INFINITY: Neural Field Modeling for Reynolds-Averaged Navier-Stokes Equations
Louis Serrano, Leon Migus, Yuan Yin, Jocelyn Ahmed Mazari, Patrick Gallinari

To cite this version:
Louis Serrano, Leon Migus, Yuan Yin, Jocelyn Ahmed Mazari, Patrick Gallinari. INFINITY: Neural Field Modeling for Reynolds-Averaged Navier-Stokes Equations. Workshop on Synergy of Scientific and Machine Learning Modeling (ICML 2023), Jul 2023, Honolulu, HI, United States. hal-04171439

HAL Id: hal-04171439
https://hal.science/hal-04171439
Submitted on 26 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
INFINITY: Neural Field Modeling for Reynolds-Averaged Navier-Stokes Equations

Louis Serrano 1  Leon Migus 1,2  Yuan Yin 1  Jocelyn Ahmed Mazari 3  Patrick Gallinari 1,4

Abstract

For numerical design, the development of efficient and accurate surrogate models is paramount. They allow us to approximate complex physical phenomena, thereby reducing the computational burden of direct numerical simulations. We propose INFINITY, a deep learning model that utilizes implicit neural representations (INRs) to address this challenge. Our framework encodes geometric information and physical fields into compact representations and learns a mapping between them to infer the physical fields. We use an airfoil design optimization problem as an example task and we evaluate our approach on the challenging AirfRANS dataset, which closely resembles real-world industrial use-cases. The experimental results demonstrate that our framework achieves state-of-the-art performance by accurately inferring physical fields throughout the volume and surface. Additionally we demonstrate its applicability in contexts such as design exploration and shape optimization: our model can correctly predict drag and lift coefficients while adhering to the equations.

1. Introduction and motivation

Numerical simulations are essential for analyzing systems governed by partial differential equations (PDEs) in fields like fluid dynamics and climate science. These simulations involve discretizing the domain and solving the equations using methods such as finite differences, finite elements, or finite volumes (Reddy, 2019; Grossmann et al., 2007; Eymard et al., 2000). Since direct numerical simulation (DNS) can be computationally expensive or intractable, it is crucial to develop computationally efficient yet accurate surrogate models to accelerate the design process. Surrogate modeling for industrial applications, however, poses several challenges. The meshes used in these applications are extensive, consisting of hundreds of thousands of cells, and they also exhibit unstructured data and involve multi-scale phenomena. A typical example is the design of airfoils which will be our application focus, although the ideas can be easily implemented for other design tasks. In this domain, a new costly simulation must be run for each mesh during the optimization process, leading to time-consuming processes. Additionally, the design process focuses on finding the optimal shape for an airfoil that minimizes the force required for flight. Experts typically maximize the lift-over-drag ratio by solving equations across the entire mesh, with particular emphasis on the surface where various multi-scale phenomena occur.

Recently, deep learning methods have emerged as promising approaches for constructing surrogate models. However, the progress in this field was initially hindered by the lack of evaluation datasets representative of real-world data. The machine learning community has begun to address this issue by developing benchmarks. In this work, we utilize the recent AirfRANS dataset (Bonnet et al., 2022), which aims to replicate real-world industrial scenarios. This comprehensive benchmark provides an evaluation framework to assess the capabilities of deep learning (DL) in modeling the two-dimensional incompressible steady-state Reynolds-Averaged Navier-Stokes (RANS) equations for airfoils. Additionally, this 2D dataset encompasses a wide range of airfoil shapes derived from NASA's early works (Cummings et al., 2015), various turbulence effects characterized by Reynolds numbers and different angles of attack.

The Navier-Stokes equations are widely used in fluid dynamics, and as a result, numerous neural network surrogates have been proposed for their modeling in different contexts. Initial attempts all relied on grid-based approaches such as convolutional Neural Networks (CNNs) (Um et al., 2020; Thuerey et al., 2020; Mohan et al., 2020; Wandel et al., 2020;
We aim at proposing a surrogate model for airfoil design within each domain. The characterization of an airfoil involves defining boundary conditions on each node within the airfoil’s domain. Normal vectors of the mesh nodes on the airfoil surface \( n(x) = (n_x(x), n_y(x)) \) specify the direction perpendicular to the airfoil surface at each node. In addition to the geometric inputs, we also have access to the inlet velocity values \( V_x \) and \( V_y \), denoting the horizontal and vertical components of the velocity, respectively. It is worth noting that, on average, a mesh consists of approximately 200,000 nodes, providing a detailed representation of the airfoil’s geometry.

