arXiv:2303.03806v2 [physics.flu-dyn] 8 Jan 2025

使用图神经网络进行数据同化闭合的主动学习

Michele Quattromini
Dipartimento di Meccanica, Matematica e Management
巴里理工大学
Bari
LISN-CNRS
巴黎萨克雷大学
奥尔赛
michele.quattromini@poliba.it

Michele Alessandro Bucci
赛峰科技，数字科学与技术 马尼莱阿莫
michele-alessandro.bucci@safrangroup.com

Stefania Cherubini Dipartimento di Meccanica, Matematica e Management 巴里理工大学 巴里 stefania.cherubini@poliba.it

Onofrio SemeraroLISN-CNRS 巴黎萨克雷大学 奥尔赛
onofrio.semeraro@universite-paris-saclay.fr

摘要

机器学习技术的普及，加上高质量实验数据和数值数据的日益丰富，显著推动了流体力学中众多应用的发展。其中值得关注的是，利用神经网络（NN）为非定常流动和湍流开发数据同化方法与闭合模型。尽管这些方法应用广泛，但往往会出现过拟合，且通常需要大量数据集，尤其是在未纳入物理约束的情况下。数值模拟中的这一问题尤为突出：鉴于计算成本高昂，建立在数据集有限时仍然有效的学习流程至关重要。为解决这些局限，我们开发了能够在低数据量极限下对未见数据进行泛化的神经网络模型，具体方法包括：i）利用图神经网络（GNN）架构将不变性融入神经网络模型；ii）设计一种自适应策略，选择学习过程中使用的数据。图神经网络尤其适用于涉及非结构化区域离散的数值模拟。我们将其与有限元（FEM）求解器相结合，用于雷诺平均纳维–斯托克斯方程的监督学习，并展示了这一方法。我们以不同雷诺数 $50 \leq Re \leq 150$ 下流经一般钝体的平均流场数据同化作为测试案例，该流动具有非定常动力学特征。结果表明，图神经网络模型能够成功预测闭合项；值得注意的是，这些性能是通过使用极为有限的数据集实现的，而该数据集由主动学习过程选取，从而确保了RANS闭合项的泛化性质。结果表明，通过主动学习流程训练的图神经网络模型，可作为卷积神经网络等灵活性较低技术的有效替代方案。

关键词 图神经网络（GNN）、代理模型、流体动力学机器学习、雷诺平均纳维–斯托克斯（RANS）方程、主动学习

1 引言

机器学习（ML）应用正在科学界各个领域广泛发展，并推动了新技术和算法的涌现；这些技术和算法汇聚于科学机器学习这一新兴领域。这一突破得益于计算机辅助活动的大幅增加、高性能计算能力的持续提升、深度学习领域的近期发展，以及数据的广泛可用性 [Goodfellow 等， 2016]。简而言之，对于回归问题，机器学习技术能够利用数据识别系统可观测量（输入）与我们希望预测的关注量（输出）之间的映射；当所分析的数据受确定性定律或统计定律支配时，原则上，这些映射对应于近似模型。其他应用还包括分类问题、聚类，以及流体力学等众多科学领域中的控制。在过去几年中，结合机器学习与标准分析技术的研究成果数量持续增长 [Raissi 等，2020年，Brunton 等，2020b,a，Garnier 等，2021年，Vinuesa 和 Brunton，2022年，Yang 等， 2024]。

这些应用自然地延伸到各类建模问题中，范围涵盖湍流的雷诺平均纳维-斯托克斯方程闭合，以及壁面模型大涡模拟 [Lapeyre 等人，2019年，Zhou 等人，2021年，Bae 和 Koumoutsakos， 2022] 或结合基于伴随方程的技术与机器学习的优化问题 [Volpiani 等人，2021年，Patel 等人， 2024]。在众多研究这一问题的作者中，Ling 和 Templeton [2015] 采用分类方法来识别不确定区域，在这些区域中，雷诺平均纳维–斯托克斯方程（RANS）的闭合项可能失效；Ströfer 和 Xiao [2021] 将数据同化与基于有限观测数据的神经网络（NN）雷诺应力建模相结合。其他方法则借助基准模型，例如 Spalart-Allmaras 闭合模型 [Singh 和 Duraisamy， 2016], 物理信息神经网络（PINNs） [Eivazi 等人， 2022], 回归方法 [Schmelzer 等人， 2020], 决策树 [Duraisamy 等人， 2019], 集成方法[McConkey 等人， 2022] 或遗传编程 [Weatheritt 和 Sandberg，2016年，Zhao 等人， 2020]。有关更全面的概述，请参阅 Duraisamy 等人 [2019] 以及 Beck 和 Kurz [2021], 的研究，其中讨论了不同的近似层级，并批判性地审视了这些方法的局限性。对于涡黏性闭合模型，Ling 等人近期的研究 [2016] 证明，张量基神经网络（TBNNs）在学习广义涡黏性模型类型方面十分有效 [Pope， 1975]。TBNN 利用 Pope 提出的张量分解方法来考虑不变性，并简化待学习参数的数量。这种建模方法引入的归纳偏置限制了 TBNN 的应用范围，使其适用于雷诺数较高、局部效应占主导地位的近均匀流 [Cai 等人， 2024]。另一方面，对于低雷诺数情形，我们可以考虑另一种方法：数据同化，其中闭合模型对应于基于伴随的方法循环中的控制参数 [Foures 等人， 2014]；该方法无需显式引入张量结构或 Boussinesq近似，因此非常适用于低雷诺数下的非均匀流。本文借鉴这些应用，主要关注低雷诺数下钝体后方发展的非定常流 $50 \leq Re \leq 150$；我们以雷诺平均纳维–斯托克斯方程（RANS）作为基准，但也可以采用其他方案，例如欧拉方程，或预解形式的线性化纳维-斯托克斯方程，此时需要调整的参数为耗散项 Morra 等人 [2019], Pickering 等人 [2021], von Saldern 等人 [2023]。与标准方法不同，本文关注监督学习策略，即通过现有数据推断并确定闭合模型。原则上，只要拥有无限量的数据，我们就可以直接从数据中识别通用闭合模型。然而在实践中，面对实际问题时，可用数据可能有限，或只有少量局部测量；这些限制会影响神经网络（NN）等方法的应用。神经网络具有较强的表达能力和通用结构，适用于各种模型，但也容易出现泛化问题。从这个意义上说，机器学习模型可能仅能代表训练过程中所用的数据集；因此，在数据有限的情况下，亟须通过在训练期间输入精心选择的数据，或通过建模提供先验知识，来规避这些问题 [Bucci 等人，2023年，Shukla 等人， 2022]。

基于这一点，我们将检验这样做是否可行：仅依赖依据合理标准选取的少量数据，便从定义在非结构网格上的数据中识别通用闭合模型。为了回答这些问题，首先需要引入图神经网络（GNN） [Hamilton， 2020]; 这种架构以节点构成的复杂多重连接网络为特征，能够自然地适用于非结构网格：GNN 中的卷积通过聚合邻近节点的信息来实现，从而克服了几何形状施加的限制，这一点不同于卷积神经网络（CNN）。此外，GNN：i）与标准网络模型相比表现出显著的泛化能力 [Sanchez-Gonzalez 等人， 2018]；ii）具有可微性；iii）能够通过离散模板直接针对算子的学习 [见 Shukla 等人， 2022]。由于具备这些特性，这种架构近年来在流体力学领域受到了关注。Lino 等人 [2023], 撰写了一篇相关综述，而 Toshev 等人 [2023] 和 Dupuy 等人 [2023],的工作则给出了具体示例：他们基于 GNN 对大涡模拟中的壁面剪切应力进行了建模。在本文中，我们受到 Donon 等人[2020], 的启发，将一种融合了置换不变性和平移不变性的基于 GNN 的架构与

统计求解器问题相结合；Donon 等人 [2020] 证明，该架构（称为 deepstatisticalsolver）具有某些通用逼近性质，并且能够进行算子学习。

在本研究中，图神经网络与采用有限元法（FEM）进行的数值模拟相结合。GNN-FEM 接口在神经网络与有限元环境之间提供双向耦合，因此能够在有限元分析的后处理中使用神经网络预测。然而，正如前文所述，如果可用数据量不足，数据规模可能成为瓶颈。此外，数据质量也会影响纯数据驱动建模中的预测性能 [Bucci 等人， 2023], 尤其是在现有数据不足以正确表示整个数据集的分布时。不一致或不平衡的数据分布可能导致模型基于无法充分代表问题底层物理机制的数据子集进行训练。因此，作为第二个基本要素，我们采用主动学习，旨在增大数据点之间的差异性，并确保模型提取出有效完成所需任务所必需的全部判别特征。据我们所知，这是流体力学领域最早应用主动学习准则进行数据选择的研究之一。关于深度学习中主动学习方法的全面概述，读者可参阅 Ren 等人。 [2021] 我们遵循 Charpiat 等人 [2019] 的工作，将与 GNN 模型权重更新相关的梯度标量积视为样本之间的相似性度量；所采用的准则能够通过引入促进数据集多样性的数据，动态扩充训练数据集。

本文其余部分安排如下。第2节和第3节阐述流体力学问题，其中讨论数值模拟所用的方程组、流动案例和计算设置（第3.3节）。第4节在理论背景的基础上，进一步讨论基于图神经网络的数据驱动建模，以及如何将这一架构与非结构网格上的数值模拟相结合。本文第二部分聚焦于结果（第5节）：图神经网络模型以由平均流（输入）和雷诺应力（目标）组成的数据集进行训练；基准则是雷诺数为$50 \leq Re \leq 150$ 时圆柱绕流中的闭合问题（第5.2节）。我们使用第5.1节定义的4种不同基准测试评估泛化性质，这些基准测试未纳入训练集。随后，将这一基准与采用不同数据选择策略训练的模型进行比较。首先，我们考虑使用随机几何形状钝体绕流数值模拟进行数据增强的案例（第5.3节）；其次，在第5.3.3节引入主动学习。最后，第6节总结全文。

2 控制方程

本文将重点研究在非定常流动状态下绕钝体发展的不可压缩二维（2D）流体流动。我们考虑不可压缩流的纳维–斯托克斯（NS）方程。将空间笛卡尔坐标记为 $\mathbf{x} = (x, y)$，速度场 $\mathbf{u}(\mathbf{x}\ t)$ 和压力场 $p(\mathbf{x}\ t)$ 的动力学由以下方程控制

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} = -\nabla p + \frac{1}{Re}\nabla^2 \mathbf{u} \tag{1a}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{1b}$$

方程 1 使用特征长度尺度 D（例如 圆柱直径）、来流均匀流的速度 U_∞ 作为参考速度，以及 ρU_∞^2 作为参考压力进行无量纲化。基于这些量，雷诺数定义为 $Re = {}^{U_\infty D}/_{\nu}$，其中 ν 为运动黏度。

作为数据同化过程的基准方程，我们将考虑雷诺平均纳维–斯托克斯方程（RANS），即 NS 方程的时间平均形式，见 Foures 等人 [2014]。请注意，尽管 RANS 方法通常用于湍流，但它也非常适用于非定常流，例如通过弱非线性分析确定非定常周期流动动力学，见 SIPP 和 LEBEDEV [2007], Manti ˇc Lugo 等人[2014]。因此，我们引入雷诺分解

$$\mathbf{u}(\mathbf{x}, t) = \overline{\mathbf{u}}(\mathbf{x}) + \mathbf{u}'(\mathbf{x}, t), \tag{2}$$

其中，向量场 $\mathbf{u} = (u, v)^T$ 被改写为时间平均速度矢量场 $\mathbf{u} = (u, v)^T$ 与脉动速度矢量场 $\mathbf{u}' = (u', v')^T$ 之和。形式上，任何非定常流都可以通过这种分解来描述。将式 2 中的雷诺分解代入纳维–斯托克斯方程（式 1）并进行时间平均，可得方程组

$$\overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p} - \frac{1}{Re}\nabla^2 \overline{\mathbf{u}} = \mathbf{f} \tag{3a}$$

$$\nabla \cdot \overline{\mathbf{u}} = 0, \tag{3b}$$

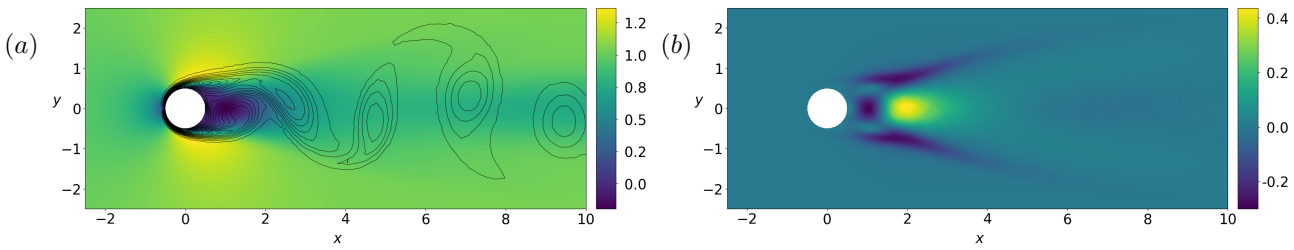

图 1：(*a*) 圆柱绕流在 $Re = 150$ 时平均流场的流向分量$\mathbf{u}$ 和涡量等值线 $\omega = \nabla \times \mathbf{u}$ 。(*b*) 对于相同工况，图中还展示了闭合项$\mathbf{f}$ 的流向分量。两种情况下均只显示了区域的一部分。

