	148
8.1 Abstract	148
8.2 Introduction	149
8.3 Materials and Methods	150
8.4 Results	150
8.5 Discussion	151

Chapter 1

Overview

Computational Fluid Dynamics (CFD) plays a crucial role in numerous applications, ranging from industry to health. In most applications, the final goal of the CFD analysis is well-defined and can range from finding the shape of an airfoil with the best lift-to-drag ratio in flow control to flood prediction in hydrology. However, a class of problems exists where the ultimate information, which is most relevant for the end-user, might not be directly provided by the CFD itself. Illustrative cases exist in the medical domain, where the doctors are interested in practical answers: is the patient healthy? Which surgery has to be performed, if any? Which, in technical terms, can be translated as: is there a flow quantity that can determine if there is a pathological condition? Can a fluid dynamic variable identify the position of a pathology? The answers to these questions should indeed be contained in the flow field; however, the explicit link between the flow field and the required information remains elusive. The complex interplay between fluid dynamics, different physiological functions, and anatomy often prevents us from clearly formulating a shape optimization problem, which would be the most intuitive solution. Without loss of generality, the recurring example adopted throughout this Thesis is the diagnosis of Nasal Breathing Difficulties (NBD). NBD are a common condition affecting the upper airways, with potential impacts on quality of sleep, work performance, and cognitive function. A range of different conditions exist with a widespread prevalence, e.g. allergic rhinitis has a prevalence of 18.7% in Europe (Canonica *et al.*, 2007) and chronic rhinosinusitis has a prevalence between 4.5% to 12% across U.S.A. and Europe (DeConde & Soler, 2016). NBD are often associated with a lifetime consumption of medical resources, thus unsurprisingly the overall healthcare burden of the chronic rhinosinusitis alone in the U.S.A. for the year 2014 was around \$22b. Furthermore, given a specific patient with NBD symptoms, there is a general lack of consensus on whether and how to perform a surgery. These challenges are reflected in the high failure rate of specific surgical manoeuvres, e.g. more than 50% of the patients that undergo a septoplasty reported that symptoms either remained or had worsened (Illum, 1997; Sundh & Sunnergren, 2015). This problem is perhaps ascribable to the lack of reliable diagnostic tools. Ear, Nose and Throat (ENT) doctors routinely perform rhinomanometry examinations, whose limit is to give only a global assessment of the flow patency by measuring the ratio between flow rate and pressure

drop (Schumacher, 2004). An alternative study is the acoustic rhinomanometry, which gives a detailed assessment of the geometry of the nasal cavities, however without evaluating the flow field (Clements & Gortds, 2005). In the last few decades, ENT doctors have started to rely more on the visual analysis of Computational Tomography (CT) scans and Magnetic Resonance Imaging. However, the link between shape and function is not clearly assessed yet (Doorly *et al.*, 2008a). This situation is complicated by the presence of large inter-subject variability unrelated to pathological conditions (Keustermans *et al.*, 2018). Since the general function of the human nose is primarily driven by fluid mechanics (air warming, humidification, and filtering), we reckon that a pipeline of CT scan reconstruction and CFD analysis presents itself as a powerful tool able to deeply improve the understanding between shape and function of the human nose. Not surprisingly, in recent years, CFD has been increasingly used as a support by ENT doctors in their diagnosis (Moreddu *et al.*, 2019; Tjahjono *et al.*, 2023), and to improve the understanding of the nasal flow physics (Calmet *et al.*, 2019; Farnoud *et al.*, 2020). However, results often do not generalize to the next patient, and the added information brought by the CFD solution is not easily readable by ENT doctors. Furthermore, the flow fields *per se* can hardly answer the fundamental questions: whether and where to perform the surgery. Therefore, we believe that pursuing a data-driven approach might have a large potential in this unexplored class of problems.

The use of ML approaches in fluid dynamics is not new, recent years have seen a surge in the use of Machine Learning (ML) techniques applied in the field of CFD (Brunton *et al.*, 2020; Vinuesa & Brunton, 2022). An important research field exists, within the framework of ML, to improve the CFD simulations, especially aimed at accelerating and replicating the Navier-Stokes (NS) solution. High-resolution flow fields can be reconstructed from coarse ones using super-resolution techniques (Fukami *et al.*, 2019; Xie *et al.*, 2018), demonstrating that ML can improve the quality of CFD simulations. Raissi *et al.* (2019) introduced a new method, the Physics-Informed Neural Network (PINN), that can be used to accelerate traditional solvers (Markidis, 2021) and for turbulence modelling (Eivazi *et al.*, 2022). Several authors tried to improve RANS modelling using ML tools (Duraisamy *et al.*, 2019; Ahmed *et al.*, 2021), e.g. Ling *et al.* (2016) embedded Galilean invariance in the prediction of the anisotropy tensor, Parish & Duraisamy (2017) used ML techniques to better calibrate existing turbulence models. LES modelling has also seen a steady increase in the use of ML techniques, examples include the modelling of subgrid-scale with neural networks (Vollant *et al.*, 2017), blind deconvolution of flow variables (Maulik & San, 2017) and estimation of unresolved subgrid-scale physics (Novati *et al.*, 2021). Reduced-order models (ROM) present themselves as an ideal tool to condensate flow information. Several techniques exist, such as proper orthogonal decomposition (Taira *et al.*, 2017), dynamic-mode decomposition (Schmid, 2010) and Variational autoencoders (Solera-Rico *et al.*, 2023). ROM techniques can be used to train ML models to predict the flow around unseen shapes (Hasegawa *et al.*, 2020), to predict turbulent shear flows (Srinivasan *et al.*, 2019), to separate physically relevant modes to better understand their physical meaning (Semeraro *et al.*, 2012) and for flow control (Semeraro *et al.*, 2011). Furthermore, ML techniques have also been used to identify governing equations (Brunton *et al.*, 2016). Although many advances have been made in recent years in the field of ML and fluid dynamics, the novelty of the present Thesis is the introduction of a new approach

that uses ML to perform inference tasks, in which information extracted from a rather standard CFD is used to improve the ability of a ML algorithm to extract a medical label. More specifically, we show that by leveraging the non-linear and convective nature of the NS equations, it is possible to amplify the consequences of the geometrical defects and make them more self-evident, aiding the performance of the inference task. We believe that this new line of work opens a new perspective of enhancing ML with CFD.

The Machine Learning problem

Using data-driven models on CFD data is particularly challenging. Due to their intrinsic cost, CFD simulations leads to hefty datasets, with few flow fields (few samples) of the order of Gigabytes of data. Such a large amount of information is not easily handled by Machine Learning (ML) models. Moreover, in the medical field, it is difficult to obtain accurate annotated data.

The output of a CFD simulation is a set of both scalar and vector quantities associated to a computational domain Ω . The CFD solution can be stacked into a matrix X of size $m \times n$, where n is the number of cells of the simulation and m the number of scalars that are the output of the simulation (typically $m = 7$, pressure, the three components of the velocity field, and their three spatial coordinates). Then, from the ML point of view, our goal is to train a model \mathcal{K} that predicts a target value Y associated to the matrix X :

$$\mathcal{K} : X \mapsto Y. \quad (1.1)$$

The target variable can be either categorical (as for a classifier that identifies the most suitable surgery for NBD), or ordinal (as for a regressor that estimates some geometric quantities associated to a pathology). To this purpose, to train a Neural Network (NN), it is necessary to provide a training set of l labelled pairs $\{(X_j, Y_j), j = 1, \dots, l\}$. However, this standard ML presents specific challenges when facing a CFD derived input, namely the large dimensionality of its input (large n) and the limited number of training samples (small l), due to the high computational cost of the simulations. Thus, we pursue the mainstream approach in ML, of reducing the size of the input array by extracting a subset f of features used for the robust training of the network. Therefore, equation 1.1 can be rewritten as:

$$\mathcal{K} : f \mapsto Y. \quad (1.2)$$

Finally, it is worth drawing a comparison between a data-driven and a more fluid dynamic oriented approach: defining a shape optimization problem (Dilgen *et al.*, 2018; Alexandersen & Andreasen, 2020). The shape optimization requires a single CFD input and a cost function to define a sensitivity map over a shape, which might indicate where to perform the surgery. By contrast, in the ML approach, we associate to each entry of the dataset X a label Y , to avoid manually defining the rule \mathcal{K} . Other data-driven methods exist, based on Reinforcement Learning (RL), in which the network interacts with the environment in a closed loop. At every chosen time step, the network is provided with a partial observation of the environment and, in response, can execute an action, whose quality is assessed by a reward system. Viquerat *et al.* (2021) use of

Deep RL to maximize the lift-to-drag ratio in an airfoil shape optimization problem. However, with each action of the agent, the flow solution is computed all over again.

In §3, the authors study three anatomies affected by complex septal deviations, and the adjoint optimization seems capable of highlighting the area of interest and improving the cost function after a cycle of virtual surgery. Each optimization cycle requires the computation of both the flow and adjoint fields. The main drawback of a shape optimization problem is the need to formulate a cost function. Given the several tasks performed by a human nose, it is not self-evident which this would be. This first approach uses the total dissipated power as the cost function since the resistance encountered by the nasal airflow clearly increases if obstructions are present; however, it does not take into account complex physiological functions like thermal exchange and humidity. Furthermore, this study gives no general conclusion on how to operate on the next patient.

Therefore, this Thesis explores the data-driven approach to the human nose pathology classification. We propose novel methodologies to build a realistic and consistent medical datasets (X and Y), extract meaningful features (f) and train a classifier (\mathcal{K}). The proposed pipelines are used for a specific medical problem but remain flexible since they are also tested on a much larger dataset of 2D airfoils, with a similar objective: identifying damaged airfoils. More information about the airfoil dataset is given in §5 and §7.

1.1 The anatomy of the nose

To establish the required nomenclature, the nasal cavities' anatomy is briefly illustrated in this section. For a full and detailed description, the reader is invited to consult papers and books devoted to the anatomy and physiology of the human nasal cavities (Jones, 2001; Wong *et al.*, 2021).

Figure 1.1, illustrates the position of the main anatomical structures that are referred along the manuscript: at the top, a typical result of a CT scan; at the bottom, the reconstructed geometry. The nasal cavities are the connecting element between the external environment and the lungs. The external part of the human nose hides the complexity of the inner structures, which serve several purposes, such as air filtering, heating, and humidification. Furthermore, the human nose contains receptors for the sense of smell and contributes greatly to the sense of taste. The nasal structure splits immediately at the two *nostrils* and keeps this left-right division in two cavities, or *fossae*, up to the *nasopharynx* where they rejoin in one channel. The wall dividing the two fossae is the *septum*, whose anterior part is made of soft cartilage, while the posterior part is made of bone. Each individual nasal cavity is occupied by three *turbinates*, namely the *inferior*, the *middle*, and the *superior* turbinates. The turbinates are bony structures that extend longitudinally and are covered by a thick layer of mucosa. The presence of the turbinates gives the characteristic hook shape, clearly visible in the coronal cross-section of the CT scan in figure 1.1. The turbinates' main function is to increase the surface area to improve thermal exchange. Around the main structure of the nose, there are a group of four air-filled spaces connected to the fossa, the *paranasal sinuses*, namely the *frontal*, *ethmoid*, *lateral*, and *sphenoid*. The function of

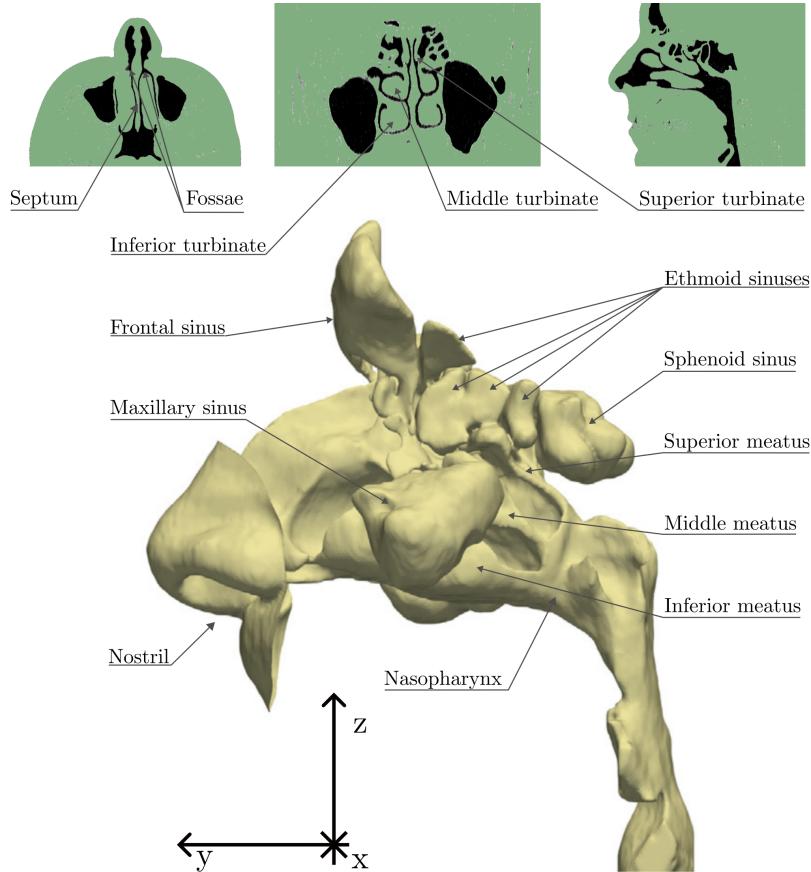


Figure 1.1: Top: CT scan of a patient, segmented for a constant value of HU (from left to right: transverse, coronal, and sagittal sections). Bottom: the reconstructed three-dimensional anatomy. The key anatomical features are indicated.

the paranasal sinuses is a topic of much debate; from a fluid dynamic point of view, the air flow through them is inconsequential, and they do not play a significant role in processing respiratory air (Cole, 1998).

1.2 Defining labels: Y

To build a dataset of l entries, it is required to produce a set of geometries with a clear label Y , which unambiguously defines a precise pathological condition. Furthermore, to better train the machine learning model, it is preferred to have several different anatomies with the exact same pathology. This last requirement is hard to meet due to the large variety of pathological conditions that might be concurrent, e.g., the same patient might have both a septal deviation and a swollen turbinate. In the following

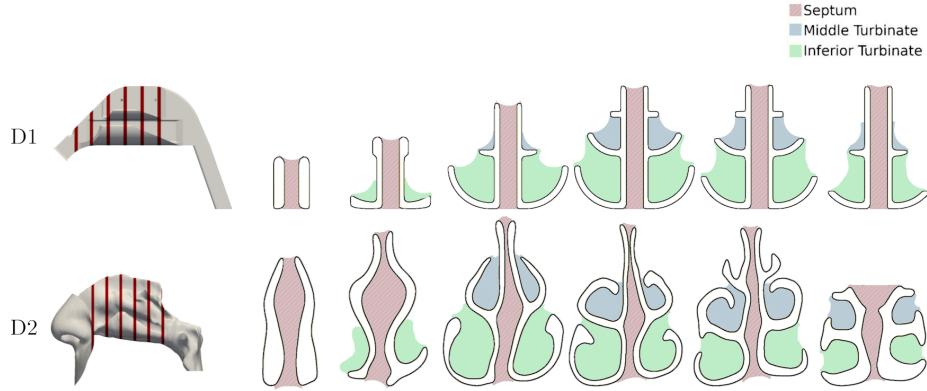


Figure 1.2: Top row: Synthetic nose anatomy D1, on the left the 3D model, highlighted in red the extracted cross-sections displayed on the right. Bottom row: realistic anatomy from dataset D2. The highlighted areas in the cross-section represent the main anatomical features, red the septum, green the inferior turbinate, blue the middle turbinate.

subsections, two different datasets are presented.

The first dataset (D1) is built to understand whether fluid dynamic features, available after a CFD analysis, are more effective than purely geometric features in the training of data-driven models. This dataset is built around an extremely simplified anatomic model, whose pathologies and anatomical variability are driven by parameters.

The second dataset (D2) addresses the automatic classification of pathologies in the human nose using flow features. The anatomies of a few healthy patients are modified with combinations of elementary pathologies; therefore, in this dataset, the anatomical variability between the patients is genuine, and the pathologies are realistic but controlled.

1.2.1 Parametric dataset D1

An intuitive way to control the dataset is to fully parameterize it, at the cost of losing some anatomical fidelity. In the simplified dataset D1, instead of using real anatomies, the geometries originate from a baseline CAD model of the human nasal cavities, which retains its main features but has the advantage of a full geometric parameterization. With the parametric approach, the control of the dataset is full, and no ambiguity ensues when assigning a label Y to the shape. The nasal literature is not new to the use of synthetic models (Naftali *et al.*, 1998; Liu *et al.*, 2009). However, we are first to build a fully parameterized model suitable for the generation of a controlled dataset. The baseline CAD model, built under the supervision of ENT surgeons, is meant to represent a healthy anatomy. The reader is invited to qualitatively compare the parametric model and a geometry reconstructed from a real CT scan, (figure 1.2). The figure highlights the main anatomical features using colours: septum, inferior, and middle turbinates. The synthetic nose is made of planar and constant-curvature surfaces; however, it replicates

the main anatomical features of a human nose, e.g., the division into two fossae and the hook-like structure of the turbinates.

The built parametric CAD is driven by eight geometrical parameters, which account for both anatomical and pathological variability. Three parameters describe the intensity of pathologies related to hypertrophies of the turbinates, while the remaining five parameters represent the physiological anatomical variability among patients. All the parameters are set to zero in the reference healthy anatomy. By varying these two groups of parameters, ninety-nine extra healthy anatomies and a hundred pathological anatomies are created, for a final dataset of two hundred geometries. Varying the parameters produces small and localized changes in the geometry; the cross-sectional area reduction compared to the reference geometry ranges between 0.7 and 14%.

1.2.2 Realistic dataset D2

When using real anatomies and real pathologies, it is not possible to fully parameterize the problem, as shown in D1. No matter the anatomical variability or the combination of pathologies, it is paramount to ensure there is no room for ambiguity in the labels. In collaboration with our team of ENT surgeons at the San Paolo Hospital University of Milan, a tree of all the possible deformations has been compiled. This ensures an unambiguous classification of the patients.

This new dataset (D2) is built upon 7 healthy anatomies, identified with P_1, P_2, \dots, P_7 , although limited in number, these anatomies do contain a fair amount of variability. Each healthy anatomy is derived from a CT scan of a patient with a normal sinonasal anatomy. Then a selection of pathologies (or a combination of them) is injected into each patient to expand the dataset to 325 different geometries. This dataset only includes pathological anatomies.

The following subsection goes into detail on how to reconstruct the geometry starting from a CT scan, how to choose an unambiguous label for realistic geometries, and how to inject consistent pathologies into the healthy anatomies.

Reconstruction of the geometry

The first step of the CFD analysis begins with the acquisition of a CT scan. The reconstructed surface quality varies greatly depending on the spatial resolution set during the acquisition (Quadrio *et al.*, 2014). Excellent spatial resolution is estimated around 0.4 mm as in Zachov *et al.* (2009), the CT scans reconstructed in this thesis have a spatial resolution of maximum 0.6 mm. The open-source software 3D-Slicer (Fedorov *et al.*, 2012) is employed to manipulate and reconstruct the CT scan images into a STereoLithography (STL) file. The only non-trivial step is the choice of the radiodensity threshold, which identifies the boundary of the surface of interest for the CFD analysis. Due to the high variability of results depending on the choice of this threshold (Quadrio *et al.*, 2016), the CT scan reconstructions in this Thesis are done at a fixed, consistent value.

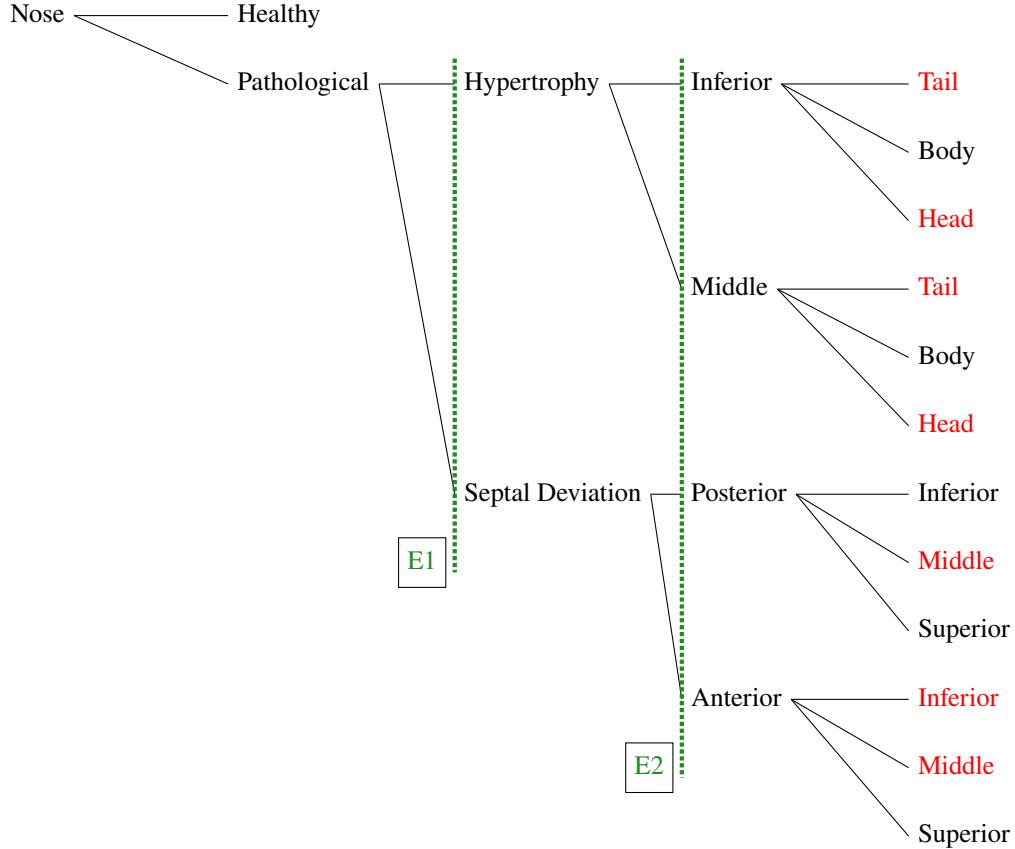


Figure 1.3: Tree of pathologies for the noses: each tree branch is one class. Every pathology is accompanied by a grade of severity (deformation amplitude). The dotted lines indicate the classification experiments. The red leaves at the rightmost level indicate pathologies that are actually considered in the present work.

The deformation tree

The tree of deformations in figure 1.3 is a convenient tool to classify the pathologies in the human nasal cavities. Each pathological patient is represented by switching on several of the leaves of the tree, while a healthy anatomy would have all the leaves off. This representation is particularly advantageous since it also provides a direct link between pathology and surgery: it shows the smallest possible pathological deformations that can be surgically corrected with a single surgical manoeuvre. Thus, the final outcome of the classification immediately provides a direct indication of what surgery should be performed to correct the anomaly.

To build the dataset, a subset of pathologies is selected from the tree, and then they are manually injected onto the healthy patient P_1 , by a sort of inverse virtual surgery.

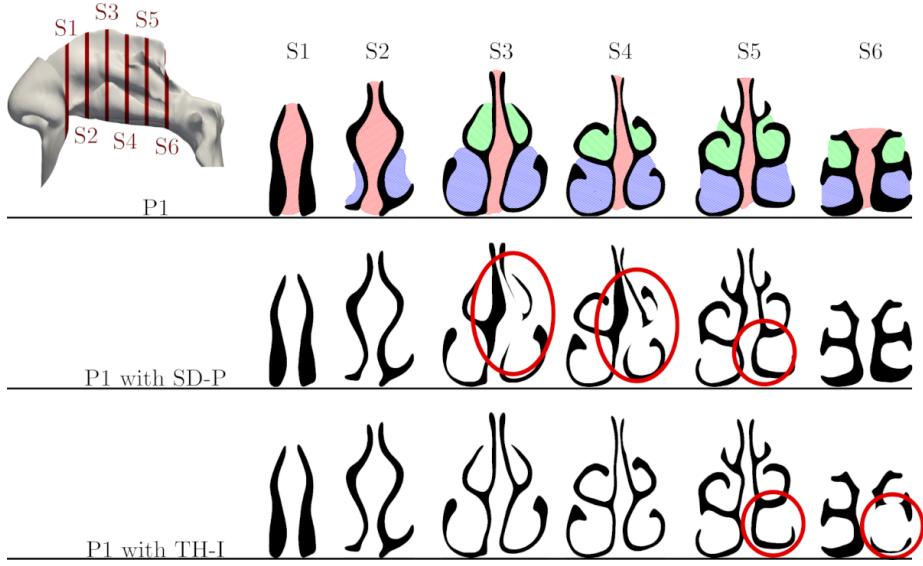


Figure 1.4: Healthy anatomy of patient P_1 (top row) versus two pathological modifications (severe septal deviation located posteriorly, SD-P, in the middle row; severe hypertrophy of the inferior turbinate, TH-I, in the bottom row). Each anatomy is described via six coronal sections, defined on the three-dimensional view. Colours indicate the main anatomical areas: the passageways are coloured black; the inferior turbinates are blue, the middle turbinates are green, the septum is light red. Red circles highlight regions altered by the pathology.

The tool of choice is the free and open-source 3D computer graphics software Blender (van Gumster, 2015). The defects studied in this work belong to two well-separated classes: septal deviations (SD) and turbinate hypertrophies (TH). The SD of major interest in this work are the anterior (SD-A) and posterior (SD-P), the TH of major interest are the ones affecting the middle (TH-M) or inferior (TH-I) turbinates. Only seven of these defects are considered in this study, as indicated by the red leaves in figure 1.3. Each pathology is characterized by two intensities (mild or severe), with the severe one leading to a contact between turbinate and septum. An example of these deformations is given in figure 1.4, which, for patient P_1 , illustrates the healthy anatomy and two of its pathological variations. A three-dimensional view defines six coronal sections, indicated with S_1, \dots, S_6 . The main anatomical regions identified in colour: septum, inferior, and middle turbinates. The six sections are equally spaced in the sagittal direction. The central row shows a severe SD-P located medially, and the bottom row is a severe TH-I affecting the turbinate tail. The anatomical changes corresponding to the pathologies are circled. The two pathologies affect the fossae in different ways; the SD restricts one fossa but enlarges the other; the TH restricts only one fossa.

The procedure of manually injecting the pathologies into the patients is extremely time-consuming; hence, to speed up the process and to ensure consistency in the design

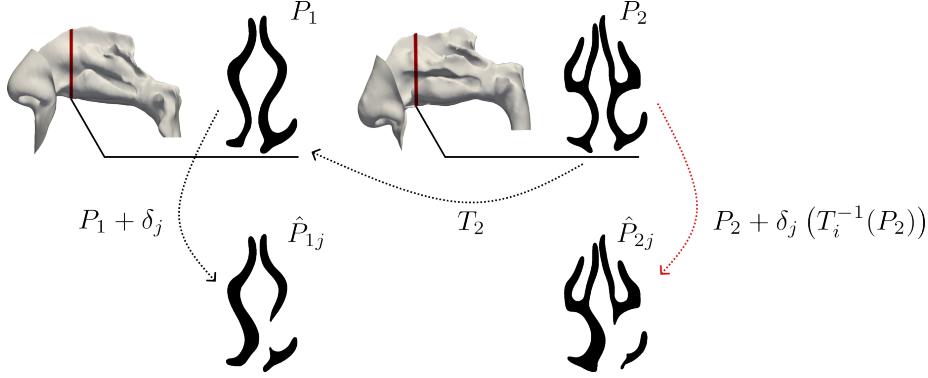


Figure 1.5: General workflow for the transport of pathological deformations. Left column: reference patient in its healthy state (top) and with the j -th pathological conditions (bottom), a cross-section (highlighted in red in the 3D view) is extracted to better understand the changes introduced by the pathology. Right column: healthy and pathological cross-section for the patient P_2 .

of the pathologies, an automatic pipeline is designed, which is described in the next section.

Consistent transport of deformations

In the previous section, the patient P_1 has been equipped with a set of deformations δ_j to obtain the pathological variants $\hat{P}_{1,j}$, where δ_j is the difference in coordinates between P_1 and $\hat{P}_{1,j}$. This section explains how to map the deformation functions to other healthy patients to automatically and consistently extend the dataset. For this purpose, a new tool called Functional Maps (FM) is introduced.

FM is a computational geometry tool that finds correspondence between shapes and allows the computation of a point-to-point map (Ovsjanikov *et al.*, 2012). It allows us to automatically apply the deformation δ_j to the generic healthy patient P_i and obtain $\hat{P}_{i,j}$. The method is first briefly introduced below and then specialized to the specific version employed here (Melzi *et al.*, 2019). A scheme of the full workflow of the data-augmentation procedure is shown in figure 1.5.

Given a pair of shapes \mathcal{M} and \mathcal{N} , the FM method finds the point-to-point map T between them. Let $\Phi^{\mathcal{M}}$ and $\Phi^{\mathcal{N}}$ be two bases defined each on his shape, and let f and g be arbitrary real-valued functions, then it is possible to write any function as the product of the chosen base and a vector of coefficients:

$$f = \sum_i a_i \Phi_i^{\mathcal{M}}, \quad g = \sum_i b_i \Phi_i^{\mathcal{N}}.$$

Ovsjanikov *et al.* (2012) shows that it is possible to link the coefficients a_i and b_i as:

$$b_j = \sum_i a_i C_{ij}.$$

Where the matrix C is the functional map. Note that C is independent from the real-valued functions f and g , but uniquely depends on the basis $\Phi^{\mathcal{M}}$ and $\Phi^{\mathcal{N}}$. This representation is flexible, in the sense that it does not require the use of a particular base. In the original paper, it is suggested to use the eigenfunctions of the Laplace-Beltrami operator, which are ordered from low to high spatial frequency. The series is truncated to the first k coefficients and a low-pass approximation of the map C is obtained.

A typical pipeline for computing a correspondence using the functional map representation is as follows:

1. Compute a base on each shape (e.g. Laplace-Beltrami)
2. Compute informative descriptors
3. Optimize the functional map C (e.g. based on descriptor functions or landmark correspondence and regularization)
4. Convert C to a point-to-point map T .

The complexity of the described pipeline depends on the size of the chosen basis and, thus, on the dimensionality of the spectral embedding. To recover an accurate pointwise correspondence, this approach is specialized into a method called *Zoom-out*. Instead of computing the full functional map C in one optimization process, Melzi *et al.* (2019) suggest starting by computing a small functional map C_0 and later extending it to a new map C_1 of size $(k_{\mathcal{M}+1} \times k_{\mathcal{N}+1})$, basically an iterative process that tends to preserve a low frequency structure. This procedure can be iterated to obtain progressively larger functional maps C_0, C_1, \dots, C_n until a sufficient large n . The map retrieved with this method is more accurate, leading to better results.

In our application, we have a set of healthy patients P_i with $i = 1, \dots, 7$. The patient P_1 is the reference geometry, which is mapped onto the other six. A set of vector fields δ_j is defined on the reference patient, describing the transformation from the healthy status to the j -th pathological condition \hat{P}_{1j} . By finding the point-to-point map between the reference patient and the i -th patient T_i , it is possible to retrieve the generic pathological patient $\hat{P}_{i,j}$ as:

$$\hat{P}_{i,j} = P_i + \delta_j \left(T_i^{-1}(P_i) \right).$$

At the end of this procedure, 325 unique geometries are obtained. A detailed explanation of the FM method for the consistent transport of deformation is available in §6.

1.3 CFD setup and flow results: X

After generating the dataset of geometries, it is necessary to set up the CFD simulations. In this Thesis, only steady inspirations are considered. This section briefly describes the

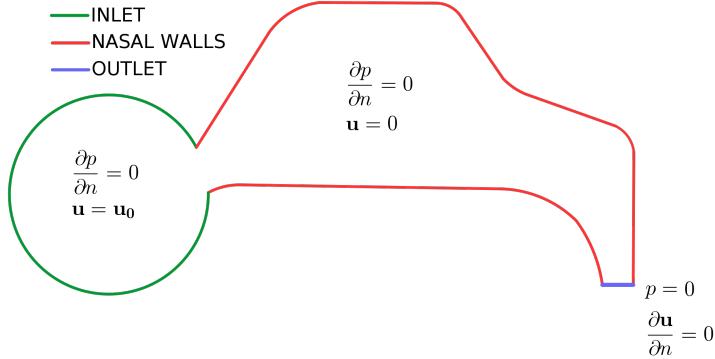


Figure 1.6: A schematic representation of the computational domain. In inspiration, the green sphere represents the inlet, the wall region is in red, and the blue area on the throat is the outlet. Boundary conditions are also reported.

setup of a CFD simulation for the human nasal cavities, followed by a brief description of the flow's main characteristics.

The computational domain is discretized using the tools available within the flow solver OpenFOAM (Weller *et al.*, 1998). The meshes presented in this thesis vary considerably depending on the final application. Following the literature standards, RANS simulations consist of a few million cells, whereas LES simulations consist of around 15 million cells. This number of cells allows us to refine enough near the solid boundaries and capture the velocity gradients without using layers. It should also be recalled that the paranasal sinuses are left out, hence the computational volume is decreased and the resolution improved for a given number of cells. The external environment is represented via a sphere placed in front of the external nose. In order to perform the CFD simulation of the human nasal cavities, several modelling choices have to be made. Even when considering an inspiration at rest, the flow of the nasal cavities is often laminar (Chung *et al.*, 2006), albeit vortical, chaotic, and three-dimensional. In the literature, some authors prefer to use a turbulence model, either RANS (Li *et al.*, 2017), LES (Calmet *et al.*, 2021) or "laminar" (Zhao *et al.*, 2004), the latter being the same steady solver used for RANS, but running without a turbulence model (under the assumption of steady flow). The flow in the nasal cavities is often enforced either by a constant pressure difference between the external environment and the throat (Radulesco *et al.*, 2019) or by a certain flow rate (Brüning *et al.*, 2020).

All the CFD simulations illustrated in this Thesis are steady, incompressible inspiration, most of which are done at a constant flow rate, since we believe that driving the flow at a constant flow rate is perhaps more physically sound. For example, a typical LES inspiration performed in this Thesis has the following setup: as for boundary conditions, the surface of the nasal cavities is considered a wall, with no-slip and no-penetration boundary conditions. Pressure is considered zero at the outlet, and its gradient is considered zero at the inlet. The inspiration rate is fixed at 280 ml/s or $2.8 \times 10^{-4} m^3/s$. Satisfactory statistics for the mean and variance fields can be obtained,

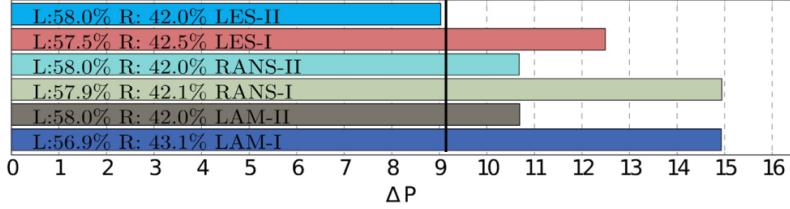


Figure 1.7: Mean pressure difference ΔP between inlet and outlet for six inspirations, LES, RANS and laminar simulations with first and second order schemes. The percentage share of the flow rate in the left (L) and right (R) fossae is also shown within each bar. The vertical line is the reference pressure difference measured by a second-order LES simulation.

at this breathing rate, after a total simulation time of 0.65 s, by discarding an initial transient of 0.05 s (Covello *et al.*, 2018). The time step for the temporal advancement of the solution varies such that the Courant–Friedrichs–Lewy number remains below unity. LES simulations compute the evolution of instantaneous, spatially filtered fields of velocity \mathbf{u} and pressure p , temporally averaging over the duration of the simulations leads to $\bar{\mathbf{U}}$ and \bar{P} . The turbulence model is the Wall-Adapting Local Eddy viscosity, or WALE (Ducros *et al.*, 1999). Figure 1.6 shows a sketch of the computational domain; the green circle represents the sphere in front of the nose, which is the inlet; the wall region is in red; and the outlet is in blue. The boundary conditions are also reported.

All the modelling choices mentioned have a deep impact on the outcome of the simulation. For example, when driving the flow with a constant flow rate, switching from first- to second-order schemes, regardless of how the turbulence is modelled, reduces the pressure drop by about 4 Pa. Switching from LAM/RANS to LES (with the same numerical scheme order) changes the pressure drop to 1.5–2.5 Pa. A brief comparison of these results is given in the right picture in figure 1.7. More details on the sensitivity of the CFD outcome to the modelling choices are given in §2.

1.3.1 The flow main features

Several authors tackled the problem of the human nasal cavities’ flow analysis and description, e.g. Doorly *et al.* (2008b); Wen *et al.* (2008); Lintermann & Schröder (2019). This section describes briefly the main flow features for the realistic dataset D2, which are the LES of steady inspirations. The description focuses mainly on the common behaviour of the flow, especially when it has already been reported in the literature.

The mean fields computed in the P_1 case are taken as an example; figure 1.8 shows the simulation results for a cross-section along one fossa in the sagittal plane. Following the path of the flow in inspiration, the air converges through the nostril to the nasal vestibule and is injected as a high-speed jet into the main cavity, where the outer air is accelerated with the velocity magnitude reaching up to 2–3 m/s. Unsurprisingly, the same area exhibits a peak of wall shear stress. Several authors recognize this region as

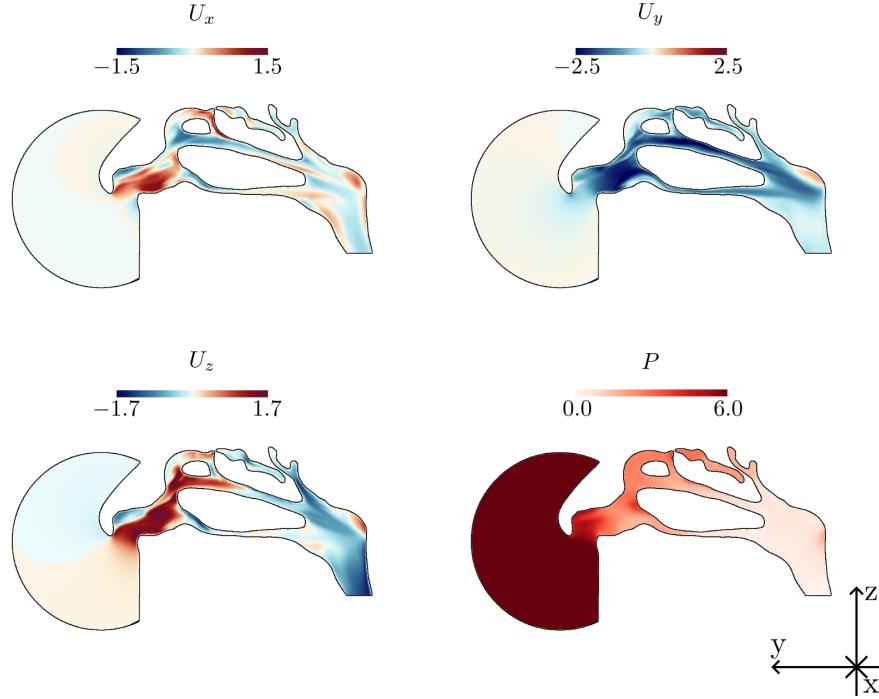


Figure 1.8: Mean velocity and pressure fields in sagittal view for the healthy patient P_1 .

one of the most influential on the inflow conditions and the formation of subsequent flow structures, being one of the most prominent non-pathological strictions along the path. The flow in the vestibular area enters almost vertically and is slowly deviated, with the most notable deviation happening in correspondence with the middle turbinate head. Then, the air flows around the turbinates through the meati and reaches the nasopharynx. Here, the flow meets another non-pathological striction, where the pressure drop can be more or less pronounced depending on the patient's anatomy. In the nasopharynx, the flow rotates downward, but also produces a recirculation (visualized by the positive U_y component) at the posterior wall of the nasopharynx. The strictions are highlighted by the pressure color map, which, relative to the level $P = 0$ set at the outlet, undergoes the largest drops immediately after the nostrils and in the nasopharynx area.

Figure 1.9 compares section S6 for healthy and several pathological conditions, to better understand the consequences of the pathology's injection. The left figure shows the velocity's magnitude for the healthy patient P_1 , on the right, the eight figures show different pathological conditions, and the color map shows the velocity difference with the healthy reference. In a healthy nose, the flow rate is usually fairly distributed between the fossae. A great flow imbalance is seen mostly for pathologies involving the anterior part of the nose, such as SD-A (figure 1.9, the first column shows the anterior SDs). Local flow unbalances are more common in TH and most noticeable if far from the anterior part (figure 1.9, the last column shows a severe TH-I of the tail). Further

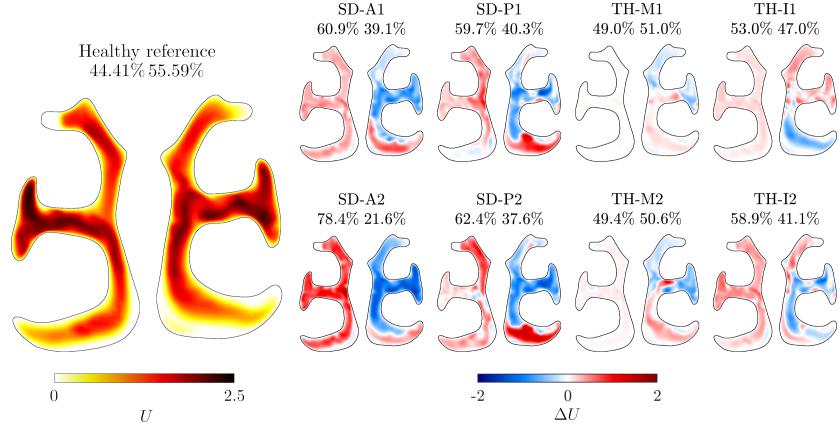


Figure 1.9: Velocity magnitude for P_1 in section S6, in comparative form between the reference anatomy and the four pathologies (SD-A, SD-P, TH-M, TH-I), at both severities. The percentages above each panel indicate the distribution of the flow rate across the two fossae.

analysis of the flow is available in §7.

This analysis helps to better understand the flow physics in the human nose; however, the CFD study of the human nasal cavities also has more practical applications. The CFD outcome can be later used not only for surgery planning or to better understand the flow field but also to prove some medical hypothesis. For example, in §8 ENT doctors check the possibility of the nasal mucosa getting infected with SARS-CoV-2 by the lungs and not vice versa.

1.4 Feature extraction procedure: f

As already discussed, the large size of the CFD dataset requires a dimensionality reduction process to extract an informative subset of X . The feature extraction process also has several benefits, such as reducing the risk of overfitting and increasing the explainability of the model. In ML, dedicated explainability methods exist, such as SHAP (Lundberg & Lee, 2017), e.g. used by (Cremades *et al.*, 2023) to identify the importance of each coherent structure.

The outcome of the feature extraction depends on the specific experiment that is performed; however, this section illustrates two techniques. The first one is to use handcrafted features §1.4.1, in which the experience matured in the flow analysis drives the engineering of the features (as in section §1.3.1). The second approach makes use of the FM method for selecting a few informative bases to describe the whole flow field §1.4.2. In both scenarios, the idea is to use Galilean invariant flow quantities due to the high geometrical complexity of the nasal cavities.

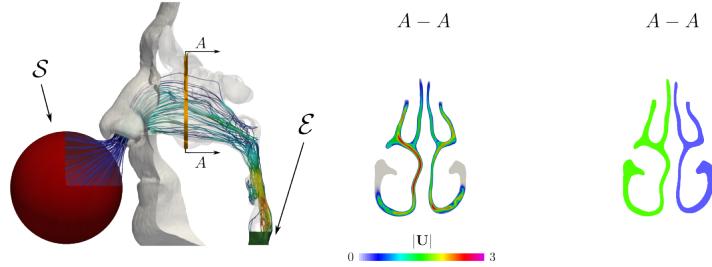


Figure 1.10: Airflow in the human nasal cavities during inspiration. Left: streamlines start from region S and end in region E . The orange slice indicates the cross-sectional cut plotted in the other two panels. Center: Mean velocity component normal to the cross-sectional cut. Right: Division of the plane into 2 regions coloured with the value of the regional average velocity.

1.4.1 Expert-driven features

Inspired by the practice of analyzing flow fields, we devised two kinds of features: integral information convected by streamlines and regional averages of flow variables. Selecting both a start region S (the inlet sphere) and an end region E (the throat), it is possible to compute the streamlines that evolve along a flow field, figure 1.10 left shows a diagram of the streamlines for a generic anatomy. It is possible to compute several flow-related integral quantities along their path, which can be treated as a distribution in the statistical sense, such as mean and variance. This operation can provide a compact and meaningful description of the flow. Streamlines entangled by vortices are inherently longer than straight streamlines. Features extracted from streamlines are very practical; they compactly convey flow information while sampling most of the volume with minimal knowledge of the geometry.

Other informative features can be extracted by averaging the flow quantities over pre-defined regions, which we call regional averages. To take into account the uneven layout of the samples, these averages are area-weighted. Figure 1.10 center illustrates the magnitude of velocity for a typical cross-section, and on the right, the same cross-section is divided into two regional averages for the modulus of velocity. Information extracted from regionally averaged features heavily depends on whether the set of selected regions is meaningful. The flow quantities used for feature extraction are scalar, such as pressure and turbulent viscosity, or the module of vectorial one, e.g. magnitude of velocity and magnitude of vorticity. Detailed information on handcrafted features is available in §5 and §7.

1.4.2 Mapped fields

This feature extraction method makes use of the FM (presented in §1.2.2), to map flow fields between different patients, thus making possible the comparison between simulations made on different domains. Intuitively, the transfer of information between two domains can be of great aid in a classification task; in fact, the use of FM for

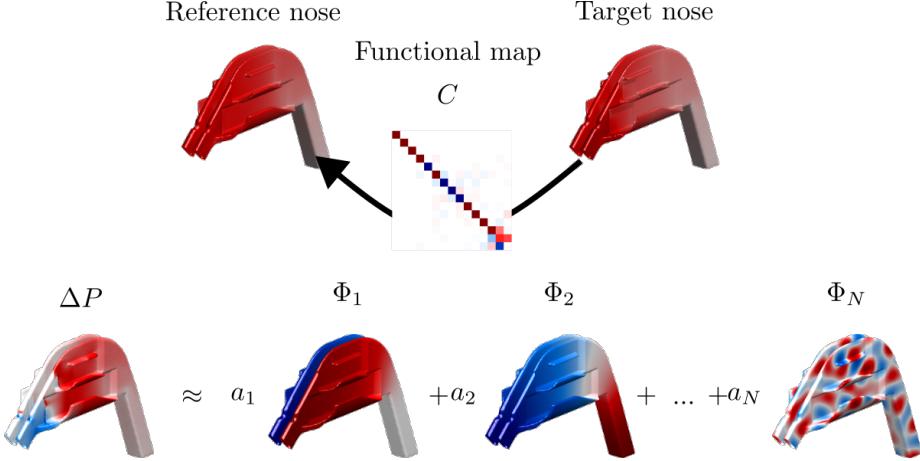


Figure 1.11: The difference between a wall-based flow quantity (say, pressure P) between the baseline anatomy and a generic anatomy. The difference ΔP is expressed as a linear combination of the eigenfunctions of the reference nose with coefficients a_i .

classification in the medical domain is not new, e.g. Magnet *et al.* (2023) uses FM for the quantification of craniofacial anomalies. Due to the purely geometric nature of the FM procedure, we aim to build a feature extraction method to understand whether flow features are more effective than geometrical ones, applying it to the synthetic dataset (D1). Figure 1.11 shows the general workflow of mapping the time-averaged pressure field of a pathological patient onto the reference one. The procedure finalizes when retrieving the coefficients that decompose the pressure difference using the Laplace-Beltrami base. Now that it is possible to decompose any field in the base of the reference geometry, the input of the NN will be the coefficients a_i .

The most immediate wall-based flow features that are more self-evident in an incompressible flow are the pressure P (F1) and the magnitude of wall shear stress τ (F2). This choice is also logical when considering the clinical experience, since the feeling of discomfort in a patient is conveyed by nerve terminations residing in the mucosa. These two fields can be mapped into the reference geometry, and once again, the coefficients can encode the difference between the i -th patient and the reference one.

With this method, purely geometrical features can also be explored. A first geometrical feature (G1) can simply be the distance between points in the baseline geometry and the corresponding ones in the modified anatomy. The coefficients encoding the distance field over the eigenmodes of the baseline geometry can be used as inputs for a classifier. Another purely geometrical feature (G2) can simply be the functional map C between the two geometries.

The result of this classification task is presented in the section §1.6.1. The next section presents how to build the classifier \mathcal{K} to use the extracted features. A more detailed presentation of the method is available in §4.

1.5 The classifier \mathcal{K}

Given the dataset with l observations, the i -th feature f vector must be associated with a target value that describes, for example, the severity of a pathological parameter. This regressor can be implemented in a NN in which either the handcrafted features or the coefficients extracted through the FM can be used as input. The most generic classifier used in the thesis is a Multi-Layer Perceptron (MLP) (Goodfellow *et al.*, 2016), which is suitable when the input is a feature vector.

In general, designing the architecture of a NN involves several choices: the number of hidden layers and nodes, the activation function, and the loss function. When our MLP is a regression network, it has an input layer whose number of nodes is equal to the length of the feature vector, three hidden layers with 30, 20, and 10 nodes each, and an output layer with only one node. The activation function is the hyperbolic tangent for all the nodes, except for the output nodes. The optimization algorithm, which updates weights and biases of the NN, is the classic Levenberg–Marquardt (Lera & Pinzolas, 2002). When the MLP is used for regression, i.e. the estimation of the pathological parameters, the output node is only one, its activation function is linear, and the loss function is the mean square error. Instead, when performing a classification task, the number of output nodes is equal to the number of classes; their activation is the sigmoid function, and the loss function is the cross-entropy.

We would like to point out that a MLP classifier is not the only possible architecture; a slice of the whole flow field can be used as the input of a Convolution Neural Network (CNN) (Gu *et al.*, 2018), which is best suited to account for local connectivity and is rotational invariant. However, giving a slice of the flow field, the CNN would not only have the flow information, but also some detail of the patient’s geometry. The objective of the thesis is to train a classifier based only on flow data, hence the decision to not use a CNN for this work.

In this Thesis, the performance of the classifier is evaluated mainly through two standard metrics: accuracy and F1-score. Accuracy is the ratio between the number of correct and total predictions, which works well as long as the dataset contains a balanced number of observations for each class. When evaluating a classification model, another metric is also used: the F1-score. It is defined as the harmonic mean of precision and recall, with precision being the ratio between true positives and total positives and recall being the ratio between true positives and the total number of relevant elements. F1-score ranges in $[0, 1]$ and should be preferred to accuracy whenever the cardinality of classes is unbalanced, since a high F1-score is obtained when both false positives and false negatives are low. This alternative metric is perhaps less intuitive but more robust.

1.6 Results of the data-driven approach

This section describes the main results of the Thesis’ work, starting with comparing flow and geometric features §1.6.1, classifying nasal pathologies, and discussing the best dataset partitioning methods for a medical dataset §1.6.2.

1.6.1 Geometric vs Flow features

The FM features described in §1.4.2 are here used to predict the pathological parameters for the dataset D1. Figure 1.12 shows the results for a regression task in which each column (q_1 , q_2 , and q_3) is one of the three pathological parameters described in §1.2.1. Each row represents one of the features: the top two rows are geometrical, and the last two flow features. The horizontal axis represents the ground truth, and the vertical axis represents the predicted value; therefore, points lying on the bisector are predicted correctly. The geometric features G1 and G2 do not perform particularly well; the very limited anatomical variability in D1 is already sufficient to throw off the model. Flow features F1 and F2 instead demonstrate good regression capabilities. In this simple dataset, the pathology is a small geometrical modification localized in space. As such, it tends to be visible only in the high-order modes of the Laplace–Beltrami basis employed to decompose the fields. Furthermore, the high-order modes tend to be noisy and dependent on the geometrical discretization. When the regression task is tackled with the use of flow features, these small changes in the geometry are amplified by the non-linear and convective nature of the Navier–Stokes equations, a small geometrical modification can radically change the flow field, making the defect more evident. This radical change in flow field can then be easily captured by the low-order modes of the Laplace–Beltrami basis. In this context, where all the geometries are generated from a single reference, the partition of the dataset is not problematic. Therefore, the classic k -fold cross-validation method is employed, using five folds.

A comprehensive description of the dataset, the dataset construction, the use of FM to extract features, and the comparison between anatomical and flow features is provided in §4.

1.6.2 Classification of nasal pathologies

The described results are obtained using the handcrafted features presented in §1.4.1 with the realistic dataset D2, performing a binary classification task, which requires determining which patients are affected by a SD from the ones with TH (E1 in figure 1.3). The complete set of results, including the multiclass experiment E2, is available in §7.

Testing on a medical dataset, such as the one on hand, may lead to different conclusions depending on how it is tested. Table 1.1 shows the classification performance for different features for a binary classification. The table compares the results between the most common k -fold cross-validation and a Leave-One-Out method (LOO). The latter requires that one patient (and all his deformations) be left out of the training set and used for testing only. Table 1.1 shows little difference in accuracy and F1-score across all the features when tested with the k -fold cross-validation method. Therefore, from here on, only LOO will be considered, since it makes the differences among features more evident.

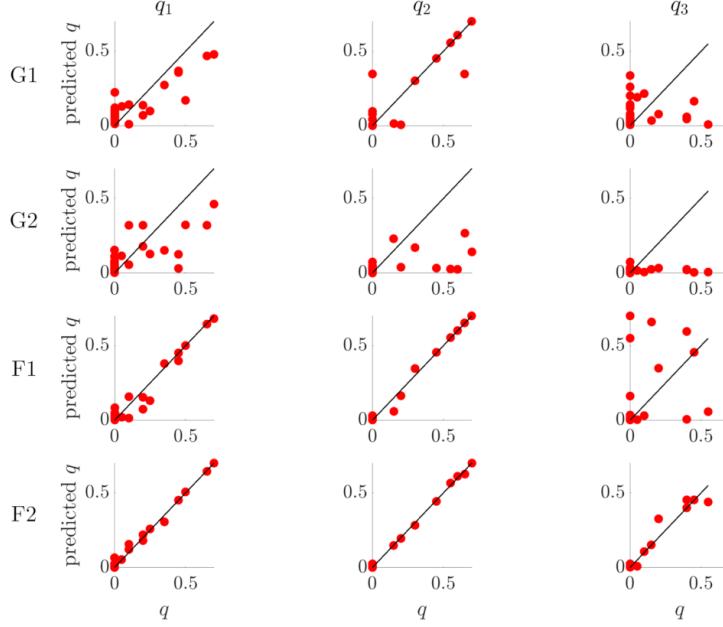


Figure 1.12: Performance of the various features (rows) in one experiment for predicting the three pathological parameters (columns). G1: distance; G2: matrix C ; F1: pressure; F2: wall shear stress. Ground truth on the horizontal axis, and predicted value on the vertical axis.

Regional averages

It seems that the most effective feature for regional averaging is the velocity magnitude U . It must be noted that in this setup, the magnitude of velocity is deeply connected to the area A : the flow is incompressible and its rate is imposed as a boundary condition; therefore, only the flow rate division between the right and left fossa is free. However, this effect is overemphasized by the use of a small dataset and by the choice of pathologies, which act in a different way on the area of the two fossae: the TH reduces the area of only one fossa, while the SD shrinks the passageways in one fossa but enlarges the other. It is worth noting that other pathologies would be transparent to the change in cross-sectional area, such as the septal perforations.

