

1 文献基本信息

作者: Sijie Yan, Yuanjun Xiong, Dahua Lin; 文献题目: 基于骨架的动作识别的时空图卷积网络; 刊物: 计算机视觉和模式识别; 刊号: ISSN(1063-6919); 时间: 25.01.2018; 格式引文: arXiv:1801.07455v2[cs.CV]

2 文献讨论的问题及其主要观点

写作目的 传统的骨架的动作识别通过设计几个手工制作的特性来捕捉关节运动的动态, 建模骨架通常依赖手工制作的零件或遍历规则。本论文解决这种方式导致表达能力有限和泛化困难问题。

研究方法 采用一个新的动态骨架模型, 称为时空图卷积网络 (ST-GCN), 此模型第一个将图 CNNs 应用于基于骨架的交流识别的系统, 该模型建立在一系列骨架图的顶部, 其中每个节点对应于人体的一个关节。有两种类型的边缘, 即空间边缘, 形成自然连接的关节和时间边缘, 连接相同的关节跨越连续的时间步。在此基础上构造了多个时空图卷积层, 使得信息可以在时空维度上进行融合。

主要结论 在两个具有挑战性的大型数据集上 (Kinetics 和 NTU-RGB+D), ST-GCN 优于先前的基于骨架的模型。此外, ST-GCN 还可以捕获动态骨架序列中的运动信息, 这与 RGB 模式是互补的。基于骨架的模型与基于框架的模型相结合, 进一步提高了动作识别的性能。

3 文献阅读后的分析

文章的题目是否清晰、准确? 本文题目为基于骨架的动作识别的时空图卷积网络, 关键字为骨架、动作识别、时空图卷积。从文献 introduction 部分的内容我们可以得知, 已经有研究人员基于骨架的图片在空间尺度上对动作识别进行研究, 而本文的工作不仅基于空间尺度, 还加入了时间尺度, 也就是题目中所提到的时空图卷积, 所以题目是清晰准确的。

文摘是否很好的提炼了文献的内容、容易理解? 本文在摘要中首先说明传统骨架动作识别方法的缺点。针对这样的缺点, 继而提出了第一个将图 CNNs 应用于基于骨架的交流识别的模型。提出这样的模型以后, 在两大风格差异较大的数据集上进行实验, 通过多组性能指数具体数据的对比, 得出新方法的性能优于传统方法的结论。很好地概括了文献的内容。

在 introduction 部分文章是否清晰地说明了写作本文的目的? 论文在 introduction 部分提到人类动作识别是近年来一个活跃的研究领域, 并且有多种形式来识别人类动作, 比如外观、光流、身体骨架。其中动态人体骨架通常传达了与其他信息十分互补的重要信息, 但是该领域受到的重视却相对较小, 本文旨在提出一种规则有效的动态骨架识别法并将他们利用于动作识别中。

本文采用的方法所解决的问题? 本文采用时序图卷积网络替代手工制作的零件或遍历规则建模骨架, 改善了传统方法表现力不足和泛化困难的问题。