其中， p 是平均压力场。在这一框架下，作为闭合项的强迫项 $\mathbf{f}$被确定为雷诺应力张量。原则上，在数据可用时，可以直接计算为

$$\mathbf{f} = -\nabla \cdot (\overline{\mathbf{u}'\mathbf{u}'}). \tag{4}$$

在实际应用中，以这种方式计算强迫项需要求解直接数值模拟，或进行时间分辨实验测量，因为脉动并不直接取决于平均流场。这个问题通常称为闭合问题。对于湍流，人们已提出多种近似模型，例如基于Boussinesq假设的模型，如$k-\epsilon$ 或 $k-\omega$ [Wilcox 等人， 1998]，以及更复杂的模型，如显式代数雷诺应力模型 [Wallin 和 Johansson， 2000]，或微分雷诺应力模型 [Cécora 等人， 2015]。本文将重点探讨 Foures 等人 [2014] 和 Patel 等人 [2024], 提出的数据同化方法，该方法不采用诸如 Boussinesq假设之类的假设。在此背景下，我们旨在将闭合项近似为优化过程中的强迫项，而不依赖基于伴随的方法循环。我们训练一个基于图神经网络（GNN）的代理模型，其输出为强迫项（式 4），输入则为平均流场$\mathbf{u}$。

3 实验设计与数值设置

本研究的主要目标是开发一种基于图神经网络的代理模型，以闭合项$\mathbf{f}$ 为目标输出、以平均流场为输入，对其进行模拟。因此，为了训练图神经网络，我们构建了一个流体动力学数据集。本节考虑不同雷诺数下钝体的几何形状，从而构建一个全面的数据集，以涵盖更多种类的流动情景。下文将在第 3.1 节介绍参考案例——即 圆柱绕流，并在第 3.2 节讨论绕随机几何形状的流动。第 3.3 节简要介绍为此开展的数值模拟，更详细的内容见附录 A。

3.1 圆柱绕流

圆柱绕流形成的非定常尾流是流体动力学文献中记录详尽的基准案例，常被作为数据同化研究中的参考案例，用于评估新技术；在临界雷诺数为 $Re_c \cong 46.7$ 之前，它表现为定常状态，达到该数值后会发生超临界 Hopf分岔 [Provansal 等人，1987年，Giannetti 和 Luchini， 2007]。在更高的雷诺数下，基流变为不稳定解，非定常流发展为一种称为冯·卡门涡街的极限环（图 1a）。对于二维（2D）情况，这种行为可持续至 $Re = 150$。超过这一点后，除周期性涡旋形成外，还开始出现不规则的速度脉动 [Anatol， 1958]；需要注意的是，在考虑三维情况时会发生转捩为湍流，这超出了本文的研究范围。相反，本文旨在定量分析在低数据量极限下利用主动学习过程训练的图神经网络模型模拟器；因此，我们将在表现出极限环行为的更简单二维场景中，对所提出的算法进行基准测试，范围为 $50 \leq Re \leq 150$；我们在上述范围内针对不同的 Re数进行数值模拟，并采用 $\Delta Re = 10$，以收集训练图神经网络模型和构建数据集所需的数据。数值模拟在时间上进行解析，而平均流场$\mathbf{u}$ 和强迫项$\mathbf{f}$ （方程4）则通过实时平均计算至收敛。作为这类数据的示例，图1（a）展示了平均流场$\mathbf{u}$ 的流向分量及其涡量等值线$\omega = \nabla \times \mathbf{u}$，而图1（$b$）展示了 $Re = 150$对应的闭合项$\mathbf{f}$ 的流向分量。

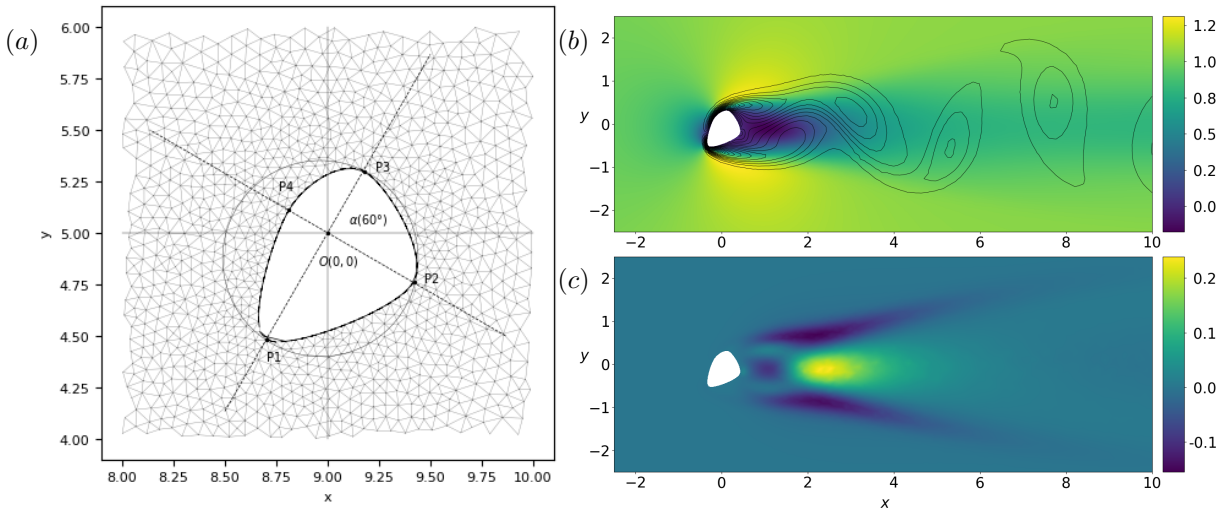

图2：在 $Re = 90$ 处，流场绕过随机生成形状的钝体。（a）钝体的几何形状$\alpha = 60°$，且钝体外接圆的直径为单位长度。所得平均流场和涡量等值线 $\omega = \nabla \times \mathbf{u}$ 如（b）所示，而（c）展示了区域一部分中的强迫应力项。

3.2 随机形状周围的流场

如前所述，通过纳入不同的几何形状可以丰富数据集；因此，除圆柱绕流外，我们还将随机形状障碍物周围的流场纳入数据集。原则上，我们应确定临界雷诺数（Re_c），在该数值下，所分析的每一种随机几何形状周围的流动都会发展为不稳定基流。这需要开展大规模数值模拟，并按照 Chiarini 等人近期工作的思路进行全面的稳定性分析 [2021, 2022]。因此，我们采取一种更务实的方法，仅保留在范围 $50 \leq Re \leq 150$ 内发展出非定常流的几何形状。

关于物体几何形状的定义，每个形状由一组样条曲线构成，这些样条曲线在 4 个控制点处连接，从而确保它们之间具有 $C1$ 连续性。控制点（图 $2a$ 中的 P1 至 P4）位于两条正交轴上，这两条轴相对于笛卡尔坐标参考系旋转了 α 角。它们与原点 O(0,0) 的距离范围为 $|s| \in [0.1, 0.6]$，离散步长为 $\Delta s = 0.1$。此外，还增加了一个自由度来控制角度 $\alpha \in [0°, 90°]$，离散步长为 $\Delta\alpha = 30°$。

最后，随机形状生成脚本会避免重复出现定义形状的同一组自由度，从而确保每种几何形状在最终数据集中只出现一次。请注意，每个随机形状的外接圆圆心不一定与坐标系原点重合。这并非偶然：在神经网络训练中，这种策略相当于在训练数据集中引入一定的位置噪声，从而促进模型更稳健地学习底层数据，并提升其泛化能力。训练数据中的这种变化还有助于降低过拟合风险，并增强神经网络将预测推广至新的、未见过的几何形状的能力。

3.3 数值设置

在数值设置中，我们采用每个生成的钝体外接圆的直径 D 作为特征尺寸。基于这一尺寸，计算域在流向的延伸范围为 $L_x = 27$，在横向的延伸范围为$L_y = 10$；坐标系原点设于 O(0，0)，沿流向距入口 $\Delta x = 9$，距对称边界 $\Delta y = 5$。

流场从入口开始演化，在入口处施加无量纲均匀速度 $\mathbf{u} = (1, 0)^T$，并以参考速度 U_∞ 进行归一化。参照 Foures 等人 [2014]，的研究，我们引入以下一组

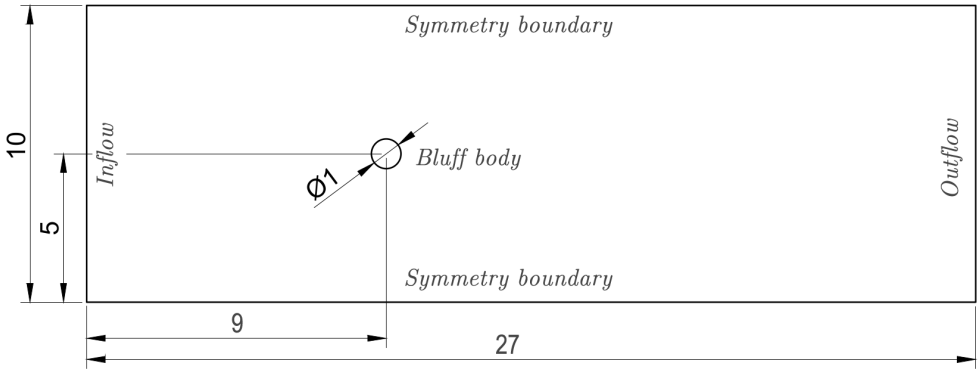

图 3：计算域几何示意图。钝体外接圆的直径、区域的高度和长度均以无量纲单位表示。

边界条件

$$\begin{cases} u = 1, \;\; v = 0 & \text{at the inlet,} \\ u = 0, \;\; v = 0 & \text{on the cylinder surface,} \\ \partial_y u = 0, \;\; v = 0 & \text{on symmetry boundaries,} \\ \dfrac{1}{Re}\partial_x u - p = 0, \;\; \partial_x v = 0 & \text{at the outlet.} \end{cases} \tag{5}$$

所有数值模拟均以 $t = 0$ 处的零流场为初始条件。所需的统计量，即平均流场和二阶矩，在数值模拟过程中实时计算。模拟的最终时间 T 由收敛准则确定，该准则基于两个连续平均流场之间差值的 L2 范数低于阈值 10^{-8}。

空间离散采用有限元法（FEM），基于 FEniCS Python 库 [Alnæs 等人，2015]；时间推进采用二阶后向差分公式（BFD）。数值模拟所用网格在障碍物附近及尾流区进行了加密，以尽可能准确地捕捉流动动力学。根据障碍物形状，整个区域的网格约包含 13500 个节点；有关所采用数值格式及数值模拟收敛性的更多细节见 A。

4 方法

本节旨在简要概述图神经网络的一些基本特征。Hamilton [2020]对该架构的主要特征进行了全面而通用的描述。这里，我们聚焦于图神经网络所依据的主要过程：利用消息传递（MP）算法，将从节点和连接中提取的信息扩散到整个网络中；第 4.1 节主要介绍消息传递，第 4.2 节则介绍数据结构。第 4.3 节描述训练过程。

4.1 神经网络架构

消息传递可分为三个基本步骤，下面结合我们的具体情况逐一说明。图 4 展示了这一过程的示意图。

1. **消息创建** – 在每个节点 i上，都会创建一个与向量 $\mathbf{h}_i$ 相关联的嵌入状态。该状态初始化为零，并会随着消息传递的进行不断嵌入信息。向量 $\mathbf{h}_i$ 的维度 d_h 在所有节点上保持不变。它是一个重要的模型超参数，与更新层的数量 k共同决定图神经网络的表达能力 [Gühring et al., 2020] ，即 模型表示一定复杂度函数的能力。嵌入状态没有直接的物理含义。
2. **消息传播**——在这一步中，信息得到传播。为了捕捉底层系统的对流/扩散动力学，即 再现雷诺平均纳维–斯托克斯方程的模型，我们双向传递消息：从每个节点 i 传递到其相邻节点 j，反之亦然，从节点 j 传递到起始

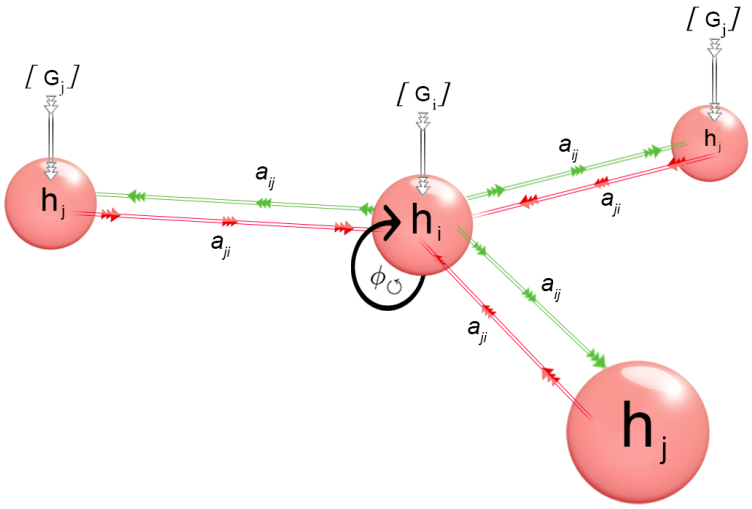

图 4：图神经网络消息传递过程示意图：给定第 i 个节点，$\mathbf{h}_i$ 是定义在 i 上的嵌入状态；a_{ij} 是节点之间的有向连接；$\mathbf{G}_i$ 表示注入每个节点的外部输入（详情见第 4.1 节）。