However, there are several interesting alternatives to the magnitude of velocity. The second-best quantity is the magnitude of the mean pressure gradient ∇P , whose regionally averaged value provides a simple estimate of the local rate of pressure loss along the axis of the nasal fossa. Hence, this quantity also relates to the variation of the cross-sectional area perpendicular to the flow field. A further quantity related to dissipative processes in the flow is the norm of the mean velocity gradients, computed

	k-fold		LOO	
	accuracy	F1	accuracy	F1
U	0.97	0.95	0.85	0.80
$ \nabla P $	0.96	0.93	0.76	0.71
$P_{in} - P$	0.91	0.87	0.76	0.64
$P_1 - P$	0.91	0.87	0.76	0.68
Ω^2	0.95	0.92	0.74	0.68
$P - P_{out}$	0.89	0.84	0.68	0.61
$P - P_6$	0.92	0.88	0.74	0.67
v_t	0.87	0.79	0.67	0.57
R	0.85	0.75	0.64	0.56

Table 1.1: Performance of various mean flow quantities (velocity magnitude, pressure gradient magnitude, pressure computed with respect to various references, norm of the velocity gradients, turbulent viscosity and nasal resistance) as input features in the binary classification of the noses dataset, experiment E1 in figure 1.3. Except for nasal resistance, the remaining inputs are the 12 left/right regional averages in S1, ..., S6.

for incompressible flows as the squared norm of the vector:

$$\Omega = \frac{1}{2} \nabla \times U.$$

The performance achieved using $|\nabla P|$ and Ω^2 is comparable.

Pressure itself has the potential to be a good indicator for pathologies since, from the pressure drops along the airflow, it is possible to spot local narrowings and therefore probable pathologies. However, in incompressible flows, only pressure differences are significant, and the choice of the reference pressure for a classification task with different anatomies is not totally irrelevant. Several choices are possible; however, the anatomy of the human nose must be taken into account since there are two important non-pathological narrowings, i.e. the nasal valve in the anterior part, between the inlet and S1 section, and the laryngeal striction at the back, between the S6 section and the outlet. Thus, when considering the pressure difference with the outlet $P - P_{out}$, the pressure drop at the nasal valve is not relevant but contains the pressure jump at the larynx. The pressure difference with the inlet $P_{in} - P$ has the opposite problem; the laryngeal striction is not relevant but contains the nasal valve jump. Another viable possibility is to exclude both non-pathological pressure drops by computing the pressure difference to either S1 or S6. However, the features in either plane are lost now because they are identically zero and the number of NN's inputs is reduced by two. As for performance, $P - P_{out}$ has the worst predictive capabilities at 68%, the other three are around 76% accurate, highlighting the negative effect of keeping the throat pressure jump. The turbulent viscosity does not perform well as a feature, and this is hardly surprising since the present database is computed with high-resolution LES. In general, using v_t as a feature would be a delicate choice, as this quantity in LES is highly mesh-dependent. For the present cases, run at the same spatial resolution, v_t is at least comparable across cases, but its values are extremely small since the simulations nearly

	Mean streamlines				Instantaneous streamlines			
	accuracy	precision	recall	F1	accuracy	precision	recall	F1
$\int \Omega^2 ds$	0.81	0.76	0.62	0.68	0.83	0.82	0.62	0.71
$\int v_t ds$	0.68	0.52	0.52	0.52	0.61	0.41	0.41	0.41
$\int \nabla P ds$	0.67	0.51	0.51	0.51	0.70	0.54	0.54	0.54
$\int U ds$	0.60	0.42	0.49	0.49	0.72	0.58	0.50	0.54

Table 1.2: Binary classification E1 of the noses dataset (LOO validation), for various flow quantities evaluated over mean (left) and instantaneous (right) streamlines.

resolve all flow scales.

Finally, a special mention is deserved by the nasal resistance R , which is in fact closely related to the hydraulic resistance of the duct and perhaps the simplest and most fundamental way of characterizing the passageways from an aerodynamic point of view. R is computed separately for the two right/left fossae as:

$$R_{r/\ell} = \frac{P_{1,r/\ell} - P_{6,r/\ell}}{Q_{r/\ell}};$$

where Q is the flow rate, and the subscript r/ℓ indicates quantities computed for the right/left fossa. In the expression above, the pressure drop is taken across S1 and S6, and thus gets rid of the extra losses localized at the nasal valve and at the larynx.

Streamlines

Selected flow quantities are integrated along each streamline, from where the streamline intersects section S1 to the point where the streamline intersects S6. The mean and the variance of these averaged values are computed over each group of streamlines; thus, streamlines end up providing four features. Table 1.2 shows the results of the binary classification experiment E1, where the four input features are computed by integrating various flow quantities over the streamlines.

Comparing these results with those reported in table 1.1 obtained with regional averages, it is noticed that the quality of the streamlines-based prediction remains generally high, even though the number of features is significantly reduced, from 12 to 4. Moreover, the best-performing features take some extra edge when computed over an instantaneous streamline rather than on a mean one. The best-performing feature is the integral of Ω^2 , which yields better results than the corresponding regionally averaged values, improving from 74% to 83%. However, at the same time, the best-performing regionally-averaged feature (U) shows a large loss of accuracy from 85% to 72%. Using instantaneous and mean streamlines involves feature-dependent differences. The quantity U^* , especially in view of the regional average, is a very large-scale quantity that is expected to change gradually between a pair of adjacent cross-sections; the non-uniform sampling of the cross-sectional area implied by integrating U over streamlines degrades performance significantly. On the other hand, the smaller-scale quantity Ω^2 is positively affected by the ability to continuously sample the volume from S1 to S6.

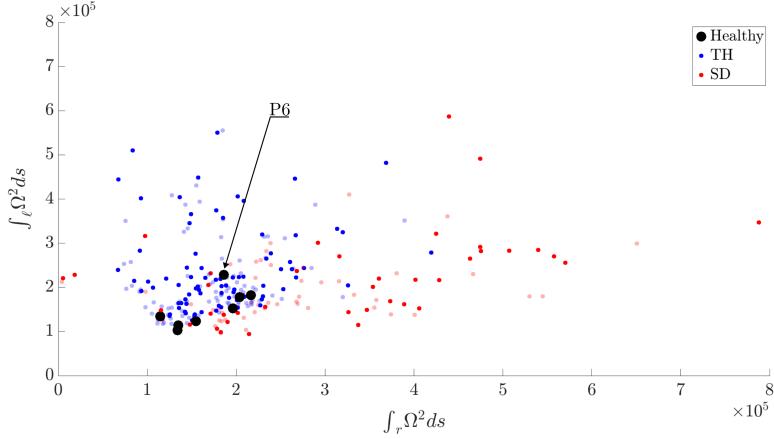


Figure 1.13: Integral of Ω^2 over instantaneous streamlines: mean values on the right/left group of streamlines are plotted on the horizontal/vertical axes. Black dots are the healthy patients, red and blue symbols represent SD and TH. Full/light colour markers indicate severe/mild pathologies.

In figure 1.13 the mean value for the integral of Ω^2 for the right (horizontal axis) and left (vertical axis) groups of streamlines is plotted. Which are two of the four features used to obtain the results in table 1.2. The figure shows that the healthy patients (shown with black dots) tend to cluster, which shows the ability of the feature to ignore anatomical variability. However, in this plot, severe (filled symbols) and mild (light symbols) pathologies seem to overlap. Inspecting the healthy cluster closely, it is possible to notice how one of the black dots is a bit further from the others. It corresponds to patient P_6 , highlighted in figure 1.13. Indeed, a clinical re-evaluation of the CT scan for this patient reveals that its anatomy is affected by a minor hypertrophy at the tail of the inferior turbinate that escaped the clinical screening during the selection of the healthy patients. This is unfortunate because it adds a small amount of unwanted noise to the entire dataset D2. However, it constitutes an indirect indication of the sensitivity of this feature, which has alerted us on a functional information that escaped the analysis of expert clinicians.

1.7 Conclusions

This Thesis describes a novel ML approach to classifying nasal pathologies using flow features. The main application scenario concerns the air flow in the human nose; however, the methodology remains flexible, being able to generically classify shapes immersed in a moving fluid. Starting from the formulation of the problem, this work briefly explores all the main requirements necessary to setup a data-driven model: the generation of the dataset, the setup of the CFD simulations, and the feature extraction process. The definition of a deformation tree, directly linking pathologies to

their surgery, allows the parameterization of the human nose, unambiguously defining a label for each anatomy. A selection of these pathologies is then injected into seven healthy anatomies to build a consistent dataset through a novel method to perform data-augmentation, which allows for different anatomies to have the exact same pathology. In the future, this procedure may be easily used to perform consistent virtual surgeries. The comparison between anatomical and flow features is also explored. Showing that when using features extracted from the flow field computed with CFD, the training of a NN becomes substantially easier in comparison to equivalent networks that rely only on geometry-based features. However, the main goal remains to build a classifier to discriminate pathological defects over data with high physiological variability. Different flow features are selected and tested in comparative form, ranging from regional averages to streamlines to fully mapped surface fields. The fluid mechanical understanding of their classification performance is sought with various success; one possible outlook is the use of more dedicated methods to explain the influence of each feature in the classification process, such as SHAP. Although the work of this Thesis is still at a preliminary stage, it is already possible to achieve high classification accuracy even when simulating a clinically challenging setting. In particular, the classification accuracy is above 80% when tested with real patients §6.

However, usability in clinical applications would require a more extended dataset, including both more anatomical variability and a greater set of pathologies.

Chapter 2

Importance of the numerical schemes in the CFD of the human nose

2.1 Abstract

Computational fluid dynamics of the air flow in the human nasal cavities, starting from patient-specific Computer Tomography (CT) scans, is an important tool for diagnostics and surgery planning. However, a complete and systematic assessment of the influence of the main modeling assumptions is still lacking. In designing such simulations, choosing the discretization scheme, which is the main subject of the present work, is an often overlooked decision of primary importance. We use a comparison framework to quantify the effects of the major design choices on the results. The reconstructed airways of a healthy, representative adult patient are used to set up a computational study where such effects are systematically measured. It is found that the choice of the numerical scheme is the most important aspect, although all varied parameters impact the solution noticeably. For a physiologically meaningful flow rate, changes of the global pressure drop up to more than 50% are observed; locally, velocity differences can become extremely significant. Our results call for an improved standard in the description of this type of numerical studies, where way too often the order of accuracy of the numerical scheme is not mentioned.

2.2 Introduction

Nasal breathing difficulties are a widespread pathological condition, accompanied by significant economical and social costs (Smith *et al.*, 2015; Rudmik *et al.*, 2015). A precise diagnosis is often difficult to achieve, corrective surgeries are sometimes required, yet after certain nose surgeries the majority of patients remains unsatisfied (Sundh & Sunnergren, 2015).

Starting about two decades ago, numerical studies of nasal airflow based on Computational Fluid Dynamics (CFD) began to increase in number and quality. Nowadays, Ear, Nose and Throat (ENT) doctors envisage the use of a detailed CFD solution to diagnose pathologies and to plan surgeries (Radulesco *et al.*, 2020; Singh & Inthavong, 2021). A recent, broad and insightful account of potential and open problems is given by Inthavong *et al.* (2019).

There is thus a growing need for a thorough validation and standardization of CFD methods and procedures. Several aspects, like the spatial resolution of the computational mesh (Frank-Ito *et al.*, 2016), or the radio-density threshold employed for CT segmentation (Zwicker *et al.*, 2018) have been specifically discussed, but a systematic assessment of the sensitivity of the CFD outcome to the various sources of uncertainty involved in the procedure is still required, noticeably so in respect to the discretization errors incurred by the numerical method. The present work describes and compares within a unified framework two major contributors to the global error in a well conducted CFD simulation: how the flow physics is modeled, and which schemes are used in the numerical solution. The former contribution has been discussed several times, while the latter has never been addressed.

CFD simulations of the nasal airflow nowadays leverage the entire spectrum of flow modeling choices, ranging from Direct Numerical Simulations (DNS) to Large-Eddy Simulations (LES) and Reynolds-averaged Navier–Stokes equations (RANS). Moreover, "laminar" simulations are also employed, where a steady RANS solver is used without a turbulence model under the assumption of steady flow. RANS assumes the flow to be turbulent, employs a (dissipative) turbulence model to describe the effect of the turbulent fluctuating field on the time-averaged motion, and only computes a time-averaged solution via a steady solver; it represents the computationally cheapest approach, with the largest amount of modeling error. DNS is at the other end of the spectrum: it solves the unsteady equations of motion without a turbulence model, because the solution takes place on a spatial mesh fine enough to resolve all the significant flow scales; the obvious downside is the computational cost. LES is midway between the two extrema, but akin to DNS: the solution is time-dependent and relatively expensive from a computational standpoint, while the role of the turbulence model, which is still required, is relatively minor and can be controlled via the size of the mesh. A further option, still used scarcely in this field, is the combined use (see e.g. Van Strien *et al.*, 2021) of RANS and LES with the so called hybrid methods, which are able to bring forth the unsteady character of the flow in the nasopharynx even at low flow rates.

The importance of flow modelling is well known. For example, Zhao and coworkers (Li *et al.*, 2017) thoroughly compared results from several RANS models, one LES model and a reference DNS, for an artificial anatomy deprived of sinuses for which prior experimental information was available. Within a commercial solver, they used second-order numerical schemes for RANS and bounded second-order schemes for LES. The laminar flow model was found to perform well, at low breathing intensity, to predict the pressure drop, but was observed to not excel at predicting local velocity profiles compared to other approaches. In fact, even for steady boundary conditions, the complex anatomy of the nasal cavity may lead to a three-dimensional and unsteady flow in the nasal fossae of a healthy subject (Churchill *et al.*, 2004) which is mostly laminar at

low flow rates (Chung *et al.*, 2006), but becomes transitional and/or turbulent at higher respiratory rates, especially in the rhinopharynx. Unsteadiness becomes locally very important, even at slow flow, in presence of anatomic anomalies (Saibene *et al.*, 2020), suggesting LES as the preferred approach, especially when particle tracking is involved (Farnoud *et al.*, 2020). While many valuable contributions (Liu *et al.*, 2007; Calmet *et al.*, 2021) employ a time-dependent solution, owing to the lower computational cost several works being published nowadays still remain of the laminar or RANS type.

Less attention has been devoted to another important design choice, whose effects are often underestimated, to the point that most papers do not even mention it: one needs to decide how to discretize the differential operators in the equations of fluid motion. In a finite-volumes CFD software (the most widespread approach), it is customary to have at least two choices available, depending on whether differential operators are discretized at first- or second-order accuracy; some codes allow to pick a different scheme for each term in the differential equations. The formal order of accuracy is the integer power of the cell size that brings the discretization error towards zero (Ferziger & Peric, 2002).

The present work introduces a comparison framework where the effects of the discretization scheme are quantified and compared to those related to the choice of the flow model (laminar, RANS or LES/DNS). Additionally, the same framework is used to quantify the effects of a computational domain truncated at the nasopharynx. Studying domain truncation is not new: e.g. Choi *et al.* (2009) did a similar study for the flow in the lungs, but only considered lower truncations below the larynx with breathing through the mouth. In the present context, and in view of the increasing availability of cone-beam CT scanners, which impart smaller radiation dosages with better spatial resolution at the cost of a smaller field of view (Tretiakow *et al.*, 2020), it is interesting to observe the effects of domain truncation just after the nasal fossae.

2.3 Methods

This paper discusses results from 24 simulations, consisting in 12 inspiration and expiration pairs where every combination of i) first- and second-order numerical schemes, and ii) laminar, RANS and LES modeling is considered. The entire study is carried out twice, on standard (CT) and truncated (TrCT) volumes. A larger LES case with second-order accuracy achieving quasi-DNS spatial resolution provides reference (inspiration only). A detailed comparison between CT and TrCT is described in the Supplementary Material, where additional details of the entire procedure are also mentioned. The various cases are indicated in this paper as for example CT-RANS-II-i, meaning CT-type scan, RANS modeling, second-order schemes, and inspiration. HRLES-II-i indicates the High-Resolution LES case. Normal breathing at rest is simulated by enforcing a steady volumetric flow rate of 280 ml/s for all cases (see e.g. Wang *et al.*, 2012). The baseline head CT scan is that of a male patient with healthy sinonasal anatomy. Figure 2.1 (top) presents the anatomy, reconstructed via standard CT segmentation procedures (Quadrio *et al.*, 2016), and also indicates where the original CT model is truncated above the epiglottis to obtain the TrCT version; the reference system used in the following is shown.

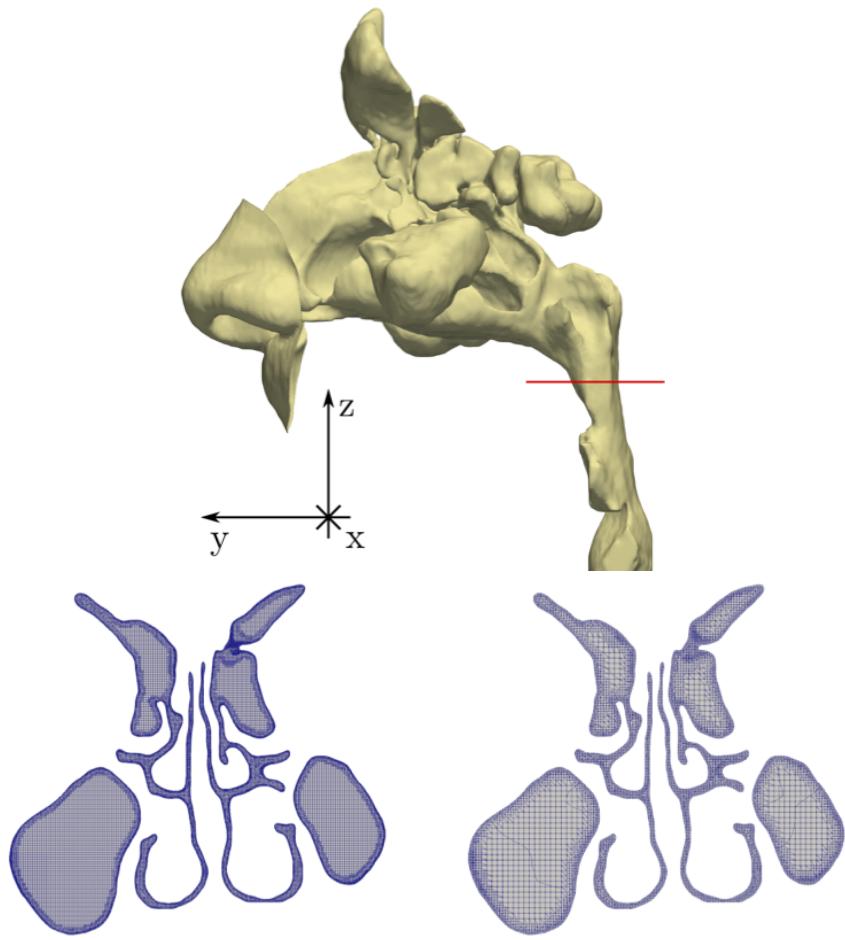


Figure 2.1: Top: three-dimensional view of the CT reconstructed anatomy, the red line is where the volume is cut to mimic the TrCT anatomy. Bottom: coronal section of the volume mesh employed for LES (left) and RANS (right) simulations. Both feature a refinement near the solid boundary.

All simulations are incompressible and carried out within the OpenFOAM (Weller *et al.*, 1998) finite-volumes software package, also used to create the volume mesh. The surface of the nasal cavities is considered as a solid wall, where no-slip and no-penetration boundary conditions are applied; pressure is set to zero at the outlet. The external ambient is represented via a sphere placed in front of the nose. RANS and LES require different meshes, and we have chosen their sizes to be typical of either approach, as determined from a broad literature scan: the RANS mesh has 3.2×10^6 cells (which drop to 2.8×10^6 for TrCT where the total volume is smaller) whereas the LES mesh has about 1.5×10^7 millions of cells (1.4×10^7 for TrCT and more than 50 millions cells for the reference HRLES). A mesh refinement analysis carried out for the RANS mesh and described in the Supplementary Material confirms its adequacy at properly describing the geometry and producing mesh-independent results. The flow is always solved down to the wall, and the use of wall functions is avoided. Figure 2.1 shows a comparison between the RANS and LES meshes.

The RANS turbulence model is the $k - \omega - SST$ model, which is quite popular in such low-Reynolds and transitional flow, and was shown by Li *et al.* (2017) to provide satisfactory results. The LES turbulence model is WALE (Wall-Adapting Local Eddy viscosity), which suits complex geometries well (Nicoud & Ducros, 1999); the high spatial resolution makes the details of the LES model relatively unimportant.

2.4 Results

The 24 cases are first compared in figure 2.2 in terms of a global quantity, i.e. the (absolute value of the) mean pressure drop ΔP between the outer ambient and the lower end of the TrCT scan, marked by the red line in figure 2.1. The percentage flow distribution in the left/right passageway is also displayed. Switching from first- to second-order schemes consistently reduces the pressure drop by about 4 Pa. RANS-I and LAM-I always predicts the highest pressure drop, followed by LES-I, RANS-II and LAM-II. LES-II, arguably the most reliable approach, provides the smallest pressure drop which is in agreement with HRLES-II. The left/right share of the flow is nearly unchanged, with about 58% passing through the left and 42% through the right, an asymmetry that Borojeni *et al.* (2020) show to be well within normal values, in light of anatomical asymmetries and the effects of the nasal cycle. Switching from LAM/RANS to LES for the same numerical scheme brings the pressure drop down by about 1.5–2.5 Pa.

Before examining how these global changes reflect locally in the mean velocity and pressure fields, the general features of the solution (which is qualitatively similar across all cases) are briefly described. The mean fields computed in the CT-LES-II case are taken as example and shown in figure 2.3. During the inspiration phase, the outer air is accelerated at the nostrils and then around the turbinates through the meati, with the velocity magnitude reaching up to 2–3 m/s. In the nasopharynx, the flow rotates downwards, but also produces a recirculation (visualized by the positive U_y component) at the posterior wall of the nasopharynx. The largest velocity values in the flow field reach up to 4–5 m/s: this happens in particular for the U_z component near the laryngeal stricture. Pressure, which is relative to the level $P = 0$ set at the outlet, undergoes the

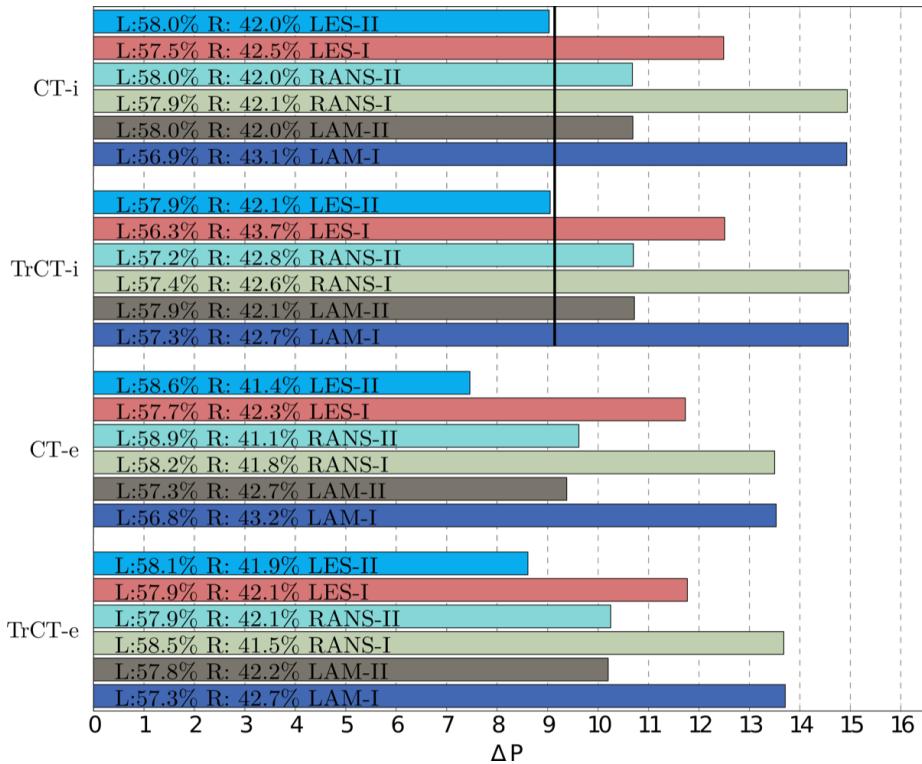


Figure 2.2: Mean pressure difference ΔP between inlet and outlet, for all the computed cases. The percentage share of the flow rate in the left (L) and right (R) fossa is also shown within each bar. For CT cases, the measurement is taken at the red line shown in figure 2.1. The vertical line is the reference pressure difference measured by HRLES-II-i.

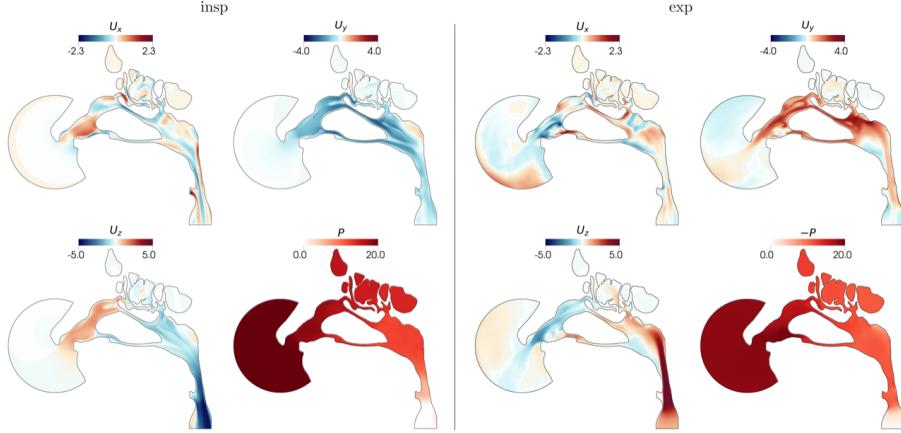


Figure 2.3: Mean velocity and pressure fields in sagittal view. Left: CT-LES-II-i; right: CT-LES-II-e.

largest drop under the epiglottis, in the lower region of the oropharynx.

During expiration, air flows through a contraction at the laryngopharynx and produces a strong vertical jet, which impacts on the rear portion of the nasopharynx, then turns horizontally to enter the fossae and eventually reaches the outer ambient. The largest component is again U_z , as shown in figure 2.3 (right), with a maximum of about 5 m/s. Pressure distribution qualitatively resembles the inspiration plot (except the direction of gradients), with the strongest drops at the larynx and in the meati.

Having illustrated the general features of the mean flow field, we can proceed now to illustrate the changes induced by the parameters of interest.

2.4.1 First- vs second-order schemes

Figure 2.4 plots the two largest Cartesian components of the difference velocity field $\mathbf{U}_{II} - \mathbf{U}_I$, with \mathbf{U}_I and \mathbf{U}_{II} being the time-averaged velocity fields computed with first- and second-order schemes, respectively.

In the RANS inspiration, differences up to 2.1 m/s are found. In the coronal view, peak differences reside in the areas with the largest rate of flow, with maxima of 1.1 m/s in the left inferior meatus and the right part of the middle meatus. The sagittal view shows significant velocity differences over the whole domain, except the external spherical volume and the sinuses. For the corresponding expiration, the coronal view shows similar differences still located in the middle meatus; the sagittal view, instead, shows a remarkable difference of 4.3 m/s in the U_z component, located in the nasopharynx. A rather similar picture is shown by the LES results, with comparable or even larger changes. To appreciate these differences, we observe that the bulk (area-averaged) velocity computed at the nostrils is 0.96 m/s.

Figure 2.5 focuses on the largest changes, occurring in the laryngeal jet, and compares its spatial structure in expiration for numerical schemes of different accuracy. (Only LES is shown, RANS is similar.) The laryngeal jet is substantially different: the

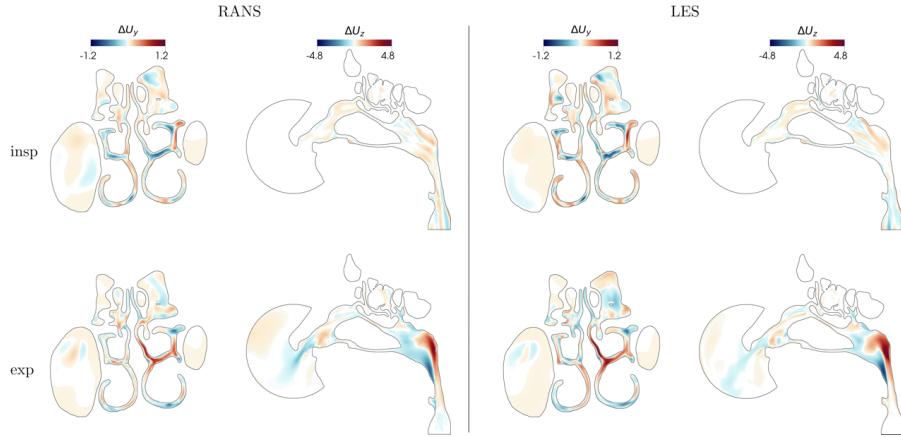


Figure 2.4: Differential velocity field $U_{II} - U_I$: RANS (left) and LES (right) for the CT anatomy.

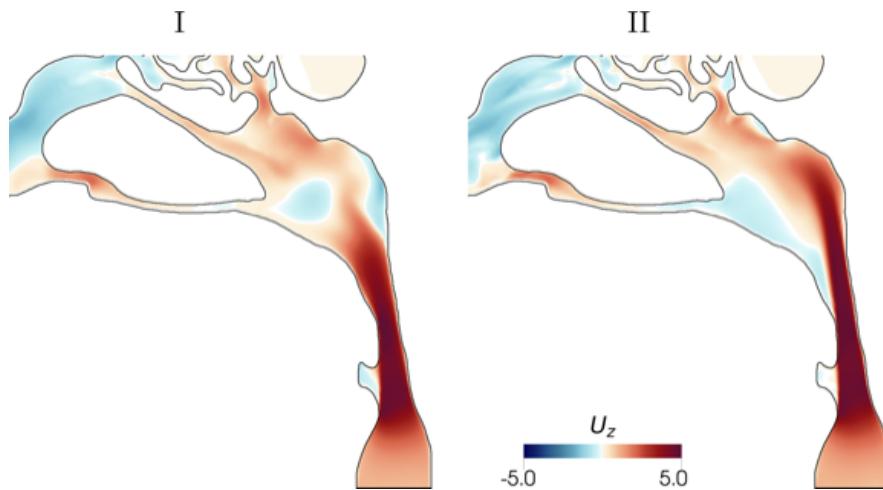


Figure 2.5: Sagittal view of CT-LES-e: U_z computed with first-order (left) and second-order (right) schemes.

lower-accuracy case shows a rather short jet that ends within the nasopharynx, whereas the higher-accuracy case presents a longer, more coherent jet that crosses the entire pharynx and impacts on the posterior wall.

2.4.2 RANS vs LES

RANS and LES results are compared via the difference of their mean velocity fields, i.e. $\mathbf{U}_{LES} - \mathbf{U}_{RANS}$. Since these differences are found to be rather independent from the numerical scheme, only cases computed at second-order accuracy are shown in figure 2.6. The horizontal component ΔU_y reaches up to 2.2 m/s in the area of the nasopharynx. In inspiration, differences are related to the shear layers detaching from the vestibular region; in expiration, differences extend to the meati. Especially during expiration, significant differences are observed in the vestibular area of the nose, of the order of 2 m/s for both velocity components.

Significant differences are also expected in the correct representation of turbulence, and in particular the field of turbulent kinetic energy k , which is entirely modelled by RANS and computed by LES. Figure 2.7 confirms that k largely differs between RANS and LES.

2.5 Discussion

The present results describe how the discretization scheme affects the CFD-computed airflow in the human nose, both globally and locally, and compares this effect to the modeling approach and to the type of CT scan.

The global effect has been quantified by measuring the pressure drop for a given flow rate. From figure 2.2, it appears that the formal order of accuracy of the discretization schemes plays a crucial role, independently from the flow model. On a given mesh, low-order numerical schemes are found to predict larger pressure drops, consistently with their more dissipative nature. Similarly, for a given numerical scheme, RANS predicts a larger pressure drop than LES, again because of the dissipative nature of the RANS turbulence models based on the concept of turbulent viscosity (Pope, 2000). The changes are substantial: at this flow rate, the pressure drops computed by a first-order RANS and by a second-order LES differ up to 6 Pa , which in the TrCT case is a difference of more than 60%. Higher-order schemes imply a larger computational cost, but marginally so: we have measured a modest 15% increase in CPU time for all the considered flow models. The large effect of the numerical scheme of choice is an important element to consider in the ongoing discussion, see e.g. Cherobin *et al.* (2020) and Berger *et al.* (2021), whether nasal resistance computed via CFD agrees with nasal resistance clinically measured with a rhinomanometer, and clearly advocates the specification of the employed numerical schemes in papers dealing with airflow in the human nose: the overestimate of the pressure drop by lower-accuracy methods would further increase the gap between the two measuring techniques, while the scatter among CFD datapoints would be most certainly reduced. Unfortunately, however, in the current literature this essential information is not reported very often.

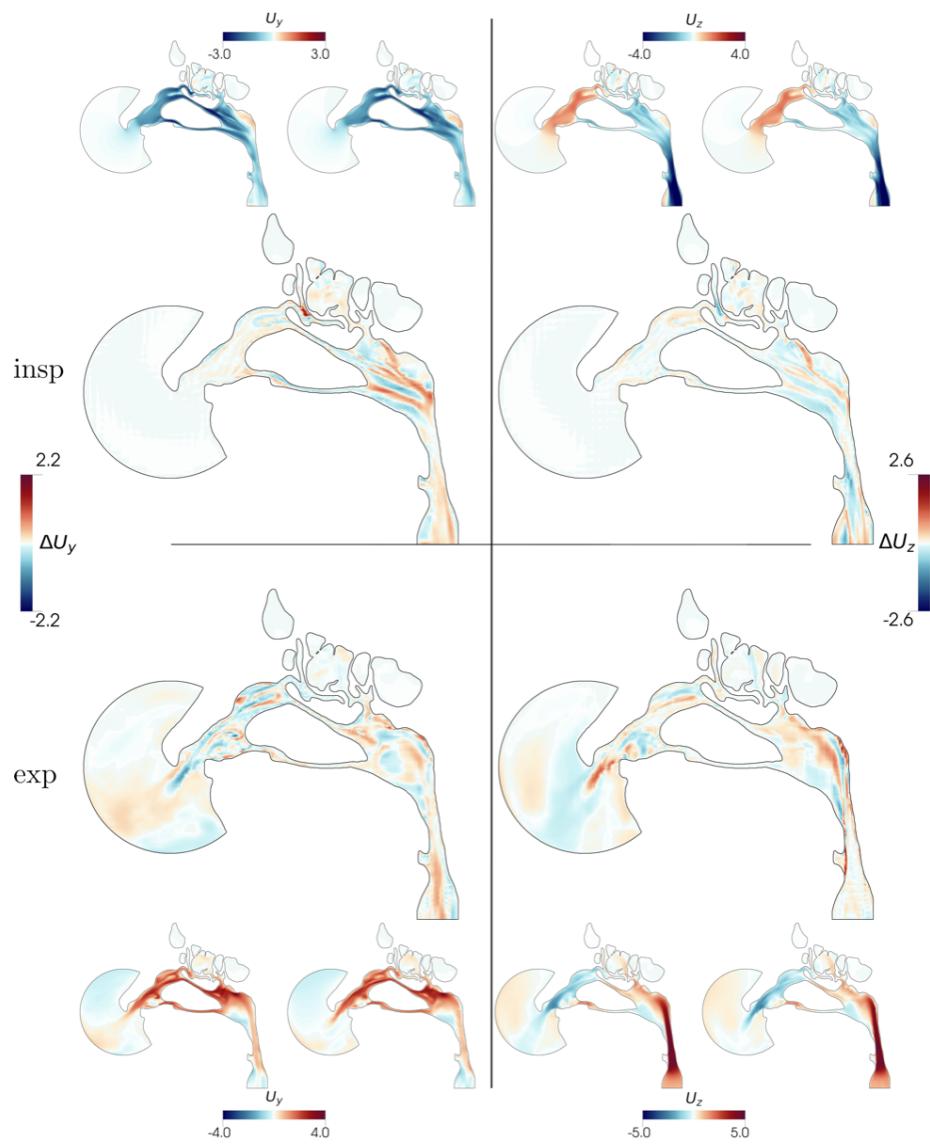


Figure 2.6: Differential velocity field $\mathbf{U}_{LES} - \mathbf{U}_{RANS}$, for CT-II cases. The left and right columns describe the U_y and U_z velocity components respectively, while the top and bottom rows concern inspiration and expiration. For each panel, the largest figure plots the difference field, while the smallest panels plot the LES (left) and RANS (right) fields from which the difference field is generated.

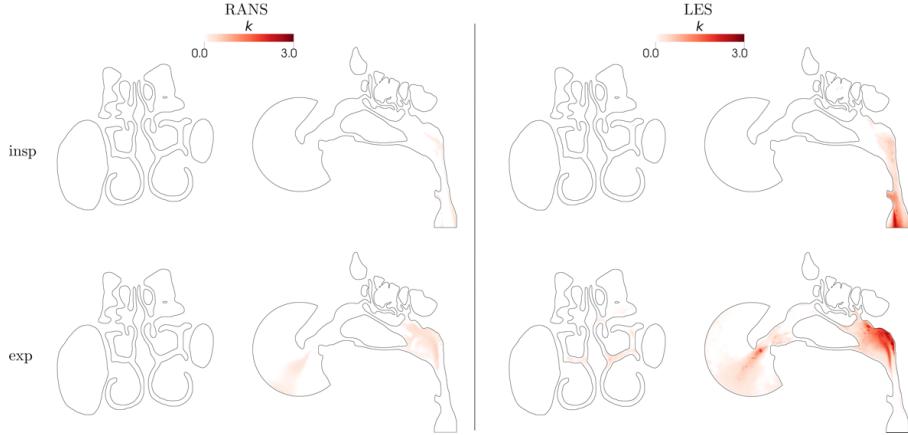


Figure 2.7: Field of turbulent kinetic energy k as computed from CT-RANS-II (left) and CT-LES-II (right).

Global differences arise as the integrated effect of a number of localized changes in the pressure and velocity fields. First-order numerical schemes misrepresent important parts of the flow physics, by for example failing to correctly capture the free shear layers in the nasopharynx during inspiration, or the massive laryngeal jet that develops during expiration. Use of CFD for detailed surgery planning would certainly benefit from a reliable representation of the whole flow physics, and thus mandates close attention to the numerical schemes employed in the CFD solution.

Flow modelling has been discussed multiple times in the past, and it comes at no surprise that laminar/RANS and LES outcomes are quite different, in terms of both pressure and velocity fields. Pressure differences indicate that RANS overestimate pressure drop by 2–4 Pa , independently from the numerical schemes; velocity differences are more delicate to interpret. The most affected flow region seems to be where free shear layers develop (the nasopharynx, and the vestibular area during expiration). Laminar/RANS modelling, although perhaps acceptable for normal sino-nasal anatomies like the present one, might become questionable once anatomic anomalies are present and disturb the flow field, leading to a more complex flow even in the relatively quiescent yet surgically delicate region of the nasal meati. Obviously, this has to be considered jointly with the different computational cost: speaking of CPU time alone, the typical mesh sizes used here lead to LES being approximately 60 times more expensive than RANS. Significant differences have been also found in the correct representation of turbulence, e.g. the turbulent kinetic energy field shown in figure 2.7, thus reinforcing the case for the inadequacy of RANS modelling whenever anatomic anomalies induce significant localized flow unsteadiness.

This study has also considered the effect of a computational domain truncated well above the larynx, as it would happen when cone-beam CT scans are used. Changing the position of the lower boundary has little influence when inspiration is computed, but expiration is much more affected: the lack of the laryngeal restriction makes the laryngeal jet impossible to predict correctly. Given the undeniable convenience of

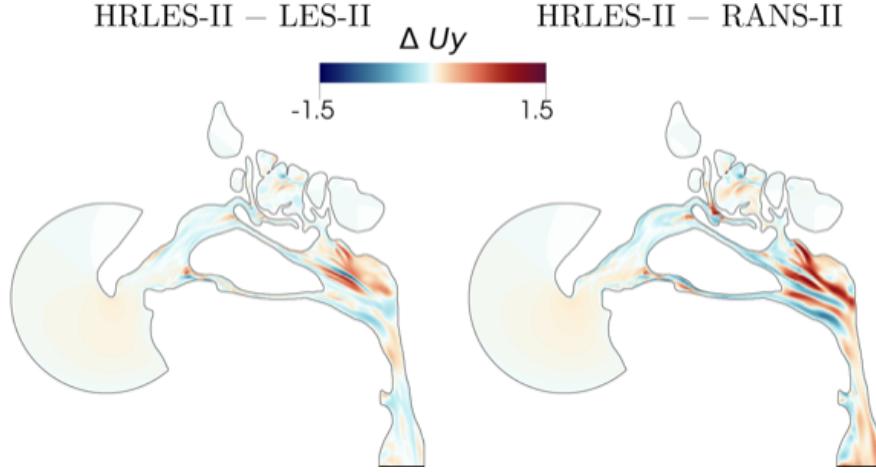


Figure 2.8: Differential velocity field (sagittal component) HRLES-II - LES-II (left) and HRLES-II – RANS-II (right).

cone-beam scans, and the importance of imparting lower radiation doses to the patient, we envisage the need for a suitable inlet boundary condition for expiration to implicitly compensate for the missing part of the domain.

Discussing differences between velocity fields would be incomplete without recalling that alternate ways exist to compare two vector fields. For example, one should be aware that looking at the Cartesian components of the velocity difference vector might misrepresent changes that would appear under different light if e.g. the modulus of the difference is considered. Also, differences should be evaluated by bearing in mind the intensity of the local mean value.

Finally, so far we have discussed "differences" with the implicit assumption that LES-II naturally represents the most accurate approach in terms of both turbulence modelling versus RANS-II and numerics versus LES-I. However, LES-II results themselves are affected by modelling and discretization error: they would become error-free only on a very fine mesh. It is thus instructive to compare LES-II with the result of HRLES-II, where the larger mesh with 50 millions cells (more than 3 times the cells of LES-II) makes it approach the DNS limit. The global result of HRLES-II was already plotted as inspiration reference in figure 2.2; now figure 2.8 clearly shows how LES-II is nearer than RANS-II to the reference, with residual errors that decrease both in spatial extension and absolute value as the spatial resolution increases and the LES modelling improves accordingly.

2.6 Conclusion

The impact of key methodological choices in the numerical simulation of the airflow in the human nasal cavities has been quantitatively assessed, by comparing the importance of the numerical scheme accuracy to that of the flow modelling. Within a well-defined

comparison framework, the output of 24 simulations has been evaluated at both the global and local level in terms of pressure losses, mean velocity and pressure fields. The choice of a laminar/RANS/LES modelling approach is very important, especially in such flows that are often laminar, albeit vortical, chaotic and three-dimensional. However, we have ascertained that the numerical scheme is even more important, leading to differences to more than 50% in global indicators (e.g. nasal resistance), and to local differences that can be extremely significant. Finally, we have also indirectly assessed that cone-beam CT scans can be used proficiently, at long as inspiration is considered; in expiration, however, the proximity of the inflow to the nasopharynx is responsible for a significant misrepresentation of the laryngeal jet that propagates up to the nostrils. Overall, the study confirms that high-fidelity and time-resolved LES/DNS computations (Calmet *et al.*, 2021) are probably necessary for a reliable simulation of the full breathing cycle at intermediate intensity, and advocates once again for high-quality numerical and experimental benchmarks, placed on the public domain and fully reproducible, to arrive at a rigorous assessment of the adequacy of the modelling choices in the CFD of the nasal airflow.

2.7 Appendix A: The numerical approach

2.7.1 Boundary conditions

Regardless of the flow modelling approach, at the boundary made by solid, rigid walls the velocity vector and the normal component of the pressure gradient are set to zero. During inspiration, the inlet is the surface of the external sphere surrounding the nose tip, and the outlet is at the throat. The required flow rate of 280 ml/s is prescribed at the inlet via an inward velocity field that is computed to be locally normal to the surface, and adjusted to provide the prescribed integral value. This value of the flow rate is representative of slow to mild breathing (Wang *et al.*, 2012; Covello *et al.*, 2018). Pressure at the inlet is given a zero-gradient condition. At the outlet, the gradient of the velocity is set to zero, and the total pressure is set to a given (zero) reference value. During expiration, the boundary conditions are reversed: the throat becomes the inlet, where the flow rate of 280 ml/s is prescribed via the normal velocity field adjusted to provide the specified integral, and pressure has zero gradient. At the outlet, i.e. the external sphere, a zero-gradient condition for velocity is accompanied by a reference zero value for the total pressure.

2.7.2 RANS model and procedures

The model of choice, already used in the past for such studies (see e.g. Liu *et al.*, 2007; Li *et al.*, 2017), is the $k-\omega$ -SST model: besides the RANS equations, it solves two additional partial differential equations for the turbulent kinetic energy k and the turbulent frequency ω , and uses the Boussinesq hypothesis to close the RANS equations via a turbulent viscosity ν_t . The turbulent frequency at the wall provides a wall constraint on the specific dissipation rate, and is imposed to match the condition by Menter *et al.* (2003). At the inlet, the flow is considered nearly non-turbulent, and

the turbulent frequency is thus set to an arbitrarily small value (unitary in the present work). At the outlet its gradient is null. The turbulent kinetic energy at the wall is zero by definition. At the inlet it is set to $k = \frac{1}{2}(I\mathbf{U})^2$ where \mathbf{U} is the local (extremely small) mean velocity and I is the turbulent intensity, which is set to 2%, resulting in an almost non-turbulent inlet. At the outlet the gradient of k is zero. The $k-\omega$ -SST model is known for its ability to provide results that are decoupled from the (necessarily approximate) far-field values of the boundary conditions.

The model coefficients are standard and taken from Menter *et al.* (2003):

$$\begin{aligned}\sigma_{k_1} &= 0.85; & \sigma_{k_2} &= 1.0; & \sigma_{\omega_1} &= 0.5; \\ \sigma_{\omega_2} &= 0.856; & \alpha_1 &= \frac{5}{9}; & \alpha_2 &= 0.44; \\ \beta_1 &= \frac{3}{40}; & \beta_2 &= 0.0828; & \beta^* &= 0.09.\end{aligned}$$

The iterative solution of the RANS equations is terminated when the residuals fall below set limits. The typical limit values for RANS-I are 10^{-9} for the residuals of the three components of velocity and pressure, and 10^{-6} for the turbulent quantities. These values are set higher for RANS-II and become 10^{-5} for every quantity. This level is considered sufficient to yield fully converged results (Zhang & Kleinstreuer, 2011).

2.7.3 LES model and procedures

The model of choice is the Wall-Adapting Local Eddy-viscosity (WALE) model, already employed in similar studies (Li *et al.*, 2017). It is an algebraic (hence not requiring additional boundary conditions) sub-grid scale model based on the work by Nicoud & Ducros (1999). The model considers both local strain and rotation rates, and has the advantage of being invariant to translation and rotation of the reference system; moreover, it requires local information only, which makes it particularly useful for complex geometries such as the present one. Compared to the widely used Smagorinsky model (Smagorinsky, 1963), by design WALE provides an eddy viscosity that approaches zero at a solid wall with the correct rate.

The time-dependent solution of the LES equations is stopped when the length of the statistical sample is enough to provide satisfactory estimates for the first and second statistical moments of the flow variables. We know by previous experience (Covello *et al.*, 2018) that, at this respiratory rate, a time integration of 0.6 seconds after properly discarding the initial transient is enough to provide a reliable mean field. In fact, a test carried out for CT-LES-II-i by adding 0.20 s of simulation time has led to change in the total pressure drop by 0.017% only.

The time-dependent LES cases are advanced in time with a second-order BDF (backward differentiation formula) scheme; the time step is dynamically varied so that the Courant–Friedrichs–Lowy number remains below one. For example, for CT-LES-II-i the average value of the timestep is 1.038×10^{-5} seconds, with maximum and minimum during the simulation being 1.226×10^{-5} and 7.809×10^{-6} seconds. The obtained average timestep is in line with the values used in literature (Li *et al.*, 2017) for comparable LES simulations.

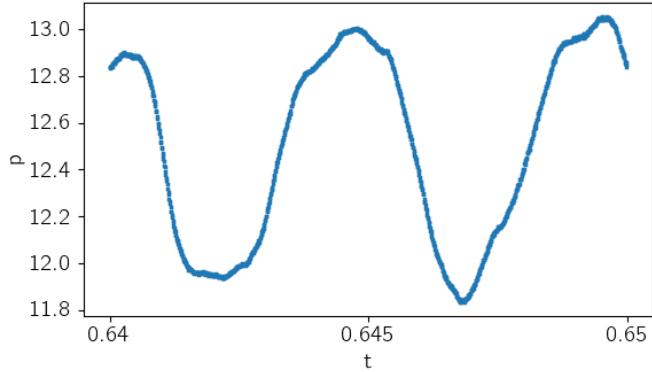


Figure 2.9: Time history (case CT-LES-II-i) over 1200 time steps of the instantaneous pressure signal measured by a probe positioned in the rhinopharynx in the point with coordinates $x = 0.001$, $y = 0.05$ and $z = 0.0145$. Each dot represents one time step.

A reliable estimate of the frequency content of the temporal signal cannot be obtained with classical turbulence arguments (e.g. comparison with the local Kolmogorov time scale), owing to the significant difference between this flow and the homogeneous isotropic turbulent flow where such arguments apply. Although the only proper assessment would derive from a frequency analysis of the temporal signal at every spatial location (a non-trivial endeavour for a non-periodic signal known over a finite time horizon), visual inspection of the time history of the local turbulent signal would provide a good proxy to establish whether the time step is small enough. Figure 2.9 plots the time history of pressure measured at one spatial location in the rhinopharynx, where most unsteadiness takes place. One immediately obtains the visual evidence that the temporal signal is extremely well sampled, reinforcing the concept (Choi & Moin, 1994; Bernardini *et al.*, 2013) that in DNS (or high-resolution LES) of wall-bounded flows the time step is usually dictated by the stability margin of the numerical scheme, well below the temporal scales of the turbulent flow.

2.8 Appendix B: Geometry and mesh quality

The CT scan is obtained from an adult male with healthy sino-nasal anatomy. The scan has an axial resolution of 0.6mm , with coronal and sagittal spacing of 0.5mm , and is representative of current clinical standards. The CT scan extends down to the larynx and includes the whole volume of the sinuses. The image set from the CT scan is processed via 3D-Slicer (Fedorov *et al.*, 2012), an open-source software for the analysis

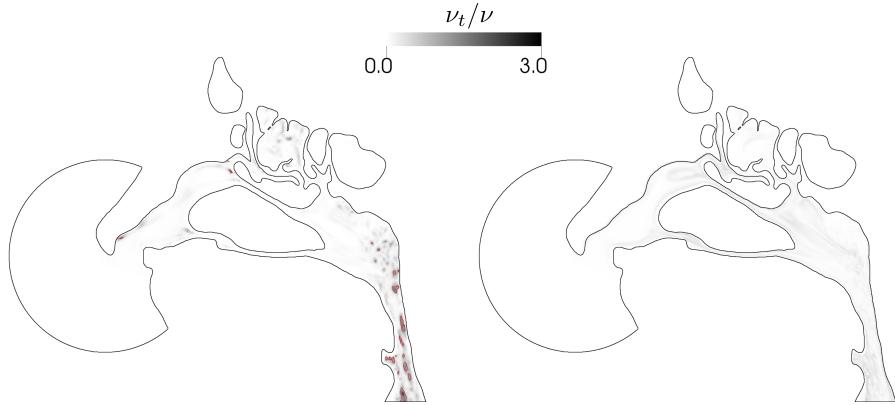


Figure 2.10: Ratio between turbulent viscosity ν_t and molecular viscosity ν for an instantaneous flow field of the case CT-LES-II-i at 15 million cells (left) and HRLES-II-i at 50 million cells (right). The red contour line marks the unitary ratio.

and visualization of medical images. Without any other manipulation or smoothing, 3D-Slicer is used, via image segmentation at uniform threshold of -220 Hounsfield units and volume reconstruction, to convert the CT images into a three-dimensional geometry, that is eventually exported as a STereoLithography (STL) file.

All the computational volume meshes are generated within the OpenFOAM toolbox (Weller *et al.*, 1998): in particular the *snappyHexMesh* tool is used to convert the STL geometry into a computational mesh. First a volume surrounding the natomy is filled with hexahedra, then a refinement is applied to the surface of the geometry such that the cells are decomposed twice in each direction (castellation phase). In this way, a background cells that intersect the STL surface is split in 64 smaller cells. The vertices of the castellated mesh are then displaced to snap onto the actual boundary (snap phase). No additional prismatic layers are added, to avoid deterioration of the mesh quality. The mesh is fine enough at the wall, with mean values of about $y^+ = 0.5$ for LES simulations and of around $y^+ = 1.09$ for the laminar/RANS simulations (with y denoting, as it is customary, the wall-normal distance, and the plus superscript indicating viscous or wall units). The mesh refinement near the wall can be seen in Fig. 1b of the main manuscript. The dictionary controlling the mesh quality is the same for LES and laminar/RANS meshes: the finer LES meshes are obtained by starting from a finer background mesh.

In terms of mesh quality parameters, for the RANS simulation (total cells about 3.2 millions), the average non-orthogonality is 11.86, well below the limit margin of 70, above which special treatment would be required. The maximum skewness is 2.46, which is above the warning threshold of 0.9 but fully acceptable in a geometry of this complexity.

For the LES mesh (total cells about 15 millions) the average non-orthogonality is 11.71, and the maximum cell skewness is 2.49, i.e. nearly identical to the RANS mesh. This is because the mesh is generated with the same target quality parameters. Figure 2.10 shows that the LES mesh is quite fine for the present problem, and yields a ratio between the turbulent viscosity ν_t and the molecular viscosity ν that, in one randomly

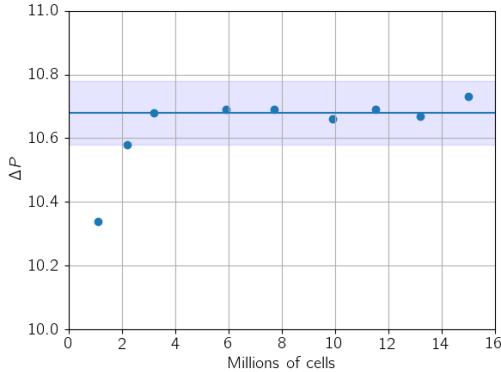


Figure 2.11: Grid convergence study for the case CT-RANS-II-i. Pressure drop (measured between external ambient and the lower end indicated by the red line in figure 1a of the main text) is plotted versus mesh size. The horizontal band shows the mean value (excluding the first two leftmost points) of the pressure drop, and the $\pm 1\%$ interval.

chosen instantaneous field, remains below 2.5. Thus the LES is certainly well resolved, especially near the wall, with the model yielding a limited contribution. (The same ratio remains well below unity everywhere for the larger HRLES mesh, which could thus be considered a fully resolved DNS.)

While the fineness of the LES mesh has been evaluated above in the context of figure 2.10, a grid refinement study has been carried out for the RANS mesh, to confirm that its size, which is typical of comparable literature studies employing RANS, is indeed adequate. The refinement study, whose results are shown in figure 2.11, proves that the considered mesh at 3.2 millions cells is already fine enough to properly capture the pressure drop between the external ambient and the outlet. Too small a mesh (below two millions cells) is clearly insufficient to describe the geometry properly, but all the finer meshes show that results are well within a $\pm 1\%$ uncertainty band.

2.9 Appendix C: CT vs TrCT

The effect of the position of the lower boundary of the computational domain is a parameter that in the main text has been discussed only in terms of global changes of the pressure drop. Here we describe in figure 2.12 the local differences, by plotting the field $\mathbf{U}_{CT} - \mathbf{U}_{TrCT}$. Since this set of differences is mostly insensitive to the turbulence modeling approach, only the LES-II simulations are shown.

The position of the lower boundary shows little or no effect when inspiration is considered, with peak differences as small as 0.2 m/s . However, during expiration large differences, of about 4.3 m/s , are observed near the inlet, because the structure of the laryngeal jet is not reproduced properly by the smaller-domain simulation. Importantly, these differences decrease with the distance from the inlet but persist quite far from it,

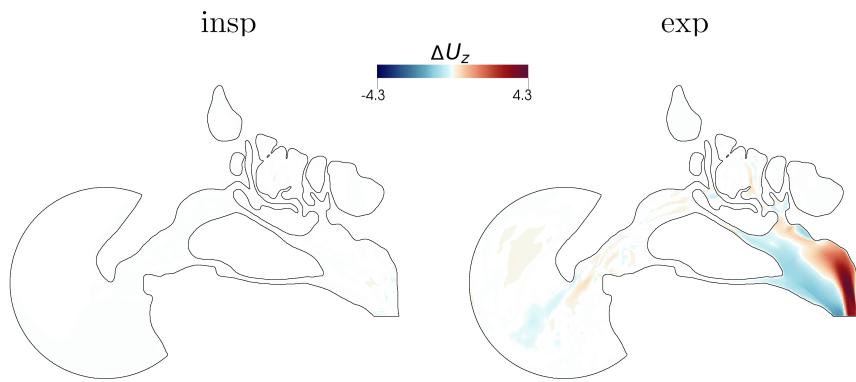


Figure 2.12: Differential velocity field $\mathbf{U}_{CT} - \mathbf{U}_{TrCT}$ (vertical component), for LES-II.

and non-negligible effects are discerned even at the nostrils.

