



基于类相关图卷积网络的单细胞蛋白质定位方法*

唐浩漾** 姚欣悦 王濛濛 杨思聪

(西安邮电大学自动化学院, 西安 710100)

摘要 **目的** 随着生物信息学的发展,涌现出大量基于深度学习的蛋白质亚细胞定位方法。这些蛋白质亚细胞定位方法(如GapNet-PL、ImPLOC等)能够较准确识别细胞群体水平的蛋白质分布模式,但在单细胞水平或复杂微环境下的定位仍存在局限性。当前蛋白质显微图像缺乏单细胞标注,仅依赖细胞群体水平的标注无法解析单细胞尺度的定位异质性,且现有大多数蛋白质亚细胞定位模型基于卷积神经网络(CNN)设计,忽略了亚细胞结构间的功能关联性,导致单细胞蛋白质亚细胞定位精度差。因此,本文提出一种基于类相关图卷积网络(CP-GCN)的单细胞蛋白质定位方法。**方法** 首先,建立类相关模块(CPM),充分提取不同亚细胞类别的语义特征。然后,设计CP-GCN网络,挖掘多细胞中蛋白质亚细胞的全局特征并捕获标签图的拓扑信息,学习多标记蛋白质类别之间的关联性。接下来,利用K-means聚类方法区分内多尺度特征,生成多细胞类激活图(CAM),根据CAM的预测区域,对单细胞图像进行伪标注,以有效区分细胞群中的异质性细胞。最后,使用伪标注训练单细胞蛋白质分类模型,实现单细胞蛋白质的精准定位。**结果** 在Kaggle2021数据集的单细胞蛋白质预测任务中,该方法的mAP指标均优于现有的蛋白质亚细胞定位方法。对生成的CAM结果进行可视化分析,证明了模型可成功定位单细胞内蛋白质亚细胞。**结论** 通过CP-GCN网络与伪标签分配策略相结合,可以有效地捕捉蛋白质图像中的异质性细胞的特征,精确定位单细胞内蛋白质位置。

关键词 单细胞蛋白质亚细胞定位, 弱监督, 图卷积网络, 类相关, 伪标签

中图分类号 TP391

DOI: 10.3724/j.pibb.2025.0047

CSTR: 32369.14.pibb.20250047

蛋白质作为生物功能的基本执行者,分布于多种亚细胞结构中执行不同的功能^[1],蛋白质定位研究是通过蛋白质染色图像观察蛋白质在亚细胞结构上的分布。蛋白质在细胞内并不是静止不动的,核糖体合成蛋白质后,被转运到不同的亚细胞中执行相应功能,并参与细胞的各种生命活动。因此,确定目标蛋白质所处的亚细胞位置对研究蛋白质功能、细胞过程和揭示疾病发展机制至关重要。当前,蛋白质定位方法主要是对同一细胞群进行蛋白质亚细胞定位,存在定位不准确的问题。随着单细胞技术的发展,研究发现,在同一群细胞中,蛋白质定位和功能等方面也存在差异,尤其当蛋白质覆盖多个亚细胞结构时,这种异质性更加显著^[2]。单细胞内蛋白质定位是深入探究细胞异质性及其功能多样性的重要途径,为揭示细胞在特定条件下的功能状态和调控机制提供有力支持。

但是,由于细胞存在异质性,蛋白质分布存在

差异,单细胞内蛋白质模式通常具有高度的复杂性和异质性,数据分析难度较高,目前仅有少量工作在单细胞水平上进行了研究。且现有的关于蛋白质定位工作的数据集主要来源于人类蛋白质图谱(human protein atlas, HPA)项目,HPA中缺乏单细胞内蛋白质亚细胞的标注信息^[3]。传统分类模型难以捕获图像中每个单细胞内的蛋白质特征,导致对无标注样本的预测准确性较低。有研究^[4]提出了一种基于卷积神经网络(convolutional neural networks, CNN)的方法,利用低层网络特征对应基本图像特征,深层网络特征分离定位类别,将2D和3D的免疫荧光(immuno fluorescence, IF)图像分割成单细胞图像,实现了细胞层级预测。也

* 陕西省自然科学基金(2025JC-YBMS-697)资助项目。

** 通讯联系人。

Tel: 029-88166460, E-mail: tanghaoyang@xupt.edu.cn

收稿日期: 2025-01-27, 接受日期: 2025-06-20

有文献提出了一个基于多实例学习框架的蛋白质亚细胞定位模型,利用无监督学习算法分配伪标签训练模型,结合了图像级和细胞级的单细胞蛋白质预测结果^[5]。但上述工作均利用CNN提取特征并进行分类,卷积层仅能捕获局部视觉特征,无法获得亚细胞之间关联性特征。

蛋白质亚细胞定位需要区别蛋白质在多种亚细胞位置上,本质是多标记分类问题,单标签预测框架难以建模生物复杂性,准确预测单细胞内蛋白质分布。针对多标签图像识别问题,Chen等^[6]提出了一种基于图卷积网络(graph convolutional network, GCN)的多标签分类模型,通过标签之间的依赖关系建模。对于高质量标注数据不足的问题,为了能够在弱监督环境下准确预测未标注样本,现在常采用两种策略。一是基于多实例学习范式训练端到端弱监督模型,如Zhu等^[7]提出了一种多实例学习框架,建模每张图像中多个细胞的蛋白质分布,但这种策略常基于独立同分布假设,没有考虑来自实例间的相关性。二是利用分类模型生成类激活映射(class activation mapping, CAM)作为伪标签,CAM是一种能够在缺乏物体位置和类别标签的监督信息的情况下,可视化深度学习模型在分类任务中关注图像区域的重要性的方法。Li等^[8]提出线上简单示例挖掘方法,通过学习CNN生成CAM和伪标签,Xie等^[9]通过自适应区域生长来细化梯度类激活映射生成的定位区域。

本文提出一种基于类相关图卷积网络(class perception-graph convolutional networks, CP-GCN)的单细胞内蛋白质亚细胞定位方法,实现对多标记蛋白质图像在单细胞水平上的定位。本文设计了CP-GCN网络,提取图像中的蛋白质语义特征,捕获多标记蛋白质亚细胞间的动态拓扑关系,同时为后续细胞级预测生成深层细胞特征,生成的细胞特征应用K-means聚类方法,利用聚类结果生成CAM,并结合细胞掩码生成伪标签,搭建单细胞蛋白质分类模型,实现单细胞内蛋白质的精准定位。