节点 i（示意图见图 4）。从数学上讲，对于每一对相连的节点 i 和 j，我们可以写出以下关系

$$\boldsymbol{\phi}_{i,j}^{(k)} = \zeta^{(k)}\left(\mathbf{h}_i^{(k-1)}, \mathbf{a}_{ij}, \mathbf{h}_j^{(k-1)}\right). \tag{6}$$

下文中，下标表示节点索引，上标表示更新索引。向量$\mathbf{h}_i^{(k-1)}$是来自前一层 $k-1$ 的嵌入状态，$\mathbf{a}^{ij}$ 表示节点 i 与节点 j 之间的有向连接，而 $\zeta^{(k)}$ 是一种通用的可微算子，例如多层感知机（MLP），参见 Goodfellow 等人 [2016]。

3. **消息聚合**——最后，在每个节点 i 上，将收集到的信息进行聚合，为每个节点生成更新后的嵌入状态$\mathbf{h}_i^{(k)}$

$$\mathbf{h}_i^{(k)} = \mathbf{h}_i^{(k-1)} + \alpha\Psi^{(k)}\left(\mathbf{h}_i^{(k-1)}, \mathbf{G}_i, \boldsymbol{\phi}_{i,\rightarrow}^{(k)}, \boldsymbol{\phi}_{i,\leftarrow}^{(k)}, \boldsymbol{\phi}_{i,\circlearrowleft}^{(k)}\right), \tag{7}$$

其中，输入为平均流场和雷诺数，包含在向量 $\mathbf{G}_i = \{\mathbf{u}\ Re\}$ 中，并在每次更新时输入；$\overline{\boldsymbol{\phi}}_i^{(k)}{}_{,\rightarrow} = \frac{1}{N_i}\sum_{j\in N_i}\boldsymbol{\phi}_{ij}^{(k)}$，是节点 i 发送给其 N_i 个邻居的消息，并对邻居数量取平均；$\boldsymbol{\phi}_i^{(k)}{}_{,\leftarrow} = \frac{1}{N_i}\sum_{j\in N_i}\boldsymbol{\phi}_{ij}^{(k)}$，是节点 i 从其 N_i 个邻居接收的消息，并对邻居数量取平均；$\boldsymbol{\phi}_i^{(k)}{}_{,\circlearrowleft} = \boldsymbol{\phi}_{ii}^{(k)}$，是节点 i 发送给自身的消息，以避免随着过程推进而丢失信息。$\Psi^{(k)}$ 是一种通用的可微算子，在此情况下也可用多层感知机近似；α 是包含在模型超参数中的松弛系数，用于在消息传递循环中，依据先前的嵌入状态对每次更新进行缩放。请注意，在 $\boldsymbol{\phi}_i^{(k)}{}_{,\rightarrow}$ 和 $\boldsymbol{\phi}_i^{(k)}{}_{,\leftarrow}$ 中使用均值算子，使我们能够针对每个第 i 个嵌入状态，收集来自邻居的信息，而每个第 i 个节点的邻居数量可能各不相同。尽管如此，此步骤中也可以引入不同的置换不变函数（最大值、求和、拼接等）。

在消息传递过程结束时，每个节点上的信息都已按上述方式处理，因此包含了来自图中任何其他点的信息，包括它们之间的相对距离。为了使每个节点上的消息都能到达网格中的其他节点，我们的直觉是——原则上——更新次数 k 应覆盖网格 [上可定义的最长测地路径（Donon 等人， 2020]）；在实践中，本研究中的更新次数是固定的，并通过遗传算法进行优化（详情参见 Akiba 等人 [2019] 和 B）。最新的嵌入状态通过解码器投影回物理状态；解码器同样由可微函数近似，在我们的案例中，该函数是一个多层感知机。

4.2 数据结构化

图神经网络与非结构网格的结合应用，基于将网格数据表示为图这一理念。在这种表示中，网格的每个节点都与图神经网络中的一个节点一一对应。为了实现这种对应关系，我们将数据组织成张量形式，使其能够作为图进行处理，同时保留网格的邻接属性。对于每个案例，我们构建：

- **节点特征矩阵 A**：该矩阵 $\mathbf{A} \in \mathbb{R}^{n_i \times d_h}$ 表示网格中每个节点的特征。其中，n_i表示网格中的节点数量，d_h 表示嵌入状态的维度，我们将有限元量投影到该状态中；每个节点均分配有唯一索引。
- **边索引矩阵 C**：矩阵 $\mathbf{C} \in \mathbb{R}^{c \times 2}$ 定义了网格中的边。其中，c 表示网格中的边数量，每一行包含一对表示相连节点的索引。
- **边属性矩阵 D**：矩阵 $\mathbf{D} \in \mathbb{R}^{c \times 2}$包含相连节点之间在 x 和 y 方向上的计算距离。这些距离包含其符号，并且 **D** 与**C**共享相同的行索引 c。

矩阵定义如下：

$$\mathbf{A}_{n_i,d_h} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,d_h} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,d_h} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_i,1} & a_{n_i,2} & \cdots & a_{n_i,d_h} \end{bmatrix},$$

$$\mathbf{C}_{c,2} = \begin{bmatrix} i_1 & j_1 \\ i_2 & j_2 \\ \vdots & \vdots \\ i_c & j_c \end{bmatrix}, \qquad \mathbf{D}_{c,2} = \begin{bmatrix} x_{i_1} - x_{j_1} & y_{i_1} - y_{j_1} \\ x_{i_2} - x_{j_2} & y_{i_2} - y_{j_2} \\ \vdots & \vdots \\ x_{i_c} - x_{j_c} & y_{i_c} - y_{j_c} \end{bmatrix}.$$

矩阵 $\mathbf{A} \in \mathbb{R}^{n_i \times d_h}$ 的每一列都被指定为一个特征向量，输入到图神经网络消息传递步骤中所采用的多层感知机结构中的单个神经元，即 ζ（见第 4.1 节）。这些特征向量在训练过程开始时被设为零（见第 4.3 节）。多层感知机结构由嵌入状态的维度 d_h 定义，而非由每个节点的特征数定义。这样，无论网格中的节点数如何变化，多层感知机架构都保持不变。实际上，改变网格中的节点数只会影响矩阵 $\mathbf{A} \in \mathbb{R}^{n_i \times d_h}$中的行索引 n_i、矩阵 $\mathbf{C} \in \mathbb{R}^{c \times 2}$中的连接数 c 以及 $\mathbf{D} \in \mathbb{R}^{c \times 2}$ 中的距离，但不会改变多层感知机结构，因为该结构由嵌入状态的维度 d_h 定义。因此，无论几何形状或节点数量如何变化，同一多层感知机架构都可用于不同的有限元模拟。此外，图神经网络的输入不包含绝对空间坐标，这意味着它不会学习空间位置与目标输出之间的特定映射。因此，图神经网络能够保持对钝体位置和网格细化的不变性，从而增强其在不同模拟设置中的泛化能力。与卷积神经网络等标准技术相比，图神经网络架构具有更强的灵活性，因此能够轻松处理非结构网格、从各种几何形状中学习，并在不同几何形状上重建强迫场。最后，该模型配备了有限元-图神经网络接口，可将图神经网络的数值结果映射为在有限元环境中计算或存储的向量场。

4.3 训练算法

下面结合图5说明本文采用的图神经网络训练框架。起点为$\mathbf{A}^0$，它表示由定义在各节点上的所有嵌入状态构成的矩阵，初始状态为零。该矩阵连同外部注入的量（**G**，即本例中的平均流场和雷诺数 Re）一起输入消息传递过程 MP^1（见第**4.1**节）。后者将输出更新后的嵌入状态张量$\mathbf{A}^1$，并将其传入解码器$\mathbf{D}^1$, 。解码器是一个可训练的多层感知机（MLP）函数，负责从嵌入状态中重建有意义的物理状态（$\hat{\mathbf{f}}^1$）；在本例中，该物理状态为雷诺平均纳维–斯托克斯方程的闭合项$\bar{}$。最后，通过损失函数 $\ell = (\hat{\mathbf{f}} - \bar{\mathbf{f}})^2$将该物理状态与来自 DNS 的真实值进行比较。重复整个过程，直至到达最后一层 k。最后一次预测（$\hat{\mathbf{f}}^k$）即为整个算法的输出。按照 Donon 等人 [2020], 提出的流程框架，为增强学习过程的稳健性，将所有中间损失值纳入全局损失函数：

$$L = \sum_{k=1}^{\bar{k}} \gamma^{\bar{k}-k} \left(\frac{1}{n_i} \sum_{i=1}^{n_i} \ell_i \right), \tag{8}$$

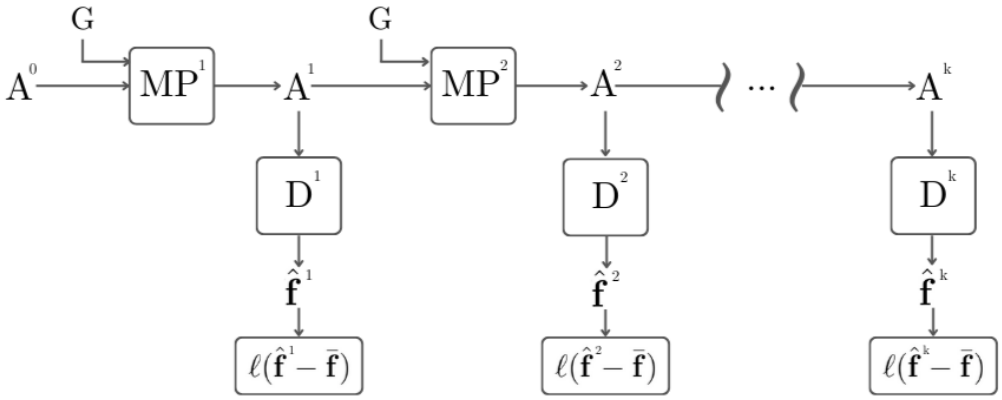

图 5：图片改编自 Donon 等人 [2020]，展示了图神经网络训练过程的整体框架。MP^k 为消息传递算法；$\mathbf{D}^k$ 为可训练的多层感知机解码器；$\mathbf{A}^k$ 为包含各节点嵌入状态的矩阵；$\mathbf{G}$ 为包含注入图神经网络的输入的向量。详情请参阅第 4.3 节。

¯其中，n_i 为节点数，k 为更新层数，γ 为用于控制与嵌入状态各次更新相关贡献重要性的超参数。消息的相对重要性由指数项 $\gamma^{\bar{k}-k}$加权，其中 $\gamma < 1$；因此，最新的更新 k 将包含最丰富的信息，并在过程中具有最高的重要性。附录B列出了优化后超参数及其对应取值。

5 结果

本节旨在介绍研究结果。我们在第 5.2 节对该方法进行概念验证；在这项初步分析中，我们使用由不同雷诺数下圆柱绕流数值模拟数据构成的数据集训练图神经网络。我们的目标是利用第 5.1 节中的基准案例，验证该方法在未见条件下的鲁棒性。在第 5.2 节中，我们重点讨论如何恰当选择数据集，以及训练数据的选择会如何影响图神经网络的泛化能力。在第 5.3 节中，我们将随机生成的钝体绕流数值模拟所得的流场纳入数据集（形状生成的详情见第 3.2 节）；第 5.3.1 节和第 5.3.2 节分别讨论数据集增强和数据集扩充的影响，而第 5.3.3 节则讨论基于主动学习准则进行的数据选择。第 5.4 节总结了这些训练方法的定量比较分析；下文将介绍本节所有比较中使用的 4 个参考案例。

5.1 测试案例

我们引入了4个基准测试，以评估图神经网络模型的性能。这些测试案例从案例1到案例4进行标记，旨在评估图神经网络预测的特定方面。我们涵盖了广泛的场景，包括雷诺数和几何形状的变化、相对于入口的位置变化以及障碍物数量的变化。

1. **案例1**：在此情景中，我们考虑圆柱绕流，其中雷诺数增加至$Re = 200$。该数值超出了训练数据集的范围（范围为 $50 \leq Re \leq 150$），用于测试模型超越训练数据进行外推的能力。2. **案例2**：在此处，我们引入一个未包含在训练数据集中的随机形状钝体绕流作为测试案例，其条件为 $Re = 120$，以评估模型对未见形状的泛化能力。3. **案例3**：本案例涉及随机形状钝体绕流的数值模拟数据，该形状未出现在训练数据集中，条件为 $Re = 100$。障碍物位于数据集中所采用参考位置的下游。4. **案例4**：本测试考虑双圆柱并列构型；流场中存在两个钝体障碍物，这与训练中采用的单个障碍物不同。雷诺数为 $Re = 90$，旨在测试诱发不同动力学特征的几何形状，即由多个相互作用的钝体产生的几何形状。

