化学所发展基于深度学习的蛋白质单分子分析新方法

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/5273.html

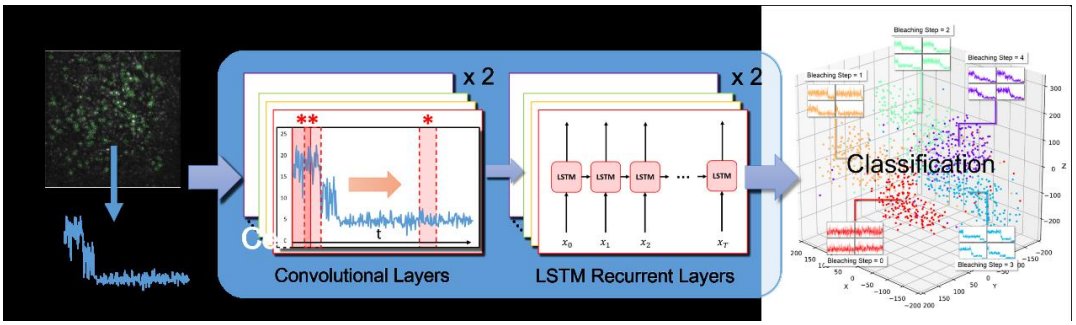
本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

化学所发展基于深度学习的蛋白质单分子分析新方法。蛋白质是生命活动的物质基础和主要承担者，许多重要的蛋白质以复合物或多聚体形式参与信号转导、离子转运、免疫响应等众多生理过程，蛋白质的化学计量组成与其生物功能的调控及多种疾病的发生发展密切相关。因此，在生理条件下定量表征蛋白质的化学计量比(亚基组成数或蛋白聚集状态)，对于研究蛋白质的相互作用、阐明蛋白质结构-功能的调控机制、研发新药等都具有十分重要的意义。

在国家自然科学基金委、科技部和中国科学院的支持下，中科院化学研究所分子纳米结构与纳米技术重点实验室方晓红课题组长期致力于发展分析活细胞体系蛋白质动态变化的单分子显微成像新方法，所建立的利用光漂白计数定量表征膜蛋白化学计量比等单分子研究方法，为化学生物学和生物医学研究提供了先进的技术(Natl. Sci. Rev. 2018, 5, 300)。近年来他们与生命科学家合作，发现了多种信号转导蛋白激活和转运新机制(Nat. Microbiology 2019, 4, 97-111; Nat. Commun. 2018, 9:11; Cell Res., 2015, 25,738-752)。

最近，他们针对单分子光漂白计数分析中检测信号弱、易被噪音掩盖和光闪烁干扰、数据量大等挑战性问题，巧妙地利用人工智能深度学习技术，提出了一个机器学习的深度神经网络架构——CLDNN，对单分子光漂白事件进行准确、高效计数。他们所设计的CLDNN可凭借其自身强大的学习能力以及特征提取能力，通过卷积层提取光漂白台阶信号，通过长短时记忆(LSTM)循环层排除荧光闪烁等干扰，实现90%以上准确率的化学计量比分析。相比于已报道的单分子信号分析方法，CLDNN具有诸多优势：(1)更高的分析准确率，尤其对于低信噪比的数据，优势更为明显;(2)更高的计算效率，可以快速处理大批量数据;(3)简单易用，无需对单分子数据做滤噪等预处理，也无需设置算法参数;(4)具有较好的推广性和扩展性，可对除训练水平外的不同信噪比数据以及包含更多光漂白事件的数据进行有效分析。此外，使用者也可以加入自己的实验数据集重新训练CLDNN，使其更加满足个性化分析需要。

CLDNN为研究人员提供了一种全新的数据分析策略，使人们可更为高效、准确、客观地分析蛋白质等生物大分子化学计量比，该架构还可应用于分析化学领域其它具有时间序列特征的数据分析，将促进人工智能技术在化学中的应用。相关研究成果近期发表于J. Am. Chem. Soc.(2019，141，6976-6985)。



CLDNN深度学习用于单分子光漂白计数分析及蛋白质化学计量比研究

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发