1 单细胞内蛋白质图像的定位方法

本文提出的单细胞蛋白质定位方法分为图像级定位预测和细胞级定位预测两个步骤(图1)。图像级定位预测的输入为整幅IF图像,通过CNN网络提取浅层蛋白质特征,构建类相关(class perception module, CPM)模块,能够区分与标签

对应的多种蛋白质亚细胞定位区域,提取含有类别信息的深层蛋白质语义特征。同时将亚细胞共现信息作为先验知识,GCN网络用于捕获亚细胞结构之间的关联性。最后融合深层亚细胞共现嵌入特征与蛋白质语义特征,得到的多源融合特征通过分类层进行预测,最终得到图像层级的预测结果。

细胞级定位预测是通过基于U-net^[10]的预训练分割模型将整幅图像中分割出的多个单细胞和利用CAM生成的单细胞对应伪标签作为输入,用于构建模型定位单细胞内蛋白质分布。首先,通过细胞分割模型将图像划分为单个细胞,并利用K-means算法对图像级预测中捕获的深层特征进行聚类。该聚类过程为优化CAM提供了结构化的信息,通过对CAM中高亮显示的区域进行伪标注,提高细胞级预测的准确性,伪标签用于构建单细胞分类模型。图像级预测提供了整体的蛋白质分布信息,而细胞级预测则对单个细胞进行了精细化分析,最后通过将图像级预测和细胞级预测结果的置信度相乘,得到每个单细胞内蛋白质亚细胞定位最终预测置信度。

1.1 基于CP-GCN网络的图像级蛋白质定位预测

本文设计的CP-GCN网络完成图像级预测,结构如图1中所示,包含3个主要部分:CNN特征提取主干网络,类别感知模块CPM,捕获标签相关性的GCN网络。选择DenseNet121^[11]和ResNet50^[12]作为CNN特征提取网络,提取多细胞蛋白质图像中的深层特征,并由CPM模块生成包含类别信息的特征表示,输入到GCN网络中,捕获全局蛋白质亚细胞特征。

1.1.1 CPM模块

由于蛋白质数据集中缺乏单细胞标注信息,传统CNN网络仅能捕获训练图像中的最显著特征,模型感知不同尺度目标能力较弱。本文设计CPM模块,对CNN提取的多个细胞特征进行深层语义感知,获得蛋白质在图像级亚细胞器中的语义特征表示。首先通过类激活映射技术定位背景的语义区域,同时引入在线注意力累积(online accumulation attention, OAA)策略优化CAM,然后利用CAM和提取的深层特征生成类别表示,准确描述输入图像中的特定亚细胞器标签相关的内容。

如图2所示,通过CNN提取输入图像的深层特征 $X \in R^{H \times W \times C}$,根据特征 X 计算类别特定的激活图 $M = [M_1, M_2, \dots, M_c] \in R^{H \times W \times D}$,再将转换后的

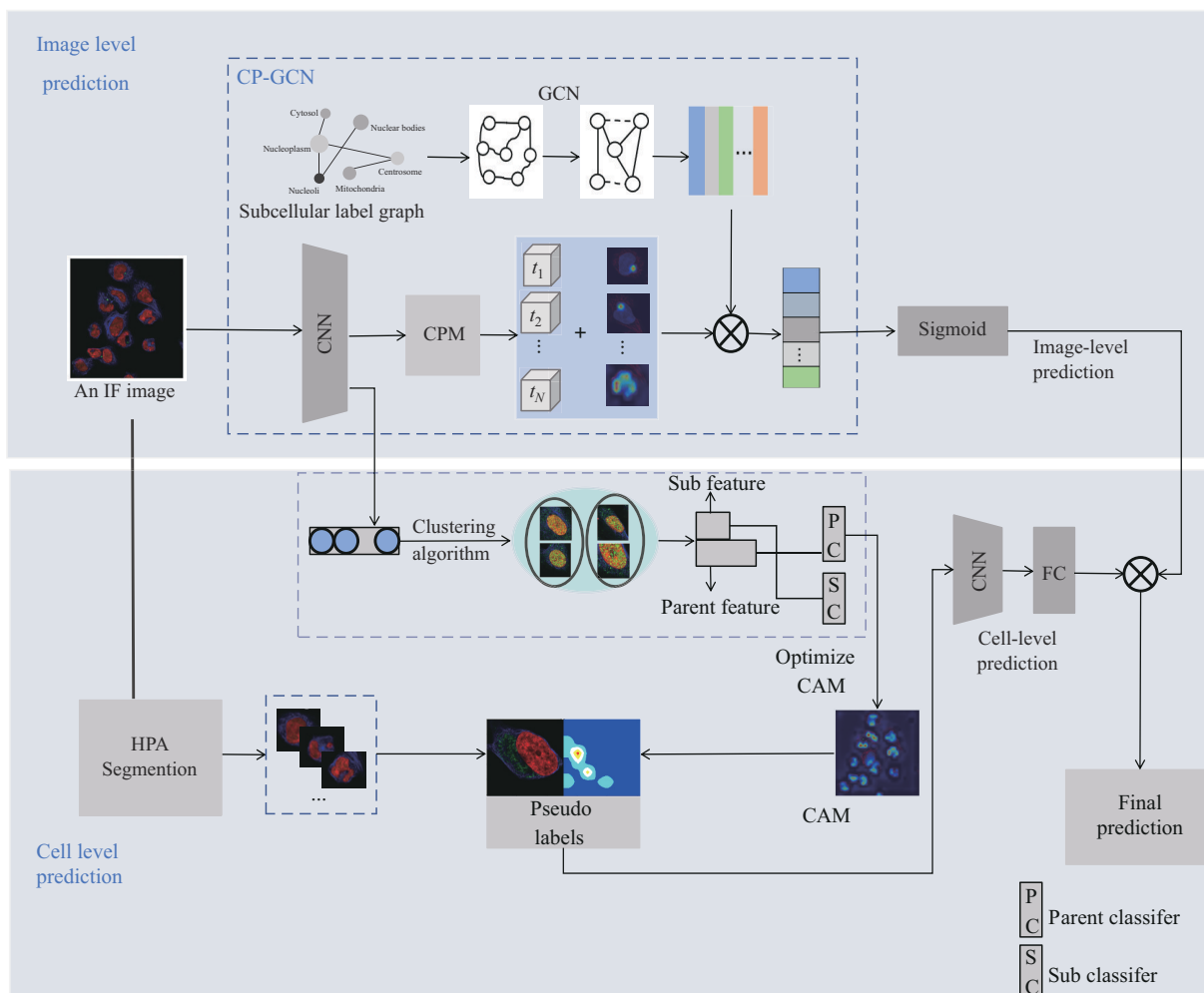


Fig. 1 Flowchart of subcellular protein localization methods in single cells

特征图 $X' \in R^{(H \times W) \times C'}$ 与激活图的加权, 按标签选择性聚合特征, 生成对应标签的内容感知类别表示 $T = [t_1, t_2, \dots, t_c] \in R^{D \times C'}$:

