arXiv:2304.02104v2 [physics.comp-ph] 6 Jun 2023

# 多孔介质中的扩散深度学习

**Krzysztof M. Graczyk**[1], *、**Dawid Strzelczyk**[1],和 **Maciej Matyka**[1]

[1]弗罗茨瓦夫大学物理与天文学院理论物理研究所，pl. M. Borna 9, 50-204，波兰弗罗茨瓦夫*

krzysztof.graczyk@uwr.edu.pl

## 摘要

我们采用卷积神经网络（CNN）预测多孔介质的基本性质。研究了两种不同类型的介质：一种模拟砂堆积体，另一种模拟源自生物组织细胞外空间的系统。我们使用格子玻尔兹曼方法生成开展监督学习所需的标注数据。我们区分两项任务。第一项任务中，基于系统几何结构进行分析的网络用于预测孔隙率和有效扩散系数。第二项任务中，网络用于重建浓度图。在第一项任务中，我们提出两种卷积神经网络模型：C-Net 和 U-Net 的编码器部分。两种网络均通过添加自归一化模块进行了改进 [Graczyk 等，Sci Rep 12, 10583 (2022)]。这些模型的预测精度尚可，但仅适用于其训练所用的数据类型。例如，在类似砂堆积体的样本上训练的模型，在类似生物组织的样本上会高估或低估预测值。在第二项任务中，我们提出采用 U-Net 架构。该架构能够准确重建浓度场。与第一项任务不同，在一种数据类型上训练的网络在另一种数据类型上也表现良好。例如，在类似砂堆积体的样本上训练的模型，在类似生物组织的样本上也能完美运行。最后，对于这两种类型的数据，我们拟合阿奇定律中的指数，以求得用于描述有效扩散随孔隙率变化关系的曲折度。

## 1 引言

复杂多孔结构中的扩散传输在自然界中无处不在[1–4]。一个典型例子是大脑，在那里，扩散是营养物质和信号传输的主要过程[5–7]。大脑的细胞外空间填充着脑细胞外间隙，位于神经毡细胞之间，可视为具有复杂拓扑结构的多孔介质。研究多孔介质中扩散有效性质的方法多种多样。应用类型决定了方法的选择[8]。应用领域涵盖多个不同区域，例如在脑研究中评估细胞外空间的曲折度[5,9]，分析锂离子电池中的扩散[10]，砂质沉积物[2] 以及分级多孔材料研究[11]。

多孔材料空隙空间中的输运性质在很大程度上取决于其孔隙极为多样的几何结构。砂岩由颗粒构成，其粒径约为数十至数百 $\mu$米，但其孔隙可能小数倍；孔隙率定义为孔隙体积与总体积之比：

$$\varphi = \frac{V_{\text{pore}}}{V}, \tag{1}$$

孔隙率可低至百分之几[12]。另一方面，生物体组织中的细胞外空间由宽度约为数十纳米的片层和隧道构成，其中有许多死端，细胞彼此缠绕[5,13,14]。因此，人们一直致力于开发工具，以预测几何结构迥异的多孔介质中的扩散输运性质。本文采用卷积神经网络研究两种截然不同类型的多孔介质：各向同性颗粒介质，例如从砂堆积体到岩石的结构；以及由通道组成、几何结构类似于脑细胞外空间的系统。

扩散输运研究涉及微观和宏观尺度。在微观尺度下，会对发生输运的孔隙空间进行显式建模。在宏观尺度下，则对多孔介质对输运分子的影响进行空间平均，并用有效系数表示。扩散传输的建模通常有两种方法。第一种方法是追踪单个粒子（例如使用随机游走）。另一种方法则是考虑通过求解扩散方程得到的浓度场。本文采用第二种方法：求解扩散偏微分方程。我们计算浓度场 $c(\vec{r}, t)$（$\vec{r}$ 表示位置向量，$t$ 表示时间），以研究这一现象。

扩散方程描述扩散物质浓度场在空间和时间上的变化。它可用于模拟两个尺度上的扩散输运。在微观尺度下，对于体扩散系数 $D_0(\vec{r})$（随空间变化），该

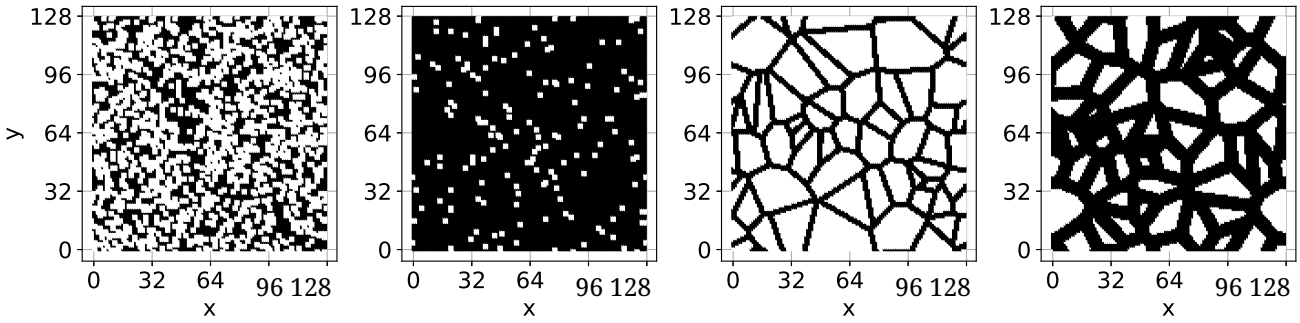


**图 1.** A型样本，具有 $\varphi = 0.42$ 和 $\varphi = 0.92$ （从左至右分别为第一个和第二个），以及 B型样本，具有 $\varphi = 0.32$ 和 $\varphi = 0.66$ （从左至右分别为第三个和第四个）。流体节点以黑色表示，固体节点以白色表示。我们展示了两种数据类型中孔隙率较低和较高的样本。

方程写作：

$$\frac{\partial c}{\partial t} = \nabla \cdot (D_0(\vec{r})\nabla c). \tag{2}$$

对于孔隙体积内部恒定的各向同性扩散率，这可简化为 $\partial c/\partial t = D_0\nabla^2 c$，且扩散通量值仅取决于通量方向上的浓度梯度。然而，在多孔介质中，扩散过程会受到阻碍[15] 或加速[16,17]，这是由于诸如物质与固体障碍物相互作用（位阻效应）或物种在孔隙中特定位置结合等现象所致。在实践中，我们用具有有效扩散系数的连续介质来表示复杂的多孔系统$D(t)$，该系数由菲克定律得到

$$\langle D_0\nabla c(\vec{r},t)\rangle = D(t)\frac{c_{\text{out}} - c_{\text{in}}}{L}, \tag{3}$$

其中，$\langle \ldots \rangle$ 表示对孔隙空间取平均，$L$为样本入口与出口之间的距离，而$c_{\text{in}}$和$c_{\text{out}}$分别为样本入口和出口处的浓度。当扩散时间足够长时，$D(t)$趋近于稳态值$D$，该值通常不同于体扩散系数$D_0$。体扩散系数$D_0$ 与有效扩散系数$D$ 之比用于定义介质的扩散曲折度：

$$\lambda = \frac{D_0}{D}. \tag{4}$$

上述量出现在适用于宏观尺度多孔介质的标量输运方程中（例如，参见 Sykova 和 Nicholson[5] 的方程 9）。多孔样品的扩散曲折度与孔隙率之间的关系仍存在争议，实验研究与理论研究之间尚有待弥合的差距[2]。

由于边界条件复杂，在大多数情况下，无法通过解析方法求解方程（2）。数值解存在若干缺点，例如时间步长呈二次方增长，或矩阵运算计算开销巨大。当需要针对某一类型的多孔介质的大量数据求解扩散方程以进行统计分析时，如 Kasthuri等人[18]，上述问题尤其难以克服。

本文采用卷积神经网络（CNN）系统对多孔介质中的扩散输运进行建模。近年来，基于深度学习的技术已用于计算多孔介质的有效性质，包括渗透率，[19,20]。Graczyk 和 Matyka[21]给出了使用卷积神经网络计算渗透率和曲折度的示例。在 Cawte 和 Bazylak[22], 的研究中，研究人员借助包括梯度提升回归、神经网络和支持向量回归在内的机器学习工具，研究了扩散和渗透率。研究人员利用扫描电子显微镜数据[23] 和基于随机场的多孔介质[24]，对有效扩散$D$进行了基于图像的预测。Kamrava 等人[25] 采用了本文也使用的 U-Net 模型，预测多孔膜中压力和速度场的空间分布。Wu 等人[26]利用基于物理信息的卷积神经网络预测了通过沃罗诺伊镶嵌生成的多孔样品的渗透率。Kamrava 等人[27] 使用一个将质量守恒和纳维–斯托克斯方程纳入学习过程的深度学习模型，模拟了流体在多孔介质中的流动。最后，Tahmasebi等人[28]全面综述了机器学习方法在多孔介质和地球科学研究中的应用。

我们的目标是开发一种深度学习（DL）系统，即深度神经网络（DNN），该系统基于系统几何结构的分析，预测孔隙率和有效扩散系数 $D$，或者稳态下的浓度场 $c(\vec{r})$。我们将

这两个目标分开处理。对于第一种情况，我们考虑两种卷积神经网络架构。其中一种改编自我们先前关于预测多孔介质中的孔隙率、渗透率和曲折度的研究[21] 。此外，两种架构都经过改造，加入了 Graczyk 提出的自归一化（SN）模块 等人[29]。改造后的架构比原始架构具有更高的精度。此外，不含 SN 模块的网络往往会对数据过拟合。最后，我们提出采用 U-Net 架构来预测浓度分布[30], ，其准确度更高，适用范围也比仅预测孔隙率和扩散系数的网络更广。

深度学习系统通过监督学习获得。在这种方法中，关键在于使用能够全面表征所研究特征的数据。我们生成两类样本：随机沉积各向同性颗粒构成的系统（模拟砂堆积体），以及由平面 Voronoi 镶嵌生成的随机系统（模拟组织的细胞外空间）。

在实践中，我们考虑具有几何结构配置的系统二维图像。然后，利用格子玻尔兹曼方法（LBM）模拟扩散过程，并获得稳态浓度场。基于此，使用公式 (3)，我们计算有效扩散系数。表示多孔介质几何结构的图像是网络的输入，网络据此预测孔隙率、扩散系数或浓度场。

