

# 基于模拟数据的 亚细胞 RNA 定位分类研究报告

## 摘要

亚细胞 RNA 定位是基因表达空间调控的一个关键机制，其机制和精确的功能作用尚未完全理解。单分子荧光原位杂交图像能够以亚细胞级精度检测单个 RNA 分子。因此，单分子荧光原位杂交需要强大的方法来定量和分类 RNA 的空间分布。本研究中介绍了一种亚细胞 RNA 定位分类方法，旨在对细胞内不同 RNA 的空间分布进行可视化分析，并展示不同亚细胞定位模式之间的差异。本研究根据收集的细胞模板，使用 Fishquant 工具生成了多种 RNA 定位模式的模拟数据。接着分别在图像和点云两个模态下开展实验，采用深度学习模型对 8 种不同定位模式进行分类。在图像模态中，选择了 Vision Transformer 作为分类模型。在点云模态中，则选用了 PointMamba 模型。实验结果表明，图像模态的模型在亚细胞 RNA 定位分类任务中表现出了更优的性能。

**关键词：**亚细胞 RNA 定位；RNA 模拟数据；深度学习

## 1 引言

信使 RNA (mRNA) 的定位对于基因表达，特别是其空间控制具有重要的功能意义。RNA 的定位与 RNA 代谢或蛋白质代谢有关。RNA 定位不仅仅是一个局部现象，而是转录组中普遍存在的机制，可能还涉及非编码 RNA 的定位。尽管这一过程至关重要，但我们对其了解仍然有限，且缺乏足够的工具来研究这一过程。RNA 的空间分布可以通过基因序列或基于图像的技术来研究。我们重点关注后者，因为它们提供了更好的空间分辨率，因此更适合用于亚细胞 RNA 定位分析。用于在完整细胞中可视化 RNA 的方法是单分子荧光原位杂交 (smFISH)，该方法有多种变体，其中一种可扩展且成本效益高的版本，我们将在以下研究中使用。在 smFISH 中，特定 RNA 种类的单个 RNA 分子通过多个荧光标记的寡核苷酸进行标记，并在显微镜下显示为明亮的衍射极限斑点，从而可以通过传统的图像分析或计算机视觉方法进行检测。通常，还会使用一些额外的荧光标记来标记相关的细胞结构，这有助于确定每个单独 RNA 相对于这些标志物的位置。通过这样的多通道显微镜图像，我们可以获得每个单独细胞及其 RNA 的坐标表示，如图 1 所示。RNA 定位导致了几种不同的模式，通常可以通过局部过度拥挤的亚细胞区域来定义。在文献 [1] 中，即使在简单的生物系统（如 HeLa 细胞）中，也描述了几种定位模式：一种随机的默认模式 (Random)，其中 RNA 均匀地定位在细胞内；RNA 聚集体 (Foci)；沿着核膜的高密度转录本 (Nuclear edge)；在细胞核内 (Intranuclear)；在细胞延伸部分 (Protrusion)；或细胞内的极生化，例如 RNA 定位在核附近 (Perinuclear)。

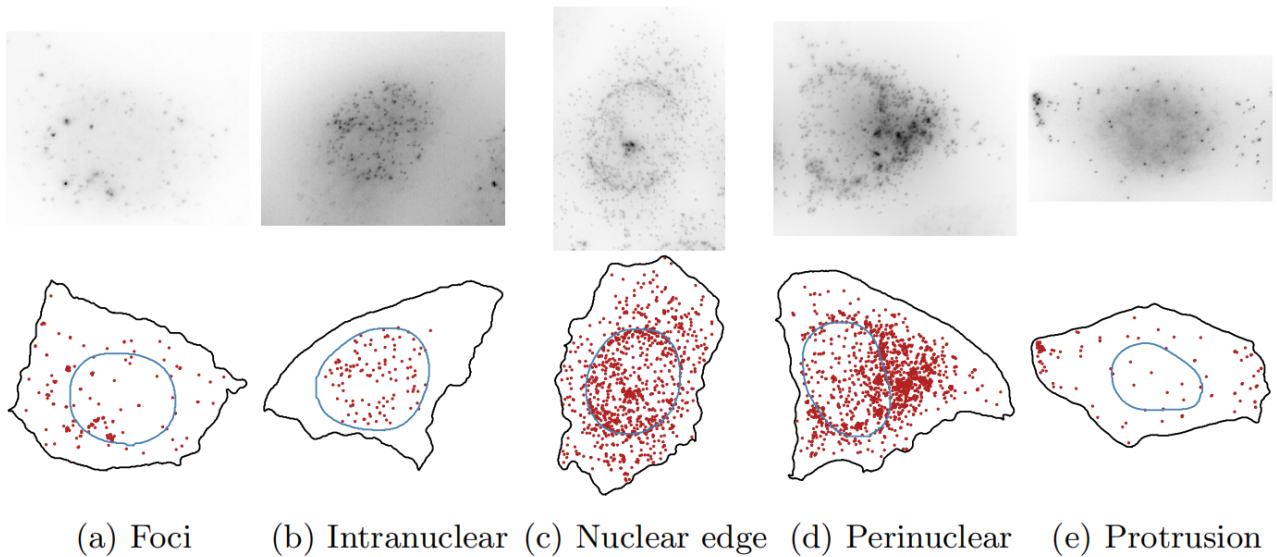


图 1. 上面的图像是不同的 RNA 定位模式的 smFISH 图像，下面的图像是其坐标表示：RNA 斑点（红色）、细胞膜（黑色）和核膜（蓝色） [1]