$$t_c = M_c^T X' = \sum_{i=1}^H \sum_{j=1}^W M_{ij}^c x'_{ij} \quad (1)$$

在分类网络收敛之前, 不同训练阶段定位到的目标区域会在语义对象各部分间漂移, 导致基于激活图生成的伪标签质量不佳。为了在训练阶段细化激活图, 引入了一种 OAA 策略^[13], 完善每个图像中每个类别的累积激活图。具体而言, 针对训练图像中的每个目标类别 c , 构建一个累积激活图 M_c , 用以存储识别出的有区分性区域。OAA 在首个训练阶段, 使用类别 c 的激活图 A_1 初始化累积激活图 M_1 , 当图像再次输入分类网络时, 采取逐元素最大操作融合 M_1 与最新激活图 A_2 , 更新累积激活图, 其公式如下:

$$M_i = \max(M_{i-1}, A_i) \quad (2)$$

式中, A_i 表示最新激活图, M_{i-1} 表示当前累积激活

图。由于分类模型可能在训练过程开始时关注噪声区域, 因此使用目标类别的预测概率大于设定阈值来决定是否积累响应激活图。

1.1.2 GCN网络模块

采用 GloVe 将来自标签的 18 个亚细胞单词 (如细胞核、线粒体) 映射到实数的向量空间, 生成 18 个词向量 $V_i \in R^{1 \times D} (i = \{1, 2, \dots, 18\})$, 每个类别标签的词向量通过查询 GloVe 的预训练词表生成, 语义相似的词在向量空间中距离相近, 然后将 18 个词向量填充到词向量矩阵 $W_0 \in R^{C \times D}$, 其中 C 为标签数量, D 为词向量的维度。利用 GCN 捕获直接相邻亚细胞标签的局部关联模式以及建模高阶亚细胞标签之间的依赖关系。首先统计训练数据集中标签对 (i, j) 的出现次数, 得到矩阵 $B_{ij} \in R^{C \times C}$, 通过公式 (3) 得到共现概率矩阵:

$$P_{ij} = \frac{B_{ij}}{N_i} \quad (3)$$

N_i 代表训练集中标签 i 的出现次数, P_{ij} 表示标签 i

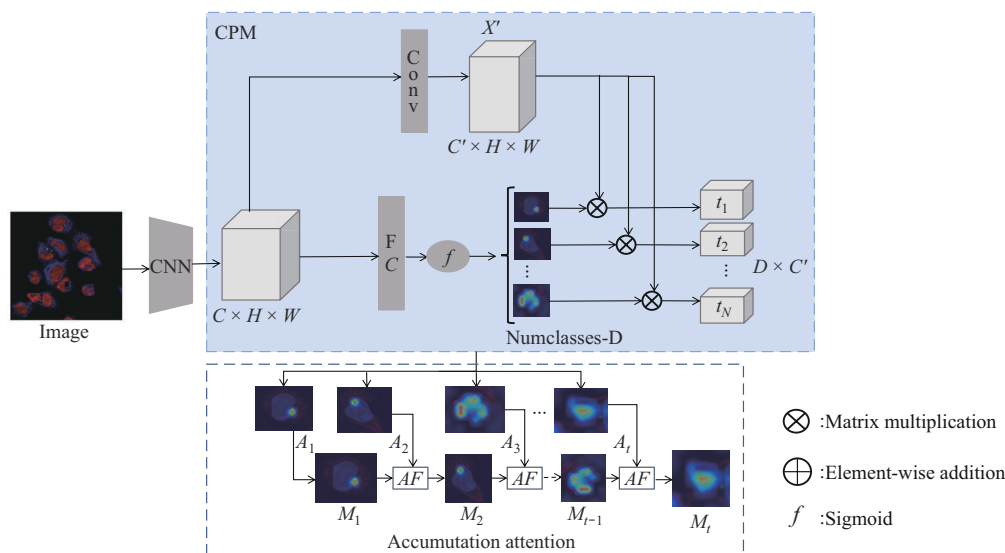


Fig. 2 CPM module

AF: activation function.

出现时, 标签 j 同时出现的概率。然而, 根据公式(3)得到的共现概率可能呈现长尾分布, 训练集和测试集中的标签共现绝对值可能存在差异。因此, 需要对共现概率矩阵进行二值化处理, 并通过设置阈值来过滤噪声边缘, 具体操作为:

$$Z_{ij} = \begin{cases} 0, & \text{if } P_{ij} < \tau \\ 1, & \text{if } P_{ij} > \tau \end{cases} \quad (4)$$

Z_{ij} 是相关性系数矩阵, τ 是阈值, 设置为0.3。

通过图卷积操作捕捉标签之间的关系, 获得更加深层的类别关联表示。GCN是一种多层网络结构, 每个卷积层仅处理节点的一阶邻域信息, 即每个节点只关注其直接相邻节点的数据。通过叠加多个卷积层, GCN能够实现多阶邻域的信息传播, 使每个节点逐渐融合更远邻的上下文信息。在进行图卷积操作之前, 邻接矩阵 A 会加上自环并进行归一化处理 $A = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}}$, 以确保数值稳定性和信息传播的有效性。在前向传播过程中, 特征从输入层开始, 经过图卷积层的运算处理, 并通过Softmax激活函数生成预测的分类概率分布, 用于进行分类。在一个两层的GCN中, 激活函数分别采用ReLU和Softmax, 网络的正向传播的公式为:

$$L_2 = f_{\text{softmax}}(A f_{\text{ReLU}}(AXW^0)W^1) \quad (5)$$

X 是每个亚细胞节点的特征值, A 是节点间的关系, 即邻接矩阵, W 为网络的权重。图卷积操作可以有效地在图结构中传播和整合节点的信息, 最终生成的亚细胞关系嵌入特征能够描述亚细胞之间的深层依赖关系。然后将图像类别特征 T 和标签共现嵌入

