

基于分层特征感知注意力与查询选择编码器的端到端目标检测

王足毅¹, 郑智萌¹, 孟濬^{1,2}, 许力^{1,2}

¹浙江大学电气工程学院, 中国杭州市, 310027

²浙江大学机器人研究院, 中国余姚市, 315400

摘要: 由于无需设计复杂人工组件且简化了检测流程, 端到端目标检测方法近年来受到广泛关注。然而, 与传统检测器相比, 这类方法存在训练收敛速度较慢、检测性能不足的问题, 究其原因是在特征融合与选择过程中算法受限于正样本监督信号不足。针对此问题, 本文提出一种用于端到端目标检测器的查询选择编码器 (QSE), 可以提升训练收敛速度与检测精度。QSE由多个编码器层组成, 且在每个编码器层后添加了轻量级网络, 以级联方式持续优化特征, 为高效训练提供更充分的正样本监督。此外, 每个编码器层引入分层特征感知注意力 (HFA) 机制, 包括层内以及跨层特征注意力, 以增强不同层级特征间的交互融合。HFA能有效抑制相似特征表征并强化判别性特征, 从而加速特征选择过程。该方法可灵活应用于基于卷积神经网络和基于Transformer的检测器; 在目标检测主流基准数据集MS COCO、CrowdHuman以及PASCAL VOC上的大量实验表明, 使用QSE的基于卷积神经网络或基于Transformer的检测器均能在更少训练周期内获得更优的端到端检测性能。

关键词: 端到端目标检测; 查询选择编码器; 分层特征感知注意力
<https://doi.org/10.1631/FITEE.2400960>