

TP-ViT: 面向视觉Transformer量化的截断均匀对数量化器与渐进式比特衰减重建方法

周喜川, 赵思桓, 丁睿, 史家宇, 聂晶, 陈黎辉, 刘海军
重庆大学微电子与通信工程学院, 中国重庆市, 400044

摘要: 视觉Transformer (ViT) 在各类人工智能驱动的计算机视觉应用中取得显著成功。然而, ViT对计算和内存需求较高, 限制了其在资源受限的边缘设备上的实际应用。尽管后训练量化 (PTQ) 通过使用少量校准数据降低模型精度, 为此提供了一种前景广阔的解决方案, 但激进的低比特量化通常会导致性能大幅下降。为应对这一挑战, 提出面向视觉Transformer量化的截断均匀对数量化器与渐进式比特衰减重建方法 (TP-ViT)。该方法是一种专为ViT设计的创新性PTQ框架, 包含两项关键技术贡献: (1) 截断均匀对数量化器——该新型量化方法能有效处理Softmax后激活中的异常值, 显著降低量化误差; (2) 比特衰减优化策略——利用过渡权重逐步降低比特精度, 在极端量化条件下仍保持模型性能。在图像分类、目标检测和实例分割任务上的综合实验表明, TP-ViT相较于当前最先进的PTQ方法具有卓越性能, 尤其在极具挑战性的3比特量化场景下表现突出。在3比特量化条件下, 本框架在ViT-small模型上实现了Top-1准确率6.18个百分点的提升。结果验证了TP-ViT的鲁棒性与普适性, 为ViT模型在边缘硬件计算机视觉应用中的高效部署铺平了道路。

关键词: 视觉Transformer; 后训练量化; 块重建; 图像分类; 目标检测; 实例分割

<https://doi.org/10.1631/ENG.ITEE.2025.0081>