亚细胞 RNA 定位是基因表达调控研究中的一个重要领域。RNA 分子在细胞中的位置不仅影响其翻译和功能，还在细胞内的生物学过程、信号传导及疾病发生等方面起着至关重要的作用。因此，研究 RNA 在细胞内的空间定位模式对于理解细胞功能和生物学现象具有重要意义。然而，由于 RNA 分子通常存在于细胞内多个不同的亚细胞区域，且其空间分布复杂，因此如何准确高效地进行亚细胞 RNA 定位的自动化分析一直是生物信息学和细胞生物学中的一个挑战。随着高分辨率显微成像技术的快速发展，细胞内 RNA 的定位和可视化已变得更加可行。同时，计算机视觉技术，特别是深度学习方法的引入，为 RNA 定位任务提供了强有力的工具。这些方法能够从大量的细胞图像数据中自动学习特征，从而实现 RNA 分子在细胞内的精确定位和分类。本文提出了一种基于模拟数据的 RNA 定位分类方法，旨在探索细胞内不同 RNA 定位模式的自动化识别和分类。本研究基于 Fishquant 工具 [2] 生成了不同 RNA 定位模式的模拟数据，并在图像模态和点云模态下分别进行了分类实验。在图像模态中，本文使用了 Vision Transformer (ViT) 模型 [3]，而在点云模态中，则采用了 PointMamba 模型 [4]。实验结果表明，图像模态模型在亚细胞 RNA 定位分类任务中表现出更为优异的性能，展示了深度学习在 RNA 定位中的巨大潜力。

## 2 相关工作

### 2.1 亚细胞 RNA 定位模式识别

亚细胞 RNA 定位作为一个多学科交叉的研究领域，近年来得到了广泛的关注。在一些疾病中，RNA 定位的异常可能导致蛋白质的错误表达和功能丧失，从而影响细胞功能。通过研究 RNA 的亚细胞定位，可以帮助理解这些疾病的发生机制。许多研究者致力于利用实验技术、计算方法以及深度学习技术，探索 RNA 在细胞内的空间分布和定位模式。

亚细胞 RNA 定位的研究主要依赖于实验方法来揭示 RNA 分子在细胞内的空间分布。传统的实验技术包括原位杂交 (ISH) 和荧光原位杂交 (FISH)，这些方法通过使用特定的探针

标记 RNA 分子并通过显微镜观察其在细胞内的分布，提供了有关 RNA 定位的直观证据。近年来，FISH 技术已经取得了较大的进展，特别是通过多重 FISH 和超分辨率显微技术，能够同时检测多个 RNA 分子的空间分布，并且提高了定位的精度。然而，这些传统技术往往依赖人工观察，并且需要大量的实验标记和成像操作，难以处理大规模数据。因此，如何将实验数据与计算方法结合，进行自动化分析，成为了亚细胞 RNA 定位研究的一个重要方向。

在以往的研究中，开发了用于分类 RNA 定位模式的手工特征 [5]。这些特征的设计受到了空间统计学文献 [6] 的启发，并从荧光显微镜图像分析管线 [7] 中借鉴而来。已有一些软件包实现了执行 smFISH 分析并计算这些手工特征的模块 [2]。然而，这些方法需要仔细设计一组特征，以对应具体的生物学问题。对于不同的研究，可能需要新的特征集。在此，我们旨在探讨一种更通用的方式来构建定位特征。

## 2.2 图像表征学习

图像表征学习旨在通过从原始像素数据中提取有效的特征表示，以便进行下游任务如分类、检测、分割等。近年来，深度学习方法尤其是卷积神经网络（CNN）在图像表征学习中取得了显著成果。经典的卷积神经网络如 LeNet [8]、AlexNet [9] 和 VGG [10]，通过层叠卷积层、池化层和全连接层，成功地将原始图像转换为高效的低维特征表示。此外，ResNet [11] 通过引入残差连接，极大地改善了深层网络的训练效果，成为了图像表征学习中的一个重要里程碑。随着网络架构的不断优化，图像表征学习逐渐向更高效、更灵活的方向发展。例如，Google 提出的 Inception 网络 [12]，通过采用多尺度卷积核，使得网络能够捕捉更多层次的信息。与此同时，Transformer 架构在自然语言处理中的成功应用促使了 Vision Transformer (ViT) [3] 等基于自注意力机制的网络模型在图像表征学习中的应用。ViT 的主要优势在于它能够对图像进行全局建模，而非仅依赖局部卷积特征，这为图像表征提供了新的视角。

## 2.3 点云表征学习

点云数据的表征学习在计算机视觉领域中逐渐成为一个重要研究方向，尤其是在三维物体识别、场景重建和机器人导航等应用中。与传统的图像数据不同，点云数据通常是由无序的三维坐标点组成，其结构复杂且不规则，这给点云的表征学习带来了挑战。近年来，深度学习方法在点云表征学习中取得了显著进展。

最早的点云表征学习方法之一是 PointNet [13]，该方法提出了一种基于对称函数的网络架构，使得点云的无序性得以处理。PointNet 通过将每个点独立处理，并使用最大池化操作聚合全局特征，成功地为点云分类和分割任务提供了有效的表征。其后，PointNet++ [14] 通过层次化的结构改进了 PointNet，能够捕捉局部几何信息，并提高了对复杂场景的处理能力。

为了进一步提高点云表征的精度和效率，许多研究者提出了基于 CNN 的点云处理方法。例如，PointConv [15] 通过设计卷积操作来处理不规则点云数据，并取得了良好的表现。与传统卷积操作不同，PointConv 基于局部点集学习卷积滤波器，能够在点云上自适应地学习局部特征。

此外，PointMamba [4] 作为一种改进的点云处理模型，针对点云数据的特点进行了优化。该模型通过引入更加灵活的特征学习策略和崭新的 Mamba 架构，使得在处理复杂的点云分布时，分类精度能够进一步提高。

