

基于 PINN 的咳嗽飞沫分布预测

1. 背景介绍

在呼吸活动中产生的含有病原体的呼吸飞沫可导致多种呼吸道传染病，如 2003 年的严重急性呼吸综合征 (SARS)、2009 年的 H1N1 流感大流行，以及 2019 年的新型冠状病毒病 (COVID-19)。以 COVID-19 为例，截至 2024 年 5 月 5 日，世界卫生组织 (WHO) 已报告 7.75 亿例 COVID-19 确诊病例，其中包括 704 万例死亡病例。因此，预防和控制呼吸道传染病的需求迫在眉睫。空气传播作为呼吸道传染病的主要传播途径之一，引起了广泛关注，并对流行病的传播和预防产生了重大影响。研究人员和公共卫生专家强调携带病毒的微小空气飞沫的重要性。流行病学研究发现，与其他途径相比，通过感染者咳嗽时排出的飞沫传播传染病的风险更大。通过科学地模拟和预测室内空间中飞沫的分布情况，可以为公共卫生战略和个人防护提供重要见解，从而实施有针对性的预防和控制措施，遏制流行病的传播并降低感染风险。

咳嗽飞沫分布预测以 CFD 仿真结果作为基础进行模型学习与预测，读取粒子发射的初始位置、粒子直径、粒子时间、速度及温度等参数，考虑蒸发效应和湍流扩散效应，对咳嗽飞沫进行了预测。

2. 模型描述

PINN 算法是一种利用神经网络求解偏微分方程的方法。其基本思想是将神经网络作为偏微分方程解的代理模型，将偏微分方程的信息编码到神经网络的损失函数中进行训练。假设存在如下一般式的偏

微分方程:

$$\frac{\partial u}{\partial t} + F(u) = 0, x \in \Omega, t \in [0, T]$$

其中, $u(x,t)$ 代表了方程的隐藏解, $F[\cdot]$ 是一个非线性微分算子。当无法求得解析解时,在传统数值方法中,可以采用有限元或者有限差分等方法对方程进行求解。而 PINN 则是使用神经网络来逼近方程的解。假设存在一个神经网络,其输入为 x 和 t ,通过链式求导法则,网络的输出为 $u_N(t,x)$,则关于方程部分的损失可以定义为如下形式:

$$r(t, x) = u_N(t, x) + F[u_N(t, x)]$$

再加上边界条件和初始条件可以将网络的总损失表示为

$$Loss = Loss_{data} + Loss_r$$

一般的物理信息神经网络结构如图 1 所示,其结构包括了神经网络和含有物理信息的损失函数,二者相互耦合,指导神经网络中的参数优化。物理信息神经网络的结构同时可以结合监督学习和无监督学习的优势,而且在网络结构中隐含了物理偏微分方程的约束,这样的结构保证了在损失下降的过程中,同时兼顾了数据信息和物理信息,因此网络的输出可以逐步向方程的隐藏解逐渐逼近。在输入了所需要的时空数据之后,首先使用一个全连接神经网络来逼近需要解函数,然后利用自动微分技术求出残差放入损失函数,该函数通常由多个项组成,并由不同系数加权,最后使用梯度下降等优化算法获得全连接神经网络的权重参数,最终训练出可以满足偏微分方程的神经网络参数以逼近解函数。由于偏微分方程在定义域和值域上对应的映射关系有无限组,因此在训练时不可能选取整个域来训练神经网络,之内选择一部分离