网络模型通过对数据进行统计分析获得，因此网络对孔隙率和扩散系数的预测存在一定的不确定性。因此，我们利用蒙特卡洛丢弃技术[31] 来估计模型预测的确定性。

Wu等人[32]开展了类似的研究。他们采用 AlexNet 和 ResNet50 卷积神经网络模型来预测扩散系数，并研究了一种数据类型，其孔隙率范围比我们分析的更广。相比之下，我们研究了另外两种多孔介质。此外，我们的神经网络模型采用了不同的架构。其中一种架构曾在我们之前的工作中进行过研究[21]。另一种是 U-Net 架构的编码器部分。我们所有的架构都包含批量归一化层，这对于获得成功的拟合结果至关重要[21]。最后，我们通过添加自归一化模块[29]对原始模型进行了修改。此外，本文还研究了 U-Net 架构在重建浓度场方面的适用性。

本文结构如下：第2节简要介绍数据的准备方法；第3.1节介绍神经网络模型。两类任务的数值结果分别在第3.2节和第3.3节呈现。本文最后在第4节总结。附录A概述了格子玻尔兹曼方法。

## 2 基于格子玻尔兹曼方法的扩散——数据准备

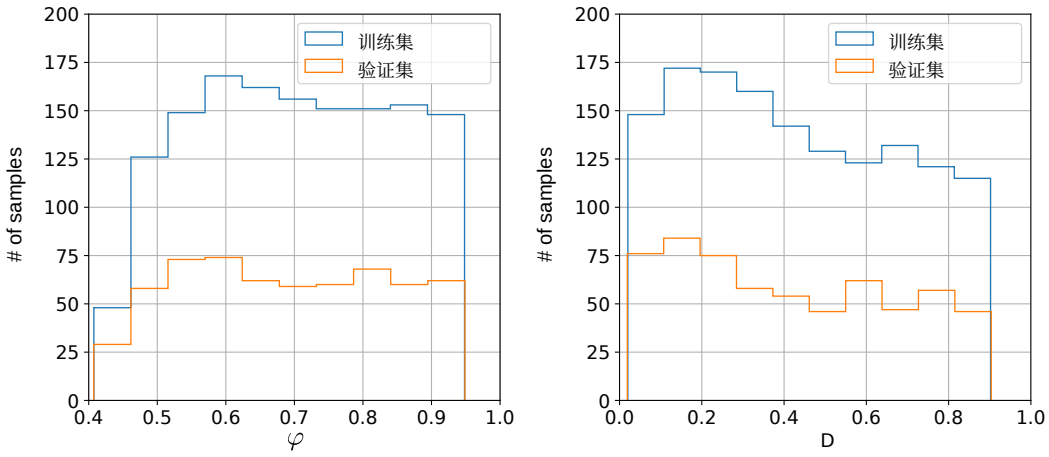


**图 2.** 使用格子玻尔兹曼方法计算得到的2017年A型样本孔隙率 $\varphi$ 和有效扩散系数$D$ 的分布（训练数据集和验证数据集分别为1412个和605个样本）。

我们将构建能够预测扩散现象的深度学习系统。为完成这项任务，我们需要用于训练深度神经网络的标注数据。我们考虑两种随机生成的多孔区域几何结构：

- A型：随机堆积各向同性颗粒构成的系统；

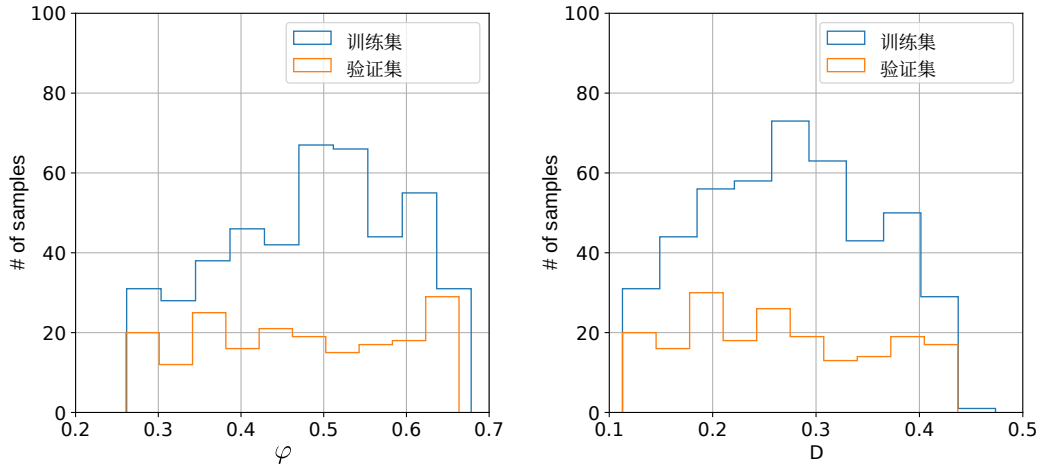


**图3.** 基于格子玻尔兹曼方法对640个B型样本（训练数据集和验证数据集分别包含448个和192个样本）得到的孔隙率 $\varphi$ 和有效扩散系数分布 $D$

- B型：基于沃罗诺伊镶嵌的样本。

每个样本的尺寸为$L \times L$，包含$L = 128$ 个节点。每个节点可以是流体节点或固体节点。所有流体节点都属于 $\mathscr{D}_f$集合，而所有固体节点都属于 $\mathscr{D}_s$集合。位于$L \times L$正方形之外的节点属于 $\mathscr{D}'$集合。

这两类系统的孔隙率均可控，范围从没有障碍物（$\varphi = 1$）到渗流阈值。A型样本是对砂堆积体、黏土或流化颗粒床等颗粒介质的理想化表示。我们通过在初始时完全充满流体的区域内放置 $3 \times 3$ 个方形障碍物来生成样本；见图 1（左侧两幅图）。障碍物位置的坐标从范围 $[0,L)$ 内的均匀分布中抽取，障碍物之间可以重叠。不断向样本构型中填入块，直至达到设定的孔隙率。随后，对左右边界之间进行渗流测试。未通过测试的样本将从数据集中剔除。

B型区域的几何形状源自生物系统，例如组织的细胞外空间；见图1（右侧两幅图）。我们通过在 $L \times L$ 的正方形区域内放置50个点来生成样本。点的位置坐标从区间 $[0,L)$ 上的实数均匀分布中抽取。随后，将该点集复制到八个相邻区域，以尽量减小边界效应。然后，基于抽取的点对 $\mathbb{R}^2$ 空间进行沃罗诺伊镶嵌，从而获得多边形的边，这些边即为孔道的中心线。离散化从全为固体的节点开始，这对应于宽度为零的通道。随后，逐步均匀增大通道宽度，并将位于任一通道内的节点标记为流体节点，见图 1。达到设定的孔隙率后，迭代过程即停止。由于边缘网络的性质，这些样本必定会从左向右渗流，因此不再进行额外测试。

我们假设障碍物及顶部/底部壁面采用反射（法向零通量）边界条件。左侧（$c([0,y],t) = 0 = c_{\text{in}}$）和右侧（$c([L,y],t) = 1 = c_{\text{out}}$）边界施加狄利克雷边界条件。

我们使用格子玻尔兹曼方法（LBM）对扩散方程进行数值求解，以获得有效扩散系数和浓度。该方法的详细信息见附录 A。每次 LBM 模拟都会持续迭代，直至连续两次迭代之间浓度场的最大变化量 $\Delta c_n$ 定义如下：

$$\Delta c_n = \max_{i:\vec{r}_i \in \mathscr{D}_f} \left( \left| c_i^n - c_i^{n-1} \right| \right) \tag{5}$$

小于 $10^{-13}$ 或最多达到 $n = 10^6$ 次迭代。随后，使用方程（$D$ ）的离散实现计算有效扩散系数3），即：

$$D = \frac{L}{N\left(c_{\text{out}} - c_{\text{in}}\right)} \sum_{\vec{r}_i \in \mathscr{D}_f} (\nabla c)|_{\vec{r}_i}, \tag{6}$$

其中，$c_i^n$表示第$i$ 个节点在第$n$次迭代时的浓度值。注意，我们选择格子玻尔兹曼方法的参数，使得$D_0 = 1$。

我们为两项任务生成数据：预测 $\varphi$ 和 $D$ 的几何结构，并根据几何结构预测浓度。对于这两项任务，我们都生成 A型和 B型数据。每个生成的数据集都划分为训练数据集和验证数据集。在图 2 和3中，我们展示了孔隙率值 $\varphi$ 和有效扩散系数 $D$在 2017个 A型样本和 640个 B型样本中的分布。

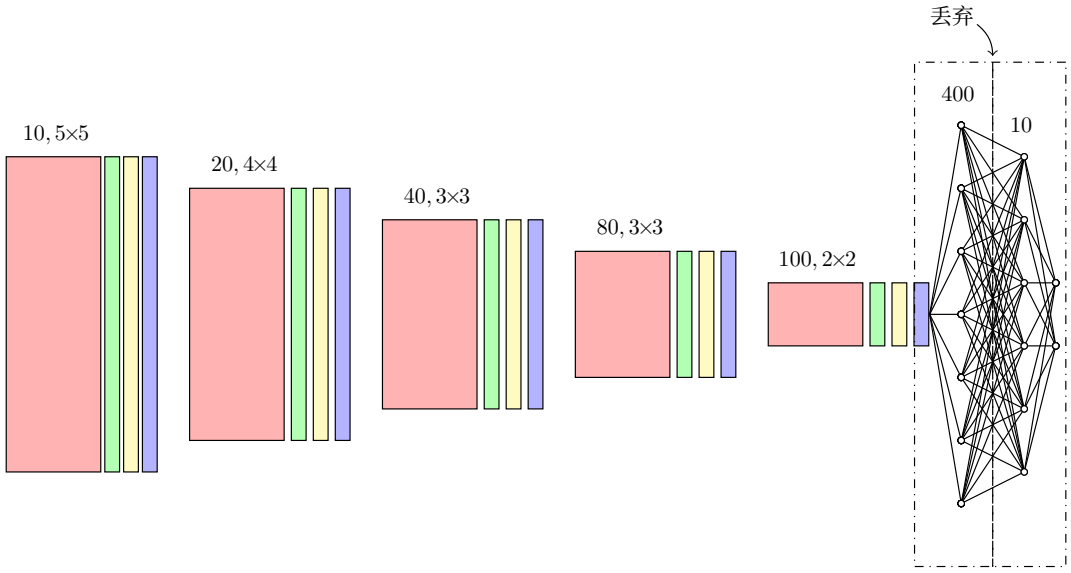


**图4.** C-Net架构。它包含五个卷积部分。每个部分由卷积层（红色，使用 ReLU）、步幅为1且填充为0、批归一化层（绿色）、卷积核大小为 $2 \times 2$（紫色）的最大池化层，以及丢弃层（黄色）组成。每个卷积层上方标注了卷积核的数量和大小。最后两个全连接层（分别含有400和10个隐藏单元）之间设有一个丢弃层。对于含有400个单元的全连接层，激活函数为双曲正切；对于含有10个单元的层，则采用线性激活函数。

