

KGMamba:基于Kolmogorov-Arnold网络优化图卷积网络和Mamba的基因调控网络预测模型

高泰, 任艳璋, 王会青, 李颖, 王彬

引用本文

高泰, 任艳璋, 王会青, 李颖, 王彬. [KGMamba:基于Kolmogorov-Arnold网络优化图卷积网络和Mamba的基因调控网络预测模型](#)[J]. 计算机科学, 2026, 53(4): 101-111.

GAO Tai, REN Yanzhang, WANG Huiqing, LI Ying, WANG Bin. [KGMamba:Gene Regulatory Network Prediction Model Based on Kolmogorov-Arnold Network Optimizing Graph Convolutional Network and Mamba](#) [J]. Computer Science, 2026, 53(4): 101-111.

相似文献推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[跨模态融合的少样本勒索软件分类器:基于预训练模型的多模态编码](#)

Cross-modal Fusion Few-sample Ransomware Classifier:Multimodal Encoding Based on Pre-trained Models

计算机科学, 2026, 53(4): 435-444. <https://doi.org/10.11896/jsjcx.250500078>

[知识辅助和强化句法驱动的方面级情感分析](#)

Knowledge-assisted and Reinforced Syntax-driven for Aspect-based Sentiment Analysis

计算机科学, 2026, 53(4): 406-414. <https://doi.org/10.11896/jsjcx.250600117>

[融合局部多视角语言特征和全局特征的对话情感四元组抽取](#)

Multi-view Local Language Feature and Global Feature Fusion for Conversational Aspect-based Sentiment Quadruple Analysis

计算机科学, 2026, 53(4): 384-392. <https://doi.org/10.11896/jsjcx.250900032>

[基于WL图核的多通道图Kolmogorov-Arnold网络](#)

Multi-channel Graph Kolmogorov-Arnold Network Based on WL Graph Core

计算机科学, 2026, 53(4): 224-234. <https://doi.org/10.11896/jsjcx.250600033>

[STWD-DLFRD:基于序贯三支决策与深度学习的多粒度虚假评论检测方法](#)

STWD-DLFRD:Multi-granularity Fake Review Detection via Sequential Three-way Decisions and Deep Learning

计算机科学, 2026, 53(4): 188-196. <https://doi.org/10.11896/jsjcx.250500088>

KG Mamba: 基于 Kolmogorov-Arnold 网络优化图卷积网络和 Mamba 的基因调控网络预测模型

高泰 任艳璋 王会青 李颖 王彬

太原理工大学计算机科学与技术学院(大数据学院) 山西 晋中 030600

(2023520656@link.tyut.edu.cn)

摘要 基因调控网络(Gene Regulatory Network, GRN)推断对于解析细胞发育机制及推动精准医学研究至关重要,但是现有深度学习方法面临计算复杂度高与全局特征捕捉不足的挑战。为此,提出一种融合 Kolmogorov-Arnold 网络(KAN)驱动的图卷积网络(KGCN)与 Mamba 模块的高效预测模型。首先,以 KAN 特有的可学习样条函数,取代图卷积网络中的多层感知器(MLP)模块。该改进在完整保留邻居节点局部特征提取能力的基础上,通过重构计算逻辑降低特征处理的冗余度,使模型计算复杂度较传统图卷积架构实现显著优化。其次,创新性地引入 Mamba 模块,通过其选择性机制优先关注对全局调控起关键作用的基因节点。两者结合实现了局部特征提取效率与全局依赖建模能力的协同优化。在公开数据集上与另外 6 种深度学习模型进行实验比较,结果显示,该模型在 AUC 和 AUPR 性能指标上都优于其他模型,同时展现出显著的鲁棒性优势和计算效率。

关键词: 基因调控网络;深度学习;Kolmogorov-Arnold 网络;图卷积网络;Mamba

中图分类号 TP183

KG Mamba: Gene Regulatory Network Prediction Model Based on Kolmogorov-Arnold Network Optimizing Graph Convolutional Network and Mamba

GAO Tai, REN Yanzhang, WANG Huiqing, LI Ying and WANG Bin

College of Computer Science and Technology(College of Data Science), Taiyuan University of Technology, Jinzhong, Shanxi 030600, China

Abstract Gene regulatory network(GRN) inference is pivotal for deciphering cell development mechanisms and propelling precision medicine research. However, existing deep learning approaches confront challenges of high computational complexity and inadequate global feature capture. To tackle this, a novel efficient prediction model integrating the Kolmogorov-Arnold network (KAN) driven graph convolutional network(KGCN) and Mamba module is proposed. Firstly, the multi-layer perceptron(MLP) in traditional graph convolution is replaced by KAN's learnable spline functions, which retain local feature extraction while reducing redundancy through restructured computation, significantly improving efficiency. Secondly, the Mamba module is innovatively incorporated to prioritize attention to gene nodes critical for global regulation via its selective mechanism. Together, these components enable a unified optimization of both local and global feature modeling. Experimental comparisons with six other deep-learning models on public datasets are performed. Results demonstrate that this model outperforms others in AUC and AUPR performance metrics, while also showcasing remarkable robustness and computational efficiency, further demonstrating the superiority of the model.

Keywords Gene regulatory network, Deep learning, Kolmogorov-Arnold network, Graph convolutional network, Mamba

1 引言

GRN 揭示多组学遗传物质间的调控相互作用,决定细胞发育及环境反应^[1]。整合基因间调控关系有助于解析细胞对外界刺激的复杂功能,并推动分子生物学与临床医学研究,如靶向药物开发、基因治疗。尽管统计与机器学习方法可利用

基因表达数据重建 GRN,但精准推断基因间的调控关系仍是生物信息学领域的难题^[2]。近几十年来,经国内外研究人员持续探索,涌现出大量基于机器学习与统计的 GRN 重建方法。研究表明,其发展与生物技术、计算机技术进步密切相关。

传统 GRN 重建方法主要采用无监督推理框架,如基于相关性的方法^[3]、基于信息论的方法^[4]、贝叶斯网络^[5]等。

到稿日期:2025-05-22 返修日期:2025-09-03

基金项目:国家自然科学基金(62176177);山西省科技合作交流专项项目(202304041101034)

This work was supported by the National Natural Science Foundation of China(62176177) and Science and Technology Cooperation and Exchange Special Projects of Shanxi(202304041101034).

通信作者:王彬(wangbin01@tyut.edu.cn)

然而,随着高通量测序技术的发展,样本与基因节点数量激增,早期模型已难以胜任。为了使用更多样本和基因节点来重建 GRN,研究者们随后提出了基于深度学习的神经网络框架。例如,利用卷积神经网络(Convolutional Neural Networks,CNN)从基因表达图像中识别基因调控相互作用。近年来,因果推断 GCN(Graph Convolutional Networks)^[6]和有向 GCN(DGCGRN)^[7]则基于 GCN,采用图嵌入技术并结合先验生物领域知识。

当前基于 GCN 推断 GRN 的模型,普遍面临两大关键挑战:1)计算复杂度随基因节点数量及图结构复杂程度急剧攀升,呈指数级增长态势,这对计算资源与时间成本提出了极高要求,极大限制了模型在大规模高维基因数据场景下的应用效能;2)模型在信息传递机制上过度聚焦局部邻域特征,对 GRN 的全局结构特性及长距离基因互作关系的捕捉能力不足,使得在解析复杂基因调控关系时,难以完整、准确地刻画基因间的多层次相互作用,进而影响预测结果的全面性与精确性^[8]。

为了突破上述方法的局限性,本文提出了一种基于 KAN(Kolmogorov-Arnold Networks)^[9]优化的 GCN 与 Mamba^[10]的 GRN 预测模型。本文的主要贡献如下:

1)首次使用 KAN 优化图卷积网络,创新性地提出 KGCN,兼具局部特征提取能力和较低的计算复杂度。

2)首次将 Mamba 引入 GRN 预测任务,利用其选择性机制优先关注对全局调控起关键作用的基因节点。

3)在基准公开数据集上进行了广泛实验,结果显示本文模型在关键预测性能指标(AUC 和 AUPR)上显著优于同类方法,且在计算效率方面实现了突破性优化,兼具高准确性与高效性。