特征 L_2 通过逐元素乘积操作进行跨模态融合:

$$S = T \cdot (L_2)^T \quad (6)$$

特征 S 描述了各类别的丰富语义信息, 对每个类别单独应用Sigmoid激活函数, Sigmoid输出表示该每个亚细胞位置蛋白质独立存在的概率, 最终获得蛋白质亚细胞图像级预测结果。

1.2 基于CAM和迁移学习的细胞级蛋白质定位预测

1.2.1 基于K-means聚类优化的CAM

基于K-means聚类算法优化的类激活图过程如图1所示, 首先提取CP-GCN网络CNN模块生成深层细胞特征, 使用K-means算法对细胞特征进行聚类, 增强分类器对细胞类别内部及类别间差异的识别能力, 通过多轮自训练聚类图像特征, 能够提高基于图像特征生成的CAM精度, 进而提高伪标签质量。设定每个类别中的聚类子类别数为 K , 并随机选择 K 个数据点作为初始聚类中心。随后, 计算剩余数据点到这些聚类中心的距离, 并按照距离最小值原则分配数据点, 迭代更新直至各簇中的数据点稳定不变。最终, 根据聚类结果, 为每个图像分配伪标签, 作为其子类的索引。获得子类别伪标签后, 联合CNN特征提取网络和父类、子类分类器网络进行自训练, 使特征和聚类结果相互增强。通过子类别伪标签增强特征, 使得同类特征更加聚集。增强后的特征促进聚类过程, 生成更准确的子类别伪标签。网络训练过程迭代进行, 直到父类分类器生成的类激活映射图精度不再上升。

1.2.2 伪标签生成

在使用CAM生成伪标签时, 先归一化CAM并调整其大小与输入图像一致。然后, 利用分割模型识别单个细胞, 将每个细胞与CAM相交。接着为每个类别初始化累加器和计数器, 分别累积类别激活值和记录像素点被细胞覆盖的次数。CAM提取细胞实例区域, 按缩放比例调整大小后, 生成单个细胞的CAM与预测概率。累加器累计类别激活值, 并更新计数器记录像素点覆盖次数。结合预测概率确定单个细胞的伪标签。

1.3 融合图像级和细胞级预测结果的蛋白质定位预测

图像级预测提供了整体的蛋白质分布信息, 反映出图像中不同细胞之间蛋白质的整体分布模式和趋势。细胞级预测则聚焦于单个细胞内的蛋白质分布情况, 捕捉到更为细致和局部的定位信息。通过将图像级和细胞级预测结果的置信度相乘, 将两者的优势有效融合, 在多细胞图像中形成对单细胞内蛋白质亚细胞定位的完整预测。

$$C_{\text{final}} = C_{\text{image}} \cdot C_{\text{cell}} \quad (7)$$

C_{image} 表示图像级预测结果的置信度, C_{cell} 表示细胞级预测结果的置信度。

2 实验结果与分析

2.1 数据集

本文实验所用数据集来自HPA数据集和2021年Kaggle举办的人类单细胞分类比赛^[14]数据集。构建训练集时采用了来自HPA和Kaggle两个数据集, 并删除损坏和重复样本筛选出54 322张高质量图像, 每张图像都有对应的图像层级标注信息, 涵盖18个亚细胞类别。

测试集采用Kaggle数据集中的公共测试集和私有测试集。公共测试集中包含559张IF图像, 其中异质性细胞较少, 私有测试集中含有1 217张图像, 相比于公共测试集, 私有测试集中图像样本里单细胞更难识别, 异质性细胞群较多。通过Kaggle测试集可以得到模型在18种亚细胞位置预测的整体性能。此外, 还采用Zhu等^[8]构建了一个包含细胞级标注的测试集, 以便更全面地评估和分析模型的单个亚细胞位置上的预测性能。

2.2 评价指标

采用多尺度评估体系评估模型性能。对每张图像的类别预测概率进行排序, 通过计算不同置信度阈值下的精度(Precision)和召回率(Recall)绘

制P-R曲线, 通过求P-R曲线下的面积, 得到18种类别的平均精度(average precision, AP), 然后取均值得到图像级均值平均精度(mean average precision, mAP); 对细胞级预测结果则匹配预测与真实掩膜的交并比(intersection over union, IoU), IoU阈值设为0.6, 通过计算不同置信度阈值下, 所有匹配细胞对中每个类别的真正例(true positive, TP)、假正例(false positive, FP)、假反例(false negative, FN)得到Precision和Recall, 然后绘制P-R曲线, 计算曲线面积得到不同类别的单细胞AP值, 取平均最终获得细胞级mAP。

2.3 实验环境

所有实验均在Ubuntu系统上进行, 使用RTX 3090 GPU进行训练和推理, PyTorch作为深度学习框架, 通过CUDA12.1加速训练, 编程语言为Python3.8。

图像级预测模型的训练, 样本输入CP-GCN网络, 批量大小为8, 采用Adam作为优化器, 初始学习率为0.001, 训练进行了50个epoch, 采用Focal loss损失函数在图像级预测模型训练过程中解决类别不平衡问题。细胞级预测模型的训练, 利用伪标签作为监督, 训练得到具有单细胞预测功能的网络模型。使用预训练ResNet50作为主干网络, epoch设置为20, 批次大小为32, 其他训练参数与图像级预测模型一致。优化CAM模型的训练, K值设置为10, 迭代次数设置为3, 批量大小为8, 初始学习率为0.001, epoch设置为20。

2.4 图像级蛋白质定位预测结果分析

为了验证网络的有效性, 本文在蛋白质亚细胞数据集上分别对DenseNet121、ResNet50、Inceptionv3^[15]、ML-GCN^[6]和本文的CP-GCN模型进行定位性能对比。图3是不同模型在18种亚细胞类别的AP分数的比较, 分析AP评分可得, 数据集的不平衡并未对模型性能产生显著影响, 但不同亚细胞模式的识别难度导致了不同类别蛋白质定位准确率差异。例如囊泡数量远超中间纤维(约40倍), 然而中间纤维的AP分数却更高, 这是因为中间纤维结构稳定, 而囊泡作为细胞内的膜性结构, 形态多变, 易与其他亚细胞结构混淆, CP-GCN模型通过结合蛋白质的局部视觉特征(如形态、纹理)和全局拓扑关系(如亚细胞结构共现模式), 能够缓解形态多样的样本较难识别的问题。对18个亚细胞位置的AP值相加并取均值得到mAP值, 其中DenseNet121、ResNet50、Inceptionv3和