Chapter 3

An adjoint-based approach for the surgical correction of nasal septal deviations

3.1 Abstract

Deviations of the septal wall are widespread anatomic anomalies of the human nose; they vary significantly in shape and location, and often cause the obstruction of the nasal airways. When severe, septal deviations need to be surgically corrected by ear-nose-throat (ENT) specialists. Septoplasty, however, has a low success rate, owing to the lack of suitable standardized clinical tools for assessing type and severity of obstructions, and for surgery planning. Moreover, the full restoration of the septal wall is often either impossible or not necessary. This paper introduces a procedure, based on advanced patient-specific Computational Fluid Dynamics (CFD) simulations, to support ENT surgeons in septoplasty planning. The method hinges upon the theory of adjoint-based optimization; by minimizing a cost function that indirectly accounts for viscous losses, a sensitivity map is computed on the mucosal wall, and the surgeon is provided with a simple quantification of how much tissue removal at each location would contribute to easing the obstruction. The optimization procedure is applied to three representative nasal anatomies, reconstructed from CT scans of patients affected by complex septal deviations. The computed sensitivity consistently identifies all the anomalies correctly. Virtual surgery, i.e. morphing of the anatomies according to the computed sensitivity, proves that the characteristics of the nasal airflow are improved

3.2 Introduction

Nasal Airway Obstructions (NAO) are one of the main medical conditions for which patients consult ear-nose-throat (ENT) doctors (Chandra *et al.*, 2009). Indeed, more than one third of the world's population is affected by NAO (Jessen & Malrn, 1997;

Li *et al.*, 2020), with a reduced nasal airflow that impacts the quality of life (Rhee *et al.*, 2003; Udaka *et al.*, 2006). Among the main aetiological factors causing NAO, septal deviations have been found to be extremely common, with a prevalence of 76% (Clark *et al.*, 2018). Under such circumstances, ENT specialists often resort to surgical corrections via septoplasty to restore a correct respiration. Even though severe complications are rare and arise in the 3% of cases only (Dabrowska-Bien *et al.*, 2018), once the actual post-surgical benefits are considered, the success rate of septoplasty falls below 50% (Tsang *et al.*, 2018) or even less (Sundh & Sunnergren, 2015), resulting in a relevant social and financial burden for the healthcare system. This state of affairs can be ascribed to the lack of standardized and reliable clinical tools to assess type, severity and consequences of obstructions, which would provide surgeons with essential information when deciding on the most appropriate surgical action (Dinis & Haider, 2002; Roblin & Eccles, 2002).

Computational Fluid Dynamics (CFD) is nowadays recognized as a valuable tool to study the nasal airflow, to quantify its characteristics and to relate them with the patient's anatomy. Techniques range from relatively inexpensive simulations that take advantage of turbulence modelling (Li *et al.*, 2017) to larger-scale and higher-fidelity calculations (Calmet *et al.*, 2016). Cases with NAO are also considered (Leong *et al.*, 2010). For instance, Refs. Chen *et al.* (2009), Liu *et al.* (2012), and Janović *et al.* (2020) used CFD to evaluate changes in the respiratory pattern when septal deviations of different type and severity occur. Refs. Malik *et al.* (2021) and Radulesco *et al.* (2019), instead, performed simulations of the nasal airflow to analyze what quantities, among those measured with existing clinical tools, are the most representative of the perception of patients. Several studies directly examined the effect of septoplasties. In particular, Campbell *et al.* (2021) focused on anterior septoplasty, and studied the CFD-computed nasal airflow of ten healthy anatomies modified to introduce NAO, to understand how surgically widening the region with minimal cross-sectional area correlates with the actual benefits of surgeries. Ramanathan *et al.* (2021), on the basis of CFD results for twelve pre- and post-operative CT scans, suggested the main regions of obstruction to be often identified by locally high pressure, velocity and shear stress. The available studies are unable to indicate how to implement the request to surgically increase the area of a certain cross-section.

All the CFD works addressing septoplasty planning rely, to a varying extent, on a subjective interpretation of the CFD results. While a strong local minimum of the cross-sectional area is certainly bound to determine an obstruction, associating NAO (and the corresponding strategy for its surgical correction) only to the minimal cross-sectional area, or to the local extremum of CFD-computed quantities, may be too simplistic, as it neglects the anatomical complexity of the cross-section itself and the non-locality of the fluid flow equations: the flow field observed in one specific location is affected by anatomy elsewhere. Moreover, the value of the minimum cross-sectional area does not inform the surgeon on the directional changes of the velocity vector induced by enlarging different portions of the minimal-area section. The same enlargement might be achieved by acting on different portions of the minimal-area section, but the effects on the airflow would be different.

The main goal of this work is to introduce a novel CFD-based procedure for patient-specific septoplasty planning. The approach includes elements from the adjoint-based

theory of shape optimization, that was introduced in fluid mechanics almost 50 years ago, see e.g. Ref.Pironneau (1974), and is becoming progressively popular in CFD. The procedure is made by two steps. In the first step, the airflow within the nasal cavities is simulated via conventional CFD. The second step uses the computed flow field to solve the optimization problem, by finding the minimum of a cost function. The end result is a sensitivity map, i.e. a distribution on the whole nasal surface of a scalar quantity that, at each point, quantifies if and to what extent the displacement of said point, as a consequence of a surgical action, is favorable or counterproductive in terms of minimization of the cost function. The adjoint-based method is exemplified via application to three nasal anatomies of patients affected by complex septal deviations. An in-depth clinical analysis by ENT surgeons confirms the surgical feasibility of the suggestions. Virtual surgery is then carried out, and confirms that the computed sensitivity leads to improved flow partitioning and nasal resistance.

The paper is structured as follows. A brief and non-technical summary of the adjoint-based optimization is given in section 3.3, followed by section 3.4 which contains a description of the numerical methods employed in the present work, including the adaptation of the theory to the specific problem. Results for the three patients are presented in section 3.5. An in-depth discussion, which includes comments on the essential aspects of the procedure, is provided in section 3.6, and section 3.7 is devoted to conclusions.

3.3 Adjoint-based optimization

Central to the present work is the adjoint optimization technique used for computing the surface sensitivity in a shape optimization problem. The main theoretical and technical aspects behind the adjoint formulation are briefly summarized below, while the tailoring of the method to septal deviations is described in section 3.4. The interested reader is referred to recent review papers (Dilgen *et al.*, 2018; Alexandersen & Andreasen, 2020) for additional information on adjoint optimization.

In general, in shape optimization one seeks the optimal shape of a two- or three-dimensional object (defined through a cloud of points) that minimizes a certain (known) cost function while satisfying a set of constraints. In the CFD context, this can be expressed as:

$$\begin{aligned} &\text{minimize } f = f(\mathbf{U}, \beta) \\ &\text{subject to } R(\mathbf{U}, \beta) = 0 \end{aligned} \tag{3.1}$$

where f is the cost function to be minimized, \mathbf{U} the set of flow variables (e.g. velocity and pressure), β the set of control variables (e.g. the wall-normal displacements of the points of the surface) which define a change in shape, and R a set of constraints. The constraints R indicate that the flow variables are not arbitrary, but must obey the differential equations governing the fluid flow. In this work, the steady incompressible Navier–Stokes equations will be considered:

$$\begin{cases} \nabla \cdot \mathbf{u} = 0 \\ (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u} \end{cases} \tag{3.2}$$

where \mathbf{u} is the velocity vector, p the pressure, and ν the kinematic viscosity of the fluid. The optimization computes the sensitivity derivatives, i.e. the gradients of the cost function with respect to the control variables. The sensitivity describes how β (a change in shape) affects f (the cost function). The visualization of β , defined on the boundary only, highlights at a glance where altering the original shape is most effective at minimizing the cost function.

Adjoint optimization excels when a simple cost function is available, and the number of control variables is large (as in the present case). Indeed, the sensitivity is obtained at a computational cost that is independent upon the number of control variables, as the procedure requires solving two systems of Partial Differential Equations (PDEs), whose computational cost is independent upon the number of elements of β . The first system is made by the usual governing equations of fluid dynamics, i.e. (3.2), whereas the second one, approximately of the same cost, includes the so-called adjoint equations, which are derived from the Navier–Stokes equations.

In practice, two approaches can be followed to derive the adjoint equations: a discrete one, known as "discretize then derive" and a continuous one, known as "derive then discretize" (Bewley, 2001). In this contribution, a continuous formulation specialized for internal flows and described by Othmer (2008) is used as the starting point for the main analytical derivations. As often assumed in shape optimization problems, in this work the cost function includes contributions only from the boundary Γ of the flow domain Ω .

The optimization problem (3.1) is first reformulated in terms of Lagrange multipliers to account for the set of constraints R . The adjoint variables are introduced: the adjoint pressure q (a scalar quantity) and the adjoint velocity \mathbf{v} (a vector). They have the same physical dimensions of their physical counterparts p and \mathbf{u} , but carry a different meaning. The resulting Lagrangian function, defined over the entire domain Ω reads:

$$L = f + \int (\mathbf{v}, q) R \, d\Omega. \quad (3.3)$$

Once L is defined, the sensitivities can be computed by starting from the total variation of L , written as:

$$\delta L = \delta_\beta L + \delta_{\mathbf{u}} L + \delta_p L \quad (3.4)$$

by separating the variations δ_β , $\delta_{\mathbf{u}}$ and δ_p .

The adjoint variables, which can take arbitrary values, are chosen such that the sum of the variations of L with respect to the state variables is null:

$$\delta_{\mathbf{u}} L + \delta_p L = 0 \quad (3.5)$$

At this point, the influence of β on the Navier–Stokes equations can be evaluated. Indeed, the previous equality leaves $\delta_\beta L$ as the only contribution to the variation of the Lagrangian function. The final expression for the sensitivities is obtained as Soto & Lohner (2004):

$$\frac{\partial f}{\partial \beta_i} \propto \mathbf{v}^i \cdot \mathbf{u}^i \quad (3.6)$$

where β_i is the displacement, in the direction normal to the surface, of every point i , and \mathbf{u} and \mathbf{v} are the direct/adjoint velocities of the fluid. It becomes evident that the

surface sensitivity embeds information from the velocity field \mathbf{u} (obtained via the usual CFD method), and from the adjoint field \mathbf{v} , computed by solving the adjoint system of PDEs.

The adjoint system is derived from Equation (3.5) by taking the required derivatives. For the present formulation, the resulting adjoint equations are:

$$\begin{cases} \nabla \cdot \mathbf{v} = 0 \\ -\nabla \mathbf{v} \cdot \mathbf{u} - (\mathbf{u} \cdot \nabla) \mathbf{v} = -\nabla q + \nu \nabla^2 \mathbf{u} \end{cases} \quad (3.7)$$

in which the adjoint variables \mathbf{v} and q are the unknown to be computed, whereas the flow variables \mathbf{u} and p are regarded as known, as they have been previously computed by solving the direct system (3.2). The adjoint equations are linear.

Since the (still unspecified) cost function is assumed to contain contributions from the domain boundary only, the adjoint system (3.7) does not depend on the boundary shape, and enjoys general validity. Details on its derivation can be found in Ref. Othmer (2008).

To solve the adjoint system, boundary conditions for the adjoint variables need to be specified. They are obtained by considering the contributions on the boundary Γ that are present in Equation (3.5) when the derivatives of the cost function are explicitly computed. Their definition depends on the specific cost function.

3.4 Methods

3.4.1 Anatomies and discretization

In this study, three CT scans are selected from a pool of cases with complex septal deviations where the most appropriate surgical approach is not self-evident. They are shown in figure 3.1. The scans, provided by the San Raffaele University Hospital, have been obtained with a standard radiological protocol through a CT scanner with an acquisition matrix of 512^2 pixels. The scanner is a GE Revolution Evo, with 128 slices. For the three patients, referred to as P1, P2 and P3, the spatial resolution of the scans in the sagittal-coronal plane are $0.39mm \times 0.39mm$, $0.31mm \times 0.31mm$ and $0.46mm \times 0.46mm$, and the gap between consecutive axial slices is $0.925mm$, $0.925mm$ and $0.4mm$

Patient P1, a 44-year-old caucasian male, presents a complex pattern of septal deviations, with a major left deviation of the quadrangular cartilage, and a small posterior deviation of the vomer bone. (As in the rest of the paper, the description adopts the point of view of the patient: hence, left/right should be intended as the patient's left/right.)It also presents a slight right deviation of the antero-superior portion of the nasal septum, at the articulation between the quadrangular cartilage and the perpendicular plate of the ethmoid bone. Patient P2, a 30-year-old caucasian male, has the anterior portion of the quadrangular cartilage deviated towards the left nasal fossa with a partial occlusion. Posteriorly, it presents an important right bone spur, bridging the middle meatus. Further posteriorly, another left bone spur reaches the posterior portion of the middle turbinate. Patient P3, a 40 year-old caucasian male, presents a nasal valve collapse (more evident in the left nostril) and a left septal deviation. The deviation involves

the quadrangular cartilage, which is dislocated laterally and reaches the left inferior turbinate. Posteriorly it presents a left condro-vomerian spur that reaches the middle meatus. There is also a minor bony spur in the right posterior nasal fossa. Patient P3 probably underwent turbinoplasty before the CT scan, as the mucosa of the right inferior turbinate is less hypertrophic than normally expected.

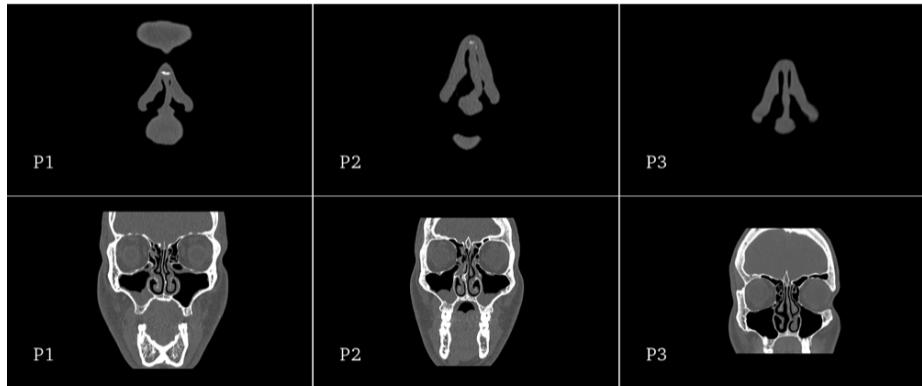


Figure 3.1: Coronal sections of the three patients, visualized in correspondence of the most significant septal deviations in the anterior (top) and posterior (bottom) regions.

Scans are segmented with the free and open-source software 3D Slicer (Fedorov *et al.*, 2012). The three-dimensional reconstruction of the boundaries of the nasal airways is obtained by applying a segmentation threshold of -475 Hounsfield units, in accordance with results by Nakano *et al.* (2013) and Zwicker *et al.* (2018). Figure 3.2 shows the final model for patient P2: the nasal airways up to the initial part of the nasopharynx and all paranasal sinuses are included. The reconstructed geometries are used as input to create a computational mesh suitable for finite-volumes discretization. In this process, a spherical volume, with a diameter of 70mm, is carefully placed around the nostrils to account for the external environment. As discussed in previous work

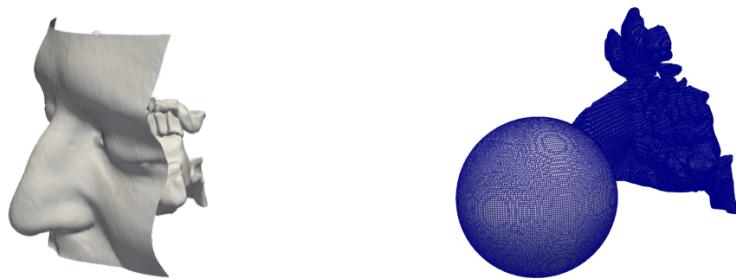


Figure 3.2: Three-dimensional geometric model for patient P2.

Figure 3.3: Computational mesh for patient P2.

	Number of cells	
	Volume Mesh	Solid Boundary
P1	5262788	950327
P2	5695738	1052087
P3	6623922	1178278

Table 3.1: Number of cells in the volumetric mesh for the three patients.

Covello *et al.* (2018), such spherical volume places the actual inflow (where boundary conditions will be applied) far enough from the critical region of the nostrils, while keeping the total volume (and thus the computational cost) under control.

Figure 3.3 shows the volume mesh obtained at the end of this procedure for P2. For a better description of the anatomy, finer cells are placed in correspondence of the nasal airways and of the paranasal sinuses, whereas a coarser grid is used for the inlet sphere. No layers are added, since the background mesh and the refining implicit in the adaptation process create fine enough cells near the boundaries. Table 3.1 reports, for each of the three patients, the number of cells for the entire volume mesh and for the boundary of the nasal airways alone. The volume mesh is made by a number of cells ranging from 5.3 to 6.6 millions, whereas about 1 million cells describe the solid boundaries, i.e. the nasal walls. The small differences between meshes for the three patients are due to the different dimensions of the anatomies, since an identical refinement strategy was used. Overall, the CT scans are of standard quality, and the size of the computational meshes is typical for comparable RANS studies of the nasal airflow (Inthavong *et al.*, 2019).

3.4.2 Direct RANS simulations

The mathematical flow model of choice is the steady incompressible RANS equations, obtained after time-averaging the Navier–Stokes equations. It represents the most common choice for such problem, and accounts for turbulence via a turbulence model. The steady Navier–Stokes equations (3.2) written for the mean fields are augmented with the divergence of the apparent stress tensor $\bar{\mathbf{u}}'\bar{\mathbf{u}}'$ called tensor of the Reynolds stresses:

$$\begin{cases} \nabla \cdot \bar{\mathbf{u}} = 0 \\ \nabla \cdot (\bar{\mathbf{u}} \bar{\mathbf{u}}) + \nabla \cdot (\overline{\mathbf{u}'\mathbf{u}'}) = \nabla \bar{p} + \nu \nabla^2 \bar{\mathbf{u}} \end{cases} \quad (3.8)$$

where an overbar $\bar{\cdot}$ indicates the time-averaging operator, and a prime denotes fluctuations. These equations become closed and thus solvable only after a turbulence model specifies the functional form of the Reynolds stress tensor in terms of the mean velocity.

The finite-volumes flow solver is OpenFOAM (Weller *et al.*, 1998). A flow rate of 15 l/min is enforced, corresponding to a mild inspiration. All modeling choices agree with relevant works that proved their accuracy at comparable flow rates, see e.g. Liu *et al.* (2007); Hörschler *et al.* (2010). The employed turbulence model is the $k - \omega$ SST turbulence model, chosen because of its ability to provide reasonable results while being numerically stable, and also because of its prevalence in the CFD

studies of the nasal airflow (Islam *et al.*, 2020). Default values for the model coefficients are used. The differential system (3.8) is provided with boundary conditions. At the inlet, the required value of the flow rate is prescribed for the velocity; no-slip and no-penetration conditions are applied at the solid boundaries representing the mucosal lining. At the outlet, located in the nasopharynx, a zero-gradient condition is applied to the velocity vector. Pressure has zero gradient at the inlet and on the solid boundaries of the nasal cavities, whereas at the outlet an arbitrary reference value $p_0 = 0$ for pressure is specified: in the incompressible setting, only pressure differences have dynamical meaning. As for the turbulent quantities, the required boundary conditions are summarized in table 3.2.

	Inlet	Outlet
k	$k = 0.01$	$\partial k / \partial \mathbf{n} = 0$
ν_t	$\partial \nu_t / \partial \mathbf{n} = 0$	$\partial \nu_t / \partial \mathbf{n} = 0$
ω	$\omega = 0$	$\partial \omega / \partial \mathbf{n} = 0$

Table 3.2: Boundary conditions for turbulent quantities: turbulent kinetic energy k , eddy viscosity ν_t and turbulence frequency ω .

3.4.3 Adjoint solution and surface sensitivity

While the procedure outlined so far is standard, computing the adjoint solution and the surface sensitivity is way less common, and to our knowledge has never been attempted in the context of the nasal airflow.

The system of the adjoint differential equations has been already presented in section 3.5.2, where the PDE to compute the adjoint velocity \mathbf{v} and the adjoint pressure q were derived from the steady Navier–Stokes equations (3.2). Here, we use the RANS equations (3.8) as primal equations, hence the adjoint equations would have to include the effect of the turbulence model of choice. However, we take advantage of the so called "frozen turbulence" assumption, according to which the variations of the turbulent quantities with respect to the control variables are negligible. Thus, the derivation of the adjoint counterparts of the complete RANS equations including the turbulence model is avoided. This reduces complexity and computational cost, while the negative consequences on the computed sensitivity map are negligible (Schramm *et al.*, 2018).

Once the adjoint system is available, the cost function f to drive the optimization must be chosen. This critical step will determine the boundary conditions for the adjoint problem. We choose the total dissipated power as the quantity to be minimized. Indeed, dissipation is linked to the resistance encountered by the nasal airflow, and clearly increases when obstructions are present. The dissipated power is written as the integral across the boundary of the net flux of mechanical energy, i.e. the sum of pressure and kinetic energy:

$$f = \int_{\Gamma} \left(p + \frac{1}{2} u^2 \right) \mathbf{u} \cdot \mathbf{n} d\Gamma \quad (3.9)$$

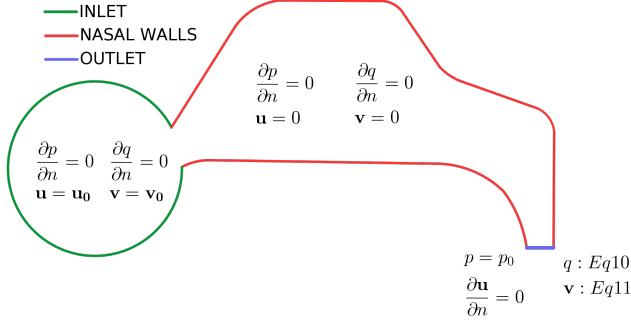


Figure 3.4: Schematic representation of the computational domain with the boundary conditions for the direct (\mathbf{u} and p) and adjoint (\mathbf{v} and q) equations.

The dissipated power equals the viscous losses, but the equivalent expression above yields an easier expression to handle. Moreover, in this way f explicitly depends on the flow variables only. (Obviously, the control variables β , albeit not explicitly appearing in the cost function, are involved in the procedure through their role in the constraints R .)

The definition of f determines the boundary conditions for the adjoint equations. Again, details on their analytical derivation can be found e.g. in Ref. Othmer (2008): here we simply list them, and represent them schematically in figure 3.4 together with those of the direct problem.

At the outer ambient and at the solid walls, the conditions for q and \mathbf{v} are identical to those for p and \mathbf{u} . At the outflow boundary, instead, the derived boundary conditions are:

$$q = \mathbf{v} \cdot \mathbf{u} + v_n u_n + \nu (\mathbf{n} \cdot \nabla) v_n - \frac{1}{2} u^2 - u_n^2 \quad (3.10)$$

and

$$0 = u_n (\mathbf{v}_t - \mathbf{u}_t) + \nu (\mathbf{n} \cdot \nabla) v_t \quad (3.11)$$

where the subscripts n and t refer to the component normal and tangential to the boundary.

3.5 Results

A brief and qualitative description of results from the standard direct RANS simulation is presented first in section 3.5.1: the newly introduced adjoint fields are shown in section 3.5.2. For brevity, only P2 is considered, as the one who presents the most evident anomalies. The sensitivity maps computed for P1, P2 and P3 are specifically addressed later in section 3.5.3.

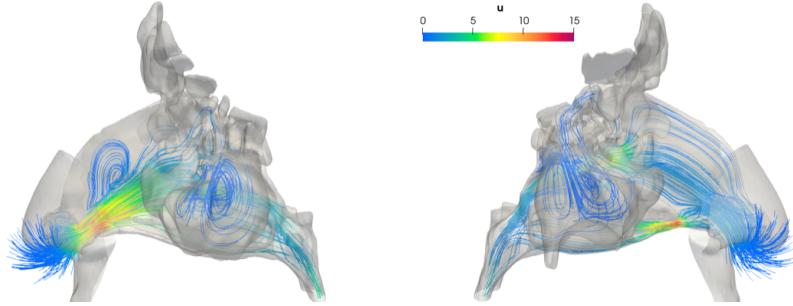


Figure 3.5: Patient P2: mean streamlines in the left and right nasal cavity (respectively), colored by the magnitude of the mean velocity.

3.5.1 RANS

Figure 3.5 shows three-dimensional views of the mean streamlines for P2, for the left and right passageways. The flow undergoes first a significant acceleration near the nose tip on both sides from the nearly still external air to more than 2 m/s . A large recirculation is then observed after the left deviation of the quadrangular cartilage. Here the velocity magnitude is around 1 m/s , which represents its minimum in both nasal fossae. A further evident consequence of the septal deviations is the flow acceleration in the restriction of the inferior right meatus. Here a very large peak value of 13 m/s for the velocity magnitude is reached. Globally, obstructions may cause an asymmetry between the left and right nasal cavities; for P2, 9.33 l/min and 5.67 l/min are the left and right flow rates, corresponding to 62.2% of the flow rate through the left passage and the remaining 37.8% through the right one.

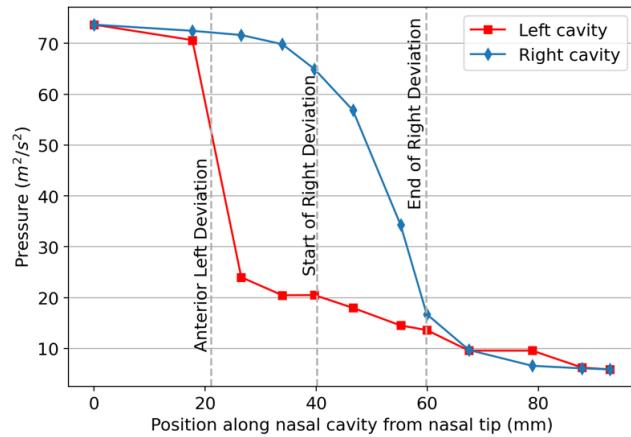


Figure 3.6: Patient P2: evolution of the mean section-averaged pressure along the right and left nasal fossae, from the nose tip to the choana.

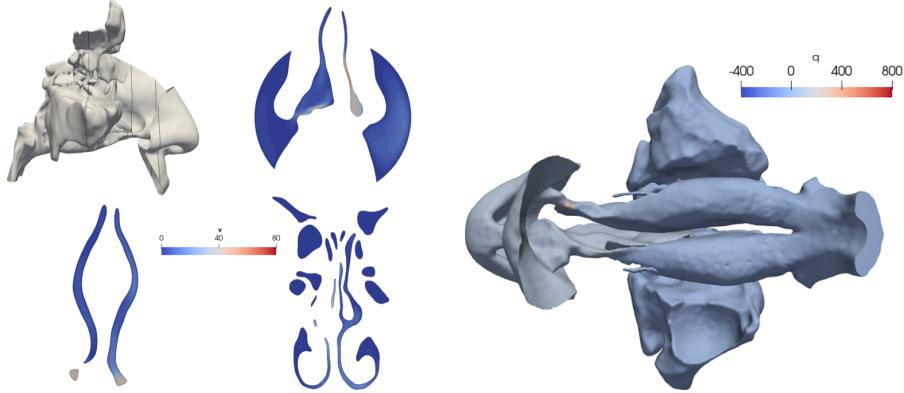


Figure 3.7: Patient P2: adjoint velocity magnitude, visualized in three sections shown in the top left panel, and taken in correspondence of major obstructions.

Figure 3.8: Patient P2: adjoint pressure.

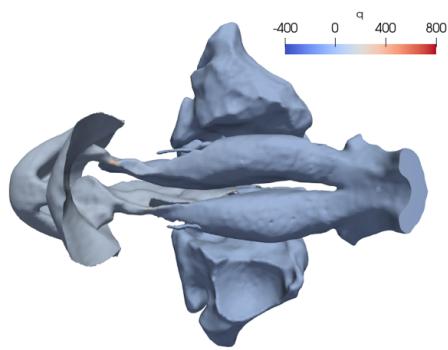


Figure 3.6 quantifies the asymmetry between the two cavities in terms of the evolution of the mean section-averaged pressure from the nasal tip to the choanae, and helps identifying regions of local pressure losses. Pressure values are computed at twelve locations, by averaging the mean pressure field over the entire local cross-sectional area. In both cavities, significant pressure losses are present in correspondence of obstructions. In the left fossa, the anterior deviation causes a sudden and localized decrease from $70 \text{ m}^2/\text{s}^2$ to $23 \text{ m}^2/\text{s}^2$; after the anomaly, pressure decreases smoothly. In the right fossa, instead, although flow perturbations are even more severe at both points where the bone spur reaches the turbinates, section-averaged losses appear to be milder; they are not concentrated at one specific section, but involve a rather large tract that extends for 25 mm along the axis of the fossa.

3.5.2 Adjoint field

The adjoint velocity and pressure fields are visualized, once again for patient P2 only, in figures 3.7 and 3.8. The adjoint velocity \mathbf{v} and the adjoint pressure q are not lending themselves to an immediate physical interpretation. They are mathematically defined fields, which depend directly on the chosen objective function (3.9): changing f would lead to computing different adjoint fields, as the adjoint equations are unchanged but their boundary conditions depend on f . Both fields are seen to assume maxima and minima in qualitative correspondence to obstructed regions, as a consequence of the choice of the dissipated power as the objective function. However, useful information for optimization can be obtained only when the adjoint field is combined with the primal field, via the surface sensitivity (3.6).

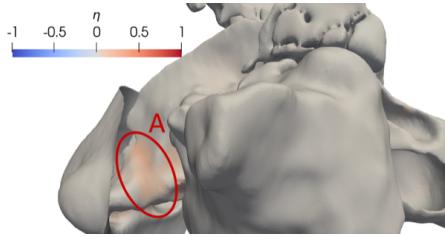


Figure 3.9: Surface sensitivity for P1, region A. Lateral view of the left side.

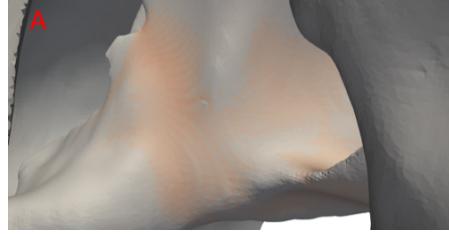


Figure 3.10: Zoom on the region marked by the red circle on figure 3.9.

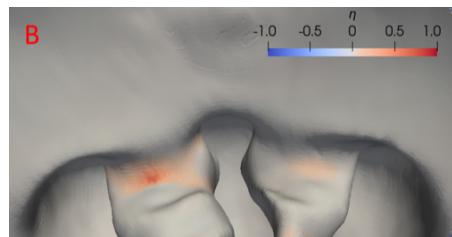


Figure 3.11: Surface sensitivity for P1, region B. View from below.

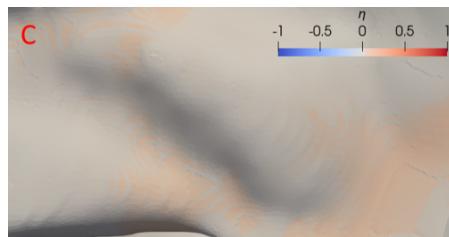


Figure 3.12: Surface sensitivity for P1, region C. View from the right of the left nasal cavity in correspondence of the left deviation.

3.5.3 Surface sensitivity

Combining the information contained in the direct and adjoint solutions into the surface sensitivity via Equation (3.6) allows one to quantify, for each point on the mucosa, the potential minimization of the cost function that can be achieved by surgery. Sensitivities, which contain the most clinically important information, are discussed below for all three patients. Owing to the linear nature of the adjoint problem, a normalized surface sensitivity (represented hereafter with the letter η) is visualized, i.e. the surface sensitivity of each cell is divided by the corresponding global maximum.

To interpret results from a surgical point of view, it should be kept in mind that the red color (i.e. high positive sensitivity) labels regions where the optimization indicates reduction of the cost function via a normal displacement of the surface which enlarges the cavity, i.e. consistent with the typical surgical action: the sensitivity takes its sign from the projection onto the wall-normal direction, which is oriented from the fluid region outwards. Regions in gray, instead, is where sensitivity is small; here the benefits of surgery towards minimization of the cost function are limited. Lastly, blue-colored regions suggest a local reduction of the cross-sectional area, and imply surgical reconstruction.

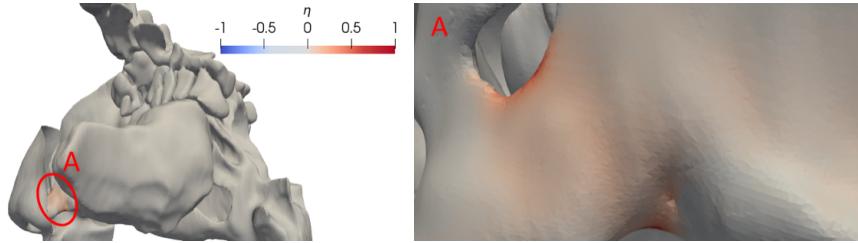


Figure 3.13: Surface sensitivity for patient P2, region A. Left: lateral view of the left side. Right: zoom on the region marked by the red circle.

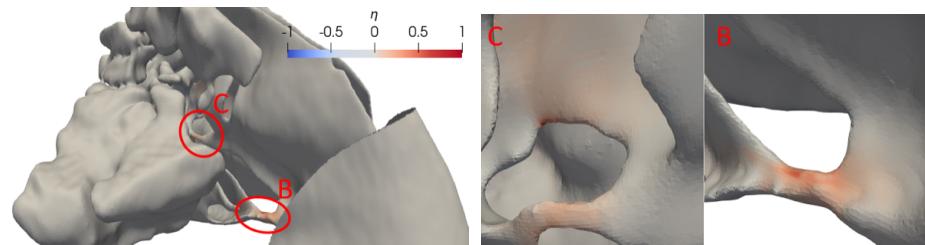


Figure 3.14: Surface sensitivity for patient P2. Left: lateral view from the right. Right: zoom on the two regions marked by the red circles.

P1

The surface sensitivity for patient P1 is shown in figures 3.9 - 3.12. For this patient, the adjoint optimization identifies three main areas with large sensitivity: region A in correspondence of the deviation of the quadrangular cartilage, region B behind the right nostril, and region C located along the left nasal fossa where the septum is deviated.

Figure 3.9 shows a lateral view of the left nasal airway, where region A is located; figure 3.10 is a zoomed-in view that highlights its characteristics. Local maxima for the sensitivity are found in correspondence of the obstruction, where cells have normalized values of about 0.2. Even though not visible, a similar red area (with lower sensitivity values) is also found on the same part of the nasal fossa but medially towards the septum. Hence, the adjoint procedure suggests an enlargement of the entire cross-sectional area in this area.

To analyze region B, figure 3.11 provides a view of the area behind the nostrils. This region hosts the cell with maximum sensitivity (normalized unitary value) on the right nostril. Nearby, a relative large area with sensitivity values above 0.25 is found. Also, a few cells with very small negative sensitivities are observed.

Finally, region C is detected in correspondence of the left deviation of the vomer bone, and is shown in figure 3.12. Here, rather small normalized sensitivities around 0.03 are obtained, suggesting a displacement that corrects the distortion. This deviation has been correctly identified, even though it does not cause a severe obstruction, and the sensitivity value is quite small.

P2

The surface sensitivity for patient P2 is shown in figures 3.13 and 3.14.

The first focuses on the sensitivity obtained in region A, corresponding to the left anterior deviation. The outcome of the adjoint here resembles that of region A for P1: the critical region is correctly identified, and the sensitivity suggests to surgically enlarge the entire cross-sectional area. However, some differences with P1 are observed when the local values of sensitivity are considered. First of all, for P2 this area corresponds to the global maximum of sensitivity. Furthermore, the area interested by large sensitivities is larger, indicating that acting on a wider area would contribute to the minimization of the cost function.

Figure 3.14 shows a global view of the nasal cavities seen from the patient's right, and emphasizes two other critical regions, noted B and C, visible in more detail in the zoomed-in views. The right bone spur obstructs the flow by touching the inferior turbinate in region B, and the middle one region C. In both cases, the adjoint procedure suggests to enlarge the restriction, but larger sensitivities are seen in correspondence of the inferior meatus, where sensitivity values of about 0.4 are found in the top part of the restriction; in other points of the same restriction the sensitivity is about 0.2. For the middle turbinate restriction, the adjoint optimization suggests to act on two locations. One corresponds to the middle meatus, where all cells have comparable sensitivity values around 0.1. The second location, instead, is the part of the left nasal cavity located between the deviated septum and the turbinate, and is characterized by normalized values around 0.04.

P3

Figures 3.15, 3.16 and 3.17 show the surface sensitivity computed for patient P3. For this patient, five critical regions are identified. Region A is the obstruction caused by the collapse of the nasal valve; regions B and C are the start and end points of the contact region between the left septal deviation and the inferior turbinate; region D is the obstruction caused by the left condro-vomerian spur that reaches the middle turbinate; region E is behind the right nostril.

Figure 3.15 plots region A. Here, the optimization suggests to operate with an outward displacement on the entire cross-section where the nasal valve is collapsed. The highest sensitivities are found in the top and bottom parts of the section, and are about 0.1, whereas the surrounding red cells are about 0.04. A corresponding red area is also found in the internal part of the nasal cavity, with a sensitivity around 0.3 as shown in figure 3.16, which is a view on the left nasal cavity as seen from the inside. This visualization, obtained via a proper clipping of the geometry, highlights also the critical regions B and C caused by the left septal deviation. First, the starting point of obstruction with the inferior turbinate can be analyzed. Here the solver computes values of the normalized sensitivity around 0.04 and this red region also involves a small portion of the meatus and the bottom part of the nasal cavity. In correspondence of the end of the contact, instead, smaller values around 0.025 are observed. Region D is highlighted in the zoomed-in view. Here the adjoint procedure gives a normalized value of 0.015.

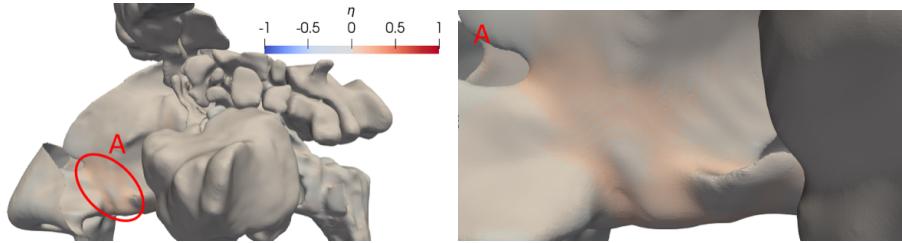


Figure 3.15: Surface sensitivity for patient P3, region A. Left: left lateral view. Right: zoom on the region marked by the red circle.

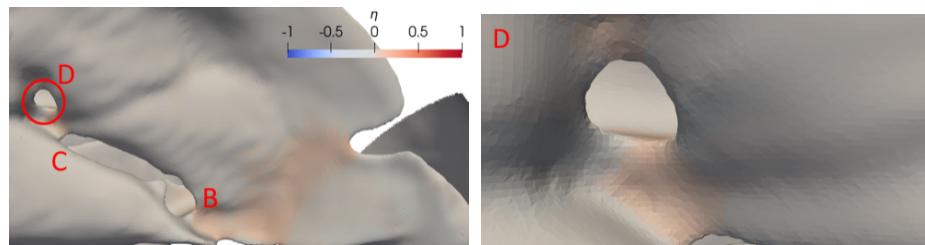


Figure 3.16: Surface sensitivity for patient P3, regions B, C and D. Left: view of the left nasal cavity from the "inside". Right: zoom on the region marked by the red circle.

Lastly, figure 3.17 shows region E, where sensitivities around 0.25 are found, with some cells peaking at about 0.35.

3.6 Discussion

We have introduced a novel approach to septoplasty planning. Besides using relatively standard tools for the execution of patient-specific CFD simulations of the nasal airflow, the procedure leverages adjoint optimization, for the first time applied in the context of the nasal flow. The former part of the procedure, which includes the segmentation of the CT scan, the creation of the computational mesh, and the set up of the CFD simulations, does not require special considerations. Although several important and

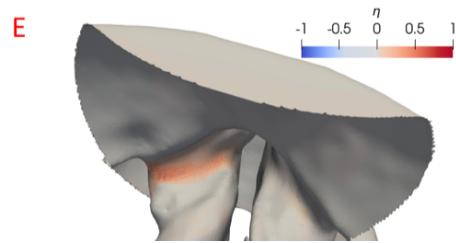


Figure 3.17: Surface sensitivity for patient P3, region E.

critical steps are involved in that part, they will not be discussed here, since abundant literature is available. We note, however, that the robustness of the direct procedure with respect to those steps carries forward to the adjoint part too.

The original part of the procedure is the application of the adjoint-based optimization theory, which indicates where, along the mucosal wall, a surgical action can be most effective. The critical sections along the airways, often but not always corresponding to local minima of the cross-sectional area, are identified: moreover, where along the contour of the section(s) should be operated upon is suggested. This descends from the ability of the adjoint optimization to combine anatomical and functional information, the latter descending from the solution of the primal problem: large sensitivities arise in locations where there is an obstruction to the flow, that is also accompanied by a potentially large flow once said obstruction is removed. Judging from anatomy alone, as surgeons traditionally and necessarily do, leads to missing the last factor. Three-dimensional maps of normalized sensitivity are immediately interpretable by an ENT surgeon: she can appreciate at a glance where it is suggested to act, and how the obstruction should be reduced to maximize the benefits and minimize the invasiveness of the procedure.

To make the optimization procedure clinically viable, it is important that the optimization procedure can be as streamlined as possible, reducing the required operator time to a minimum. Presently, the segmentation of the CT scan is the only step that requires external supervision to check the quality of the reconstruction. As such, it is the most expensive part in terms of operator time, and it requires 10-20 minutes, depending on the expertise of the operator. The steps after segmentation can be easily automated, as they can be set up once and for all, independently from a specific anatomy. In terms of computing time, on a conventional laptop it takes about 10 hours to compute the surface sensitivity map starting from the reconstructed geometry. More performing hardware and parallel computing can reduce these figures almost at will.

The procedure has been preliminarily applied to three selected cases of patients presenting complex and diverse septal deviations. The adjoint formulation has correctly highlighted every portion of the nasal septum that was previously identified during a pre-operative inspection of the CT scans carried out by experienced ENT surgeons; the computed sensitivity was very low over the turbinates, paranasal sinuses and at the nasopharynx, in agreement with the clinical observation that no patient presents turbinate hypertrophy. The informed opinion of the ENT surgeons is that the surgical actions suggested by the adjoint are deemed surgically feasible, with the septal wall being displaced outwards and the section being locally enlarged. The only exception is in P3, where in region A a displacement of the collapsed nasal valve was suggested. The collapse of the nasal valve is a frequent dynamical occurrence caused by fluid-structure interaction e.g. during intense inspiration or sniffing, but is not usually observed for normal breathing, and is not surgically corrected. Hence, although enlarging the nostrils is definitely meaningful in purely fluid dynamical terms, the sensitivity map obtained around the nasal valve should be considered with care.

A noteworthy point concerns the robustness of the results. The sensitivities discussed above have been obtained by employing CT scans of standard quality and different spatial resolution: the axial spacing ranges from a rather low 0.925 mm for P1 and P2 to an average 0.4 mm for P3, thus including most – if not all – the routinely

acquired scans: adjoint optimization can be applied to ordinary scans without the need of a dedicated protocol.

The outcome of the optimization is also robust with respect to the discretization of the numerical simulations. This has been assessed by re-running all cases on a coarser mesh, where the total number of cells is nearly halved (3.06 millions for P1, 3.17 millions for P2, and 4.1 millions for P3), finding that the indications provided by the adjoint optimization and by the wall sensitivity remain unchanged. That the optimization is robust with respect to (reasonable) variations of the spatial discretization is by no means an obvious result.

Also, the outcome of the optimization is robust with respect to the choice of the turbulence model. Simulations were repeated (for P3 only) by using two different models, namely the $k - \epsilon$ model and the $v^2 - f$ RANS model, in addition to the $k - \omega$ SST. Results compared in terms of normalized sensitivity are unchanged within a relative 1% in every cell of the domain.

3.6.1 Validation by virtual surgery

With CFD, virtual surgery can be employed to conveniently assess possible surgical actions. After computing sensitivities, the original or pre-op anatomies have been thus modified according to the indications of the adjoint procedure, and post-op anatomies were created. The direct problem was thus solved again on the post-op anatomies. Virtual surgery was carried out by morphing the pre-op anatomies, i.e. every point of the nasal walls was displaced in the wall-normal direction according to the local value of the normalized sensitivity. An arbitrary scaling factor of 1 mm has been used to scale the maximum normalized sensitivity. Such displacements are quite small; the maximum value has been chosen to guarantee a balance between the possibility of measuring an improvement and the need of controlling the quality of the final STL file automatically generated with morphing.

	Before			After		
	L [%]	R [%]	R_{nose}	L [%]	R [%]	R_{nose}
P1	32.5	67.5	0.091	32.2	67.8	0.081
P2	62.2	37.8	0.360	60.6	39.4	0.311
P3	22.9	77.1	0.037	23.5	76.5	0.033

Table 3.3: Flow partitioning and nasal resistance computed for all three patients after optimization

Table 3.3 reports, for all the pre-op and post-op anatomies, the computed values of flow partitioning, and the corresponding nasal resistance, defined as $R_{nose} = \Delta p/Q$, where Q is the volumetric flow rate expressed in ml/s , and Δp is the pressure drop, expressed in Pa , from the outer ambient to the nasopharynx. Different types of obstructions lead to a rather wide range of nasal resistances; flow partitioning too goes from highly asymmetrical (P3) to nearly normal (P2). In fact, none of these quantities is a general and robust indicator for NAO.

The post-op anatomies indicate a rebalancing of the quantity of air passing through the two nasal fossae for P2 and P3; for P1, instead, a minor (less than 0.3%) deterioration of the flow symmetry is observed. At least at the first step of the optimization procedure, this outcome is certainly possible, as the objective function does not target flow symmetry directly. Hence, it constitutes a further indirect confirmation of the weak and indirect link between flow partitioning and NAO.

The nasal resistance, though, does decrease in all cases. Since the inlet flow rate was not changed, these variations are exclusively due to a reduction of the pressure difference between the nostrils and the nasopharynx. Changes in the nasal resistance are quite significant, in the order of 5-10%, which is to be evaluated against the very small maximum displacement, set at 1 mm , which corresponds to a typical or averaged displacement, on the virtually operated areas, of less than 50 μm . The surgery suggested by the adjoint thus achieves large improvements of the nasal resistance with minimally invasive surgical corrections.

3.6.2 The choice of the cost function

Selecting the cost function is the most delicate step of the procedure, and one that is not entirely free from empiricism. In this work, the use of the dissipated power in Equation (3.9) was found to yield satisfactory results, with the identification and the characterization of all major anomalies reported by ENT doctors. However, this specific cost function would most probably not generalize to different obstructive problems, and even with septal deviations it remains to be assessed whether it represents the best possible option. Alternative cost functions that could describe additional functionalities of the nasal airways should be considered in future analyses. The physiology of the nose is complex, and its many functions should be properly evaluated. In this regard, the linearity of the adjoint equations might be exploited to define a more general cost function, written as the sum of multiple contributions to represent the different feelings and needs of the patients. This would also require the redefinition of both the direct and the adjoint equations, accounting e.g. for heat exchange and humidification. Work is ongoing to make the optimization procedure fully aware of the rich physics of the airflow in the nose.

3.7 Conclusions

Septoplasty is known to be sometimes ineffective at relieving patients from symptoms induced by a deviated nasal septum. This is related, at least in part, to the lack of standardized patient-specific tools to evaluate each septal deviation, and to provide the surgeon with functional information.

In this study, we have introduced time a CFD procedure that augments the usual numerical study of the nasal airflow with an adjoint-based optimization, thus becoming an effective tool for surgery planning. Adjoint-based optimization is used to compute sensitivity derivatives for a cost function that expresses the dissipated power and indirectly accounts for the nasal resistance. The procedure naturally outputs the quantitative

information needed by surgeons to decide where their efforts should be preferentially placed.

The validity of the adjoint-based procedure has been proved by applying it to three nasal anatomies affected by complex septal deviations. Inspection of the computed normalized surface sensitivity maps has demonstrated the ability of the method to automatically identify all the functionally important anatomical alterations. Furthermore, the surgical approach suggested by the adjoint formulation has been validated by ENT surgeons. The robustness of the procedure with respect to several aspects of the computational procedure has also been proved.

Further progress is certainly needed for this CFD method to become a clinical tool that ENT surgeons can use in their daily practice. The formulation itself in terms of surface sensitivity has alternatives, and work is underway to assess what is the best approach for the specific optimization problem. Moreover, a cost function of more general validity should be conceived, depending on the generality one intends to achieve. However, the present work represents a significant step towards a robust and patient-specific approach for computer-assisted septoplasty planning.

Chapter 4

Comparing flow-based and anatomy-based features in the data-driven study of nasal pathologies

4.1 Abstract

In several problems involving fluid flows, Computational Fluid Dynamics (CFD) provides detailed quantitative information, and often allows the designer to successfully optimize the system, by minimizing a cost function. Sometimes, however, one cannot improve the system with CFD alone, because a suitable cost function is not readily available: one notable example is diagnosis in medicine. The field of interest considered here is rhinology: a correct air flow is key for the functioning of the human nose, yet the notion of a functionally normal nose is not available, and a cost function cannot be written. An alternative and attractive pathway to diagnosis and surgery planning is offered by data-driven methods. In this work, we consider the machine-learning study of nasal pathologies caused by anatomic malformations, with the aim of understanding whether fluid dynamic features, available after a CFD analysis, are more effective than purely geometric features in the training of a neural network for regression. Our experiments are carried out on an extremely simplified anatomic model and a correspondingly simple CFD approach; nevertheless, they demonstrate that flow-based features perform better than geometry-based ones, and allow the training of a neural network with fewer inputs, a crucial advantage in fields like medicine.

4.2 Impact Statement

Machine-learning (ML) algorithms and Computational Fluid Dynamics (CFD) techniques are often discussed together in the recent scientific literature in fluid mechanics. However, ML is always used as a tool to perform a better/cheaper/faster CFD. In this work, we explore the potential of the inverse approach, in which CFD provides useful information to a ML model. In an idealized three-dimensional problem, flow features restricted to the boundary of the computational domain and derived from CFD are shown to be more informative than the geometry of the boundary itself, leading to a better ML classifier, which can be trained with fewer labeled data.

The application described in the paper is of the medical type, and concerns rhinology, where large amounts of accurately labeled data are not always available. Since the larger information content of flow-based features derives from the non-linear relationship between geometry and the corresponding flow field, the present result is relevant to other flow problems addressed with CFD where the lack of a clearly defined cost function suggests a data-driven approach.

4.3 Introduction

With the continuous development of computing hardware and software, Computational Fluid Dynamics (CFD) is becoming increasingly useful in several applications, extending from industry to health. CFD, ranging from the cheaper and lower-fidelity flow models like the Reynold-averaged Navier–Stokes equations (RANS) to the opposite extreme of the highly accurate direct numerical simulation, is a useful tool to improve the design of industrial systems, by e.g. increasing the aerodynamic efficiency of an airplane, reducing the aerodynamic drag of a vehicle, or enhancing the mixing in a fluidic system.

Sometimes, however, the CFD solution, albeit informative, does not explicitly provide an immediate means to improve the system. This is often the case in the medical field. The specific example considered in this work concerns the air flow in the human nose. The nasal cavities are the connecting element between the external ambient and the lungs, and serve a number of additional functions, which include smell, filtering and humidifying the incoming air, and heating/cooling it to the correct temperature. Most of these functions are directly driven by the anatomical shape of the nasal cavities. In fact, nowadays Ear, Nose and Throat (ENT) surgeons routinely take their surgical decisions mostly based on the analysis of Computed Tomography (CT) scans, which provide a detailed view of the anatomy of the nasal cavities. In principle, the nose flow can be well described by CFD, which is indeed increasingly used to support ENT doctors in their diagnosis (Moreddu *et al.*, 2019; Tjahjono *et al.*, 2023), and to improve our understanding of the complex physics of the nose flow (Calmet *et al.*, 2019; Farnoud *et al.*, 2020). Yet, the basic questions routinely asked by the ENT doctors (e.g. whether to perform a surgery on a given patient, and where) cannot straightforwardly be answered by CFD alone. Designing a surgery can be considered to be akin to a shape optimization problem. Unfortunately, the mathematical and numerical tools available for shape optimization cannot be deployed to solve the nose problem,

because the goal is not self-evident, owing to the lack of a functionally normal reference nose. A huge anatomical variability among healthy anatomies is present, which makes the discrimination between healthy and pathological cases far from obvious.

Several studies have attempted to develop a robust workflow for a CFD analysis of the nose flow (Quadrio *et al.*, 2014; Tretiakow *et al.*, 2020), and to understand what is a healthy airflow (Zhao & Jiang, 2014; Borojeni *et al.*, 2020) via multi-patient analyses. However, as seen from a clinical perspective, the present state of affairs remains unsatisfactory. There is evidence that the rate of failure of certain surgical corrections is extremely high: for the correction of septal deviations, for example, more than 50% of the patients report poor postoperative satisfaction ratings (Rhee *et al.*, 2003; Sundh & Sunnergren, 2015; Tsang *et al.*, 2018). Although there is general agreement (Inthavong *et al.*, 2019) that CFD offers a significant potential for improved surgery planning, this potential remains as yet largely untapped.

An interesting approach to diagnose pathologies and suggest surgeries relies on the use of Machine Learning (ML) techniques. The central question that we are going to address in this paper is whether CFD-computed flow information can be useful in this process, and possibly be more effective than the geometrical information embedded in the CT scan.

ML in fluid mechanics has recently seen a huge activity and recorded significant progresses (see e.g. the review by Vinuesa & Brunton, 2022); however, very little information is available in the literature regarding the combined use of ML and CFD in rhinology, if exception is made for our own preliminary study (Schillaci *et al.*, 2021b). The use of artificial intelligence and machine learning techniques has been limited so far to classification of images derived from CT scans (Crowson *et al.*, 2020). In closing their paper, Lin *et al.* (2020) mention that putting together CFD and ML would be an interesting future avenue for research. A very recent study by Jin *et al.* (2023) uses CFD and ML approaches, but only one at a time, and there is no attempt to combine them in any way.

The main goal of the present work is to answer the question whether flow features derived from CFD of the nasal airflow can be useful for a ML-based analysis. The possibility that CFD-based features outperform geometrical ones is rooted in the nature of the highly non-linear Navier–Stokes equations, which link anatomy and the ensuing CFD solution. A positive answer would carry general interest in all those situations where ML is used in the context of flow systems to replace optimization, since a cost function is not readily available.

To answer the question above, we consider synthetic (but realistic) healthy and pathological nasal anatomies, created with CAD, and a simple CFD solution of the flow within them (computed with the Reynolds-averaged Navier–Stokes equations and a standard turbulence model). An inference model made by a standard neural network is trained to understand whether each synthetic nasal anatomy is affected by a pathology, and to predict its severity. Two alternative approaches are employed. One resembles the approach currently followed by ENT doctors for their diagnosis, and is solely based upon geometric/anatomic information. The other, instead, relies on flow features extracted from CFD.

By comparing the performance of the two types of features, we intend to understand whether CFD, albeit somewhat costly, carries potential advantages, like e.g. increased

accuracy or the need for less observations, over anatomy alone. In particular, the ability to train the ML model with information derived from fewer patients would be extremely important, owing to the difficulty of obtaining highly informative and accurately labeled training data in health-related applications.