The primary objective of the design optimization process is to maximize the lift-over-drag coefficient ratio, which serves as the key performance metric. To achieve this, we place significant emphasis on evaluating the relative errors in both the drag and lift coefficients, as well as assessing the Spearman correlation between predicted and actual values.

Rather than directly predicting the drag and lift values, our approach focuses on inferring various fluid fields associated with the airfoil’s geometry. This includes calculating the velocities \((v_x, v_y)\), pressure \(p\), and turbulent kinematic viscosity \(\nu_t\) on the mesh nodes, following the experimental protocol proposed in (Bonnet et al., 2022). Therefore the inputs of our surrogate model are \(\{V_x, V_y, d(x), n_x, n_y\}_{i=1}^{n_{tr}}\), and the outputs are \((v_x|x, v_y|x, p|x, \nu_t|x)\) \(i=1\). The output physical fields provide valuable insights into the underlying behavior of the fluid and its interaction with the airfoil’s geometry. The drag and lift coefficients are calculated based on the predictions of the trained model while respecting the form of the RANS equations. This approach enables us to obtain a comprehensive understanding of the underlying fluid behavior and its relationship with the airfoil’s geometry, thereby ultimately enhancing the accuracy of drag and lift estimation.

2.2. Model

We present INFINITY: Implicit Neural Fields for INterpreting geometrY and inferring phYSics.

Modulated INR In our model, we will treat each geometric input \((d)\) or physical output function \((v, p, \nu)\) separately and each will be modeled by an INR. Let us then consider a generic function \(u\), which will represent either an input geometric field or an output physical field defined over a domain \(\Omega\) or at its boundary \(\partial\Omega\). Let us denote \(u_i\) the function corresponding to a specific airfoil example. \(u_i\) will be represented by an INR \(f_{\theta_{ui}, \phi_{ui}}\) with two sets of parameters: parameters \(\theta_{ui}\) shared by all the \(u_i\), and modulation parameters \(\phi_{ui}\) specific to each individual function \(u_i\). In our airfoil example, \(\phi_{ui}\) enables the INR to handle different geometries. Overall, this decomposition allows the modulated INR to capture both shared characteristics among the example’s functions \(u_i\) and the unique

\(d(x)\) provides the distance from each node to the surface of the airfoil.

Leveraging recent advances in implicit neural representations (INRs) (Sitzmann et al., 2020; Mildenhall et al., 2021), which have shown successful applications in physics problems (Yin et al., 2022), we introduce INFINITY, a model that utilizes coordinate-based networks to encode geometric information and physical fields into concise representations. INFINITY establishes a mapping between variables representing the problem’s geometry and the corresponding physical fields, within this representation space. It possesses several unique features: (i) it is robust to varying mesh sampling, allowing for adaptability to different geometries, (ii) it effectively captures multi-scale phenomena, resulting in state-of-the-art scores for both volume and surface evaluations, (iii) as a continuous surrogate model, it can be used to accelerate the evaluation of different meshes during the design process, leading to significant speed-up. Importantly, we verified that INFINITY’s field predictions accurately produce the correct lift and drag forces clearly outperforming all the baselines.
properties of each one. INFINITY leverages latent representations inferred from the modulation spaces of the INRs. These latent representations, denoted as $z_{ui}$, are compact codes that encode information from the INRs’ parameters. They serve as inputs to a hypernetwork $h_u$, which computes the modulation parameters $\phi_{ui} = h_u(z_{ui})$.

In this work we use Fourier Features (Tancik et al., 2020) as an INR backbone and apply shift modulation (Perez et al., 2018): $f_{\theta, \phi_{ui}}(x) = W_L(\chi_{L-1} \circ \chi_{L-2} \circ \cdots \circ \chi_0(x)) + b_L$, with $\chi_j(\eta_j) = \sigma(W_j \eta_j + b_j + (\phi_{ui}))$. We note $\eta_0 = x$ and $(\eta_j)_{j\geq 1}$ the hidden activations throughout the network. Hence, the parameters $\theta = (W_j, b_j)_{j=0}^L$ are shared between all examples and the modulation $\phi_{ui} = ((\phi_{ui})_j)_{j=0}^{L-1}$ is specific to a single example. We compute the modulation parameters $\phi_{ui} = ((\phi_{ui})_j)_{j=0}^{L-1}$ from $z$ with a linear hypernetwork.