本研究不仅在模型结构上实现创新,也具有重要的理论与应用价值。在理论层面,KGMamba 首次将 KAN 引入图神经网络领域,为图结构中的非线性建模提供了全新范式,有助于推动图学习算法在高维稀疏数据上的发展。同时,Mamba 模块以状态空间机制替代传统注意力机制,提升了远程依赖建模能力,为解决 GCN 局部性瓶颈提供了新思路。在应用层面,GRN 的精确推断对识别关键调控因子、研究疾病机制及寻找潜在药物靶点具有重要意义。KGMamba 在处理大规模、高稀疏单细胞 RNA 测序(scRNA-seq)数据时表现良好,具有广泛的应用前景,尤其适用于精准医学、癌症研究、干细胞发育等复杂生物系统的调控网络构建。

2 相关工作

2.1 基因调控网络预测研究现状

目前,推断 GRN 的方法通常可分为无监督方法和有监督方法。无监督方法主要包括基于信息论^[11]、模型^[12]和机器学习^[13]的 3 类方法。基于信息论的方法,如互信息 MI^[14]、Normi^[15]等通过基因相关性分析,假设相关性强度与调控可能性呈正相关;基于模型的方法,如 LogBTF^[12]和 MetaSEM^[16]通过拟合基因表达谱构建模型以重建 GRN,其中 LogBTF 结合正则化逻辑回归与布尔阈值函数,MetaSEM 利用结构方程模型和元学习应对高维稀疏数据挑战。另外,

GCHMM-DBN 模型^[17]通过在 HMM-DBN 的基础上增加全局耦合超参数,在所有时间段中共享具有相似高斯分布的噪声方差超参数和信噪比超参数,实现了回归参数的全局耦合。基于机器学习的方法,如 GENIE3^[18]采用随机森林独立推断基因 GRN,以特征重要性为调控分数。此外,GABR^[19]将基因表达数据转化为原子序列,通过基因关联熵有效识别基因之间的因果关系,并采用启发式搜索策略构建基因关联贝叶斯调控网络,巧妙地将信息论、模型构建(贝叶斯网络模型)和机器学习进行了融合。尽管这些方法无需标签信息,但普遍难以有效处理 scRNA-seq 数据的噪声、缺失值、高稀疏性和高维度问题,且多数仅适用于小数据集,大规模数据可能导致计算时间过长。

最近,深度学习技术在 GRN 推断的有监督方法中应用广泛。早期的尝试,如 DGRNS^[20]结合循环神经网络与 CNN 提取时间特征和空间特征以推断 GRN;CNCC^[21]将基因对表达值转为直方图并用 CNN 分类,但受传递相互作用问题影响。为缓解该问题,DeepDRIM^[22]引入相邻基因信息作为额外输入。

随着自然语言领域技术的进步,基于 Transformer 的 GRN 推断方法也被开发。GRN-Transformer^[23]构建轴向 Transformer 弱监督框架,从 scRNA-seq 数据推断细胞类型特异性 GRN 并推导通用 GRN;STGRNS^[24]通过基因表达基序技术将基因对编码为子向量,并输入 Transformer 以推断 GRN。此外,GCN 也被应用于 GRN 预测。Ji 等^[6]结合因果信息与 GCN,通过转移熵和重构层减少多轮邻居聚合导致的信息损失,并利用高斯核自编码器提取可分离特征。GCN 作为 GRN 推断中主流的图结构学习方法,其优化方向主要集中在特征聚合机制和邻居选择策略上。GraphSAGE^[25]引入采样机制,解决了大规模图计算的效率问题;GAT^[26]和 GATv2^[27]通过引入注意力权重,在一定程度上提升了特征选择的灵活性。然而,这些方法在提升建模能力的同时显著增加了模型复杂度,且依赖 MLP 模块作为非线性映射器,存在表达能力受限、冗余度高等问题。相比之下,KGCN 通过将 MLP 替换为 KAN 样条函数模块,在保证特征表达能力的同时显著减少了参数量,具备更强的非线性拟合能力与轻量化计算特性。

2.2 Kolmogorov-Arnold 网络

KAN 最近在机器学习领域引起了广泛关注^[28]。与传统多层感知机(MLP)不同,KAN 在边上部署可学习的单变量函数,而非固定激活函数与静态权重。这一特性使得 KAN 通过一维样条函数实现了复杂非线性关系的建模,其参数效率显著高于 MLP,尤其在高维数据场景下表现突出。在个性化推荐这一重要应用领域,Su 等^[29]也通过 KAN 网络替代传统 MLP,结合 BERT 微调技术与评分、评论多源特征融合策略,在 Yelp、Amazon 等具有高稀疏度的数据集上实现了更优的评分预测性能,进一步从跨任务角度验证了 KAN 在复杂非线性建模、参数效率提升上的优势。KAN 尽管刚被提出,但已被应用于诸多领域,如卫星图像分类任务^[30]和谣言检测^[31]。到目前为止,大多数研究集中在时间序列数据上^[32]。此外,KAN 还与时间序列预测任务中常用的 LSTM 网络和

Transformer 结合使用。

然而, KAN 在 GCN 中的应用仍属空白。本文首次将 KAN 引入 GCN, 提出 KAN 优化的 GCN 模块(KGCN), 通过将 GCN 中的 MLP 层替换为 KAN 层, 显著降低模型参数量与计算复杂度, 同时利用 KAN 的灵活非线性建模能力增强局部基因调控特征的提取效率。这一创新为 GRN 预测任务提供了更轻量化且高效的解决方案。

2.3 状态空间模型

自 S4^[33] 模型首次提出状态空间序列建模以来, 其通过频域参数化捕捉长距离依赖, 但在训练过程中计算复杂度较高, 需大量内存与 FFT 操作。S5^[34] 在此基础上引入精简架构, 虽将复杂度降为线性, 但在实际任务中面临表示能力下降的问题。H3^[35] 进一步提升了表达力, 但结构更为复杂。上述模型均主要针对语言或序列数据设计, 难以直接迁移至高稀疏、非结构对称的 GRN 任务中。

与之相比, Mamba^[10] 在结构上引入选择性状态机制, 通过输入依赖的参数生成方式, 自适应聚焦关键特征区域, 抑制噪声干扰, 提升建模鲁棒性。同时, Mamba 摒弃了传统 S4 和 S5 中的复杂频域运算, 采用结构化卷积设计, 训练效率更高、资源占用更低^[36]。本文首次将 Mamba 引入 GRN 预测中, 结合 KGCN 实现了局部-全局建模能力的融合互补, 突破了神经网络与状态空间建模分离发展的局限, 为复杂调控关系建模提供了新范式。

3 KGMamba 预测模型

3.1 问题描述

本文将 GRN 预测任务视为一个给定基因表达数据和 GRN 的二元分类问题。给定基因表达数据集 $\{E_1, E_2, \dots,$

$E_n\}$, 其中 E_m 表示第 m 个基因的基因表达, 对于每个基因 $E_m = \{x_1, x_2, \dots, x_n\}$, 它由基因 m 在 n 个样本中的表达值组成。同时, 提供已知的 GRN $\{(g_i, g_j)\}$, 该数据可看作图数据, 节点代表基因, 边代表基因之间的相互作用关系。任务目标是根据上述两种数据提取基因的特征向量, 对于任意基因对, 利用两个基因各自的表达特征预测两者之间是否存在调控关系。

3.2 框架描述

首先, 在数据预处理模块中, 针对原始 scRNA-seq 数据, 选取前 500 和 1000 个高度变异基因。

其次, 使用 KGCN 通过高效聚合一阶邻居节点的特征信息, 精准提取局部邻域的关键信息来捕捉基因间的依赖关系, 同时借助 KAN 中可学习的样条函数, 在保留 GCN 对图结构依赖关系建模能力的同时, 进一步捕捉基因间的复杂非线性相互作用, 有效降低了模型的计算复杂度, 更使网络在处理高维基因数据时, 兼具对局部信息的精细提取能力与强大的特征表达能力, 从而更精准地解析基因间的复杂调控关系。

然后, 引入 Mamba 模块, 利用其独特的选择性机制与对长程依赖的捕捉能力, 不仅能动态聚焦高维稀疏数据中的关键基因信息, 更可对全局范围内的基因关联进行有效提取。在其后加入归一化步骤, 解决因特征尺度差异导致的训练不稳定问题, 实现对基因依赖关系的高效建模。