The structure of the work is as follows. After this Introduction, §4.4 describes the generation of a suitable set of geometries, the setup of CFD simulations, the extraction of information from the CFD solution, and the neural network used for regression. Results presented in §4.5 are used to critically discuss the comparison between geometrical and CFD features. Lastly, concluding remarks are put forward in §4.6.

4.4 Methods

This Section illustrates the entire workflow, and describes in §4.4.1 how the parametric CAD geometries of the noses are created, in §4.4.2 how the CFD simulations are set up, in §4.4.3 how different anatomies and flow solutions are compared, and in §4.4.4 how a neural network is designed and trained to predict nasal pathologies.

A CAD-based, synthetic reference nose model is built first. The model presents the fundamental advantage of being parametric; changing the numerical values of a small set of geometric parameters allows us to introduce anatomical variability, related to both physiological inter-subject differences and pathological conditions. Using this parametric nose model, 200 distinct anatomic shapes are generated. For each shape, a point-to-point mapping is computed between each nose and the reference nose model.

A CFD simulation is then carried out for each geometry. Thanks to the previously computed mapping, both flow and geometrical features of each geometry can be mapped back to the reference one, so that any (flow or geometrical) feature of any model can be observed on the reference one. At this point, a neural network is trained to perform a regression on the geometrical parameters of each nose and to predict the value of the pathological parameters.

4.4.1 The anatomies

All the anatomies considered in the present work originate from a reference CAD-based simplified model of the human nasal cavities, already introduced by Schillaci *et al.* (2021*b*), which lends itself to a simple geometrical parametrization. The use of simplified model geometries for the study of the flow in the human nasal cavities is not new: for example Liu *et al.* (2009) employed a model obtained by averaging together the CT scans of 30 patients. Our CAD-based approach relates more closely to that by Naftali *et al.* (1998), who used a CAD nose-like model to reproduce the essential features of the nasal cavities. However, we are first to build a fully parametrized model, that is required for the generation of a complete and controlled dataset.

The reference CAD nose model is meant to represent a healthy anatomy, and realistically mimics the major anatomical structures of a real nose. A qualitative comparison between the nasal cavities of a real patient (top) and our simplified baseline model (bottom) can be seen in figure 4.1. The CAD model, developed in collaboration with a group of ENT surgeons, is designed to reconcile the opposite requirements of

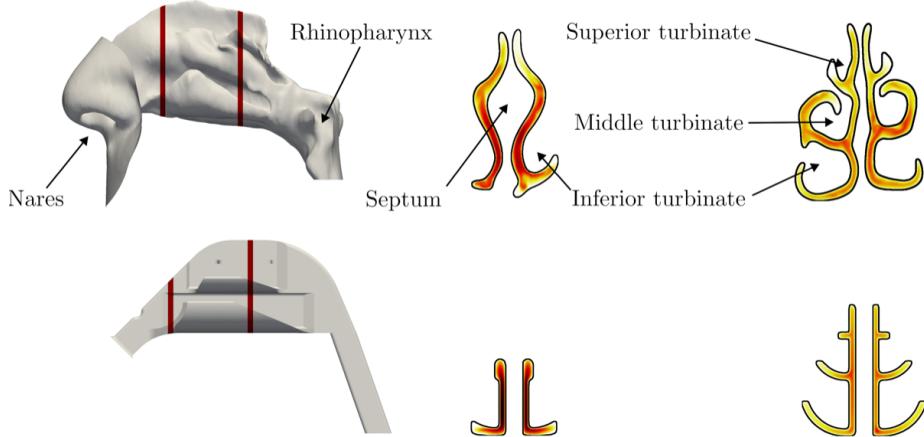


Figure 4.1: Comparison between a real nasal anatomy (top) and the simplified CAD model used in the present work (bottom). The three-dimensional view on the left marks the two coronal sections plotted on the right. The colormap represents the CFD solution in terms of the magnitude of the velocity vector. The top solution results from averaging a time-averaged LES solution (not reported in this paper), and the bottom one is a RANS solution. Although the CAD model is highly simplified, the major features are in line with the real anatomy.

simplicity and clinical significance. Its planar or constant-curvature surfaces are indeed highly idealized, but the model replicates in a quantitatively accurate way the crucial anatomical features, such as the dimensions of the septum between the two fossae, the hook-like structure of the inferior and middle turbinates, and the thickness of the passageway in the most critical areas of the nasal fossae. This is an essential prerequisite to provide deformations with clinical significance. The comparison (shown in figure 4.1) of the flow field computed with an high-fidelity approach on a patient-specific anatomy confirms the suitability of the present model.

The nasal geometry begins anteriorly with the nares, and ends posteriorly with the rhinopharynx and the hypopharynx. The nasal septum, a thin structure lying approximately on the median plane, separates the nasal cavities in two halves, the left and right nasal fossae. Each fossa has a cross-sectional shape that changes significantly along the nasal vestibules, developing as a narrow channel of convoluted shape, medially bounded by the septum; the particular conformation of the fossae is considered to improve humidification and heat exchange. Three long and curled bony structures, the (inferior, middle and superior) turbinates, define the cross-sectional shape of the fossae. The turbinates unfold roughly parallel to the flow, and are attached to the lateral walls of the nose. The inferior turbinate is the largest, running almost the entire way from the vestibule to the rhinopharynx.

The parametric CAD model contains eight numerical parameters, whose range of variation is meant to account at the same time for the physiological and the pathological variability of real anatomies. Three parameters q_1 , q_2 and q_3 correspond to three

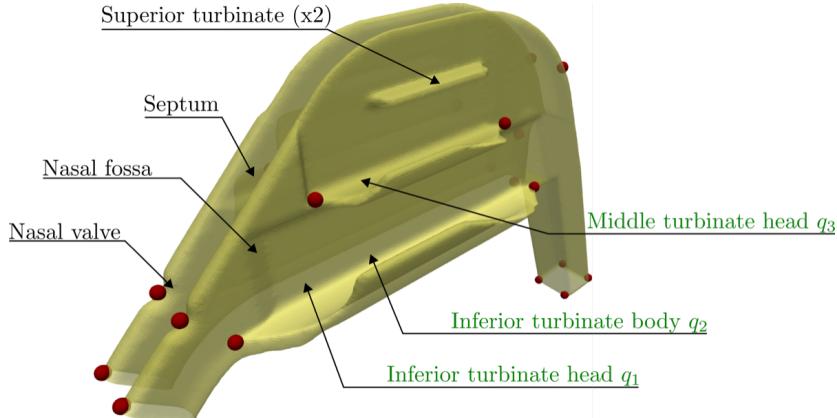


Figure 4.2: Nose model and position of its parametric modifications; black text labels indicate physiological variations, and green text highlights pathologies. The red points are landmarks used for functional mapping.

clinically sensitive regions of the nasal cavities, and describe the intensity of selected pathologies, related to an hypertrophy of the turbinate; the remaining five parameters are clinically harmless, and represent the physiological variability among healthy anatomies. The parameters describe anatomical changes or their arbitrary combinations; their numerical values are expressed in millimeters and are changed in steps of 0.05 mm. In the healthy reference anatomy, all the parameters are set to zero. Figure 4.2 shows where in the model the parameters act to modify the reference nose.

Ninety-nine extra healthy anatomies are created from the reference one by varying the values of the five healthy parameters. They alter (see figure 4.2) the vertical and longitudinal position of the superior turbinate, the axial position of the nasal valve, the thickness of the septum in the area of the nasal valve, and the thickness of the septum in correspondence to the head of the inferior turbinate. Varying the parameters produces localized and relatively small changes in the geometry: the strongest change leads to a 14% reduction in cross-sectional area in the most affected section, while the smallest change decreases the area by 0.7% only.

One hundred pathological anatomies are created by varying the values of the remaining three pathology-related parameters. They are designed to mimic in a quantitatively reliable way a common condition called turbinate hypertrophy, a swelling of the turbinates which produces a constriction of the meati, up to a point where the airway may, in the most severe cases, become completely obstructed. The obvious consequence is reduced nasal patency and difficulty to breathe. Such hypertrophies have a number of causes, from allergic rhinitis to inflammation of the sinuses, and affect one or more turbinates in either of the nasal fossae.

The reference CAD shape is modified by altering the values of the three pathological parameters; their numerical values are set under tight supervision of ENT doctors, and remain within clinically meaningful values, to ensure that deformations are realistic, notwithstanding the idealized model. Parameter q_1 varies between 0 and 0.7 and mimics

an hypertrophy of the head (anterior portion) of the inferior turbinate; parameter q_2 varies between 0 and 0.7 and mimics an hypertrophy of the body (intermediate portion) of the inferior turbinate; parameter q_3 varies between 0 and 0.55 and mimics an hypertrophy of the head of the middle turbinate. All the pathologies are applied to the right fossa.

An important point to stress is that the parameterization of the geometry provides an unambiguous label for each case. The label consists in numerical values of the three pathological parameters: in other words, for each case it is precisely known which pathology is at play, and how much is its severity. This characteristic, that will be essential when training the inference model, is impossible to obtain when working with real anatomies.

4.4.2 Simulations

A CFD solution is computed for each of the 200 distinct anatomies. In view of the simplified geometries employed here, a basic RANS flow model is employed: the finite-volume package OpenFOAM (Weller *et al.*, 1998), with its SIMPLE solver and a first-order discretization are used to arrive quickly at a converged solution. In doing so, we follow standard modeling and discretization choices, which are summarized below.

The computational grid is generated with the utilities available in OpenFOAM. An uniform background mesh with cubic cells of side length 1 mm is created first; the mesh is then refined further and adapted to the surface. The final mesh is rather coarse, in line with the RANS approach, and consists of around 1.1 millions cells. Values in the nasal flow literature range between 0.1 up to 44 millions cells (Inthavong *et al.*, 2018), with the finest meshes being generally used in highly resolved LES studies (Calmet *et al.*, 2016; Covello *et al.*, 2018), meanwhile typical, modern RANS simulations range between 1 and 4 millions elements (Liu *et al.*, 2007; Wen *et al.*, 2008).

The dataset is built for a steady inspiration, which is considered as the most clinically representative breathing condition. The inspiration is driven by a pressure difference of $20 \text{ m}^2/\text{s}^2$ between the inlet and the outlet section, which for the reference geometry corresponds to a flow rate of about 178 ml/s . This value corresponds to an inspiration at rest or at mild physical activity (Wang *et al.*, 2012). As in the vast majority of CFD studies in this field (Radulesco *et al.*, 2019), the nasal walls are considered as rigid, thus neglecting the compliancy of the tissues (which is very small at this low breathing rate) and the modifications of the erectile mucosal tissue during the nasal cycle, known to cyclically alter the shape of the passageways over a time scale of few hours (Patel *et al.*, 2015). The velocity has zero gradient at the inlet and outlet sections. A no-slip boundary condition for the velocity components is applied at the walls, whereas for pressure a zero-gradient condition is enforced. The turbulence model of choice is $k - \omega - SST$ (Menter *et al.*, 2003), which is commonly employed in such simulations (Li *et al.*, 2017). The model solves two additional differential equations, one for the turbulent kinetic energy k and one for the turbulent frequency ω . At the inlet k is set by assuming just 1% turbulent intensity, zero gradient is used at the outlet, and $k = 0$ is set at the walls. For the turbulence frequency ω , at the inlet $\omega = 1\text{s}^{-1}$ is prescribed, at the outlet a null gradient is imposed, and at the wall the value is set as in Menter (1994).

The outcome of a typical simulation for the healthy anatomy is compared in figure

4.1 with the temporally averaged solution obtained on a real anatomy with an higher-fidelity (and significantly more expensive) CFD method, namely Large Eddy Simulation (LES), where a WALE turbulence model is used (Ducros *et al.*, 1999), for the same flow conditions. The LES mesh is of about 12.8 millions cells. Although the comparison clearly has to be intended in a qualitative sense only, it is seen that most of the flow passes through the same regions, in particular in the meatus between the septum and the inferior and middle turbinate for all the section. This confirms the suitability of the simplified model for the purpose of the present work.

We stress once again that, in the present study, seeking the highest fidelity in the solution is not our primary concern. Hence, the solution method (the RANS equations), the numerical schemes (first order) and the quality of the mesh (fairly coarse) are all standard, and meant to generate quickly and cheaply a dataset of reasonable size. Since each case carries around 100 *MB* of data, the full dataset has a total size of 20 GB. To put these numbers into perspective, the Google Open Images Dataset V6 dataset (Kuznetsova *et al.*, 2020) consists of around 9×10^6 images and is made by about 561 GB of data (including labels). In other words, our dataset is only one order of magnitude smaller, but consists of four order of magnitudes less observations, which emphasizes the high dimensionality of a typical CFD dataset.

4.4.3 Functional maps

A correspondence needs to be determined between the reference nose and each of the other noses in the full set. This correspondence is computed with a tool derived from computational geometry and called functional maps (FM) (Ovsjanikov *et al.*, 2012); functional mapping is briefly introduced below, and then specialized to the implementation employed here (Melzi *et al.*, 2019).

Functional mapping provides an efficient method to estimate the correspondence between two shapes, as well as the correspondence of functions represented on them. It is a relatively new tool, that has been recently introduced for solving shape classification problems (Magnet *et al.*, 2023). Rather than directly estimating point-to-point correspondence between shapes, FM registers functional spaces defined over the two shapes. The multi-scale basis for the function space on each shape is given by the finite (truncated) set of eigenfunctions Φ_j , $j = 1 \dots N$ of its Laplace–Beltrami operator. Once the basis is known, any function f defined on the shape is approximated by the following linear combination of eigenfunctions

$$f = \sum_{j=1}^N \gamma_j \Phi_j.$$

In general, the functional mapping between two shapes is described by a matrix A , whose elements describe how each eigenfunction on one shape is expressed as a linear combination of the eigenfunctions on the other shape. We refer the interested readers to the original paper by Ovsjanikov *et al.* (2012), or to the recent contribution by Magnet & Ovsjanikov (2023), for detailed information on functional mapping.

Given two shapes, each functional map is computed by solving a least squares minimization problem. To retrieve more precise maps, in our case twenty landmarks are

identified for each of the shapes and used as descriptors. Landmarks, shown in figure 4.2 for the reference nose, are points selected on the geometry because of their anatomical significance; they are often used in the ENT practice (see e.g. Denour *et al.*, 2020) to help dealing with different anatomies in the context of CT analysis or registration. As for the basis, the eigenfunctions of the Laplace–Beltrami operator used here possess the convenient characteristic of bringing out the dominating "frequencies" of the shape; therefore, they naturally provide a multi-scale representation of the geometry. Note that each nasal geometry has its own basis, but the more two geometries are similar, the more their bases are similar. Hence, the matrix A becomes more diagonally dominant when two geometries are alike.

The specific FM implementation used here is called zoom-out, and has been introduced by Melzi *et al.* (2019). Instead of computing the map for the full set of basis functions, with zoom-out one computes first a smaller map and a smaller matrix that involves fewer basis elements (say, 10); the map is then iteratively extended, by adding rows and columns to A , up to the desired size. We have determined that, with the shapes of interest, truncating the modal expansion to $N = 150$ eigenfunctions provides satisfactory results. Since the mapping is always between the reference shape and every other shape, 199 maps in total are computed.

Computing the correspondence between a generic shape and the reference one using FM involves the following steps:

1. Compute (once) the truncated Laplace–Beltrami base on the reference shape;
2. Compute the truncated Laplace–Beltrami base on every other shape;
3. Position the 20 landmarks on each shape;
4. Compute the functional map matrix A by solving a least-squares optimization;
5. Convert the matrix A into a point-to-point map.

Once the matrix A is available for the generic i -th nose, the workflow, graphically sketched in figure 4.3, starts from the corresponding CFD solution, from which relevant flow quantities (for example, the pressure field p_i , or the skin-friction field τ_i) are computed at the wall. Thanks to FM, the wall field p_i is transported back to the reference nose to yield the field \hat{p}_i ; the difference field $\Delta p_i = p - \hat{p}_i$ can be expanded by using the Laplace–Beltrami basis Φ of the reference anatomy and the corresponding coefficients γ_j , that will be used later in the regression. Note that eigenfunctions of the reference nose only are involved in the latter expansion, and that only wall-based quantities are mapped, thus eliminating any issue regarding the continuity equation. Vector field are transformed component-wise.

4.4.4 The classifier

Given the dataset with $\ell = 200$ observations, a vector of input features (be it geometrical or derived from the flow solution) must be associated to a target value which describes the pathology through the numerical values of each the pathology parameters q . Carrying out the proper association is the task of a regressor, usually implemented as some

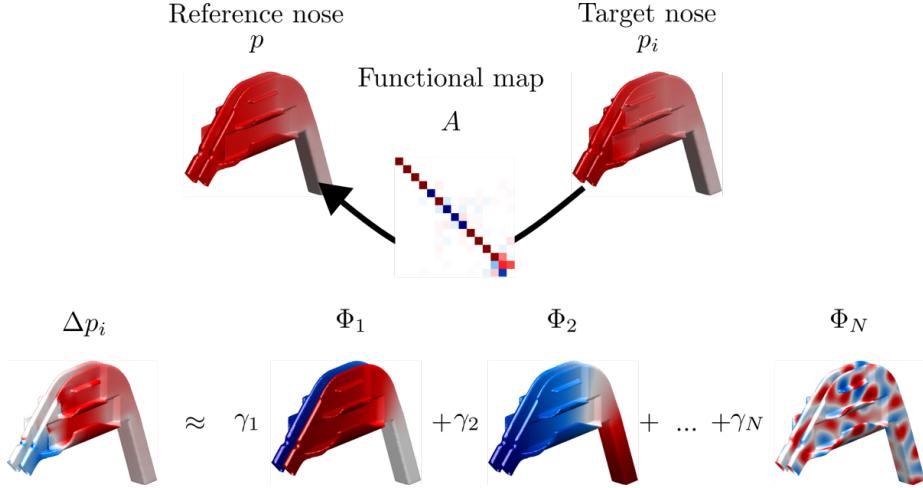


Figure 4.3: Difference of a wall-based flow quantity (pressure p in this figure) between the reference and the generic i -th anatomies. Pressure p_i on the boundary of the i -th nose is mapped to the baseline nose as \hat{p}_i . The difference $\Delta p_i = p - \hat{p}_i$ is expressed as a linear combination of the eigenfunctions of the reference nose with coefficients $\gamma_j, j = 1 \dots N$.

kind of neural network (NN). The raw data from CFD for each observation are available in each cell of the discretized domain. Since the cardinality of raw data is much larger than ℓ , a process of dimensionality reduction of the input, called feature extraction, is necessary to balance the number of observations with the number of inputs. The outcome of the feature extraction process depends on the specific experiment; hence, feature extraction will be described later in §4.5, where the various experiments are discussed.

We train two different NN models, depending on the input data: a Multi-Layer Perceptron (MLP) (Goodfellow *et al.*, 2016) when the input is a feature vector, and a Convolutional Neural Network (CNN) (Lecun *et al.*, 1998) when the input is the matrix A , that we treat as spatial data. CNN is chosen because of its ability to capture patterns in maps; it is widely used in the field of image recognition, and it has also been applied to fluid dynamics in recent years (see e.g. Fukami *et al.*, 2020), due to its capability to deal with spatially coherent information.

In general, designing the architecture of a NN involves several choices, e.g. deciding on the number of hidden layers and nodes, the activation function and the loss function. Our MLP is a regression network; it has an input layer, whose number of nodes is equal to the length of the feature vector, three hidden layers (with 30, 20 and 10 nodes each) and an output layer with one node only. The activation function is the hyperbolic tangent for all the nodes, except for the output nodes, which has a linear activation function. Since the goal is to predict the numerical values of the parameters which quantify the severity of the pathology, we adopt the mean square error as loss function. Lastly, the optimization algorithm, which updates weights and biases of the NN, is the

classic Levenberg–Marquardt (Lera & Pinzolas, 2002). Reduction in the number of inputs is obtained by extracting the 20 most informative features via the Lasso method (Tibshirani, 1996).

Our CNN has a rather standard architecture. The functional map, i.e. matrix A , is passed into a convolutional block, which consists of a 3-by-3 convolution layer, then a dropout layer randomly deactivates 20% of the weights in its layer. Afterwards the data is normalized by a batch normalization layer and then fed into a hyperbolic tangent layer; finally data are reduced in size through a maxpooling layer with a 2-by-2 filter. The dropout and batch normalization layers are applied to avoid overfitting. After the convolution block, data are flattened into a vector and input into a fully-connected (FC) block, consisting of 20 perceptrons, dropout layers, batch normalization and hyperbolic tangent layers. The size of the FC block is halved until a single output node remains. The weights of the CNN are optimized with the widely used Adam method (Kingma & Ba, 2017), owing to its efficiency and stability.

For both NN architectures, a reliable assessment of the error over the entire dataset is obtained with the k -fold cross-validation method (James *et al.*, 2021). The dataset is partitioned over $k = 5$ folds: each has 140 cases used for training, 20 for validation and 40 for testing. The 40 simulations used for testing do not overlap over the 5 folds, so that the performance is assessed over the whole dataset, albeit by training 5 different NN (with the same architecture). To avoid the potential bias of considering just a specific run, this operation is run 100 times per feature and per pathology, and eventually the average absolute error is computed.

4.5 Experiments

Results of the regression experiments are now presented in comparative form, for geometrical and flow features. The goal of the experiments consists in retrieving the numerical value of the three pathological parameters.

In consideration of the relatively small size of the database, instead of training a single NN to predict the three parameters in a single attempt, we opt for training one NN for each of the three parameters.

4.5.1 Geometrical features

Multiple options exist to select geometrical features. In this work, we consider two features, identified with G1 and G2. The geometry-based feature G1 is simply the displacement between each point on a certain nose model and the corresponding point on the reference shape. Feature G2, instead, is the matrix A obtained via FM when registering each nose with the reference one. Owing to the different nature of G1 and G2, a different NN architecture is used, namely a CNN for G2 and a MLP for G1.

Feature G1 (shown in figure 4.4) is simply the distance between points on the reference nose and the corresponding ones on a modified shape. Since the modified shape is only varied through local changes, and thus does not need registration, unchanged points lead to zero displacement. The i -th nose is mapped on the reference nose, and the pointwise distance is a scalar field defined on the surface on the reference shape.

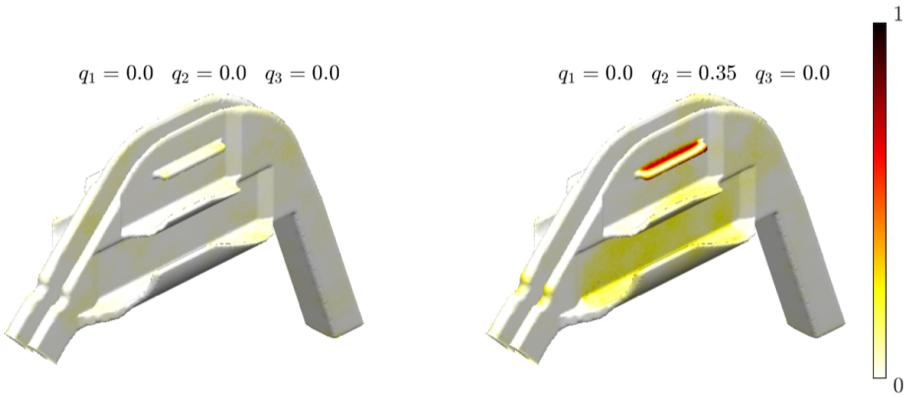


Figure 4.4: Pointwise distance (feature G1) for a healthy nose (left) and a pathological nose (right). The numerical values of the three pathological parameters are reported above each panel. The right plot shows large values of G1 which highlight a non-pathological parameter (namely the position of the superior turbinate), and non-zero but smaller values in the area interested by the pathological parameter q_2 (which mimics hypertrophy of the body of the inferior turbinate). The colormap units are mm .

This field is then decomposed into the Laplace–Beltrami basis Φ_j of the reference nose, arranged in a rectangular matrix, and the vector of Laplace–Beltrami coefficients γ_j . The ensuing overdetermined linear system is solved by using the Lasso method with a penalization constant. A careful choice of the constant leads to a solution with 20 coefficients only.

G1 is effective at spotting differences introduced by changes in the parameters. Figure 4.4 indeed shows that G1 peaks exactly where the geometrical variations have been introduced. A quantity like G1 is expected to be subject to a small amount of noise, since each nose undergoes its own meshing process. Even though a certain portion of the surface is unaltered, its points would modify their position slightly when a different mesh is computed. Nevertheless, figure 4.4 shows that the noise remains more than acceptable, so that the deterministic geometrical changes are clearly highlighted: G1 peaks at the superior turbinate (determined by a non-zero value of a non-pathological parameter), and is also large in correspondence of the inferior meatus, defined by the inferior turbinate, which is affected by a pathology since $q_2 = 0.35$.

Feature G2, instead, involves the functional map between the two shapes, expressed via its matrix A . Owing to the usual need to avoid overfitting, the CNN is not given the full matrix A , but only a square sub-matrix A_1 of size 20. Hence, we are only comparing the first 20 eigenfunctions between the two anatomies. We stress again that A , since it is based on geometry only, does not contain information concerning the flow field, and is computed without the need of a CFD solution.

Matrix A appears to have a more diagonal structure when the shape corresponds to a healthy nose, compared to pathological cases. This can be confirmed by looking at figure 4.5, which portraits in graphical form the structure of the sub-matrix A_1 : each

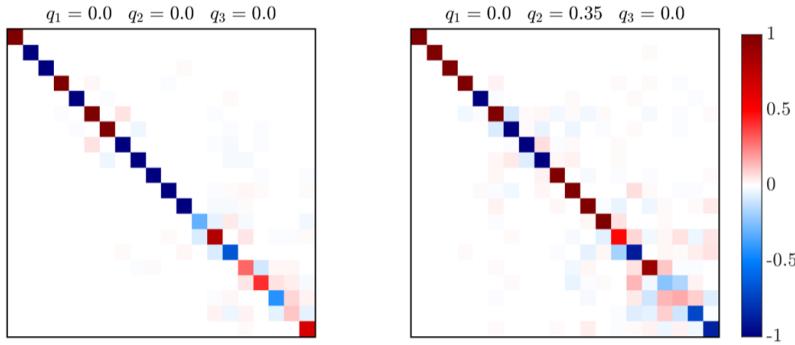


Figure 4.5: Functional map (feature G2), represented by the sub-matrix A_1 , for a healthy nose (left) and a pathological nose (right); each element is color-coded after normalization at unitary maximum. The numerical values of the three pathological parameters are reported above each panel. Healthy anatomies tend to produce more diagonal maps, whereas pathologies alter the diagonal structure of the matrix.

square represents an element, whose value is encoded in its color. A diagonal matrix implies that each eigenfunction in one geometry is fully described by the corresponding eigenfunction in the other geometry. When off-diagonal elements are non-zero, one eigenfunction on one nose becomes a linear combination of several eigenfunctions on the other.

4.5.2 Flow features

Flow features are restricted at the wall; this choice is supported by clinical considerations: the feeling of discomfort is conveyed by nerve terminations residing in the mucosal lining (Sozansky & Houser, 2014). The most straightforward wall-based features one can resort to in an incompressible flow are wall pressure p and the magnitude τ of the wall shear stress (Bewley & Protas, 2002). As before, the difference of either quantity between the reference case and each mapped-back case is computed and expended in the reference basis: input to the NN are the first 20 coefficients of the expansion. In sharp contrast to geometry-based features, that have no need for CFD, these features are based on CFD but do not retain direct information about the geometry: only the eigenfunctions of the Laplace–Beltrami basis on the reference nose are used in the procedure.

Figure 4.6 illustrates the first flow-based feature F1, given by wall-pressure: it shows the wall-pressure field p for the baseline nose (left), the transformed pressure field \hat{p}_i for the i -th nose, affected by a severe hypertrophy of the head of the inferior

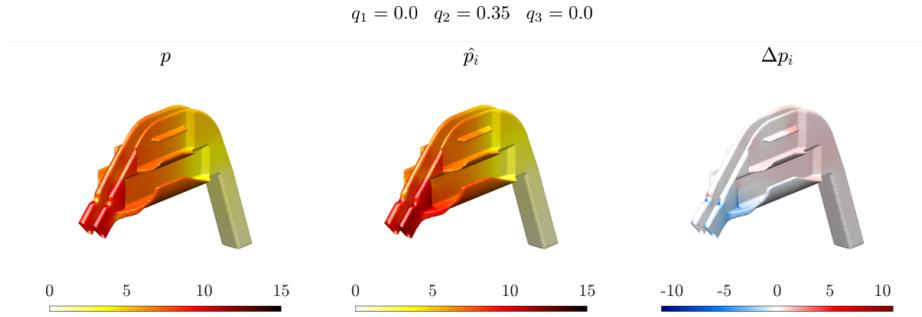


Figure 4.6: Wall pressure p (feature F1). Left: p on the reference anatomy. Centre: \hat{p}_i from a pathological anatomy (hypertrophy of the inferior turbinate) mapped back on the baseline. Right: Δp_i . The numerical values of the three pathological parameters are reported above each panel. Colormap units are Pa .

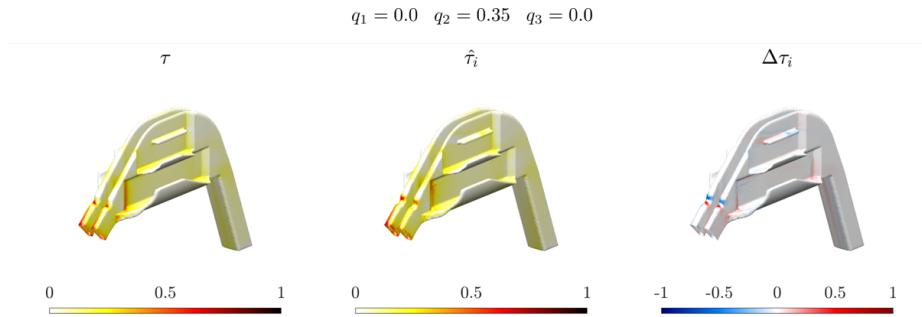


Figure 4.7: Wall shear stress τ (feature F2). Left: τ on the reference anatomy. Centre: $\hat{\tau}_i$ from a pathological anatomy (hypertrophy of the inferior turbinate) mapped back on the baseline. Right: $\Delta\tau_i$. The numerical values of the three pathological parameters are reported above each panel. Colormap units are Pa .

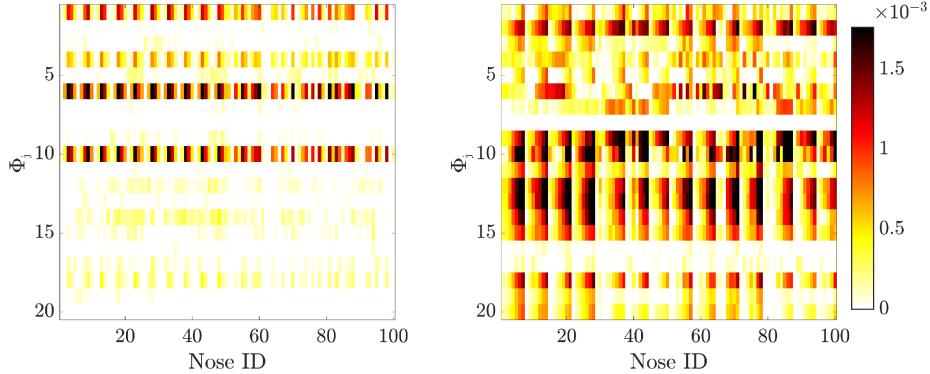


Figure 4.8: Wall shear stress coefficients using the first 20 Laplace–Beltrami eigenfunctions, for the 100 healthy noses (left) and the 100 pathological ones (right). Nose ID is on the horizontal axis, and mode number on the vertical axis (starting from top).

turbinate (center), and the corresponding difference Δp_i represented on the baseline nose. Although the two pressure fields p and p_i appear very similar, the difference field shows values of about 5 Pa, which is a clinically significant value whose order of magnitude captures the variance in nasal resistance for patients with hypertrophy.

Analogously, figure 4.7 describes the second flow-based feature F2 for the magnitude of the wall shear stress τ .

It is instructive to observe how wall-based flow quantities are reconstructed using the Laplace–Beltrami basis. As an example, for the wall shear stress, figure 4.8 compares the 100 healthy noses (left) with the 100 pathological ones (right): for each nose, the coefficients of the first 20 eigenfunctions of the Laplace–Beltrami expansion of the baseline anatomy are plotted, with color indicating the magnitude of the coefficient. It can be noticed that the pathological cases involve contributions from a higher number of modes, when compared to the healthy cases. Furthermore, certain eigenfunctions (e.g. modes 2,3,7,9) are minimally involved in the description of the flow solution for healthy noses, but become important to reconstruct the fields pertaining to pathological cases.

4.5.3 Performance and discussion

Results of the regression experiments are presented in table 4.1, in terms of the average error over the whole dataset. Distinct experiments are carried out for each pathology. The first, striking and most important observation is that geometry-based features produce an error which is about one order of magnitude larger than that obtained with flow-based features. Within flow-based features, the wall shear stress achieves a better performance over pressure in a consistent way over the set of pathologies. The pathology parameter q_3 (hypertrophy of the middle turbinate) turns out to be the hardest to predict via flow features (although its error when geometric features are used is the lowest among the three pathologies): this is reasonable, because an hypertrophy of the middle turbinate affects the anatomy in a region that is not crucial in terms of distribution of

Feature	NN type	q_1	q_2	q_3
G1: distance	MLP	0.156	0.116	0.094
G2: matrix A_1	CNN	0.148	0.117	0.099
F1: pressure	MLP	0.032	0.023	0.062
F2: wall shear stress	MLP	0.019	0.019	0.041

Table 4.1: Average value (in millimeters) of the test error computed over 100 runs, for each of the three pathological parameters q_1 , q_2 and q_3 . The test error of one run contains the average results of the 5 NN derived from k -fold cross-validation with $k = 5$. Flow-based features show a consistently lower error than geometry-based features.

the flow (see figure 4.1). However, even for such an unfavorable situation, flow features do perform significantly better the geometric ones.

The reason for the superiority of flow-based features can be traced back to the pathology being a geometrical modification which is typically quite localized in space, at least in our simplified model. As such, pathologies are small-scale features that tend to be visible only as higher-order modes of the Laplace–Beltrami basis. However, higher-order modes might be *per se* quite noisy, and dependent on the geometrical discretization. This difficulty could be further emphasized by the pathologies considered here, which aim at being clinically faithful and as a consequence sometimes involve geometric modifications which are small in absolute terms, often a fraction of a millimeter. If pathologies are examined in terms of the corresponding flow field, instead, these small changes are passed through the "filter" of the Navier–Stokes equations: a small and localized geometrical change in a sensitive position leads to a larger, more easily identified modification of the flow field. This, in turn, tends to appear within lower-order modes of the Laplace–Beltrami basis, and thus becomes easier to capture.

Figures 4.9 and 4.10 provide a closer look at the results, by focusing on a single experiment where one NN is tested on 1/5 of the dataset (40 noses). The specific test set and NN are chosen randomly and are thus representative of the whole set. Figure 4.9, where predicted and true values of the pathologic parameters are plotted side by side for each nose, emphasizes how the geometric features G1 and G2 do not perform particularly well: it seems that the limited anatomical variability considered in the present work is already enough to throw off the model. It should be noted, though, that the model prediction is not meaningless: non-zero predictions are often associated to non-zero ground truth, in particular for the parameter q_2 (hypertrophy of the body of the inferior turbinate), which involves a larger surface. Both G1 and G2 are unable to predict the values of hypertrophy on the middle turbinate. Flow features F1 and F2, instead, demonstrate good regression capabilities, especially when used to predict pathologies of the inferior turbinate. For the middle turbinate, the prediction accuracy decreases. Between the two flow features, wall shear stress has a small, but significant edge over pressure.

Figure 4.10 is an alternate view of the same dataset, and shows the correlation between the predicted (vertical axis) and true (horizontal axis) labels, for each feature and pathology parameter. The superiority of flow-based features with respect to geometry-based ones is even more evident. In particular, the figure suggests that F2,

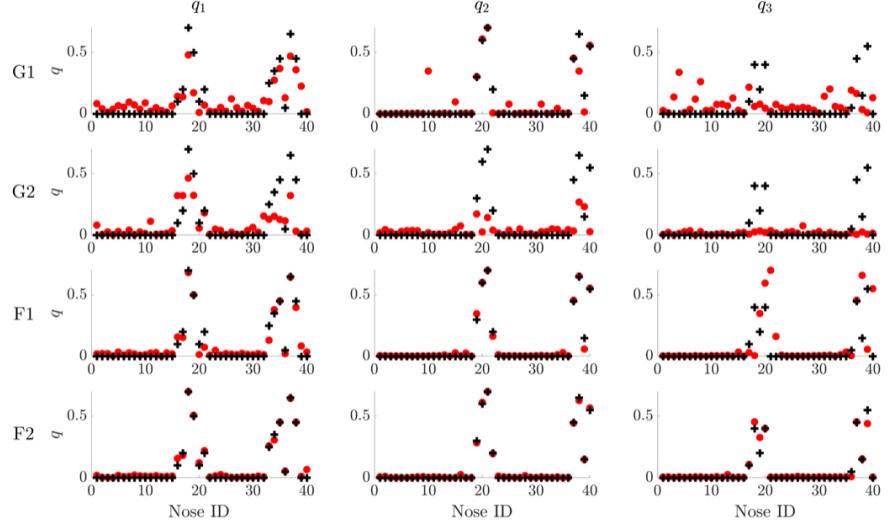


Figure 4.9: Performance of the various features (rows) in one experiment for predicting the three pathological parameters (columns): predicted values (red symbols) and ground truth (black symbols) for various nose IDs.

the distribution of wall shear stress, is the most effective feature at predicting all the considered pathologies.

4.6 Conclusions and outlook

This work has introduced and discussed a novel interaction between Computational Fluid Dynamics (CFD) and Machine Learning (ML). The key conclusion is that CFD-computed information may harbour extremely informative features, and may thus be useful to ML in the execution of classification and regression tasks.

The problem of interest is the flow in the human nose, and the classification of anatomic pathologies, in view of clinical decisions concerning functional surgery of the human upper airways. In this context, the strategic objective is learning to automatically discriminate physiological inter-subject anatomic variations from variations related to a pathological condition. Two major difficulties in this endeavor consist in the large anatomical variability which exists across healthy noses, and in the cost of obtaining the large number of annotated observations that is typically required to train ML algorithms. The latter issue, in particular, renders the standard approach of building a deep neural network (which in principle could learn directly from anatomies) highly impractical.

We have shown that an alternate solution strategy is possible: when using features extracted from the flow field computed with CFD, the training of a neural network becomes substantially easier in comparison to equivalent networks that rely on geometry-based features. The non-linearity of the Navier–Stokes equations, together

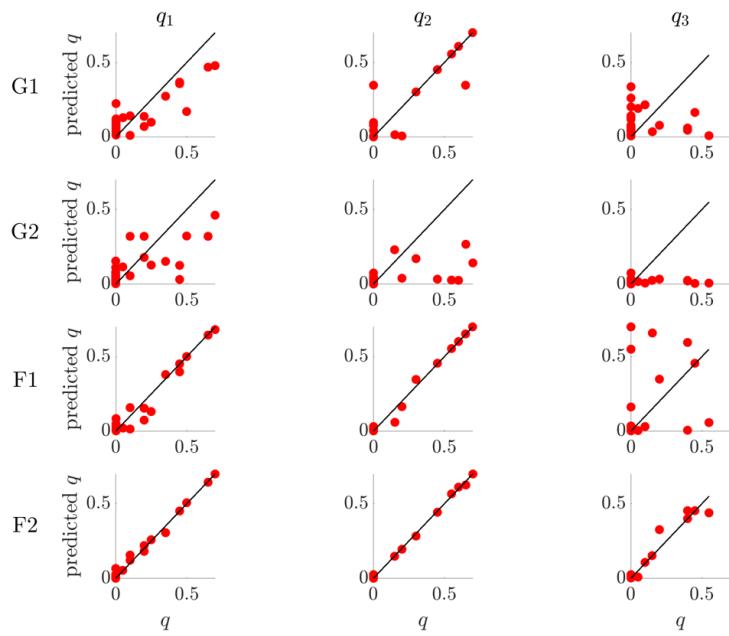


Figure 4.10: Performance of the various features (rows) in one experiment for predicting the three pathological parameters (columns). Ground truth on the horizontal axis, and predicted value on the vertical axis.

with the convective nature of the flow, is such that extracting significant information from the flow field is more effective than looking at the small anatomical changes that are behind that information. While a very large number of annotated CT scans (hence, a large amount of purely geometrical information) could in principle lead to a successful ML classification procedure, relying on the computed flow field is an interesting and effective alternative whenever, as in the medical field, annotated CT scans are difficult to obtain, and thus necessarily available in limited quantity. Albeit the present model only considers localized geometrical deformations, it is found that CFD-based features outperform both small-scale geometrical features like G1 and the large-scale ones conveyed by G2.

The present work and the ensuing conclusions are obviously limited by the extreme simplification of the anatomical model, and by the corresponding low-fidelity CFD approach, based on RANS simulations only. It is important to keep in mind that this work does not aim at introducing a clinically usable tool: in a realistic setting, the full parametrization of the entire geometry would be impossible, and alternative ways for representing pathologies would be needed. However, thanks to the careful design of the reference nose and of its pathologies, which are clinically significant, we are confident that the main conclusions are robust and will continue to apply even when the underlying anatomies become more realistic or, eventually, are derived from CT scans. The results of this work are motivating our ongoing research efforts (Schillaci *et al.*, 2022) for the classification of nasal pathologies, where the nose model is substituted with real patient-specific anatomies. Additional difficulties are encountered, like e.g. the need to avoid a full parametrization of the anatomy, but the present results support the design of a procedure based on CFD-computed features. At the same time, these conclusions may be of general interest, and pave the way to the use of fluid mechanical features as input to improved ML methods.

Chapter 5

Inferring Functional Properties from Fluid Dynamics Features

5.1 Abstract

In a wide range of applied problems involving fluid flows, Computational Fluid Dynamics (CFD) provides detailed quantitative information on the flow field, at variable level of fidelity and computational cost. However, CFD alone cannot predict high-level functional properties that are not easily obtained from the equations of fluid motion. In this work, we present a data-driven framework to extract these additional information, such as medical diagnostic output, from CFD solutions. This is a challenging task because of the huge data dimensionality of CFD, and the limited training data that can be typically gathered due to the large computational cost of CFD. By pursuing a traditional Machine Learning (ML) pipeline of pre-processing, feature extraction, and model training, we demonstrate that informative features can be extracted from CFD data. Two experiments, pertaining to different application domains, support our claim that the convective properties implicit into a CFD solution can be leveraged to retrieve functional information that does not admit an analytical definition. Despite the preliminary nature of our study and the relative simplicity of both the geometrical and CFD models, for the first time we demonstrate that the combination of ML and CFD can diagnose a complex system in terms of high-level functional properties.

5.2 Introduction

Computational Fluid Dynamics (CFD), i.e., solving the differential equations of the fluid motion with the aid of a digital computer, plays a crucial role in a large number of applications, ranging from industry to health. Nowadays CFD is relied upon as much as (sometimes more than) the traditional wind-tunnel testing, and its accuracy (determined by the amount of discretization as well as by the models employed) can be increased at will, provided the computational cost remains affordable.

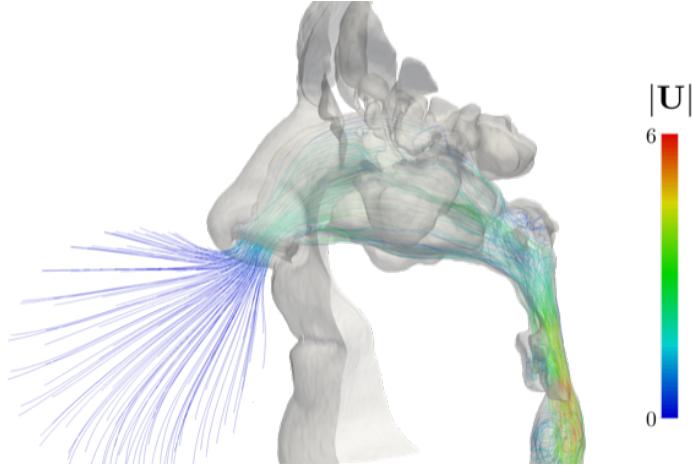


Figure 5.1: CFD solution of the airflow in the upper human airways during an inspiration: streamlines are colored with the magnitude of the local velocity in m/s .

Quite often, however, the final goal of the CFD analysis, i.e. the diagnosis of the system, remains elusive: the ultimate information that is relevant for the end-user might not be directly provided by the CFD itself, or might not be expressed as a function of the CFD solution. In particular, the complex interplay between fluid dynamics and the geometry of interest, prevents us to formulate (and solve) the design of the best geometry as an optimization problem involving CFD outcomes. Illustrative cases exist in the medical domain (Choi *et al.*, 2016), and we consider the diagnosis of Nasal Breathing Difficulties (NBD) as a running example. NBD represent an extremely widespread pathological condition of the human upper airways and often requires corrective surgery: a precise diagnosis is troublesome and the failure rate of surgery is up to 50% (Sundh & Sunnergren, 2015; Illum, 1997). A detailed CFD solution for the nasal airflow for a specific patient (see figure 5.1), is certainly important and useful to diagnose NBD, but *per se* it does not help the surgeon to make a rational decision as to whether and how to perform a specific surgical manoeuvre. Several other similar examples could be made, ranging from flood control in rivers, to aerodynamics in the transport sector, to a large number of industrial problems such as probe placement in wind tunnels. In fluid dynamics, the strong non-linearity of the governing equations makes a small geometrical detail potentially result in significant flow changes far away (for example a small imperfection on the wing surface can compromise the aerodynamic performance of the entire aircraft). On the other hand, a large geometrical modification sometimes leads to little or no consequences (for example a large deviation of the nasal septum may be compatible with normal breathing). The diagnosis of these complex systems can benefit from CFD outcomes, as for instance to determine whether and where to perform surgery, where to best prevent coastline erosion, where to optimally place a probe. The answers to these questions are indeed contained within and dictated by the CFD-computed flow field, but an analytical link between the flow field itself and the required information is not available. We believe that pattern recognition techniques

(Bishop, 2006) and data-driven models in particular (Hastie *et al.*, 2001) have a large potential in this relatively unexplored class of problems.

Using data-driven models on CFD data is particularly challenging for several reasons. First and foremost, there is a dimensionality problem: CFD invariably leads to large data sets, which are costly to produce and difficult to analyse. Such a huge amount of data is not amenable to be directly handled by Machine Learning (ML) models. To set the stage, we mention that a simple two-dimensional CFD simulation of the time-averaged flow field around an airfoil – i.e. a basic configuration of aeronautical interest addressed with the simplest of the CFD approaches – requires the discretization of space into no less than 10^6 cells. Since several flow variables (two velocity components, pressure, auxiliary turbulence variables) are computed for each cell, a single CFD simulation easily produces hundreds of Megabytes of data. This figure grows by orders of magnitude when three-dimensional configurations are considered, and/or higher-fidelity simulations are used. Furthermore, it is very difficult to gather large training sets of annotated simulations, due to their large computational cost and the difficulty of gathering a representative set of experts' decisions in domains such as medicine.

Here we propose a ML methodology to diagnose a complex system whose physics is governed by fluid dynamics. The class of problems we consider relies on the ability of the flow field to convey information, especially of the geometrical type, from an *a priori* unknown location to a predetermined sensing location. Crucially, the success of this endeavour hinges upon the convective properties of the flow. In particular, we aim at using data-driven models to arrive at important information that cannot be computed via the simulation itself, such as a diagnostic output in medicine. We identify and describe low-dimensional features that can be realistically extracted from CFD data and then used in a ML pipeline. These features, namely the field values measured at predefined locations or streamlines arrival time, will be demonstrated to be effective in two different application scenarios, where they enable accurate inference of the target variable even with rather small training sets. Since, to the best of our knowledge, no CFD dataset on parametric geometries is publicly available to date, we develop two case studies in distant application domains: studying the airflow in the human nose, and the airflow around a two-dimensional section of an airplane wing. The airfoils dataset is publicly available for download at <https://doi.org/10.5281/zenodo.4106752>.

For reasons related to the computational cost of creating the database, both experiments are quite simplified in terms of geometric and CFD models, without compromising the validity of the ML procedure. Both problems share an identical structure, insofar as the interest lies in retrieving non-local information (pathological anatomic anomalies of the airways, or shape characteristics of the profile) from simple features extracted from the computed flow field.

5.3 Related work

In the last 5-10 years, the application of ML to fluid mechanics has bloomed. This is witnessed by the quantity and quality of the published material. Recent researches and authoritative surveys can be found in Kutz (2017); Brenner *et al.* (2019); Duraisamy

et al. (2019); Brunton *et al.* (2020); Raissi (2018). Most often ML is used to model fluid equations using CFD as input (or, equivalently, physical realizations of a flow), and expecting fluid mechanical quantities as output. Hence ML models are often used to predict the complex input-output relationship typical of fluid flows governed by highly non-linear equations. To the best of our knowledge, however, there is no previous work that shares our goal of inferring quantities that cannot be computed by the CFD itself.

A clearly identifiable strand of work aims at improving turbulence models (Duraisamy *et al.*, 2019), which is needed in CFD approaches where the small-scale details and the unsteady behaviour of a turbulent flow cannot be computed. Indeed, a universal and accurate turbulence model is still lacking. Recent developments are leading to bound uncertainties in existing turbulence models via physical constraints and to adopt statistical inference to characterize the empirical coefficients of existing models. Among the several examples, Ling *et al.* (2016) were the first to employ a deep neural network to enforce a correction to the popular Spalart-Allmaras RANS turbulence model (Spalart & Allmaras, 1992), by embedding the required Galilean invariance into the model-predicted tensor of the turbulent stresses. Along similar lines, Wang *et al.* (2017) used random forests to identify large discrepancies in model-based turbulent stresses. Fukami *et al.* (2020) applied supervised ML to solve a number of regression problems for reconstruction and estimation. Example applications were the estimation of time-varying force coefficients and flow reconstruction from a limited number of sensors. They also considered convolutional neural networks for super-resolution, training the ML model with direct numerical simulations to extract key features from the training data.

Another class of works attempts to bypass the use of the differential equations that govern the fluid motion to get rid of the simulation stage altogether. For example, a physics-informed deep-learning framework was developed (Raissi *et al.*, 2020) to learn the velocity and pressure fields from the flow visualizations; it shows potential also for biomedical applications, in cases where quantitative measurements are unavailable. Srinivasan *et al.* (2019) illustrated the potential of neural networks to predict the dynamical evolution of a simple model of a temporally-evolving turbulent shear flow, training multilayer perceptron and long short-term memory networks.

It is important to note that in all the aforementioned works the fluid dynamics quantities are used as both input *and* output of the ML algorithm. In other words, ML is typically used as a surrogate of the Navier–Stokes governing differential equations, either to speed up or replace the computation, or to improve the turbulence modeling required by CFD.

5.4 Problem Formulation

The output of a CFD simulation is a set of scalar or vector fields defined over a domain $\Omega \subset \mathbb{R}^3$ which in CFD always undergoes discretization, for example into many small volumes or a computational mesh. These fields are obtained by solving the discretized Navier–Stokes equations (sometimes in a simplified form supplemented by a turbulence model) together with boundary conditions applied at the geometrical boundary $\Gamma \subset \mathbb{R}^3$. For instance, for the human nose, Γ includes the internal geometry of the nasal cavities

extracted from the CT scan of the patient, as shown in figure 5.1.

A CFD simulation results in several output fields, which in general are also time-dependent. However, the present work only considers time-averaged quantities, in particular the vector field of the mean velocity $\mathbf{U}(x, y, z)$ and the scalar field of the mean pressure $p(x, y, z)$ (expressed in a Cartesian reference system without loss of generality):

$$\mathbf{U}(x, y, z) = \begin{bmatrix} u(x, y, z) \\ v(x, y, z) \\ w(x, y, z) \end{bmatrix}, \quad p(x, y, z). \quad (5.1)$$

All the flow quantities referring to the generic i -th cell resulting from the discretization of Ω can be stacked into a vector $\mathbf{Q}_i \in \mathbb{R}^4$:

$$\mathbf{Q}_i = \begin{bmatrix} u_i \\ v_i \\ w_i \\ p_i \end{bmatrix}, \quad (5.2)$$

where for conciseness $u_i = u(x_i, y_i, z_i)$ being $(x_i, y_i, z_i) \in \Omega$ the cell center. Since the spatial domain Ω is discretized over n cells, which in our elementary case studies is already $n \sim 10^6$, the CFD output is a (very large) matrix $\mathbf{C} \in \mathbb{R}^{4 \times n}$, which contains all the flow quantities in every cell.

Our goal is to train a model \mathcal{K} that predicts a target value Y associated to the matrix \mathbf{C} provided by CFD:

$$\mathcal{K} : \mathbf{C} \mapsto Y. \quad (5.3)$$

The target variable can be either categorical (as for a classifier that identifies the most suitable surgery for NBD), or ordinal/real (as for a regressor that estimates some geometric quantities from Γ). To this purpose, we assume that a training set of l labelled pairs $\{(\mathbf{C}_j, Y_j), j = 1, \dots, l\}$ is provided.

The major challenges to be addressed in our settings are *i*) the large dimensionality of each input (namely large n); and *ii*) the limited number of training samples l , due to the high computational cost of each CFD simulation. To tackle the latter challenge, we opted for a computationally cheap CFD approach – i.e. solving the Reynolds Averaged Navier–Stokes (RANS) equations. The available alternatives would lead to a prohibitive computational cost for dataset generation, even though more accurate results may contain additional important information. RANS equations are fast to solve (around 10-12 computing hours per case in our simple 3D application), but they only provide information on the mean fields.

5.5 Proposed solution

We describe now our approach for training a model and performing inference over the CFD output \mathbf{C} . It consists of a concatenation of rather customary steps of ML pipelines (Bishop, 2006), namely *pre-processing*, *feature extraction*, and *model training*; however, the first two steps are customized to CFD data and are therefore described in detail below.

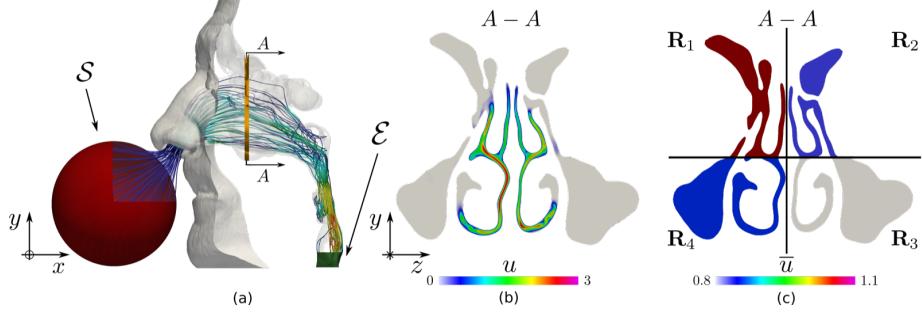


Figure 5.2: Airflow in the human nasal cavities during inspiration. (a) Streamlines start from region \mathcal{S} and end in region \mathcal{E} . The orange slice indicates the cross-sectional cut plotted in panels (b) and (c). (b) Mean velocity component normal to the cross-sectional cut. (c) Division of the plane in 4 regions $\{R_{1-4}\}$, colored with the value of the regional average velocity \bar{u}_k .

5.5.1 Pre-processing

The CFD output \mathbf{C} is first pre-processed to compute *streamlines*. By definition, streamlines are locally tangent to the velocity vector and can be thought of as massless tracer paths. A number of streamlines is drawn connecting a start region $\mathcal{S} \subset \Omega$ to an end region $\mathcal{E} \subset \Omega$. For example, figure 5.2 shows streamlines for the nasal airflow starting from \mathcal{S} , a spherical surface placed in front of the nostrils, and ending at \mathcal{E} , a plane crossing the downstream end of the computational domain, beneath the larynx. Figure 5.3 shows the streamlines pattern for the two-dimensional flow around an airfoil: in this case, \mathcal{S} is a vertical line upstream of the profile and \mathcal{E} is a similar line placed downstream. Streamlines provide a compact view of the flow field in the domain Ω , and can highlight vortical structures, recirculation zones, and high-velocity regions (where the streamlines approach each other).