## 3 深度学习揭示的扩散现象

在本文中，我们利用深度神经网络构建计算系统，用于预测：

I 孔隙率，$\varphi$，以及有效扩散系数$D$，见公式4；

II 系统中的浓度分布 $c(\vec{r}, t)$。

在这两项任务中，神经网络都以系统几何结构的图像作为输入。在第一项任务中，神经网络预测 $\varphi$、$D$；在另一项任务中，则预测最终的浓度图。

在以下各节中，我们将分别讨论任务I和任务II。任务I的深度神经网络获取流程可概括如下：

(i) 考虑给定类型的数据样本；

(ii) 为每个样本指定孔隙率和扩散系数；

(iii) 使用训练数据集训练深度神经网络，并使用验证数据集进行验证

(iv) 使用其他数据类型测试所得的深度神经网络。

对于任务II，神经网络会重建浓度场。因此，数据包含成对的图像：显示系统几何结构的输入图像，以及显示浓度场的输出图像。

完整分析在 PyTorch[33] 中进行。

### 3.1 深度神经网络模型

对于任务I，我们考虑两种类型的神经网络架构，即：

- C-Net：一种与我们在此前论文[21] 中使用的神经网络结构相似的神经网络，用于预测多孔介质的孔隙率、渗透率和（水力）曲折度。
- U-Net-Half：它是 U-Net 神经网络[30] 的解码器部分，用于重建介质中的浓度分布图。

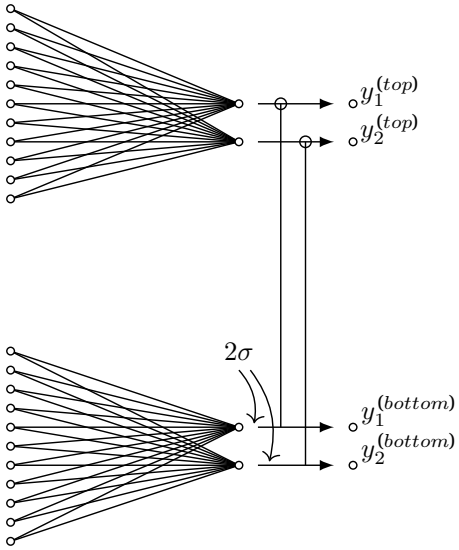


**图 5.** C-Net（或 U-Net-Halfnetworks）的自归一化模块。C-Net（或 U-Net-Half）的最后一个隐藏层（图4中标出的部分）被 SN 模块替换。

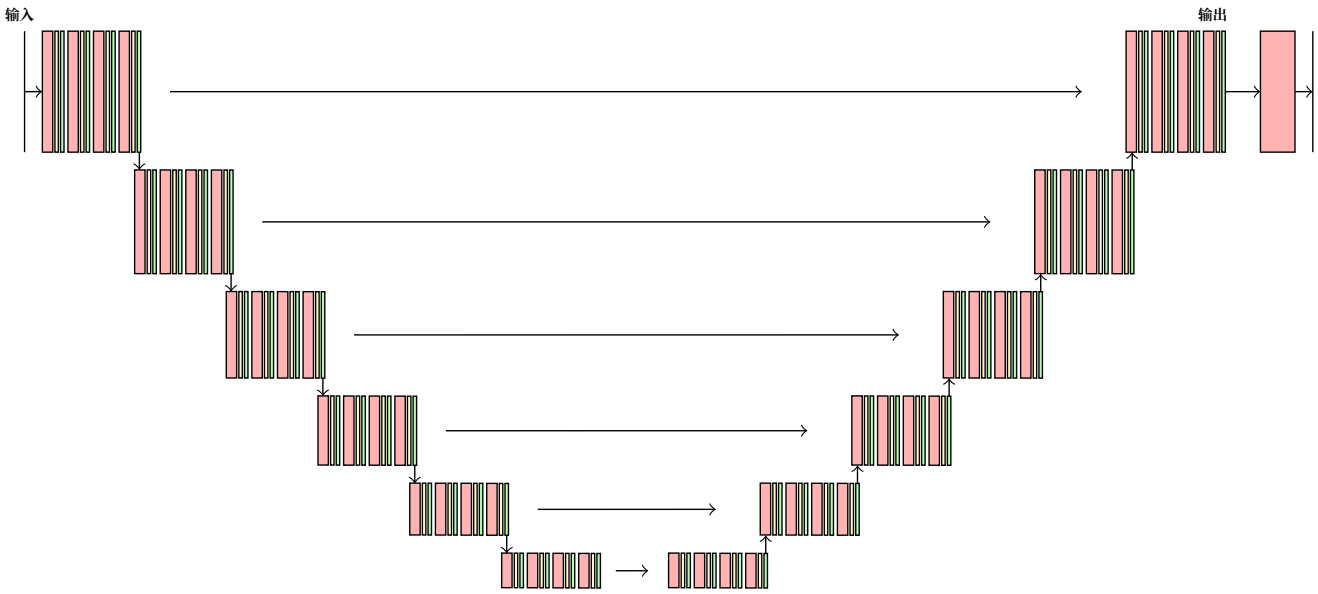


**图 6.** U-Net 架构。每个下采样容器由（$N = 4$）个块组成，每个块包含一个卷积层（使用 ReLU 激活，包含 $3 \times 3$ 个大小为 size 的卷积核）、丢弃层（黄色）和批量归一化层（绿色）。卷积层的填充为 1。每个下采样块之后都有一次最大池化（卷积核大小为 $2 \times 2$），以竖直箭头表示。类似地，每个上采样容器由四个与下采样块结构相同的块组成。经过这四个块后，进行一次上采样（倍率为 2），以竖直箭头表示。请注意，上采样容器连接前一个容器的输出，以及下采样部分中对应块的输出。在神经网络的最后一部分，设有一个下采样容器（不含最大池化）和一个卷积层，用于生成与输入图大小相同的单通道密度图。

C-Net 是一个包含五个块的卷积神经网络；见图 4。神经网络顶部依次连接两个全连接层，分别包含 400 和 10 个隐藏单元，激活函数分别为双曲正切和线性函数。该神经网络以尺寸为 $128 \times 128$、具有相应几何结构配置的输入图像作为输入。其输出为二维向量 $(\varphi, D)$。每个卷积块都包含一个卷积层（使用 ReLU 激活），后接批量归一化、丢弃层和最大池化层（使用 $2 \times 2$ 卷积核）。卷积块之后是两个全连接层。第一个全连接层之后设有一个丢弃层。卷积层分别包含 10、20、40、80 和 100 个卷积核，尺寸依次为 $5 \times 5$、$4 \times 4$、$3 \times 3$、$3 \times 3$ 和 $2 \times 2$。关于 C-Net 结构的更多详情，请参见图 4 的图注。

U-Net-half 是 U-Net（下文将介绍）的编码器部分，后接两个与 C-Net 中结构相同的全连接层。

最后，对于任务I，我们还考虑了另外两种神经网络架构，它们分别由 C-Net 和 U-Net-Half 改造而来：将最后一个隐藏层替换为自归一化（SN）模块。这是 Graczyk 等人[29] 提出的结构，用于校正神经网络的输出。SN 机制的动机源于这样一种观察：通常很容易得到定性预测良好的深度神经网络，但在定量层面上，系统会出现系统性的高估或低估。自归一化模块可校正网络响应，并提高预测精度。该模块是系统的一部分，其参数也需要优化。

SN 模块如图所示5。实际上，使用的是两个大小相近、彼此并行且相互独立的全连接层，而非一个全连接层。顶部全连接层的输出为二维，并与底部全连接层的相应输出相乘，即

$$(y_1^{(top)}, y_2^{(top)}) \to (y_1^{(top)} y_1^{(bottom)}, y_2^{(top)} y_2^{(bottom)}), \tag{7}$$

其中，$(y_1^{(bottom)}$，$y_2^{(bottom)})$ 为底部输出，其中

$$y_i^{(bottom)} = 2\sigma(..), \tag{8}$$

$\sigma$ 为 sigmoid 函数。按照这样的定义，$y^{(bottom)}{}_i$ 起到归一化因子的作用，其取值范围为 0 到 2。原则上，这些因子应接近 1。为了确保神经网络具有适当的归一化，我们在损失函数中加入一个额外的惩罚项，即

$$L(s_i) \to L(s_i) + (1 - y_1^{(bottom)})^2 + (1 - y_2^{(bottom)})^2, \tag{9}$$

其中，$L(s_i)$ 是第 $s_i$ 个样本的损失。对于 $L(...)$，我们的分析采用均方误差（MSE）损失。

对于任务 II，我们建议采用 U-Net 架构，该架构旨在解决生物医学图像分割问题[29,30]。这种神经网络架构曾用于根据膜的形态预测其流动特性[25]，并模拟多孔材料中的流体流动[27]。其结构比 C-Net 神经网络更复杂；见图 6。其结构可分为编码器和解码器两部分。编码器由六个容器依次组成，每个容器将输入下采样两倍。每个容器包含四个块，每个块包括一个卷积层（采用 ReLU 激活）、丢弃和批量归一化层。经过四个块后，最大池化层会对输入进行下采样。解码器包含六个对应的上采样容器，每个容器将输入尺寸放大两倍。上采样容器由四个卷积块（与编码器中的结构相同）以及其后的上采样层组成。具有相同输出和输入尺寸的编码器与解码器容器相连接。最后，在解码器部分之后还有一个由卷积层组成的块（含批归一化层和丢弃层）。图 6 详细描述了 U-Net 的结构。请注意，每个下采样或上采样容器包含 $N = 4$ 个卷积块。在初步分析中，我们从 $N = 1$ 开始，经过多次测试后，得到最优值 $N = 4$。