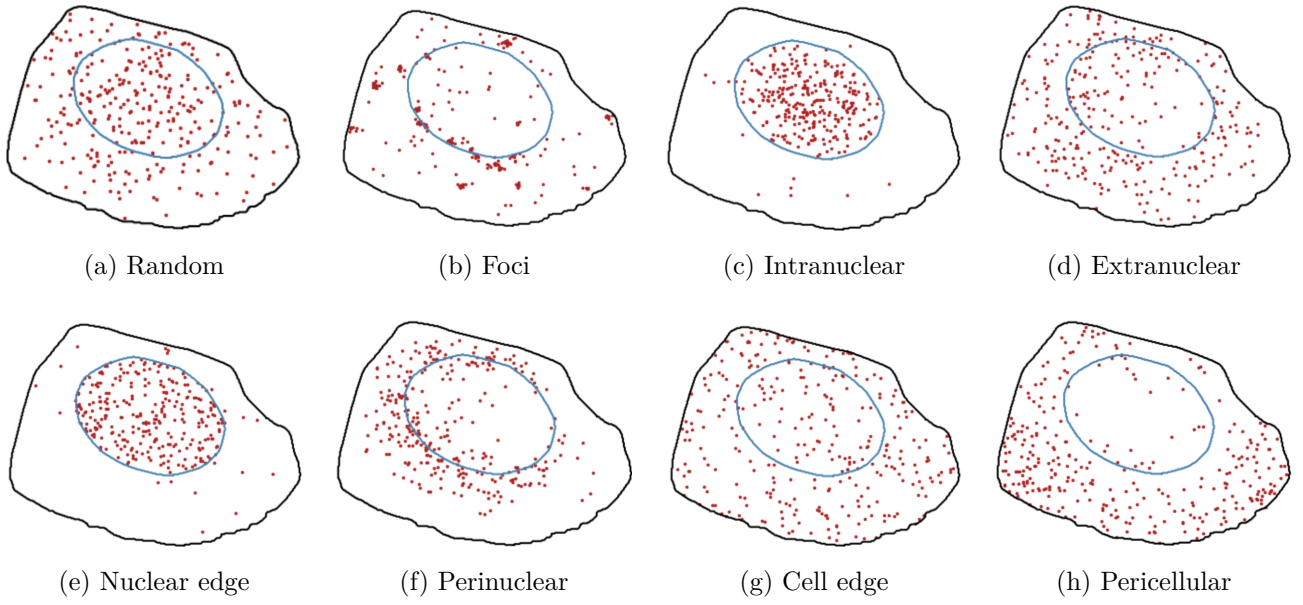


图 2. 亚细胞 RNA 定位分类介绍

### 3 本文方法

#### 3.1 模拟数据的生成

##### 3.1.1 亚细胞 RNA 定位分类介绍

如图 2 所示，常见的亚细胞 RNA 定位总共有 8 种，下文依次进行介绍：

- 随机 (Random)：RNA 在细胞内均匀。
- 焦点 (Foci)：RNA 聚集成簇。
- 核内 (Intranuclear)：主要在细胞核内部分布。
- 核外 (Extranuclear)：主要在细胞核外部均匀分布。
- 核膜 (Nuclear edge)：沿着核膜的转录本高密度分布。
- 核周 (Perinuclear)：在靠近细胞核的位置分布。
- 细胞膜 (Cell edge)：沿着细胞膜的转录本高密度分布。
- 细胞周质 (Pericellular)：主要在细胞质聚集分布。

##### 3.1.2 亚细胞 RNA 定位数据的生成

为了进行亚细胞 RNA 定位的分析，本研究使用了 Fishquant 工具，该工具是基于 RNA 原位杂交技术的定量分析工具。Fishquant 能够从 FISH 图像中提取 RNA 分子的亚细胞定位信息，并生成定量数据，以便进行进一步分析和可视化。



模拟可以用于验证分析管线的不同步骤，从斑点检测到定量研究 RNA 定位的统计框架。如上所述，我使用该软件包验证了斑点检测和密集区域分解方法。我通过三步模拟真实的 sm-FISH 图像。首先，我根据特定的细胞模版中生成 2D 或 3D 的斑点坐标，这些坐标可以是随机的，或显示特定的亚细胞 RNA 定位模式。此外，还可以添加聚集的 RNA。第二，我通过将 RNA 斑点建模为高斯函数，模拟符合相应定位模式分布的 RNA 坐标。第三，我向图像中添加噪声背景。

我总共生成了三种类型的数据：首先是散点图，每个 RNA 分子的位置通过散点图的形式显示在图像上，然后将相应细胞膜和核膜也绘制在图像上；其次是核密度（KDE）图，根据细胞内的 RNA 分子坐标计算 KDE，然后将 KDE 密度、细胞膜和核膜绘制在图像上；最后是点云，在细胞膜与核膜上均匀采样若千的点，然后与该细胞内 RNA 的点合并为一个点云数据。

### 3.2 ViT

ViT 是基于 Transformer 架构 [16] 的图像处理模型，其基本思想是将图像划分为固定大小的图像块，并通过 Transformer 进行特征学习。这种方法与传统的 CNN 不同，ViT 采用的是全局自注意力机制，能够捕捉长程依赖关系，从而在多个视觉任务中取得了显著效果。

#### 3.2.1 图像块化处理

如图 3 所示，ViT 的输入首先是一个二维图像，尺寸为  $H \times W \times C$ ，其中  $H$  是图像的高度， $W$  是宽度， $C$  是通道数。为了将图像适配到 Transformer 架构，ViT 将图像划分为若干个大小为  $P \times P$  的图像块。假设图像被划分为  $N = HW/P^2$  个图像块，每个图像块的尺寸为  $P \times P \times C$ 。每个图像块经过展平后，将变为一个一维向量，尺寸为  $P^2 \times C$ 。接下来，通过一个线性映射层将每个图像块嵌入到一个固定维度的空间中，即

$$z_i = W_{patch} \cdot Flatten(X_i) + b_{patch} \quad (1)$$

其中， $X_i \in R^{P \times P \times C}$  表示第  $i$  个图像， $z_i \in R^D$  是其嵌入后的表示， $W_{patch} \in R^{D \times (P^2 C)}$  是线性变换矩阵， $b_{patch}$  是偏置项， $D$  是嵌入空间的维度。

将所有图像块嵌入后，形成一个序列  $Z = [z_1, z_2, \dots, z_N] \in R^{N \times D}$ ，作为 Transformer 的输入。