Each streamline is defined by its tangent, which is locally parallel to the velocity vector \mathbf{U} . Hence, once the velocity field is known, streamlines are computed by selecting s locations over the region \mathcal{S} , and by numerically integrating their trajectory. In detail, we set an initial location for the k -th streamline $(x_0^k, y_0^k, z_0^k) \in \mathcal{S}$ and initialize its velocity as $\mathbf{U}(x_0^k, y_0^k, z_0^k)$. Then, trajectory is integrated until the end region \mathcal{E} is reached; the velocity \mathbf{U} is obtained by linear interpolation out of the mesh nodes.

5.5.2 Feature Extraction

Due to its large size, the CFD output \mathbf{C} cannot be fed to the classifier \mathcal{K} directly. Therefore, we perform feature extraction to dramatically reduce the number of inputs of the classifier, while preserving the information content of the CFD. We propose two kinds of expert-driven features, which are inspired by engineering practice in the analysis of flow fields: distribution of *streamline arrival times* and *regional averages* of flow variables.

Distribution of Streamline Arrival Times

Once the s streamlines connecting \mathcal{S} to \mathcal{E} have been computed, we measure the time required to travel from \mathcal{S} to \mathcal{E} along each streamline at the local mean velocity. The arrival times are then considered as realizations of a random variable with unknown distribution, of which we estimate mean μ_1 and centered moments up to fifth order, i.e. μ_2, \dots, μ_5 . The statistics of the arrival times provide an extremely compact and meaningful description of the flow. For example, streamlines entangled by vortices would take longer to reach \mathcal{E} than straight streamlines; similarly, streamlines passing through highly turbulent regions would result in outliers with respect to the distribution of normal trajectories. Besides arrival times, additional quantities can be extracted from streamlines, e.g. by integrating flow quantities (like velocity or pressure) along the streamlines and computing the sample moments of their distribution.

Features extracted from streamlines are very practical, since they compactly convey flow information while sampling most of the volume Ω with minimal knowledge of the geometry Γ . In fact, only the initial and final regions \mathcal{S} and \mathcal{E} need to be identified: no accurate registration is required for the rest of the surface.

Regional Averages

Other informative features can be extracted by averaging the flow quantities over r pre-defined regions $\mathbf{R}_k \subset \Omega$, $k = 1, \dots, r$. To take into account the uneven layout of the samples in Ω , these averages are volume-weighted. For example, the region-averaged pressure p over region \mathbf{R}_k is referred to as \bar{p}_k and is defined as

$$\bar{p}_k = \frac{\sum_i p_i V_i}{\sum_i V_i} \quad (5.4)$$

where the index i includes all the cells $(x_i, y_i, z_i) \in \mathbf{R}_k$, and V_i denotes their volumes.

Figure 5.2(a) illustrates a thin orange slice ($A - A$) as a meaningful choice for a set $\{\mathbf{R}_k\}$. This coronal section (figure 5.2(b)) intersects large areas exhibiting little or no flow (the paranasal sinuses), as well as narrower areas delimited by the turbinates, where most of the flow rate is concentrated. Figure 5.2(c) shows how this section has been divided into four regions ($k = 1, \dots, 4$), with the color indicating the computed value \bar{u}_k in each region.

Information conveyed in regionally-averaged features obviously depends on whether the set of selected regions $\{\mathbf{R}_k\}$ is meaningful. The selection of these regions might not be straightforward in the medical domain, where \mathbf{R}_k typically refers to landmarks that cannot be detected automatically or that require sophisticated registration procedures to align the input surface Γ with a common reference where regions can be defined.

Regional averages mimic procedures often used in wind-tunnels measurement campaigns, where probes like hot-wire anemometers or Pitot tubes are placed in the flow beforehand. Our experiments suggest that averages over a few significant regions in Ω might be discriminative enough to solve our inference problems.

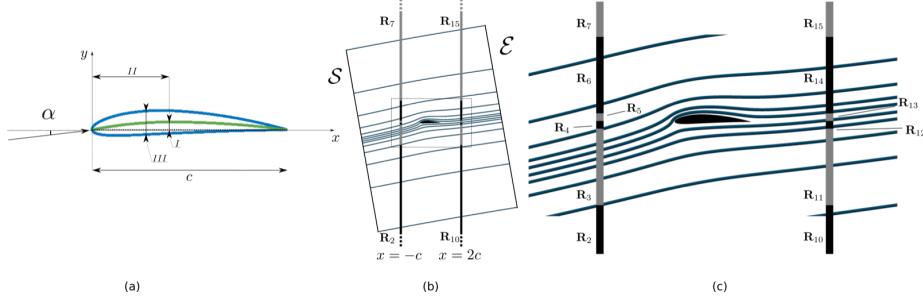


Figure 5.3: Flow field around an airfoil at incidence (flow is from left to right). (a) Sketch of the airfoil, indicating chord c (the segment connecting the leading edge to the trailing edge), angle of incidence α formed between chord and free-stream velocity, the leading edge at $x = 0$, and the trailing edge at $x = c$. The green line is the camber line. First number of NACA code: maximum camber I . Second number of the NACA code: position II of maximum camber along the cord. Third number of the NACA code: maximum thickness III . (b) Streamlines connecting start region S to end region E , with part of the regional sets \mathbf{R}_k (which in a two-dimensional case reduce to lines). (c) Zoom around the airfoil. Smaller regions around $y = 0$ like $\{\mathbf{R}_{4-5}\}$ can be appreciated.

5.5.3 Model training

The pre-processing and feature-extraction steps map the output $\mathbf{C} \in \mathbb{R}^{4 \times n}$ of each CFD to a feature vector $\mathbf{f} \in \mathbb{R}^m$, which stacks m features being either the streamline moments or the regional averages of velocity and pressure. Overall, we expect $m \ll 4 \times n$, so that these two steps yield a substantial reduction in the dimensionality of the problem. Depending on the nature of the target variables, any classifier or regressor \mathcal{K} can be trained from the set of labeled feature vectors $\{(\mathbf{f}_j, \mathbf{Y}_j), j = 1, \dots, l\}$. In the experiments described below, we adopt Neural Networks trained to perform regression over the space of target variables and we show that a limited number of features is often enough to provide very accurate predictions.

5.6 Experiments

We describe two experiments to show that a handful of informative features are sufficient to infer quantities that cannot be computed directly from a CFD simulation. To demonstrate the flexibility of the method presented in section 5.5, the two case studies belong to distant application domains: prediction of geometrical parameters of an airfoil (subsection 5.6.1) and prediction of the severity of an anatomical anomaly of a human nose (subsection 5.6.2). From a fluid-dynamic perspective, the two case studies are far away from each other: the airfoil case is two-dimensional and involves an external fully turbulent flow, in which the inertia forces dominate. The human nose case is three-dimensional and involves an internal, mostly laminar or transitional flow. However, in both cases the goal is to retrieve geometrical information from far away CFD data.

The numerical simulations are carried out with OpenFOAM (Weller *et al.*, 1998), a popular open-source C++ CFD toolbox. We choose the most simple and computationally affordable CFD approach by solving the Reynolds-Averaged Navier–Stokes (RANS) equations using the Spalart-Allmaras (Spalart & Allmaras, 1992) turbulence model to generate the airfoil dataset, and the $k - \omega$ SST turbulence model (Menter, 1994) to generate the human nose dataset.

Overall, the best features are found to be the regional averages, with accuracy varying according to the distance between $\{\mathbf{R}_k\}$ and the geometry of interest. Table 5.1 shows that in the airfoil dataset, the overall accuracy exceeds 95% when the regional sets are not too far away from the profile.

5.6.1 Prediction of Geometrical Features of an Airfoil

Dataset and task

We consider a popular family of airfoils four digit NACA (National Advisory Committee for Aeronautics). Our goal is to train a multivariate regressor \mathcal{K} to predict the NACA numbers, i.e. the shape of the airfoil itself, starting from the CFD solution.

The shape of a NACA airfoil is described by their four-digits code, which corresponds to three integer numbers, and the length of the chord c (see figure 5.3 a). The first number in the NACA code corresponds to the first digit (integer, [0-9]) and quantifies the maximum camber of the airfoil in units of $c/100$; the second number corresponds to the second digit (integer, [0-9]) and locates the point of maximum camber along the chord measured from the leading edge, expressed in $c/10$; the third number has two digits (integer, [05-50]) and quantifies the maximum thickness of the airfoil expressed in $c/100$.

The two-dimensional CFD domain Ω is centered on the airfoil and has a radius larger than $500c$; the angle of incidence α (figure 5.3 a) is set at 10 degrees, the free-stream velocity is 30 m/s . A database of CFD solutions is built by considering 3025 different combinations of digits, hence 3025 different airfoil shapes.

Feature Extraction

Streamlines connecting \mathcal{S} to \mathcal{E} are shown in figure 5.3 (a). \mathcal{S} is a straight segment of length $10c$ orthogonal to the free-stream velocity, whose center is $3c$ distant from the leading edge; \mathcal{E} is identical to \mathcal{S} with center shifted $3c$ downstream from the trailing edge. Along \mathcal{S} , the streamlines starting points (x_0^k, y_0^k, z_0^k) are non-uniformly spaced, with finer spacing towards the center, as shown in panel (c) of figure 5.3.

To extract region-averaged flow quantities, 24 regions $\{\mathbf{R}_k\}$ are selected, consisting of eight portions of three vertical lines drawn perpendicular to the airfoil chord. The first eight segments for $1 \leq k \leq 8$ lay on a vertical line placed at $x = -c$ upstream of the airfoil; eight segments for $9 \leq k \leq 16$ lay on a vertical line placed $1c$ downstream (figure 5.3 b,c), and the eight segments for $17 \leq k \leq 24$ lay on a vertical line placed $10c$ downstream the airfoil trailing edge. On each segment, the regions are symmetrically placed with respect to $y = 0$, and their boundaries have y coordinates

Features	I		II		III		a [%]
	$ e $	σ	$ e $	σ	$ e $	σ	
μ_{1-5}	0.24	1.16	0.41	1.16	0.89	11.54	60.79
\bar{p}_{1-8}	0.04	0.30	0.06	0.29	0.03	0.16	99.34
\bar{v}_{1-8}	0.04	0.30	0.06	0.21	0.04	0.39	97.45
\bar{p}_{9-16}	0.06	0.16	0.11	0.31	0.06	0.16	96.39
\bar{v}_{9-16}	0.03	0.32	0.04	0.13	0.04	0.54	99.47
\bar{p}_{17-24}	0.15	0.37	0.27	0.70	0.15	0.70	86.25
\bar{v}_{17-24}	0.15	0.43	0.26	0.60	0.12	0.29	85.71

Table 5.1: Interpolation experiments for the airfoil dataset. Training set dimension for regional averages: 484, training set dimension for streamlines: 2000.

of $[-500, -10, -1, -0.1, 0, 0.1, 1, 10, 500]$. Note that in figure 5.3 the most rearward segment and the regions farthest from the profile are not displayed.

Model Training and Performance Assessment

We train a three-layers neural network to estimate the three numbers in the NACA code. This is a regression network with 3 output neurons, one per each number of the NACA code. Since the estimated numbers are not necessarily integers, they are rounded to yield the output code. We adopt different splitting criteria in training and test set, considering both interpolation (table 5.1) and extrapolation (table 5.2). As a figure of merit, we primarily consider $|e|$, the mean absolute error over each estimated code and also the classification accuracy a , the percentage of correctly estimated codes.

k-fold cross-validation experiment

The goal of this experiment is to identify the most informative features and assess our regression performance when varying the dimension of the training set. Features are initially grouped according to classical fluid dynamics practices, and are then selected by performing a 5-fold split over the whole training set. In particular, we select three sections, up and downstream the airfoil, where to extract features from pressure and velocity measures.

Table 5.1 shows the mean absolute error $|e|$ and the standard deviation σ for each NACA number, as well as the classification accuracy a . The network is trained by minimizing the mean square error of the estimated NACA numbers. When 8 regionally-averaged flow features are used with a training set of only 484 airfoils, the neural network achieves very small absolute errors and an overall accuracy between 85% and 99% on the NACA code (cfr. last column of table 5.1). The relatively large range in accuracy suggests that some regional averages are more informative than others. In particular, regions closer to the airfoil like $\bar{p}_{1-8}, \bar{v}_{1-8}$ (located at $x = -c$) and $\bar{p}_{9-16}, \bar{v}_{9-16}$ (located at $x = 2c$) achieve higher prediction scores than those further away (like $\{\mathbf{R}_{13-18}\}$ placed at $x = 11c$). This is not surprising since all the flow variables become more uniform as the distance from the airfoil increases: thus, spatial information conveyed by

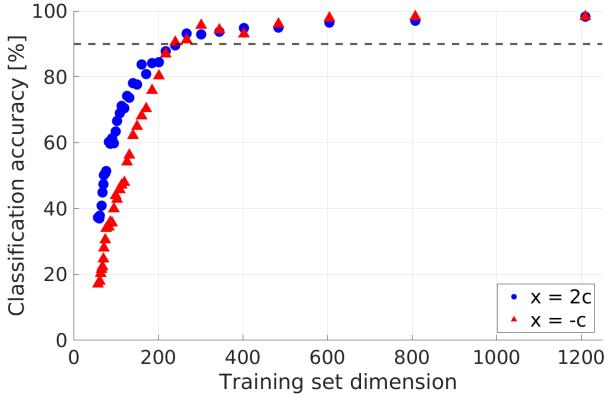


Figure 5.4: Classification accuracy a versus dimension of the training set, with features \bar{p}_{1-8} , measured on regions at $x = -c$, and \bar{p}_{9-16} measured at $x = 2c$

each flow variables decreases with the distance from the airfoil. The statistical moments of arrival times provide fairly good predictive capabilities too, especially for the first number in the NACA code. Even with a training set of 2000 airfoils, $|e|$ is relatively low for the second and third NACA numbers.

Based on these results, we restrict to regional average features extracted from pressure for studying how the performance varies as a function of the training set size. The above experiment is repeated by progressively reducing the training set size, to investigate how this solution would perform when – owing to their computational cost – only a few CFD simulations are available for training. We split the dataset into N equal segments and separately perform training and testing on each segment through a 5-fold cross validation. This procedure allows us to reliably compute the standard deviation of the regression error.

Figure 5.4 illustrates the accuracy of the network classification as a function of training set size, with features \bar{p}_{1-8} and \bar{p}_{9-16} , and indicates that about 300 training samples are enough to achieve 90% accuracy. This plot confirms that, at least when the training set is small, features located downstream ($x = 2c$) are more informative than those upstream ($x = -c$) at the same distance.

Extrapolation

In this experiment we assess the model performance at predicting NACA numbers that are out of the range of training samples. All the entries corresponding to a subrange of the third NACA number, which has a range of 05–50, are removed from the training set. In the first experiment, we test the range 30–40, and in the second experiment, we skip an internal subrange testing 05–15 and 45–50. Every experiment is carried out 5 times, to average the results. Table 5.2 shows that the first extrapolation experiments are very close to the previous k-fold cross validation tests, even though the training set is four times larger than in the k-fold cross-validation case (table 5.1). Little

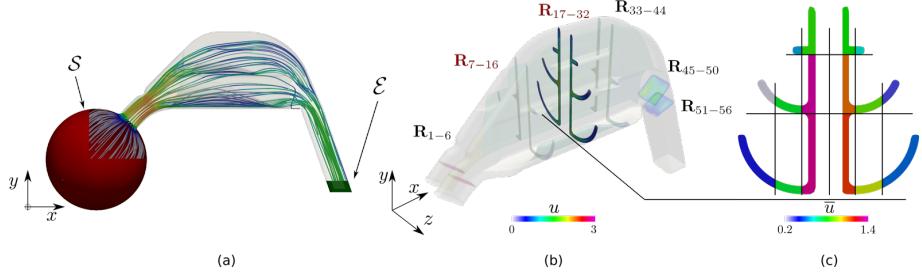


Figure 5.5: Simplified model of the human nose. (a) CAD geometry which excludes paranasal sinuses, and placement of regions \mathcal{S} and \mathcal{E} . (b) Cut planes for regional averages. Pathologies, if present, are applied in the space between the sections highlighted in red. (c) Regional averages of the x velocity component in the region set $\{\mathbf{R}_{17-32}\}$.

difference is observed when velocity or pressure are chosen as a feature, with the far downstream regions at $x = 11c$ consistently performing slightly worse than the other two regions. The second experiment is obviously more extreme. Velocity seems to be more informative than pressure as a feature. The far regions at $x = 11c$ lead to worse performance than the others placed closer to the airfoil.

Features	I		II		III		a [%]	
	$ e $	σ	$ e $	σ	$ e $	σ		
Inner	\bar{p}_{1-8}	0.03	0.04	0.07	0.08	0.03	0.04	99.97
	\bar{v}_{1-8}	0.03	0.06	0.06	0.15	0.03	0.05	98.36
	\bar{p}_{9-16}	0.06	0.08	0.08	0.12	0.06	0.08	99.79
	\bar{u}_{9-16}	0.06	0.08	0.08	0.15	0.05	0.08	98.83
	\bar{v}_{9-16}	0.05	0.08	0.07	0.45	0.05	0.08	98.83
	\bar{p}_{17-24}	0.1	0.14	0.19	0.28	0.1	0.15	92.83
Outer	\bar{v}_{17-24}	0.08	0.12	0.16	0.24	0.09	0.12	95.34
	\bar{p}_{1-8}	0.55	1.75	1.22	3.63	1.54	6.37	76.92
	\bar{v}_{1-8}	0.12	1.00	0.14	0.68	0.19	1.30	95.27
	\bar{p}_{9-16}	0.69	2.47	1.58	5.34	1.52	6.06	75.24
	\bar{u}_{9-16}	0.21	1.02	0.27	0.82	0.29	1.66	85.45
	\bar{v}_{9-16}	0.15	0.92	0.20	1.15	0.19	1.04	92.42
	\bar{p}_{17-24}	1.72	5.81	4.39	15.88	1.54	5.27	52.33
	\bar{v}_{17-24}	0.41	1.29	0.81	2.45	0.41	1.32	61.82

Table 5.2: Extrapolation experiments for the airfoil dataset. Training set dimension: 2400

5.6.2 Prediction of pathologies in a simplified human nose

Dataset and task

Figure 5.5 (a-c) illustrates the simplified model used to build the CFD database for the human nose. This model replicates all the essential features of a human nose as represented in figure 5.1 and 5.2, but at the same time involves a CAD-based simplified shape which, for example, does not include paranasal sinuses. A key advantage of the simplified CAD model is its parametrization, which is used to implement controlled variations of the basic anatomy. The CFD dataset has been created by defining and modifying 7 geometrical parameters of the baseline model. These parameters mimic anatomical variations observed by Ear-Nose-Throat (ENT) doctors in their clinical practice. In particular, four of them result in "healthy" anatomical alterations of the human noses, namely that ENT doctors deem not to affect the normal breathing function. The other three parameters mimic pathological conditions at different levels of severity. These are the anterior hypertrophy of the Inferior Turbinate, the hypertrophy of the whole Inferior Turbinate, and the hypertrophy of the anterior head of the Middle Turbinate. These parameters affect the shape between the sections labeled in red in figure 5.5 (b). While our CAD model is certainly overly simplified compared to a CT scan of the human nose and the variety of pathologies, the size of the CFD simulations is instead comparable to those that can be derived from a CT scan. To take into account anatomical variability of human noses, the CFD dataset is generated from 200 unique combinations of these 7 parameters. We address the task of estimating the three *pathological* parameters, and to this purpose we train a neural network having 3 hidden layers and 3 output neurons.

Feature Extraction

The regional averages are computed over the sections shown in figure 5.5 (b). The six cross-sectional planes are perpendicular to the mean flow and are further subdivided in several regions (from 6 to 16 each, depending on the surface area). The results of the experiments are reported in table 5.3, in terms of streamlines arrival time and pressure regional averages. Most of the regional averages achieve a small regression error, such as for $\{\mathbf{R}_{17-32}\}$ which lie on a cross plane that directly "sees" the modification of the turbinates. Since, as in the airfoils case, regional averages of velocity are found to perform similarly to regional averages of pressure, they have not been reported.

Model Training and Performance Assessment

In a real scenario, there is of course no guarantee to know data directly from the region where the patient's pathology is present, since this is *a priori* unknown. Thus, the most significant results in table 5.3 are those concerning features extracted from regions far from where the pathological alterations have been applied. For example, the regional averages from regions $\{\mathbf{R}_{33-44}\}$, have a mean absolute error varying between 0.0185 and 0.0570 mm, considering that the severity of the pathologies varies with a step of 0.05mm, it is a good result. Obviously the error is expected to be smaller when CFD

information is extracted right from the regions where the pathology is present: for this reason, these values are greyed out in table 5.3.

This demonstrates that the ML algorithm is actually able to make accurate predictions, taking advantage of the fluid dynamic ability to transport information along the flow. Indeed the regions close to the throat such as $\{\mathbf{R}_{45-50}\}$ and $\{\mathbf{R}_{51-56}\}$ still produce rather low inference error, taking into account how far these are from the position where pathological alterations have been introduced. In comparison, streamlines arrival times do not achieve good performance, with an error of over 18.56 mm in a reference domain of 12 mm . The hypertrophy of the head of the Middle Turbinate appears to be more difficult to predict; most likely, this is due to the fact that the Middle Turbinate is interested by a smaller fraction of the global flow rate, hence its influence on the overall flow is smaller.

5.7 Conclusions

We have demonstrated that ML can effectively predict functional properties of complex fluid mechanical systems, when the knowledge of the flow field does not immediately provide required high-level diagnostic information. We exploit the convective properties of the fluid flow by identifying a small set of informative features extracted from CFD simulations, which provide accurate predictions of geometrical information. The required training sets are relatively small: this is an extremely important characteristic, owing to the large cost of CFD and the difficulty in gathering annotated data from experts, especially in domains such as medicine.

The flexibility of the proposed approach is demonstrated by dealing with two rather simplified examples, pertaining to applications as diverse as industry and health: the airflow around wing sections (where the goal is the prediction of the airfoil type) and the airflow within a model of the human nose (where the goal is to predict pathological anatomic deformation leading to breathing difficulties).

We identify two types of features that are potentially very informative and reconcile the massive dimensionality of a CFD dataset within a ML pipeline. One hinges upon the reconstruction of streamlines in the flow field and the integration of flow quantities

Features	Inf. Turb. Head		Inf. Turb. Body		Middle Turb. Head	
	$ e [\text{mm}]$	σ	$ e [\text{mm}]$	σ	$ e [\text{mm}]$	σ
μ_{1-5}	4.478	5.581	18.556	22.214	6.008	7.3033
\bar{p}_{1-6}	0.113	0.181	0.083	0.140	0.087	0.1307
\bar{p}_{7-16}	0.023	0.038	0.012	0.022	0.020	0.038
\bar{p}_{17-32}	0.017	0.028	0.014	0.023	0.031	0.047
\bar{p}_{33-44}	0.019	0.032	0.019	0.032	0.057	0.099
\bar{p}_{45-50}	0.034	0.056	0.014	0.026	0.064	0.110
\bar{p}_{51-56}	0.038	0.060	0.018	0.029	0.072	0.119

Table 5.3: Interpolation experiments for the human nose dataset. Rows corresponding to features extracted on a pathological section are greyed out

along them. The other consists in averages of fluid dynamic variables over suitable regions in the flow field. Their relative merit has been assessed, with regional averages performing better than streamlines, although this is deemed to depend on the type and quality of the CFD analysis. In fact, the steady nature of the CFD simulation used here fails at providing the streamlines with the information required to successfully solve the addressed regression problems. This is particularly apparent in the human nose, where streamlines computed using RANS simulation differ much from the true ones. In contrast, streamlines are more informative in the airfoil scenario, since the flow is essentially steady. We believe that the use of unsteady CFD on an unsteady problem will unlock their full potential.

Ongoing work concerns designing effective features for addressing real-world medical scenarios, where we plan to combine ML and CFD to infer diagnostic information. Furthermore the construction of a more realistic database, using geometries from CT-scans, is ongoing. In particular, we will use our framework for surgery planning in the ENT domain, where high-fidelity and time-resolved CFD simulations will be used to analyze patient-specific CT scans. A wider target consists in adapting our framework to handle measurements derived from experimental fluid mechanics data. This opens plenty of relevant applications, such as identifying anomalies due to damages or detecting ice formation over airfoils.

Chapter 6

Data Augmentation Based on Computational Geometry for Neural Network Training in Medical Flow Field Classification

6.1 Abstract

Machine Learning (ML) applied to Computational Fluid Dynamics (CFD) has recently blossomed, primarily because of its potential to speed up simulations by replacing numerical solutions of physical equations. However, the dominant use of ML, which is to infer expert labels that cannot be computed from explicit equations, has been less investigated in CFD. One of the major limitations of using ML to classify CFD flow fields is the lack of large and annotated training sets. 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 method to augment a few training data requiring little expert supervision, and an automated procedure to extract CFD-ready surfaces from CT scans. In particular, we leverage computational geometry techniques to synthetically generate training samples by mapping deformation functions (defined by experts over a reference surface) to any healthy individual. This procedure provides a training set characterized by large anatomical variability (since we can leverage different healthy individuals) and unequivocal labels (since we control the transfer of pathologies). Our method allows us to generate a large training set starting from a few healthy individuals, and train a Neural Network (NN) that classifies two pathologies: septal deviation and turbinate hypertrophy. We show that a model trained exclusively on the augmented data can successfully generalize to identify pathologies in real patients.

6.2 Introduction

In recent years, the use of ML in fluid mechanics, and in particular in CFD domains, has grown exponentially, thanks to the improvement in terms of computational power and the potential of deep learning in a number of applications (Brunton *et al.*, 2020; Panchigar *et al.*, 2022). ML is widely used in fluid dynamics to estimate sophisticated non-linear relations between input and output embedded in the equations of motion, speeding up or sometimes replacing numerical simulations. In the context of ML-CFD interplay, a host of works have been proposed: turbulence model improvement by informing models with data available in literature (Duraisamy *et al.*, 2019); reconstruction, estimation, and super-resolution using Convolutional Neural Networks (CNN) (Fukami *et al.*, 2020); physics-informed deep learning aimed to reduce the computational burden of CFD simulations (Raissi *et al.*, 2020), and several others. However, the mainstream use of ML, which is to infer labels that cannot be computed from explicit equations, has been poorly investigated in CFD. In fact, training a neural network (NN) directly taking a CFD input is challenging, as CFD data is difficult to gather and returns a huge amount of data that cannot be directly fed to a ML model. Nevertheless, CFD data can convey unique information to the diagnosis of complex fluid dynamics systems, and for this reason, it is important to design methods to effectively extract information from the flow fields.

In this work, we address the classification of pathologies from their effect on the flow field. The ear, nose and throat (ENT) surgeons face in fact an impressive anatomical variability when diagnosing nasal pathologies and can only make decisions based on visual analysis of CT scans without taking advantage of fluid dynamics information. Fluid dynamics can provide unique information concerning how pathologies affect the flow field. Yet, the native form of CFD data is not immediately interpretable and a tool that analyzes these data and returns useful information to a surgeon would help in increasing the rate of success of diagnoses. In this respect, ML could be the solution.

This work is inspired by the paper by Schillaci *et al.* (2021b), which demonstrates that the combination of ML and CFD enables diagnosing a complex system in terms of high-level functional properties. Experiments in Schillaci *et al.* (2021b) were conducted in two different scenarios: the first concerning the prediction of aeronautical airfoil shapes from the flow field, and the second regarding the classification of shapes (mimicking pathologies) in simplified human upper airways from the internal flow. An important limitation of this latter scenario is the dataset: given the difficulty in gathering large and realistic training sets of annotated simulations in domains such as medicine, the dataset was generated ad hoc for this work using elementary geometrical shapes. Real human anatomies extracted from real CT scans, however, are much more complex and exhibit larger anatomical variability. The generation of a *labeled realistic* training set of CFD data computed in real upper airways is the goal of our work.

The major challenge when training classifiers taking as input CFD flow fields is that CFD data shows impressive variability with respect to small geometric differences, thus ML models trained on CFD data might poorly generalize. The only way to train an accurate model is to collect a large set of annotated flow fields that is sufficiently representative of the problem. To overcome the difficulty in collecting large datasets of CFD simulations, we have designed a data augmentation procedure that enables expanding a

small training to a much bigger one, representing entirely the huge variability that the CFD data shows.

Training Data Augmentation methods (Shorten & Khoshgoftaar, 2019) are the standard solution to deal with scarce data: augmentation techniques artificially increase the training dataset size, usually by either data warping or oversampling. In a CFD-ML framework however, data augmentation is not straightforward: being the CFD data subject to strict physical laws and precise boundary conditions, standard augmentation techniques for images (Yang *et al.*, 2023) could return inconsistent flow fields.

Our solution is to perform data augmentation on the surfaces on which CFD simulations are performed, that is, surfaces extracted from patients' CT scans. The complexity of these 3D surfaces and the difficulty in dealing with them, make the goal of augmenting the training set challenging. We exploit a computational geometry tool as the functional correspondence (Ovsjanikov *et al.*, 2012, 2017) to map a set of deformations functions defined by ENT experts mimicking pathologies on healthy individuals. Our method is to produce a *wide*, *reliable* and *labeled* database of *synthetic* pathological anatomies in a field such as the medical one, where annotations (in our case pathologies) do not always admit an objective and unique definition, but are instead subjective to experts' perspective and costly to produce. Furthermore, to reduce the cost of fluid dynamics simulations, we also use functional correspondence to *clean* the surfaces extracted from CT scans, and remove negligible parts (regions where the flow is so slow that can be neglected) via shape registration (Cosmo *et al.*, 2016) with a reference surface.

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 data-augmentation method, starting from the CT scans of healthy individuals. Overall, 277 surfaces are produced considering different severity levels and locations of pathologies. We then test the classifier trained on this dataset both on synthetic samples and on a set of 10 real patients. We demonstrate that our augmentation method sufficiently generalizes the training set while drastically reducing the man-hours needed to prepare the samples. We can in fact successfully identify pathologies in new patients, although using a training set based on very few healthy individuals.

6.3 Related Work

In the past 5-10 years, the application of machine learning (ML) in the field of fluid mechanics has experienced significant growth. This is evident from the increasing quantity and quality of published material (Brenner *et al.*, 2019; Brunton *et al.*, 2020). Most often, ML is used to model fluid equations using CFD as input and expecting fluid mechanical quantities as output.

One of the mainstream uses of ML for CFD focuses on bypassing the use of differential equations that govern fluid motion using physics-informed neural networks (Raissi *et al.*, 2020; Wu *et al.*, 2022b), eliminating the need for numerical simulations. Another area of work is aimed at improving turbulence models as there is no universal one. Several examples highlight these efforts. Ling *et al.* (2016) were pioneers in utilizing deep neural networks to introduce a correction to the widely used Spalart-Allmaras turbulence model. Fukami *et al.* (2020) used supervised ML models in various

regression problems related to the reconstruction and estimation of flow fields. They also explored the use of convolutional neural networks for super-resolution, training the ML model with direct numerical simulations to extract key features from the training data.

In the aforementioned works, fluid dynamic quantities are used as both the input and output of the NN algorithm, and ML is aimed at improving or speeding up what CFD can already do. Our goal is to use ML to infer quantities that cannot be computed by CFD itself, as this has a great potential in diagnosing complex systems. Unfortunately, this problem has been much less investigated, and to the best of our knowledge, the only attempt is from Schillaci *et al.* (2021b), where the authors demonstrate that the combination of ML and CFD can effectively diagnose fluid dynamics systems in two *toy examples*, requiring few training data. Thanks to our augmentation method, however, we can train ML models to solve real-world problems concerning the diagnosis of systems by their fluid dynamical properties.

Data augmentation on CFD data is a delicate theme, as evidenced by the paucity of papers on the subject. Since CFD flow fields are subject to strict physical laws and precise boundary conditions, the use of rotations and translations as in images would not be sufficiently informative, while other affine transformations would return unrealistic or inconsistent flows. In this context, Wu *et al.* (2022a) proposed an approach based on GANs (Generative Adversarial Networks) to generate synthetic flow field data, which is based on two generators and two discriminators. The first generator is trained with the available data to reproduce flow fields. The second generator is introduced for data augmentation. Once trained, the first generator can synthesize a discriminative flow field for any given flow condition in the design space. Differently, Abucide-Armas *et al.* (2021) proposed a data augmentation technique that considers the similarity principle of fluid dynamics. This technique generates different synthetic cases to increase training and validation data, keeping the Reynolds number constant and guaranteeing the fulfillment of the boundary conditions. In these works, however, augmentation focuses on the flow field, producing consistent flows only in simple applications, which add little information. From the perspective of inferring pathologies from the upper airways of real patients, these techniques are not sufficient to define the effect of pathologies on the internal flows, i.e., these cannot directly augment the functional effect that pathologies have on the flow field. The augmentation method we propose is designed to overcome this problem: we do not augment directly the CFD data, but we act on the geometries in which CFD data is computed, such that we can safely assign a label (i.e., pathology) to each augmented data.

6.4 Problem Formulation

We address the problem of training a model \mathcal{K} that classifies CFD data of the human upper airways to recognize pathologies belonging to a pre-defined label set $\mathcal{Y} : \{y_j\}_{j=1,\dots,M}$. Thus, \mathcal{K} operates as follows

$$\mathcal{K} : \mathcal{F} \mapsto \mathcal{Y}, \quad (6.1)$$

where \mathcal{F} is the domain where the results of each CFD simulation live.

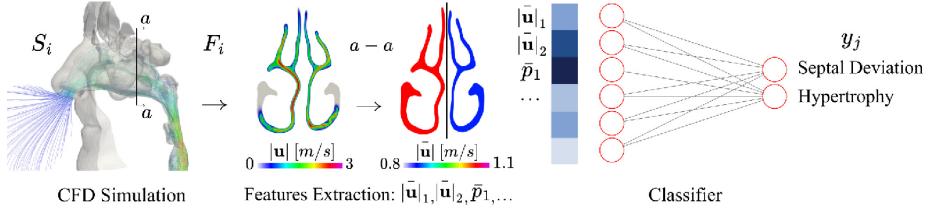


Figure 6.1: Taken with modifications from Schillaci *et al.* (2021b). Scheme of the inference problem. CFD simulations are performed inside the surfaces $\{S_i\}_{i=1,\dots,N}$ producing the CFD flow fields $\{F_i\}_{i=1,\dots,N}$. From each F_i we extract a feature vector, which as recommended in Schillaci *et al.* (2021b), is made of regional averages of CFD quantities computed on transversal sections (e.g., the figure shows the velocity magnitude $|\bar{\mathbf{u}}|$ that is averaged to $|\bar{\mathbf{u}}|$ on the section a-a). These features are used to train and test a classifier in inferring pathologies.

As illustrated in figure 6.1, each CFD simulation F_i is computed from a surface $S_i \subset \mathbb{R}^3$ extracted from the CT scan, and F_i consists of a very large matrix containing 3D coordinates and fluid dynamics quantities. In more detail, each flow field $F_i \subseteq \mathbb{R}^{7 \times n_i}$ is a matrix having n_i columns, where n_i is the number of cells in which the simulation domain Ω is discretized. For our application, n_i can easily range in the order of $n_i \approx 10^7$. The j -th column $F_i[:, j]$ stacks the 3D coordinates and the 4 outputs of the CFD simulation for the j -th cell, namely:

$$F_i[:, j] = [x_j, y_j, z_j, p_j, u_j, v_j, w_j]^T \in \mathbb{R}^7, \quad j = 1, \dots, n_i \quad (6.2)$$

where x_j, y_j, z_j are the coordinates in the \mathbb{R}^3 space, u_j, v_j and w_j are the 3 components of the velocity vector and p_j is the pressure, simulated via CFD.

The *training set* of the model \mathcal{K} consists of *labeled* pairs $\{(F_i, y_i)\}$. Here we address the problem of generating a large training set via augmentation starting from a small set $\{S_i\}_{i=1,\dots,N}$ of surfaces of healthy patients.

6.5 Methodology

Our goal is to design a methodology to generate a labeled and realistic training set containing annotated CFD flow fields, requiring little expert supervision since this is a very expensive and time-consuming task. Our solution is illustrated in figure 6.2 and summarized in Algorithm 1. This consists of defining, together with the ENT experts, a surface $R \subset \mathbb{R}^3$ representing a *reference* healthy nose, and a set of *surface deformation functions* $\{(\delta_j, y_j)\}_{j=1,\dots,N}$. Each deformation δ_j modifies R to obtain a variant R_j^* , which is affected by the pathology y_j . Deformation functions have been defined by ENT experts by manually modifying 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 $\{(\delta_j, y_j)\}_{j=1,\dots,N}$ together with computational geometry techniques to generate a large training set of annotated

Algorithm 1 Our augmentation method.

- 1: **Input:** Reference surface R , set of deformation functions $\{(\delta_j, y_j)\}_{j=1,\dots,M}$, set of CT scans $\{T_i\}_{i=1,\dots,N}$.
- 2: **Output:** feature vectors with their associated labels y_j .
- 3: **for** $i = 1, \dots, N$ **do** ▷ For each CT scan T_i
- 4: Segment surface \tilde{S}_i from CT scan T_i . ▷ CT scan segmentation
- 5: Obtain S_i , the cleaned version of \tilde{S}_i through registration with R . ▷ Preprocessing
- 6: Estimate the mapping $M_i : R \rightarrow S_i$.
- 7: **for** $j = 1, \dots, M$ **do** ▷ For each pathology δ_j
- 8: Obtain $S_{i,j}^*$ by applying the deformation δ_j to S_i via the mapping M_i . ▷ Augmentation via deformation mapping
- 9: Obtain $F_{i,j}$ through CFD using $S_{i,j}^*$ as boundaries. ▷ CFD simulations
- 10: Extract transversal sections from $F_{i,j}$.
- 11: Compute regional averages values of $|\hat{\mathbf{u}}|$, $|\nabla \hat{p}|$, \hat{e} , k . ▷ Features extraction
- 12: Define the feature vector of the classifier.
- 13: **end for**
- 14: **end for**

pathologies starting from a small set $\{\tilde{S}_i\}_{i=1,\dots,N}$ of surfaces acquired from *healthy* individuals. In particular, we assume that each surface from a healthy patient can be seen as a deformed variant of R , and we adopt functional maps (Ovsjanikov *et al.*, 2012) to establish point-to-point mappings between R and each \tilde{S}_i . This mapping is used for two purposes: *i*) preprocess each surface \tilde{S}_i to obtain a cleaned surface S_i , automatically cutting off negligible parts that can corrupt CFD simulations (Algorithm 1, line 5), and *ii*) automatically apply deformation δ_j to the cleaned surface S_i (Algorithm 1, lines 7 - 13), obtaining $S_{i,j}^*$, 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 for all the pathologies, we obtain a set of surfaces $\{S_{i,j}^*\}_{i=1,\dots,N; j=1,\dots,M}$ with their associated labels. We use these surfaces to run $N \times M$ CFD simulations $\{F_{i,j}\}_{i=1,\dots,N; j=1,\dots,M}$ representing the training set of our classifiers.

Our method guarantees diversity in CFD data by promoting anatomical variability from healthy individuals, at the same time we can easily collect a set of CFD simulations whose annotations are unambiguous and automatically defined. It is also important to emphasize that our method allows for combining different pathologies and controlling their severity, noticeably widening the training set size. Each step of our methodology is described in detail below.

6.5.1 Selection of the reference surface R

The reference surface R is one of the most important components of our method. It should be considered as a reference for the other surfaces in terms of anatomical features and the absence of geometric anomalies. To define the reference surface R , two otolaryngologists reviewed and diagnosed a database of CT scans provided by ASST

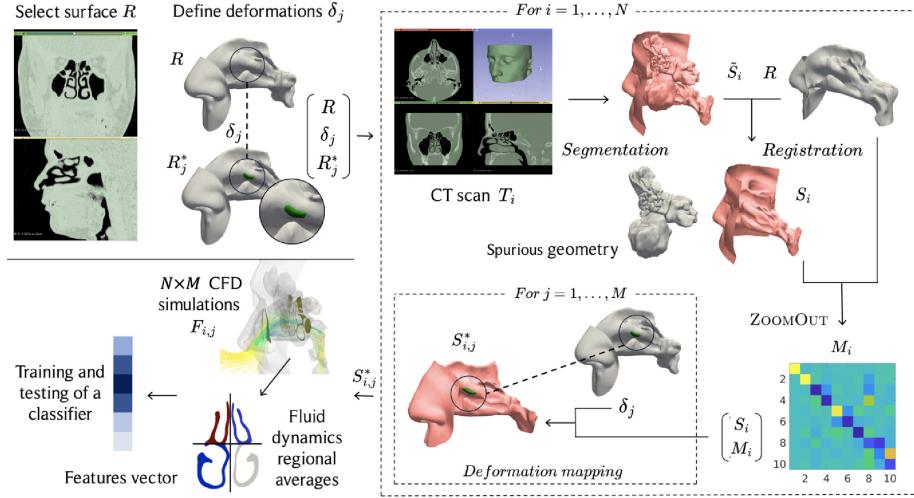


Figure 6.2: A scheme illustrating our methodology. At the top-left corner, we illustrate two preliminary steps, which consist of selecting a reference nose R and defining the collection of deformation functions $\{\delta_j\}_{j=1,\dots,N}$. The deformation function δ_j transforms R into a variant affected by the pathology y_j . On the right side of the figure, we show the procedure for generating the synthetic pathological surfaces $S_{i,j}^*$. For each of the N CT Scans from healthy individuals T_i , we extract the surface \tilde{S}_i by thresholding the CT Scan data. We then register R over \tilde{S}_i to clean \tilde{S}_i from spurious parts, i.e., the paranasal sinuses, obtaining surface S_i . The bottom-right corner shows the compact representation of mapping M_i . M_i associates at each point of R a point of S_i and enables the transfer of functions δ_j from R to S_i producing $S_{i,j}^*$, the variant of S_i affected by pathology δ_j . We then assess the performance of our method by running CFD simulations on each $S_{i,j}^*$ and using CFD data to train and test a classifier, as shown in the bottom-left corner.



(a) Paranasal sinuses are visible in green.

(b) Surface without paranasal sinuses.

Figure 6.3: The figure shows in green the paranasal sinuses which are the parts we aim to remove. On the left-hand side the *complete* geometry extracted from the CT scan. On the right-hand side, the sinuses are removed from the geometry.

Santi Paolo e Carlo in Milan, and found the CT scan of a patient that is unambiguously diagnosed to have no pathologies and to show a well defined CT scan of high quality and resolution.

As described in section 6.5.3, the surface extracted from the CT scan contains unnecessary parts for a CFD simulation, namely the paranasal sinuses (depicted in green in figure 6.3), where the flow is so slow that can be neglected (Jin *et al.*, 2006). Removing these parts should reduce the computational time of the CFD simulations. For these reasons, a hand-crafted procedure was performed by experts to make the reference surface R simpler, cutting off the paranasal sinuses.

6.5.2 Definition of Deformation Functions

The main intuition behind our method is to synthetically define surfaces affected by pathology y_j by manually determining a set of surface deformation functions $\delta_j : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ on the reference R . Deformation functions $\{(\delta_j, y_j)\}_{j=1,\dots,M}$ are designed by ENT surgeons. In particular, we asked surgeons to transform the reference nose R into a variant affected by pathology y_j . We refer to the pathological variants of R as R_j^* . We thus obtain deformations functions $\delta_j(x, y, z)$ by manipulating the surface of R and tracking the displacements of all its vertices. $\delta_j(x, y, z)$ associates to each point (x, y, z) of R a vector $(\Delta x, \Delta y, \Delta z)$ that corresponds to the displacement registered while obtaining R_j^* .

We consider two different pathologies: *septal deviation* and *turbinate hypertrophy*, each with different localization and degrees of severity. Whereby, several deformation functions δ_j can correspond to the same pathology δ_j . A deviated septum is a condition in which the nasal septum (the bone and cartilage that divides the nose in half) is bent, and in severe cases, breathing through the nose can be difficult. In turbinate hypertrophy, on the other hand, the tissue on the side walls of the nose is too large, causing nasal obstruction. Figure 6.4 shows a section of the reference geometry R in which the

inserted hypertrophy of the middle turbinate blocks the internal airway. We emphasize that our design choice in defining pathologies allows us to combine them, by simply superposing different deformations on the same surface.

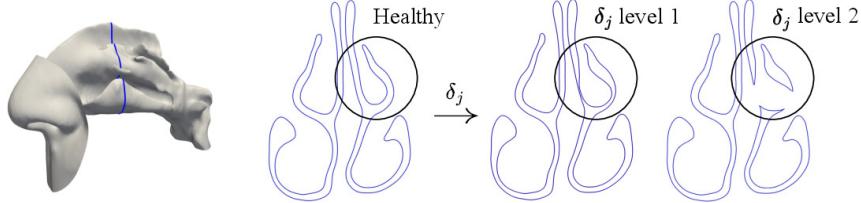


Figure 6.4: Left: section of R used to visualize the pathology. Center: R healthy section. Right: turbinate hypertrophy has been inserted by experts with two levels of severity.

6.5.3 Extraction of the surface \tilde{S}_i from the CT scan

An essential step of our method is to extract the surfaces from the set of healthy CT scans $\{T_i\}_{i=1,\dots,N}$. A CT scan T_i results from a diagnostic imaging technique that uses a combination of X-rays and reconstruction algorithms to get information about the inside of the body. Voxels covered by biological tissue or bones can be rather easily distinguished from empty areas, as the CT scan determines a local value (CT number, expressed in terms of Hounsfield Unit [HU]) proportional to the density of the tissue. The CT data consists of a voxel-reconstructed 3D volume of the scanned object.

The surface \tilde{S}_i that we extract from T_i represents the biological tissue of the upper airways. To extract \tilde{S}_i we *segment* the CT scan data T_i by setting a threshold on the CT number, keeping the surface \tilde{S}_i and excluding internal cavities. The threshold is set to a value that results in a satisfactory discrimination between air and biological material (Quadrio *et al.*, 2016). Some ambiguities can occur in small regions where nasal mucus may accumulate and increase the local density. These small errors need to be evaluated and corrected by experts, making the extraction of surfaces really time-demanding. Our solution reduces the number of times this iteration is performed since it generates multiple examples from the same scan of a healthy individual, therefore we strongly reduce manual corrections, significantly decreasing the time and effort involved. Each surface represents the boundary of the CFD simulations.

6.5.4 Cleaning \tilde{S}_i via registration with R

The manual removal of the sinuses could reduce the computational burden of CFD simulations. However, this could be a very time-consuming procedure as it should be manually performed. Moreover, manual removal should be carried out by experts so that it could become extremely expensive. Our intuition is to *clean* the surface \tilde{S}_i by identifying the sub-portion that best resembles the reference surface R . We do so by a *shape-registration* method based on *functional mapping*, and in particular, we

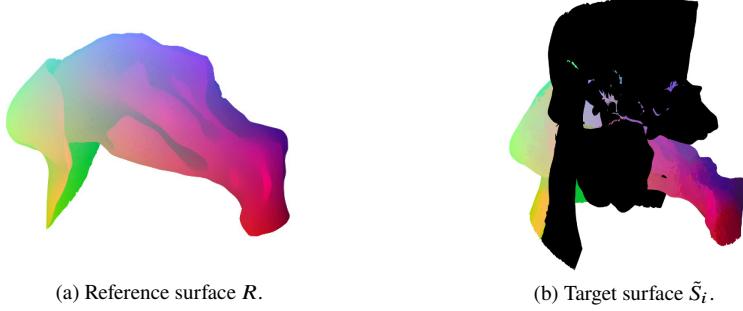


Figure 6.5: A glimpse of the registration of \tilde{S}_i with R . The identified sub-portion of \tilde{S}_i that is closest to R is visible with colors on the right-hand side. The black portion of \tilde{S}_i contains the vertices with 0 mask value. The color allows to see which point are in correspondence in the two surfaces.

refer to the deformable object recognition and dense correspondence in cluttered 3D scenes presented by Cosmo *et al.* (2016). The inputs to the method are the surface \tilde{S}_i , which includes the paranasal sinuses (green parts in figure 6.3), and the reference R which instead does not. Our idea is that the paranasal sinuses can be considered as clutter when matching \tilde{S}_i with R . Thus, we use a robust registration technique for deformable surfaces that exploits a point-to-point mapping and determines the subset of the \tilde{S}_i that corresponds to the same anatomy of R . By this registration, we obtain S_i , the identified sub-portion of \tilde{S}_i , and a mapping that encodes the correspondence between \tilde{S}_i and R . The matching parts are represented as binary indicator functions on the respective surfaces, called segmentation masks. As R is the source, no vertices have to be eliminated, while the mask of \tilde{S}_i takes 0 value for points that have to be cut off. A glimpse of the registration is shown in figure 6.5, where vertices whose mask value is null are depicted in black, and vertices whose mask value is unitary are shown in color. The following paragraph gives an idea of what is functional mapping and how shape-registration works.

Background on Functional Correspondence and Shape-registration Shape matching poses a significant challenge in computer graphics and geometry processing. Its goal consists of determining a correspondence M among points on two 3D shapes. This task becomes notably complicated when the shapes undergo non-rigid deformations. We head to this intricate scenario: we aim in fact to estimate the matching between two separate yet similar surfaces, i.e., the reference R , and each surface S from a healthy patient, assuming that each surface S can be seen as a deformed variant of R .

We tackle this problem with the approach proposed by Ovsjanikov *et al.* (2012), based on functional mapping. Functional maps are constructed upon a basis $\Phi = \{\phi_i\}$ that spans a subset of the functional space of each 3D shape, such that a function f can be expressed as $f = \sum_i a_i \phi_i$, where a_i are the projections of f on Φ . Given a pair of shapes R and S , and the respective two basis Φ and Ψ , the core intuition of

the functional map framework is that instead of estimating directly M , we search for correspondence among functions defined over the shapes expressed in terms of Φ and Ψ . Then, once the functional correspondence C has been estimated, we extract the corresponding point-wise map M . This approach allows functional correspondences to be encoded in a linear transformation defined by a compact matrix denoted as $C = [c_{ij}]$. Namely, given two corresponding functions f and g defined over R and S (i.e., such that $g = M(f)$, where $M : R \rightarrow S$ is the point-wise mapping between R and S), we can write that:

$$\mathbf{b} = C \cdot \mathbf{a}, \quad (6.3)$$

where $\mathbf{a} = [a_i]$ and $\mathbf{b} = [b_i]$ are the projections of f and g on the bases $\Phi = \{\phi\}_i$ and $\Psi = \{\psi\}_j$ defined respectively on R and S . The selection of the basis proves to be a key factor for the optimization problem, and the majority of approaches have opted for the eigenfunctions of the Laplace-Beltrami operator to define functional bases. Matrix C is usually estimated solving an optimization problem that takes as input landmarks, corresponding segments, and descriptors as corresponding functions defined on the shapes to be matched, and finding the version of C that best preserves the functional constraints (in the least square sense). The simplest method for recovering C is to solve the following optimization problem:

$$C = \arg \min_X \| XA - B \|^2 + \alpha \| \Lambda^S X - X \Lambda^R \|^2, \quad (6.4)$$

where A and B are the matrices whose columns contain the coefficients of the functions to be matched expressed in the basis Φ or Ψ , Λ^R and Λ^S are diagonal matrices of eigenvalues of Laplacian operator, and α is a scalar weight parameter. The second term in equation 6.4 accounts for the commutativity of C with respect to linear operators. The dimensions of matrix C corresponds to the number of the basis functions within each shape, and in practice, we consider only the first l functions of the bases, meaning that matching two shapes simplifies to the estimation of a $l \times l$ matrix C . This framework does not only provides an efficient representation of the shape correspondence but is also a practical means to transfer functions from one shape to another. We can transfer function f defined over R to S by using the point-wise map $M : R \rightarrow S$, as $f_S : f(M^{-1}(S))$.

The setting described above provides a satisfactory estimate of matrix C when the two shapes to be matched are rather similar. However, in our case one of the two shape, R , is matched only to a part of S , and Problem 6.4 does not return good result. S contains in fact a big number of outliers for the matching process, i.e., the paranasal sinuses that are visible in green in figure 6.3. We therefore address a variant of Problem 6.4 proposed by Cosmo *et al.* (2016), called shape-registration, to recognize the sub-portion of S that best resembles the shape of R . This method proved to be effective in the case of non-rigid deformations and in the presence of spurious geometry (for us, the paranasal sinuses visible in green in figure 6.3) or missing parts. Together with C , the registration returns the segmentation masks u and v , two binary functions that identify the matching parts between R and S . The optimization problem is regularized by several terms and priors concerning the structure of C , u , and v .

In the case of complex surfaces, the matching process of corresponding functions can become extremely computationally intensive. For this reason, in section 6.5.5, we

resort to the method presented by Melzi *et al.* (2019) called ZoomOUT, to compute the mapping M between R and S , the key element that allows us to map deformations δ_j from R to S . ZoomOUT introduces at each iteration additional frequencies of the Laplace-Beltrami eigenfunctions, adding samples in the spectral domain to improve the point-wise mapping. This method takes an initial $l_0 \times l_0$ low resolution map C_0 or a point to point map M_0 , and extends it to a new map C_l of size $(l+1) \times (l+1)$ by iteratively converting the point-wise map M_l in C_l , and recovering M_l . At each iteration the value l is updated to $l+1$. Instead of solving problem 6.4, ZoomOUT approaches the following one:

$$M_l(p) = \arg \min_q \| C_l(\Psi(q))^T - \Phi(p))^T \|_2, \forall p \in R, \quad (6.5)$$

where $\Phi(p)$ and $\Psi(q)$ denote respectively the p^{th} and the q^{th} rows of the matrix of eigenvectors Φ defined on R , and Ψ defined over S . Problem 6.5 returns a point-to-point map $M : R \rightarrow S$.

6.5.5 Mapping Deformations from R to S_i

As described in section 6.5.2, ENT surgeons define a collection of deformation functions $\{\delta_j\}_{j=1,\dots,M}$ over the reference surface R . Each deformation function δ_j corresponds to a single pathology and is manually defined by asking the ENT surgeons to transform the reference nose R into a variant affected by pathology y_j . The main contribution of our work is the transfer of deformation δ_j over the surface of different healthy individuals S_i , via correspondence with R . On the one hand, this approach is very practical, since deformation functions $\{\delta_j\}_{j=1,\dots,M}$ do not need to be defined over each healthy individual. On the other hand, mapping deformation functions requires establishing a point-to-point correspondence between R and S_i , which is far from being simple, given the deformable nature of these surfaces as well as the intrinsic differences between R and any target patient.

We define the point-to-point correspondence $M_i : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ between the reference R and the target S_i , which associates a point of S_i to each point of R , such that $S_i = M_i(R)$. We estimate M_i using a functional mapping (Ovsjanikov *et al.*, 2012), a computational geometry technique that allows registration of non-rigid surfaces. More specifically, we resort to ZoomOUT, the iterative scheme presented in Melzi *et al.* (2019) to achieve robust matching estimates. ZoomOUT takes as input R and S_i , and returns the high-resolution mapping M_i iteratively refined using a spectral up-sampling technique. More details can be found in paragraph 6.5.4.

Once the point-to-point mapping between R and S_i has been defined, it is possible to apply the deformations δ_j to any surface S_i from a healthy patient, even though δ_j is defined on R . In fact, for $i = 1, \dots, N$ and for $j = 1, \dots, M$ we can compute the deformed variants $S_{i,j}^*$ as:

$$S_{i,j}^* = S_i + \delta_j(M_i^{-1}(S_i)). \quad (6.6)$$

This equation describes the mapping process, defined for each deformation δ_j , to bring a function defined on the surface of R to the surface S_i . In practice, each point (x, y, z) on the surface S_i , thus $S_i(x, y, z)$, is modified by summing a vector that corresponds

to how δ_j would modify the corresponding point in R , namely, $M_i^{-1}(S_i(x, y, z))$. By applying this procedure for each δ_j and each S_i , we generate a set of $N \times M$ annotated surfaces in which we can perform CFD simulations. Moreover, to further extend the training set, we linearly combine forms of the same pathology with different locations and severity, namely:

$$S_{i,(j_1,j_2)}^* = S_i + \alpha\delta_{j_1}(M_i^{-1}(S_i)) + \beta\delta_{j_2}(M_i^{-1}(S_i)), \quad (6.7)$$

where α and β controls the intensity of the deformations. We stick to unitary values for α and β , as the severity level of our pathologies is directly defined by experts. We stress that our method generates a wide set of annotated CFD flow fields requiring little expert supervision, drastically reducing the time that experts would spend in both preparing the surfaces and diagnosing CT scans.