最后, 将基因对的特征向量进行点积运算, 并经 ReLU 激活函数处理, 生成基因调控得分矩阵, 以此预测基因间的调控关系。

本框架主要由 3 个模块组成: 数据预处理模块、特征提取模块和关系预测模块。本文提出的 KGMamba 模型框架如图 1 所示。

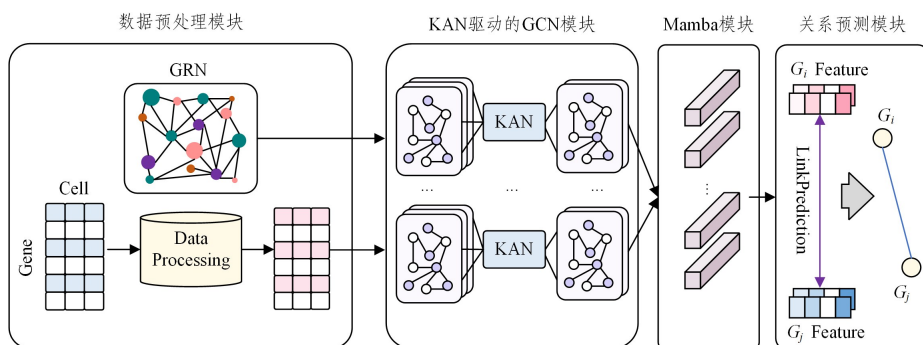


图1 KGMamba 模型

Fig.1 KGMamba model

3.3 数据预处理模块

为处理 scRNA-seq 数据的规模冗余性, 采用成熟方法^[37] 进行预处理: 过滤低表达基因(表达细胞比例小于 10%), 基于 Bonferroni 校正筛选显著高变异基因(P 小于 0.01)。基因表达水平通过对数转换进行归一化。对单细胞表达数据执行上述操作后, 得到特征矩阵 $X \in \mathcal{R}^{N \times M}$, 其中 N 为基因数量, M 为细胞数量。为评估计算方法在不同规模网络中的有效性, 遵循 Pratapa 等^[37] 提出的方法, 按基因方差排序选取高变异转录因子及前 500/1000 个基因, 构建 14 个不同规模的数据集。

3.4 特征提取模块

3.4.1 Kolmogorov-Arnold 网络

作为 MLP 的替代方案, KAN 已引起广泛关注。如图 2 所示, 该模型基于 Kolmogorov-Arnold 表示定理, 该定理断言任何多元函数 $f: [0, 1]^d \rightarrow \mathbf{R}$ 均可表示为:

$$f(x) = \sum_{i=1}^{2d+1} \varphi_i \left(\sum_{j=1}^d \phi_{ij}(x_{ij}) \right) \quad (1)$$

其中, 所有函数 φ 和 ϕ 都是一元函数, 求和是唯一的多元运算。

式(1)可看作一个两步过程。首先, 将一组不同的一元非线性激活函数应用于输入的每个维度, 然后将这些函数的输

出进行求和。作者基于这种理解定义了一个 KAN 层,它是一个从空间 $A \subseteq \mathbf{R}^d$ 到另一个空间 $B \subseteq \mathbf{R}^{d'}$ 的映射(在用途上与 MLP 层相同)。这样的一层由 $d \times d'$ 个可训练函数构成。那么,对于 $x \in A$,其映射结果 x' 的计算式如下:

$$x_i' = \sum_{j=1}^d \phi_{ij}(x_j) \quad (2)$$

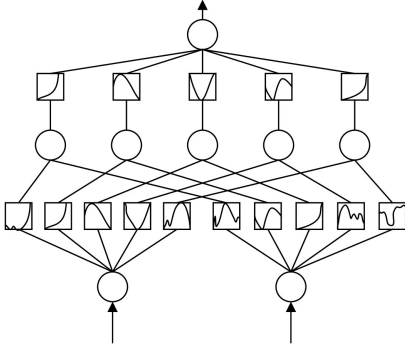


图2 Kolmogorov-Arnold 网络

Fig.2 Kolmogorov-Arnold network

堆叠两个这样的层,第一层输入维度为 d ,输出维度为 $(2d+1)$,第二层输入维度为 $(2d+1)$,输出维度为 1,由此推导的模型是一个通用函数逼近器。它似乎比 MLP 更具复杂度优势,因为表示从 $[0,1]^d$ 到 \mathbf{R}^d 的任意多元函数所需的一元函数数量最多为 $(2d^2+d) \times d'$,而 MLP 的通用逼近定理则可能需要无限数量的神经元。然而,正如原论文所述,此类一元函数的行为可能具有任意复杂性,从而导致其不可表示和不可学习。

MLP 通过堆叠有限宽度的层来放宽无限宽度的约束。类似地,KAN 通过堆叠较低复杂度的 KAN 层来放宽对非线性函数的任意复杂性约束。因此,函数的输出如式(3)所示:

$$y = \text{KAN}(x) = \phi_L \circ \phi_{L-1} \circ \dots \circ \phi_1(x) \quad (3)$$

其中, ϕ_1, \dots, ϕ_L 为 KAN 层。

KAN 的原论文采用样条函数,即可训练的分段多项式函数作为非线性组件。这使得 KAN 能够在参数数量相对较少的情况下保持高表达能力,但代价是需要满足一定的局部平滑性。因此,第 l 层由一个 $d_l \times d_{l-1}$ 的样条网格构成。每个样条使用的次数(称为样条阶数)以及每个函数使用的样条数量(称为网格大小)均为架构的超参数。尽管如此,KAN 允许任何函数作为其基础元素。因此,替代实现可使用不同的基函数,例如径向基函数,这类基函数尤其能缓解样条函数的一些计算瓶颈。

3.4.2 Kolmogorov-Arnold 网络驱动的图卷积网络模块

为获取基因的初始特征,对原始单细胞表达数据执行预处理操作。随后,利用基于 KAN 驱动的 GCN,通过挖掘基因交互图的潜在结构来学习基因特征。将先验网络表示为 $G = \{V, \xi\}$,其中 $V \in \mathbf{R}^N$ 是包含 N 个节点的集合, ξ 是边的集合。在图编码器的框架下,主要目标是通过迭代聚合邻居节点的特征,学习图 G 中每个节点 v_i 的特征。形式上,基于 GCN 的图编码器的第 l 层可描述为:

$$h_i^l = \text{AGGREGATE}(\{h_j^{(l-1)}; v_j \in N_i\}) \quad (4)$$

节点 v_j 在第 $l-1$ 层的特征表示为 h_j^{l-1} 。网络中节点 v_j

的一阶邻居(包括节点 v_j 本身)表示为 N_i 。为获得节点 v_j 在第 l 层的表示,图编码器在计算该层特征时引入聚合函数 $\text{AGGREGATE}(\cdot)$,将其一阶邻居集合 N_j 中各节点在第 $l-1$ 层的特征进行汇聚,从而完成对 v_j 特征的更新。它是一个通用组件,可以使用各种图神经架构来定义,包括流行的 GCN。

GCN 已成为分析图结构数据的强大工具,在各个研究领域展现出显著的有效性。鉴于此,本研究利用 GCN 通过整合一阶邻居特征数据来获取基因特征,并将此过程描述为一个 GCN 层。首先,假设网络中的每个节点都是自连接的。这一假设能够定义一个归一化的邻接矩阵,表示为 $\tilde{\mathbf{A}}$ 。矩阵 $\tilde{\mathbf{A}}$ 的计算方式为 $\tilde{\mathbf{A}} = \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{1/2}$,其中 \mathbf{A} 表示图的邻接矩阵,其中元素 A_{ij} 表示节点 v_i 和 v_j 之间存在一条边。对角矩阵 \mathbf{D} 的构造方式是使其对角元素 D_{ij} 等于 \mathbf{A} 中对应行的和。本文利用 GCN 作为聚合函数来更新节点特征。给定在 $l-1$ 层的节点特征矩阵 \mathbf{H}^{l-1} ,使用式(5)计算 GCN 的 l 层的特征矩阵 \mathbf{H}^l :

$$\mathbf{H}^l = \text{ReLU}(\tilde{\mathbf{A}} \mathbf{H}^{l-1} \mathbf{W}^{l-1} + \mathbf{b}^{l-1}) \quad (5)$$