ML-GCN 网络分别为 77.80%、79.42%、81.47%、81.05%，本文的 CP-GCN 模型定位性能最佳，mAP 值为 83.57%。为了验证 CPM、GCN 网络模块对蛋白质亚细胞定位任务中的影响，进行了消融实验。首先仅使用 CNN 预测蛋白质亚细胞位置，然后逐步加入 CPM、GCN 模块，保持其他训练参数不变，用 mAP 值作为标准依次验证各个模块性能

对比，消融实验的结果如表 S1 所示。可以看出，仅采用 CNN 时 mAP 值为 0.795 5，加入 CPM 模块捕获蛋白质亚细胞不同空间位置上的语义特征，mAP 值上升了 1.46%。加入 GCN 模块，图注意力网络在捕获蛋白质亚细胞间关系发挥了重要作用，mAP 值提高了 2.56%。

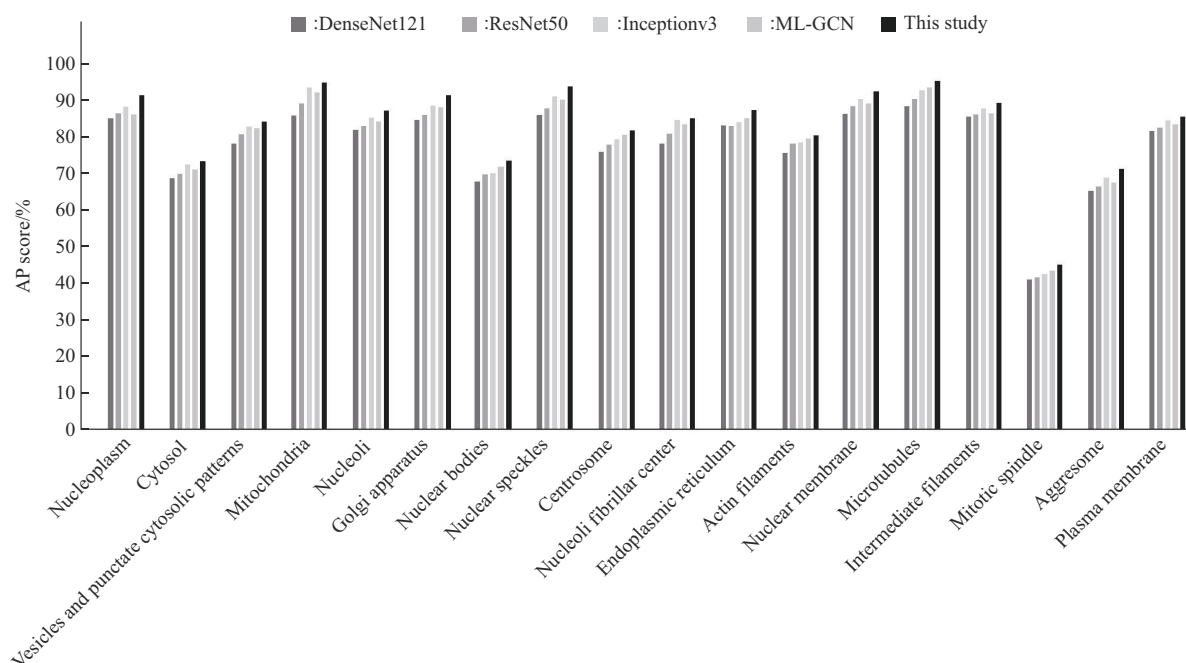


Fig. 3 Comparison of AP scores across subcellular localization categories

如图4所示，采用 Grad-CAM^[16] 进行可视化，在定位图中突出显示目标类别预测的特征。本文提出的网络能有效识别该类蛋白质的亚细胞类别。第一组图像展示了一张多标签 IF 图像，图像中蛋白质同时出现在核质（nucleoplasm）和有丝纺锤体（mitotic spindle）两个亚细胞结构中，通过生成的两张 CAM 图，展示了所预测的图像类别和捕获的类别特征。第二组展示的蛋白质样本定位在中心体（centrosome）和核质（nucleoplasm）两种亚细胞器中。本文提出的 CP-GCN 网络能准确区分不同亚细胞位置模式，保证了细胞级预测模型所获取的伪标签的准确性。

2.5 细胞级蛋白质定位预测结果分析

2.5.1 优化 CAM 模型实验结果

首先研究 K-means 聚类算法的聚类子类别数 K 对网络模型性能的影响，进行了关于 K 值的消融研究。如表 1 所示，对比了未使用聚类方法 ($K=1$) 与所提出方法使用不同聚类数 $K = \{5, 10, 20\}$ 的

F1-score。结果表明，所提出方法在不同 K 值变化范围内表现出稳健性，并且始终优于未使用 K-means 聚类的方法 ($K=1$)。这些结果验证了引入聚类信息对于网络学习和增强特征表达的必要性和重要性。考虑到计算效率和准确性，最终在所有实验中选择 $K=10$ 作为聚类数默认值。

图 5 展示了在初始轮次和第三轮生成的 CAM 的结果对比。从图中可以看出，多次聚类 and 迭代训练对同类蛋白质的不同状态或形态特征进行了细粒度区分和学习，这种细粒度学习增加了网络的学习难度，但也显著提高了类激活图的质量，为后续单细胞伪标签的分配提供了可靠的依据。

2.5.2 单细胞蛋白质定位模型实验结果

为了验证伪标签生成质量，将基于伪标签构建的单细胞蛋白质亚细胞定位模型与基于图像级标签构建的单细胞蛋白质亚细胞定位模型进行对比实验，基于图像级标签构建的单细胞模型直接使用图像级标签作为细胞级标签，在 Annotated Test Set 上

