

说明书摘要

本发明提出了一种基于可扩展表征学习的行人重识别方法。主要包括视觉表征学习模块、文本表征学习模块和行人检索模块。视觉表征学习模块利用来自两个空间的 ViT 结构的视觉编码特征增强图像区分能力，并通过因果注意力机制融合；文本表征学习模块借助示范样本，同时采用因果注意力机制，消除数据集中存在的偏差，从而提高行人重识别的性能，使行人检索结果更加准确。