其中, ReLU 是激活函数, \mathbf{H}^{l-1} 表示模型在第 $l-1$ 层的输出。需要注意的是, \mathbf{H}^0 对应输入特征矩阵 \mathbf{X} 。可训练权重矩阵和偏置向量分别表示为 \mathbf{W}^{l-1} 和 \mathbf{b}^{l-1} 。在 GCN 的第 l 层之后,最后一层的输出被用作最终的基因特征,表示为 \mathbf{H} ,它是一个 $N \times d_1$ 大小的矩阵。这里, N 代表基因的数量, d_1 代表基因特征的维度。

为了充分利用其表达能力并减小计算复杂度,由于 GCN 是基于 MLP 构建的,可以用 KAN 替换 MLP。使用式(6)来计算第 l 层:

$$h_v^l = \text{KAN}^l(h_v^{l-1}) = \sigma((1+\epsilon) \cdot h_v^{l-1} + \sum_{u \in N(v)} h_u^{l-1}) \quad (6)$$

其中, h_v^l 表示节点 v 在第 l 层的特征表示; h_v^{l-1} 表示节点 v 在前一层的特征表示; $N(v)$ 是节点 v 的邻居节点集合; ϵ 是用于调整自连接影响的小常数; σ 为激活函数。

3.4.3 Mamba 模块

状态空间序列模型和 Mamba 从线性系统中获取灵感,旨在将表示为 $x(t) \in \mathbf{R}$ 的一维函数或序列,通过隐藏空间 $h(t) \in \mathbf{R}^N$ 映射到 $y(t)$ 。在这种情况下, $\mathbf{A} \in \mathbf{R}^{(N \times N)}$ 充当演化参数,而 $\mathbf{B} \in \mathbf{R}^{(N \times 1)}$ 和 $\mathbf{C} \in \mathbf{R}^{(1 \times N)}$ 充当投影参数。该系统可以用式(7)进行数学表示:

$$h'(t) = \mathbf{A}h(t) + \mathbf{B}x(t) \quad (7)$$

$$y(t) = \mathbf{C}h'(t)$$

如图 3 所示, Mamba 模型作为连续系统的离散对应模型,引入时间尺度参数 Δ ,将连续参数 \mathbf{A} 和 \mathbf{B} 转换为它们的离散对应形式 $\tilde{\mathbf{A}}$ 和 $\tilde{\mathbf{B}}$ 。这种转换中常用的方法是零阶保持(ZOH)方法,其形式定义如下:

$$\tilde{\mathbf{A}} = \exp(\Delta \mathbf{A}) \quad (8)$$

$$\tilde{\mathbf{B}} = (\Delta \mathbf{A})^{-1} (\exp(\Delta \mathbf{A}) - 1) \cdot \Delta \mathbf{B}$$

这个线性系统的离散表示如下:

$$h_t = \tilde{\mathbf{A}}h_{t-1} + \tilde{\mathbf{B}}x_t \quad (9)$$

$$y_t = \mathbf{C}h_t$$

最后,通过全局卷积得出输出:

$$\tilde{\mathbf{K}} = (\tilde{\mathbf{C}}\mathbf{B}, \tilde{\mathbf{C}}\mathbf{A}\mathbf{B}, \dots, \tilde{\mathbf{C}}\mathbf{A}^{M-1}\mathbf{B}) \quad (10)$$

$$\mathbf{y} = \mathbf{x} * \tilde{\mathbf{K}}$$

其中, M 表示 \mathbf{x} 的序列长度, $\tilde{\mathbf{K}} \in \mathbf{R}^M$ 表示一个结构化卷积核。

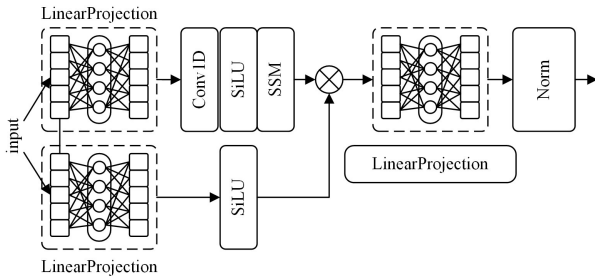


图 3 Mamba 块

Fig. 3 Mamba block

3.5 关系预测模块

获得基因特征矩阵对于预测潜在的基因调控依赖关系很有价值。KGMamba 模型对基因之间的特征向量进行点积

运算,以生成全面的基因调控依赖关系,如式(11)所示:

$$\mathbf{R} = \text{ReLU}(\mathbf{H}\mathbf{H}^T) \quad (11)$$

其中,ReLU 是激活函数;分数矩阵 \mathbf{R} 表示重建的 GRN,该矩阵的每个元素对应于一对基因的调控作用分数。

4 实验设计与结果分析

4.1 数据集

本文采用 BEELINE^[37] 提供的 7 种 scRNA-seq 数据集,涵盖人类胚胎干细胞(hESC)、人类成熟肝细胞(hHEP)、小鼠树突状细胞(mDC)、小鼠胚胎干细胞(mESC)、小鼠红系造血干细胞(mHSC-E)、小鼠粒单核细胞系造血干细胞(mHSC-GM)和小鼠淋巴细胞系造血干细胞(mHSC-L)。此外,还使用了细胞类型特异性染色质免疫共沉淀测序 ChIP-seq^[38]、非细胞类型特异性转录调控网络 ChIP-seq^[39] 以及来自 STRING 数据库^[40] 的功能相互依赖关系作为真实网络。对于 mESC 数据集,还可获得功能 LOF/GOF 真实网络^[41]。具体统计信息如表 1 所列。

表 1 实验数据

Table 1 Experimental data

细胞类型	细胞	STRING			Non-specific ChIP-seq		
		调控因子	基因	密度	调控因子	基因	密度
hESC	758	343(351)	517(709)	0.024(0.021)	301(309)	764(1152)	0.015(0.013)
hHEP	425	409(414)	656(889)	0.028(0.024)	331(339)	833(1224)	0.015(0.013)
mDC	383	264(273)	486(679)	0.038(0.032)	251(255)	642(977)	0.019(0.016)
mESC	421	495(499)	648(799)	0.024(0.021)	518(524)	895(1220)	0.015(0.013)
mHSC-E	1071	156(161)	300(427)	0.029(0.027)	147(150)	447(678)	0.022(0.019)
mHSC-GM	889	92(100)	206(357)	0.039(0.037)	91(85)	533(302)	0.028(0.03)
mHSC-L	847	39(40)	74(86)	0.047(0.045)	37(39)	169(199)	0.046(0.042)

细胞类型	细胞	STRING			Non-specific ChIP-seq		
		调控因子	基因	密度	调控因子	基因	密度
hESC	758	34(34)	815(1260)	0.164(0.166)	—	—	—
hHEP	425	30(31)	874(1330)	0.380(0.377)	—	—	—
mDC	383	20(21)	447(687)	0.085(0.082)	—	—	—
mESC	421	88(89)	975(1383)	0.345(0.378)	34(34)	775(1099)	0.159(0.154)
mHSC-E	1071	28(32)	690(1172)	0.572(0.559)	—	—	—
mHSC-GM	889	21(22)	618(1089)	0.534(0.557)	—	—	—
mHSC-L	847	16(16)	525(640)	0.525(0.507)	—	—	—

4.2 实验设置

在基准网络中记录基因对间有向调控依赖关系(正样本,标记为 1),不存在的基因对为负样本(标记为 0)。根据网络密度可知,负样本量远超正样本。随机选取已知基因对的 3/5 为正训练集,剩余 1/5 和 1/5 分别为正测试集与正验证集。考虑到负样本集的高维度,训练时按 Zhu 等^[42] 的建议将正基因对(g_1, g_2)与负对(g_1, g_3)配对,共享 g_1 ,从剩余负样本中均匀随机采样选取硬负样本,来确保各数据集正负样本比例与网络密度一致。

本文模型实验的所有参数为:训练迭代次数 $epoch = 70$;所有模型 $batch_size = 64$;学习率 $learning_rate = 0.0001$;GCN 架构为两层,隐藏层大小分别为 256 和 128;Mamba 架构配置为 12 层。尽管这些参数值是根据经验确定的,但在下文对 GCN 和 Mamba 的层数进行了参数分析。

本节结果完全基于测试集预测,为保证评估公平无偏,所有监督方法使用相同训练集和验证集,以便在相同测试集上比较模型性能,消除数据划分差异。

4.3 评价指标

将预测分类结果与实际分类结果进行比较来评价算法的有效性。针对二分类问题,采用接收者操作特征曲线下面积(Area Under the Receiver Operating Characteristic Curve, AUC)与精确召回曲线下面积(Area Under the Precision-Recall Curve, AUPR)作为评价指标。它们的具体计算式如下。