#### 3.2.2 位置编码

由于 Transformer 架构缺乏处理序列中元素位置的能力，因此需要引入位置编码。ViT 通过将每个图像块的位置信息编码为位置向量并与图像块的嵌入向量相加来保留空间位置信息。位置编码的计算公式为

$$z'_i = z_i + PE_i \quad (2)$$

其中， $z'_i$  是加入了位置编码后的图像块嵌入向量， $PE_i \in R^D$  是第  $i$  个图像块的位置信息。

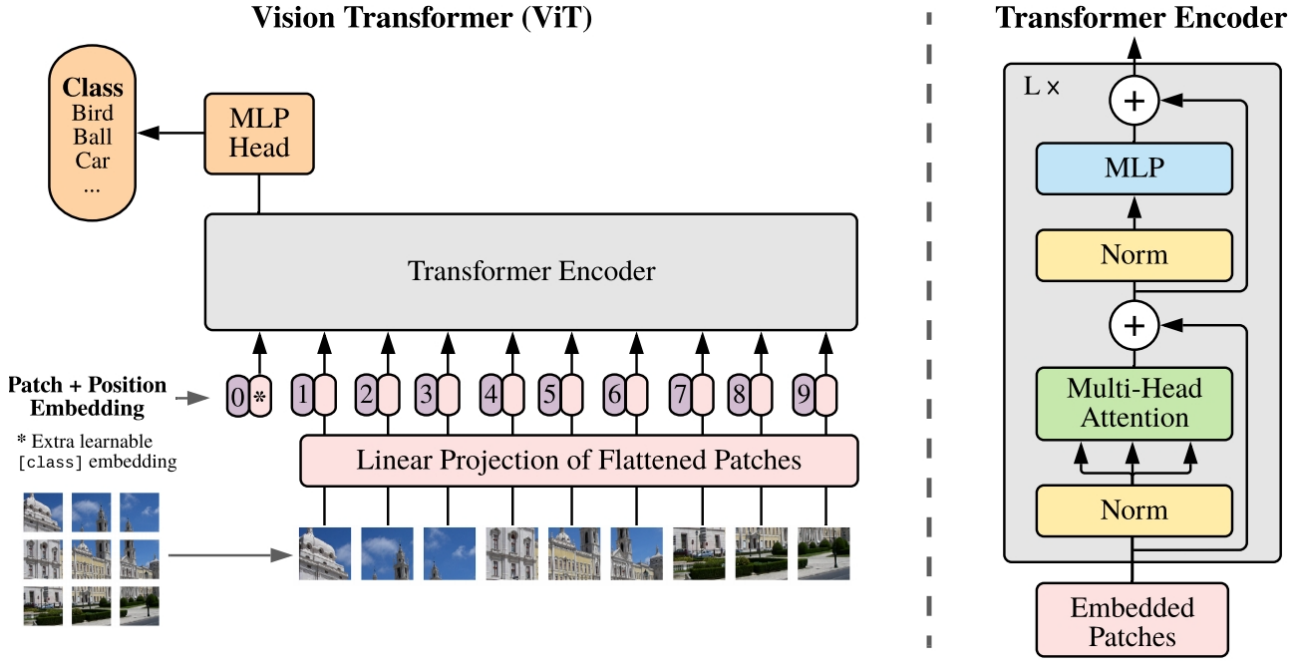


图 3. ViT 架构 [3]

### 3.2.3 Transformer 编码器

ViT 采用标准的 Transformer 编码器结构，包含多个自注意力 (Self-Attention) 层和前馈神经网络 (Feed-Forward Neural Network, FFN)。

首先计算自注意力层，计算查询 (Query)、键 (Key) 和值 (Value)：

$$Q = ZW_Q, K = ZW_K, V = ZW_V \quad (3)$$

其中， $W_Q, W_K, W_V$  是训练得到的权重矩阵， $Q, K, V \in R^{N \times D}$  分别是查询、键和值矩阵。接着计算注意力权重并得到加权值：

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{D}}) \quad (4)$$

其中， $\frac{QK^T}{\sqrt{D}}$  是查询和键的点积除以  $\sqrt{D}$ ，目的是防止内积值过大而导致梯度消失，softmax 用于计算归一化的注意力权重。

最后经过一个前馈神经网络，由两层全连接层组成，其计算公式为：

$$FFN(z) = max(0, zW_1 + b_1)W_2 + b_2 \quad (5)$$

其中， $W_1, W_2$  是全连接层的权重矩阵， $b_1, b_2$  是偏置项。

### 3.3 PointMamba

受控制理论的启发，状态空间模型表示一个连续系统，通过隐式潜在状态  $h_t \in R^N$  将一个状态  $x_t$  映射到  $y_t$ 。为了将 SSM 集成到深度模型中，S4 [17] 使用四个参数 (A, B, C 和采样步长  $\Delta$ ) 来定义该系统。序列到序列的变换定义为：

$$h_t = \bar{A}h_{t-1} + \bar{B}x_t, y_t = Ch_t + Dx_t \quad (6)$$

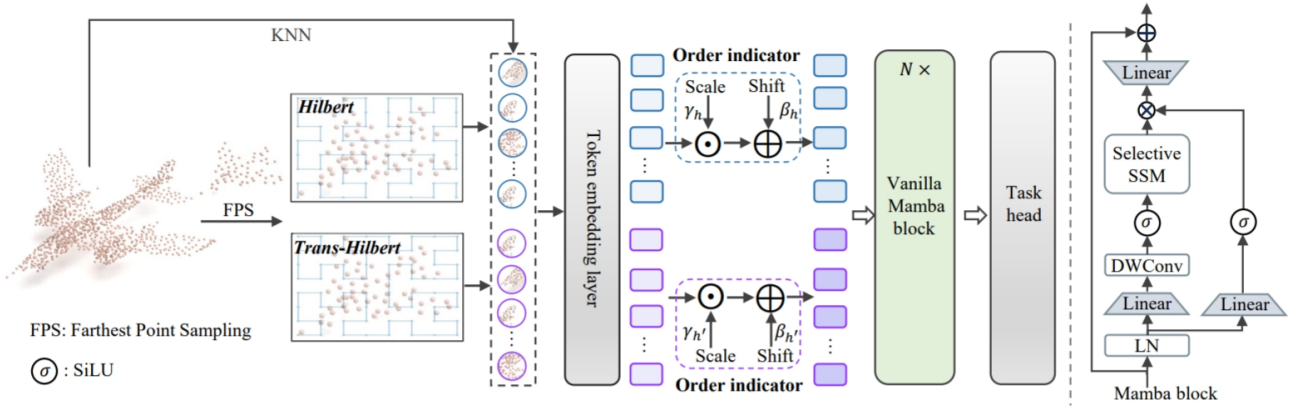


图 4. PointMamba 架构 [4]

其中,  $C \in R^{1 \times N}$  是映射参数,  $D \in R^{1 \times N}$  代表残差连接, 参数  $\bar{A}, \bar{B}$  代表使用了零阶保持离散化:

$$\bar{A} \in R^{N \times N} = \exp(A\Delta), \bar{B} \in R^{N \times 1} = (A\Delta)^{-1}(\exp(A\Delta) - I) \cdot \Delta B \quad (7)$$