6.5.6 CFD Simulations

The numerical simulations that return the flow fields inside $S_{i,j}^*$ are carried out using OpenFOAM (Weller *et al.*, 1998), a popular open-source C++ CFD toolbox. We choose a high-fidelity CFD approach by performing Large Eddy Simulations (LES). LES return filtered quantities obtained by applying a spatial low-pass filter to the flow field. This means that any scalar field χ can be split up into a filtered $\hat{\chi}$ and sub-filtered χ' (denoted with a prime) portion, such as $\chi = \hat{\chi} + \chi'$, where the hat represents the filtering operation. The filtering process is crucial, as it distinguishes between the turbulent flow's large-scale motions and small-scale motions. The main objective of LES is to accurately capture large-scale turbulent structures while approximating smaller turbulent scales. The difference between the actual velocity and the filtered velocity represents the subgrid-scale (SGS) velocity. This SGS turbulence is addressed through various models to consider the impact of unresolved small-scale motions on the resolved large-scale motions. Among these, we adopt the WALE (Wall Adapting Eddy Viscosity) (Ducros *et al.*, 1998) turbulence model. We run the simulations on a 0.65 seconds inspiration (Islam *et al.*, 2020), setting initial and boundary conditions on the velocity, pressure, and eddy viscosity, and fixing the flow rate at the inlet to 16 l/min (in accordance with the work by Wexler *et al.* (2005)). Since the meshes contain a large number of cells (around 12 million), we run each simulation in parallel using 96 cores and 160 GB of RAM. The simulations are performed with the same setting for all synthetic samples generated with our method and the test data, i.e., a set of new pathological and never-seen-before individuals. Figure 6.6 shows how and where we enforce boundary conditions: we position a sphere around the nostrils to reproduce a *closed mask*. The sphere, together with the surface $S_{i,j}^*$, creates a closed volume containing the domain of the CFD simulation of the internal flow. We can identify three regions on the surface, corresponding to the sphere (in red), the throat (in green), and the internal surface of the nasal airways (in blue). On these, we set various boundary conditions, namely:

Sphere : fixed inlet flow rate for the velocity, zero-gradient for the pressure and the eddy viscosity.

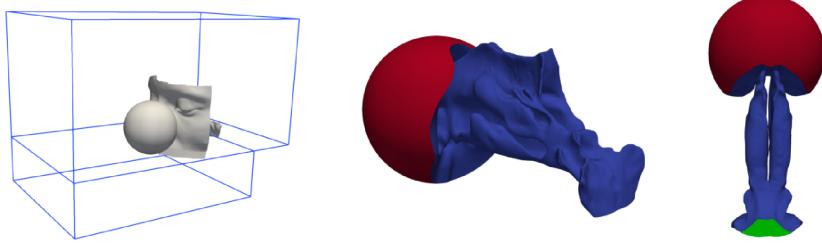


Figure 6.6: Visualization of the simulation setup. On the left-hand side the sphere closing the nostrils is visible, as well as the bounding box of the simulation domain. At the center and on the right-hand side, the regions are shown in colors: in red, the sphere; in blue, the side walls; in green the throat.

Internal surface : non-slip, non-penetration for the velocity, zero-gradient for the pressure and the eddy viscosity.

Throat : zero-gradient for the velocity and the eddy viscosity, reference zero value for the pressure.

6.6 Experiments

To demonstrate the effectiveness of our augmentation method, we train different classifiers exclusively using CFD data $\{F_{i,j}\}_{i=1,\dots,N;j=1,\dots,M}$ computed on a set of surfaces $\{S_{i,j}^*\}_{i=1,\dots,N;j=1,\dots,M}$ augmented by the proposed method. According to the research conducted in Schillaci *et al.* (2021b), which was centered on excessively simplified nose models, from each surface $S_{i,j}^*$ we extract 6 transversal sections having right and left-hand sides. Then, we compute the regional averages of some CFD quantities that define the feature vectors to be fed to the classifier. We assess the performance in two different settings: *i*) testing on synthetic samples generated by the proposed augmentation procedure, adopting a Leave One Patient Out Cross Validation (LOPO-CV), and *ii*) testing on a set $\{D_i\}_{i=1,\dots,10}$ of new *pathological* and *never-seen-before* patients. Each D_i undergoes the same procedure for surface extracting (section 6.5.3) and cleaning (section 6.5.4), and then the same CFD simulations as the healthy surfaces. We show that, despite the fact that $\{S_{i,j}^*\}_{i=1,\dots,N;j=1,\dots,M}$ is augmented from very few patients, our augmentation method sufficiently generalizes the classifiers, which are able to effectively identify pathologies on real patients.

6.6.1 Training set generation

We generate the training set by applying our method to the CT scans of 6 healthy individuals and then performing CFD simulations for each surface $S_{i,j}^*$. The CT scans of these patients come from a database made available by ASST *Santi Paolo e Carlo*

and have been diagnosed as *healthy* by ENT experts. CT scans have also been checked for good quality and resolution before applying our augmentation procedure.

We consider 2 pathologies, i.e., *septal deviation* and *turbinate hypertrophy*. For each pathology, ENT experts manually defined a set of deformation functions $\{(\delta_j, y_j)\}$ as described in section 6.5.2 in different locations and with two levels of severity, for a total of 6 forms of septal deviation and 9 of turbinate hypertrophy. Moreover, to further extend the training set, we combined forms of the same pathology with different locations and severity, generating 12 combinations of septal deviations and 18 combinations of turbinate hypertrophy. Eventually, we augment the initial set of 6 individuals to a set of 277 synthetic samples, on which we perform CFD simulations as described in section 6.5.6.

6.6.2 Features extraction

Due to its large size, CFD flow fields cannot be directly fed to a NN. Therefore, we perform feature extraction to reduce the dimension of the input vector. Henceforth, all the quantities have to be considered as *averaged in time*. The features used within this work are inspired by engineering practice in the analysis of flow fields, namely *regional averages* of time-averaged fluid dynamics quantities. Regional averages have been proved to be effective (Schillaci *et al.*, 2021b) in classifying pathologies in simplified human airways. In each surface $S_{i,j}^*$, we identify 6 transversal sections, the first and the last representing the end and the beginning of the olfactory region respectively, and the other 4 as equally spaced between the previous two (see figure 6.7). We then compute the regional averages of the following fluid dynamics quantities:

- i) $|\hat{\mathbf{u}}|$, the module of the time-averaged filtered velocity vector, such that:

$$|\hat{\mathbf{u}}| = \sqrt{(\hat{u}_x^2 + \hat{u}_y^2 + \hat{u}_z^2)}. \quad (6.8)$$

The filtered velocity offers understanding of the macroscopic flow patterns and can be used to analyze turbulence characteristics on the resolved scales.

- ii) $|\nabla \hat{p}|$, the module of the time-averaged filtered pressure gradient, namely:

$$|\nabla \hat{p}| = \sqrt{(\nabla_x \hat{p}^2 + \nabla_y \hat{p}^2 + \nabla_z \hat{p}^2)}. \quad (6.9)$$

This value delivers comprehension of how pressure changes over space considering the influence of resolved-scale flow patterns, offering information about acceleration, convergence, divergence, energy transfer, and the interaction between pressure and velocity fluctuations.

- iii) $\hat{\epsilon}$, half of the module squared of the vorticity $\hat{\omega} = \nabla \times \hat{\mathbf{u}}$, namely:

$$\hat{\epsilon} = \frac{1}{2} (\hat{\omega}_x^2 + \hat{\omega}_y^2 + \hat{\omega}_z^2). \quad (6.10)$$

This feature contains information regarding the rotational behavior of the fluid flow on the resolved scales. The filtered vorticity can reveal regions where the flow might be undergoing instability due to high vorticity gradients, indicating potential formation or breakdown of turbulent structures.

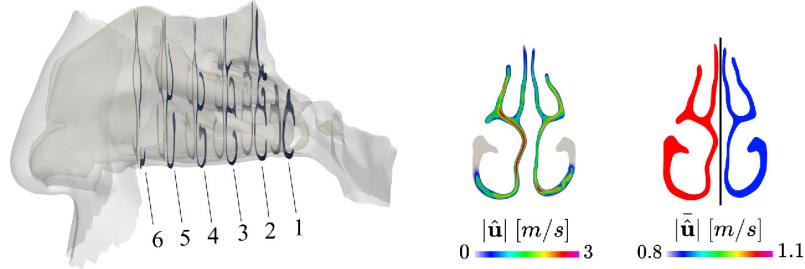


Figure 6.7: Sections used during the features extraction. The left-hand side shows the positions of the 6 sections we considered. On the right-hand side, the module of the velocity is averaged on the left and right half-sections.

iv) k , half the sum of squared velocity fluctuations, namely:

$$k = \frac{1}{2}(u'_x{}^2 + u'_y{}^2 + u'_z{}^2), \quad (6.11)$$

The apex denotes the time-averaged fluctuating velocity components, namely, the components of the resolved Reynolds stress tensor. k is the resolved part of Turbulent Kinetic Energy (TKE) and contains information regarding the level of turbulence of the flow, representing the energy associated with the resolved-scale turbulent structures.

The quantities listed above are expected to test the classifiers with features of different natures: on one hand, $|\hat{\mathbf{u}}|$ can be related to a geometric feature, as in general, the flow rate Q can be expressed in terms of averaged velocity and sectional area as

$$Q = A \cdot u_{mean}, \quad (6.12)$$

where Q is the fixed flow rate, A is the area of the considered region, and u_{mean} is the regional average of $|\hat{\mathbf{u}}|$ on A . On the other hand, $|\nabla \hat{p}|$, \hat{e} , and k are pure fluid dynamic quantities, normally, they are directly uncorrelated to geometry. These quantities are meant to demonstrate that pure fluid dynamic quantities can effectively be used to identify pathologies through their effect on the internal flow.

The regional average \bar{q} of a generic quantity q on the n -th region \mathcal{R}_n is computed weighting values using the uneven cells area. Namely:

$$\bar{q}_n = \frac{\sum_h q_h A_h}{\sum_h A_h}, \quad (6.13)$$

where index h includes all the cell centers coordinate $(x_h, y_h, z_h) \in \mathcal{R}_n$, and A_h denotes their areas. Regional averages are calculated separately for the left and right regions of each section. Therefore, we boil down each simulation $F_{i,j}$ to 48 regional averages of fluid dynamics quantities, 12 values per each quantity listed above. We then use these values separately or jointly as training data for different classifiers, which we test in section 6.6.4.

6.6.3 Model training

We train 5 different types of models to classify feature vectors made by regional averages of fluid dynamics quantities. 4 classifiers take as input a vector of 12 features extracted from the same fluid dynamic quantity. We refer to these as $\mathcal{K}_{|\dot{\mathbf{u}}|}$, $\mathcal{K}_{|\nabla \dot{\rho}|}$, \mathcal{K}_e , and \mathcal{K}_k . We adopt multilayer perceptions (MLP) as classifiers, and after a simple hyperparameters tuning, we found that the best solution has 4 hidden layers of 60, 40, 20, and 10 neurons respectively, for a total of 4261 trainable parameters. The 5th classifier, which we named \mathcal{K}_{Full} , is a MLP taking as input all the 48 regional averages and contains four hidden layers with 120, 60, 30, and 10 neurons respectively, counting 15291 parameters. The output layer consists of a neuron providing a value between 0 (hypertrophy) and 1 (septal deviation). All neurons implement the ReLU activation function. We trained by optimizing the binary cross-entropy as loss function on our augmented training set, setting a maximum number of epochs to 1000 and an early stopping criterion on the loss function. The 4 feature vectors have been standardized to have zero mean and unitary standard deviation. To evaluate the accuracy, the training set in each epoch is divided into two parts: 85% of the samples are used for training, and the remaining 15% to evaluate the score.

6.6.4 Results

We test the classifiers firstly on synthetic data generated with our method from healthy individuals, adopting a Leave One Patient Out Cross Validation (LOPO-CV), and secondly, on a set $\{D_i\}_{i=1,\dots,10}$ of 10 never-seen-before pathological individuals. These patients have been diagnosed by experts to clearly show either hypertrophy or septal deviations. $\{D_i\}_{i=1,\dots,10}$ contains 5 surfaces showing hypertrophy and 5 showing septal deviation. Each surface D_i followed the procedures described in section 6.5 of segmentation from CT scan, CFD simulation, and feature extraction. Through the LOPO-CV, we aim to demonstrate that the classifiers effectively generalize on synthetic samples (i.e., on CFD data computed with our augmentation method) which differ from the ones used for training. The test on real patients is instead designed to demonstrate that we can identify pathologies on real scans from never-seen-before patients, using a NN trained only with synthetic samples. The scores of the classifiers can be seen in table 6.1 and in figure 6.8.

The LOPO-CV consists of iteratively excluding from the training set all the data generated by our procedure from a single healthy individual, and then training on the remaining ones. The excluded patient is used as test data, and eventually, we average the inference results over all the patients. This experiment shows that the classifiers achieve high classification accuracy, reaching the best score (i.e., the highest achieved performance in the classification task) of 89% accuracy on synthetic data, with a peak of 96% accuracy in identifying hypertrophies. High performance on *synthetic* data, i.e., properly predicting pathologies on the testing patient of the LOPO-CV, denotes that the classifiers accurately learned how deformations δ_j affect CFD data, and that they properly generalize on synthetic data excluded from the training. It is important to note that $\mathcal{K}_{|\dot{\mathbf{u}}|}$, $\mathcal{K}_{|\nabla \dot{\rho}|}$, and \mathcal{K}_e perform effectively in terms of classification accuracy, while \mathcal{K}_k shows the worst result. This can be pointed to the limited correlation between the

Classifier	LOPO-CV	Test on new patients
$\mathcal{K}_{ \hat{\mathbf{u}} }$	89%	7\10
$\mathcal{K}_{ \nabla \hat{p} }$	84%	8\10
$\mathcal{K}_{\hat{\epsilon}}$	84%	8\10
\mathcal{K}_k	66%	6\10
\mathcal{K}_{Full}	86%	8\10

Table 6.1: Scores and accuracy of the classifiers.

geometric deformation associated with pathologies, and its effect on the resolved part of the Reynolds stress tensor, collected in k .

The test on real patients is conducted maintaining the same NN architectures, but differently to LOPO-CV, the classifiers are trained on all the synthetic data and tested on real pathological individuals. This test set a best score of 8\10 correctly classified pathologies. Furthermore, $\mathcal{K}_{|\nabla \hat{p}|}$ and \mathcal{K}_{Full} set a score of 5\5 on hypertrophies. This shows which is the real potential of our augmentation method: given the huge anatomical variability, and considering that we apply our augmentation method to a set of only 6 healthy patients, the good performance of this test demonstrates that our method can successfully generalize to real individuals while training only on synthetic data. This is an important result since we simulate a real-world scenario where ENT doctors are called to diagnose never-seen individuals without preliminary information. Similarly to the LOPO-CV, \mathcal{K}_k shows the worst classification result, thereby confirming that k is less correlated with geometric deformation compared to other fluid dynamic quantities.

Figure 6.8 shows that hypertrophies are generally more accurately identified than septal deviations. Hypertrophies, in fact, usually modify the anatomies causing nasal obstructions that strongly affect the internal flow. Septal deviations, on the other hand, can in some cases consist of a bent nasal septum that does not necessarily show evident differences in the flow field with respect to a healthy individual. Namely, the fluid dynamic features associated with a septal deviated anatomy can be similar to healthy ones, making the classification more prone to errors.

The different classification accuracy shown by our classifiers denotes that some features are more informative than others in this specific task. We stress a key aspect: we are not surprised that $\mathcal{K}_{|\hat{\mathbf{u}}|}$ can accurately classify geometric deformations that directly modify the shape of these sections. As we mentioned in section 6.6.2, in fact, we can directly relate $|\hat{\mathbf{u}}|$ to a geometric quantity, that is the area of the transversal section (recalling that $Q = A \cdot u_{mean}$). However, being $|\nabla \hat{p}|$, $\hat{\epsilon}$, and k pure fluid dynamic quantities, their success in classifying deformations is less obvious. In particular, $\mathcal{K}_{\hat{\epsilon}}$ and $\mathcal{K}_{|\nabla \hat{p}|}$ show satisfactory classification performance, and demonstrate that pathologies convey unique information to the flow field which can be helpful in identifying such pathologies through data-driven models.

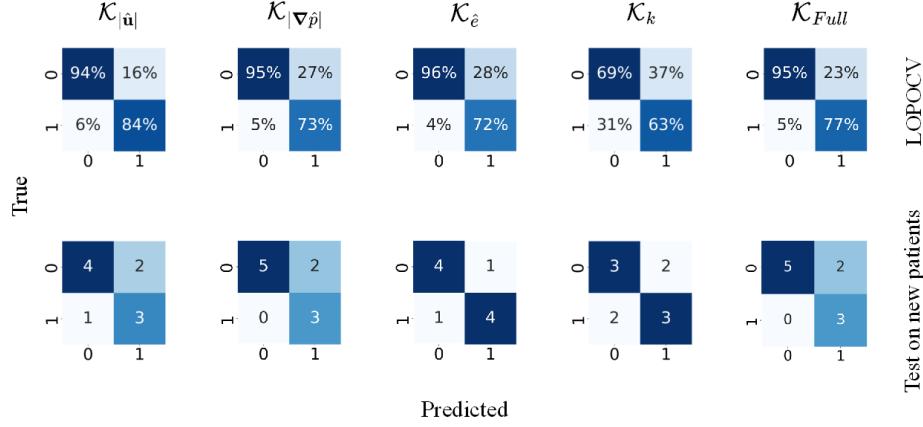


Figure 6.8: Performance of our classifiers in terms of confusion matrices. Top row: the scores of the LOPO-CV. Bottom row: the performance of the test on new patients. 0 and 1 correspond respectively to hypertrophy and septal deviation.

6.7 Conclusions

We demonstrated that we can effectively identify pathologies in real patients by training a NN with CFD data computed by our procedure. Our method effectively generates wide, realistic, and labeled CFD training data, augmenting a set of a few healthy individuals, while requiring little expert supervision. The classifiers trained on augmented data properly generalize, as evidenced by the classification performance that reaches up to 90% accuracy on synthetic samples and 80% accuracy on never-seen-before patients. One of the main strengths of our method is the automatic procedure to prepare surfaces, which strongly reduced manual intervention in processing real anatomies. Indeed, the effort in building the test set of 10 pathological individuals is in proportion significantly higher than building the set $\{S_{i,j}^*\}_{i=1,\dots,N; j=1,\dots,M}$ of 277 surfaces by augmenting the initial 6 patients. Our method holds great potential, as we can now realistically develop a tool, with little effort, that allows otolaryngologists to access fundamental functional information that is directly conveyed by the flow field. This information can be used to improve surgical decisions and thus increase the success rate of diagnoses.

A natural direction of development of this work is to design new and more informative features, such as considering fluid dynamic values computed along streamlines. We could adapt this framework to simpler CFD models without compromising the performance. Another hint to extend this work could be to upgrade the classifiers to deep ones, bypassing the feature extraction procedure. We might also consider using computational geometry tools as functional mapping to encode new geometric features that, together with the fluid dynamic ones, can improve the classification accuracy.

This work opens plenty of relevant real-world applications. As an example, a realistic scenario consists of adapting our framework to deal with measurements derived from experimental fluid mechanics data. One such problem can be the identification of anomalies due to damages or the detection of ice formation over aeronautical airfoils.

We could in fact use our method to generate a database of damaged airfoils or ice-covered airfoils, and try to identify the anomaly from experimental data simulated by CFD.

Chapter 7

Enhancing Machine Learning with Computational Fluid Dynamics

7.1 Abstract

This paper discusses the automatic classification of geometries starting from the surrounding flow field, thus inverting the prevailing approach in which Machine Learning (ML) models are ancillary to Computational Fluid Dynamics (CFD). The specific application is the classification of anatomic pathologies in the human nose; an equivalent, simpler problem presented in parallel is the identification of simple geometrical defects inserted onto NACA airfoils. ML could in principle learn to connect shapes (anatomies) and defects (pathologies) directly, but at the expense of a large number of high-quality labeled data for training; this is difficult to achieve in medicine. Feeding the ML model with informative features mediated by fluid mechanics is more efficient, and reduces the necessary number of observations significantly. The anatomies of few healthy patients are modified with various combinations of well-defined elementary pathologies; a CFD dataset is then built with Large Eddy Simulations. Informative fluid mechanical features are extracted from the flow field, and used to train a neural network and classify pathologies. For both airfoils and noses, results are encouraging, and confirm how highly informative fluid mechanical features can be successful for classification, whenever a large number of annotated geometrical data is not available. The fluid mechanical meaning of the most informative features is discussed.

7.2 Introduction

In the wide realms of science and technology, several physical problems involving moving fluids exist where the design goal of improving the performance of a system requires optimization. The discipline of flow control (see e.g. Gad-el-Hak, 2000,

for a broad review) considers well-defined technological objectives, which can be as diverse as finding the shape of an airfoil with the best lift-to-drag ratio via adjoint techniques (Jameson, 1988), reducing pressure losses in ducts via active or passive control (Wojewodka *et al.*, 2018), or reducing separation on lifting surfaces or behind bluff bodies (Choi *et al.*, 2008; Leschziner *et al.*, 2011).

In several fluid-related situations, however, an optimization problem simply cannot be posed, even though the design objective is self-evident, because a suitable cost function is not readily available. An important example is the vast class of clinical situations related to the human health in which fluid mechanics is important for maintaining or reinstating the well-being of the patient (e.g. breathing, circulation of blood and other biological fluids): the well-being can hardly be translated into a mathematically described cost function. This is a typical application where artificial intelligence in its widest sense has a great potential: in fact, the perspective of machine learning in medicine is bright (Goecks *et al.*, 2020).

Since one decade or so, the interest for the combined application of Computational Fluid Dynamics (CFD) and Machine Learning (ML) techniques has raised. We do not attempt to provide here a mini-summary, as several excellent and up-to-date expert reviews are already available (see, among others, Brunton *et al.*, 2020; Brunton, 2021; Vignon *et al.*, 2023). Important research threads exist where neural networks are being designed to mimic, replicate, or improve the solution of the Navier–Stokes equations (Kim *et al.*, 2021; Kochkov *et al.*, 2021); to improve data-driven solutions by using laws of physics as an additional constraint (Raissi *et al.*, 2019); to develop data-driven turbulence models (Ling *et al.*, 2016; Duraisamy *et al.*, 2019; Liu *et al.*, 2020); to construct a reduced-order modelling for unsteady flows (Hasegawa *et al.*, 2020); even to extract the governing equations of a dynamical system (Brunton *et al.*, 2016).

The present work introduces a new approach that uses ML to solve a classification problem, in which information extracted from (standard) CFD simulations are used to improve the ability of a ML model to extract information, by leveraging the non-linearity of the Navier–Stokes equations that lie behind the CFD solution. Although geometry itself already contains the required information to address the classification problem, we posit that the CFD solution computed for that geometry enables distilling such information into a more self-evident state. The title of the present contribution reverses on purpose that of the work by Vinuesa & Brunton (2022), i.e. "Enhancing Computational Fluid Dynamics with Machine Learning", and emphasizes our stance: CFD is employed here as a powerful feature extractor tool to render the ML problem more tractable, by exploiting the flow solution, governed by the non-linear Navier–Stokes equations which act as a filter on to the geometry.

The approach is presented in the context of a specific health-related application, i.e. the flow within the human nose, which is a long-standing research effort of the authors. The aim is to support the medical doctor with surgical decisions regarding the frequently necessary corrections of the nasal breathing difficulties which affect so many of us, because of pathologies or malformations of the upper respiratory airways. The various functions of the human nose (transport of mass, but also air warming, conditioning, humidifying and filtering, and guiding the odorant particles towards the olfactory mucosa) are primarily driven by fluid mechanics, and the convoluted, extremely variable shape of the nasal cavities determines the functioning of the nose.

However, the link between shape and function is not clearly assessed yet (Doorly *et al.*, 2008a). The Ear, Nose and Throat (ENT) surgeons face an impressive inter-subject anatomical variability, and, in a standard clinical setting, they take surgical decisions mostly relying on the visual analysis of the patient's anatomy, acquired by a CT scan. Since extreme anatomies sometimes happen to be asymptomatic, while other apparently "normal" anatomies lead to severe symptoms, it is well known to ENT doctors that the problem of describing the functionally average nose is far from being solved (Brüning *et al.*, 2020). The clinical path leading to surgical decisions is often quite subjective, and many surgical maneuvers simply do not achieve the expected goal: an impressive example is the surgical correction of septal deviations, where up to 50% of the patients report dissatisfaction after surgery (Illum, 1997; Sundh & Sunnergren, 2015).

The sole diagnostic tool that contains fluid mechanics information is rhinomanometry (Clements & Gortds, 2005), a non-invasive exam which provides an objective measurement of the nasal resistance, i.e. the ratio between the pressure difference (measured between the external ambient and the pharynx) required to create a certain flow in the nose and the flow itself. Unfortunately, when the patient reports stuffed nose and troubled breathing (i.e. insufficient nasal patency), it is well known (Radulesco *et al.*, 2019; Barbarite *et al.*, 2021; Na *et al.*, 2022) that the cause is often way less trivial than too large a nasal resistance. For example, the so called empty nose syndrome (Houser, 2007) is a recognized iatrogenic disorder in which nasal resistance lowered by surgery is accompanied by a subjective sense of poor nasal breathing.

Great hopes were recently spurred by the functional information obtained by CFD. Indeed, literature over the last 10–15 years witnesses a surge of CFD studies (Leong *et al.*, 2010; Zubair *et al.*, 2012; Quadrio *et al.*, 2014; Radulesco *et al.*, 2019; Singh & Inthavong, 2021) that start from the CT scan of the patient, and reach a complete characterization of the flow field. Unfortunately, bar some self-evident cases where CFD is hardly necessary, often times the outcome of a CFD analysis, accurate as it might be, is simply unable to assess whether or not the patient is healthy, where a pathology is located, and what surgery needs to be performed. The fundamental reason for such failure can be ultimately traced to the lack of a reference healthy nose, and to the ensuing inability to conceive a suitable cost function (based on some distance from the healthy nose) to express the goal of the optimization in mathematical terms.

In such scenario, a ML model might, in principle, be able to create the correct link between the anatomy of a specific patient and the possible presence of a pathology, its location and its severity. However, because of the complexity of the nasal anatomy and of the possible deformations, together with the huge variability of healthy anatomies, this would require a (very) deep neural network, whose training would necessitate a large number of properly annotated cases. Such a training set does not currently exists, and is not expected to exist in the near future, owing to the cost of highly informative and properly labelled data in medicine. In fact, so far ML has seen no use in rhinology, if exception is made for an attempt at automated classification of CT images of the paranasal sinuses (Chowdhury *et al.*, 2019; Crowson *et al.*, 2020).

In this paper, we intend to leverage patient-specific CFD to improve the feasibility of the process, by bringing the required number of annotated samples down to an acceptable level. A preliminary study (Schillaci *et al.*, 2023), where extremely simplified nose models and a basic computational approach were employed to assess whether CFD

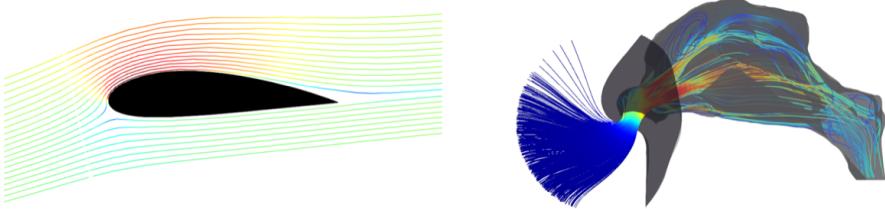


Figure 7.1: Left: RANS solution of a non-defective NACA four-digits airfoil at incidence. Right: LES solution of a healthy nose during inhalation. In both cases, the mean streamlines are plotted, with color (blue to red) encoding the local velocity magnitude.

information can in principle outperform geometric information, motivates the present research, by suggesting that flow information might indeed lead to reasonable prediction accuracy with a database of reasonable size. In the present paper, we describe the construction of a database of a few hundreds large-eddy simulations (LES) for patient-specific nasal anatomies. The database is then used to experiment with a ML pipeline made by hand-crafted feature extraction, selection and classification, in which a simple shallow neural network is eventually trained to classify nasal pathologies. To the best of our knowledge, this is the first reported attempt at automatic classification of shapes of solid bodies immersed in a moving fluid, by using solely information gathered from the surrounding flow field. A few works exist which pursue a not unrelated approach. For example Alsalman *et al.* (2018) classified vortex wakes behind airfoils using information extracted from sensors on the wake, and Wang & Hemati (2019) classified wakes from hydrodynamic signal measurement. In all these cases, the classification applies to the flow field, and not to the type of geometry that produced it. Similarly, Hasegawa *et al.* (2020) trained a reduced-order model for generic bluff-body shapes, aiming at prediction of flow statistics.

For illustrative purposes, the main application of the flow in the human nose is presented alongside with a simpler problem, characterized by the same main features but geometrically and computationally easier to handle: the classification of geometrically defective two-dimensional airfoils belonging to the NACA four-digits family. The different airfoils within the family parallels the anatomical variability of the healthy noses. The CFD setup for an acceptable solution is rather established, and a RANS approach provides the flow solution at moderate computational cost and complexity. The NACA airfoils are given "pathologies", consisting in a set of predetermined "defects", and the CFD solution is used to infer such geometrical anomalies. The two problems possess an evident parallelism: both involve a significant "healthy" geometrical variability, which makes it difficult to detect a pathology. The goal, in both cases, is learning to classify pathologies with CFD quantities as inputs. A sketch of the two problems is provided in figure 7.1.

The paper is structured as follows. The next §7.3 presents the methodology employed to define the baseline geometry and to apply deformations; it also provides details on the meshing procedure and the simulations setup, including boundary condi-

tions and turbulence models. It is then explained how the needed features are defined and extracted, how the classification task is posed in mathematical terms, and what is the architecture of the neural network. Airfoils and noses are presented within an unified framework. The next §7.4 presents the result of several classification experiments, whereas in §7.5 we address the fluid mechanical meaning of various features, and discuss the robustness of the obtained results. The work closes with a discussion and a brief outlook in §7.6.

7.3 Methods

This Section describes the two data sets, i.e. the airfoils and the human nasal cavities, and their usage. The parallelism between them will be emphasized when describing the geometrical aspects of the procedure, the computational approach adopted to obtain the flow field, and the process of feature extraction up to the design of the classifier. In the last subsections, the ML pipeline – which is identical for the two datasets – will be described, by touching upon the structure of the shallow neural network, and by describing how the datasets are partitioned for the experiments.

The output of a CFD simulation is a set of scalar and vector quantities known in discrete form over a spatial domain \mathcal{V} . These quantities are obtained by numerically solving the Navier–Stokes equations, possibly in one of their derivative forms (in this work: the Reynolds-averaged Navier–Stokes equations for the airfoils, and the LES spatially filtered Navier–Stokes equations for the noses; both require turbulence modelling), and are either instantaneous or averaged in time. Since each neuron of a neural network can only deal with scalars as input and output, the velocity vector must be evaluated component-wise. In this work, we opt to only consider vector magnitudes, as this guarantees a rotation-invariant measure (which is a crucial detail once it is realized that the reference frame of the nose anatomies acquired via CT scans varies on a case-by-case basis, see later §7.3.1).

A dataset is built by running ℓ distinct CFD simulations, referred to as instances or observations. Each observation contains the complete field of m scalars (in the present cases $m \leq 4$; the scalars are the velocity components and pressure) in discrete form throughout the computational domain \mathcal{V} , discretized into a number n of cells, or discretization points (in this work, only finite-volumes discretization is considered, hence the volume is discretized into cells). The number n may differ among observations.

The CFD solution is stacked into one big data array X of size $m \times n$. The typical value for the total number of cells n in a CFD simulation is large: here, the two-dimensional RANS of the airfoils have $n = O(10^6)$, and the three-dimensional LES of the noses have $n = O(10^7)$. The output of a single observation is thus very large, and cannot be handed over directly to the neural network.

Our goal is to train a classifier \mathcal{K} that associates the array X from one observation to a (small) set Y of target variables. For classification, the target variables are categorical, i.e. are two (for binary classification) or more (for multiclass classification) labels from a predefined list of geometrical properties (i.e. types of deformation, or nose

pathologies). Formally, the problem is written as:

$$\mathcal{K} : X \mapsto Y. \quad (7.1)$$

This is a standard ML problem, which presents specific difficulties when X is derived from CFD. Since, as mentioned above, the size of X is typically very large (because such is n), the number ℓ of labeled instances is unavoidably much smaller than $m \times n$, and the learning process becomes next to impossible. Thus, we adopt the mainstream approach in ML of reducing the size of the input array, by extracting from X a small (compared to $m \times n$) set f of hand-crafted features, used for the robust training of the classifier. Once f is extracted from X , the classification problem (7.1) is reformulated as:

$$\mathcal{K} : f \mapsto Y. \quad (7.2)$$

7.3.1 The geometry: discretizing \mathcal{V}

The first logical step of the procedure consists in the definition and discretization of the computational domain \mathcal{V} .

Airfoils

The airfoils dataset includes regular (healthy) and geometrically defective (pathological) airfoils. The regular airfoils belong to the family of NACA 4-digits profiles, and are defined by a numerical code which specifies three parameters: the maximum camber (encoded by the first digit), its chordwise position (encoded by the second digit), and the maximum thickness (encoded by the last two digits). Once these parameters are known, the airfoil shape is obtained analytically.

A single two-dimensional airfoil is described by placing one thousand points on its contour, with clustering at the leading and trailing edges. The computational domain surrounding the airfoil is discretized with cells arranged in an O-type grid, created following the procedure described by Sorenson (1980). The grid includes the points defined on the airfoil; the number of cells in the radial direction and their expansion rate are chosen by enforcing the size of the first cell and the radius of the grid, which is about $500c$, where c is the airfoil chord. The typical size of the resulting two-dimensional mesh is about half a million cells. Robust and consistent grids are generated for all the profiles; the parameters of the mesh generator are unchanged across cases.

The healthy geometrical variability in the airfoils dataset is obtained by spanning the NACA 4-digits family of profiles. However, extreme geometries (where for example the airfoil is abnormally thick or extremely thin) are discarded, so that in the end 3025 unique airfoil geometries are retained. The entire dataset, which includes airfoils and the corresponding flow solutions, is available via a public repository (Schillaci *et al.*, 2021a). Further details concerning the database are provided by Schillaci *et al.* (2021b).

Noses

The noses dataset includes healthy and pathological patient-specific anatomies. The healthy patients are only 7, identified with P1, P2, ..., P7. Their number is small, and

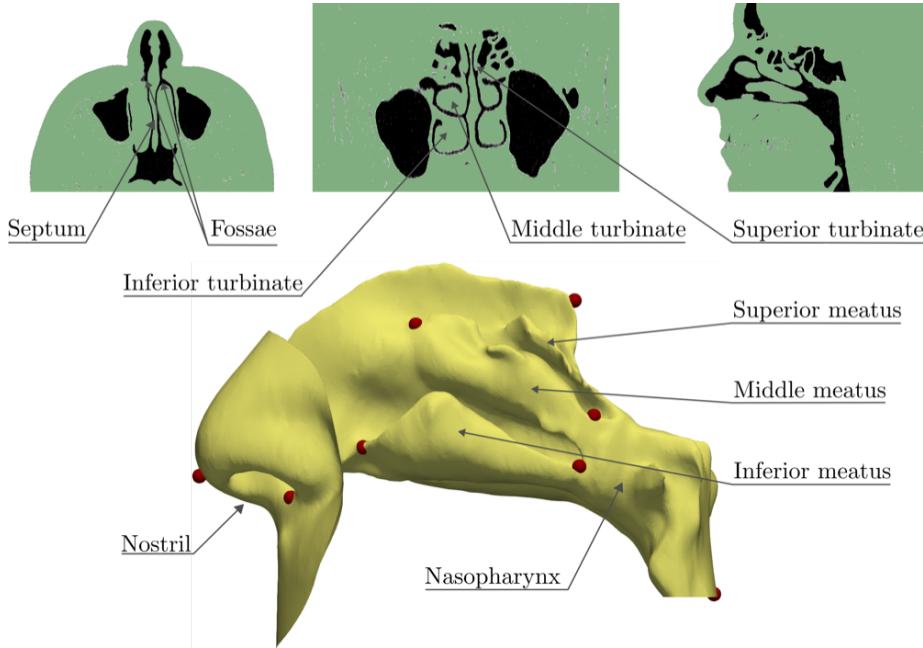


Figure 7.2: Top: CT scan of patient P1, segmented for a constant value of *HU* (from left to right: transverse, coronal and sagittal sections). Bottom: the reconstructed three-dimensional anatomy, after manual removal of the nasal sinuses. The key anatomical features are indicated. Red dots indicate the anatomical landmarks employed for registration.

this certainly constitutes an important limitation of the present study, motivated by the need to keep the computational cost under control. It should be mentioned, though, that these anatomies do contain a fair amount of variability: the most serious consequence of having few healthy patients is that the healthy class has much fewer samples than the pathological ones.

Each healthy anatomy is derived from a Computed Tomography (CT) scan of a patient with a normal sinonasal anatomy, selected after approval by the internal Institutional Review Board (IRB) of the San Paolo Hospital, University of Milan. The CT scanner is equipped with a 64-row multi-detector CT (VCT, General Electric Healthcare, Wisconsin, USA), and the patient receives a 1.9 *mSv* effective dose. Each scan is made by several hundreds DICOM images, slicing the volume of interest with axial planes; the number of pixels in each image, as well as the axial gap separating consecutive slices, dictate the spatial resolution of the scan, which must be sufficient to ensure a good reconstruction of the geometrical details. The radiological protocol, specifically designed for this study and consistent across the scans, ensures that the resolution is always better than 0.6 *mm* for each spatial direction, which grants an accurate description of the complex anatomy of the nasal airways (Quadrio *et al.*, 2016).

The open-source software 3DSlicer (Fedorov *et al.*, 2012) is then used to convert the CT images into an accurate three-dimensional geometrical model of the air-tissue interface, which defines the computational volume \mathcal{V} . The segmentation step requires the choice of a threshold to discriminate between tissue and air. The gray scale of CT scans encodes an integer value that quantifies the radiodensity of each voxel of the medium in terms of Hounsfield units or HU . The same threshold value is used for all scans for the reconstruction of the boundary. After segmentation, the reconstructed boundary is converted into a three-dimensional stereolithography (STL) file. Figure 7.2 illustrates the starting and end point of this procedure, which also includes the removal of paranasal sinuses: these large volumes contribute little to the flow in the nasal cavities while complicating geometry and meshing.

The anatomy of the nasal cavities is quite complex; excellent descriptions of its anatomy and functions are available elsewhere (Jones, 2001; Wong *et al.*, 2021). For the sake of the present work, we just mention that nasal cavities are the connecting element between the external environment and the lungs, and serve several purposes, such as air filtering, heating and humidification, smell, etc. To establish the required nomenclature, and in reference to figure 7.2, we mention that at the nose tip the nasal airway splits in correspondence of the two *nostrils*, and develops into the left and right *fossae*, separated by the *septum*, made by cartilage (anteriorly) and bone (posteriorly). The space between the septum and the lateral wall of the nose is occupied by the *turbinates*, bony structures extending longitudinally and covered by a thick layer of mucosa, that give the cross-section its characteristic hook shape, visualized in the coronal view of the CT scan in figure 7.2. Three turbinates are present for each fossa, namely the *inferior*, the *middle* and the *superior* turbinate; they serve the main physiological purpose of increasing the surface area for a better thermal exchange. The turbinates extend down to the *nasopharynx*, where the septum ends and the two fossae rejoin.

After the STL anatomy is created with CT scan segmentation, it undergoes a registration process, to provide a broadly consistent orientation across the cases. In fact, while the CT scanner is the same across the scans, the precise positioning of the patient's head varies from scan to scan; registration helps reducing the consequent variability in the reference frame and aligns all patients, by using P1 as a reference. Registration is based upon a set of 14 anatomical landmarks, customarily employed in rhinology (Denour *et al.*, 2020); the alignment takes place by minimizing the mean square error between the set of landmarks of each patient and the set of P1. Figure 7.2 also plots the location of some of the 14 landmarks employed for registration. Lastly, before creating the computational mesh, a watertight spherical volume is positioned just outside the nose tip, to represent the external ambient. The complete volume is then discretized on a mesh obtained with the meshing tool available within the flow solver OpenFOAM. The mesh consists of about 15 millions cells; no layers are used, since the uniform background mesh gets automatically refined enough near the solid boundaries to accommodate the largest velocity gradients.

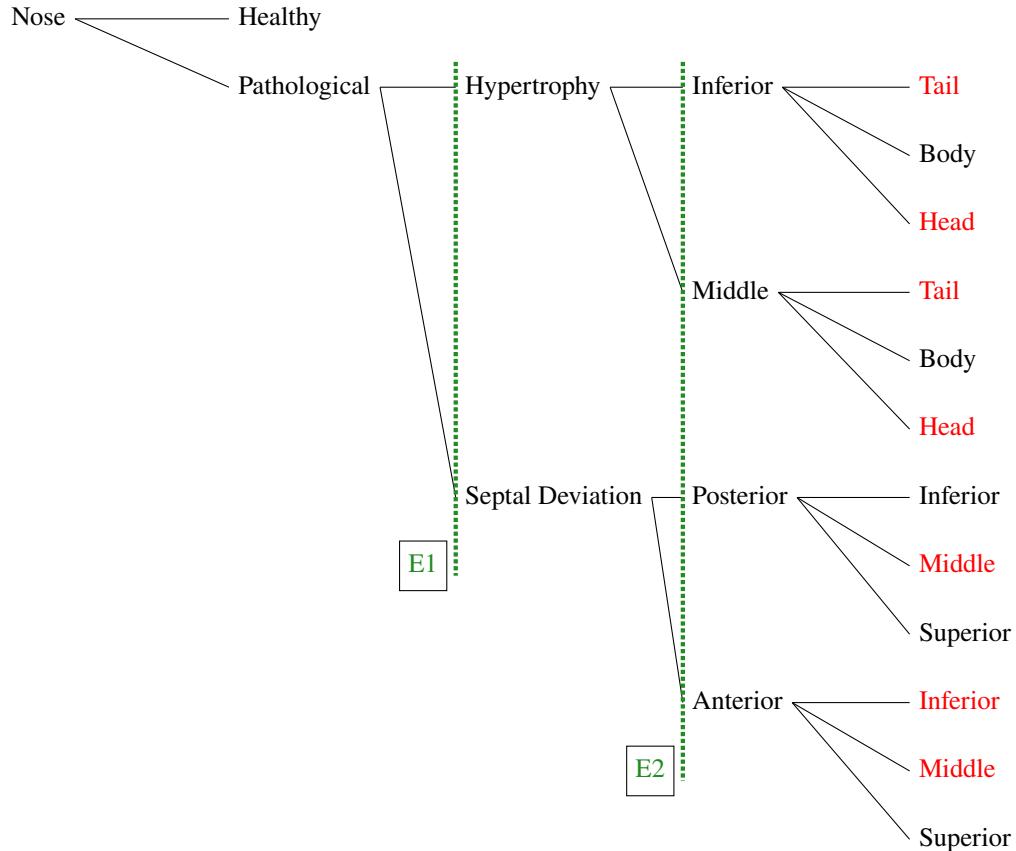


Figure 7.3: Tree of deformations/pathologies for the airfoils (top) and for the noses (bottom): each tree branch is one class. Every pathology is accompanied by a grade of severity (deformation amplitude). The dotted lines indicate the classification experiments discussed in §7.4. The red leaves at the rightmost level indicates pathologies that are actually considered in the present work.

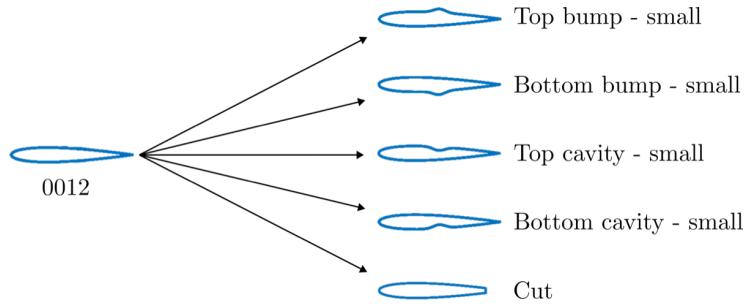


Figure 7.4: The regular (healthy) NACA 0012 airfoil, and its possible deformations.

7.3.2 The pathologies: defining Y

By design, the two datasets contain "healthy" and "pathological" geometrical variabilities. The former is represented by the 3025 standard airfoils within the NACA 4-digits family, and by the 7 healthy patients with a normal nasal anatomy. The latter involves geometrical defects, whose automatic classification based on the CFD output represents the goal of the present work. A complete view of all the geometrical defects is provided in the tree of pathologies schematically shown in figure 7.3, which includes profiles and noses.

Airfoils

The regular airfoils are modified by adding one or more geometrical defects taken from a set of five, sketched in figure 7.4 and arranged hierarchically in the top portion of figure 7.3. The "cut" defect is located at the trailing edge, which becomes truncated; the remaining four are $0.2 x/c$ wide and develop between $0.4 x/c$ and $0.6 x/c$ of the airfoil, and consists in either "cavity" or "bump", placed on "top" or "bottom". All the defects but "cut" have an associated intensity, i.e. the maximum vertical displacement taking the values of $0.05c$, $0.1c$ and $0.15c$. As mentioned, due to the small vertical displace the thinnest airfoils are not included in the dataset. Defects are added to airfoils in various combinations, leading to 4661 additional geometries, which bring the total number of observations for the airfoils dataset to $\ell = 7686$.

Noses

To introduce pathological variability in a controlled and parametrized way, the pathological anatomies are not obtained from CT scans of real patients, but created by injecting selected pathologies into the healthy anatomies. Pathologies are always applied to the right meatus.

This essential step of the procedure has required an extremely tight collaboration between all the authors. First, the list of pathologies shown at the rightmost part in figure 7.3, i.e. the "leaves" of the pathology tree, has been compiled. Each leaf is defined as the smallest deformation of an otherwise normal nasal anatomy which can be surgically corrected with a single surgical maneuver. This definition presents the

significant advantage that the final outcome of the classification immediately provides a direct indication of what surgery should be performed to correct the anomaly. Next, the deformation is manually injected onto the healthy patient P1, by a sort of inverse virtual surgery; the tool of choice is the free and open-source 3D computer graphics software Blender (van Gumster, 2015). Maximum care is given to the design of utterly realistic deformations; this involves paying particular attention to the three-dimensional distribution of the surface alterations, as well as to the quantitative measure of displacements. For example, it is known that some defects, when present in one meatus, induce (smaller) adaptations into the other meatus: although defects are always injected on the right, there is often a need to also alter the left meatus. The process of manual creation of deformations is extremely time-consuming. However, after deformations have been satisfactorily created manually for patient P1, the procedure for replicating them on the other patients can be automated, by taking advantage of functional maps (Ovsjanikov *et al.*, 2012). Functional mapping, described in greater detail by Magnet & Ovsjanikov (2023), is a computational geometry tool which allows us to seamlessly transfer the function defining pathology over the shape (anatomy) of patient P1 to the shape of any other patients. The interested reader is referred to the original work for further details.

As illustrated in figure 7.3, the defects considered in this work belong to two well separated classes: septal deviations (SD) and turbinate hypertrophies (TH). SD can be located either anteriorly (SD-A) or posteriorly (SD-P), and also differ depending on their axial location (superior, middle or inferior). TH may affect either the middle (TH-M) or the inferior (TH-I) turbinate, and according to their coronal position can be placed at the turbinate head, body or tail. Only seven of these defects, indicated by the red leaves in figure 7.3, are considered in this study. Each is characterised by two intensities (mild or severe), with the severe one leading to a contact between turbinate and septum. Since in some cases more than one pathology is present at the same time, in the end 17 distinct combinations of pathologies are considered. In total, the database includes 270 pathological and unequivocally labelled anatomies, added to the 7 healthy ones.

Figure 7.5 illustrates, for patient P1, the healthy anatomy and two pathological variations. A three-dimensional view defines six coronal sections, indicated with S1, ..., S6. The anatomy in these cross-sections is visualized, with the main anatomical regions identified in colour. The six sections are equally spaced in the sagittal direction; the first and last one are identified by the landmarks at the head and tail of the inferior turbinate. The central row shows a severe SD-P located medially, and the bottom row is a severe TH-I affecting the turbinate tail. The anatomical changes corresponding to the pathologies are highlighted. They actually deform a significant portion of the original anatomy. For SD, especially in the severe cases, the restriction of the airway passage on the affected fossa is also accompanied by a small enlargement of the opposing one; hypertrophy of the turbinates, instead, tends to involve a single fossa, where the cross-sectional area is reduced, while the opposite one remains untouched.

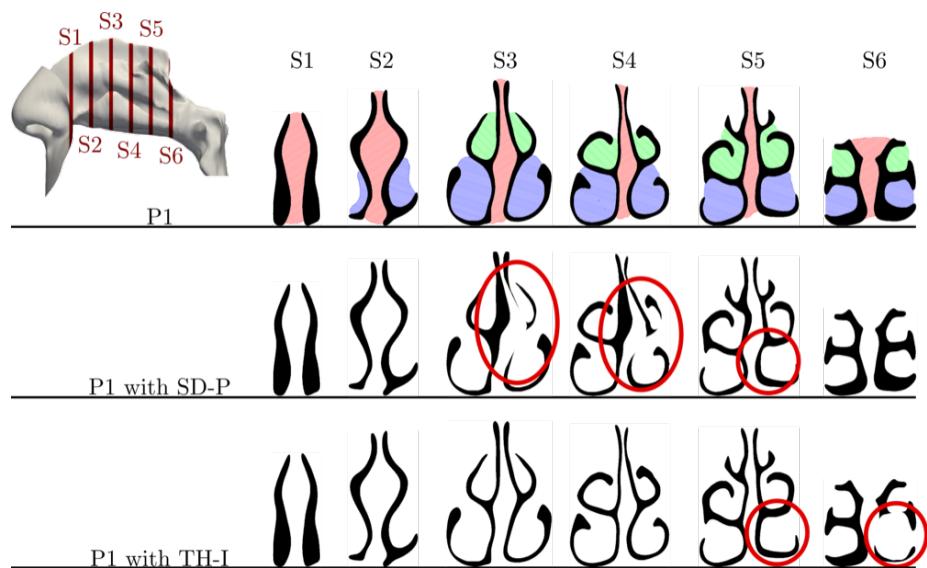


Figure 7.5: Healthy anatomy of patient P1 (top row) versus two pathological modifications (severe septal deviation located posteriorly, SD-P, in the middle row; severe hypertrophy of the inferior turbinate, TH-I, in the bottom row). Each anatomy is described via six coronal sections, defined on the three-dimensional view. Colours indicate the main anatomical areas: the passageways are coloured in black; the inferior turbinates in blue, the middle turbinates in green, the septum in light red. Red circles highlight regions altered by the pathology.

7.3.3 The simulations: computing X

The simulation strategy used to compute the flow fields and to extract flow-based features f differs for the two datasets. The two-dimensional cases of the airfoils dataset are simulated with a low-cost and low-fidelity approach, in which RANS equations are solved with a turbulence model. The three-dimensional cases belonging to the geometrically complex noses dataset employ well-resolved Large Eddy Simulation (LES). Both datasets are computed using the finite-volume library OpenFOAM (Weller *et al.*, 1998).

Airfoils

Each airfoil is placed in a uniform free stream U_∞ at an angle of attack of 10° . The Reynolds number, defined with the profile chord c as the reference length, U_∞ as the reference velocity and the kinematic viscosity ν of the fluid, is set at $Re \equiv U_\infty c / \nu = 3 \times 10^6$.

The two-dimensional simulations solve the incompressible RANS equations. All the differential operators are discretized at second-order accuracy. The turbulence model is the Spalart-Allmaras model (Spalart & Allmaras, 1992), with standard values of the model constants. On the airfoil surface the no-slip condition for the velocity, the zero-normal gradient condition for the pressure and zero turbulent viscosity are imposed. On the far boundary, free-stream boundary conditions are imposed, with a uniform turbulent viscosity of 1×10^{-5} . A simulation is considered at convergence once the drag coefficient varies less than 5×10^{-6} over an interval of 100 iterations. RANS simulations directly provide the mean velocity and pressure fields, \mathbf{U} and P , as well as quantities related to the turbulence model. Results are validated by comparing the drag coefficients of some popular airfoils against experimental data obtained from Abbott & von Doenhoff (2012). On average, computing a single case requires about 15 core minutes on 16 cores with Intel Xeon E5-2697 CPUs at 2.30 GHz. The dataset containing unmodified airfoils is available from the public repository at <https://doi.org/10.5281/zenodo.4106752>.

Noses

Each nasal anatomy is simulated for a steady inspiration at a rate of 280 ml/s or $2.8 \times 10^{-4} \text{ m}^3/\text{s}$, which corresponds to a restful breathing (Wang *et al.*, 2012).

The three-dimensional numerical simulations solve the incompressible LES equations. Each differential operator is discretized at second-order accuracy: Schillaci & Quadrio (2022) have recently demonstrated in this flow the importance of using at least second-order discretization, which is at least as important as a time-resolved approach to obtain quantitatively accurate results in this flow. The chosen LES turbulence model is the Wall-Adapting Local Eddy viscosity or WALE (Ducros *et al.*, 1999), which suits well complex geometries and is algebraic, hence does not need additional boundary conditions. We know after previous experience (Covello *et al.*, 2018) that satisfactory statistics for the mean and variance fields can be obtained, at this breathing rate, after a total simulation time of 0.65 s , by discarding an initial transient of 0.05 s . The time step for the temporal advancement of the solution varies such that the Courant–Friedrichs–Lewy number remains below unity.

At the spherical inlet, the required flow rate is imposed as a velocity boundary condition, and a zero-gradient condition is used for pressure. At the outlet, placed in the throat, a reference pressure of 0 Pa is imposed, with zero-gradient for velocity. At the wall the no-slip condition for the velocity and the zero-gradient condition for the pressure are imposed. LES simulations compute the evolution of instantaneous, spatially filtered fields of velocity and pressure, \mathbf{u} and p ; temporally averaging over the duration of the simulations leads to \mathbf{U} and P .

The mesh, consisting of about 15 millions cells, provides a spatial resolution comparable to the most resolved LES simulations of this flow (see e.g. Calmet *et al.*, 2021); it should also be recalled that the paranasal sinuses are left out, hence the computational volume is decreased and the resolution improved for a given number of cells. Owing to the very fine mesh, the role of the LES model becomes marginal. In a representative case, the maximum (in space and time) ratio between the turbulent viscosity and the molecular viscosity is only 4.4, whereas the mean value of the ratio is about 10^{-2} ; this confirms that the presents simulations are very well resolved LES, and are not far from a direct numerical simulation.

On average, computing a single case, including the discarded initial transient, requires about 4600 core hours on 96 cores, with Intel Xeon E5-2697 CPUs at 2.30 GHz.

7.3.4 Feature extraction: computing f

The size of the CFD output X is much larger than the number of available observations ℓ . This is especially true for the nose dataset, where the small total number of observations $\ell = 277$ must be compared with the large size of the CFD output, with $m = 4$ variables per cell and $n \approx 1.5 \times 10^7$ cells. Avoiding the direct use of the full CFD-computed flow field becomes mandatory in practice, and one has to resort to compact and informative features to shrink the number of inputs to the classifier, while preserving as much as possible of the information content of the CFD solution.

Choosing the appropriate feature is a non-trivial problem. In previous work (Schillaci *et al.*, 2021*b*), we preliminarily discussed various hand-crafted features that lend themselves to compacting CFD information for the purpose of predicting geometrical information with a relatively small training set. Building on that work, here we consider two distinct frameworks to extract features from the flow field, namely regional averages and streamlines.

The regional or local average of a fluid variable, indicated with an asterisk superscript as e.g. in P^* , is the spatial average of a space-varying quantity over a predefined portion (e.g. a plane or a line, or portions thereof) of the computational domain. Averaging quantities over a small number of predefined regions offers the key advantage of sampling the whole flow field without reference to the (unknown) position of the geometrical defect. A different answer to the same problem is provided by streamlines. Owing to the convective character of the flow, streamlines departing from the inflow cross the entire computational volume before exiting through the outflow. Hence, streamlines can be computed without reference to the defect, and inspected for flow information accumulated throughout the domain.