1) AUC: ROC 曲线下的面积。ROC 曲线纵轴是“真正例率”(True Positive Rate, TPR),横轴是“假正例率”(False Positive Rate, FPR)。

$$TPR = \frac{TP}{TP + FN} \quad (12)$$

$$FPR = \frac{FP}{TN + FP} \quad (13)$$

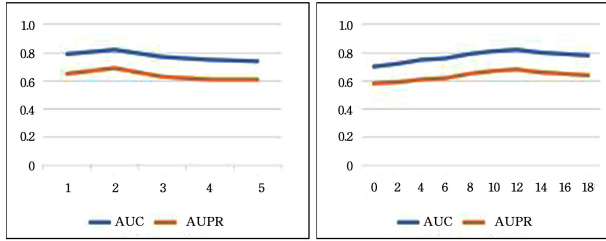
2) AUPR: PR 曲线下的面积, 用于评估模型在正负样本不平衡场景下识别正样本的能力。PR 曲线的纵坐标为查准率 P(Precision), 横坐标为查全率 R(Recall)。

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

4.4 超参数敏感性分析

GCN 与 Mamba 层数分别控制模型对图结构与全局依赖的建模能力, 是影响 KGMamba 性能的关键因素。为评估模型预测 GRN 的有效性, 在 7 个由细胞类型特异性网络组成的基准数据集上进行了评估。通过改变参数值对其进行微调, 并深入分析其对 GRN 预测的影响。7 个数据集的平均性能如图 4 所示。



(a) GCN 的层数

(b) Mamba 的层数

图 4 KGMamba 对关键结构参数的敏感性分析

Fig. 4 Sensitivity analysis of KGMamba to key structural parameters

GCN 的层数决定了 KGMamba 对基因间局部图结构信息的聚合深度, 是建模节点邻域相互作用的核心因素。为探究该参数对模型性能的影响, 将 GCN 层数系统地由 1 调整至 5。模型性能先升后降, 在 2 层时达到最优。当层数超过 2 后, 性能呈现递减趋势, 可能是由于过多的图卷积操作导致局部特征过度平滑, 甚至引入噪声干扰, 进而引发过拟合问题。

随后探究 Mamba 层数对模型捕捉基因间全局依赖关系的影响。实验中, Mamba 层数的探索范围为 0 至 18, 最佳性能出现在 12 层。继续提升层数后, 性能虽有波动但整体趋于稳定。这一现象表明, 适度增加 Mamba 层数可增强长程依赖建模能力, 但超过临界值后, 额外层数的信息增益逐渐饱和, 甚至因模型复杂度提升而产生冗余, 导致性能提升受限。

4.5 模型训练收敛性与预测稳定性分析

为验证 KGMamba 在训练过程中的收敛性与稳定性, 使用 STRING 真实 GRN 在多个数据集上进行了系统实验, 包括典型训练 Loss 分析、多数据集性能趋势观察, 以及预测结果的稳定性评估。

图 5 展示了 KGMamba 在 mESC 数据集上的训练与验证 Loss 随 epoch 的变化趋势。可见模型在前 30 到 50 轮快速下降, 第 65 轮后趋于平稳, 表明该模型训练过程稳定, 具备良好的收敛性。

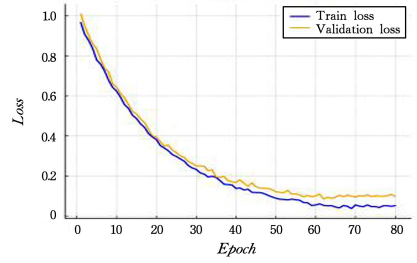


图 5 训练与验证 Loss 随 epoch 变化的趋势

Fig. 5 Training and validation loss over epochs

图 6 进一步给出了 KGMamba 在 7 个数据集上的平均性能表现。随着训练的进行, AUC 与 AUPR 均稳步提升, 在第 70 轮前后分别稳定在 0.83 与 0.36 左右, 说明模型在多种数据集上均具有稳定收敛趋势。

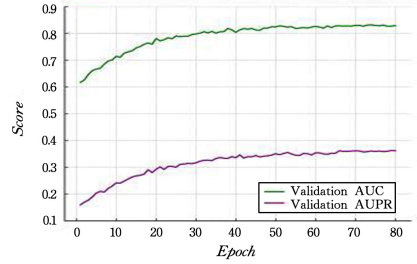


图 6 验证集 AUC 与 AUPR 随 epoch 变化的趋势

Fig. 6 Validation AUC and AUPR over epochs

为评估 KGMamba 在不同运行下的预测结果稳定性, 在 7 个数据集上分别进行了 5 次独立实验, 统计每个数据集上的 AUC 与 AUPR 均值及标准差。图 7 给出了 KGMamba 模型在各数据集上的平均性能以及误差棒表示的波动范围。结果显示, 所有数据集上的标准差均较小 ($AUC < 0.005$, $AUPR < 0.006$), 表明模型在多次训练下具有良好的稳定性与一致性。

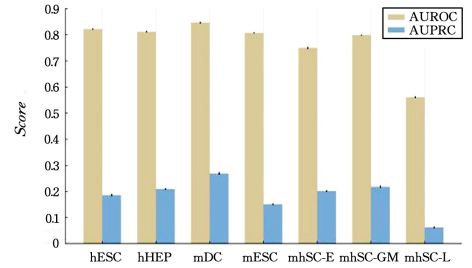


图 7 KGMamba 在 7 个数据集上预测性能的均值 ± 标准差 (AUC, AUPR)

Fig. 7 Mean \pm standard deviation of AUC and AUPR across 7 datasets

4.6 不同模型的对比结果

将 KGMamba 模型与当前主流的模型进行对比, 结果如图 8 所示, 具体数据如表 1 所列。

为评估 KGMamba 模型的性能, 利用表 1 中概述的数据集, 并通过 AUC 和 AUPR 指标评估其有效性。将 KGMamba 方法与 6 种最先进的算法进行基准测试, 包括 DGCGRN^[7], GNNLink^[43], GNE^[44], CNNC^[21], STGRNS^[24] 和 GRN-Transformer^[23]。

图 8(a) 和图 8(b) 分别给出了各模型获得的 AUC 和

AUPR 值,结果表明,KGMamba 模型在 scRNA-seq 数据集上始终优于所有基线方法的性能。KGMamba 获得的 AUC 值平均比次优模型 DGCGRN 高出约 3.87%。同时,KGMamba 在 44 个基准数据集上获得的 AUPR 值平均比次优模型 DGCGRN 高出约 14.31%。此外,在所有基准测试中,KGMamba 更是大幅超越了 GNE, CNNC, STGRNS 和 GRN-

Transformer。

为了评估 KGMamba 与其他模型在 AUC 和 AUPR 上的统计显著性,对不同模型的结果使用 Wilcoxon 符号秩检验进行了统计分析。如表 2 所列,KGMamba 与所有对比模型之间均存在显著差异($p < 0.05$),说明其性能提升具有统计学支持,差异并非由偶然因素引起。