如图 4 所示, PointMamba 首先利用最远点采样 (Farthest Point Sampling, FPS) 选择关键点。具体来说, 给定输入点云  $P \in R^{M \times 3}$ , 其中  $M$  是点的数量, 应用 FPS 从原始点云  $P$  中采样  $n$  个关键点, 得到  $p \in R^{n \times 3}$ 。通常情况下, 采样的关键点  $p$  的顺序是随机的, 没有特定的顺序。对于先前基于 Transformer 的方法, 这不是一个重要问题, 因为 Transformer 在处理序列数据时是顺序无关的: 在自注意力机制中, 给定位置的每个元素可以通过注意力权重与序列中的所有其他元素进行交互。然而, 对于选择性状态空间模型, 我们认为由于单向建模的特点, 难以建模无结构的点云。因此, 我们提出利用空间填充曲线将无结构的点云转换为规则的序列。具体而言, 我们选择了两种具有代表性的空间填充曲线来扫描关键点: 希尔伯特曲线及其转置变体, 称为 Trans-Hilbert。与随机序列相比, 像希尔伯特曲线这样的空间填充曲线可以很好地保留空间局部性, 即在扫描的 1D 序列化点序列中, 邻近的关键点在 3D 空间中的位置通常是几何上接近的。我们认为, 这一特性确保了点与点之间的空间关系在很大程度上得以保持, 这对于点云数据中的特征表示和分析至关重要。作为补充, Trans-Hilbert 表现相似, 但从不同的线索扫描, 这可以提供关于空间局部性的多样化视角。通过应用希尔伯特和 Trans-Hilbert 曲线到关键点上, 我们得到两个不同的点序列化  $p_h$  和  $p'_h$ , 这些用建立点 token。

在获得两个序列化的关键点  $p_h$  和  $p'_h$  后, 我们利用 KNN 算法为每个关键点选择  $k$  个最近邻, 形成  $n$  个 token 补丁  $T_h \in R^{n \times k \times 3}$  和  $T'_h \in R^{n \times k \times 3}$ , 补丁大小为  $k$ 。为了聚合局部信息, 每个补丁中的点通过减去关键点来进行归一化, 得到相对坐标。然后使用轻量级 PointNet 将无偏的局部补丁映射到特征空间, 得到序列化点嵌入  $E_0^h \in R^{n \times C}$  和  $E_0^{h'} \in R^{n \times C}$ , 其中前者是基于希尔伯特的, 后者是基于 Trans-Hilbert 的。

将两种序列化的点令牌直接输入 Mamba 编码器可能会引起混淆, 因此 PointMamba 提出了一个非常简单的顺序指示器, 用于指示使用的扫描策略:

$$Z_0^h = E_0^h \odot \gamma_h + \beta_h, Z_0^{h'} = E_0^{h'} \odot \gamma_{h'} + \beta_{h'} \quad (8)$$

在获取 token  $Z_0$  后, 我们将其输入到包含  $N$  个 Mamba 块的编码器中, 以提取特征。具体而言, 对于每个 Mamba 块, 采用层归一化、选择性 SSM 和深度卷积。

### 3.4 可视化方法

#### 3.4.1 主成分分析

主成分分析 (Principal Component Analysis, PCA) 是一种广泛应用于数据降维和特征提取的线性变换技术。PCA 的主要目标是通过正交变换, 将数据集从高维空间映射到低维空间, 同时尽可能保留原始数据中的方差信息。PCA 通过寻找数据中的主成分, 即数据的最大变异方向, 来实现这一目标。

PCA 的第一步是计算数据集的协方差矩阵。假设数据集为  $X \in R^{n \times d}$ , 其中  $n$  是样本数,  $d$  是特征数。首先, 对每个特征进行均值中心化处理, 即将每一列减去其均值。然后, 计算数据的协方差矩阵  $C$ :

$$C = \frac{1}{n-1} X^T X \quad (9)$$

协方差矩阵反映了数据中不同特征之间的线性关系。矩阵中的每个元素  $C_{ij}$  表示第  $i$  个特征与第  $j$  个特征之间的协方差。如果两个特征之间的协方差为正, 则它们在同一方向上变动; 若为负, 则表示它们在相反的方向上变动; 若为零, 则表明两者之间没有线性关系。

接下来, 通过对协方差矩阵进行特征值分解, 得到数据的主成分。特征值分解的目的是将数据投影到新的坐标系中, 其中新的坐标轴 (即主成分) 按照数据方差的大小依次排列。具体来说, 协方差矩阵  $C$  的特征值和特征向量通过以下关系获得:

$$C v_i = \lambda_i v_i \quad (10)$$

其中,  $\lambda_i$  是第  $i$  个特征值,  $v_i$  是对应的特征向量。特征值表示了每个主成分所代表的方差大小, 而特征向量则是数据集中的主成分方向。

通过特征值的大小, 我们可以确定每个主成分的重要性。较大的特征值对应的数据主成分能保留更多的方差信息。因此, 选择特征值较大的主成分可以最大化数据的变异性。常见的做法是选择前  $k$  个最大特征值对应的特征向量, 构成新的低维空间。特征值的选择可以通过累计方差贡献率来指导, 通常选择前  $k$  个主成分, 使得累计方差贡献率达到某一预设的阈值 (例如 95%)。

最后, 将数据投影到选定的主成分上。若我们选择了前  $k$  个主成分对应的特征向量  $v_1, v_2, \dots, v_k$ , 则可以通过以下方式将原始数据映射到新的低维空间:

$$X' = X V_k \quad (11)$$

其中,  $V_k$  是包含前  $k$  个特征向量的矩阵,  $X$  是降维后的数据。在这个新的低维空间中, 数据的主要特征被保留, 而噪声和冗余信息被削减。

#### 3.4.2 UMAP

Uniform Manifold Approximation and Projection (UMAP) [18] 是一种用于降维和数据可视化的非线性算法。UMAP 的主要目标是保持高维数据中的局部结构, 并在低维空间中尽可能准确地重建数据的全局结构。它是基于流形学习的思想, 并且在高维数据的可视化、聚类和分类等任务中具有广泛的应用。