散的映射关系,也就是训练点来指导神经网络来完成训练。对训练点的选取方式,会影响到网络的泛化能力,应该选择尽可能的包括定义域和值域之间映射关系的所有特征。

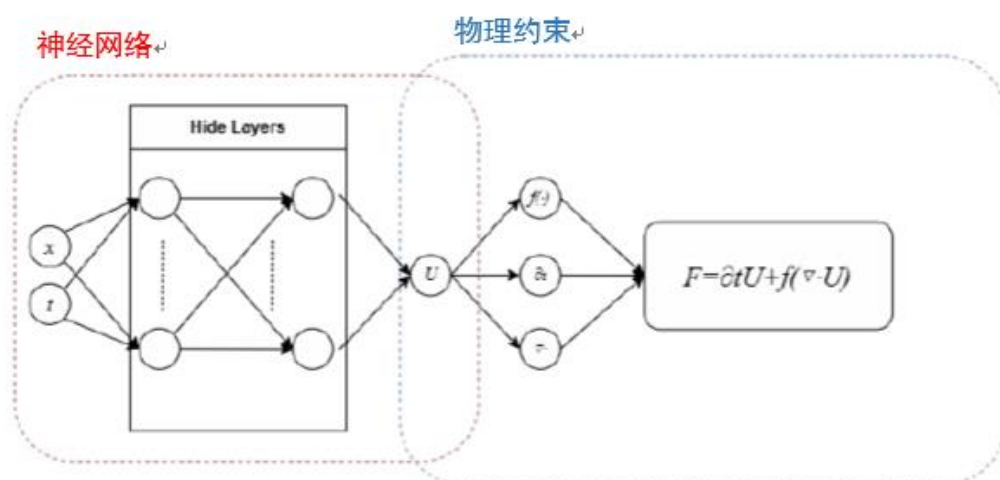


图 1 一般物理信息神经网络结构

3. 训练过程

以部分训练过程为例：

```
Epoch 1/100
22938/22938 [=====] - 650s 28ms/step - loss:
0.2369 - val_loss: 0.4507
Epoch 2/100
22938/22938 [=====] - 644s 28ms/step - loss:
0.2364 - val_loss: 0.4540
Epoch 3/100
22938/22938 [=====] - 635s 28ms/step - loss:
0.2367 - val_loss: 0.4482
Epoch 4/100
22938/22938 [=====] - 634s 28ms/step - loss:
0.2364 - val_loss: 0.4508
Epoch 5/100
22938/22938 [=====] - 633s 28ms/step - loss:
0.2364 - val_loss: 0.4540
Epoch 6/100
22938/22938 [=====] - 634s 28ms/step - loss:
0.2363 - val_loss: 0.4566
Epoch 7/100
```

```

22938/22938 [=====] - 633s 28ms/step - loss:
0.2358 - val_loss: 0.4562
Epoch 8/100
22938/22938 [=====] - 630s 27ms/step - loss:
0.2361 - val_loss: 0.4537
Epoch 9/100
22938/22938 [=====] - 632s 28ms/step - loss:
0.2358 - val_loss: 0.4542
Epoch 10/100
22938/22938 [=====] - 633s 28ms/step - loss:
0.2362 - val_loss: 0.4511
Epoch 11/100
22938/22938 [=====] - 636s 28ms/step - loss:
0.2356 - val_loss: 0.4548
Epoch 12/100
22938/22938 [=====] - 637s 28ms/step - loss:
0.2351 - val_loss: 0.4509
Epoch 13/100
22938/22938 [=====] - 634s 28ms/step - loss:
0.2354 - val_loss: 0.4569
Epoch 14/100
22938/22938 [=====] - 634s 28ms/step - loss:
0.2355 - val_loss: 0.4551
Epoch 15/100
22938/22938 [=====] - 634s 28ms/step - loss:
0.2354 - val_loss: 0.4519
Epoch 16/100
22938/22938 [=====] - 632s 28ms/step - loss:
0.2350 - val_loss: 0.4561
Epoch 17/100
22938/22938 [=====] - 639s 28ms/step - loss:
0.2351 - val_loss: 0.4534
Epoch 18/100
22938/22938 [=====] - 633s 28ms/step - loss:
0.2348 - val_loss: 0.4538
Epoch 19/100
22938/22938 [=====] - 629s 27ms/step - loss:
0.2348 - val

```

最终训练全部结束后的损失记录如下图 2。

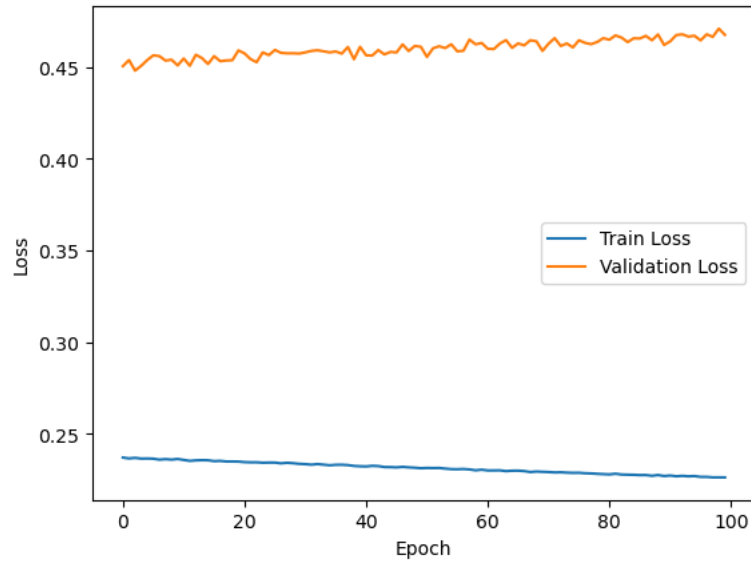


图 2 预测模型训练损失

4. 预测结果

该任务首先对数据进行了归一化处理，随后按时间顺序划分训练集、验证集和测试集。构建 PINN 模型后定义了一系列物理约束的损失函数，为 PINN 模型加载之前训练好的模型权重，给 X 和 Y 方向上的误差赋予更高的权重，使用较小的学习率进行微调，随后对模型持续进行训练。最后对咳嗽飞沫进行预测，更新颗粒直径，考虑蒸发效应；更新颗粒随机游走，考虑湍流扩散效应，计算未来位置并加入随机游走和湍流扩散，最后输出可视化预测结果与实际值的对比。

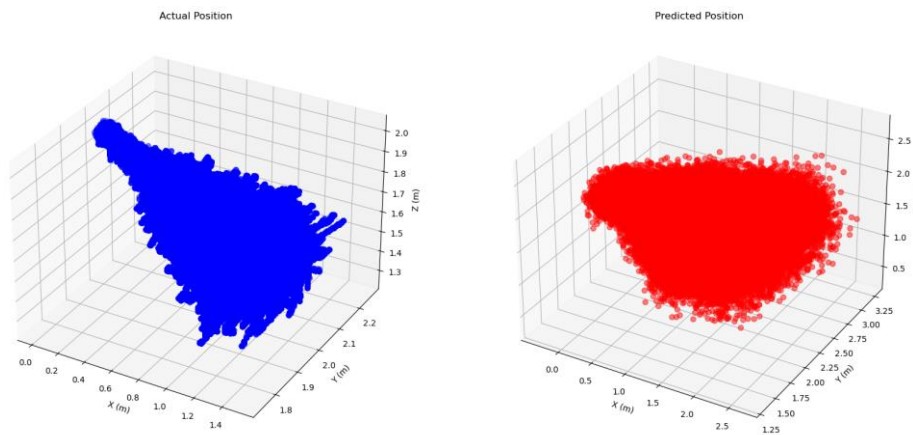


图 3 预测结果与实际值对比