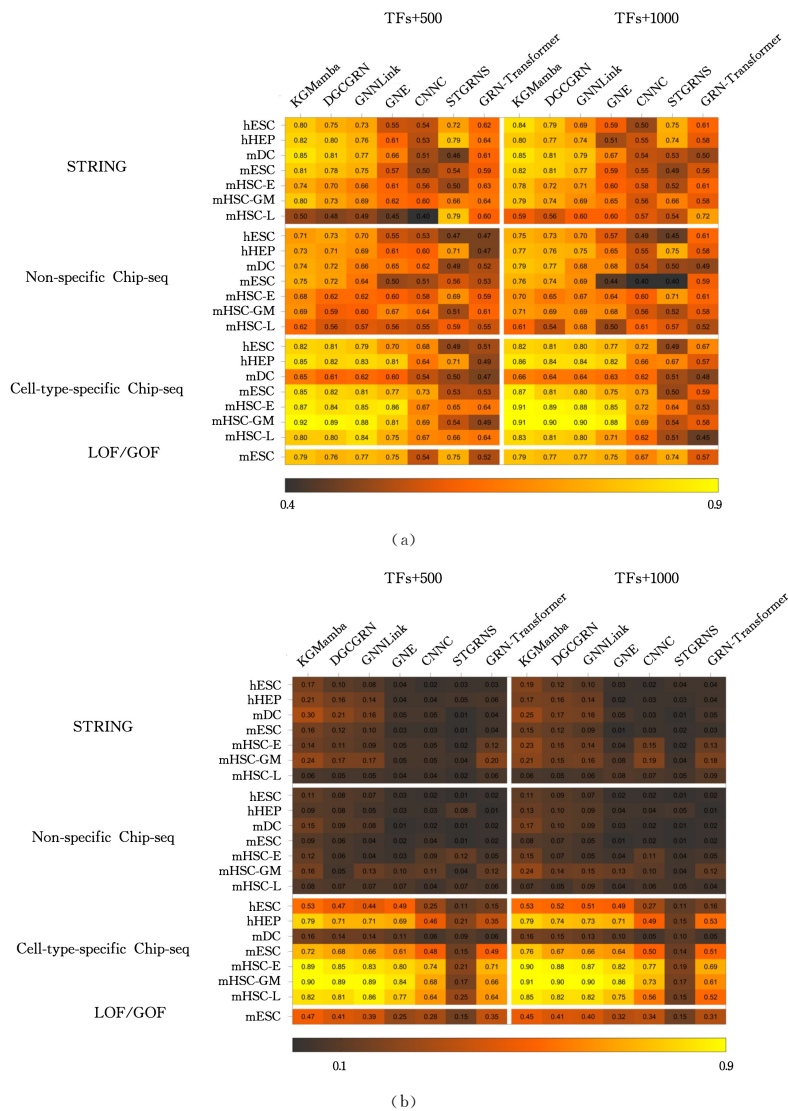


图 8 各模型的对比

Fig. 8 Comparison of each model

表 2 KGMamba 与其他模型的 AUC 和 AUPR Wilcoxon 显著性检验结果

Table 2 Wilcoxon significance test results for AUC and AUPR between KGMamba and other models

	DGCGRN	GNNLink	GNE	CNNC	STGRNS	GRN-Transformer
P 值(AUC)	0.000105	0.000015	0.000047	0.000114	0.000146	0.000033
P 值(AUPR)	0.000058	0.000161	0.000041	0.000057	0.000059	0.000061

4.7 KGMamba 模型的鲁棒性

本文模型依赖于 KAN 驱动的 GCN 和 Mamba 来有效捕捉基因的表达特征,进一步探索训练数据规模对监督模型性能的影响,即特征是否能被有效提取。在 7 个不同的数据集上进行实验,这些数据集包含细胞类型特异性的 GRN 真实来源。为了比较不同方法的鲁棒性,从监督训练数据中随机抽取 30%,40%,50%和 60%的数据。图 9

给出了 DGCGRN 模型在细胞类型特异性 GRN 中,从 30%到 60%的训练数据集规模下的性能。正如预期,GRN 预测时的性能通常随着监督训练数据量的增加而提升。值得注意的是,当使用 50%的训练数据时,KGMamba 和 DGCGRN 方法的性能都达到了稳定状态。与 DGCGRN 方法相比,KGMamba 也实现了更优越的性能,AUC 平均提高 4.16%,AUPR 平均提高 6.24%。这些改

进突显了 KGMamba 即使是利用小的训练数据集,在 GRN

推断中也取得了不错的结果。

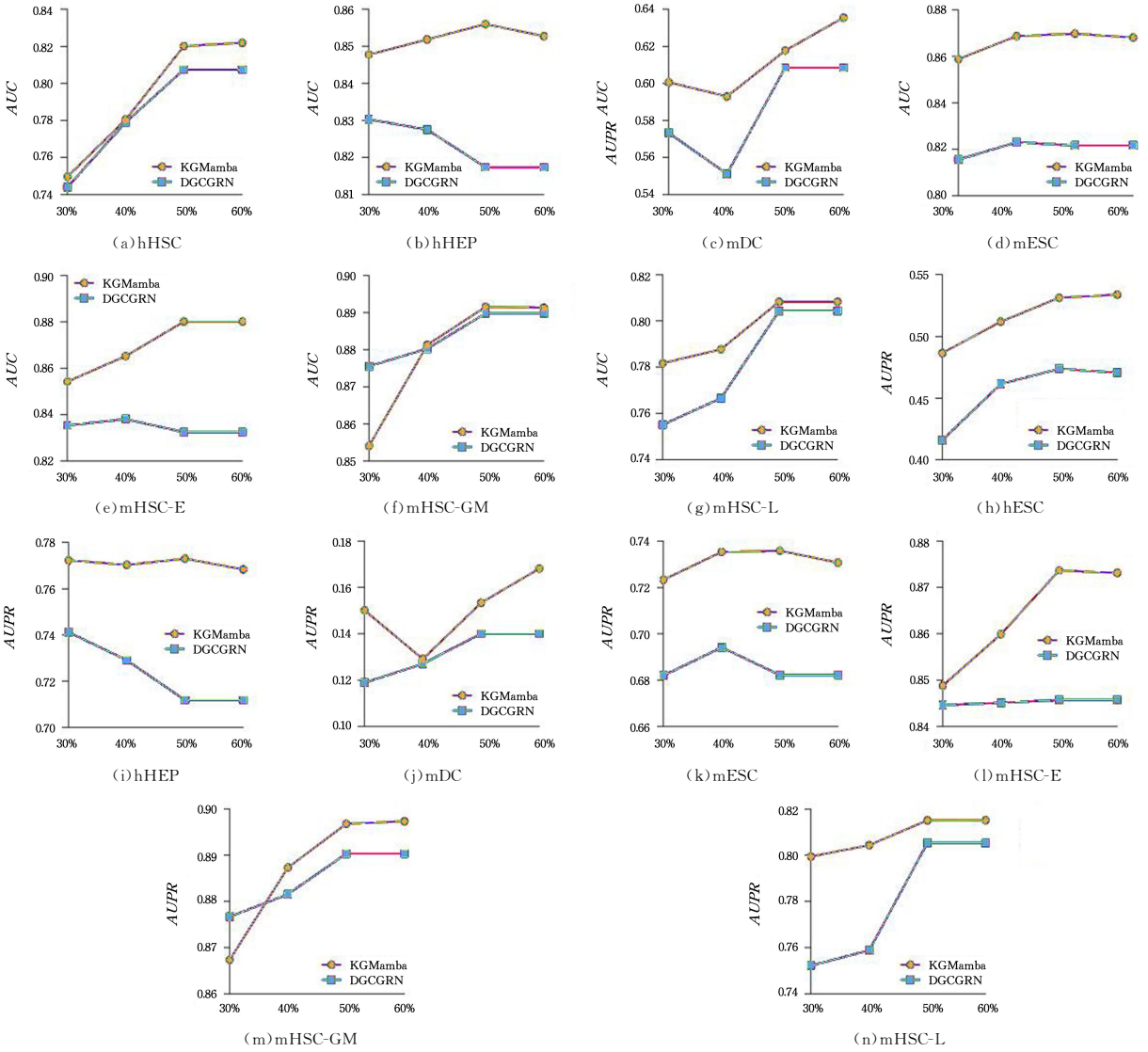


图 9 训练集规模对模型的影响

Fig. 9 Impact of training set size on the model

4.8 数据稀疏性对 KGMamba 模型的影响

在单细胞基因表达数据中,常常因基因表达检测不完整而产生零值,致使许多基因被赋予零表达水平。为解决这一问题,近年来已开发出各种先进的计算方法,用于估计 scRNA-seq 数据中缺失的零值。在本研究中,利用 DeepImpute^[45]对原始数据集中的零值进行插补,然后通过测量 7 个具有细胞类型特异性网络的原始数据集的 AUC 和 AUPR,

来评估插补的效果。

KGMamba 方法在每个数据集上执行 10 次的结果如图 10 所示。可以看出,与未处理数据相比,利用 DeepImpute 方法估计原始单细胞数据集中零元素的 KGMamba 方法,展现出与其相当的预测性能,表明 scRNA-seq 数据中零值的存在对 MKGRN 方法的预测准确性的影响微乎其微。