为了估计网络预测的不确定性，我们采用蒙特卡洛（MC）丢弃法[31,34]。Graczyk等人[29] 采用此方法来估计微生物学数据分析中网络预测的不确定性。简而言之，为了获得网络响应的 $1\sigma$ 不确定性，我们在保持丢弃层处于激活状态的情况下运行神经网络$M = 20$ 次。然后，模型响应取所有网络响应的平均值，而 $1\sigma$ 不确定性则由标准差得出。

### 3.2 基于深度学习的孔隙率和扩散系数

我们的首要任务是构建一个系统，根据系统几何结构的输入图像预测孔隙率和扩散系数。在这里，我们区分了模型开发的两种情形。第一种情形中，神经网络使用 A 型数据进行训练，验证数据集也属于同一类型。随后，我们验证模型在 B 型数据上的表现。第二种情形是指使用 B 型数据集进行训练和验证，并在 A 型数据集上测试的模型。

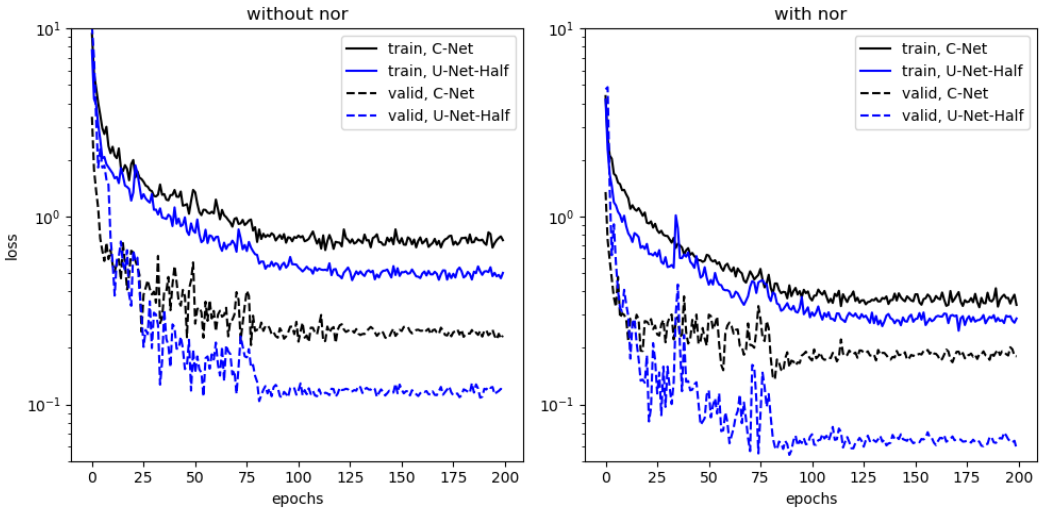


**图 7.** 神经网络训练过程中的均方误差损失。虚线和实线分别表示验证数据集和训练数据集上的误差。蓝色和黑色曲线分别表示 U-Net-Half 和 C-Net 的结果。所有网络均包含丢弃层，且 $p = 0.1$。不使用/使用自归一化模块的模型结果分别显示在左图/右图中。

| | 带 SN 的 C-Net | 带 SN 的 U-Net-Half |
|---|---|---|
| 均方误差（A型数据） | 0.73 | 0.72 |
| 均方误差（B型数据） | 0.33 | 0.29 |

**表 1.** 在蒙特卡罗丢弃框架下运行的模型在验证数据集上计算得到的均方误差。

请注意，损失由均方误差（MSE）给出。在数据增强中，我们利用了这样一个性质：对输入图像进行水平翻转不会改变系统的孔隙率和扩散系数。在图 7中，我们展示了训练过程中损失值的变化情况。这些图展示了第一种情形下所讨论的所有神经网络模型的结果——训练数据集和验证数据集均为 A 型。在本分析中，我们将丢弃层中的 $p = 0.1$ 设为该值。结果表明，这是丢弃机制的最优值。

对图 7 的详细分析表明，添加自归一化模块能够提升网络的性能。事实上，采用 SN模块的模型在验证数据集上计算出的误差始终低于原始网络。另一方面，C-Net 网络计算出的误差始终高于 U-Net-Half 模型，这可以理解为 U-Net-Half 的表现优于 C-Net。为更仔细地研究这一问题，我们引入

$$\overline{\chi}^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{(\varphi^i_{pred} - \varphi^i_{true})^2}{\Delta\varphi^{i2}} + \frac{(D^i_{pred} - D^i_{true})^2}{\Delta D^{i2}}\right], \tag{10}$$

其中，$\varphi^i_{pred/true}$ 和$D^i_{pred/true}$ 分别表示第$i$个样本的孔隙率和扩散系数预测值/真实值。 $\chi^2$ 包含关于不确定性、 $\Delta\varphi$以及 $\Delta D$的信息，这些信息通过蒙特卡洛丢弃获得。

对于在 A型数据集上训练的 C-Net 和 U-Net-Half（均采用 SN 模块）模型，我们得到 $\chi^2 = 1058.4$ 和 $\overline{\chi}^2 = 1465.4$ （在 A型验证数据集上计算）。在这种情况下，U-Net-Half 比 C-Net 更容易对数据过拟合。不出所料，C-Net 神经网络的参数数量远少于 U-Net-Half。事实上，前者的大小约为 300 kB，而后者约为 3。3 MB。因此，采用 SN 模块的 C-Net 是在 A型数据集上预测孔隙率和扩散系数的更优模型。然而，当模型在 B 型数据集上训练时，在验证数据集上计算得到的 $\chi^2$取值为：C-Net 和 U-Net-Half（均采用 SN 模块）分别为 342。7 和 237。0。因此，对于 B 型数据，U-Net-Half 的确定性更高。此外，在 B 型数据集上训练的模型，其预测准确度高于在 A 型数据集上训练的模型。表 1总结了这一特性，其中给出了所分析的两种情形下的均方误差。

让我们更详细地讨论一下拟合结果的质量。在图8中，我们展示了验证数据集的预测结果

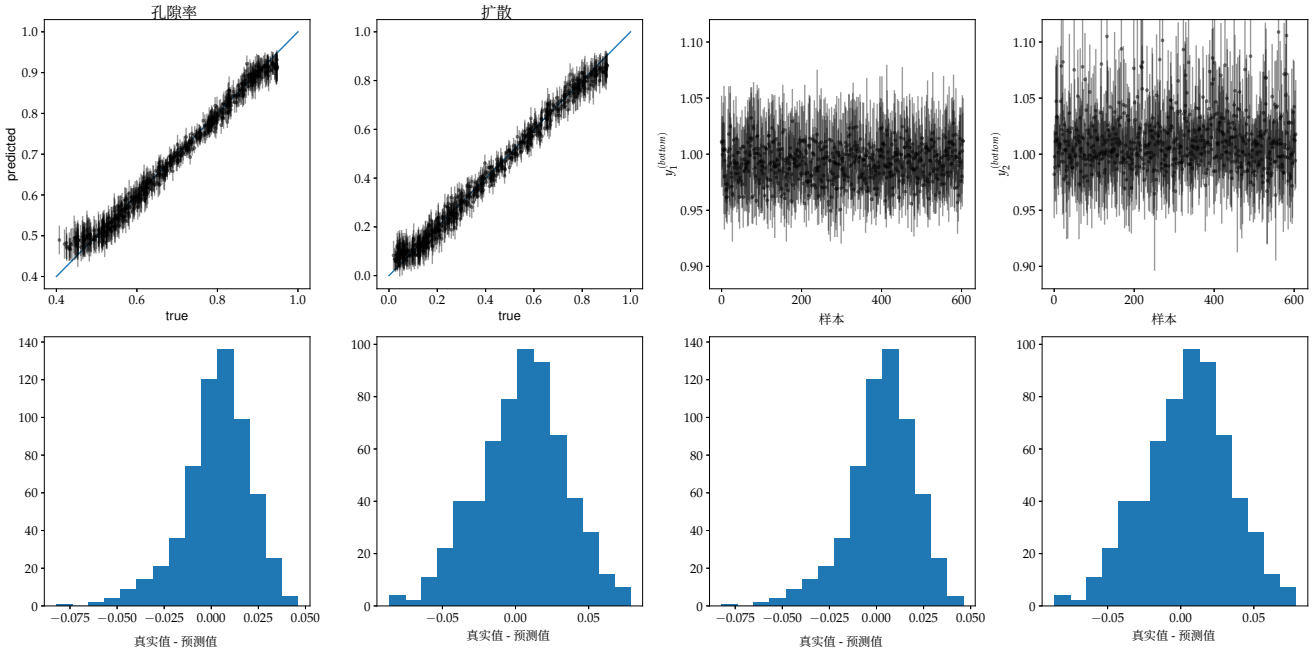


**图8.** 使用 A型数据训练的带 SN模块 的 C-Net。顶行（从左起第一幅和第二幅图）展示了 A型验证数据中孔隙率和扩散系数预测值与真实值的对比。顶行的第三幅和第四幅图展示了 SN模块 的归一化输出，即$y_1^{bottom}$和 $y_2^{bottom}$，参见公式 7。底行给出了相应的真实值与预测值之差直方图。不确定性是通过蒙特卡洛丢弃计算的，丢弃率为$p = 0.1$。