图6显示了针对4个基准测试案例1–4，通过数值模拟计算得到的平均流场速度的流向分量$\mathbf{u}$、涡量 $\omega = \nabla \times \mathbf{u}$ 的等值线（左列），以及强迫应力$\mathbf{f}$ 的流向分量（右列）。

将使用两个度量指标来评估图神经网络的性能。第一个指标是比较真实值$\mathbf{f}$ 与图神经网络对各个案例的预测，并将其视为基于 $L2$-范数的相对误差 ε，定义如下：

$$\varepsilon = \frac{||\mathbf{f} - \hat{\mathbf{f}}||}{||\mathbf{f}||} = \frac{\left[\int_\Omega (\mathbf{f} - \hat{\mathbf{f}})^2 d\Omega\right]^{1/2}}{\left(\int_\Omega \mathbf{f}^2 d\Omega\right)^{1/2}}. \tag{9}$$

第二个度量指标旨在 评估平均流重构 $\hat{}\mathbf{u}$ 的准确性，该重构由图神经网络预测得出
，并将其与真实值进行比较。其定义如下：

$$\delta = \frac{||\mathbf{u} - \hat{\mathbf{u}}||}{||\mathbf{u}||} = \frac{\left[\int_\Omega (\mathbf{u} - \hat{\mathbf{u}})^2 d\Omega\right]^{1/2}}{\left(\int_\Omega \mathbf{u}^2 d\Omega\right)^{1/2}}. \tag{10}$$

$\hat{}$ 在前述方程中，Ω 是计算域（第 3.3 节），$\mathbf{f}$ 是图神经网络预测，$\mathbf{f}$ 是来自 DNS 的雷诺平均纳维–斯托克斯方程 $\hat{}$ 的闭合项。将 $\mathbf{f}$ 代入公式 3 并求解反问题，可以得到 $\hat{}\mathbf{u}$，即基于图神经网络预测对平均流场的重构。

5.2 概念验证训练：圆柱绕流

这里，我们考虑基线结果，在下文中将其称为概念验证（PC）。训练数据集包含通过直接数值模拟获得的圆柱绕流数据；该数据集由 11 对平均流场$\mathbf{u}$ （输入）和雷诺强迫应力$\mathbf{f}$ （目标）组成，范围为 $50 \leq Re \leq 150$，步长为 $\Delta Re = 10$。如前所述，数据集中的所有案例都表现出钝体尾流中的卡门涡街失稳。

图 7 展示了该训练方法的训练损失曲线和验证损失曲线。训练损失的变化表明，该模型能够捕捉流体流动的动力学特征，并学习训练数据集中的模式。然而，在训练损失曲线和验证损失曲线之间可以观察到明显差距。这种差异表明模型可能存在对训练数据过拟合的问题。考虑到训练数据集完全由经过相同几何形状的数值模拟获得的数据组成，这一问题并不意外。因此，当图神经网络模型使用未见过的形状或流体动力学条件进行测试时，其性能会下降。

关于测试案例，案例1（图 $8a$）中，图神经网络预测在较高雷诺数 $Re = 200$ 下能够准确再现流体结构，$\varepsilon = 0.1153$。图 8（b）展示了案例2，其中图神经网络预测与直接数值模拟数据之间的差值主要出现在近尾流区。该区域对不稳定动力学的发展至关重要，因为它对应于所谓的波发生器区域 [Giannetti 和 Luchini,2007]；对于这类非标准几何形状的案例，其流场涡脱落行为可能与标准圆柱案例中观察到的情况大不相同。本案例中的 ε 范数为 $\varepsilon = 0.2995$。图 8（c）中的案例3用于测试钝体发生变化时图神经网络的泛化能力；主要差异同样出现在钝体的近尾流区。本案例中的 ε 范数为 $\varepsilon = 0.2084$。最后，图 8（d）中的案例4尤具挑战性，因为它测试了存在多个钝体时的泛化能力。图神经网络模型的预测捕捉到了每个圆柱强迫应力的主要特征。然而，差异主要出现在两个圆柱之间的区域，并延伸至远尾流区，在那里两个圆柱的雷诺应力场会发生相互作用。本案例中的 ε 范数为 $\varepsilon = 0.8704$，显著高于前几个案例。

在此前讨论的所有案例中，都可以观察到预测结果存在一些数值误差，但这些误差主要归因于统计噪声，或神经网络模型固有的典型误差。神经网络本身确实存在统计不确定性，这源于其训练算法的随机性（例如权重随机初始化、训练过程中的批次选择），以及用于表征底层物理的模型所具有的近似性质。除了这一点之外，基于这些基准案例的初步分析表明，通过采用包含更多种钝体形状的训练集，图神经网络模型的性能还可以进一步提升。因此，接下来我们将研究数据增强的影响：在训练数据集中加入随机形状钝体，以增强模型的鲁棒性和泛化能力。

5.3 数据增强与主动学习：流体流过随机几何形状

下面，我们研究数据量和质量对图神经网络模型训练过程及其泛化性质的影响。本研究分为若干个不同步骤。

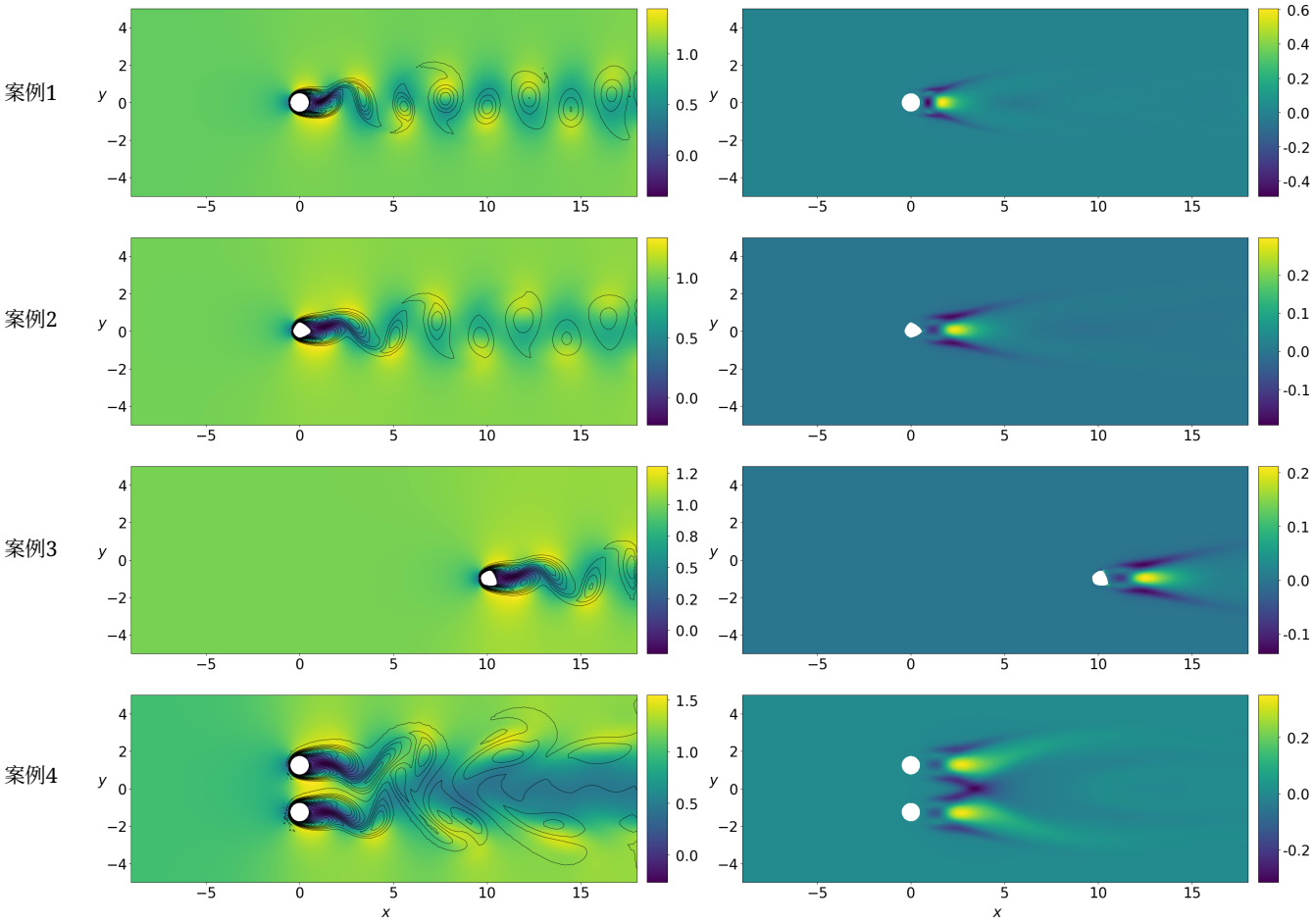

图6：作为第5节所讨论学习策略的基准所用的测试案例。从上至下：案例1，$Re = 200$处的圆柱绕流；案例2，$Re = 120$处随机形状钝体绕流；案例3，计算域中偏移至$Re = 100$处的随机形状钝体绕流；案例4，$Re = 90$处双圆柱并列构型的绕流。对于每个案例，左列显示平均流场的流向分量$\mathbf{u}$及涡量等值线$\omega = \nabla \times \mathbf{u}$，右列显示强迫应力的流向分量$\mathbf{f}$。

1. **数据增强（DA）**（第5.3.1节）；第一种方法涉及数据集增强，即将随机形状钝体案例纳入现有数据集，同时保持与PC案例相同的数据集规模，即即11个案例。该方法旨在使模型训练过程中使用的几何形状更加多样化，同时不增加数据量。2. **数据集扩充**（第5.3.2节）；第二种方法通过纳入33个随机形状钝体案例来扩充数据集。我们通过进行k折验证，根据模型对特定构型的敏感性评估所用数据的质量。3. **主动学习（AL）数据选择**（第5.3.3节）；在最后一种方法中，通过依据相似性准则逐步添加所选数据来构建训练集。目标是开发一个代理模型，使其能够基于给定数据集，尽可能好地泛化到未见案例。

5.3.1 使用随机形状进行数据增强

在研究的这一阶段，我们采用分层随机抽样准则，从第 3.2 节所述随机生成的构型中选取形状。具体而言，我们针对区间 $50 \leq Re \leq 150, \Delta Re = 10$ 内的每个雷诺数生成完整的形状集合，并为每个雷诺数随机选取 1 个形状，共选出 11 个形状，组成新的训练数据集。所选几何形状各不相同。在

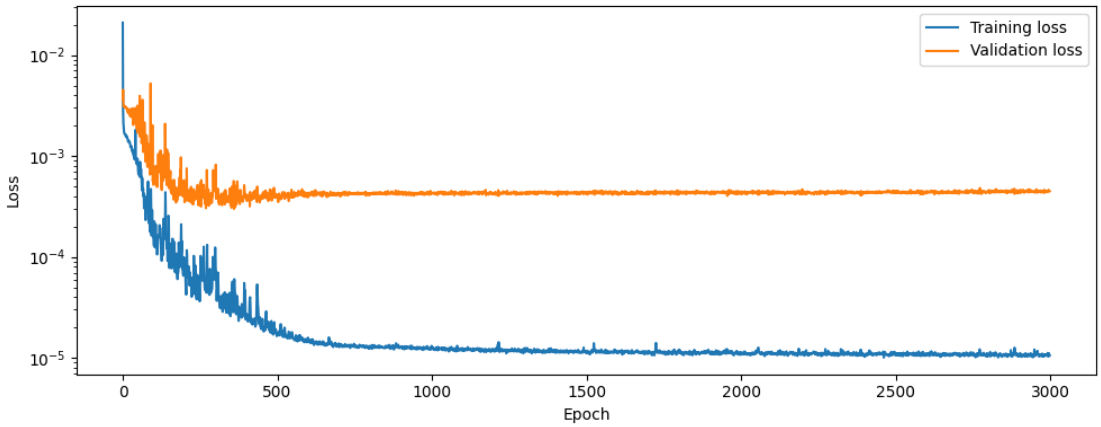

图7：概念验证基线的训练损失和验证损失曲线，训练至 3000 轮。训练数据集由 11 组平均流场-强迫项对组成，这些数据在不同雷诺数下叠加，范围为$50 \leq Re \leq 150$ 至 $\Delta Re = 10$。验证集由第5.1节中介绍的测试案例构成。