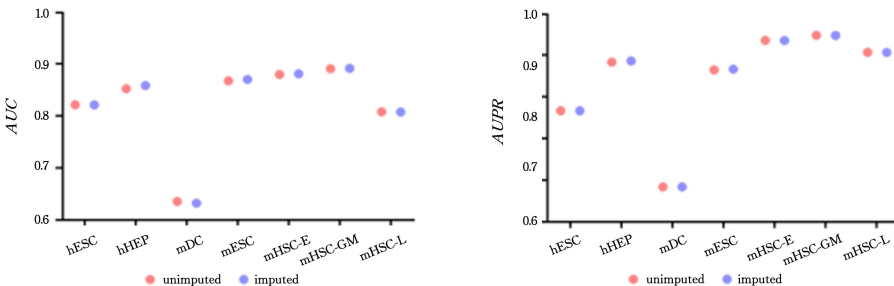


图 10 数据中的零值对模型的影响

Fig. 10 Impact of zero values in the data on the model

4.9 模型的计算复杂度分析

本文研究了每种方法在 7 个 scRNA-seq 数据集上的运行时间。表 3 列出了每种方法在 7 个具有细胞类型特异性网络的不同大小的单细胞数据集上的平均运行时间。KGMamba 在 TFs+500 和 TFs+1000 数据集上展现出较短的运行时间。值得注意的是,本文方法在 TFs+500 数据集上的运行时间排名第一,在 TFs+1000 数据集上则位居第二。这是因为 KGMamba 采用 GCN 作为交互图编码器,仅设置两层 GCN 隐藏层,参数量少,属于轻量级模型架构,这种简化设计

表 3 实验模型的运行时间比较

Table 3 Running time comparison of experimental models

时间	KGMamba	GNNLink	DGCGRN	GRN-Transformer	GNE	STGRNS	CNNC
TFs+500	4.29 s	4.4 s	4 min 45 s	5 min 34 s	6 min 12 s	20 min 51 s	1 h 10 min
TFs+1000	10.21 s	8.59 s	15 min 6 s	27 min 41 s	34 min 18 s	57 min 15 s	18 h 41 min

4.10 消融实验

为评估 KGMamba 模型中每个模块的贡献,使用 STRING 数据设计并进行了两组消融实验。第一组实验研究了移除关键模块对模型性能的影响。结果如图 11 所示,在 AUC 指标上,KGMamba 达到了最高值,为 0.8,其次是 KGCN,为 0.76,单独使用 Mamba 为 0.7。在 AUPR 指标上,KGMamba 的表现也最佳,达到 0.17,而单独使用 KGCN 和 Mamba 分别约为 0.15 和 0.11。这些结果表明,KGCN 和 Mamba 模块的集成显著提升了模型的性能。移除任一模块都会导致性能下降,表明两个模块都发挥着关键作用。

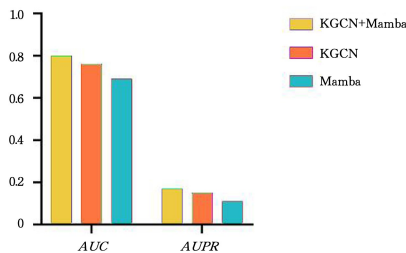


图 11 各模块对模型的影响

Fig. 11 Impact of each module on the model

第二组实验探索了使用传统 MLP 层和所提出的 KAN 之间的性能差异。结果如表 4 所列,使用 KAN 的模型的运行时间短于使用 MLP 的模型。这表明,与传统 MLP 层相比,KAN 可以降低模型的计算复杂度。

表 4 Kolmogorov-Arnold 网络对模型的影响

Table 4 Impact of Kolmogorov-Arnold network on the model

Methods	Time/s
KAN	24.52
MLP	29.23

4.11 案例分析

在 GRN 推断中,评估模型是否能够重现已有的生物学调控机制对于验证其预测可靠性至关重要。为此,基于 KGMamba 推断了 mESC 调控网络,研究了两个关键多能性调控因子 Pou5f1 和 Sox2 的上游调控因子,相关结果如图 12 所示。值得注意的是,这些调控关系均已被 TRRUSTv2^[46] 数据库

使其在运行速度上具备天然优势。而 KGMamba 为实现更精准的基因调控关系建模,在模型结构与功能设计上更为复杂,例如将 KAN 融入图卷积网络的特征映射过程以保持非线性表达并降低时间复杂度,同时引入 Mamba 进行全局依赖建模以增强对长程调控关系的捕捉能力,这种对性能的优化在一定程度上增加了计算开销。尽管如此,KGMamba 仍以相对较低的计算成本实现了更优的预测性能,在复杂生物网络推断的准确性与算法效率之间取得了较好的平衡,展现了其在实际应用中的潜力。

验证,表明 KGMamba 能挖掘出具有生物学意义的调控通路。

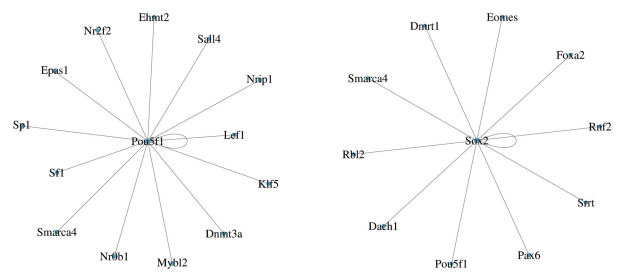


图 12 KGMamba 推断网络中 Pou5f1 和 Sox2 的调控因子

Fig. 12 Regulatory factors of Pou5f1 and Sox2 in the network inferred by KGMamba

此外,提取了连接靶基因数量最多的前 20 个调控因子,并对其预测的下游目标基因进行了 GO 功能注释,重点关注其在“干细胞群体维持”(GO:0019827)中的富集情况。如图 13 所示,在这 20 个调控因子中有 11 个(包括 Pou5f1, Prdm14, Klf4, Nanog, Sox2, Esrrb, Tbx3, Zfp281, Sall4, Stat3, Myc)被 GO 功能注释为与“干细胞群体维持”(GO:0019827)相关,表明 KGMamba 能够识别出在干细胞多能性维持中发挥核心作用的关键调控因子。

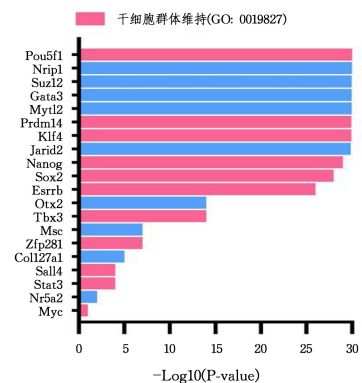


图 13 前 20 个调控因子的预测靶基因富集分析

Fig. 13 Enrichment analysis of predicted target genes of the top 20 regulators

结束语 本文提出了 KGMamba 模型,通过引入 KAN 对 GCN 进行轻量化改造,并结合 Mamba 模块增强全局依赖

建模能力,有效提升了基因调控网络预测的准确性与效率。在公开数据集上进行的系统评估表明,该模型在 AUC 和 AUPR 上均优于现有主流方法,且具有良好的训练稳定性与计算效率。此外,通过对预测结果中连接靶基因最多的调控因子进行 GO 功能注释分析,验证了模型对多能性维持相关关键调控因子的识别能力,进一步增强了结果的生物学解释性。

尽管本文方法取得了较为显著的实验结果,但仍存在一些局限性值得深入探讨。1)模型仅依赖 scRNA-seq 数据构建 GRN,尚未融合 ATAC-seq,ChIP-seq 等多组学信息,导致对复杂调控机制的刻画仍有局限;2)模型尚未在异质性更强的跨物种或跨细胞类型背景下进行迁移泛化测试,泛化能力仍有待提升;3)尽管通过富集分析验证了预测结果的生物学相关性,但模型仍缺乏结构化、可追溯的机制解释能力,难以直接揭示调控通路或因果结构。此外,在更大规模数据下的资源消耗与推理效率仍有进一步优化空间。

针对上述问题,未来研究可从以下几个方向展开:1)引入多模态数据融合机制,提升模型对真实调控关系的建模能力;2)探索迁移学习、小样本适应等策略,增强模型在多样化生物背景下的适应性与泛化性;3)结合图可解释学习框架,提升模型对关键调控路径的结构化解释能力,拓展其在机制发现与实验辅助中的应用价值;4)进一步精简网络结构或采用稀疏计算策略,以实现更高效的推理过程,满足大规模组学数据的实际需求。