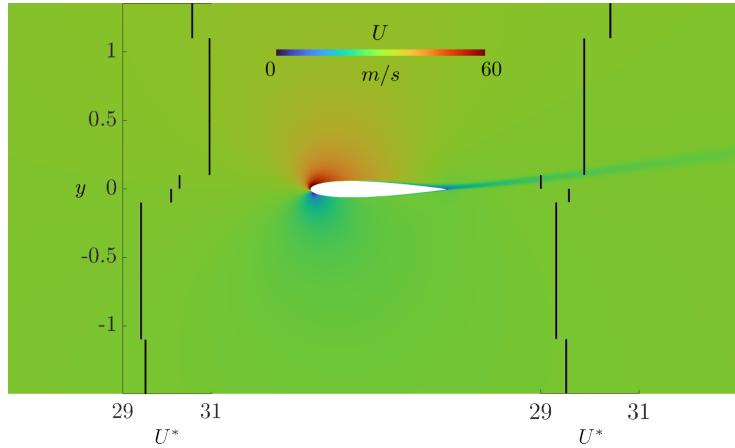


Figure 7.6: Flow field around an airfoil at incidence. The color map shows the time-averaged velocity. One chord upstream and one chord downstream the two piecewise-constant plots illustrate the evolution of the regionally-averaged velocity U^* along the line.

Airfoils

For the two-dimensional flow field around the airfoils, regional averages are computed over two one-dimensional subdomains, made by straight lines drawn orthogonally to the chord of the airfoil. These lines are positioned $1c$ upstream of the leading edge, and $1c$ downstream the trailing edge. As shown in figure 7.6, each line is subdivided in 8 parts, corresponding to the following intervals for the y/c coordinate: $[-500, -100, -10, -1, -0.1, 0, 0.1, 1, 10, 100, 500]$. Overall, 16 segments become available for the extraction of the regionally averaged value of flow variables.

Mean streamlines, instead, start from one thousand seed points on the chord-normal straight line placed $1c$ upstream of the airfoil; seeds are placed to be approximately symmetric with respect to the stagnation streamline. Streamlines are then subdivided in "top" and "bottom" groups, depending on their position relative to the stagnation streamline; their spatial extension is truncated by defining a normalized curvilinear coordinate such that $s = 0$ corresponds to the line $1c$ upstream, and $s = 1$ to the line $1c$ downstream.

Noses

Selecting the subdomains for regional averaging obviously affects the information content of the extracted feature. While for the profiles one can safely guess that the information fades away with distance, the situation becomes more complicated for the noses, where there is a fair amount of "healthy" anatomical variability, and the registration is only approximate. To identify corresponding sections across different anatomies, we leverage again the anatomical landmarks already described in the context of figure 7.2, and identify six equally spaced coronal cross-sections, from S1 to S6, in such a

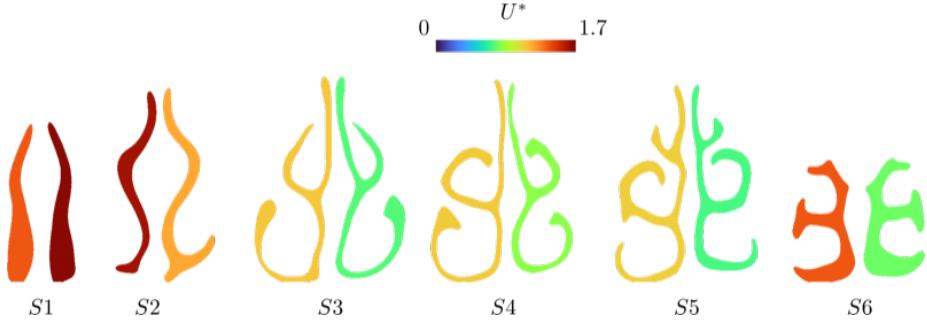


Figure 7.7: Healthy patient P1: regional average U^* of the velocity magnitude, across the sections S1 to S6, for a total of 12 values.

way that S1 corresponds to the head of the inferior turbinate, and S6 to its tail. Each section is then divided into left and right portions, and regional averages are computed on these, for a total of 12 features. Figure 7.7 shows a typical distribution of these values, for the healthy patient P1.

As for the streamlines, they are computed, based on either the mean or an instantaneous flow field, by randomly seeding the spherical boundary inlet with one thousand points, and computing the 1000 corresponding streamlines. They are then sorted into "left" and "right" groups, depending on the meatus they are passing through. A normalized curvilinear coordinate s is defined on each of them, with $s = 0$ at the intersection with S1, and $s = 1$ at the intersection with S6.

7.3.5 Neural network: defining \mathcal{K}

The classifier \mathcal{K} is a neural network (Goodfellow *et al.*, 2016); its architecture is identical for the two datasets. A neural network (NN) is made by several layers of connected neurons: the input layer, one or more hidden layers, and the output layer. Each neuron is connected with all the neurons of the previous layer. Such connection is associated to a learnable parameter (weight). On top of that, there is another parameter associated to each neuron, the bias. Each neuron computes a linear combination of the values of the previous layer, and then feed this result to a non-linear activation function. To tune the parameters of the NN during training, a cost (or loss) function is used to quantify its accuracy, and to optimize weights and biases via minimization. The NN is defined by its hyperparameters and of neurons per layers, moreover it is possible to select a few activation functions, the loss function and the backpropagation algorithm used to train the NN. Our NN has a number of input nodes that equals the number of input features. The input layer is followed by three fully connected hidden layers, made by 30, 20 and 10 neurons respectively, and by the output layer, whose number of nodes is equal to the number of classes. For all but the output nodes, the activation function is an hyperbolic tangent, widely used in classification tasks (Bishop, 2006). The activation function of the output layer is the sigmoid function. The loss function is the binary cross-entropy (Goodfellow *et al.*, 2016), which assigns a penalty that increases logarithmically as the

predicted probability diverges from the actual label. Lastly, the weight and bias values are updated according to the scaled conjugate gradient backpropagation algorithm (Møller, 1993).

7.3.6 Neural network: dataset partition and training

When training the NN, the dataset is partitioned into three parts: the training, the validation and the test set. The training set, usually the largest one, is used to tune the parameters of the net; the validation set is used to decide when training can be stopped; the test set is used to assess the accuracy of the NN over a set of instances unavailable to both the training and validation sets. More than one strategy can be employed to use the entire dataset effectively; and the way the dataset is partitioned for training, validation and test can noticeably impact the results (Kutz, 2017).

In this work, we consider two partition strategies, described for example by James *et al.* (2021). One is the most widespread partition strategy: the k -fold cross-validation method, which divides the dataset into k non-overlapping folds, trains and validates the NN on $k - 1$ folds, and reserves the k -th fold for testing. This procedure shuffles and spans the entire dataset at least once, and is a good choice in most scenarios. However, given the medical background of the present work, we adopt the Leave-One-Out (LOO) cross-validation. With LOO, one patient (and all the derived anatomies) is not included in the training and validation sets, and is only used for testing. In this way, during testing the NN is always confronted with entirely new cases, mimicking the scenario where new patients come up and are classified based on a pre-existing NN. LOO provides a significant check of the ability of the NN to generalize to unseen patients. Unfortunately, LOO partitioning is less obvious for the airfoils dataset. We accomplish it here by dividing the airfoils into groups based on thickness intervals, while making sure that the training set remains balanced.

7.4 Results

This Section describes the results of the six (binary and multiclass) classification experiments E1, E2, … E6 shown schematically in figure 7.3 and pertaining to the two datasets. As stated above in §7.3.6, datasets are partitioned using either the k -fold and LOO strategies, to emphasize the importance of the chosen validation experiment. On the other hand, in this Section only the regional averages of velocity magnitude are used as input to the NN, thus the six left/right pairs of values for U^* in correspondence of the six cross-sections S1, …, S6 depicted in figure 7.7. An in-depth discussion of the results, including the critical comparison of various features with different fluid mechanical significance and the use of streamlines, is deferred to §7.5.

Table 7.1 shows the results of the six experiments. The performance of the classifier is expressed via two standard metrics: accuracy and F1-score. Accuracy is the ratio between the number of correct and total predictions: it is a simple and intuitive metric, which works well as long as the dataset contains a balanced number of observations for each class. The F1-score, instead, is an alternate metric which is perhaps less intuitive but more robust. It is defined as the harmonic mean of precision and recall, precision

Exp	Dataset	classes	ℓ	inputs	Accuracy		F1	
					k -fold	LOO	k -fold	LOO
E1	Airfoils	2	7686	8	0.81	0.75	0.85	0.82
E2	Airfoils	3	6625	8	0.82	0.76	0.73	0.64
E3	Airfoils	4	6204	8	0.83	0.78	0.67	0.56
E4	Airfoils	6	5779	8	0.88	0.84	0.64	0.53
E5	Noses	2	270	12	0.97	0.85	0.95	0.80
E6	Noses	4	154	12	0.86	0.79	0.73	0.57

Table 7.1: Results of the six classifications experiments shown in figure 7.3, assessed in terms of accuracy and F1 score, using the regionally-averaged quantity U^* as input feature. Two cross-validation strategies (k -fold and LOO) are considered.

being the ratio between true positives and total positives, and recall the ratio between true positives and the total number of relevant elements. F1 ranges in $[0, 1]$ and should be preferred to accuracy whenever the cardinality of classes is unbalanced, since an high F1-score is obtained when both false positives and false negatives are low.

Table 7.1 also includes the number of observations ℓ . While the binary classification experiments E1 and E5 use all the observations of the dataset, the remaining experiments are multiclass; owing to the limited number of instances, we limit predictions to single-label cases. This implies that observations with multiple labels, present in both databases, need to be discarded, leading to a reduction of the number of usable observations. Moreover, in multiclass experiments a compound average is carried out, i.e. the relevant metrics are computed on class-averaged values.

7.4.1 Airfoils

The first experiment E1, as illustrated in figure 7.3, is a binary classification problem on the whole airfoils dataset, including $\ell = 7686$ samples: the task is simply to predict whether the airfoil defective or not. Defects appear either singularly or in arbitrary combinations. The regional velocity averages are computed on the vertical line placed $1c$ downstream the trailing edge, and divided in 8 segments, thus leading to 8 input nodes; the other vertical line is not considered, because of its limited information content. For a binary classification, the output layer consists of a single neuron. With k -fold cross-validation and $k = 7$, the prediction accuracy is 81%, while LOO achieves a lower accuracy of 75%. This outcome, fully confirmed with the remaining experiments, indicates that a NN trained with k -fold cross-validation always performs better than one trained with LOO, because the training set is very similar to the test one. Unfortunately, k -fold is not plausible when the background application is of the medical type.

E2 is a multiclass classification problem. Moving one step right in figure 7.3, the goal of the classifier is the discrimination among three classes (healthy, body-defective and tail-defective), thus the output layer has 3 neurons. To avoid sample associated to multiple defect types, airfoils with defects on both body and tail are excluded, which shrinks the dataset size down to $\ell = 6625$ samples. Compared to the binary classification E1, accuracy remains nearly the same, but F1-score deteriorates significantly, owing to

Class	accuracy	precision	recall	F1
Healthy	0.60	0.59	0.80	0.68
Top bump	0.88	0.03	0.00	0.01
Top cavity	0.90	0.21	0.00	0.01
Bottom bump	0.90	0.52	0.33	0.40
Bottom cavity	0.88	0.36	0.34	0.35
Cut	0.87	0.35	0.43	0.39

Table 7.2: Results of the airfoil multi-class classification E4, with LOO cross-validation.

the increased difficulty of the task. This effect is evident with both partition methods. The almost unchanged accuracy exemplifies how this metric is potentially misleading whenever classes are not properly balanced. Indeed, in E2 there is a relative increase in weight of the healthy class which slightly outnumber the others, as defective cases with multiple defects are pruned to avoid non-unique labels.

E3 is another multiclass classification problem on the airfoil dataset, corresponding to a further step down the tree of figure 7.3. The total number of classes is now four: healthy, top defect, bottom defect, and tail defect. The dataset is further reduced in size to discard cases with non-unique labels, and includes $\ell = 6204$ observations. The trend noticed with E2 is confirmed: increasing the complexity of the classification task leads to small changes in accuracy, which remains high at 83% with k -fold and 78% with LOO, but F1-score deteriorates further, regardless of the partitioning method.

E4 is the most detailed multiclass classification. The classifier is requested to identify each defect, which brings the total number of output neurons to 6 (corresponding to the 5 defects and the non-defective class). Once again, cases with multiple defects are discarded, leading to $\ell = 5779$. The same effect seen in E2 and E3 is amplified here: accuracy increases (this time more significantly), while the F1 metrics deteriorates. In this experiment, the healthy class alone contains more than one half of the total number of observations for the entire set of six classes: thus, the NN simply tends to avoid the selection of underrepresented labels, explaining the generally increased accuracies.

Table 7.2 provides classification results without the compound average, and provides information about the predictability of each defect. An additional effect becomes evident: for aerodynamic reasons that are easy to understand, defects placed on the suction side are less visible in the flow field. It should be remembered that the simulations are carried out at a constant, positive angle of incidence. Hence, although the very high accuracy, the F1-score for these two defects drops to nearly zero: the NN does not capture true positives for the top defects, because it never indicates positives.

7.4.2 Noses

A major difference between the noses and the airfoils datasets lies in the number of observations, which is approximately one order of magnitude smaller for the noses, with $\ell = 270$. More importantly, owing to the procedure described in §7.3.2 and employed to build the dataset, all the pathological cases are derived from 7 healthy anatomies only. Because of the insufficient cardinality of the healthy class, the 7 healthy anatomies are

Patient	accuracy	precision	recall	F1
P1	0.85	0.75	1	0.86
P2	0.93	0.89	0.94	0.91
P3	0.83	1	0.61	0.76
P4	0.76	0.6	1	0.75
P5	0.93	1	0.77	0.87
P6	0.88	1	0.62	0.76
P7	0.82	0.57	0.89	0.70

Table 7.3: Results of the nose binary classification E5, with LOO cross-validation.

Class	accuracy	precision	recall	F1
Anterior septal deviation, SD-A	0.91	0.82	0.91	0.86
Posterior septal deviation, SD-P	0.90	0.30	0.11	0.16
Middle turbinate hypertrophy, TH-M	0.67	0.47	0.51	0.49
Inferior turbinate hypertrophy, TH-I	0.71	0.51	0.51	0.51

Table 7.4: Results of the nose multi-class classification E6, with LOO cross-validation.

not included in the dataset, although this means giving up to what is perhaps the most interesting classification task, i.e. the discrimination between healthy and pathological noses.

The task of E5 is the binary classification of the pathology, i.e. to predict whether the pathology is a septal deviation or a turbinate hypertrophy. Because of the small number of observations, only 12 input features are used (the regional average of the velocity magnitude in the left and right part of sections S1, ..., S6). Table 7.1 shows that the prediction accuracy is 97% with k -fold cross-validation (with $k = 7$) and 85% with LOO. Like in the previous experiments, k -fold tends to yield higher values for accuracy and F1-score in comparison to LOO.

Table 7.3 expands the classification results, and aggregates the classification performance on samples generated from each patient. The performance is quite stable over all the patients, with the accuracy comfortably lying above 82%, except for patient P4, that shows a small reduction in accuracy down to 76%.

Lastly, E6 is a multiclass classification problem for the noses dataset, and descends one more level down than E5 in the pathology tree of figure 7.3. The classifier has to discriminate among four classes: anterior and posterior septal deviations (SD-A and SD-P), and hypertrophy of the middle and inferior turbinates (TH-M and TH-I). Once again, observations with ambiguous labels must be pruned, thus lowering the number of observations down to $\ell = 154$ samples. When considering the overall performance, the accuracy is 86% with k -fold cross-validation and 79% for LOO. Compared to the previous experiment, the F1-score drops noticeably, meaning that there is a stronger tendency to wrongly label the pathologies, as in the airfoil dataset. Results can also be analyzed per each class, as in table 7.4. It turns out that septal deviations are easier to identify than turbinate hypertrophies. In particular, SD-A presents high accuracy and F1-score, whereas SD-P has a comparable accuracy, but its F1-score is extremely low at

16%. This implies that the classifier tends to provide a negative response, and this can be traced back to the fewer number of occurrences for SD-P in the dataset. Hypertrophy of the turbinates is consistently less predicted by the classifier. In particular, TH-M has the worst performance with accuracy at 67% and F1-score at 49%.

7.5 Discussion

To the best of our knowledge, the present work is first to attempt the automatic classification of shapes of solid bodies immersed in a moving fluid, by using solely information gathered from the surrounding flow field. There are a few works which bear significant similarity; for example Alsalmal *et al.* (2018) classified vortex wakes behind airfoils using information extracted from sensor on the wake, and Wang & Hemati (2019) classified wakes from hydrodynamic signal measurement. However, in those cases the classification applies to the flow field, and not to the type of geometry that produced it.

In this section the classification experiments described above are further analysed, with a focus on the nose dataset. In particular, the observation that the classifier successfully discriminates nasal pathologies will be integrated with a critical discussion of the choice of the flow feature(s) used as input, and of their fluid mechanical significance. Choosing the right feature is a critical step, especially for health-related applications, which are bound to have a limited number of labeled samples. This hinders the use of deep neural networks, which would otherwise dispense from selecting a specific feature *a priori*, and might in principle lead to satisfactory results based on the anatomy alone. Hence, features are selected manually in a somewhat heuristic manner; it is crucial to combine fluid mechanical insight and medical expertise to understand when and why they are successful, so that the procedure can at the same time be successful and guarantee explainability.

7.5.1 The regional average of the velocity magnitude

The classification experiments described above in §7.4 have employed U^* as input feature, i.e. the regional average of the velocity magnitude over six predetermined coronal sections, separated into the left and right fossae, for a total of 12 inputs.

The first question that comes to mind is whether the geometric information of the area A of the airways in the same sections would be an equivalent or better feature. In preliminary work (Schillaci *et al.*, 2023), on the basis of an extremely simplified version of the present problem, we have suggested that flow-based features might present substantial advantages compared to geometry-based ones, and would be better at training a NN for a given number of observations. However, for the present setup U^* is strongly connected to A , since the flow rate is enforced as a boundary condition, and the flow is incompressible: owing to mass conservation, only the flow partition between right and left is actually free.

The binary classification experiment of the noses dataset, i.e. E5, is thus repeated by using the area A of the 12 coronal subsections as input. The accuracy with LOO partitioning drops from 85% with U^* to 81% with A : the flow-based feature remains superior, but the information conveyed by the cross-sectional areas is not insignificant.

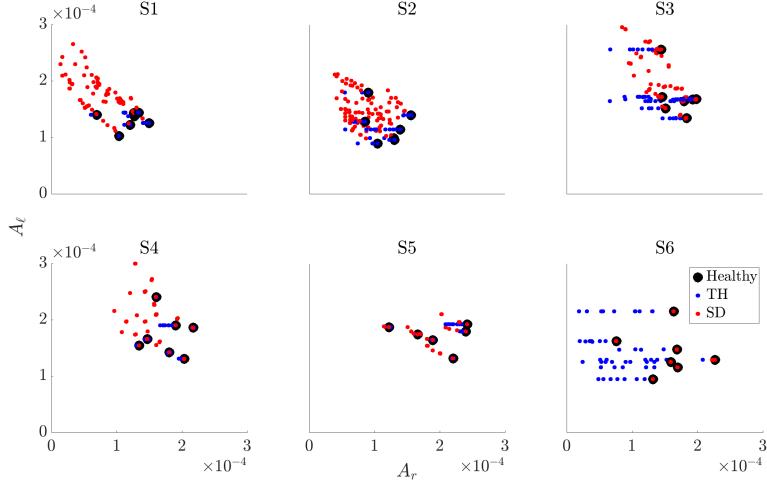


Figure 7.8: Evolution of the cross-sectional area (A_r of the right fossa on the horizontal axis, A_ℓ of the left fossa on the vertical axis) from S1 to S6. Black dots are the healthy patients, red and blue symbols represent SD and TH.

We believe, however, that the reasonably good performance of A on such a small dataset is overemphasized by the simplified setting of the experiments. In fact, pathologies beyond those considered here do exist (e.g. the perforation of the nasal septum) which do not alter the cross-sectional area, but modify the flow field significantly: these would be plainly invisible to any area-based feature. On the other hand, SD and TH leave a clear footprint on A , especially when little anatomical variation is available.

Figure 7.8 illustrates how, in the present setup with two pathologies only, area changes possess a pattern that leads to a reasonable discrimination between the two pathologies. The figure plots the evolution from S1 to S6 of the right cross-sectional area A_r (horizontal axis) and the left cross-sectional area A_ℓ (vertical axis), for healthy and pathological cases. We recall that pathologies are always injected on the right side. The seven black dots correspond to the healthy patients, and illustrate – within the limits of 7 observations only – the natural variability of healthy anatomies. The color of the dot identifies the pathology, with red corresponding to SD and blue to TH. SD modify both fossae, reducing the right area and increasing the left one; the red dots thus tend appear to the top and to the left of the corresponding healthy anatomy. TH instead affect the right fossa only, and the blue dots align horizontally below the reference black dots. This nearly deterministic pattern, however, is expected to disappear as soon as more variability is added to the dataset.

A further limitation of the area feature is that pathologies sometimes alter A in a very localized way; pathologies located far from the six predetermined sections would thus be little or no visible at all to area-based features. The SD-A evidently affect S1, which in turns is only marginally influenced by the TH-I, as most of the blue points overlap to their healthy reference. A similar observation holds for SD-P in S5. On

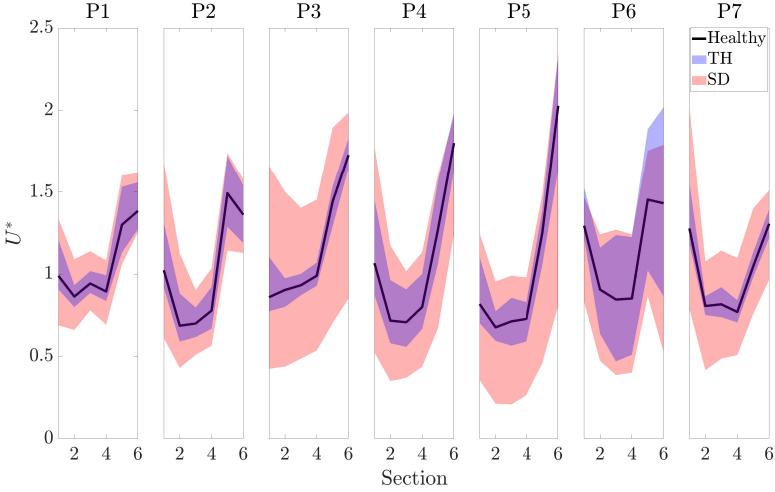


Figure 7.9: Evolution of U^* along the coronal sections S_1, \dots, S_6 , for each group of patients derived from the same healthy anatomy. The horizontal axis of each panel is the section ID. The black line plots the average of the left/right U^* for the healthy patient; color bands are drawn from the average value of the positive/negative excursions from the mean for the pathological patients. Blu is for TH, red for SD.

S_6 , instead, turbinate hypertrophies clearly stand out, at least the subset of TH which affect the turbinate tail, whereas no SD is visible: only blue dots scatter in the plot. To confirm quantitatively that area changes are more local than flow changes, a dedicated classification experiment, similar to E5, is carried out which only uses information extracted by the last two cross-sections S_5 and S_6 ; there, only a subset of pathologies (namely SD-P and TH-I at the tail) are directly observed with the area. Indeed, for U^* accuracy remains high at 87% with 82% F1, whereas it drops to 69% with 60% F1-score when A is used as a feature. It is confirmed that area is a local information, while flow-based features keep memory of the geometric changes upstream.

Figure 7.9 compares, for each patient, the evolution from S_1 to S_6 of the regional average U^* for cases with SD and TH. Each of the seven frames combines together one patient and the corresponding set of pathological modifications, with SD in red and TH in blue. The figure portraits with a black line the left-right average of U^* for the healthy patient; the color bands indicate the average value of the positive and negative excursions across all the corresponding pathological cases. The red bands are generally wider, indicating that SD induce larger excursions of U^* in both fossae. However, this is clearly patient-dependent, with for example P3 showing extremely significant excursions, and P6 having TH excursions that are sometimes larger than SD ones.

Results of E6, reported above in table 7.4, show that SD is the easiest pathology to classify. The underlying reason can be appreciated by looking at figure 7.10, where a two-dimensional time averaged map of U in section S_6 of the healthy anatomy for P1 is compared with the analogous quantity computed for (single) pathologies derived from

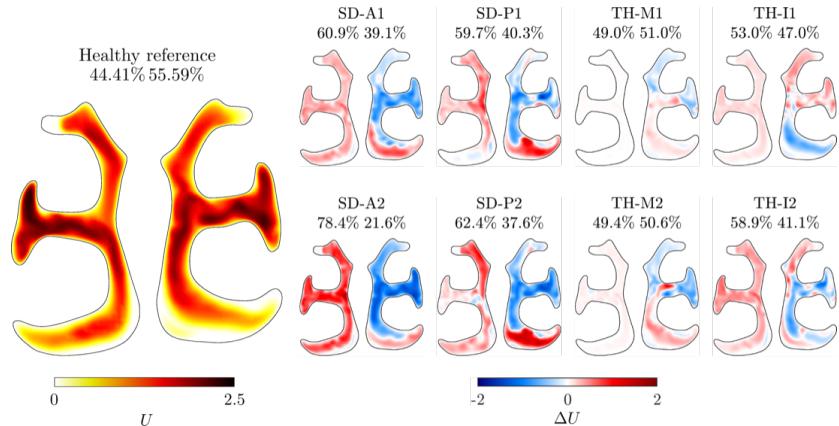


Figure 7.10: Velocity magnitude for P1 in section S6, in comparative form between the reference anatomy and the four pathologies (SD-A, SD-P, TH-M, TH-I), at both severities. The percentages above each panel indicate the distribution of the flow rate across the two fossae.

P1, allowing to directly observe the effect of a pathology on the mean flow field.

Figure 7.10 shows that SD alter the flow field in S6 significantly, starting from the unbalance of the flow rate, which is significantly less symmetrical than the reference anatomy already at mild severity. In contrast, mild TH leave the flow rate distribution quite balanced. Besides shifting part of the flow rate to the left fossa, SD alter the flow in the right fossa by increasing it in the lower part of the cross-section, between the middle and inferior turbinate, a consequence of the downward deflection of the flow by the upstream septal deviation. The septal deviations placed posteriorly, though, are quite close to S6, and produce some effect on the left fossa too. TH have a less pronounced effect in S6: the left fossa at most shows a small and uniform velocity increase, and no clear pattern is visible on the right. This explains why, as seen from table 7.4, TH are most difficult to predict. These flow field modifications have been verified to be qualitatively consistent across all the considered patients.

It should be noted that such a relatively obvious comparison to highlight the direct consequence of a certain pathology on the flow field has never been carried out so far. As long as physiological variability is not removed, the analytical power of CFD is unable to provide a clear picture. For example, one of the best available recent attempts (Casey *et al.*, 2017) compared the flow in a batch of 15 healthy patients and 15 pathological ones, without the ability to discriminate pathologies. The comparison, moreover, was carried out in a statistical sense only, and resulted in correlating the well-being of the patient with the changes of flow rate in the inferior/middle/superior portion of the coronal cross-section of the narrowest nasal cavity. What makes such comparison possible here with relative ease is, once again, the strategy employed to inject realistic pathologies on top of a healthy anatomy: the procedure described in §7.3.2 enables the comparison of two anatomies which are realistic and also identical but for the pathology of interest.

	k-fold				LOO			
	accuracy	precision	recall	F1	accuracy	precision	recall	F1
U	0.97	0.97	0.93	0.95	0.85	0.75	0.87	0.80
$ \nabla P $	0.96	0.96	0.91	0.93	0.76	0.61	0.84	0.71
$P_{in} - P$	0.91	0.86	0.88	0.87	0.76	0.63	0.65	0.64
$P_1 - P$	0.91	0.90	0.84	0.87	0.76	0.63	0.74	0.68
Ω^2	0.95	0.96	0.89	0.92	0.74	0.59	0.80	0.68
$P - P_{out}$	0.89	0.88	0.80	0.84	0.68	0.53	0.70	0.61
$P - P_6$	0.92	0.91	0.85	0.88	0.74	0.59	0.77	0.67
v_t	0.87	0.89	0.72	0.79	0.67	0.52	0.64	0.57
R	0.85	0.91	0.64	0.75	0.64	0.49	0.66	0.56

Table 7.5: Performance of various mean flow quantities (velocity magnitude, pressure gradient magnitude, pressure computed with respect to various references, norm of the velocity gradients, turbulent viscosity and nasal resistance) as input features in the binary classification of the noses dataset, experiment E5 in figure 7.3. Except for nasal resistance, the remaining inputs are the 12 left/right regional averages in S1, ..., S6.

Overall, the observation that each pathology tends to produce a relatively localized and deterministic effect on the flow field paves the way for a more informative use of the concept of regional averages, which extends and improves the approach by Casey *et al.* (2017). As suggested already by Schillaci *et al.* (2021b), one could improve the present classification strategy by further dividing each of the left and right cross-sections in the inferior/middle/superior portions. Each cross-section would thus yield not just 2, but 6 regionally averaged values of U^* . This is not attempted here, because of the limited size of the dataset; however, increasing the number of inputs by a factor 3 implies a larger dataset by an equivalent factor, and such a factor appears not unreasonable, especially against the substantial increase of informative content.

7.5.2 Other features based on regional averaging

So far, the only quantity considered for regional averaging has been the magnitude U of the mean velocity vector. Obviously, alternative features based on different flow quantities are possible, and multi-feature classification can be attempted. For the binary classification E5 of the noses dataset, table 7.5 introduces additional flow quantities, which are then regionally averaged as before in S1, ..., S6; their performance is compared to that of U^* . Flow quantities are chosen among scalars and magnitudes of vectors, for reasons related to invariance; moreover, we require the feature to be the same across all the inputs. None of these is actually a hard requirement for a NN.

First of all, table 7.5 shows the classification performance when features are extracted from different quantities and confirms that k -fold cross-validation leads to an excellent classification accuracy for any feature, to the point that there is little difference across the various features. Hence, from here on, only LOO as the worst case will be considered, since it makes differences among features more evident.

The most effective feature for regional averaging is still the velocity magnitude U ,

but interesting alternatives are available. The second-best quantity is the magnitude of the mean pressure gradient ∇P , whose regionally-averaged value provides a simple estimate of the local rate of pressure loss along the axis of the nasal fossa, which is approximately normal to the section. Hence, this quantity too relates to the longitudinal variation of the cross-sectional area. The two pathologies considered in this work affect the cross-sectional area in a distinct way, as discussed above in the context of figure 7.8, and this information passes on to the pressure gradient. At this moment, we are unable to predict to what extent this specific feature will generalize. A further quantity that is physically related to dissipative processes in the flow is the norm of the mean velocity gradients, computed for incompressible flows as the squared norm of the vector:

$$\boldsymbol{\Omega} = \frac{1}{2} \nabla \times \mathbf{U}.$$

The performance achieved using $|\nabla P|$ and Ω^2 are comparable.

Pressure itself has the potential for being a good indicator for pathologies, as the evolution of pressure along the passageways keeps track of localized narrowings. However, in incompressible flows only pressure differences are significant; the choice of the reference pressure is not irrelevant. Simply using the regional average P^* of pressure implies using pressure differences with the outlet, because of the reference pressure $P_{out} = 0$ at the outflow boundary. Using the regional average of $P - P_{out}$ as a feature leads to a 68% of accuracy. The anatomy of the human nose includes two important non-pathological narrowings, i.e. the nasal valve in the anterior part, placed upstream of S1, and the laryngeal striction at the back, placed downstream of S6. The significant pressure jumps caused by these narrowings might confuse the classifier through their non-pathological variability. The feature $P^* - P_{out}^*$ is not sensitive to the pressure jump at the nasal valve, but contains the pressure jump at the larynx. The dual approach of including the nasal valve jump while excluding the larynx jump leads to the feature $P_{in}^* - P^*$, which improves the accuracy to 76%, suggesting that the nasal valve jump, placed upstream of the surgically interesting area, has a larger predictive capability.

Both the extremal pressure jumps can be removed from the picture, by computing pressure differences with respect to either S1 or S6. In this way, however, the features in either plane become identically zero, and the global number of non-trivial inputs reduces to ten. Both features $P_1^* - P^*$ and $P^* - P_6^*$ performs closely to $P_{in}^* - P^*$, the first one with just slight better F1-score, the latter with slightly worse accuracy and F1-score. This result remarks the importance of excluding the throat jump; however, losing two features is detrimental to the prediction.

The turbulent viscosity does not perform well as a feature, and this is hardly surprising since the present database is computed with high-resolution LES. In general, using ν_t as a feature would be a delicate choice, as this quantity in LES is highly mesh-dependent. For the present cases, run at the same spatial resolution, ν_t is at least comparable across cases, but its values are extremely small, since the simulations nearly resolve all flow scales.

Finally, a special mention is deserved by the nasal resistance R , which is in fact the hydraulic resistance of the duct and perhaps the simplest and most fundamental way of characterizing the passageways from an aerodynamic point of view. R is computed

	Mean streamlines				Instantaneous streamlines			
	accuracy	precision	recall	F1	accuracy	precision	recall	F1
$\int \Omega^2 ds$	0.81	0.76	0.62	0.68	0.83	0.82	0.62	0.71
$\int v_t ds$	0.68	0.52	0.52	0.52	0.61	0.41	0.41	0.41
$\int \nabla P ds$	0.67	0.51	0.51	0.51	0.70	0.54	0.54	0.54
$\int U ds$	0.60	0.42	0.49	0.49	0.72	0.58	0.50	0.54

Table 7.6: Binary classification E5 of the noses dataset (LOO validation), for various flow quantities evaluated over mean (left) and instantaneous (right) streamlines.

separately for the two right/left fossae as

$$R_{r/\ell} = \frac{P_{1,r/\ell}^* - P_{6,r/\ell}^*}{Q_{r/\ell}};$$

where Q is the flow rate, and the subscript r/ℓ indicates quantities computed for the right/left fossa. In the expression above, the pressure drop is taken across S1 and S6, and thus gets rid of the (unwanted) extra losses localized at the nasal valve (upstream S1) and at the larynx (downstream S6).

While R is certainly an adequate quantity for the characterization of a duct, its classification performance is quite poor: with 64% accuracy and 56% F1-scores, it is the feature that performs worst. This can be, at least in part, attributed to the fact that R provides only two inputs. However, there is a more substantial argument: the functionally healthy nose is, in fact, way more elusive than a nose with low R (Radulescu *et al.*, 2019); the unsatisfactory performance of R does not surprise the informed ENT doctor. However, it is undeniable that the hydraulic characteristics of the nose is certainly part of the picture, and this will be confirmed later on in §7.5.4, where it will be shown that using R on top of other features provide the NN with useful extra information.

7.5.3 Streamlines

Probing the flow field by means of (mean or instantaneous) streamlines is an interesting option made possible by the convective nature of the flow. Using the streamlines dispenses with the need to arbitrarily prescribe a set of cross-sections to extract features from, and provides an intuitive and convenient way of automatically sampling the entire flow field.

As mentioned in §7.3.4, one thousand points on the inlet sphere are randomly selected to be the seeds for streamlines. These are computed on a provided, mean or instantaneous, complete velocity field, and traced towards the outlet, where they separated into "left" and "right" groups. A selected flow quantity is then integrated along each streamline, from the point at $s = 0$, where the streamline intersects S1, to the point at $s = 1$, where the streamline intersects S6. The mean and the variance of these averaged values are computed over each group of streamlines; thus, streamlines end up providing four features.

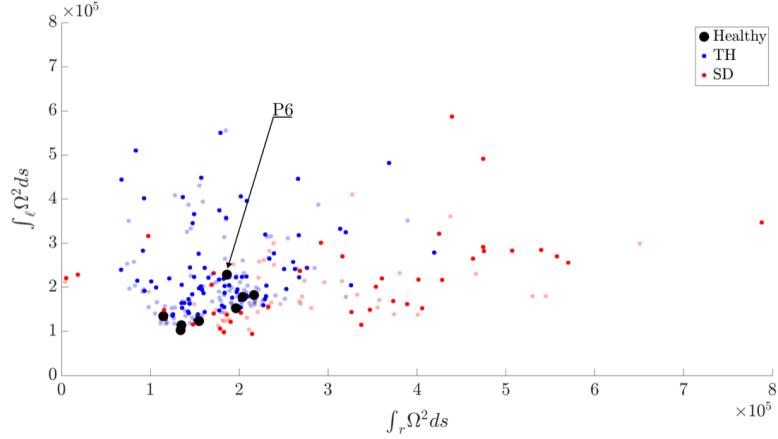


Figure 7.11: Integral of Ω^2 over instantaneous streamlines: mean values on the right/left group of streamlines are plotted on the horizontal/vertical axes. Black dots are the healthy patients, red and blue symbols represent SD and TH. Full/light colour markers indicate severe/mild pathologies.

Table 7.6 shows results of the binary classification experiment E5, where the four input features are computed by integrating various flow quantities over the streamlines. Comparing these results with those reported in table 7.5 obtained with regional averages, it is noticed that the quality of the streamlines-based prediction remains generally high, even though the number of features is significantly reduced, from 12 to 4. Moreover, the features that perform well apparently take some advantage from being computed over an instantaneous streamline rather than on a mean one. The best performing feature is the integral of Ω^2 , which yields significantly better results than the corresponding regionally-averaged values, improving from 74% to 83%. However, a striking result is the large loss of accuracy of U , which was the best feature once regionally averaged, whereas it yields only 72% accuracy when integrated over the instantaneous streamlines (and 60% over the mean ones). Using instantaneous and mean streamlines involves feature-dependent differences. The quantity U^* , especially in view of the regional average, is a very large-scale quantity that is expected to change gradually between a pair of adjacent cross-sections; the non-uniform sampling of the cross-sectional area implied by integrating U over streamlines degrades performance significantly. On the other hand, the smaller-scale quantity Ω^2 is positively affected by the ability of sampling continuously the volume from S1 to S6.

In figure 7.11 the mean value for the integral of Ω^2 for the right (horizontal axis) and left (vertical axis) groups of streamlines is plotted. Healthy patients (shown with black dots) tend to cluster, regardless of their significant anatomical differences. The colored symbols scattered around the healthy cluster correspond to pathological cases, and are organized in a way that illustrates why this feature is rather effective at discriminating between SD (red symbols) and TH (blue symbols). At odds with the area, for which

	pathology on random side		pathology on right side	
	accuracy	F1	accuracy	F1
U	0.86	0.80	0.85	0.80
$ \nabla P $	0.70	0.65	0.76	0.71
Ω^2	0.72	0.65	0.74	0.68
v_t	0.62	0.52	0.67	0.57
A	0.83	0.78	0.81	0.77

Table 7.7: Effect of random right-left switch for E5 (LOO validation).

mild and severe pathologies are by definition not overlapping, here it is interesting to note that severe (filled symbols) and mild (light symbols) pathologies are intermingled. A close inspection reveals that one black dot is a bit further from the healthy cluster than the others. It corresponds to patient P6, which has been already noted, when commenting figure 7.9, to be peculiar in terms of U^* . Indeed, a clinical re-evaluation of the CT scan for P6 reveals that its anatomy is affected by a minor hypertrophy at the tail of the inferior turbinate, that escaped the clinical screening during the preliminary stage of the study. This is unfortunate, because it adds a small amount of unwanted noise to the entire dataset; at the same time, however, it constitutes an indirect indication of the sensitivity of this feature which has alerted us on a functional information that escaped the analysis of expert clinicians.

7.5.4 Robustness of results

The final set of considerations revolves around the robustness of our results. A first concern is the potential bias deriving from the pathologies being always injected to the right fossa, a deterministic pattern which is certainly absent in reality. To assess its importance, the experiments are repeated by switching left and right (for the whole six sections) for one randomly selected half of the observations. This is accomplished quite simply, by switching the six input nodes pertaining to the right regionally-averaged values with the six pertaining to the left. Results for the binary classification experiment E5 with random right/left switching are reported in table 7.7. Small differences because of statistical noise are expected. In fact, we attribute the minor improvement of U^* , or the minor deterioration of Ω^2 to such noise. However, other features such as $|\nabla P|$ and v_t present a measurable degradation. From the ML point of view, this features still is unable to generalise, and this may be a symptom that will be cured by a larger dataset.

Another interesting question concerns the size of the dataset. More precisely, since the dataset is clearly way too small, the question concerns the adequacy of the balance between the number of input features and the number of available observations. In an ideal setting, the performance of the classifier should keep improving while the size of the dataset grows. However, since each new patient brings in a different anatomy, the robustness of the classifier against the anatomical variability is not obvious: adding a new patient might even deteriorate the performance, provided that not enough anatomical variability was learned. To address this concern within the constraints of the available dataset, the number of patients can be gradually reduced. Hence, the

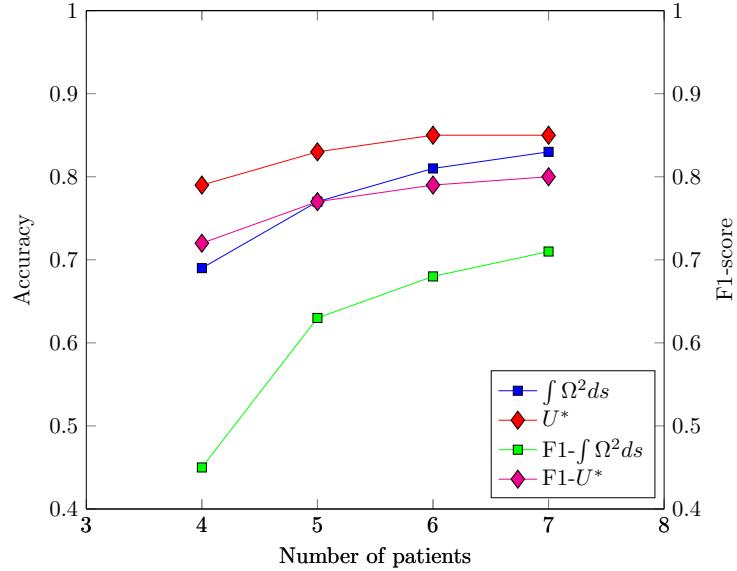


Figure 7.12: Maximum accuracy (filled symbols) and F1 score (empty symbols) versus the number of healthy patients in the dataset, experiment E5. The considered features are U^* (squares) and integral of Ω^2 over instantaneous streamlines (circles).

classification experiment E5 is repeated first with 6 patients, then 5 and finally 4. Performance is plotted in figure 7.12, for U^* and the integral of Ω^2 on instantaneous streamlines. The outcome is clear: every new patient leads to increased accuracy, with a learning curve that monotonically increases with the number of patients used for training. The growth rate, though, does saturate, especially the one for U^* , suggesting that the chosen balance between the number of features and the size of the dataset is appropriate.

The last comment is about our choice of presenting, so far, the analysis of separate features. This choice is motivated by the intent to preserve explainability of the results, and also to avoid overfitting. However, bar these considerations and the necessary balance between the number of features and the size of the dataset, nothing precludes the use of multiple, combined features.

The noses dataset is definitely still of inadequate size for any production-type study, and we believe that it is pointless, at this stage, to try and consider the best way of combining together the available features. However, a simple test which does not alter the number of features significantly is checking whether the use of lateral nasal resistance (two additional inputs only) can improve performance of regional averaging. The test is simple, and the result remarkable: repeating E5 with the left/right values of R used together with the 12 values of U^* increases the accuracy from 85% to 92%, and F1 from 80% to 89%. This is an impressive improvement, and provides just a glimpse of the further steps forward that will become possible once the methodological approach presented here is consolidated. Moreover, it comes as a confirmation that

the hydraulic resistance, while not being the indicator of a healthy nose, is definitely important within the many functions of a healthy human nose.

7.6 Conclusion

This work has described a neural network aimed at the automatic classification of shapes of solid bodies immersed in a moving fluid, based upon information gathered from the surrounding flow field only. Two example applications have been discussed, concerning the flow around airfoils and the air flow in the human nose. Besides the obviously different geometrical complexity, the two applications also differ in the flow model employed for the CFD solution: the flow around the two-dimensional airfoils is solved with RANS, while the flow in the complex three-dimensional anatomy of the noses is solved with high-resolution LES. In both cases, the natural or "healthy" variability of the shapes considered in the dataset is accompanied by a more significant type of "pathological" variability, given by geometrical defects of the airfoils or by alterations of the otherwise healthy anatomies which imply impaired breathing. The goal of the classifier is to discriminate pathological defects and physiological variability; the main conclusion of the work is that features not based directly on the geometry itself, but indirectly derived from the CFD-computed flow field that develops around the geometry, are advantageous in terms of efficiency of training and accuracy of classification.

Two large databases have been built, containing $\mathcal{O}(10^4)$ observation for the airfoils but only $\mathcal{O}(10^2)$ for the noses. The latter contains a limited amount of healthy variability, as it is derived from the CT scans of 7 healthy patients only. For the airfoil dataset, it is relatively easy to provide accurate labels for the geometrical defects added to standard NACA four-digits airfoils. Unique labelling of the noses, instead, is far from trivial. We have therefore identified a set of elementary pathologies, defined in terms of the surgical manoeuvre required to correct them, which we inject onto the otherwise healthy anatomy with a process of virtual anti-surgery. Regardless of the efficient solution implemented for carrying out hundreds of such virtual anti-surgeries, the ability to provide an anatomy with a unique label is an important achievement of the present work, which achieves this by working with anatomies that are at the same time patient-specific (extremely realistic) and artificial (not derived from a CT scan directly).

How to translate the massive amount of data contained in the CFD solution into a handful of informative feature has been the core of the work. Flow features with fluid mechanical significance have been selected and tested in comparative form, spanning from regional averages (where flow variables are averaged over portions of the computational domain) to the line integral of quantities over (mean or instantaneous) streamlines. A fluid mechanical understanding of their classification performance has been sought. Although the work is still at a preliminary stage, very high classification accuracies are already possible in the clinically challenging problem of the noses. Using regionally-averaged velocities complemented by the lateral nasal resistances (for a total of 14 inputs only) yields a classification accuracy well above 90%. It has been shown, moreover, that we are still limited by the small size of the dataset.

For it to become usable in clinical applications, the procedure definitely needs a more extended dataset; however, there is no fixed rule to make an easy prediction on the

final size of the dataset. This will depend on the desired accuracy and the complexity of the net: a NN that makes a finer classification will inevitably have more parameters to tune, thus requiring a larger dataset. Furthermore, right now, it is not easy to gauge how much anatomical variability is present in the dataset. A test is run in section §7.5.4, where the robustness of the classifier against anatomical variability is tested by adding new patients one by one. The accuracy and F1-score keep improving when enlarging the dataset, showing that for this set of features, enough anatomical variability has been learned. However, these robustness results are not conclusive yet, hence the requirement to extend the dataset in order to contain more anatomical variability (more healthy patients) and cover the entire set of elementary pathologies. However, an important step that needs to be addressed before scaling up the procedure concerns the degree of fidelity required by the CFD model to provide highly informative features. In fact, once it is realized that the ultimate interest is not the accurate prediction of the flow field, there is no reason to assume that high-fidelity methods as DNS or LES are required for an accurate classification of nasal pathologies.

Chapter 8

Through The Back Door: Expiratory Accumulation of SARS-CoV-2 in the Olfactory Mucosa as Mechanism for CNS Penetration

8.1 Abstract

Severe Acute Respiratory Syndrome Coronavirus 2 is a respiratory virus supposed to enter the organism through aerosol or fomite transmission to the nose, eyes and oropharynx. It is responsible for various clinical symptoms, including hyposmia and other neurological ones. Current literature suggests the olfactory mucosa as a port of entry to the Central Nervous System, but how the virus reaches the olfactory groove is still unknown. Because the first neurological symptoms of invasion (hyposmia) do not correspond to first signs of infection, the hypothesis of direct contact through airborne droplets during primary infection and therefore during inspiration is not plausible. The aim of this study is to evaluate if a secondary spread to the olfactory groove in a retrograde manner during expiration could be more probable.

Four three-dimensional virtual models were obtained from actual Computed Tomography scans and used to simulate expiratory droplets. The volume mesh consists of 25 million of cells, the simulated condition is a steady expiration, driving a flow rate of 270 ml/s, for a duration of 0.6 seconds. The droplet diameter is of $5 \mu\text{m}$.

The analysis of the simulations shows the virus to have a high probability to be deployed in the rhinopharynx, on the tail of medium and upper turbinates. The possibility for droplets to access the olfactory mucosa during the expiratory phase is lower than other nasal areas, but consistent.

The data obtained from these simulations demonstrates the virus can be deployed in the olfactory groove during expiration. Even if the total amount in a single act is scarce, it must be considered it is repeated tens of thousands of times a day, and the source of contamination continuously acts on a timescale of several days. The present results also imply Central Nervous System penetration of Severe Acute Respiratory Syndrome Coronavirus 2 through olfactory mucosa might be considered a complication and, consequently, prevention strategies should be considered in diseased patients.

8.2 Introduction

Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) is a respiratory virus, still widely spreading throughout the globe that is presumed to enter the organism through aerosol or fomite transmission to the nose, eyes and oropharynx (van Doremalen *et al.*, 2020; Kwok *et al.*, 2015). Presentation ranges from respiratory symptoms including cough and fever to neurological symptoms like headache, dizziness and hyposmia, showing different target organs of the virus (Cui *et al.*, 2020; Tong *et al.*, 2020). Current literature has recently started to study access points into the Central Nervous System (CNS) and the anatomical proximity between neurons, nerve fibres and the mucosa within the olfactory groove (Doty, 2008); the reported clinical-neurological signs related to alteration in smell suggest SARS-CoV-2 exploits this neuro-mucosal interface as a port of entry. Even though early reports (Bulfamante *et al.*, 2020) are indeed supporting this hypothesis through autopsy sampling, no literature exists as to how the SARS-CoV-2 reaches the mucosa at the level of the olfactory cleft, and whether the olfactory mucosa involvement is a direct consequence of viral particle deposition or due to a secondary viral invasion of these tissues during the infection.

From other respiratory viruses we know that aerosols, which are responsible for the transmission of airborne microorganisms, consist of small droplet nuclei ($1\text{--}5\ \mu\text{m}$) or droplets ($>5\ \mu\text{m}$) (Zemouri *et al.*, 2017); these have specific characteristics regarding their distribution inside the nose and respiratory tract. Considering that hyposmia more commonly follow the first signs of presentation of infection (Lechien *et al.*, 2020; Giacomelli *et al.*, 2020), the hypothesis of direct contact through airborne at the stage of primary infection and therefore during inspiration is not plausible.

The second hypothesis of a secondary spread to the olfactory groove in a retrograde manner during for example expiration in an already challenged organism appears more likely (Zhao *et al.*, 2004). This would make CNS penetration a complication secondary to e.g. pulmonary infection, thus opening the field to so far unconsidered preventative measures.

Our group has therefore used computational fluid dynamics to study distribution of airflow and deposition of supposed infectious sub-micron droplets during breathing, to better understand the possible routes of infection and penetration inside the nasal cavity and the olfactory mucosa.

8.3 Materials and Methods

This study was granted exemption from the Institutional Review Board of the San Paolo Hospital, Milano, Italy, due to its retrospective nature and is based on a set of Computational Fluid Dynamics (CFD) simulations of breathing, where only expiration is considered. The Large Eddy Simulation (LES) technique on Computed Tomography (CT) scan reconstruction of nasal anatomy is used. LES is a high-cost and high-fidelity CFD approach, which allows fine control over the modelling error in dealing with complex and possibly turbulent flows.

LES numerical simulations were performed starting from a set of four CT scans, whose sinonasal anatomy was defined by consensus by all authors as lacking any appreciable anatomic anomaly (i.e. a straight septum, normotrophic turbinates with orthodox bending, symmetrical distribution of anatomical features among the two sinonasal emi-systems).

CT scans have a 512×512 matrix with a 0.49×0.49 mm spatial resolution in the sagittal-coronal plane and a 0.625 mm gap between consecutive axial slices, with 250-350 native images for each case. More details on CT image processing, choice of the threshold value and 3D reconstruction, carried out via the software 3D Slicer (Fedorov *et al.*, 2012), have been already reported in the literature documenting the entire procedure (Saibene *et al.*, 2020; Buijs *et al.*, 2019; Covello *et al.*, 2018; Quadrio *et al.*, 2014, 2016). The CFD simulations were carried out with the finite-volumes OpenFOAM software package (Weller *et al.*, 1998).

The CFD analysis of each of the four cases (patients from P1 to P4) was conducted on a finely discretized volume mesh consisting of 25 million of cells, yielding extreme accuracy. The simulated condition is a steady expiration driving a flow rate of 270 ml/s, which corresponds to low to medium intensity breathing (Calmet *et al.*, 2016), for a duration of 0.6 seconds. A considerable number of droplets with a diameter of $5 \mu\text{m}$, in accordance with the expected droplet size described above (Giacomelli *et al.*, 2020), were placed at the lower boundary of the computational domain and allowed to enter as time progresses. Droplets are transported by the airflow, and most of them are exhaled after travelling through the nasal chamber, becoming responsible for the potential airborne contagion. Due to their inertia, however, a fraction of the droplets deposits on the mucosal lining of the nose. The simulations identify the deposited droplets, and therefore provide a quantitative representation of the deposition pattern, highlighting areas of preferential deposition during expiration. Particular attention is given to the particles that reach the olfactory slit, qualitatively sketched in figure 8.1. This study provides information on the preferential site of adhesion of expiratory droplets to the olfactory mucosa and computes the spatially varying degree of probability for a droplet to deposit in a specific location instead of being convected to the external ambient.

8.4 Results

The simulations portrait the preferential sites of droplet deposition on the nasal mucosa during expiration. It is clearly visible in figures 8.2, 8.3 and 8.4 that, although virus deposition is prevalent in the nasal vestibule and rhinopharynx. Some droplets indeed

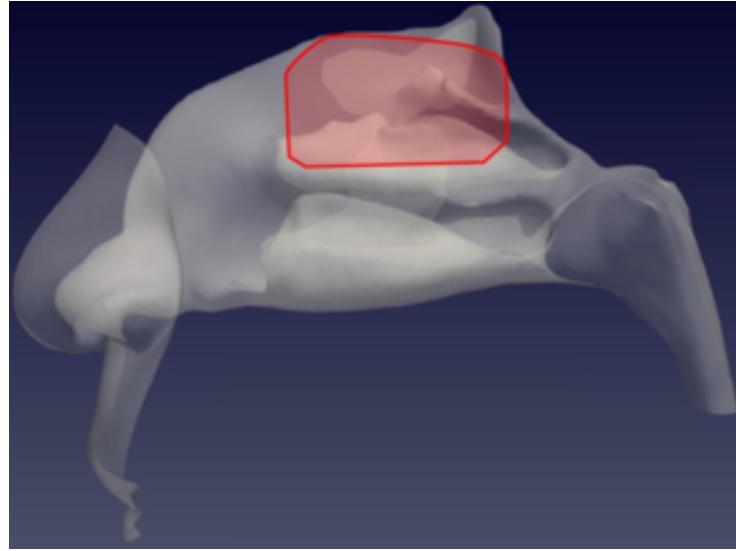


Figure 8.1: Sagittal view of the left nostril. The olfactory slit is highlighted in red.

do deposit in the area corresponding to the olfactory mucosa. Moreover, as expected, interindividual differences are visible. Droplets have been emphasized with a red dot.

The analysis of sagittal sections (figure 8.2, where only the left nasal fossa is shown for clarity) shows the virus to have a high probability to be deployed in the rhinopharynx, on the tail of medium and upper turbinates. The possibility for droplets to access the olfactory mucosa during the expiratory phase is of primary interest. The evaluation of axial projections (figure 8.3) confirms a high concentration of particles in the posterior segments in addition to a better visualization of particle distribution between medial and lateral compartments. Although heterogeneously, it can be observed how the particles, and consequently the virus, are more likely to settle in the medial quadrants of the nasal cavities than in lateral ones. During the expiratory phase, the particles have a significant probability to impact and adhere to the mucosa of the ethmoidal rostrum. Finally, the analysis of coronal projections (figure 8.4) confirms previous observations, although it demonstrates a better visualization of the septal rostrum region and of both portions of the rhinopharynx. Coronal projections confirm indeed some particles reach the olfactory slit.