UMAP 的理论基础是流形假设, 即高维数据通常位于一个低维流形上。流形是一种局部看起来像欧几里得空间的非线性空间, 这使得高维数据可以在低维空间中得到有效的表示。



UMAP 通过构建数据的局部邻域关系来学习数据的流形结构，从而实现降维。在高维空间中，数据点通常只有局部关系，而全局结构则较难直接建模。UMAP 通过优化局部结构的保留，同时减少全局结构的失真，取得了良好的降维效果。

UMAP 首先构建一个高维空间中的局部邻域图。具体来说，对于数据集中的每个点，UMAP 使用 K 近邻算法或基于距离的半径邻域算法来定义其邻居。UMAP 通过计算每个数据点到其邻居的距离，并在距离较近的邻居之间建立高权重的连接。这些邻接关系用于构建图的边。UMAP 在这一阶段基于对称化的距离度量和局部密度来捕捉数据的局部几何结构。

在低维空间中，UMAP 采用随机初始化的低维表示，并通过优化目标函数来使低维空间的结构尽可能保持高维空间中的局部邻域关系。优化目标函数由两部分组成：局部邻域的保留和数据点之间的吸引力和排斥力。具体而言，UMAP 采用了一个基于 Frobenius 范数的损失函数，该函数衡量高维空间和低维空间中点对之间的相似度差异。通过最小化该损失函数，UMAP 能够在低维空间中找到一个嵌入，使得高维空间中的局部关系在低维空间中得到最好的保留。

UMAP 使用梯度下降法对目标函数进行优化，利用变步长的优化策略来加速收敛过程。该优化过程中，UMAP 不仅通过邻接点的关系来更新低维表示，还引入了拉普拉斯平滑项，以进一步减少全局结构的扭曲。最终，通过多次迭代优化，UMAP 能够在低维空间中得到一个理想的表示，保持数据的局部结构，并同时尽量减少全局结构的失真。

## 4 实验

### 4.1 实验设计

为了模拟与细胞形态一致的模式，我使用了实际的细胞模板进行 Fishquant 定位模拟。我收集了 318 个细胞模版作为数据集，数据集包括细胞膜、细胞核膜的坐标，用于构建真实的细胞形态。为了避免由于细胞形状导致的信息泄露，我将 318 个细胞模版以 6:2:2 的比例进行划分，分别用于训练集、验证集和测试集的模拟数据生成，每个模版生成 8 份数据，分别对应 8 种定位模式。

我使用 Fishquant 的 sim-fish 库，在细胞中模拟 8 种定位模式的 RNA 分布，每个细胞的每种定位模式都模拟出 1024 个 RNA 坐标。具体来说，每个数据都可以控制符合该定位模式的点的百分比，这里我设定的是在 [60% 70% 80% 90%] 这四个百分比中随机选取一个百分比进行单个数据生成。最终每个细胞生成了三种类型数据：散点图、核密度采样图和点云。

对于图像模态，采用 ViT-base 进行训练和推理，对训练集图像采用水平翻转和竖直翻转进行数据增强；对于点云模态，采用 PointMamba 进行训练和推理，对训练集点云采用旋转进行数据增强。训练使用交叉熵损失函数，使用 AdamW 优化器 [19] 进行参数优化。验证集和测试集使用准确率来进行评测，每个训练轮次结束时，都会使用验证集进行一次评测，如果验证集的准确率连续 5 个轮次未提升，则直接停止训练，缓解过拟合问题。最后使用 UMAP 和 PCA 前 7 维热图进行可视化。

### 4.2 实验环境

本实验在 Python 3.9.20, pytorch 1.13.1, sim-fish 0.2.0 版本下运行。

## 5 实验结果

散点图数据的损失变化图、验证集准确率变化图、UMAP 可视化图和 PCA 热图，如图 5 所示。验证集最佳准确率为 99.6875%，测试集准确率为 99.5625%。

核密度图数据的损失变化图、验证集准确率变化图、UMAP 可视化图和 PCA 热图，如图 6 所示。验证集最佳准确率为 99.5%，测试集准确率为 99.5%。

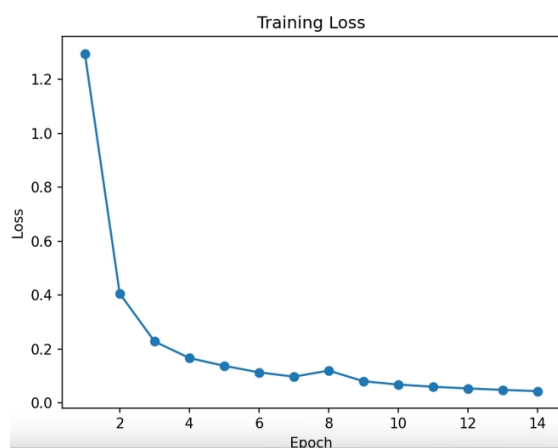
点云数据的损失变化图、验证集准确率变化图、UMAP 可视化图和 PCA 热图，如图 7 所示。验证集最佳准确率为 80.7617%，测试集准确率为 82.5195%。

可以看到三种类型数据的 8 个定位模式在 UMAP 图和 PCA 热图中能很明显地区分开，并且散点图的效果最好。

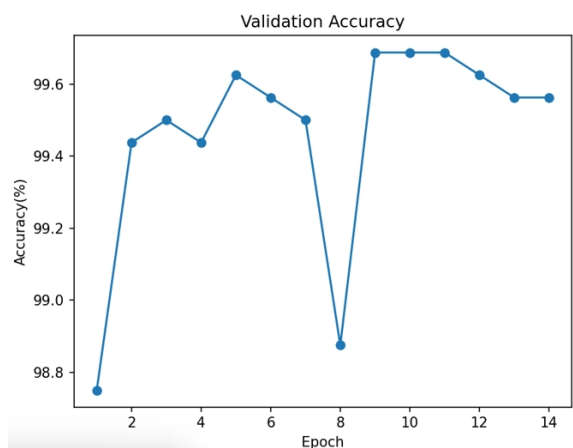
## 6 总结与展望

亚细胞 RNA 定位是基因表达空间调控的一个关键机制，对于帮助理解这些疾病的发生机制等任务起到极其关键的作用。本研究根据收集的细胞模板，使用 Fishquant 工具生成了多种 RNA 定位模式的模拟数据，接着分别在图像和点云两个模态下开展实验，采用深度学习模型对 8 种不同定位模式进行分类，最终验证了散点图的效果是最好的。