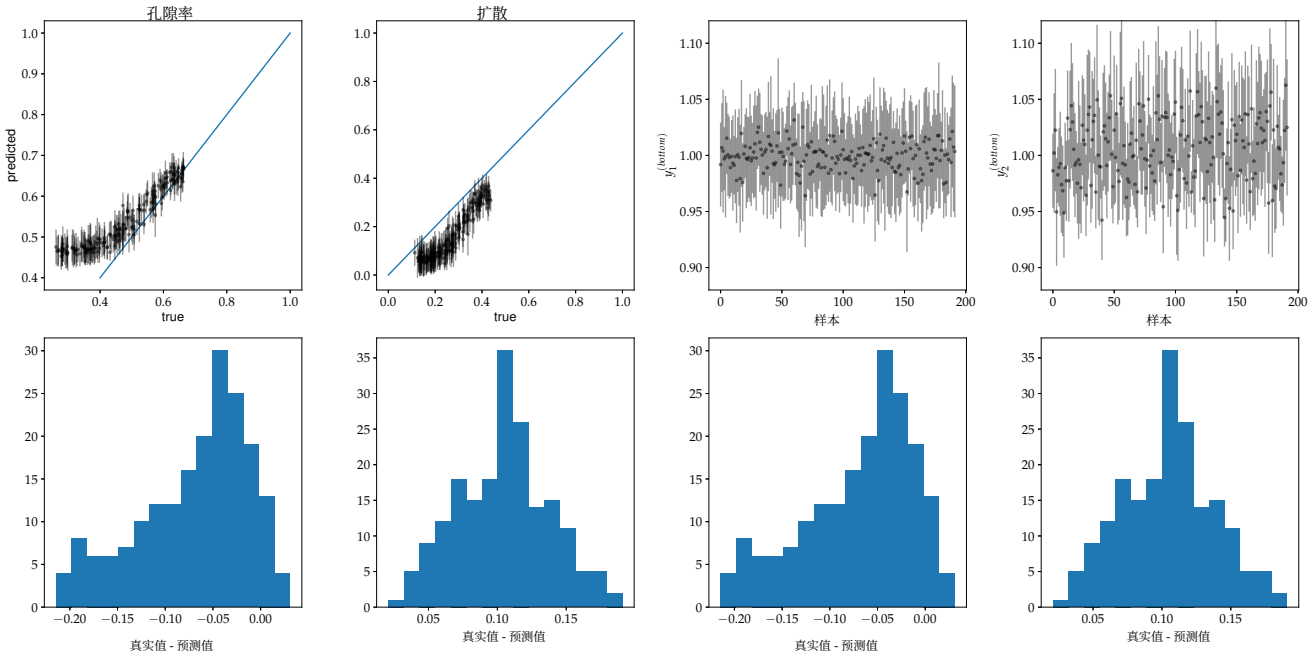


**图9.** 说明与图8 相同，但预测值是针对 B型验证数据得出的。

针对同类型数据训练的 C-Net 对 A 型数据集的预测结果。预测值与真实值($\varphi$，$D$) 吻合良好。但对于 B 型数据集的预测则并非如此，见图9。该模型高估了孔隙率

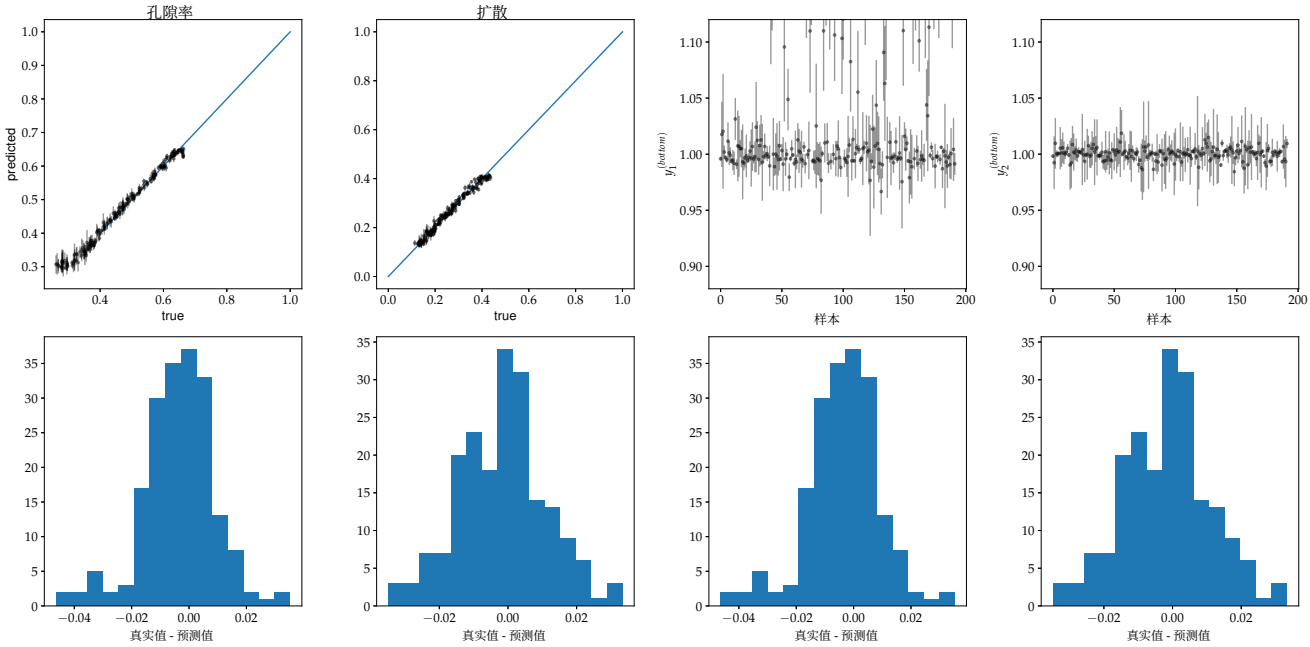


**图 10.** 图注同图 8，但对应于使用 B 型数据集训练的 U-Net-Half（带 SN 模块），以及对 B 型验证数据集所做的预测。

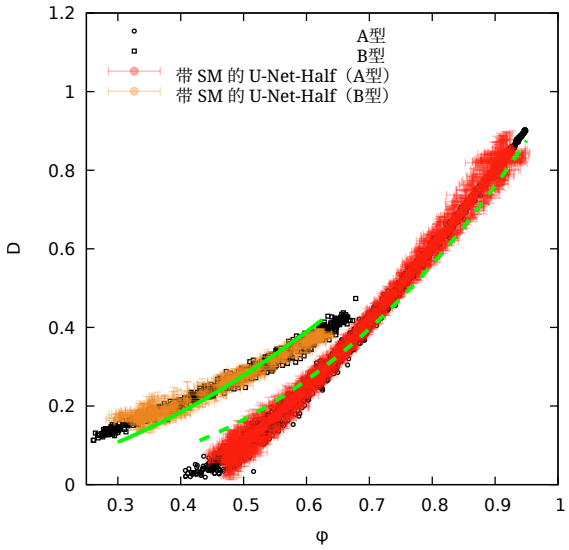


**图 11.** 有效扩散与真实孔隙率的关系。空心符号表示格子玻尔兹曼方法解的数据，带误差棒的实心圆表示两个 U-Net-Half（带 SN 模块）网络的预测结果，这两个网络分别使用 A型或 B 型数据集训练。预测分别针对 A型或 B型验证数据集。绿色线条表示根据 Archie 定律拟合的有效扩散与孔隙率的函数关系，见公式 11。

并在这种情况下低估扩散系数。不过，对于具有 $\varphi > 0.4$ 的样本，神经网络的预测值在两倍标准差范围内保持一致。请注意，A型数据不包含具有 $\varphi < 0.4$ 的样本。事实上，生成孔隙率小于 0.4 的 A型数据颇具挑战，因为这接近离散重叠四边形的渗流阈值，如 Koza 等人[35] 所示。此外，生成 A型样本的算法几乎不可能生成 B型数据的样本。因此，使用 A型数据训练的神经网络无法处理 B型数据也就不足为奇了，尤其是在 A型样本无法覆盖的孔隙率范围内。

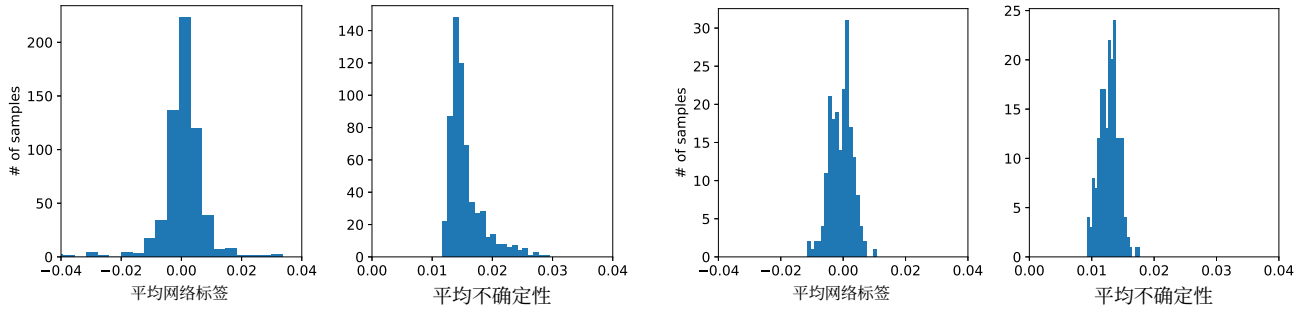


**图 12.** 展示了在 A 型数据上训练的 U-Net 模型对 A 型验证数据集（右侧两幅图）和 B 型验证数据集（左侧两幅图）中浓度的预测结果汇总。

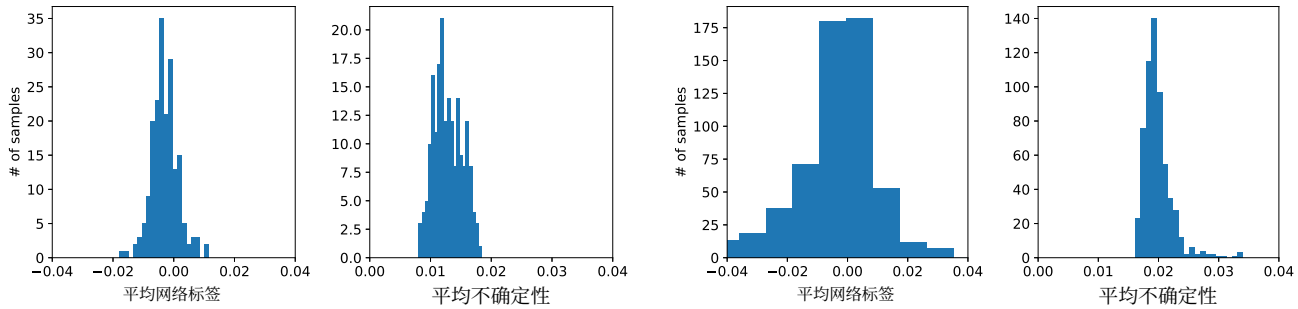


**图 13.** 展示了在 B 型数据上训练的 U-Net 模型对 B 型验证数据集（右侧两幅图）和 A 型验证数据集（左侧两幅图）中浓度的预测结果汇总。

在第二种情形中，模型使用 B型数据进行训练。如上所述，在这种情况下，U-Net-Half 的效果优于 C-Net架构。图 10 展示了 U-Net-Half（带 SN 模块）对 B型验证数据集的预测结果。神经网络能够很好地重现数据，即使在低孔隙率范围内也是如此。然而，该模型无法处理 A型数据样本（论文中未展示）。