With the learned shared parameters $(\theta_u, w_u)$, the modulated INR enables two processes: decoding and encoding (see Figure 1). Decoding refers to mapping a given code $z_{ui}$ to the corresponding INR function $f_{\theta_u, \phi_{ui}}$, where $\phi_{ui} = h_u(z_{ui})$, while encoding involves generating a code $z_{ui}$ given a function $u_i$, providing a compact representation of the function within the modulation space of the INR.

To obtain the compact code $z_{ui}$ for reconstructing the original field $u_i$ using the INR, an inverse problem is solved through a procedure called auto-decoding. The objective is to compress the necessary information into $z_{ui}$ such that the reconstructed value $\hat{u}_i(x) = f_{\theta_u, \phi_{ui}}(x)$ approximates the original value $u_i(x)$ for all $x \in X_i$. The approximate solution to this inverse problem is computed iteratively through a gradient descent optimization process:

$$
\begin{align*}
    z_{ui}^{(0)} &= 0, \\
    z_{ui}^{(k+1)} &= z_{ui}^{(k)} - \alpha \nabla_{z_{ui}} L_{\mu_{ik}}(f_{\theta_u, \phi_{ui}}(z_{ui}), u_i), \\
    \text{with } \phi_{ui}^{(k)} &= h_u(z_{ui}^{(k)}) \text{ for } 0 \leq k \leq K - 1.
\end{align*}
$$

where $\alpha$ is the inner loop learning rate, $K$ the number of inner steps, and $L_{\mu_{ik}}(u_i, \tilde{u}_i) = \mathbb{E}_{x \sim \mu_i}[(u_i(x) - \tilde{u}_i(x))^2]$ where $\mu_i$ is a measure defined through the observation grid $X_i \mu_i(\cdot) = \sum_{x \in X_i} \delta_x(\cdot)$, with $\delta_x(\cdot)$ the Dirac measure.

As indicated before, we treat each input and output function independently: there are two input functions denoted as $(d, n)$ and four output functions denoted as $(v_x, v_y, p, v_t)$. Each $u_i \in \{d, n, v_x, v_y, p, v_t\}$ is represented by a modulated INR $f_{\theta_u, \phi_{ui}}$, where $u_i$ stands for a field specific to an airfoil example. INFINITY then learns a mapping between the latent representations of the geometric input fields and the latent representations of the physics output fields.

**Inference** As illustrated in Figure 1, INFINITY follows a three-step procedure: encode, process, and decode.

- **Encode:** Given the geometric input functions $d_i, n_i$ and the corresponding INR learned parameters, respectively $\theta_d, w_d, \theta_n, w_n$, functions $d_i, n_i$ are encoded into the latent codes $z_{d_i}, z_{n_i}$ according to Equation (1). Since we can query the INRs anywhere within the domain, we can hence freely encode functions without mesh constraints. This lets us freely encode inputs with different geometries.

- **Process:** Once we obtain $z_{d_i}$ and $z_{n_i}$, we can infer the latent output codes $(\tilde{z}_{v_x}, \tilde{z}_{v_y}, \tilde{z}_p, \tilde{z}_{v_t}) = g_{\psi}(z_{d_i}, z_{n_i}, V_{x_i}, V_{y_i})$. We consider here that $g_{\psi}$