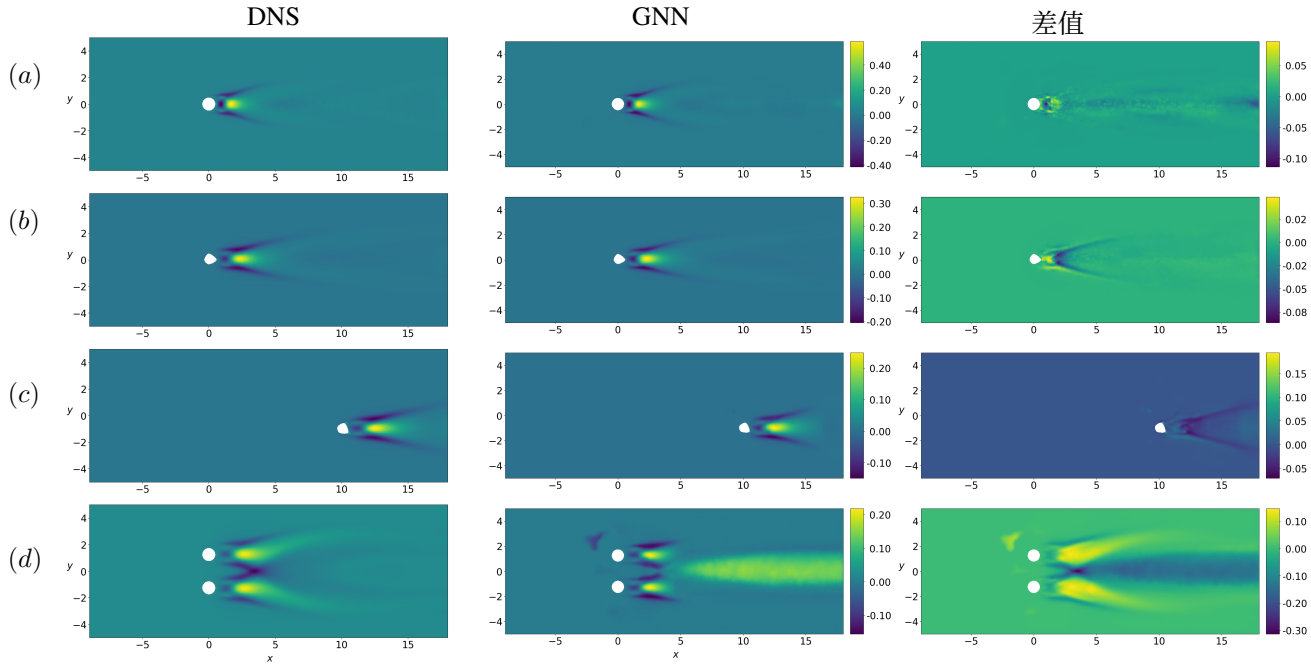

图8：雷诺应力张量流向分量的对比。图神经网络模型使用由 11 个圆柱形钝体案例组成的数据集进行训练，范围为 $50 \leq Re \leq 150$、 $\Delta Re = 10$。（a）案例1， $Re = 200$，圆柱钝体形状；（b）案例2， $Re = 120$，随机形状钝体；（c）案例3， $Re = 100$，随机形状偏移钝体；（d）案例4， $Re = 90$，绕流双并列圆柱。

这些选定形状上进行直接数值模拟，以获得用于训练图神经网络的平均流场（输入）和强迫应力（目标）。

图9展示了采用第二种训练方法时的训练损失曲线和验证损失曲线。训练损失持续下降，体现了图神经网络模型的学习效果。值得注意的是，尽管训练损失曲线与验证损失曲线之间仍存在显著差距，但与最初的训练方法相比（图7），这一差距有所缩小。这一差距的缩小表明过拟合程度有所降低，也说明在训练数据集中引入更多样的形状和流动条件，有助于提升图神经网络模型的泛化能力。

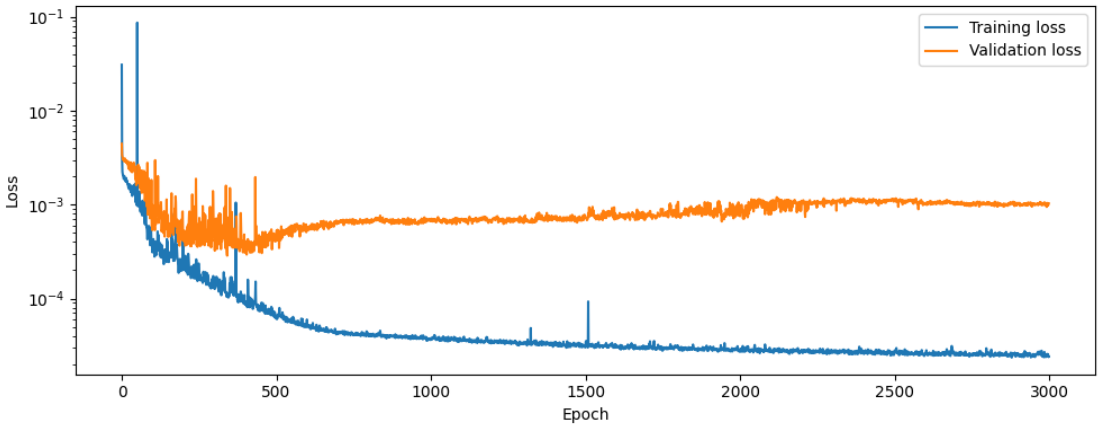

图9：在数据集中引入11种随机几何形状流动案例进行数据增强时的训练损失和验证损失曲线。雷诺数的变化范围为 $50 \leq Re \leq 150$，其中 $\Delta Re = 10$。训练持续至3000个训练轮次。验证集由第5.1节中给出的测试案例构成。

关于图神经网络预测的测试案例，与第5.2节讨论的训练方法相比，案例1（图 $10a$）在圆柱附近的强迫项预测偏差更大，$\varepsilon = 0.5033$。对此的一种解释是，PC训练数据集仅包含圆柱几何形状；因此，最终的图神经网络模型专门用于预测圆柱绕流。相比之下，当前训练策略涉及绕随机形状的流动：一方面，这种选择通过增强泛化能力，提高了图神经网络捕捉多种流动条件的灵活性；另一方面，它也同时降低了圆柱形状案例的预测精度。

案例2（图 $10b$）和案例3（图 $10c$）的差值场主要呈现数值噪声和统计噪声，而预测结果很好地再现了雷诺应力张量的主要特征。与案例1中观察到的情况相反，这种训练方法能够得到一个高效的图神经网络模型，用于预测流经随机形状钝体的流动特征。两个测试案例的误差范数相近，案例2为 $\varepsilon = 0.2063$，案例3为 $\varepsilon = 0.1999$。最后，案例4的误差主要位于尾流区，与PC案例中观察到的情况相似（读者可对比图 $8d$ 和图 $10d$）。然而，值得注意的是，与第一种训练方法中观察到的误差相比，这些误差在定量上小得多，其中 $\varepsilon = 0.6312$。

5.3.2 数据集扩充

提升神经网络模型泛化能力的一种常见方法，是增加模型训练所用数据的数量。这种方法可以降低神经网络模型对小型数据集中特定条件过拟合的风险，也更有可能捕捉到底层现象。因此，我们研究了将训练数据集中的数据量增加至原来的三倍所产生的影响。采用第 5.3.1 节所述的相同分层随机抽样方法，我们在参考区间内选取的 11 个雷诺数中，每个雷诺数对应 3 个案例，从而构建了一个包含 33 个案例的训练数据集。为了分析图神经网络模型的泛化能力对训练数据集的敏感性，我们采用 k 折交叉验证。首先，将训练数据集划分为 k 组，并训练 k 个不同的模型。对于每个模型，我们使用 $k-1$ 组作为训练数据集，并将剩余的第 k 组作为验证数据集。尽管这种方法在实际应用中并不可行，但它对于初步评估数据质量对最终预测的影响非常有帮助。

图 11(a) 展示了 10 个模型的训练曲线，每个模型对应 k 折中的一折；图 11(b) 则展示了它们相应的验证曲线。无论选择哪个数据集，训练曲线都较为接近，这表明所选图神经网络架构的表现相当一致。然而，验证损失变化显著。这种差异可能归因于训练数据集的分布。不一致或不平衡的数据分布可能导致模型基于无法充分代表整体数据集的数据子集进行训练，从而对验证结果产生负面影响。为应对这一挑战并确定最有效的数据集，下面引入主动学习。

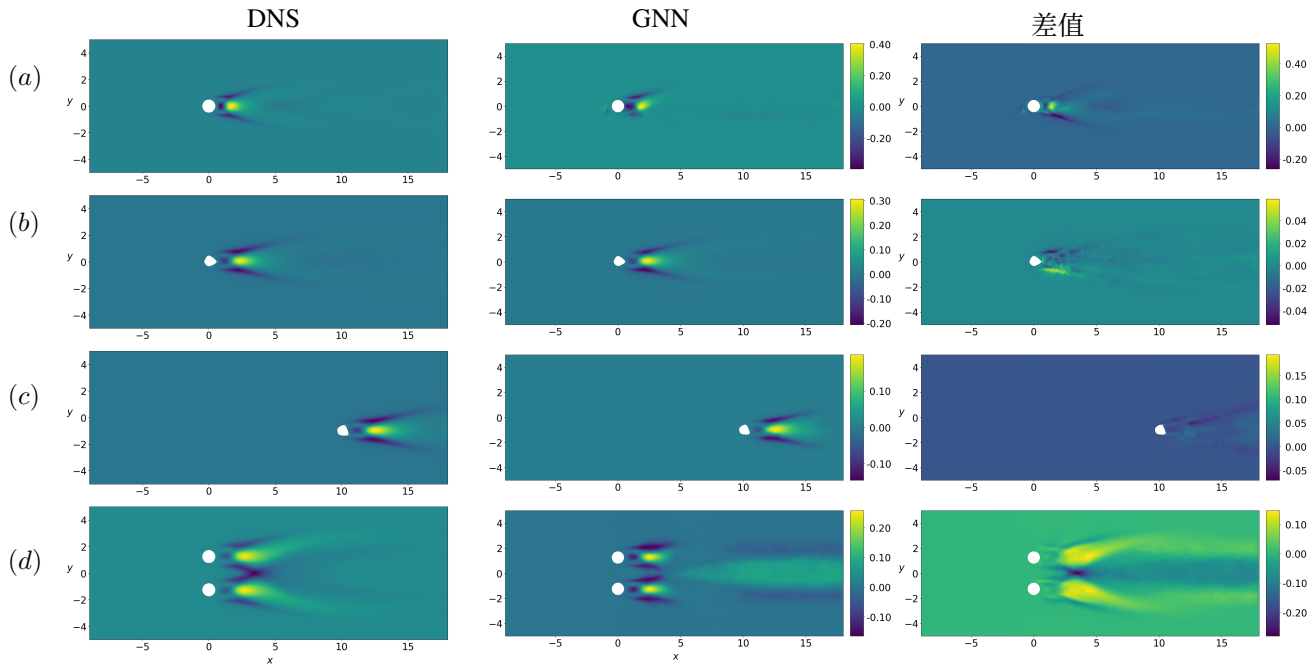

图 10：雷诺应力张量流向分量的对比。图神经网络模型使用由 11 个随机形状钝体案例组成的数据集进行训练，范围为 $50 \leq Re \leq 150, \Delta Re = 10$。(*a*) 案例1，$Re = 200$，圆柱钝体形状；(*b*) 案例2，$Re = 120$，随机形状钝体；(*c*) 案例3，$Re = 100$，随机形状偏移钝体；(*d*) 案例4，$Re = 90$，绕流双并列圆柱。

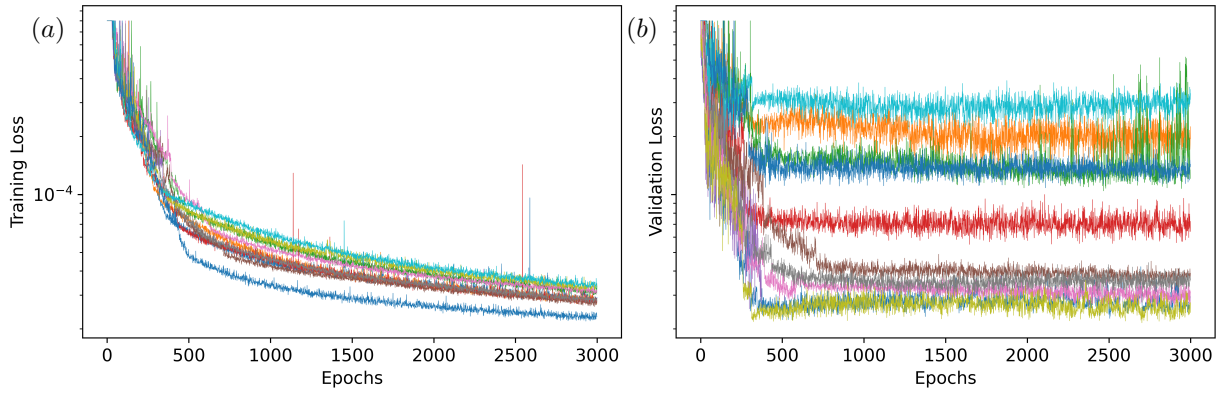

图 11：使用 $10\,k$ 折交叉验证训练至 3000 个训练轮次的模型的训练损失曲线和验证损失曲线。(*a*) 显示训练曲线，(*b*) 显示每一折对应的验证曲线。