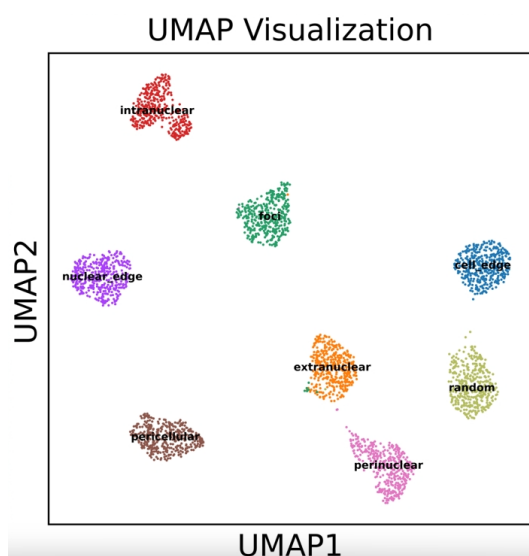
但是本研究是在模拟数据上进行的，未来可以将应用范围拓展到现实，在真实的数据集上进行训练以及验证，用于协助各种生物工作的开展。



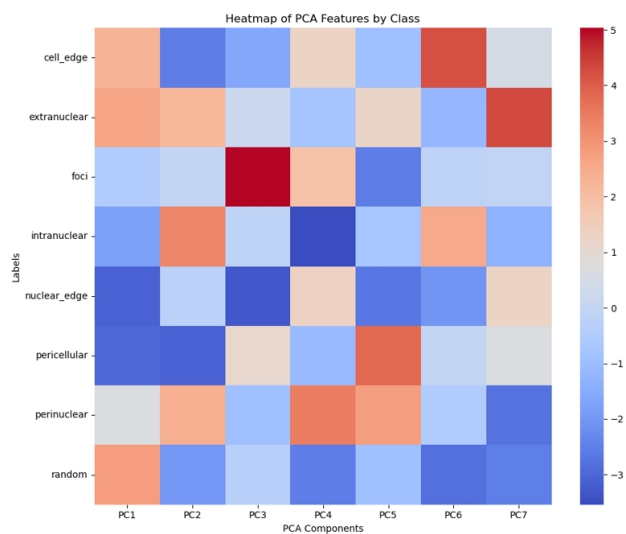
(a) 损失曲线



(b) 验证集准确率

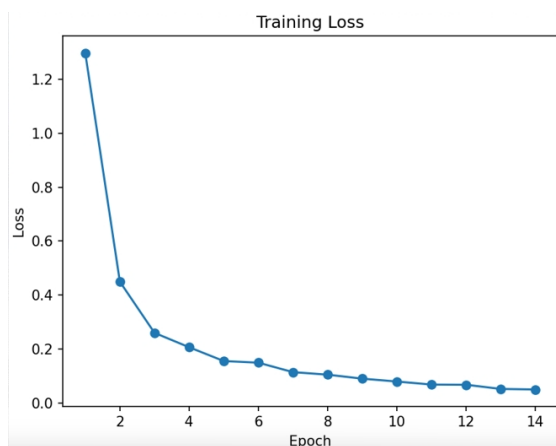


(c) UMAP 可视化

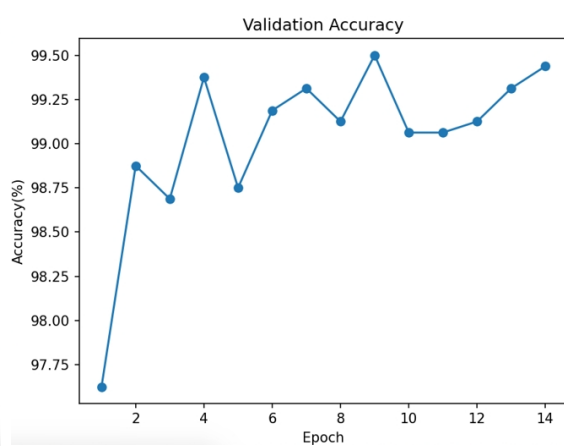


(d) PCA 前 7 维热图

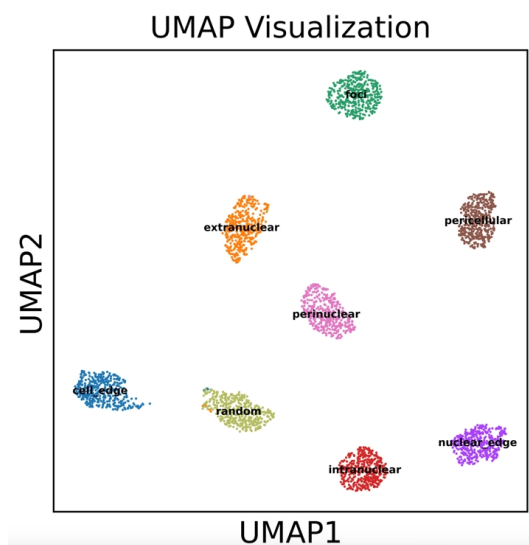
图 5. 散点图实验结果



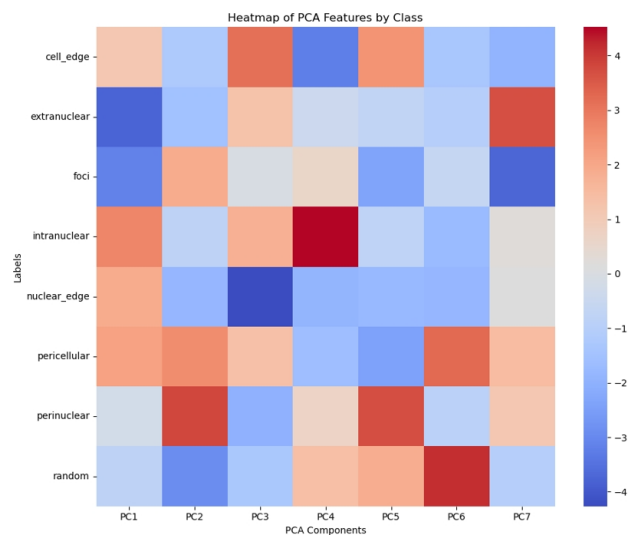
(a) 损失曲线



(b) 验证集准确率



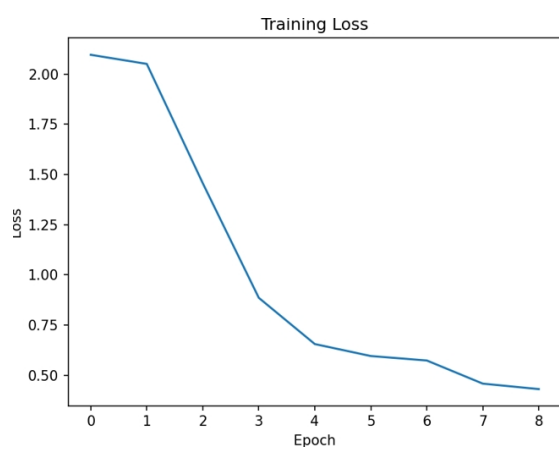
(c) UMAP 可视化



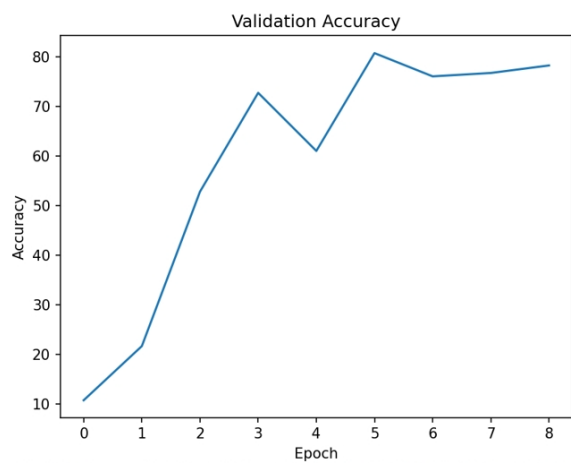
(d) PCA 前 7 维热图

图 6. 核密度图实验结果

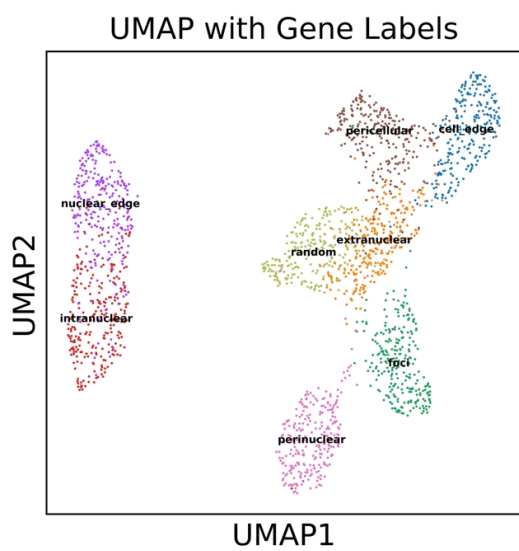




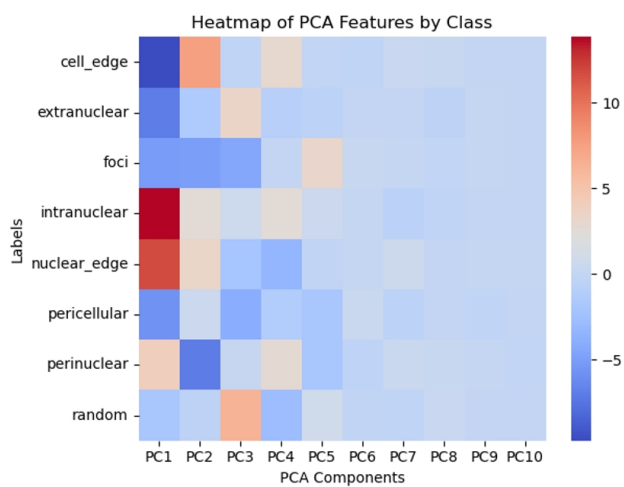
(a) 损失曲线



(b) 验证集准确率



(c) UMAP 可视化



(d) PCA 前 7 维热图

图 7. 点云实验结果

## 参考文献

- [1] Arthur Imbert, Florian Mueller, and Thomas Walter. Pointfish: Learning point cloud representations for rna localization patterns. In *European Conference on Computer Vision*, pages 487–502. Springer, 2022.
- [2] Arthur Imbert, Wei Ouyang, Adham Safieddine, Emeline Coleno, Christophe Zimmer, Edouard Bertrand, Thomas Walter, and Florian Mueller. Fish-quant v2: a scalable and modular tool for smfish image analysis. *Rna*, 28(6):786–795, 2022.
- [3] Alexey Dosovitskiy. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [4] Dingkan Liang, Xin Zhou, Wei Xu, Xingkui Zhu, Zhikang Zou, Xiaoqing Ye, Xiao Tan, and Xiang Bai. Pointmamba: A simple state space model for point cloud analysis. *arXiv preprint arXiv:2402.10739*, 2024.
- [5] Racha Chouaib, Adham Safieddine, Xavier Pichon, Arthur Imbert, Oh Sung Kwon, Aubin Samacoits, Abdel-Meneem Traboulsi, Marie-Cécile Robert, Nikolay Tsanov, Emeline Coleno, et al. A dual protein-mrna localization screen reveals compartmentalized translation and widespread co-translational rna targeting. *Developmental cell*, 54(6):773–791, 2020.
- [6] Nico Battich, Thomas Stoeger, and Lucas Pelkmans. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nature methods*, 10(11):1127–1133, 2013.
- [7] Walter Andrew Shewhart, Samuel Stanley Wilks, CE McCulloch, and SR Searle. *Wiley series in probability and statistics*. Wiley Online Library, 1986.
- [8] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.
- [10] Karen Simonyan. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

- [12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
- [13] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017.
- [14] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30, 2017.
- [15] Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 9621–9630, 2019.
- [16] A Vaswani. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017.
- [17] Albert Gu, Karan Goel, and Christopher Ré. Efficiently modeling long sequences with structured state spaces. *arXiv preprint arXiv:2111.00396*, 2021.
- [18] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- [19] I Loshchilov. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.