作为最后的交叉检验，我们给出了扩散系数随真实孔隙率变化的曲线，见图 11。通过格子玻尔兹曼方法（LBM）获得的数据与分别在 A 型和 B 型数据上训练的两个 U-Net-Half（带 SN 模块）模型的预测结果基本一致。

数据显示，扩散系数与孔隙率之间存在相关性。在图 11中，我们还绘制了 $D(\phi)$ 的拟合结果。为拟合数据，我们假设 $D$ 是曲折度平方的倒数，其定义见[36]。然后，我们使用阿奇定律来表示曲折度，$\lambda$:

$$\lambda^2 = \varphi^{1-n}, \tag{11}$$

其中，$n$ 为阿尔奇指数[36]。我们发现，A型数据（虚线）对应的$n = 3.6$ 和B型数据（实线）对应的$n = 2.8$，与 Boudreau[36] 针对细粒沉积物报告的$n = 2.14$ 吻合良好。

### 3.3 使用 U-Net 重建浓度

本小节总结任务II的研究结果。我们的目标是构建这样一种神经网络：以包含系统几何构型的图像为输入，以显示几何结构和浓度分布的图像为输出。训练数据由成对图像组成：输入几何结构图和浓度图。

与任务I类似，我们考虑两种情形。首先，使用A型数据训练和验证神经网络，然后在B型数据集上测试模型。其次，使用B型数据训练和验证神经网络，再在A型数据集上测试模型。与先前的研究一样，我们采用蒙特卡洛丢弃法（其中 $p = 0.1$）来估计不确定性。

为了评估每个图片样本得到的结果，我们计算输入图片与输出图片之间的差值，以及所有节点上的均值。请注意，对于每个位置，我们都会得到 $1\sigma$ 不确定性；为估算输出图片的平均不确定性，我们计算所有节点上的均值，见图12 和13。

如图 12所示，在 A 型数据上训练的 U-Net 能够很好地预测浓度场。与任务I不同，该模型对 B 型数据也能很好地处理。实际上，在任务 I 中，在 A 型数据上训练的模型无法预测孔隙率和扩散

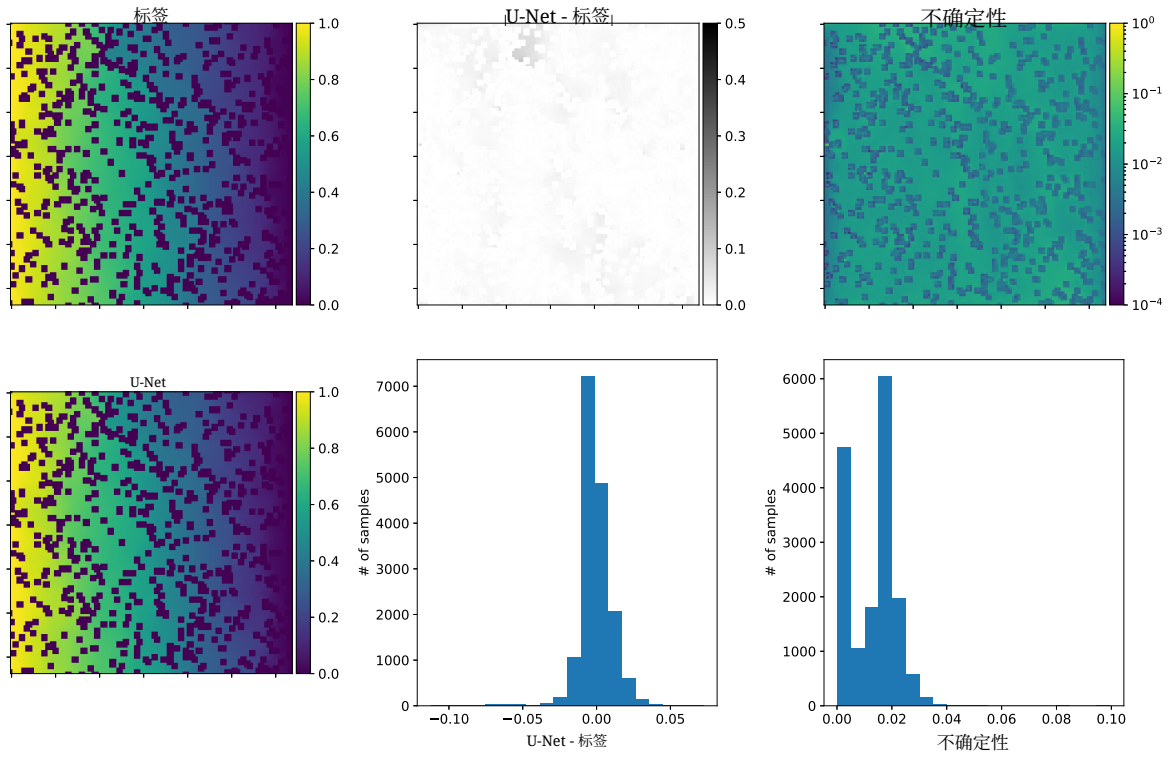


**图14**。 在 A 型数据上训练的 U-Net 对 A 型输入样本中浓度的预测。上排从左至右依次为：输入图、输入与网络响应之间的差异，以及各节点的不确定性分布图。下排从左至右依次为：网络预测结果、输入与网络输出之间差异的直方图，以及各节点不确定性的直方图。输运方向为从左向右。请注意，障碍物所在节点的误差最小；而在孔隙空间中，不确定性相对较大，但分布均匀。

系数，而这只有 B 型样本才达到 $\varphi < 0.4$。此外，在 B 型数据样本上训练的 U-Net 能够准确预测 B 型和 A 型数据样本的浓度场，见图 13。

我们通过展示 U-Net 在选定样本上的表现，结束对数值结果的讨论。图 14 和 15 分别展示了在 A 型数据上训练的神经网络，以及其对 A 型和 B 型数据样本的浓度场预测结果。图 16和17 展示了类似的结果，但对应于在 B 型数据上训练的神经网络，以及其对 B 型和 A 型样本的预测结果。每幅图包含两排图像。第一排依次为真实浓度场分布图、真实值与预测值之间的绝对误差分布图，以及不确定性分布图。第二排包含预测浓度场分布图，以及绝对误差和不确定性的直方图。可以看出，与格子玻尔兹曼方法解中的最大浓度值相比，这四种情况下的网络预测误差都很小。对障碍物形状和位置映射结果的定性比较表明，网络能够很好地完成这项任务。每幅图中的误差（真实值减预测值）分布都集中在零附近。对于大多数样本，空隙空间中的网络预测不确定性比障碍物所在位置更显著；参见图 14-17。

## 4 总结

我们开发了深度学习模型，用于预测多孔介质中扩散的基本性质。我们考虑了两种不同类型的卷积神经网络。两者都以系统几何结构的图像作为输入，但其中一种预测孔隙率和扩散系数，另一种则重建系统的几何结构和浓度图。

对于第一个任务，我们讨论了卷积神经网络的两种变体：C-Net 和 U-Net-Half。这两种变体都采用了 Graczyk 提出的自归一化模块 等人[29]。我们表明，采用 SN 模块的模型比未采用该模块的模型性能更好。对于第二个任务，我们提出采用 U-Net 架构。该模型能够完美重建所有类型数据的浓度场。

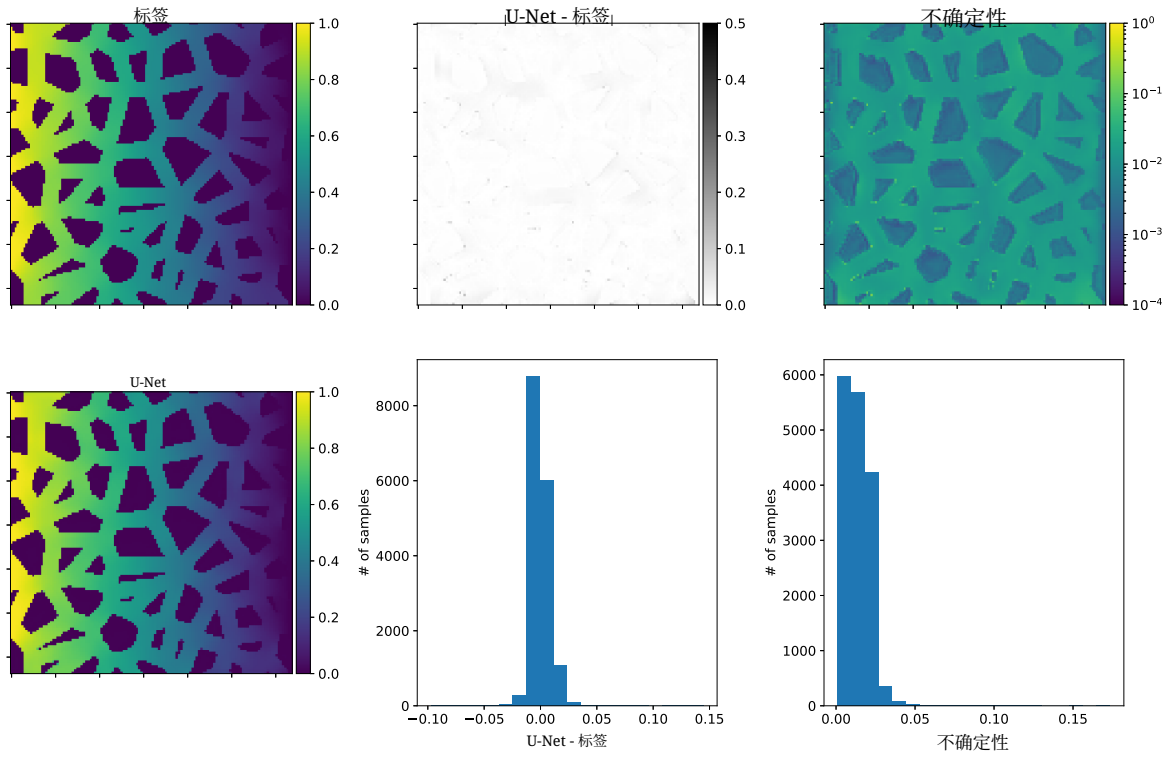


**图15**。与图14 相同，但此处展示的是使用A型数据训练的模型对B型输入样本所作的预测。

我们考虑了两种类型的数据：一种模拟砂堆积体，另一种模拟源自生物组织细胞外空间的系统。我们的目标之一是验证在一种数据上训练的神经网络能否用于分析另一种数据，反之亦然。预测孔隙率和扩散系数的模型仅在其训练所用的数据类型上表现良好。模型对孔隙率和扩散系数的预测均处于 $1\sigma$置信水平界限内。