5.3.3 主动学习数据选择

主动学习能够动态调整数据集，将目标聚焦于对模型泛化能力贡献最大的 数据。我们从图神经网络的角度采用一种相似性准则，遵循 Charpiat 等人 [2019] 的研究。其核心思想是，避免将从泛化角度看不会对模型带来任何显著改进的样本纳入训练数据集；相反，为了提高图神经网络模型的多样性，我们需要选择从图神经网络角度看彼此差异尽可能大的样本。我们可以通过标量积比较每一对可用数据的代价函数相对于图神经网络 θ 可学习参数的向量梯度，从而实现这一目标。作为比较度量指标，我们采用余弦相似度角

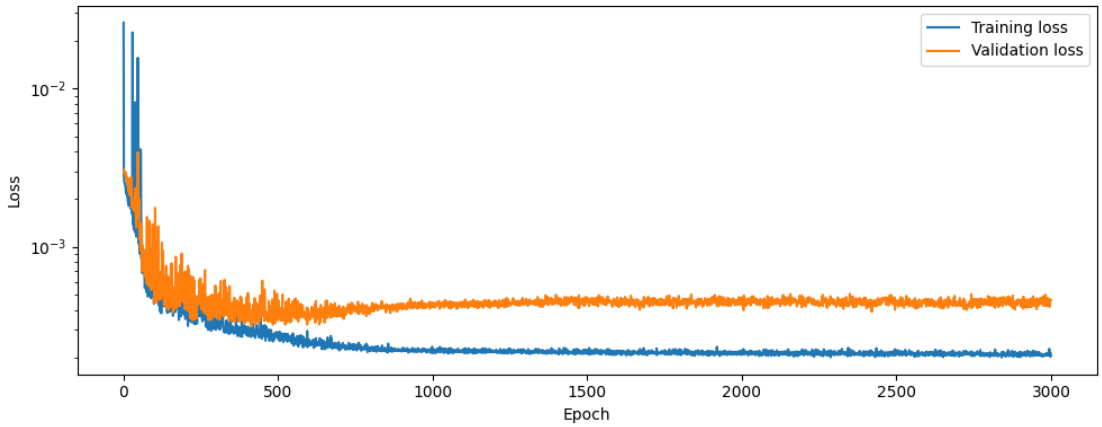

图12：展示了相似度阈值为 0。8 时，主动学习训练至 3000 个训练轮次的训练损失曲线和验证损失曲线。训练数据集由 6 个随机形状钝体构成，其雷诺数在区间 $50 \leq Re \leq 150$内变化。验证集由第5.1节中介绍的测试案例构成。

距离

$$\cos(\beta) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}, \tag{11}$$

其中，$\mathbf{a} \in \mathbb{R}^m$ 和$\mathbf{b} \in \mathbb{R}^m$ 是两个一般的 m维向量， β 是它们之间的夹角。基本思想是，每个梯度向量都会将神经网络的状态引向解空间中的一个特定方向。因此，如果多个数据点的梯度指向同一方向，将它们纳入训练可能会造成冗余。相反，我们的目标是选择并纳入训练数据集尽可能多的不同方向，以便更全面地探索解空间（另见 C）。

当达到预先设定的相似度阈值时，选择过程便停止；只有相似度低于该阈值的数据对才会被添加到数据集中。相反，相似度高于阈值的案例会被舍弃。不过，相似度阈值并无直接且易于解释的含义，需要根据所需性能将其作为过程的输入参数进行调节。通常，相似度阈值越低，纳入训练数据集的案例就越多。这虽然有助于提升模型应对多样化场景的能力，但也可能使训练过程更加复杂、耗时。另一方面，过于严格的相似度阈值可能会排除有潜在价值的训练数据。因此，需要在训练计算成本与模型泛化性能之间权衡，选择最优的相似度阈值。

在本研究中，我们考察了三种不同的相似度阈值，即 $\cos(\beta) \in [0.7, 0.8, 0.9]$。这种方法能够观察不同程度的数据纳入对模型训练和性能的影响。图13和图12所示的结果采用了 0。8 的相似度阈值，这使得训练数据集中仅选取了6个随机钝体案例；其他阈值下的结果详见第5.4节）。图12给出了训练和验证损失曲线：与先前学习方法的结果相比，两条曲线之间观察到的差距减小，凸显出过拟合问题显著减轻。训练从训练数据集中的单个案例开始，具体为 $Re = 120$下的圆柱案例。

一旦向量梯度收敛（见C），便从训练集中选择一个新案例；由于该过程是一种在每一步扩大数据集的搜索并纳入算法，因此可以采用两种不同的方法。第一种策略受课程学习框架 [Bengio et al., 2009], 的启发，其中每个选定案例都会加入持续进行的训练。因此，模型会逐步更新。本研究采用的第二种方法是，每当数据集因加入一个新数据对而扩大时，便重新初始化图神经网络权重。作出这一选择的依据在于解空间的动态变化：当训练集中加入新数据时，解空间会发生改变；从这个意义上说，无法保证新的极小值相较于图神经网络的初始化点，会比相较于先前的极小值“更接近”。尽管如此，我们观察到，对于所分析的流动案例，这两种训练策略在最终结果上产生的差异可以忽略不计。

考虑用于模型评估的基准测试，案例1 中图神经网络预测的雷诺应力张量与直接数值模拟结果之间的 $\varepsilon = 0.5507$ 。对于随机形状钝体案例（案例2），误差为 $\varepsilon = 0.1518$ ；对于下游偏移随机形状钝体（案例3），误差为$\varepsilon = 0.3595$ ；对于两个并列

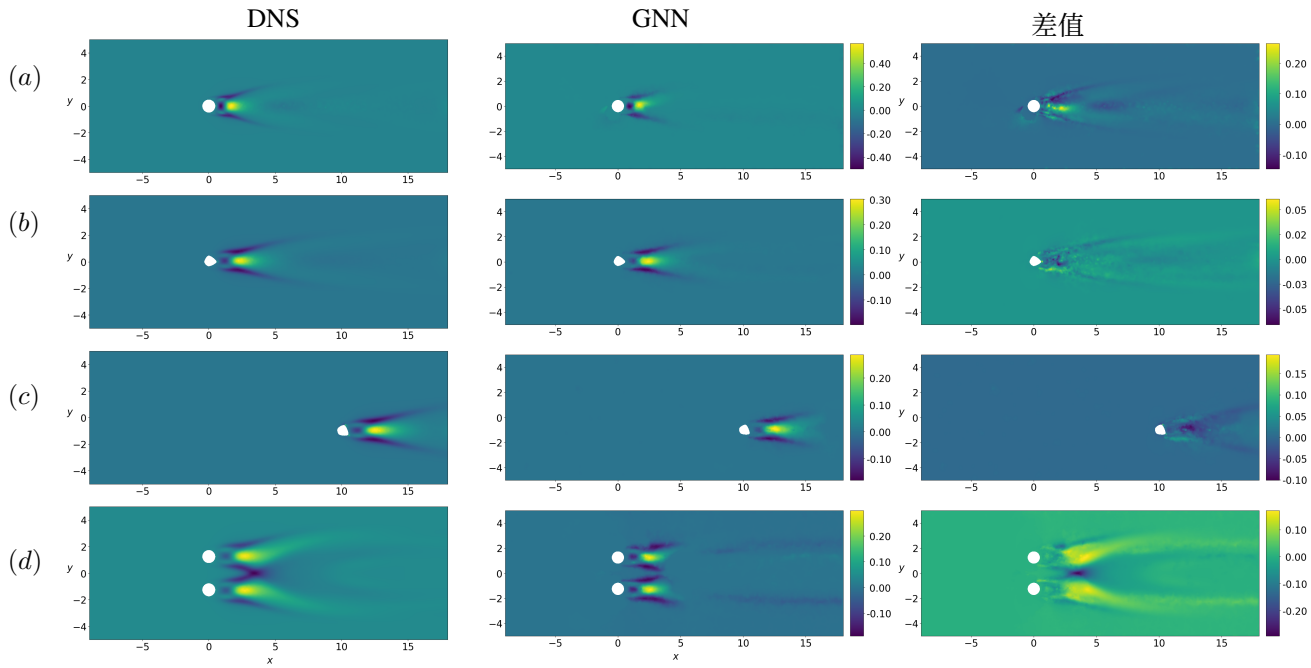

图13：雷诺应力张量流向分量的比较。图神经网络模型使用一个数据集进行训练，该数据集由通过主动学习方法选取的6个随机形状钝体案例组成，范围为$50 \leq Re \leq 150, \Delta Re = 10$。（$a$）案例1，$Re = 200$处的圆柱绕流；（$b$）案例2，$Re = 120$处随机形状钝体的绕流；（$c$）案例3，$Re = 100$处随机形状钝体向下游平移；（$d$）案例4，$Re = 90$处绕流双并列圆柱。

圆柱构型（案例4），误差为。结果保持了解相对于 y 轴的对称性，仅在远尾流区域观察到轻微的不一致；我们注意到，误差集中在钝体正下游的区域，但我们认为这些误差源于数值因素或随机性，而不是对流体结构预测不准确。

5.4 定量比较

总之，我们在表1中总结了结果。对前两种训练方法（PC和DA）进行直接比较发现，使用随机几何形状扩充训练数据集能够增强图神经网络的泛化能力。这种情况并不出人意料，不过在较高雷诺数下进行外推时（案例1）观察到了例外：PC方法的表现优于DA。正如前文所述，这一现象可以这样理解：采用PC方法训练的图神经网络模型专门用于预测圆柱绕流，因为训练时使用的是这一特定几何形状。

另一方面，与 PC 和 DA 方法相比，主动学习训练方法能够获得泛化能力总体更优的图神经网络模型，尤其是在案例2和案例4中；而在案例3中，性能基本相当。主动学习方法在全部4个测试案例中的平均误差均有所降低，表明模型的整体性能得到改善。关于训练数据的主动选择，一个重要方面是数据量。当阈值为 0.8 时，仅使用 6 对数据即可取得良好的性能，不过也可以观察到，训练所用数据对的总数会随所选阈值而大幅变化。然而，必须指出，在所有主动学习案例中，引入选择准则都能确保所选数据点带来良好的泛化性能。

与现有文献中的研究相比，我们工作的显著之处在于，我们的方法仅需少得多的训练案例，便能达到相当的准确度。例如，在 [Chen 等人 2021] 的研究中，作者针对 $Re = 10$下圆柱周围的稳态不可压缩流，在训练数据集中使用了 2000 个案例。相比之下，我们在所用的最大数据集中仅使用 19 个案例，仍能推广到不同的雷诺数和钝体位置，并取得相当的准确度。在 [Lee 和 You 2019]的研究中，训练数据集使用了 $500k$ 个案例，但其采用卷积神经网络预测非定常流。值得注意的是，图神经网络架构能够推广到不同的几何形状和 Re 数，而 [Lee 和 You 2019]的研究并未探讨这一点。最后，在 [Thuerey 等人 2020],的研究中，他们采用图神经网络来预测时间平均稳态流。该数据集由 12800 个数据点组成，而图神经网络

案例	PC		DA		AL 0.7		AL 0.8		AL 0.9	
	ε	δ	ε	δ	ε	δ	ε	δ	ε	δ
案例*1*	**0.1153**	*0.0118*	0.5033	0.1987	0.8671	0.1859	0.5507	0.1165	0.4132	0.1162
案例*2*	0.2995	0.0280	0.2063	0.0160	0.1256	*0.0079*	0.1518	0.0166	**0.1124**	0.0092
案例*3*	0.2084	0.0131	**0.1999**	0.0090	0.3058	0.0107	0.3595	0.0183	0.2123	*0.0082*
案例*4*	0.8707	0.2683	0.6312	0.1701	0.6751	0.1930	**0.5243**	*0.1565*	0.5719	0.1653
训练数据	11		11		3		6		19	

表 1：对 4 个基准案例进行比较，涵盖三种不同的训练方法：概念验证（PC）训练（第 5.2 节）、数据增强（DA）训练（第 5.3.1 节）和主动学习（AL）策略（第 5.3.3 节）。对于后者，我们考虑三个相似度阈值 $\cos(\beta) \in [0.7, 0.8, 0.9]$。所选度量指标 ε和 δ 分别在式 9 和式 10 中定义。最后一行报告了训练过程中使用的样本对数量。

该模型的复杂度超过 $30M$ 个参数，对于大多数实际应用而言难以承受。相比之下，我们的图神经网络架构最多约有 $900k$ 个参数。总之，我们认为，图神经网络模型与主动学习相结合，凭借较低的数据需求，使我们的方法更易于应用，也更适用于更广泛的场景。

6 结论

机器学习技术在流体力学中的应用往往存在一些不足，例如过拟合、对未见案例的预测缺乏鲁棒性，以及对数据的需求量大。本文提出了一种新方法，将基于图神经网络（GNN）的神经网络架构、基于有限元法（FEM）的数值求解器和主动学习流程相结合，以应对其中的一些局限。本文考虑一种不依赖优化过程的数据同化方案，并以雷诺平均纳维–斯托克斯（RANS）方程作为基准方程。图神经网络模型经过训练后用作代理模型，以预测通过监督学习得到的强迫项/闭合项，而给定的平均流则作为输入。与文献中的其他方法相比，图神经网络架构尤其适用于本研究，因为它能够适应非结构网格，并具有较强的泛化能力。此外，与所需数据量更大的替代架构相比，该架构能够在低数据量极限下以较少的数据完成训练。