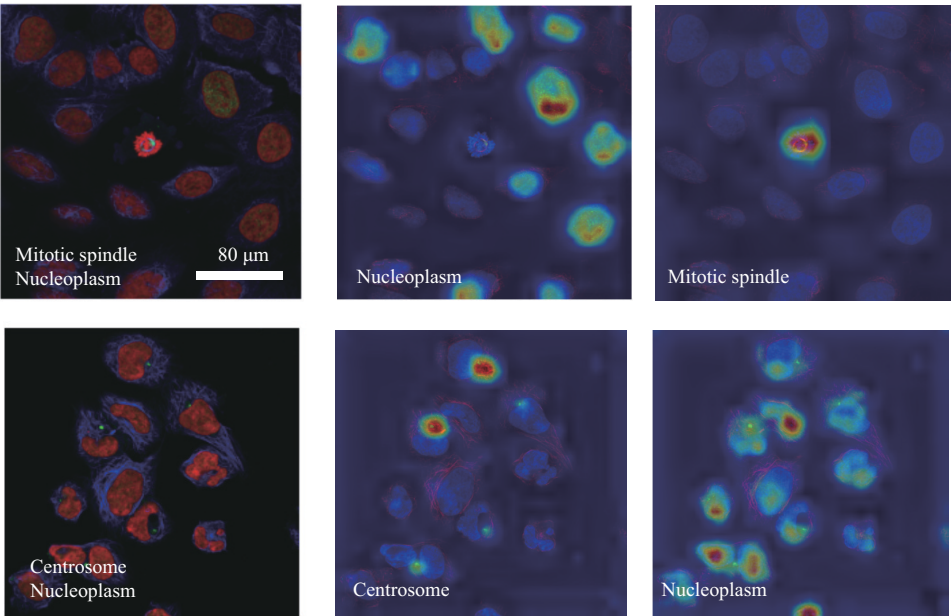


Fig. 4 Class activation maps for multi-label IF image

The class activation maps for mitotic spindle, nucleoplasm, and centrosome highlighted the corresponding subcellular regions exhibiting these expression patterns in the images. These activation maps comprehensively marked all cells bearing the correct labels, demonstrating that the image-level model can effectively discriminate between distinct protein expression patterns.

Table 1 F1-scores under different K values

K	F1-score
1	0.632
5	0.668
10	0.679
20	0.670

对比了两种模型在 10 种亚细胞位置上的 AP 分数，模型均基于预训练的 ResNet50 框架，实验结果见表 2，分析表中数据可得图像级别的标签不足以代表每个细胞的标签。因此，本文通过 CAM 生成伪标签，并利用伪标签构建了单细胞蛋白质分类模型。

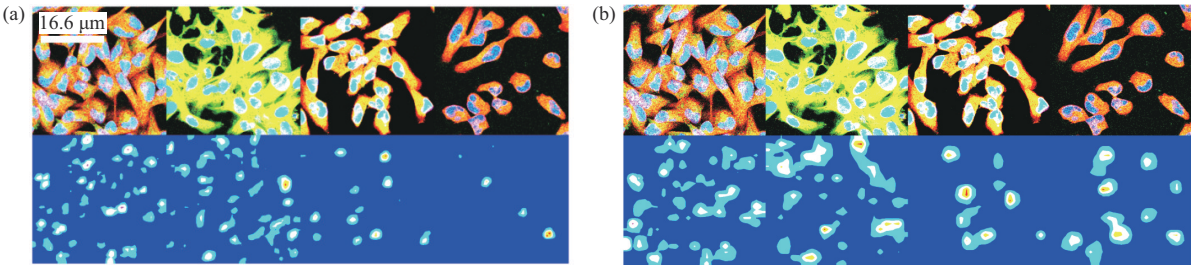


Fig. 5 Visualization of CAM in the first (a) and third (b) epoch

After iterative optimization training, the generated class activation maps demonstrate significantly improved spatial coverage.

Table 2 AP scores of different single-cell protein localization models across 10 subcellular locations on the Annotated Test Set

Class	Nucleoplasm	Nuclear membrane	Nucleoli	Nucleoli fibrillar center	Nuclear speckles	Nuclear bodies	Endoplasmic reticulum	Golgi apparatus	Mitochondria	Aggresome
Image label model	0.689	0.595	0.812	0.606	0.561	0.543	0.762	0.639	0.781	0.842
Pseudo label model	0.718	0.629	0.863	0.647	0.612	0.613	0.850	0.680	0.859	0.822

表2中结果显示, 基于伪标签训练的模型能实现比图像标签训练的模型更好的结果。但由于样本的图像级标注信息与实际图像中某些单细胞内的蛋白质表达不一致, 在生成单细胞CAM过程会存在粗糙且有噪声, 例如在含有聚集体标签的部分样本中, 图像中仅有少量细胞内蛋白质的在聚集体结构中表达, 与标注一致, 剩余细胞中蛋白质未表达, 这会导致生成伪标签过程中引入了噪声信息, 从而影响了部分细胞的伪标签准确性。但相比于大部分基于图像级标签构建的的单细胞模型的预测结果改善较高, 因此结合CAM的方法可以为单细胞图像分配更准确的标签。

2.6 融合预测结果分析

CP-GCN网络模型提供图像级预测结果, 然后基于伪标签训练的网络模型则提供细胞级预测结果, 最后通过将图像级预测和细胞级预测结果融合, 在测试集上进行验证, 采用的对比方法是当前单细胞蛋白质定位工作的最优模型公平激活细胞网络 (fair cell activation network, FCAN), 实验结果表明, 本文方法实现了更精准的单细胞内蛋白质亚细胞位置预测。

对比结果如表3所示, 私有测试集图像中细胞之间的异质性更显著, 本文模型在该测试集上的验证结果比当前最优模型具有更高的准确性, 且模型在私有测试集上的最终 mAP 值仅比公开测试集上

Table 3 mAP score of the final ensemble model

Model	Public test set	Private test set
Image level prediction	0.554 1	0.544 0
Cell level prediction	0.539 6	0.522 3
Fusion final prediction	0.567 6	0.558 2
FCAN ^[15]	0.569 5	0.554 1

的结果低 0.014 1, 而最优模型在私有测试集与公开测试集上的结果差异为 0.029 0, 本文模型适应性强, 面对不同数据分布时表现鲁棒, 单细胞分类性能优于现有方法。模型性能较优在于提出的CP-GCN网络和分配单细胞伪标签的组合方法, 首先CP-GCN网络能够较好地捕获类别标签之间的关系, 私有测试集上最终 mAP 值预测结果为 0.544 0, 进而通过生成伪标签扩大了单细胞之间的差异, 提高了CP-GCN模型在测试集上的预测性能, 因此结合两个模型在私有的测试集上的最终预测结果为 0.558 2, 均比单一模型预测结果的性能更好。

为了判断模型在预测时关注的图像区域是否与生物学相关的亚细胞位置有关, 使用Grad-CAM算法来可视化对预测重要的区域。图6a为单个细胞的类激活映射图, 热图完全覆盖了图像中的蛋白质通道, 并且能够突出相应类别的特征区域, 因此模型在单细胞中能够完整识别目标区域, 捕捉的局部特征具有明确的生物学意义。如图6b所示, 单细胞内蛋白质覆盖多个亚细胞区域, 生成的多个