与任务I不同，为任务II（浓度图重建）开发的模型能够很好地处理训练过程中未包含的数据类型。对于B型数据（即源自生物组织细胞外空间的系统），浓度图重建的准确度最高。在这种情况下，平均误差小于1%，而平均图重建不确定性略高。

## 数据和材料的可用性

本研究所用数据集可按需提供，请联系Krzysztof Graczyk。

## 致谢

D.S. 和 M.M. 得到国家科学中心的支持。本研究由波兰国家科学中心通过 Weave 计划下的 OPUS 项目资助，项目编号为 2021/43/I/ST3/00228。

K.M.G. 在 2020—2026年期间部分获得弗罗茨瓦夫大学“卓越计划——研究型大学”计划的支持。

该出版物部分由弗罗茨瓦夫大学“卓越倡议——研究型大学计划”资助。本研究全部或部分由国家科学中心资助（2021/43/I/ST3/00228）。

为实现开放获取，作者已将 CC-BY 公共版权许可应用于由本次投稿产生的任何作者接受稿（AAM）版本。

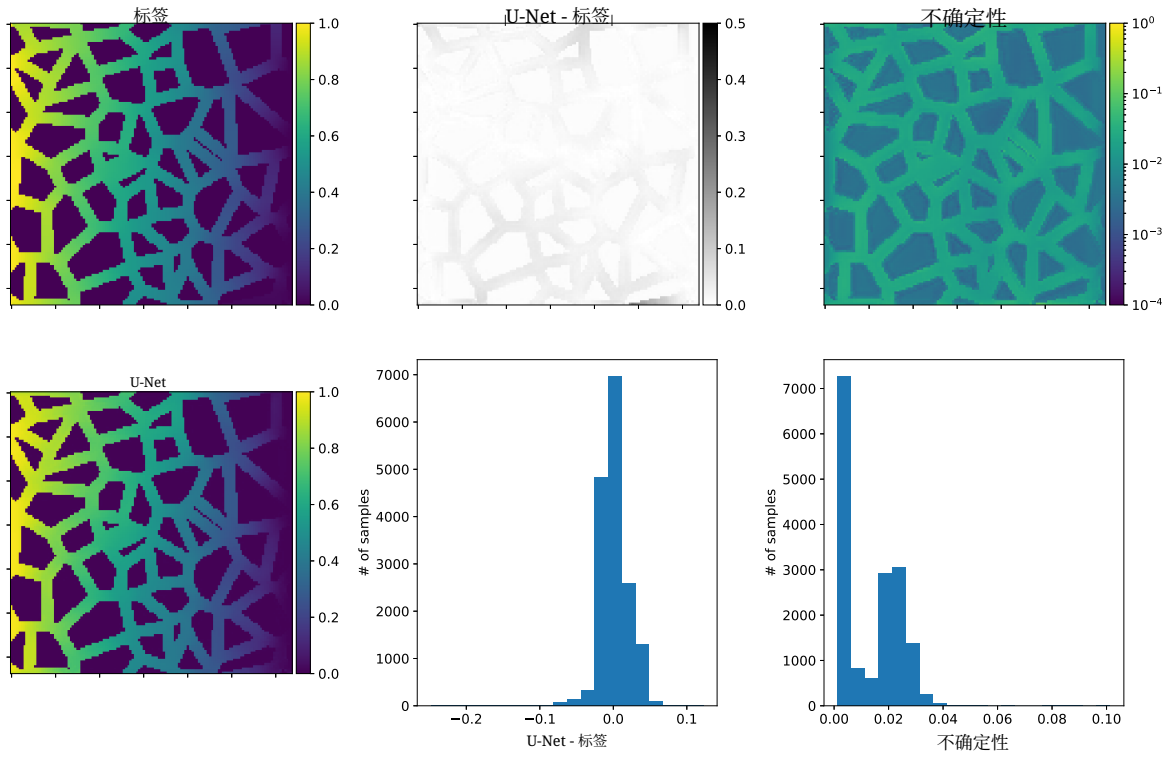


**图16。** 与图 14 相同，但针对使用B型数据训练的模型，以及针对B型输入样本所作的预测。

## 格子玻尔兹曼方法

晶格 B　oltzmann 方法[37,38] 是用于求解离散玻尔兹曼输运方程的数值工具：

$$\frac{\partial f_k}{\partial t} + \vec{\xi}_k \cdot \nabla f_k = \Omega(f_k^{\mathrm{eq}}, f_k); \quad f_k = f_k(\vec{r}, t) \tag{12}$$

其中，$f_k$ 是沿第$k$个粒子速度离散后的分布函数，$\Omega$ 是碰撞算子。根据所采用的离散速度集合 $\vec{\xi}_k$ 和使用的平衡函数 $f^{\mathrm{eq}}{}_k$，方程解的某些矩也同时是宏观输运方程的解。例如，在具有以下九个离散速度的二维空间（$\vec{r} \in \mathbb{R}^2$）中：

$$\begin{array}{lll} \xi_0 = [0,0]^T & \xi_1 = [1,0]^T & \xi_5 = [1,1]^T \\ & \xi_2 = [0,1]^T & \xi_6 = [-1,1]^T \\ & \xi_3 = [-1,0]^T & \xi_7 = [-1,-1]^T \\ & \xi_4 = [0,-1]^T & \xi_8 = [1,-1]^T \end{array} \tag{13}$$

以及平衡分布函数：

$$f_k^{\mathrm{eq}} = \omega_k c, \quad \text{where} \quad \omega_0 = 4/9, \quad \omega_{1,2,3,4} = 1/9, \quad \omega_{5,6,7,8} = 1/36, \tag{14}$$

可得到扩散分子的浓度场（方程（2）的解），它是分布函数的零阶矩：

$$c(\vec{r}, t) = \sum_{k=0}^{8} f_k(\vec{r}, t) \tag{15}$$

在我们的分析中，我们使用式（13）给出的离散速度集合、式（14）给出的平衡分布，以及由 Bhatnagar、Gross 和 Krook 提出的碰撞算子[39]（BGK 碰撞）：

$$\Omega_{\mathrm{BGK}}(f_k^{\mathrm{eq}}, f_k) = \frac{1}{\tau}(f_k^{\mathrm{eq}} - f_k), \tag{16}$$

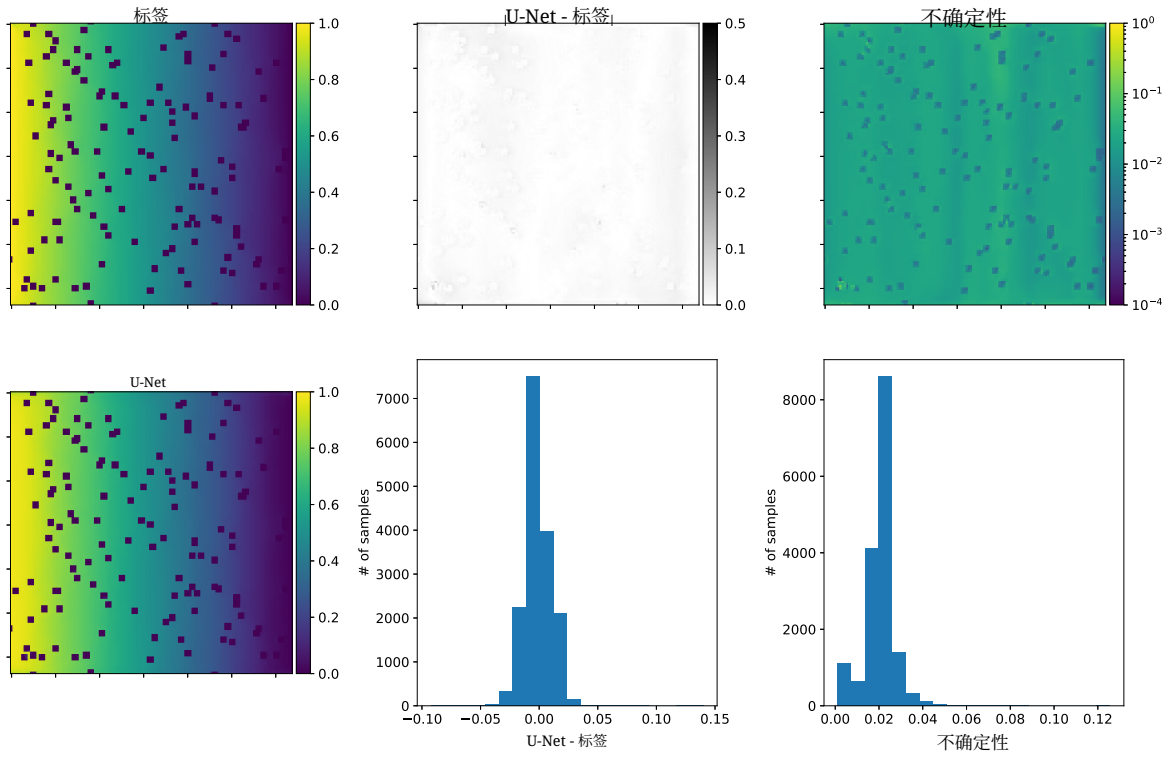


**图 17.** 与图 14 相同，但针对使用 B 型数据训练的模型，以及对 A 型输入样本所作的预测。

其中，$\tau$ 是分布函数达到平衡态的弛豫时间。因此，格子玻尔兹曼方法模型的体扩散系数为

$$D_{0,\mathrm{LBM}} = \frac{1}{3}\left(\tau - \frac{1}{2}\right). \tag{17}$$

我们使用上述量来计算所求解扩散方程的宏观扩散系数：

$$D_0 = D_{0,\mathrm{LBM}}\frac{\delta x^2}{\delta t}, \tag{18}$$

其中 $\delta x$ [m] 和 $\delta t$ [s] 分别为 LBM 在物理空间和时间上的离散化参数

我们采用拉格朗日输运求解离散玻尔兹曼方程，其中碰撞后分布函数 $f_k^*$ 从出发节点 $\vec{r} - \xi_k \delta x$ 输运至到达节点 $\vec{r}$：

$$\underbrace{f_k(\vec{r},t+1) = f_k^*(\vec{r}-\vec{\xi}_k,t)}_{\text{streaming}} = \underbrace{f_k(\vec{r}-\vec{\xi}_k,t) + \frac{1}{\tau}\left(f_k^{\mathrm{eq}}(\vec{r}-\vec{\xi}_k,t) - f_k(\vec{r}-\vec{\xi}_k,t)\right)}_{\text{collision}}. \tag{19}$$