我们开发了有限元法与图神经网络环境之间的双向接口，用于将FEM矢量场转换为神经网络结构能够处理的数值张量，反之亦然，并在整个过程中保留关键信息。作为测试平台，我们重点研究了低雷诺数下绕障碍物的二维不可压缩流，具体范围为 $50 \leq Re \leq 150$。在这些雷诺数范围内，障碍物的存在会触发不稳定性，并在非定常流中发展。我们首先研究了范围为 $50 \leq Re \leq 150$ 的圆柱几何体，将其作为初始基准，以评估基于该训练数据集的模型在多大程度上也可用于未见案例。不出所料，在圆柱案例中，模型对未见雷诺数表现良好。另一方面，当考虑随机几何形状的钝体周围流动时，可以观察到预测精度不足和过拟合现象。

为克服这些局限性，我们从数据的数量和质量两个方面探讨了训练数据对图神经网络泛化能力的影响。首先，我们考虑了一个扩展数据集。结果表明，数据的质量和规模会显著影响模型能够泛化到的未见案例范围。特别是，将多样化的流体流动条件纳入训练数据集，能够提升绝大多数案例的整体泛化能力。最后，我们提出了一种基于梯度相似性分析的主动学习数据选择准则，旨在构建一个能够扩展数据分布的数据集。据作者所知，这是流体力学领域较早的应用之一：通过系统地选择数据，在最大化预测质量的同时，尽量减少训练集中的数据量，从而提升神经网络模型预测未见案例的泛化能力。结果清楚地表明，在保持训练集数据量相当有限的情况下，仍有可能提高模型的性能，包括其泛化能力。当训练代理模型的计算资源有限，需要在预测精度与训练计算成本之间寻求权衡时，该准则尤其具有相关性。需要强调的是，与机器学习文献中出现的其他方法相比，训练过程中使用的数据集相对较小，因为其中规模最大的一个数据集包含的快照对少于20组

圆柱	C_d 平均值	C_l 振幅
Etienne 和 Pelletier	1.36	0.67
Li 等人	1.34	0.69
Liu 等人	1.31	0.69
当前结果	1.32	0.73

表 2： C_d 平均值与 C_l 振幅的比较

快照。这些数据集通过一种能够最小化所需数据点数量的准则进行选择，并且该准则具有足够的稳健性，可以应用于更大的数据集，在保留关键信息的同时高效地对其进行缩减。

我们目前正在通过与同化循环相关的伴随方程，将物理约束纳入学习循环，以此拓展研究成果。通过利用物理约束，这种整合旨在进一步提升图神经网络的预测性能和泛化能力。这些数据同化方案的最终目标是调整训练方法，使其适用于只能获取有限或受污染流场测量数据的情形，例如基于稀疏探针测量、含噪或不完整数据的情形。

致谢 M. Quattromini 的博士研究员奖学金由意大利大学部提供。本研究部分经费来自国家复苏与韧性计划（NRRP）任务 4、组成部分 2、投资 1.3——2022年10月11日第 1561 号招标公告，项目代码 PE0000021，项目名称为 Network 4 Energy Sustainable Transition – NEST，由欧盟 NextGenerationEU 资助。我们谨此感谢法国国家研究署（ANR）通过 ANR-21-REASON 项目资助提供的支持。

附录 A 数值模拟详情

时间分辨数值模拟使用基于 FEniCS [Alnæs 等人 2015]库的 Python 代码进行。图神经网络模型的输入和输出通过在线对这些数据求平均获得。有关数值方案的更多详情，请参阅 Zienkiewicz 等人 [2014]；主脚本基于 Guégan [2022], 中介绍的代码，其中对相关内容进行了深入讨论。此处我们概述核心内容，并详述验证过程。从数值角度来看，空间离散通过引入
基于有限元法（FEM）的弱形式。具体而言，所用有限元为 Taylor-Hood单元，其中速度采用二阶 P2 单元，压力采用一阶 P1 单元。所实现的空间积分方案如下

$$\begin{cases} \left(\dfrac{3\mathbf{u}^n - 4\mathbf{u}^{n-1} + \mathbf{u}^{n-2}}{2\Delta t}\right) + (\mathbf{u}^{n-1}\cdot\nabla)\mathbf{u}^n + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n-1} & \\ \qquad -(\mathbf{u}^{n-1}\cdot\nabla)\mathbf{u}^{n-1} - \dfrac{1}{Re}\Delta\mathbf{u}^n + \nabla p^n & = \mathbf{0} \\ \nabla\cdot\mathbf{u}^n & = 0, \end{cases} \tag{12}$$

其中， $\mathbf{u} = (u,\ v)^T$ 表示速度矢量， p 表示压力， Re 表示雷诺数。时间推进采用二阶后向差分公式（BFD）： n 上标表示某一量在当前时刻的取值， $n-1$ 表示其在前一时间步的取值， $n-2$ 表示其在前两个时间步的取值。时间步长——记为 Δt ——的选取旨在确保针对每个随机生成形状满足 Courant-Friedrichs-Lewy条件， $CFL \leq 0.5$ 。对流项采用Newton和Picard方法处理，从而在时间迭代的每一步求解一个线性方程组。基础网格基于我们作为参考案例的圆柱绕流所采用的网格；在障碍物周围最敏感的区域以及圆柱后方的尾流区域对网格进行加密[Giannetti 和 Luchini， 2007]：这些区域需要较高的网格加密度，以提高计算在不稳定性的正确频率和增长率方面的精度。除众所周知的动力学特征外，丰富的文献也有助于为验证目的进行详细比较。因此，采用圆柱绕流对数值求解器和网格进行了验证，尤其是通过比较阻力系数 C_d 和升力系数 C_l。图14展示了 C_d 和 C_l随时间的演化，而表2则给出了（时间平均） C_d 和C_l（振幅）与文献数值之间的数值比较。第一个系数与文献中的数值符合得非常好，而发现 C_l 略高：这可能与所选数值计算域在 y 方向上略小有关，因而产生了一些数值阻塞效应；尽管如此，就本研究的性质而言，结果令人满意。最后，所选钝体的特征机制与圆柱绕流中观察到的现象相似，因此具有可比的时间和空间尺度；相应网格的加密程度也与参考案例所采用的网格相近。

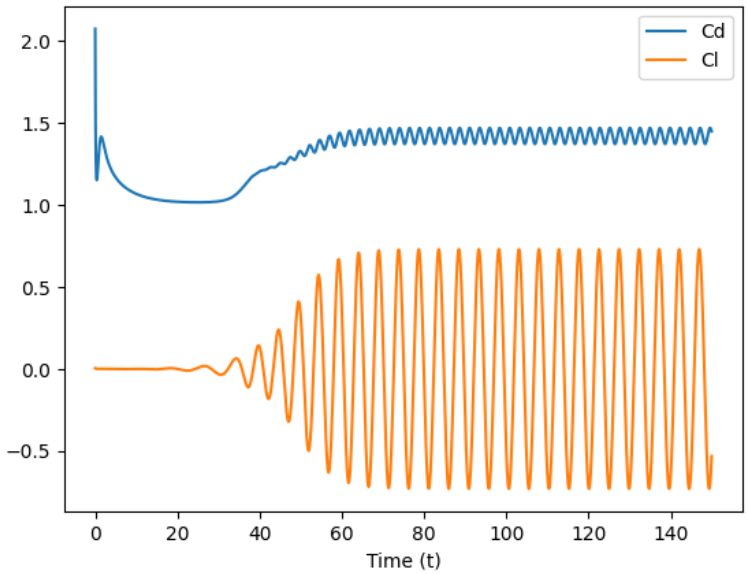

图14：圆柱绕流的阻力系数 C_d 和升力系数 C_l。

附录B 超参数优化

众多参数 定义神经网络的结构，通常用术语超参数来表示。我们可以
区分 m odel 超参数和过程超参数。

- 模型超参数决定了神经网络的容量，即 模型表示高复杂度函数的能力。因此，模型容量与逼近多种非线性函数的可能性直接相关。

- 过程超参数决定了训练阶段。调整这些超参数会显著改变训练时长、计算成本，以及模型演化过程中权重的调整方式。

这些术语在训练模型之前预先定义，因此需要手动调整，因为它们无法在训练过程中调整或学习。通常，在深度学习中这不会造成问题，因为只要神经网络的表达能力（例如 神经元数量、层数等）足以表示所研究问题的复杂性即可。另一方面，由于图神经网络经过训练是为了完成特定任务，而且其推断的计算成本必须相较于当今最先进的湍流模型仍可承受，因此我们尽可能使图神经网络保持简约。

为此，需要对定义架构的超参数进行优化。在处理神经元数量或层数等整数时，无法使用标准的基于梯度的优化器。为此，可以使用无梯度算法。现有专门的库能够通过尝试所有可能的超参数组合，并适当剪枝其中表现不佳的组合，自动完成调优过程。在本文中，我们采用了 Optuna [Akiba等人， 2019], 这是一个开源软件包，能够高效地结合搜索与剪枝算法。通过探索复杂的解超空间，可以找到若干超参数组合，其中在所监测的验证指标方面优于其他组合的组合由以下参数集给出

1. 嵌入维度，35 2. 图神经网络层数， $k = 403$. 更新松弛权重， $\alpha = 6 \times 10^{-1}4$. 损失函数权重， $\gamma = 0.15$. 学习率， $LR = 3 \times 10^{-3}$，作为最大值/初始值。

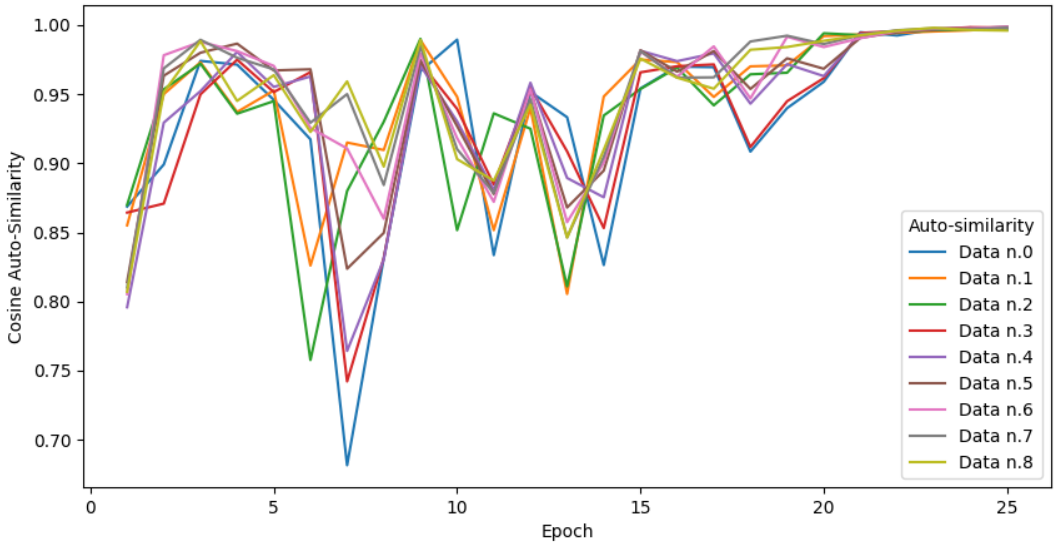

图 15：使用平均绝对误差（MAE）对训练数据集中的 9 个不同案例进行多个训练轮次的梯度自相似性收敛情况可视化。

附录C 相似性判据算法详情

旨在从神经网络的角度比较不同数据的相似性判据算法，基于分析度量函数（式 14）相对于神经网络 θ 参数的向量梯度。具体而言，两个通用的 m 维向量$\mathbf{a} \in \mathbb{R}^m$ 和$\mathbf{b} \in \mathbb{R}^m$ 之间的相似性比较通过余弦相似度计算，其定义如下

$$\cos(\beta) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}, \tag{13}$$

其中，β 是两个向量之间的夹角。在此情境下，我们使用的度量指标是平均绝对误差（MAE），它是一个分段线性函数，定义如下：

$$MAE = \sum_{i=1}^{n_i} |x_i - y_i|, \tag{14}$$

其中，x_i 是节点 i上的神经网络预测，y_i 是真实值，n_i 是节点数 .