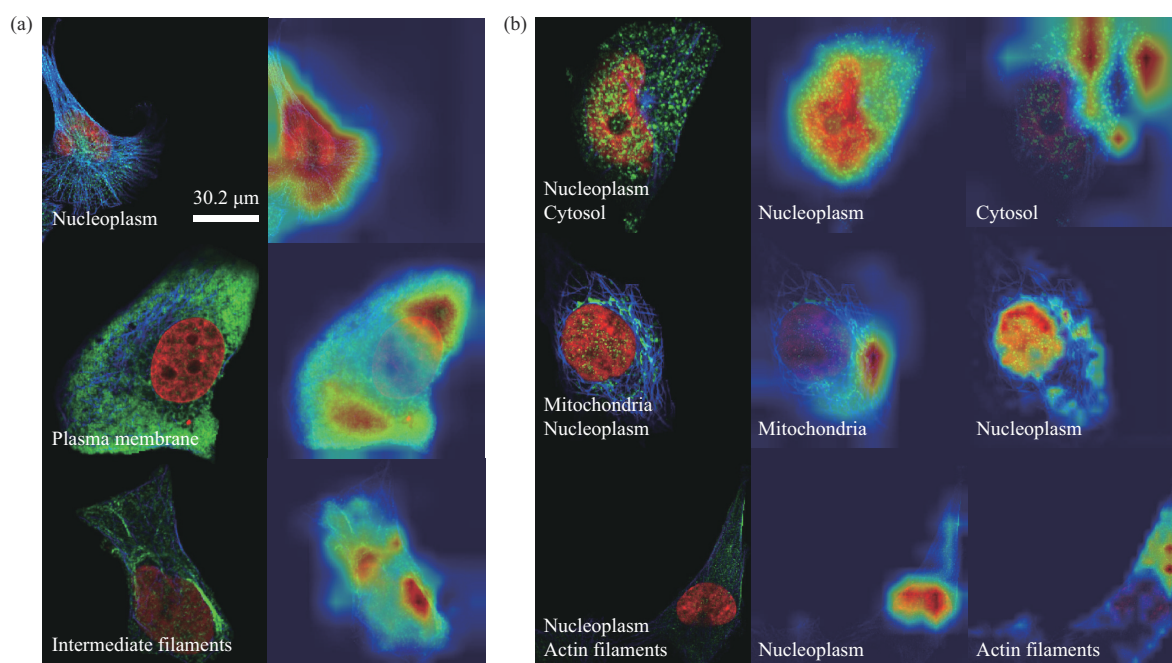


Fig. 6 The CAM for single cells can effectively highlight distinctive protein feature patterns of specific subcellular categories
(a) single-label; (b) multi-label.

CAM表明模型能准确区分不同亚细胞位置模式。

3 结 论

针对单个细胞内蛋白质难以精确定位的问题, 本文提出一种全局类别感知和捕捉标签关联的图像级弱监督单细胞蛋白质亚细胞定位方法, 在图像级标签的不完全约束下实现单细胞内蛋白质精准定位。设计CP-GCN网络建立了图像中蛋白质亚细胞之间的长距离依赖关系, 同时伪标签生成策略提高了对单个细胞生成伪标签的准确度。实验结果表明, 本方法相较于现有单细胞蛋白质亚细胞定位方法在HPA数据集上展现出显著的定位效果和鲁棒性提升。尽管本文提出的伪标签生成策略能够有效标注单细胞内蛋白质亚细胞位置, 但图像中存在蛋白质弱表达区域, 模型难以识别从而影响伪标签生成质量。未来可通过开发基于教师-学生框架的渐进式伪标签优化方法, 通过提高模型对弱表达区域的预测一致性, 从而增强模型对弱表达区域的感知能力, 优化伪标签生成方法。综上所述, 本研究为蛋白质亚细胞定位技术提供了一种创新的解决方案, 有助于加速未知蛋白质的表征和对细胞功能的理解, 从而提高对疾病相关表型的认识并推动药物研究的发展。

附件 见本文网络版 (<http://www.pibb.ac.cn>, <http://www.cnki.net>):
PIBB_20250047_Table_S1.pdf