8.5 Discussion

Regardless of the mechanism for viral transmission (direct respiratory, aerosol or fomite), first access to the nose must happen through inspiration. Once the virus has gained entry to the sinonasal cavity, however, many potential mechanisms concur to further diffusion, among which transport, local replication, and invasion of proximal structures. The ability of SARS-CoV-2 to bind the angiotensin converting enzyme

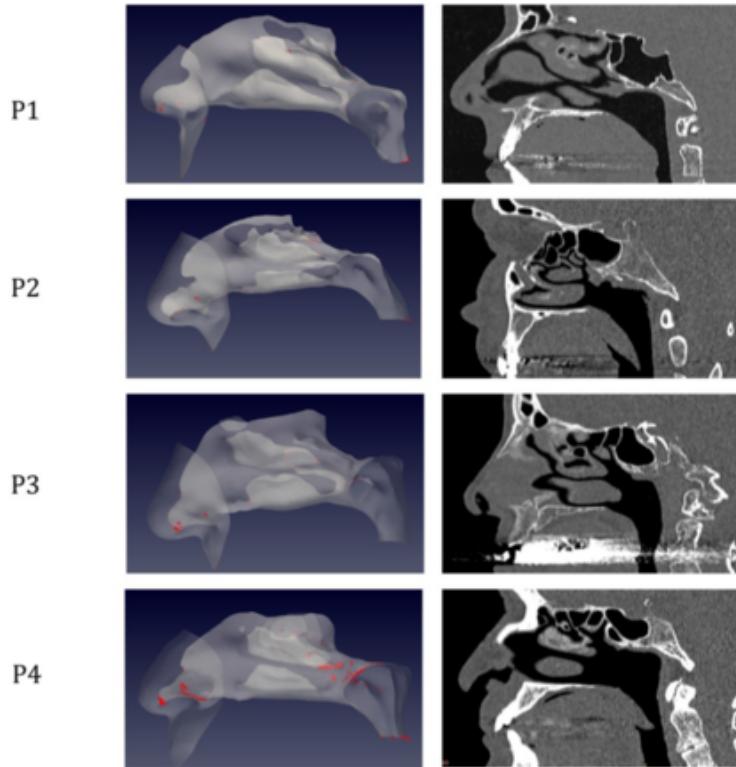


Figure 8.2: Distribution of droplet deposition during expiration, in sagittal projection. 3D models of the nasal fossae (left column), obtained from the CT, which are shown in the right panel. Particle size is increased to improve clarity.

receptor, enter the respiratory epithelium cells and thereby initiate its replication has been thoroughly demonstrated (Hui *et al.*, 2020; Hoffmann *et al.*, 2020).

Respiratory droplets containing viral particles are unable to massively reach the olfactory cleft, which should not be therefore considered a primary target for Coronavirus Disease 2019 (COVID-19) infection. The droplet ability to deposit on the olfactory cleft is a direct function of the particle size, given that the olfactory cleft is anatomically developed to receive smaller particles like odorants, while droplets carrying the viral load can be larger (Zhao *et al.*, 2004). Such an ineffective viral deposition onto the olfactory mucosa, coupled with the known defensive mechanisms employed by the olfactory mucosa to protect from environmental noxae (Hui *et al.*, 2020; Herbert *et al.*, 2012), make the direct infection of the olfactory cleft by SARS-CoV-2 at the time of primary entry into the organism unlikely at best. Conversely, cumulative exposure of the olfactory cleft to expiratory droplets from the lower respiratory tract in an already diseased organism may be more likely. The viral load in the lung is much higher than a one-time aerosol reaching the nose, the act of expiration is repeated tens of thousands of times a day, and the source of contamination continuously acts on a timescale of

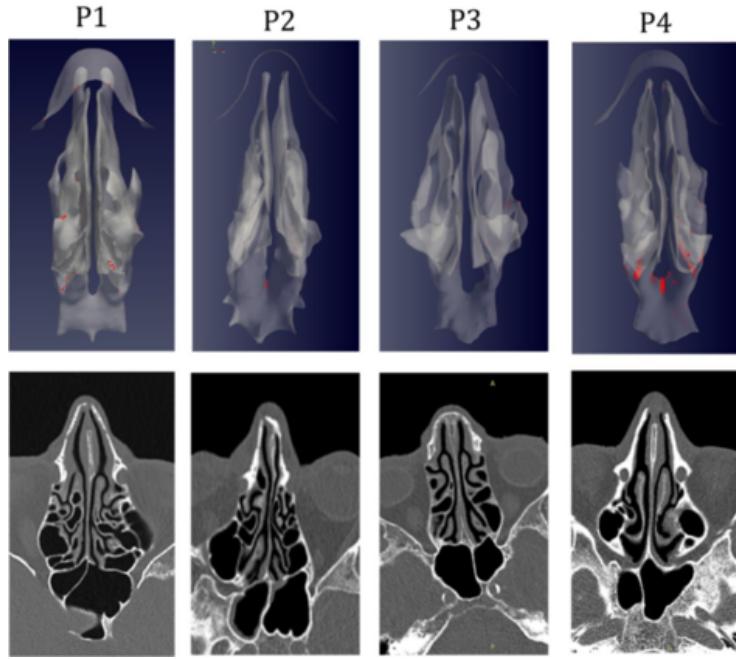


Figure 8.3: As in figure 8.2, but axial projection

several days. This route to the olfactory cleft and maybe to CNS may also explain the time lag between first symptoms and first neurological impairments including hyposmia (Ahmad & Rathore, 2020; Niazkar *et al.*, 2020).

The present results also imply CNS penetration of SARS-CoV-2 through olfactory mucosa might be a complication of an already present infection of the lower respiratory tract. Hence, prevention of the olfactory mucosa penetration by the virus should be considered in diseased patients. High volume nasal washes, usually performed with saline, can be used to reduce the adherence of viral parts emitted from the lower respiratory tract towards the nasal cavity, thus weakening the virus ability to spread to the olfactory mucosa. Indeed, other authors have advocated use of nasal lavages in SARS-CoV-2 infection as a preventive measure (Singh *et al.*, 2020), and prior studies on viral upper respiratory tract infections with hypertonic saline showed reduced viral shedding and patient infectivity (Ramalingam *et al.*, 2019) in the already diseased. Prior studies on viral upper respiratory tract infections with hypertonic saline showed decreased viral shedding and reduced patients' infectivity (Ramalingam *et al.*, 2019) in the already diseased. Others propose other types of medications: in Pianta *et al.* (2020) inhalation of acetic acid is suggested being effective in shortening the duration of symptoms. The present study – besides suggesting the olfactory region as a target for inhibition of the secondary viral infection which endangers the CNS – provides further support for the effectiveness of such preventive measures, since a diffuse droplet deposition takes place in the nasal fossae, that can be easily reached by washing or by other nasal medications.

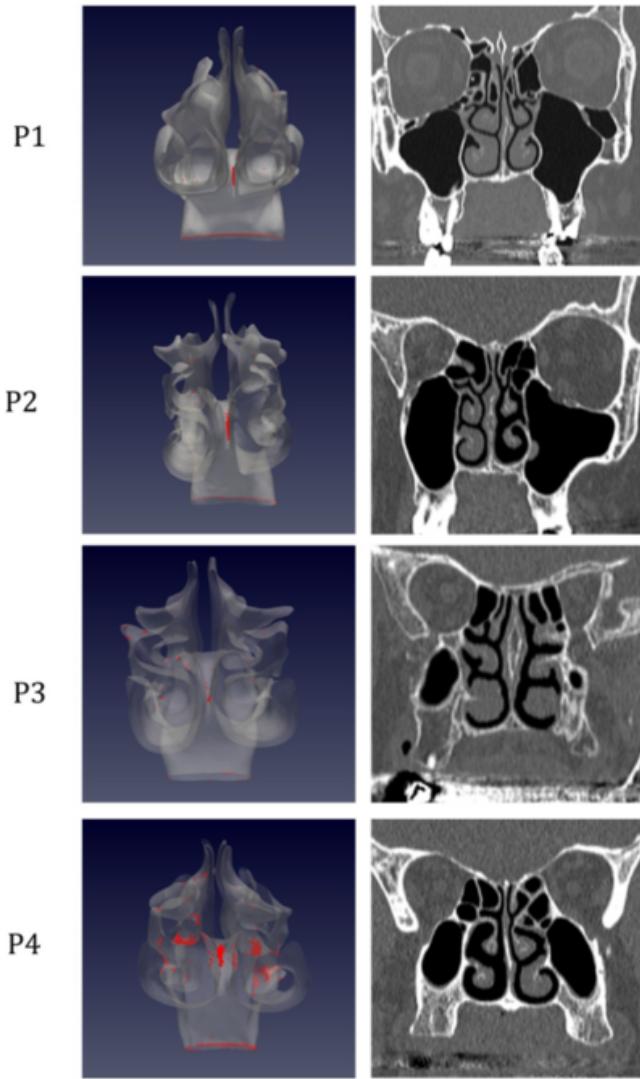


Figure 8.4: As in figure 8.2, but coronal projection.

Further studies are thus required to focus on nasal washes not only as a preventive measure to infection but also as a means to inhibit the secondary spreading of the virus to the olfactory mucosa and therefore to the CNS for COVID-19 patients. These studies should clinically quantify the ability of nasal washes, i.e. a simple and non-invasive treatment, to halt the progression of the disease by containing its complications involving the CNS.

Bibliography

- ABBOTT, I.H. & von DOENHOFF, A.E. 2012 *Theory of Wing Sections: Including a Summary of Airfoil Data*. Courier Corporation.
- ABUCIDE-ARMAS, A., PORTAL-PORRAS, K., FERNANDEZ-GAMIZ, U., ZULUETA, E. & TESO-FZ-BETOÑO, A. 2021 A Data Augmentation-Based Technique for Deep Learning Applied to CFD Simulations. *Mathematics* **9** (16), 1843.
- AHMAD, I. & RATHORE, F.A. 2020 Neurological manifestations and complications of COVID-19: A literature review. *Journal of Clinical Neuroscience* **77**, 8–12.
- AHMED, S.E., PAWAR, S., SAN, O., RASHEED, A., ILIESCU, T. & NOACK, B.R. 2021 On closures for reduced order models—A spectrum of first-principle to machine-learned avenues. *Physics of Fluids* **33** (9), 091301.
- ALEXANDERSEN, J. & ANDREASEN, C.S. 2020 A Review of Topology Optimisation for Fluid-Based Problems. *Fluids* **5** (1), 29.
- ALSALMAN, M., COLVERT, B. & KANSO, E. 2018 Training bioinspired sensors to classify flows. *Bioinspiration & Biomimetics* **14** (1), 016009.
- BARBARITE, E., GADKAREE, S.K., MELCHIONNA, S., ZWICKER, D. & LINDSAY, R.W. 2021 Computational Fluid Dynamics Modeling of Nasal Obstruction and Associations with Patient-Reported Outcomes. *Plastic and Reconstructive Surgery* **148** (4), 592e–600e.
- BERGER, M., GIOTAKIS, A.I., PILLEI, M., MEHRLE, A., KRAXNER, M., KRAL, F., RECHEIS, W., RIECHELMANN, H. & FREYSINGER, W. 2021 Agreement between rhinomanometry and computed tomography-based computational fluid dynamics. *International Journal of Computer Assisted Radiology and Surgery* **16**, 629–638.
- BERNARDINI, M., PIROZZOLI, S., QUADRI, M. & ORLANDI, P. 2013 Turbulent channel flow simulations in convecting reference frames. *Journal of Computational Physics* **232** (1), 1–6.
- BEWLEY, T. 2001 Flow Control: New Challenges for a New Renaissance. *Progress in Aerospace Sciences* **37**, 21–58.
- BEWLEY, T. & PROTAS, B. 2002 Skin friction and pressure: The "footprints" of turbulence. In *Proceedings of the 3rd Symposium on Smart Control of Turbulence, Tokyo*.

- BISHOP, C.M. 2006 *Pattern Recognition and Machine Learning*. Springer.
- BOROJENI, A.A.T., GARCIA, G.J.M., MOGHADDAM, M.G., FRANK-ITO, D.O., KIMBELL, J.S., LAUD, P.W., KOENIG, L.J. & RHEE, J.S. 2020 Normative ranges of nasal airflow variables in healthy adults. *International Journal of Computer Assisted Radiology and Surgery* **15** (1), 87–98.
- BRENNER, M.P., ELDREDGE, J.D. & FREUND, J.B. 2019 Perspective on machine learning for advancing fluid mechanics. *Physical Review Fluids* **4** (100501), 1–7.
- BRÜNING, J., HILDEBRANDT, T., HEPPT, W., SCHMIDT, N., LAMECKER, H., SZENGEL, A., AMIRIDZE, N., RAMM, H., BINDERNAGEL, M., ZACHOW, S. & GOUBERGRITS, L. 2020 Characterization of the Airflow within an Average Geometry of the Healthy Human Nasal Cavity. *Scientific Reports* **10** (1), 1–12.
- BRUNTON, S.L. 2021 Applying machine learning to study fluid mechanics. *Acta Mechanica Sinica* **37** (12), 1718–1726.
- BRUNTON, S.L., NOACK, B.R. & KOUMOUTSAKOS, P. 2020 Machine Learning for Fluid Mechanics. *Annual Review of Fluid Mechanics* **52** (1), 477–508.
- BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **113** (15), 3932–3937.
- BUIJS, E.F.M., COVELLO, V., PIPOLO, C., SAIBENE, A.M., FELISATI, G. & QUADRI, M. 2019 Thermal water delivery in the nose: Experimental results describing droplet deposition through computational fluid dynamics. *ACTA Otorhinolaryngologica Italica* **39** (6), 396–403.
- BULFAMANTE, G., CHIUMELLO, D., CANEVINI, M.P., PRIORI, A., MAZZANTI, M., CENTANNI, S. & FELISATI, G. 2020 First ultrastructural autoptic findings of SARS -Cov-2 in olfactory pathways and brainstem. *Minerva Anestesiologica* **86** (6), 678–679.
- CALMET, H., GAMBARUTO, A.M., BATES, A.J., VÁZQUEZ, M., HOUZEUX, G. & DOORLY, D.J. 2016 Large-scale CFD simulations of the transitional and turbulent regime for the large human airways during rapid inhalation. *Computers in Biology and Medicine* **69**, 166–180.
- CALMET, H., INTHAVONG, K., EGUZKITZA, B., LEHMKUHL, O., HOUZEUX, G. & VÁZQUEZ, M. 2019 Nasal sprayed particle deposition in a human nasal cavity under different inhalation conditions. *PLOS ONE* **14** (9), e0221330.
- CALMET, H., INTHAVONG, K., OWEN, H., DOSIMONT, D., LEHMKUHL, O., HOUZEUX, G. & VÁZQUEZ, M. 2021 Computational modelling of nasal respiratory flow. *Computer Methods in Biomechanics and Biomedical Engineering* **24** (4), 440–458.
- CAMPBELL, D.A., MOGHADDAM, M.G., RHEE, J.S. & GARCIA, G.J.M. 2021 Narrowed Posterior Nasal Airway Limits Efficacy of Anterior Septoplasty. *Facial Plastic Surgery & Aesthetic Medicine* **23** (1).

- CANONICA, G.W., BOUSQUET, J., MULLOL, J., SCADDING, G.K. & VIRCHOW, J.C. 2007 A survey of the burden of allergic rhinitis in Europe. *Allergy* **62** (s85), 17–25.
- CASEY, K.P., BOROJENI, A.A.T., KOENIG, L.J., RHEE, J.S. & GARCIA, G.J.M. 2017 Correlation between Subjective Nasal Patency and Intranasal Airflow Distribution. *Otolaryngology—Head and Neck Surgery* **156** (4), 741–750.
- CHANDRA, R.K., PATADIA, M.O. & RAVIV, J. 2009 Diagnosis of Nasal Airway Obstruction. *Otolaryngologic Clinics of North America* **42** (2), 207–225.
- CHEN, X.B., LEE, H.P., CHONG, V.F.H. & WANG, D.Y. 2009 Assessment of septal deviation effects on nasal air flow: A computational fluid dynamics model. *American Laryngological Rhinological and Otological Society* **119**, 1730–1736.
- CHEROBIN, G.B., VOEGELS, R.L., PINNA, F.R., GEBRIM, E.M.M.S., BAILEY, R.S. & GARCIA, G.J.M. 2020 Rhinomanometry Versus Computational Fluid Dynamics: Correlated, but Different Techniques. *Am J Rhinol Allergy* **35** (2), 245–255.
- CHOI, E., BAHADORI, M.T., SCHUETZ, A., STEWART, W.F. & SUN, J. 2016 Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. *JMLR workshop and conference proceedings* **56**, 301–318.
- CHOI, H., JEON, W.-P. & KIM, J. 2008 Control of Flow Over a Bluff Body. *Annual Review of Fluid Mechanics* **40** (1), 113–139.
- CHOI, H. & MOIN, P. 1994 Effects of the Computational Time Step on Numerical Solutions of Turbulent Flow. *J. Comp. Phys.* **113**, 1–4.
- CHOI, J., TAWHAI, M.H., HOFFMAN, E.A. & LIN, C.L. 2009 On intra- and intersubject variabilities of airflow in the human lungs. *Physics of Fluids* **21** (10), 101901.
- CHOWDHURY, N.I., SMITH, T.L., CHANDRA, R.K. & TURNER, J.H. 2019 Automated classification of osteomeatal complex inflammation on computed tomography using convolutional neural networks. *International Forum of Allergy & Rhinology* **9** (1), 46–52.
- CHUNG, S.K., SON, Y.R., SHIN, S.J. & KIM, S.K. 2006 Nasal Airflow during Respiratory Cycle. *American Journal of Rhinology* **20** (4), 379–384.
- CHURCHILL, S.E., SHACKELFORD, L.L., GEORGI, J.N. & BLACK, M.T. 2004 Morphological variation and airflow dynamics in the human nose. *American Journal of Human Biology* **16** (6), 625–638.
- CLARK, D.W., DEL SIGNORE, A.G., RAITHATHA, R. & SENIOR, B.A. 2018 Nasal Airway Obstruction: Prevalence and Anatomic Contributors. *Ear, Nose & Throat Journal* **97** (6), 173–176.
- CLEMENTS, P.A. & GORTDS, F. 2005 Standardisation Committee on Objective Assessment of the Nasal Airway, IRS, and ERS Consensus report on acoustic rhinometry and rhinomanometry. *Rhynology* **43** (3), 169–179.

- COLE, P. 1998 Physiology of the nose and paranasal sinuses. *Clinical Reviews in Allergy & Immunology* **16** (1), 25–54.
- COSMO, L., RODOLÀ, E., MASCI, J., TORSERLO, A. & BRONSTEIN, M.M. 2016 Matching Deformable Objects in Clutter. In *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 1–10.
- COVELLO, V., PIPOLO, C., SAIBENE, A., FELISATI, G. & QUADRI, M. 2018 Numerical simulation of thermal water delivery in the human nasal cavity. *Computers in Biology and Medicine* **100**, 62–73.
- CREMADES, A., HOYAS, S., DESHPANDE, R., QUINTERO, , LELLEP, M., LEE, W.J., MONTY, J., HUTCHINS, N., LINKMANN, M., MARUSIC, I. & VINUESA, R. 2023 Identifying regions of importance in wall-bounded turbulence through explainable deep learning, arXiv: 2302.01250.
- CROWSON, M.G., RANISAU, J., ESKANDER, A., BABIER, A., XU, B., KAHMKE, R.R., CHEN, J.M. & CHAN, T.C.Y. 2020 A contemporary review of machine learning in otolaryngology–head and neck surgery. *The Laryngoscope* **130** (1), 45–51.
- CUI, C., YAO, Q., ZHANG, D., ZHAO, Y., ZHANG, K., NISENBAUM, E., CAO, P., ZHAO, K., HUANG, X., LENG, D., LIU, C., LI, N., LUO, Y., CHEN, B., CASIANO, R., WEED, D., SARGI, Z., TELISCHI, F., LU, H., DENNENY, J.C., SHU, Y. & LIU, X. 2020 Approaching Otolaryngology Patients During the COVID-19 Pandemic. *Otolaryngology–Head and Neck Surgery* **163** (1), 121–131.
- DABROWSKA-BIEN, J., SKARZYNSKI, P.H., GWIZDALSKA, I., LAZECKA, K. & SKARZYNSKI, H. 2018 Complications in septoplasty based on a large group of 5639 patients. *European Archives of Oto-Rhino-Laryngology* **275** (7), 1789–1794.
- DECONDE, A.S. & SOLER, Z.M. 2016 Chronic Rhinosinusitis: Epidemiology and Burden of Disease. *American Journal of Rhinology & Allergy* **30** (2), 134–139.
- DENOURL, E., ROUSSEL, L.O., WOO, A.S., BOYAJIAN, M. & CROZIER, J. 2020 Quantification of Nasal Septal Deviation With Computed Tomography Data. *Journal of Craniofacial Surgery* **31** (6), 1659–1663.
- DILGEN, C.B., DILGEN, S.B., FUHRMAN, D.R., SIGMUND, O. & LAZAROV, B.S. 2018 Topology optimization of turbulent flows. *Computer Methods in Applied Mechanics and Engineering* **331**, 363–393.
- DINIS, P. B. & HAIDER, H. 2002 Septoplasty: Long-term evaluation of results. *American Journal of Otolaryngology* **23** (2), 85–90.
- DOORLY, D., TAYLOR, D.J., GAMBARUTO, A.M., SCHROTER, R.C. & TOLLEY, N. 2008a Nasal architecture: Form and flow. *Philosophical Transaction of the Royal Society* **366**, 3225–3246.
- DOORLY, D.J., TAYLOR, D.J. & SCHROTER, R.C. 2008b Mechanics of airflow in the human nasal airways. *Respiratory Physiology & Neurobiology* **163** (1), 100–110.

- DOTY, R.L. 2008 The olfactory vector hypothesis of neurodegenerative disease: Is it viable? *Annals of Neurology* **63** (1), 7–15.
- DUCROS, F., FERRAND, V., NICOUD, F., WEBER, C., DARRACQ, D., GACHERIEU, C. & POINSOT, T. 1999 Large-Eddy Simulation of the Shock/Turbulence Interaction. *Journal of Computational Physics* **152** (2), 517–549.
- DUCROS, F., NICOUD, F. & POINSOT, T. 1998 Wall-Adapting Local Eddy-Viscosity models for simulations in complex geometries. *Numerical Methods for Fluid Dynamics VI*.
- DURAISAMY, K., IACCARINO, G. & XIAO, H. 2019 Turbulence Modeling in the Age of Data. *Annual Review of Fluid Mechanics* **51** (1), 357–377.
- EIVAZI, H., TAHANI, M., SCHLATTER, P. & VINUESA, R. 2022 Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations. *Physics of Fluids* **34** (7), 075117.
- FARNOUD, A., TOFIGHIAN, H., BAUMANN, I., GARCIA, G.J.M., SCHMID, O., GUTHEIL, E. & RASHIDI, M.M. 2020 Large eddy simulations of airflow and particle deposition in pulsating bi-directional nasal drug delivery. *Physics of Fluids* **32** (10), 101905.
- FEDOROV, A., BEICHEL, R., KALPATHY-CRAMER, J., FINET, J., FILLION-ROBIN, J.-C., PUJOL, S., BAUER, C., JENNINGS, D., FENNESSY, F., SONKA, M., BUATTI, J., AYLWARD, S.R., MILLER, J.V., PIEPER, S. & KIKINIS, R. 2012 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* p. 22770690.
- FERZIGER, J.H. & PERIC, M. 2002 *Computational Methods for Fluid Dynamics*, 3rd edn. Springer.
- FRANK-ITO, D.O., WOFFORD, M., SCHROETER, J.D. & KIMBELL, J.S. 2016 Influence of Mesh Density on Airflow and Particle Deposition in Sinonasal Airway Modeling. *J Aerosol Med Pulm Drug Deliv* **29** (1), 46–56.
- FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2019 Super-resolution reconstruction of turbulent flows with machine learning. *Journal of Fluid Mechanics* **870**, 106–120.
- FUKAMI, K., FUKAGATA, K. & TAIRA, K. 2020 Assessment of supervised machine learning methods for fluid flows. *Theoretical and Computational Fluid Dynamics* **34** (4), 497–519.
- GAD-EL-HAK, M. 2000 *Flow Control – Passive, Active and Reactive Flow Management*. Cambridge University Press, Cambridge.
- GIACOMELLI, A., PEZZATI, L., CONTI, F., BERNACCHIA, D., SIANO, M., ORENI, L., RUSCONI, S., GERVASONI, C., RIDOLFO, A.L., RIZZARDINI, G., ANTINORI, S. & GALLI, M. 2020 Self-reported Olfactory and Taste Disorders in Patients With Severe Acute Respiratory Coronavirus 2 Infection: A Cross-sectional Study. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **71** (15), 889–890.

- GOECKS, J., JALILI, V., HEISER, L.M. & GRAY, J.W. 2020 How Machine Learning Will Transform Biomedicine. *Cell* **181** (1), 92–101.
- GOODFELLOW, I., BENGIO, Y. & COURVILLE, A. 2016 *Deep Learning*. MIT Press.
- GU, J., WANG, Z., KUEN, J., MA, L., SHAHROUDY, A., SHUAI, B., LIU, T., WANG, X., WANG, G., CAI, J. & CHEN, T. 2018 Recent advances in convolutional neural networks. *Pattern Recognition* **77**, 354–377.
- HASEGAWA, K., FUKAMI, K., MURATA, T. & FUKAGATA, K. 2020 Machine-learning-based reduced-order modeling for unsteady flows around bluff bodies of various shapes. *Theoretical and Computational Fluid Dynamics* **34** (4), 367–383.
- HASTIE, T., FRIEDMAN, J. & TIBSHIRANI, R. 2001 *The Elements of Statistical Learning*. New York, NY: Springer.
- HERBERT, R.P., HARRIS, J., CHONG, K.P., CHAPMAN, J., WEST, A.K. & CHUAH, M.I. 2012 Cytokines and olfactory bulb microglia in response to bacterial challenge in the compromised primary olfactory pathway. *Journal of Neuroinflammation* **9** (1), 109.
- HOFFMANN, M., KLEINE-WEBER, H., SCHROEDER, S., KRÜGER, N., HERRLER, T., ERICHSEN, S., SCHIERGENS, T., HERRLER, G., WU, N.H., NITSCHE, A., MÜLLER, M., DROSTEN, C. & PÖHLMANN, S. 2020 SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181** (2), 271–280.
- HÖRSCHLER, I., SCHRÖDER, W. & MEINKE, M. 2010 On the assumption of steadiness of nasal cavity flow. *Biomechanics* **43**, 1081–1085.
- HOUSER, STEVEN M. 2007 Surgical Treatment for Empty Nose Syndrome. *Archives of Otolaryngology—Head & Neck Surgery* **133** (9), 858–863.
- HUI, K.P.Y., CHEUNG, M.C., PERERA, R.A.P.M., NG, K.C., BUI, C.H.T., HO, J.C.W., NG, M.M.T., KUOK, D.I.T., SHIH, K.C., TSAO, S.W., POON, L.L.M., PEIRIS, M., NICHOLLS, J.M. & CHAN, M.C.W. 2020 Tropism, replication competence, and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva: An analysis in ex-vivo and in-vitro cultures. *The Lancet Respiratory Medicine* **8** (7), 687–695.
- ILLUM, P. 1997 Septoplasty and compensatory inferior turbinate hypertrophy: Long-term results after randomized turbinoplasty. *European Archives of Oto-Rhinolaryngology* **254** (1), S89–S92.
- INTHAVONG, K., CHETTY, A., SHANG, Y. & TU, J. 2018 Examining mesh independence for flow dynamics in the human nasal cavity. *Computers in Biology and Medicine* **102**, 40–50.
- INTHAVONG, K., DAS, P., SINGH, N. & SZNITMAN, J. 2019 In silico approaches to respiratory nasal flows: A review. *Journal of Biomechanics* **97**, 109434.

- ISLAM, M.S., PAUL, G., ONG, H.X., YOUNG, P.M., GU, Y.T. & SAHA, S.C. 2020 A Review of Respiratory Anatomical Development, Air Flow Characterization and Particle Deposition. *International Journal of Environmental Research and Public Health* **17** (2), 380.
- JAMES, G., WITTEN, D., HASTIE, T. & TIBSHIRANI, R. 2021 *An Introduction to Statistical Learning: With Applications in R*. New York, NY: Springer US.
- JAMESON, A. 1988 Aerodynamic design via control theory. *Journal of Scientific Computing* **3** (3), 233–260.
- JANOVIĆ, N., ĆOĆIĆ, A., STAMENIĆ, M., JANOVIĆ, A. & DJURIĆ, M. 2020 Side asymmetry in nasal resistance correlate with nasal obstruction severity in patients with septal deformities: Computational fluid dynamics study. *Clinical Otolaryngology* **45** (5), 718–724.
- JESSEN, M. & MALRN, L. 1997 Definition, prevalence and development of nasal obstruction. *Allergy* **52** (s40), 3–6.
- JIN, H.H., FAN, J.R., ZENG, M.J. & CEN, K.F. 2006 Large eddy simulation of inhaled particle deposition within the human upper respiratory tract. *J. of Aerosol Science* **38**, 257–268.
- JIN, X., LU, Y., REN, X., GUO, S., JIN, D., LIU, B., BAI, X. & LIU, JUNXIU 2023 Exploring the Influence of Nasal Vestibule Structure on Nasal Obstruction Using CFD and Machine Learning Method. *Medical Engineering & Physics* p. 103988.
- JONES, N. 2001 The nose and paranasal sinuses physiology and anatomy. *Advanced Drug Delivery Reviews* **51** (1), 5–19.
- KEUSTERMANS, W., HUYSMANS, T., DANCKAERS, F., ZAROWSKI, A., SCHMELZER, B., SIJBERS, J. & DIRCKX, J. J. J. 2018 High quality statistical shape modelling of the human nasal cavity and applications. *Royal Society Open Science* **5** (12), 181558.
- KIM, H., KIM, J., WON, S. & LEE, C. 2021 Unsupervised deep learning for super-resolution reconstruction of turbulence. *Journal of Fluid Mechanics* **910**.
- KINGMA, D.P. & BA, J. 2017 Adam: A Method for Stochastic Optimization, arXiv: 1412.6980.
- KOCHKOV, D., SMITH, J.A., ALIEVA, A., WANG, Q., BRENNER, M.P. & HOYER, S. 2021 Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences* **118** (21), e2101784118.
- KUTZ, J.N. 2017 Deep learning in fluid dynamics. *Journal of Fluid Mechanics* **814**, 1–4.
- KUZNETSOVA, A., ROM, H., ALLDRIN, N., UIJLINGS, J., KRASIN, I., PONT-TUSSET, J., KAMALI, S., POPOV, S., MALLOCI, M., KOLESNIKOV, A., DUERIG, T. & FERRARI, V. 2020 The Open Images Dataset V4. *International Journal of Computer Vision* **128** (7), 1956–1981.

- KWOK, Y.L.A., GRALTON, J. & McLAWS, M.L. 2015 Face touching: A frequent habit that has implications for hand hygiene. *American Journal of Infection Control* **43** (2), 112–114.
- LECHIEN, J.R., CHIESA-ESTOMBA, C.M., DE SIATI, D.R., HOROI, M., LE BON, S.D., RODRIGUEZ, A., DEQUANTER, D., BLECIC, S., EL AFIA, F., DISTINGUIN, L., CHEKKOURY-IDRISSI, Y., HANS, S., DELGADO, I.L., CALVO-HENRIQUEZ, C., LAVIGNE, P., FALANGA, C., BARILLARI, M.R., CAMMAROTO, G., KHALIFE, M., LEICH, P., SOUCHAY, C., ROSSI, C., JOURNE, F., HSIEH, J., EDJLALI, M., CARLIER, R., RIS, L., LOVATO, A., DE FILIPPIS, C., COPPEE, F., FAKHRY, N., AYAD, T. & SAUSSEZ, S. 2020 Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): A multicenter European study. *European Archives of Oto-Rhino-Laryngology* **277** (8), 2251–2261.
- LECUN, Y., BOTTOU, L., BENGIO, Y. & HAFFNER, P. 1998 Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86** (11), 2278–2324.
- LEONG, S.C., CHEN, X.B., LEE, H.P. & WANG, D.Y. 2010 A review of the implications of computational fluid dynamic studies on nasal airflow and physiology. *J. of Rhinology* **48**, 139–145.
- LERA, G. & PINZOLAS, M. 2002 Neighborhood based Levenberg-Marquardt algorithm for neural network training. *IEEE Transactions on Neural Networks* **13** (5), 1200–1203.
- LESCHZINER, M., CHOI, H. & CHOI, K.-S. 2011 Flow control approaches in aerodynamics: Progress and prospects. *Phil. Trans. R. Soc. A* **369** (1940), 1349–1351.
- LI, C., JIANG, J., DONG, H. & ZHAO, K. 2017 Computational modeling and validation of human nasal airflow under various breathing conditions. *Journal of Biomechanics* **64**, 59–68.
- LI, C.H., KAURA, A., TAN, C., WHITCROFT, K.L., LEUNG, T.S. & ANDREWS, P. 2020 Diagnosing nasal obstruction and its common causes using the nasal acoustic device: A pilot study. *Laryngoscope Investigative Otolaryngology* **5** (5), 796–806.
- LIN, H.-F., HSIEH, Y.-C. & HSIEH, Y.-L. 2020 Factors Affecting Location of Nasal Airway Obstruction. In *2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, pp. 21–24.
- LING, J., KURZAWSKI, A. & TEMPLETON, J. 2016 Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics* **807**, 155–166.
- LINTERMANN, A. & SCHRÖDER, W. 2019 A Hierarchical Numerical Journey Through the Nasal Cavity: From Nose-Like Models to Real Anatomies. *Flow, Turbulence and Combustion* **102** (1), 89–116.
- LIU, B., TANG, J., HUANG, H. & LU, X.-Y. 2020 Deep learning methods for super-resolution reconstruction of turbulent flows. *Physics of Fluids* **32** (2), 025105.

- LIU, T., HAN, D., WANG, J., TAN, J., ZANG, H., WANG, T., LI, Y. & CUI, S. 2012 Effects of septal deviation on the airflow characteristics: Using computational fluid dynamics models. *Acta Oto-Laryngologica* **132** (3), 290–298.
- LIU, Y., JOHNSON, M.R., MATIDA, E.A., KHERANI, S. & MARSAN, J. 2009 Creation of a standardized geometry of the human nasal cavity. *Journal of Applied Physiology* **106** (3), 784–795.
- LIU, Y., MATIDA, E. A., GU, J. & JOHNSON, M. R. 2007 Numerical simulation of aerosol deposition in a 3-D human nasal cavity using RANS, RANS/EIM, and LES. *Journal of Aerosol Science* **38** (7), 683–700.
- LUNDBERG, S.M. & LEE, S. 2017 A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*, , vol. 30. Curran Associates, Inc.
- MAGNET, R., BLOCH, K., TAVERNE, M., MELZI, S., GEOFFROY, M., KHONSARI, R.H. & OVSJANIKOV, M. 2023 Assessing craniofacial growth and form without landmarks: A new automatic approach based on spectral methods. *Journal of Morphology* **284** (8), e21609.
- MAGNET, R. & OVSJANIKOV, M. 2023 Scalable and Efficient Functional Map Computations on Dense Meshes. *Computer Graphics Forum* **42** (2), 89–101.
- MALIK, J., SPECTOR, B.M., WU, Z., MARKLEY, J., ZHAO, S., OTTO, B.A., FARAG, A.A. & ZHAO, K. 2021 Evidence of Nasal Cooling and Sensory Impairments Driving Patient Symptoms With Septal Deviation. *The Laryngoscope* **132** (3), 509–517.
- MARKIDIS, S. 2021 The Old and the New: Can Physics-Informed Deep-Learning Replace Traditional Linear Solvers? *Frontiers in Big Data* **4**.
- MAULIK, R. & SAN, O. 2017 A neural network approach for the blind deconvolution of turbulent flows. *Journal of Fluid Mechanics* **831**, 151–181.
- MELZI, S., REN, J., RODOLÀ, E., SHARMA, A., WONKA, P. & OVSJANIKOV, M. 2019 ZoomOut: Spectral upsampling for efficient shape correspondence. *ACM Transactions on Graphics* **38** (6), 155:1–155:14.
- MENTER, F.R. 1994 Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications. *AIAA Journal* **32** (8), 1598–1605.
- MENTER, F.R., KUNTZ, M. & LANGTRY, R. 2003 Ten Years of Industrial Experience with the SST Turbulence Model. *Turbulence, Heat and Mass Transfer* **4**, 8.
- MØLLER, M.F. 1993 A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* **6** (4), 525–533.
- MOREDDU, E., MEISTER, L., PHILIP-ALLIEZ, C., TRIGLIA, J.M., MEDALE, M. & NICOLAS, R. 2019 Computational Fluid Dynamics in the assessment of nasal obstruction in children. *European Annals of Otorhinolaryngology, Head and Neck diseases* p. 6.

- NA, Y., KIM, Y.-J., KIM, H.Y. & JUNG, Y.G. 2022 Improvements in airflow characteristics and effect on the NOSE score after septoturbinoplasty: A computational fluid dynamics analysis. *PLOS ONE* **17** (11), e0277712.
- NAFTALI, S., SCHROTER, R.C., SHINER, R.J. & ELAD, D. 1998 Transport Phenomena in the Human Nasal Cavity: A Computational Model. *Annals of Biomedical Engineering* **26** (5), 831–839.
- NAKANO, H., MISHIMA, K., UEDA, Y., MATSUSHITA, A., SUGA, H., MIYAWAKI, Y., MANO, T., MORI, Y. & UEYAMA, Y. 2013 A new method for determining the optimal CT threshold for extracting the upper airway. *Dentomaxillofacial Radiology* **42** (3), 26397438.
- NIAZKAR, H.R., ZIBAEE, B., NASIMI, A. & BAHRI, N. 2020 The neurological manifestations of COVID-19: A review article. *Neurological Sciences* **41** (7), 1667–1671.
- NICOUDE, F. & DUCROS, F. 1999 Subgrid-Scale Stress Modelling Based on the Square of the Velocity Gradient Tensor. *Flow, Turbulence and Combustion* **62** (3), 183–200.
- NOVATI, G., DE LAROUSSILHE, H.L. & KOUMOUTSAKOS, P. 2021 Automating turbulence modelling by multi-agent reinforcement learning. *Nature Machine Intelligence* **3** (1), 87–96.
- OTHMER, C. 2008 A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows. *International Journal for Numerical Methods in Fluids* **58** (8), 861–877.
- OVSJANIKOV, M., BEN-CHEN, M., SOLOMON, J., BUTSCHER, A. & GUIBAS, L. 2012 Functional maps: A flexible representation of maps between shapes. *ACM Transactions on Graphics* **31** (4), 30:1–30:11.
- OVSJANIKOV, M., CORMAN, E., BRONSTEIN, M., RODOLÀ, E., BEN-CHEN, M., GUIBAS, L., CHAZAL, F. & BRONSTEIN, A. 2017 Computing and processing correspondences with functional maps. In *ACM SIGGRAPH 2017 Courses*, pp. 1–62. New York, NY, USA: Association for Computing Machinery.
- PANCHIGAR, D., KAR, K., SHUKLA, S., MATHEW, R.M., CHADHA, U. & SELVARAJ, S.K. 2022 Machine learning-based CFD simulations: A review, models, open threats, and future tactics. *Neural Computing and Applications* **34** (24), 21677–21700.
- PARISH, E.J. & DURAISAMY, K. 2017 Non-Markovian closure models for large eddy simulations using the Mori-Zwanzig formalism. *Physical Review Fluids* **2** (1), 014604.
- PATEL, R.G., GARCIA, G.J.M., FRANK-ITO, D.O., KIMBELL, J.S. & RHEE, J.S. 2015 Simulating the Nasal Cycle with Computational Fluid Dynamics. *Otolaryngology–Head and Neck Surgery* **152** (2), 353–360.
- PIANTA, L., VINCIGUERRA, A., BERTAZZONI, G., MORELLO, R., MANGIATORDI, F., LUND, V. J. & TRIMARCHI, M. 2020 Acetic acid disinfection as a potential adjunctive therapy for non-severe COVID-19. *European Archives of Oto-Rhino-Laryngology* **277** (10), 2921–2924.

- PIRONNEAU, O. 1974 On optimum design in fluid mechanics. *Journal of Fluid Mechanics* **64** (1), 97–110.
- POPE, S.B. 2000 *Turbulent Flows*. Cambridge University Press, Cambridge.
- QUADRIOS, M., PIPOLI, C., CORTI, S., LENZI, R., MESSINA, F., PESCI, C. & FELISATI, G. 2014 Review of computational fluid dynamics in the assessment of nasal air flow and analysis of its limitations. *European Archives of Oto-Rhino-Laryngology* **271** (9), 2349–2354.
- QUADRIOS, M., PIPOLI, C., CORTI, S., MESSINA, F., PESCI, C., SAIBENE, A.M., ZAMPINI, S. & FELISATI, G. 2016 Effect of CT resolution and radiodensity threshold on the CFD evaluation of nasal airflow. *Medical & Biological Engineering & Computing* **54**, 411–419.
- RADULESCO, T., MEISTER, L., BOUCHET, G., GIORDANO, J., DESSI, P., PERRIER, P. & MICHEL, J. 2019 Functional relevance of computational fluid dynamics in the field of nasal obstruction: A literature review. *Clinical Otolaryngology* **44** (5), 801–809.
- RADULESCO, T., MEISTER, L., BOUCHET, G., VAROQUAUX, A., GIORDANO, J., DESSI, P., PERRIER, P. & MICHEL, J. 2020 Computational fluid dynamics and septal deviations—Virtual surgery can predict post-surgery results: A preliminary study including two patients. *Clinical Otolaryngology* **45** (2), 286–291.
- RAISSI, M. 2018 Deep hidden physics models: Deep learning of nonlinear partial differential equations. *The Journal of Machine Learning Research* **19** (1), 932–955.
- RAISSI, M., PERDIKARIS, P. & KARNIADAKIS, G. E. 2019 Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* **378**, 686–707.
- RAISSI, M., YAZDANI, A. & KARNIADAKIS, G.E. 2020 Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367** (6481), 1026–1030.
- RAMALINGAM, S., GRAHAM, C., DOVE, J., MORRICE, L. & SHEIKH, A. 2019 A pilot, open labelled, randomised controlled trial of hypertonic saline nasal irrigation and gargling for the common cold. *Scientific Reports* **9** (1), 1015.
- RAMANATHAN, M., RAMESH, P., AGGARWAL, N., PARAMESWARAN, A., SAILER, H.F. & GEORGE, A.E. 2021 Evaluation of airflow characteristics before and after septoplasty in unilateral cleft patients with a deviated nasal septum: A computational fluid dynamics study. *International Journal of Oral and Maxillofacial Surgery* **50** (4), 451–456.
- RHEE, J.S., BOOK, D.T., BURZYNSKI, M. & SMITH, T.L. 2003 Quality of Life Assessment in Nasal Airway Obstruction. *The Laryngoscope* **113** (7), 1118–1122.
- ROBLIN, D.G. & ECCLES, R. 2002 What, if any, is the value of septal surgery? *Clinical Otolaryngology and Allied Sciences* **27** (2), 77–80.

- RUDMIK, L., SOLER, Z.M., MACE, J.C., SCHLOSSER, R.J. & SMITH, TIMOTHY L. 2015 Economic evaluation of endoscopic sinus surgery versus continued medical therapy for refractory chronic rhinosinusitis. *The Laryngoscope* **125** (1), 25–32.
- SAIBENE, A.M., FELISATI, G., PIPOLO, C., BULFAMANTE, A.M., QUADARIO, M. & COVELLO, V. 2020 Partial Preservation of the Inferior Turbinate in Endoscopic Medial Maxillectomy: A Computational Fluid Dynamics Study. *American Journal of Rhinology and Allergy* **34** (3), 409–416.
- SCHILLACI, A., BORACCHI, G., PIPOLO, C. & QUADARIO, M. 2022 A CFD-augmented machine-learning approach for the classification of nasal pathologies. In *EUROMECH Fluid Mechanics Conference*. Athens (Greece).
- SCHILLACI, A., HASEGAWA, K., PIPOLO, C., BORACCHI, G. & QUADARIO, M. 2023 Comparing flow-based and anatomy-based features in the data-driven study of nasal pathologies. *Flow* (submitted).
- SCHILLACI, A. & QUADARIO, M. 2022 Importance of the numerical schemes in the CFD of the human nose. *Journal of Biomechanics* **138**, 111100.
- SCHILLACI, A., QUADARIO, M. & BORACCHI, G. 2021a A database of CFD-computed flow fields around airfoils for machine-learning applications. <https://doi.org/10.5281/zenodo.4106752>.
- SCHILLACI, A., QUADARIO, M., PIPOLO, C., RESTELLI, M. & BORACCHI, G. 2021b Inferring Functional Properties from Fluid Dynamics Features. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 4091–4098. Milan, Italy.
- SCHMID, P. J. 2010 Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **656**, 5–28.
- SCHRAMM, M., STOEVESANDT, B. & PEINKE, J. 2018 Optimization of Airfoils Using the Adjoint Approach and the Influence of Adjoint Turbulent Viscosity. *Computation* **6** (1), 5.
- SCHUMACHER, M.J. 2004 Nasal dyspnea: The place of rhinomanometry in its objective assessment. *American Journal of Rhinology* **18** (1), 41–6.
- SEMERARO, O., BAGHERI, S., BRANDT, L. & HENNINSON, D.S. 2011 Feedback control of three-dimensional optimal disturbances using reduced-order models. *Journal of Fluid Mechanics* **677**, 63–102.
- SEMERARO, O., BELLANI, G. & LUNDELL, F. 2012 Analysis of time-resolved PIV measurements of a confined turbulent jet using POD and Koopman modes. *Experiments in Fluids* **53** (5), 1203–1220.
- SHORTEN, C. & KHOSHGOFTAAR, T.M. 2019 A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data* **6** (1), 60.

- SINGH, N.P. & INTHAVONG, K. 2021 Can computational fluid dynamic models help us in the treatment of chronic rhinosinusitis. *Current Opinion in Otolaryngology & Head and Neck Surgery* **29** (1), 21–26.
- SINGH, S., SHARMA, N., SINGH, U., SINGH, T., MANGAL, D.K. & SINGH, V. 2020 Nasopharyngeal wash in preventing and treating upper respiratory tract infections: Could it prevent COVID-19? *Lung India : Official Organ of Indian Chest Society* **37** (3), 246–251.
- SMAGORINSKY, J. 1963 General Circulation Experiments with the Primitive Equations. *Monthly Weather Review* **91** (3), 99–164.
- SMITH, K.A., ORLANDI, R.R. & RUDMIK, L. 2015 Cost of Adult Chronic Rhinosinusitis: A Systematic Review. *The Laryngoscope* **125** (7), 1547–1556.
- SOLERA-RICO, A., VILA, C.S., GÓMEZ, M.A., WANG, Y., ALMASHJARY, A., DAWSON, S.T.M. & VINUESA, R. 2023 \$\beta\$-Variational autoencoders and transformers for reduced-order modelling of fluid flows, arXiv: 2304.03571.
- SORENSEN, R.L. 1980 A computer program to generate two-dimensional grids about airfoils and other shapes by the use of Poisson's equation. *Tech. Rep. A-8178*.
- SOTO, O. & LOHNER, R. 2004 On the Boundary Computation of Flow Sensitivities. In *42nd AIAA Aerospace Sciences Meeting and Exhibit*. Reno, Nevada.
- SOZANSKY, J. & Houser, S.M. 2014 The physiological mechanism for sensing nasal airflow: A literature review. *International Forum of Allergy & Rhinology* **4** (10), 834–838.
- SPALART, P. & ALLMARAS, S. 1992 A one-equation turbulence model for aerodynamic flows. In *30th Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics.
- SRINIVASAN, P.A., GUASTONI, L., AZIZPOUR, H., SCHLATTER, P. & VINUESA, R. 2019 Predictions of turbulent shear flows using deep neural networks. *Physical Review Fluids* **4** (5), 054603.
- SUNDH, C. & SUNNERGREN, O. 2015 Long-term symptom relief after septoplasty. *European Archives of Oto-Rhino-Laryngology* **272** (10), 2871–2875.
- TAIRA, K., BRUNTON, S.L., DAWSON, S.T.M., ROWLEY, C.W., COLONIUS, T., McKEON, B.J., SCHMIDT, O.T., GORDEYEV, S., THEOFILIS, V. & UKEILEY, L.S. 2017 Modal Analysis of Fluid Flows: An Overview. *AIAA Journal* **55** (12), 4013–4041.
- TIBSHIRANI, R. 1996 Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58** (1), 267–288.
- TJAHJONO, R., SALATI, H., INTHAVONG, K. & SINGH, N. 2023 Correlation of Nasal Mucosal Temperature and Nasal Patency—A Computational Fluid Dynamics Study. *The Laryngoscope* **133** (6), 1328–1335.

- TONG, J.Y., WONG, A., ZHU, D., FASTENBERG, J.H. & THAM, T. 2020 The Prevalence of Olfactory and Gustatory Dysfunction in COVID-19 Patients: A Systematic Review and Meta-analysis. *Otolaryngology–Head and Neck Surgery* **163** (1), 3–11.
- TRETIAKOW, D., TESCH, K., MEYER-SZARY, J., MARKIET, K. & SKOREK, A. 2020 Three-dimensional modeling and automatic analysis of the human nasal cavity and paranasal sinuses using the computational fluid dynamics method. *European Archives of Oto-Rhino-Laryngology* **278**, 1443–1453.
- TSANG, C.L.N., NGUYEN, T., SIVESIND, T. & CERVIN, A. 2018 Long-term patient-related outcome measures of septoplasty: A systematic review. *European Archives of Oto-Rhino-Laryngology* **275** (5), 1039–1048.
- UDAKA, T., SUZUKI, H., KITAMURA, T., SHIOMORI, T., HIRAKI, N., FUJIMURA, T. & UEDA, N. 2006 Relationships Among Nasal Obstruction, Daytime Sleepiness, and Quality of Life:. *The Laryngoscope* **116** (12), 2129–2132.
- VAN DOREMALEN, N., BUSHMAKER, T., MORRIS, D.H., HOLBROOK, M.G., GAMBLE, A., WILLIAMSON, B.N., TAMIN, A., HARcourt, J.L., THORNBURG, N.J., GERBER, S.I., LLOYD-SMITH, J.O., DE WIT, E. & MUNSTER, V.J. 2020 Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *New England Journal of Medicine* **382** (16), 1564–1567.
- VAN GUMSTER, J. 2015 *Blender for Dummies*. John Wiley & Sons.
- VAN STRIEN, J., SHRESTHA, K., GABRIEL, S., LAPPAS, P., FLETCHER, D.F., SINGH, N. & INTHAVONG, K. 2021 Pressure distribution and flow dynamics in a nasal airway using a scale resolving simulation. *Physics of Fluids* **33** (1), 011907.
- VIGNON, C., RABAULT, J. & VINUESA, R. 2023 Recent advances in applying deep reinforcement learning for flow control: Perspectives and future directions. *Physics of Fluids* **35** (3), 031301.
- VINUESA, R. & BRUNTON, S.L. 2022 Enhancing computational fluid dynamics with machine learning. *Nature Computational Science* **2** (6), 358–366.
- VIQUERAT, J., RABAULT, J., KUHNLE, A., GHRAIEB, H., LARCHER, A. & HACHEM, E. 2021 Direct shape optimization through deep reinforcement learning. *Journal of Computational Physics* **428**, 110080.
- VOLLANT, A., BALARAC, G. & CORRE, C. 2017 Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures. *Journal of Turbulence* **18** (9), 854–878.
- WANG, D.Y., LEE, H.P. & GORDON, R. 2012 Impacts of Fluid Dynamics Simulation in Study of Nasal Airflow Physiology and Pathophysiology in Realistic Human Three-Dimensional Nose Models. *Clinical and Experimental Otorhinolaryngology* **5** (4), 181–187.

- WANG, J.X., WU, J.L. & XIAO, H. 2017 Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Physical Review Fluids* **2** (3), 034603.
- WANG, M. & HEMATI, M.S. 2019 Detecting exotic wakes with hydrodynamic sensors. *Theoretical and Computational Fluid Dynamics* **33** (3), 235–254.
- WELLER, H.G., TABOR, G., JASAK, H. & FUREBY, C. 1998 A Tensorial Approach to Computational Continuum Mechanics using Object-Oriented Techniques,. *Computers in Physics* **12** (6), 620–631.
- WEN, J., INTHAVONG, K., TU, JIYUAN & WANG, S. 2008 Numerical simulations for detailed airflow dynamics in a human nasal cavity. *Respiratory Physiology & Neurobiology* **161** (2), 125–135.
- WEXLER, D., SEGAL, R. & KIMBELL, J. 2005 Aerodynamic effects of inferior turbinate reduction. *Arch Otolaryngol Head Neck Surg.* **131**, 1102–1107.
- WOJEWODKA, M.M., WHITE, C., SHAHPAR, S. & KONTIS, K. 2018 A review of flow control techniques and optimisation in s-shaped ducts. *International Journal of Heat and Fluid Flow* **74**, 223–235.
- WONG, E., SIU, J., DOUGLAS, R. & SINGH, N. 2021 Anatomy and Physiology of the Human Nose. In *Clinical and Biomedical Engineering in the Human Nose: A Computational Fluid Dynamics Approach* (ed. K. Inthavong, N. Singh, E. Wong & J. Tu), pp. 9–29. Singapore: Springer.
- WU, H., LIU, X., AN, W. & LYU, H. 2022a A generative deep learning framework for airfoil flow field prediction with sparse data. *Chinese Journal of Aeronautics* **35** (1), 470–484.
- WU, P., PAN, K., JI, L., GONG, S., FENG, W., YUAN, W. & PAIN, C. 2022b Navier–stokes Generative Adversarial Network: A physics-informed deep learning model for fluid flow generation. *Neural Computing and Applications* **34** (14), 11539–11552.
- XIE, Y., FRANZ, E., CHU, M. & THUEREY, N. 2018 tempoGAN: A temporally coherent, volumetric GAN for super-resolution fluid flow. *ACM Transactions on Graphics* **37** (4), 95:1–95:15.
- YANG, S., XIAO, W., ZHANG, M., GUO, S., ZHAO, J. & SHEN, F. 2023 Image Data Augmentation for Deep Learning: A Survey, arXiv: 2204.08610.
- ZACHOV, S., MUIGG, P., HILDEBRANDT, T., DOLEISCH, H. & HEGE, H.C. 2009 Visual exploration of nasal airflow. *IEEE Transaction on Visualization and Computer Graphics* **15** (6), 1407–1414.
- ZEMOURI, C., SOET, H., CRIELAARD, W. & LAHEIJ, A. 2017 A scoping review on bio-aerosols in healthcare and the dental environment. *PLOS ONE* **12** (5), e0178007.

- ZHANG, Z. & KLEINSTREUER, C. 2011 Laminar-to-turbulent fluid–nanoparticle dynamics simulations: Model comparisons and nanoparticle-deposition applications. *International Journal for Numerical Methods in Biomedical Engineering* **27** (12), 1930–1950.
- ZHAO, K. & JIANG, J. 2014 What is normal nasal airflow? A computational study of 22 healthy adults. *International Forum of Allergy & Rhinology* **4** (6), 435–446.
- ZHAO, K., SCHERER, P.W., HAJILOO, S.A. & DALTON, P. 2004 Effect of anatomy on human nasal air flow and odorant transport patterns: Implications for olfaction. *Chemical Senses* **29**, 365–379.
- ZUBAIR, M., ABDULLAH, M.Z., ISMAIL, R., SHUAIB, I. L., HAMID, S.A. & AHMAD, K.A. 2012 Review: A Critical Overview of Limitations of CFD Modeling in Nasal Airflow. *Journal of Medical and Biological Engineering* **32** (2), 77–84.
- ZWICKER, D., YANG, K., MELCHIONNA, S., BRENNER, M. P., LIU, B. & LINDSAY, R. W. 2018 Validated reconstructions of geometries of nasal cavities from CT scans. *Biomedical Physics & Engineering Express* **4** (4), 045022.