最后，我们假设障碍物表面以及顶部和底部边界处均满足反射边界条件，并采用反弹格式实现：

$$f_k(\vec{r},t+1) = f_{k'}^*(\vec{r},t) \quad \text{when} \quad \vec{r}+\vec{\xi}_{k'} \in \mathscr{D}_s \tag{20}$$

其中 $k'$ 表示与 $k$ 相反的格子矢量，即 $\xi_{k'} = -\xi_k$。在左侧和右侧边界（分别为入口和出口）假设浓度恒定（狄利克雷边界条件），分别等于 $c([0,y],t) = c_{\mathrm{in}} = 1$ 和 $c([L,y],t) = c_{\mathrm{out}} = 0$，通过将未知分布函数设为以下值来实现：

$$f_k = \left(c_B - \sum_{i\in k_{\mathscr{D}}} f_i\right)\frac{\omega_k}{\sum_{i\in k_{\mathscr{D}'}}\omega_i} \quad \text{for} \quad \vec{r}+\vec{\xi}_{k'} \in \mathscr{D}', \tag{21}$$

其中$c_B$表示边界处（0 或 1）的规定浓度，对$k_{\mathscr{D}}$和$k_{\mathscr{D}'}$分别求和表示对位于区域内部和外部的晶格邻居求和 $\mathscr{D}$。因此，式（21）相当于根据位于外部节点的晶格权重，将来自 $\mathscr{D}$ 的浓度进行分配 $\omega_k$ 。 $\mathscr{D}$。

## 参考文献


1. Bell, J. & Grosberg, P. 多孔材料中的扩散。《自然》189, 980–981 (1961)。2. Shen, L. & Chen, Z. 曲折度对扩散影响的批判性综述。Chem. Eng. Sci. 62, 3748–3755 (2007)。3. Kuhn, T. 等. 对活体斑马鱼胚胎中 nodal 和 lefty 的单分子追踪支持受阻扩散模型。Nat. communications 13, 1–15 (2022)。4. Muñoz-Gil, G. 等. 反常扩散解码方法的客观比较。Nat. Commun. 12, 6253, DOI: 10.1038/s41467-021-26320-w (2021)。5. Syková, E. & Nicholson, C. 大脑细胞外空间中的扩散。Physiol. reviews 88, 1277–1340 (2008)。6. Nicholson, C. 脑组织中的扩散及相关传输机制。Reports on progress Phys. 64, 815 (2001)。7. Postnikov, E. B., Lavrova, A. I. & Postnov, D. E. 脑细胞外空间中的输运：扩散，但属于哪种扩散？Int. J. Mol. Sci. 23, 12401 (2022)。8. Tartakovsky, D. M. & Dentz, M. 多孔介质中的扩散：现象与机制。Transp. Porous Media 130, 105–127 ( 2019)。9. Chen, K. C. & Nicholson, C. 脑细胞形状的变化会产生残余细胞外空间体积，并解释渗透压挑战期间的曲折度行为。Proc. Natl. Acad. Sci. 97, 8306–8311 (2000)。10. Weber, R. M., Korneev, S. & Battiato, I. 用于估算锂离子电池有效性质的均质化信息卷积神经网络。Transp. Porous Media 1–22 (2022)。11. Wernert, V. 等. 分级多孔材料的曲折度：扩散实验与随机游走模拟。Chem. Eng. Sci. 264, 118136 (2022)。12. Li, H., Li, H., Gao, B., Wang, W. & Liu, C. 不同粒径砂岩的孔隙特征与微观结构研究。J. Appl. Geophys. 136, 364–371, DOI: https://doi.org/10.1016/j.jappgeo.2016.11.015 (2017年)。13. Kinney, J. P. 等. 细胞外片层和隧道调节海马神经毡中的谷氨酸扩散。J. Comp. Neurol. 521, 448–464 (2013)。14. Godin, A. G. 等. 单纳米管追踪揭示活体脑细胞外空间的纳米尺度组织结构。Nat. Nanotechnol. 12, 238–243, DOI: 10.1038/nnano.2016.248 (2017年)。15. Tartakovsky, A. M., Tartakovsky, D. M. & Meakin, P. 用于多孔介质中流动与输运的随机朗之万模型。Phys. Rev. Lett. 101, 044502, DOI: 10.1103/PhysRevLett.101.044502 (2008)。16. Kalz, E. 等. 碰撞增强奇扩散系统中的自扩散。Phys. Rev. Lett. 129, 090601, DOI: 10.1103/ PhysRevLett.129.090601 (2022)。17. Alexandre, A., Mangeat, M., Guérin, T. & Dean, D. S. 黏滞性如何加快受限系统中的扩散。Phys. Rev. Lett. 128, 210601, DOI: 10.1103/PhysRevLett.128.210601 (2022)。18. Kasthuri, N. 等. 新皮层体积的饱和重建。Cell 162, 648–661, DOI: 10.1016/j.cell.2015.06.054 (2015)。19. Kamrava, S., Tahmasebi, P. & Sahimi, M. 利用深度学习将多孔介质形态与其宏观渗透率联系起来。Transp. Porous Media 131, DOI: 10.1007/s11242-019-01352-5 (2020)。20. Santos, J. E. 等. 应用于复杂三维多孔介质流体流动的高效多尺度神经网络。Transp. porous media 140, 241–272 (2021)。21. Graczyk, K. M. & Matyka, M. 利用深度学习从图像预测多孔介质的孔隙率、渗透率和曲折度。Sci. reports 10, 1–11 (2020)。22. Cawte, T. & Bazylak, A. 利用机器学习方法准确预测多孔纤维材料的传输性质。Electrochem. Sci. Adv. e2100185 (2022)。23. Ranjan Sethi, S., Kumawat, V. & Ganguly, S. 基于卷积神经网络、从显微镜图像预测有效扩散率。J. Appl. Phys. 131, 214901 (2022)。

24. Röding, M., Wåhlstrand Skärström, V. & Lorén, N. 具有预设扩散率的各向异性旋节体材料的逆向设计。Sci. Reports 12, 17413, DOI: 10.1038/s41598-022-21451-6 (2022)。25. Kamrava, S., Tahmasebi, P. & Sahimi, M. 利用深度学习，基于物理和图像预测复杂多孔膜及材料中的流体流动与输运。J. Membr. Sci. 622, 119050, DOI: https://doi.org/10.1016/j.memsci. 2021.119050 (2021)。26. Wu, J., Yin, X. & Xiao, H. 从图像看渗透率：利用卷积神经网络进行快速预测。Sci. Bull. 63, 1215–1222, DOI: https://doi.org/10.1016/j.scib.2018.08.006 (2018)。27. Kamrava, S., Sahimi, M. & Tahmasebi, P. 将控制方程与深层机器相结合，模拟复杂多孔材料中的流体流动。npj Comput. Mater. 7, 127, DOI: 10.1038/s41524-021-00598-2 (2021)。28. Tahmasebi, P., Kamrava, S., Bai, T. & Sahimi, M. 地球与环境科学中的机器学习：从小尺度到大尺度。《水资源进展》142, 103619, DOI: https://doi.org/10.1016/j.advwatres.2020.103619 (2020)。29. Graczyk, K. M., Pawłowski, J., Majchrowska, S. & Golan, T. 用于计数微生物对象的自归一化密度图（sndm）。Sci. Reports 12, 10583, DOI: 10.1038/s41598-022-14879-3 (2022)。30. Ronneberger, O., Fischer, P. & Brox, T. U-net：用于生物医学图像分割的卷积网络 (2015)。1505.04597。31. Gal, Y. & Ghahramani, Z. 将丢弃作为贝叶斯近似：在深度学习中表征模型不确定性 (2015)。1506.02142。32. Wu, H., Fang, W.-Z., Kang, Q., Tao, W.-Q. & Qiao, R. 利用深度学习从图像预测多孔介质的有效扩散率。Sci. Reports 9, 20387, DOI: 10.1038/s41598-019-56309-x (2019)。33. Paszke, A. 等. PyTorch：一种命令式、高性能的深度学习库。载于《神经信息处理系统进展》第32卷，8024–8035页 (Curran Associates, Inc., 2019)。34. Gal, Y. & Ghahramani, Z. 将丢弃作为贝叶斯近似：附录 (2015)。1506.02157。35. Koza, Z., Kondrat, G. & Suszczy´nski, K. 晶格上重叠方形或立方体的渗流。J. Stat. Mech. Theory Exp. 2014, P11005 (2014)。36. Boudreau, B. P. 细粒未固结沉积物的扩散曲折度。《地球化学与宇宙化学学报》60, 3139–3142, DOI: https://doi.org/10.1016/0016-7037(96)00158-5 (1996)。37. Krüger, T. 等. 《格子玻尔兹曼方法：原理与实践》（Springer Cham，2016）。38. Succi, S. 《格子玻尔兹曼方程：流动物质复杂状态》（牛津大学出版社，2018）。39. Bhatnagar, P. L., Gross, E. P. & Krook, M. 气体中的碰撞过程模型 I：带电和中性单组分系统中的小振幅过程。《物理评论》94, 511–525, DOI: 10.1103/PhysRev.94.511 (1954)。

## 作者贡献声明

K.M.G.、D.S. 和 M.M. 构思了本项目。K.M.G. 设计并开展了深度学习分析，提出了使用带自归一化模块的卷积神经网络和 U-Net 架构。D.S. 采用格子玻尔兹曼方法设计并开展了扩散模拟，并准备了数据。所有作者均参与了文本撰写和结果讨论。MM 推导出了阿奇定律中的指数。

## 补充信息

**利益冲突：**作者声明不存在任何利益冲突。