参 考 文 献

- [1] NGUYEN H, TRAN D, TRAN B, et al. A comprehensive survey of regulatory network inference methods using single cell RNA sequencing data [J]. *Briefings in Bioinformatics*, 2021, 22(3): bbaa190.
- [2] MARBACH D, COSTELLO J C, KÜFFNER R, et al. Wisdom of crowds for robust gene network inference [J]. *Nature Methods*, 2012, 9(8): 796-804.
- [3] RICE J J, TU Y, STOLOVITZKY G. Reconstructing biological networks using conditional correlation analysis [J]. *Bioinformatics*, 2005, 21(6): 765-773.
- [4] FAITH J J, HAYETE B, THADEN J T, et al. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles [J]. *PLoS Biology*, 2007, 5(1): 54-66.
- [5] FRIEDMAN N. Inferring cellular networks using probabilistic graphical models [J]. *Science*, 2004, 303(5659): 799-805.
- [6] JI R, GENG Y, QUAN X. Inferring gene regulatory networks with graph convolutional network based on causal feature reconstruction [J]. *Scientific Reports*, 2024, 14(1): 21342.
- [7] WEI P J, GUO Z, GAO Z, et al. Inference of gene regulatory networks based on directed graph convolutional networks [J]. *Briefings in Bioinformatics*, 2024, 25(4): bbae309.
- [8] KOMMU S, WANG Y, WANG Y, et al. Gene Regulatory Network Inference with Joint Representation from Graph Neural Network and Single-Cell Foundation Model [J/OL]. <https://www.biorxiv.org/content/10.1101/2024.12.16.628715v1.full.pdf>.
- [9] LIU Z, WANG Y, VAIDYA S, et al. Kan: Kolmogorov-Arnold networks [J]. *arXiv:2404.19756*, 2024.
- [10] GU A, DAO T. Mamba: Linear-time sequence modeling with selective state spaces [J]. *arXiv:2312.00752*, 2023.
- [11] YANG B, XU Y, MAXWELL A, et al. MICRAT: a novel algorithm for inferring gene regulatory networks using time series gene expression data [J]. *BMC Systems Biology*, 2018, 12: 19-29.
- [12] LI L, SUN L, CHEN G, et al. LogBTF: gene regulatory network inference using Boolean threshold network model from single-cell gene expression data [J]. *Bioinformatics*, 2023, 39(5): btad256.
- [13] HUYNH-THU V A, GEURTS P. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data [EB/OL]. *Scientific Reports*, 2018, 8: 3384. <https://www.nature.com/articles/s41598-018-21715-0>.
- [14] MARGOLIN A A, NEMENMAN I, BASSO K, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context [C]// *BMC Bioinformatics*. 2006: 1-15.
- [15] ZENG Y, HE Y, ZHENG R, et al. Inferring single-cell gene regulatory network by non-redundant mutual information [J]. *Briefings in Bioinformatics*, 2023, 24(5): bbad326.
- [16] ZHANG Y, WANG M, WANG Z, et al. MetaSEM: gene regulatory network inference from single-cell RNA data by meta-learning [J]. *International Journal of Molecular Sciences*, 2023, 24(3): 2595.
- [17] MA M Y, SUN J X, HU C L. Modeling Gene Regulatory Networks with Global Coupling Parameters [J]. *Computer Science*, 2023, 50(S2): 138-144.
- [18] HUYNH-THU V A, IRRTHUM A, WEHENKEL L, et al. Inferring regulatory networks from expression data using tree-based methods [J]. *PLoS One*, 2010, 5(9): e12776.
- [19] LI Z J, LIAO S, LIU A F, et al. Gene Association Analysis Algorithm for GeneRegulatory Network [J]. *Computer Engineering and Applications*, 2025, 61(3): 155-165.
- [20] ZHAO M, HE W, TANG J, et al. A hybrid deep learning framework for gene regulatory network inference from single-cell transcriptomic data [J]. *Briefings in Bioinformatics*, 2022, 23(2): bbab568.
- [21] YUAN Y, BAR-JOSEPH Z. Deep learning for inferring gene relationships from single-cell expression data [J]. *Proceedings of the National Academy of Sciences*, 2019, 116(52): 27151-27158.
- [22] CHEN J, CHEONG C W, LAN L, et al. DeepDRIM: a deep neural network to reconstruct cell-type-specific gene regulatory network using single-cell RNA-seq data [J]. *Briefings in Bioinformatics*, 2021, 22(6): bbab325.
- [23] SHU H, DING F, ZHOU J, et al. Boosting single-cell gene regulatory network reconstruction via bulk-cell transcriptomic data [J]. *Briefings in Bioinformatics*, 2022, 23(5): bbac389.
- [24] XU J, ZHANG A, LIU F, et al. STGRNS: an interpretable transformer-based method for inferring gene regulatory networks from single-cell transcriptomic data [J]. *Bioinformatics*,

- 2023,39(4):btad165.
- [25] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017;1025-1035.
- [26] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. arXiv:1710.10903, 2017.
- [27] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks? [J]. arXiv:2105.14491, 2021.
- [28] SOMVANSHI S, JAVED S A, ISLAM M M, et al. A survey on Kolmogorov-Arnold network [J]. arXiv:2411.06078, 2024.
- [29] SU Z, CHEN M, AI J, et al. Research on Recommendation Algorithm Based on Kolmogorov-Arnold Network-driven Scoring and Review Fusion [J]. Journal of Chinese Computer Systems. 2025,46(11):2600-2609.
- [30] CHEON M. Kolmogorov-Arnold network for satellite image classification in remote sensing [J]. arXiv:2406.00600, 2024.
- [31] JIANG W, CHEN T, GAO X, et al. Epidemiology-informed network for robust rumor detection [C]//Proceedings of the ACM on Web Conference 2025. 2025;3618-3627.
- [32] XU K, CHEN L, WANG S. Kolmogorov-Arnold Networks for Time Series: Bridging Predictive Power and Interpretability [J]. arXiv:2406.02496, 2024.
- [33] GU A, GOEL K, RÉ C. Efficiently modeling long sequences with structured state spaces [J]. arXiv:2111.00396, 2021.
- [34] SMITH J T H, WARRINGTON A, LINDERMAN S W. Simplified state space layers for sequence modeling [J]. arXiv:2208.04933, 2022.
- [35] MEHTA H, GUPTA A, CUTKOSKY A, et al. Long range language modeling via gated state spaces [J]. arXiv:2206.13947, 2022.
- [36] QU H, NING L, AN R, et al. A survey of Mamba [J]. arXiv:2408.01129, 2024.
- [37] PRATAPA A, JALIHAI A P, LAW J N, et al. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data [J]. Nature Methods, 2020, 17(2):147-154.
- [38] MOORE J E, PURCARO M J, PRATT H E, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes [J]. Nature, 2020, 583(7818):699-710.
- [39] GARCIA-ALONSO L, HOLLAND C H, IBRAHIM M M, et al. Benchmark and integration of resources for the estimation of human transcription factor activities [J]. Genome Research, 2019, 29(8):1363-1375.
- [40] SZKLARCZYK D, GABLE A L, LYON D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets [J]. Nucleic Acids Research, 2019, 47 (D1): D607-D613.
- [41] XU H, BAROUKH C, DANNENFELSER R, et al. ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells [J]. Database, 2013, 2013; bat045.
- [42] ZHU X, JING X Y, ZHANG F, et al. Distance learning by mining hard and easy negative samples for person re-identification [J]. Pattern Recognition, 2019, 95:211-222.
- [43] MAO G, PANG Z, ZUO K, et al. Predicting gene regulatory links from single-cell RNA-seq data using graph neural networks [J]. Briefings in Bioinformatics, 2023, 24(6): bbad414.
- [44] KC K, LI R, CUI F, et al. GNE: a deep learning framework for gene network inference by aggregating biological information [J]. BMC Systems Biology, 2019, 13:1-14.
- [45] ARISDAKESSIAN C, POIRION O, YUNITS B, et al. DeepImpute: an accurate, fast, and scalable deep neural network method to impute single-cell RNA-seq data [J]. Genome Biology, 2019, 20:1-14.
- [46] HAN H, CHO J W, LEE S, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions [J]. Nucleic Acids Research, 2018, 46 (D1): D380-D386.



GAO Tai, born in 2000, postgraduate, is a member of CCF (No. Z6724G). His main research interests include deep learning and bioinformatics.



WANG Bin, born in 1986, Ph.D, professor, Ph.D supervisor. His main research interests include deep learning, brain imaging research, medical images and bioinformatics.

(责任编辑:何杨)