Figure 1. The inference of INFINITY proceeds in three steps. 1. We encode the distance function $d$ and the normal components $n_x, n_y$ into the latent representations $z_d$ and $z_n$. 2. We process these codes along with the inlet velocities $V_x, V_y$ to obtain the predicted output codes $\tilde{z}_{v_x}, \tilde{z}_{v_y}, \tilde{z}_p, \tilde{z}_{v_t}$ corresponding respectively to velocity, pressure and viscosity. 3. The processed codes are decoded with the modulated INRs, which can be queried directly at any mesh position $x \in X$.
We outline the training pipeline of a modulated INR in Algorithm 1. Once the different INRs have been fitted, we encode the functions into the input codes \( z \) and \( u \). During the training of the INRs we force the auto-decoding process to take only a few gradient steps to encode the geometric functions or physical fields. This enhances the INR capability to encode new geometrical inputs in a few steps at test time, and also reduces the space size of the target output codes. This regularization prevents the different INRs to memorize the training sets into the individual codes. In order to obtain a network that is capable of quickly encoding new geometrical and physical inputs, we employ a second-order meta-learning training algorithm based on CAVIA (Zintgraf et al., 2019). Compared to a first-order scheme such as Reptile (Nichol et al., 2018), the outer loop back-propagates the gradient through the \( K \) inner steps, consuming more memory. Indeed, we need to compute gradients of gradients but this yields higher reconstruction results with the modulated INR. We experimentally found that using 3 inner-steps for training, or testing, was sufficient to obtain very low reconstruction errors for the geometric or physical fields. Using more inner-steps would result in a higher computation cost with only a marginal gain in reconstruction capacity. We outline the training pipeline of a modulated INR in Algorithm 1. Once the different INRs have been fitted, we encode the functions into the input codes \( z_{d,i} \), \( z_{n,i} \) and target codes \( z_{t,x}, z_{t,y}, z_{t,z} \). The training of \( g_\psi \) is performed in the small dimensional \( z \)-code space, and is supervised through the MSE loss with the target codes.

### 3. Experiments

#### Baselines
We use the same baselines as Bonnet et al. (2022); GraphSAGE (Hamilton et al., 2017), a PointNet (Qi et al., 2017), a Graph U-Net (Gao and Ji, 2019) and a MLP. Those baselines have been initially chosen as they process in different ways the inputs. The results are given for the setup “full data regime” of AirfRANS, using 800 samples for training and 200 for testing.

#### Results
In Table 1, the INFINITY model demonstrates superior inference capabilities on the volume and surface compared to the baselines. Indeed, it achieves significantly lower error values on the volume velocity and pressure fields, while exhibiting an order-of-magnitude lower MSE on the surface pressure. This substantial gain in prediction power translates to order of magnitude lower relative errors on the drag and lift forces, accompanied by high positive Spearman correlations. These results indicate a strong alignment between INFINITY’s predictions and the true drag and lift forces. Consequently, INFINITY emerges as the only model capable of predicting accurately physical fields on the volume and surface while maintaining coherent and accurate drag and lift estimations. On the downside, the INFINITY model has a longer inference time compared to GraphSAGE and PointNet. However, this increased inference time is still within an acceptable range, considering its superior performance and that a numerical solver needs approximately 20 minutes to complete a simulation. Furthermore, it is counterbalanced by the ability to query the full mesh directly, in stark contrast to graph-based methods that necessitate sub-sampling to process the inputs.

### 4. Conclusion
We introduce INFINITY, a model that utilizes coordinate-based networks to encode geometric information and physical fields into compact representations. INFINITY establishes a mapping between geometry and physical fields within a reduced representation space. We validated our model on AirfRANS, a challenging dataset for the Reynolds-Averaged Navier-Stokes equation, where it significantly outperforms previous baselines across all relevant performance metrics. At post-processing stage, the predicted fields yield accurate lift and drag forces. This validates INFINITY's potential as a surrogate design model, where it could be plugged in any design optimization or exploration loop.
Table 1. Test results on AirfRANS. Mean squared error (MSE) on normalized fields expressed with factor ($\times 10^{-2}$) for the volume and ($\times 10^{-1}$) for the surface. Relative errors $C_D$, $C_L$ on the drag and lift and Spearman correlations $\rho_D$, $\rho_L$ on the drag and lift. The results from the baselines are taken from Bonnet et al. (2022).

Broader impact

This work could serve multiple purposes, including:

- Enhancing surrogate solvers for Computational Fluid Dynamics (CFD) engineers, facilitating efficient design iterations.

- Cost and risk reduction in prototyping new plane designs, mitigating real-world study expenses.

Acknowledgements

We acknowledge the financial support provided by DL4CLIM (ANR-19-CHIA-0018-01) and DEEPNUM (ANR-21-CE23-0017-02) ANR projects, as well as the program France 2030, project 22-PETA-0002.

References