平均绝对误差（MAE）的选择对我们的分析至关重要。在使用均方误差（MSE）的情形下，我们通过实证观察到，向量梯度自相似性的方向在连续训练轮次中会出现明显振荡。相反，使用 MAE 时，向量梯度的自相似性往往趋近于 1，这表明对于所分析的数据而言，其在解空间中的方向是稳定的。图 15 展示了向量梯度自相似性随训练轮次变化而持续收敛的情况，训练过程中包含训练数据集中的 9 个案例。训练从一个特定的初始数据集开始；当训练数据集中的每个数据点的自相似性收敛程度都超过预设阈值 0。99 时，训练便会停止，此时我们可以认为训练数据集中每个实例的向量梯度方向都已稳定。

下一步是评估训练数据集内的样本与数据集外样本之间的相似度。此举旨在找出未纳入训练集的样本中多样性最高的样本，并将其添加到训练数据集中，以提高多样性。首先，图神经网络额外运行 10 个训练轮次，以获取每个训练数据集外样本的向量梯度。然后，通过交叉计算每个训练集内样本与每个训练集外样本之间的相似度，得出相似度矩阵。相似度分数最低的样本也是多样性最高的样本，可利用现有数据提高训练数据集的多样性。请注意，这些值使用 z分数进行归一化

$$z = \frac{S - \mu}{\sigma}, \tag{15}$$

其中， S 表示特定数据点的相似度分数， μ 是所有相似度分数的均值， σ 是这些分数的标准差。总之，此处概述的方法利用神经网络模型对数据行为的相似性进行比较和评估，是一种稳健的方法，从而提高训练过程的有效性。

致谢

M. Quattromini 的博士研究员奖学金由意大利大学部资助。本研究部分由国家复苏与韧性计划（NRRP）任务 4、组成部分 2、投资 1.3——2022年10月11日第 1561 号招标公告资助，项目代码 PE0000021，项目名称为 Network 4 Energy Sustainable Transition（NEST），资金来自欧盟 NextGenerationEU。谨此感谢法国国家研究署（ANR）通过 ANR-21-REASON 项目资助提供的支持。

参考文献

T. Akiba、S. Sano、T. Yanase、T. Ohta 和 M. Koyama。Optuna：下一代超参数优化框架。载于第*25*届计算机协会国际知识发现与数据挖掘会议论文集，2019年。M. Alnæs、J. Blechta、J. Hake、A. Johansson、B. Kehlet、A. Logg、C. Richardson、J. Ring、M. E Rognes 和 G. N Wells。FEniCS 项目 1.5版。数值软件档案，3(100)，2015年。R. Anatol。NACA报告1191。1958年。H Jane Bae 和 Petros Koumoutsakos。用于湍流壁面模型的科学多智能体强化学习。《自然·通讯》，13(1):1443，2022年。Andrea Beck 和 Marius Kurz。机器学习方法在湍流建模中的应用展望。*GAMM-* 通讯，44(1):e202100002，2021年。Yoshua Bengio、Jérôme Louradour、Ronan Collobert 和 Jason Weston。课程学习。载于第*26*届国际机器学习年会论文集，第41–48页，2009年。S. L. Brunton、B. R. Noack 和 P. Koumoutsakos。用于流体力学的机器学习。《流体力学年度评论》，52:477–508，2020a。Steven L Brunton、Maziar S Hemati 和 Kunihiko Taira。流体动力学中机器学习与数据驱动方法专题。理论与计算流体动力学，34(4):333–337，2020b。Michele Alessandro Bucci、Onofrio Semeraro、Alexandre Allauzen、Sergio Chibbaro 和 Lionel Mathelin。用于动力系统数据驱动建模的课程学习。《欧洲物理学杂志*E*》，46(3):12，2023年。Jiayi Cai、Pierre-Emmanuel Angeli、Jean-Marc Martinez、Guillaume Damblin 和 Didier Lucor。重新审视用于雷诺应力建模的张量基神经网络：在平面通道流和方形管道流中的应用。《计算机与流体》，第106246页，2024年。René-Daniel Cécora、Rolf Radespiel、Bernhard Eisfeld 和 Axel Probst。用于航空学的微分雷诺应力建模。*AIAA* 期刊，53(3):739–755，2015年。Guillaume Charpiat、Nicolas Girard、Loris Felardos 和 Yuliya Tarabalka。从神经网络视角看输入相似性。载于 H. Wallach、H. Larochelle、A. Beygelzimer、F. d'Alché-Buc、E. Fox 和 R. Garnett（编），神经信息处理系统进展，第32卷。Curran Associates, Inc.，2019年。网址：https://proceedings. neurips.cc/paper_files/paper/2019/file/c61f571dbd2fb949d3fe5ae1608dd48b-Paper.pdf。J Chen、E Hachem 和 J Viquerat。用于预测随机二维形状周围层流的图神经网络。《流体物理学》，33(12):123607，2021年。A. Chiarini、Maurizio Quadrio 和 F. Auteri。流过矩形柱体的定常流的线性稳定性。《流体力学杂志》，929，2021年12月。doi: 10.1017/jfm.2021.819。A. Chiarini、M. Quadrio 和 F. Auteri。对称钝体绕流失稳的新尺度律。《流体力学杂志》，936:R2，2022年。doi: 10.1017/jfm.2022.99。Balthazar Donon、Zhengying Liu、Wenzhuo Liu、Isabelle Guyon、Antoine Marot 和 Marc Schoenauer。深度统计求解器。神经信息处理系统进展，33:7910–7921，2020年。Dorian Dupuy、Nicolas Odier、Corentin Lapeyre 和 Dimitrios Papadogiannis。使用图神经网络在大涡模拟中建模壁面剪切应力。《数据中心工程》，4:e7，2023年。

K. Duraisamy、G. Iaccarino 和 H. Xiao。数据时代的湍流建模。《流体力学年度评论》，51:357–377，2019年。H. Eivazi、M. Tahani、P. Schlatter 和 R. Vinuesa。用于求解雷诺平均纳维–斯托克斯方程的物理信息神经网络。《流体物理学》，34(7):075117，2022年。D. P. G. Foures、N. Dovetta、D. Sipp 和 P. J. Schmid。一种用于雷诺平均纳维–斯托克斯方程驱动的平均流重构的数据同化方法。《流体力学杂志》，759:404–431，2014年。Paul Garnier、Jonathan Viquerat、Jean Rabault、Aurélien Larcher、Alexander Kuhnle 和 Elie Hachem。流体力学深度强化学习综述。《计算机与流体》，225:104973，2021年。F. Giannetti 和 P. Luchini。圆柱尾流首次失稳的结构敏感性。《流体力学杂志》，581:167–197，2007年。I. Goodfellow、Y. Bengio 和 A. Courville。深度学习。麻省理工学院，2016年。Thibaut Guégan。通过机器学习进行湍流控制。普瓦捷大学，2022年。Ingo Gühring、Mones Raslan 和 Gitta Kutyniok。深度神经网络的表达能力。*arXiv* 预印本 *arXiv:2007.04759*，34，2020 年。W. L. Hamilton。图表示学习。人工智能与机器学习综合讲座，14 (3):1–159，2020年。Corentin J Lapeyre、Antony Misdariis、Nicolas Cazard、Denis Veynante 和 Thierry Poinsot。训练卷积神经网络以估计湍流亚网格尺度反应速率。《燃烧与火焰》，203:255–264，2019年。Sangseung Lee 和 Donghyun You。使用深度学习对圆柱周围非定常流进行数据驱动预测。《流体力学杂志》，879:217–254，2019年。J. Ling 和 J. Templeton。用于预测高雷诺平均纳维–斯托克斯不确定性区域的机器学习算法评估。《流体物理学》，27(8):085103，2015年。Julia Ling、Andrew Kurzawski 和 Jeremy Templeton。利用具有嵌入式不变性的深度神经网络进行雷诺平均湍流建模。《流体力学杂志》，807:155–166，2016年。Mario Lino、Stathi Fotiadis、Anil A Bharath 和 Chris D Cantwell。用于流体动力学模拟的当前与新兴深度学习方法。《英国皇家学会会刊*A*辑》，479(2275):20230058，2023年。Vladislav Manti ˇc Lugo、Cristóbal Arratia 和 Fran çois Gallaire。圆柱尾流中涡脱落非线性饱和的自洽平均流描述。《物理评论快报》，113:084501，2014年8月。Ryley McConkey、Eugene Yee 和 Fue-Sang Lien。关于用于预测性湍流建模的机器学习辅助各向异性映射的泛化性。《国际计算流体动力学杂志》，36(7):555–577，2022年。Pierluigi Morra、Onofrio Semeraro、Dan S Henningson 和 Carlo Cossu。关于雷诺应力在湍流壁面约束流预解分析中的相关性。《流体力学杂志》，867:969–984，2019年。Yusuf Patel、Vincent Mons、Olivier Marquet 和 Georgios Rigas。用于平均流重构的湍流模型增强的物理信息神经网络。《物理评论：流体》，9(3):034605，2024年。Ethan Pickering、Georgios Rigas、Oliver T Schmidt、Denis Sipp 和 Tim Colonius。用于湍流射流相干结构基于预解模型的最优涡黏度。《流体力学杂志》，917:A29，2021年。Stephen B Pope。一种更一般的有效黏度假设。《流体力学杂志》，72(2):331–340，1975年。M. Provansal、C. Mathis 和 L. Boyer。贝纳德–冯·卡门不稳定性：瞬态和受迫状态。《流体力学杂志》，182:1–22，1987年。doi: 10.1017/S0022112087002222。Maziar Raissi、Alireza Yazdani 和 George Em Karniadakis。隐式流体力学：从流动可视化中学习速度和压力场。《科学》，367(6481):1026–1030，2020年。Pengzhen Ren、Yun Xiao、Xiaojun Chang、Po-Yao Huang、Zhihui Li、Brij B Gupta、Xiaojiang Chen 和 Xin Wang。深度主动学习综述。计算机协会计算综述，54(9):1–40，2021年。A. Sanchez-Gonzalez、N. Heess、J. T. Springenberg、J. Merel、M. Riedmiller、R. Hadsell 和 P. Battaglia。作为可学习物理引擎、用于推断和控制的图网络。载于国际机器学习会议，第4470–4479页，2018年。M. Schmelzer、R. P. Dwight 和 P. Cinnella。利用稀疏符号回归发现代数雷诺应力模型。《流动、湍流与燃烧》，104:579–603，2020年。K. Shukla、M. Xu、N. Trask 和 G. E. Karniadakis。用于物理信息神经网络与图网络的可扩展算法。《数据中心工程》，3，2022年。

A. P. Singh 和 K. Duraisamy。利用场反演量化湍流闭合中的函数误差。《流体物理学》，28(4):045110，2016年。DENIS SIPP 和 ANTON LEBEDEV。基流与平均流的全局稳定性：一种通用方法及其在圆柱绕流和开放腔流中的应用。《流体力学杂志》，593:333–358，2007年。doi: 10.1017/S0022112007008907。C. A. Ströfer 和 H. Xiao。基于间接观测对湍流模型进行端到端可微学习。理论与应用力学快报，11(4):100280，2021年。Nils Thuerey、Konstantin Weißenow、Lukas Prantl 和 Xiangyu Hu。用于翼型绕流雷诺平均纳维–斯托克斯模拟的深度学习方法。*AIAA* 期刊，58(1):25–36，2020年。Artur P Toshev、Gianluca Galletti、Johannes Brandstetter、Stefan Adami 和 Nikolaus A Adams。利用 E (3) 等变图神经网络学习拉格朗日流体力学。载于几何信息科学国际会议，第 332–341 页。施普林格，2023年。Ricardo Vinuesa 和 Steven L Brunton。利用机器学习增强计算流体力学。《自然·计算科学》，2(6):358–366，2022年。Pedro Stefanin Volpiani、Morten Meyer、Lucas Franceschini、Julien Dandois、Florent Renac、Emeric Martin、Olivier Marquet 和 Denis Sipp。机器学习增强的湍流建模：用于大范围分离流的 RANS 模拟。《物理评论：流体》，6(6):064607，2021年。Jakob G von Saldern、Johann M Reumschüssel、Thomas L Kaiser、Oliver T Schmidt、Peter Jordan 和 Kilian Oberleithner。线性化平均场方法的自洽闭合建模。载于*AIAA AVIATION 2023* 论坛，第 4351 页，2023年。Stefan Wallin 和 Arne V Johansson。用于不可压缩和可压缩湍流的显式代数雷诺应力模型。《流体力学杂志》，403:89–132，2000年。Jack Weatheritt 和 Richard Sandberg。将一种新型进化算法应用于雷诺平均纳维–斯托克斯方程应力–应变关系的代数修正。计算物理学杂志，325:22–37，2016年。David C Wilcox 等。计算流体力学湍流建模，第 2 卷。DCW industries，拉加拿大，加利福尼亚州，1998年。Sunwoong Yang、Ricardo Vinuesa 和 Namwoo Kang。改进图 U-Net，实现网格无关时空流动预测，2024年。Yaomin Zhao、Harshal D Akolekar、Jack Weatheritt、Vittorio Michelassi 和 Richard D Sandberg。利用计算流体力学驱动的机器学习开发 RANS 湍流模型。计算物理学杂志，411:109413，2020年。Zhideng Zhou、Guowei He 和 Xiaolei Yang。基于神经网络的壁面模型，用于周期性山丘上湍流的大涡模拟。《物理评论：流体》，6(5):054610，2021年。O.C. Zienkiewicz、R.L. Taylor 和 P. Nithiarasu，编。流体动力学有限元法。巴特沃思-海涅曼，牛津，第七版，2014年。ISBN 978-1-85617-635-4。doi: https://doi. org/10.1016/B978-1-85617-635-4.00016-9。URL https://www.sciencedirect.com/science/article/pii/ B9781856176354000169。