参 考 文 献

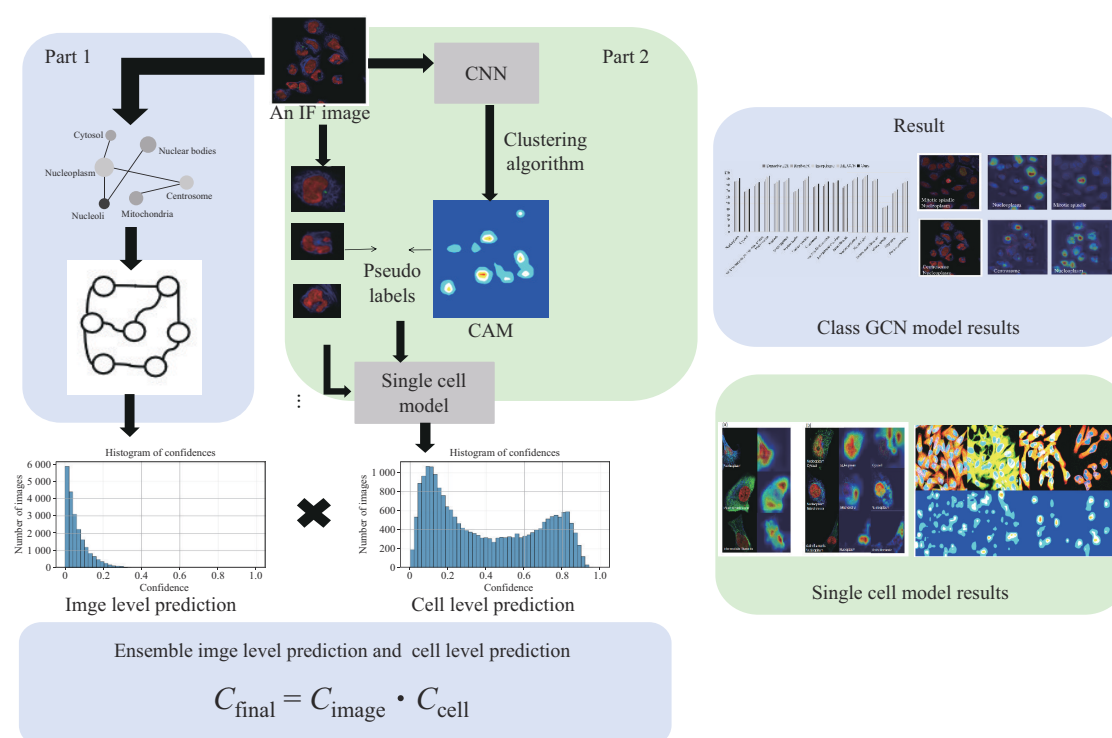
- [1] Uhlen M, Fagerberg L, Hallström M B, *et al.* Tissue-based map of the human proteome. *Science*, 2015, **347**(6220):1260419
- [2] Christopher J A, Stadler C, Martin C E, *et al.* Subcellular proteomics. *Nat Rev Methods Prim*, 2021, **1**(1): 32
- [3] Barbe L, Lundberg E, Oksvold P, *et al.* Toward a confocal subcellular atlas of the human proteome. *Mol Cel Proteomics*, 2008, **7**(3): 499-508
- [4] Pärnamäa T, Parts L. Accurate classification of protein subcellular localization from high-throughput microscopy images using deep learning. *G3 (Bethesda)*, 2017, **7**(5): 1385-1392
- [5] Su R, He L, Liu T, *et al.* Protein subcellular localization based on deep image features and criterion learning strategy. *Brief Bioinform*, 2021, **22**(4): bbaa313
- [6] Chen Z M, Wei X S, Wang P, *et al.* Multi-label image recognition with graph convolutional networks//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 5172-5181
- [7] Zhu X L, Bao L X, Xue M Q, *et al.* Automatic recognition of protein subcellular location patterns in single cells from immunofluorescence images based on deep learning. *Brief Bioinform*, 2023, **24**(1): bbac609
- [8] Li Y, Yu Y, Zou Y, *et al.* Online easy example mining for weakly-supervised gland segmentation from histology images//Navab N, Hornegger J, Wells W M, *et al.* International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer Nature Switzerland, 2022: 578-587
- [9] Xie Y, Zhang Z, Chen S, *et al.* Detect, Grow, Seg: a weakly supervision method for medical image segmentation based on bounding box. *Biomed Signal Process Control*, 2023, **86**: 105158
- [10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation//Navab N, Hornegger J, Wells W M, *et al.* Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). Cham: Springer, 2015: 234-241
- [11] Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 2261-2269
- [12] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778
- [13] Zhou B, Khosla A, Lapedriza A, *et al.* Learning deep features for discriminative localization//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2921-2929
- [14] Le T, Winsnes C F, Axelsson U, *et al.* Analysis of the human protein atlas weakly supervised single-cell classification competition. *Nat Methods*, 2022, **19**(10): 1221-1229
- [15] Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the inception architecture for computer vision//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2818-2826
- [16] Selvaraju R R, Cogswell M, Das A, *et al.* Grad-CAM: visual explanations from deep networks via gradient-based localization//IEEE. 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 618-626

Single-cell Protein Localization Method Based on Class Perception Graph Convolutional Network*

TANG Hao-Yang**, YAO Xin-Yue, WANG Meng-Meng, YANG Si-Cong

(College of Automation, Xi'an University of Posts & Telecommunications, Xi'an 710100, China)

Graphical abstract



Abstract Objective This study proposes a novel single-cell protein localization method based on a class perception graph convolutional network (CP-GCN) to overcome several critical challenges in protein microscopic image analysis, including the scarcity of cell-level annotations, inadequate feature extraction, and the difficulty in achieving precise protein localization within individual cells. The methodology involves multiple innovative components designed to enhance both feature extraction and localization accuracy. **Methods** First, a class perception module (CPM) is developed to effectively capture and distinguish semantic features across different subcellular categories, enabling more discriminative feature representation. Building upon this, the CP-GCN network is designed to explore global features of subcellular proteins in multicellular environments. This network

* This work was supported by a grant from Natural Science Foundation of Shaanxi Province (2025JC-YBMS-697).

** Corresponding author.

Tel: 86-29-88166460, E-mail: tanghaoyang@xupt.edu.cn

Received: January 27, 2025 Received: June 20, 2025

incorporates a category feature-aware module to extract protein semantic features aligned with label dimensions and establishes a subcellular relationship mining module to model correlations between different subcellular structures. By doing so, it generates co-occurrence embedding features that encode spatial and contextual relationships among subcellular locations, thereby improving feature representation. To further refine localization, a multi-scale feature analysis approach is employed using the K-means clustering algorithm, which classifies multi-scale features within each subcellular category and generates multi-cell class activation maps (CAMs). These CAMs highlight discriminative regions associated with specific subcellular locations, facilitating more accurate protein localization. Additionally, a pseudo-label generation strategy is introduced to address the lack of annotated single-cell data. This strategy segments multicellular images into single-cell images and assigns reliable pseudo-labels based on the CAM-predicted regions, ensuring high-quality training data for single-cell analysis. Under a transfer learning framework, the model is trained to achieve precise single-cell-level protein localization, leveraging both the extracted features and pseudo-labels for robust performance. **Results** Experimental validation on multiple single-cell test datasets demonstrates that the proposed method significantly outperforms existing approaches in terms of robustness and localization accuracy. Specifically, on the Kaggle 2021 dataset, the method achieves superior mean average precision (mAP) metrics across 18 subcellular categories, highlighting its effectiveness in diverse protein localization tasks. Visualization of the generated CAM results further confirms the model's capability to accurately localize subcellular proteins within individual cells, even in complex multicellular environments. **Conclusion** The integration of the CP-GCN network with a pseudo-labeling strategy enables the proposed method to effectively capture heterogeneous cellular features in protein images and achieve precise single-cell protein localization. This advancement not only addresses key limitations in current protein image analysis but also provides a scalable and accurate solution for subcellular protein studies, with potential applications in biomedical research and diagnostic imaging. The success of this method underscores the importance of combining advanced deep learning architectures with innovative training strategies to overcome data scarcity and improve localization performance in biological image analysis. Future work could explore the extension of this framework to other types of microscopic imaging and its application in large-scale protein interaction studies.

Key words single-cell protein subcellular localization, weak supervision, graph convolutional network, class perception, pseudo label

DOI: 10.3724/j.pibb.2025.0047

CSTR: 32369.14.pibb.20250047