

基于骨架的行为识别算法 测试数据可视化

郭佳俊 21215074

2022 年 6 月 25 日

NTU RGB+D 60^[1] 是行为识别任务常用的数据集之一，包含 60 个类别共 56880 个样本，每个样本都有相应的 RGB 视频、深度视频、红外视频以及三维骨架序列数据。基于骨架的行为识别任务利用其中的三维骨架信息进行动作的分类。由于类别数目较多，在使用混淆矩阵展示测试结果时，矩阵元素数太多会导致难以找对感兴趣的类别的分类情况，所以本文尝试使用和弦图来展示测试结果，结合过滤筛选等交互方式，帮助用户对数据进行简单的分析¹。

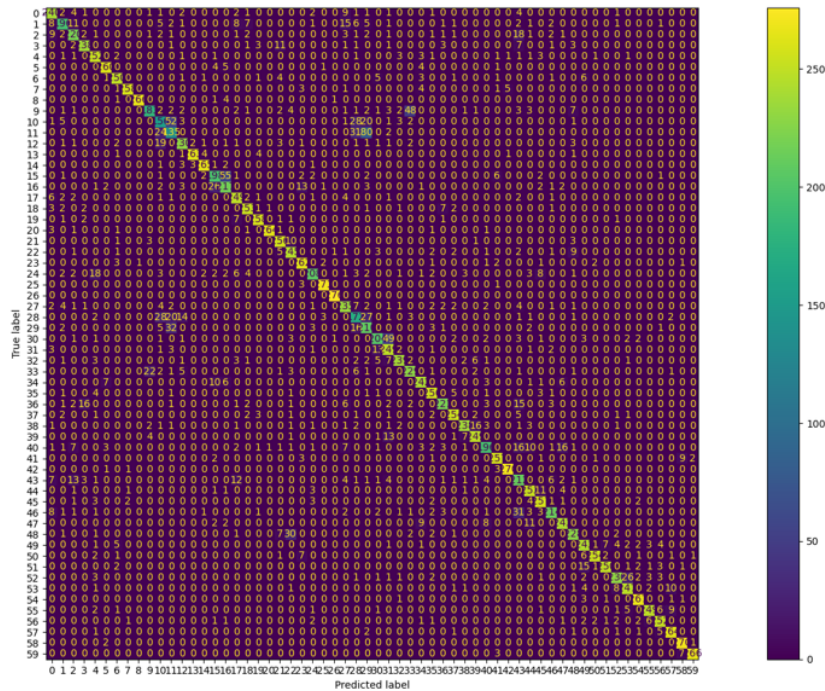


图 1 混淆矩阵

本文使用的基于骨架的行为识别模型是 Richly Activated GCN v2 (RA-GCN v2)^[2]，它将类激活图 (Class Activation Map, CAM)^[3] 的概念应用到了图卷积中，将网络最后一层全连接层的权重与前一层图卷积的特征图的加权和，上采样到与输入骨架序列相同的尺寸，作为骨架点的激活情况，以展示各个时刻各个骨架点对分类结果的影响程度。

¹完整代码位于 <https://github.com/guojj33/skeleton-vis>，
在线演示位于 <https://guojj33.github.io/skeleton-vis/> (纯前端实现)

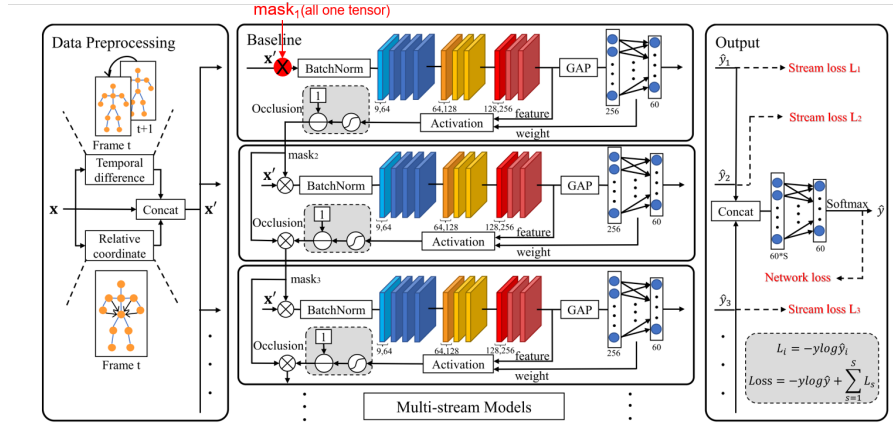


图 2 Richly Activated GCN v2 模型结构

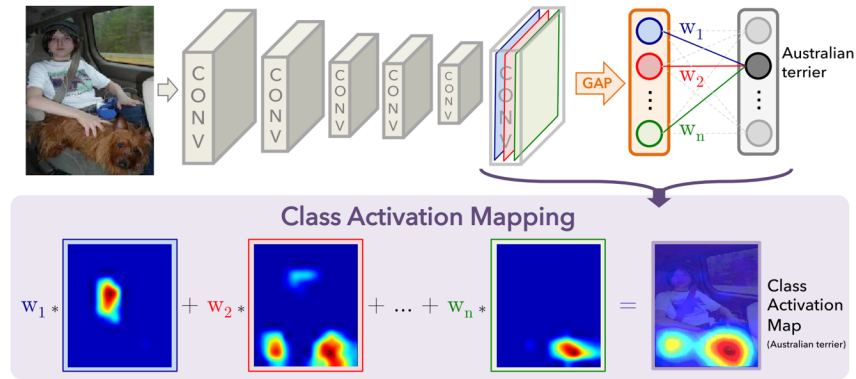


图 3 Class Activation Map 计算方法

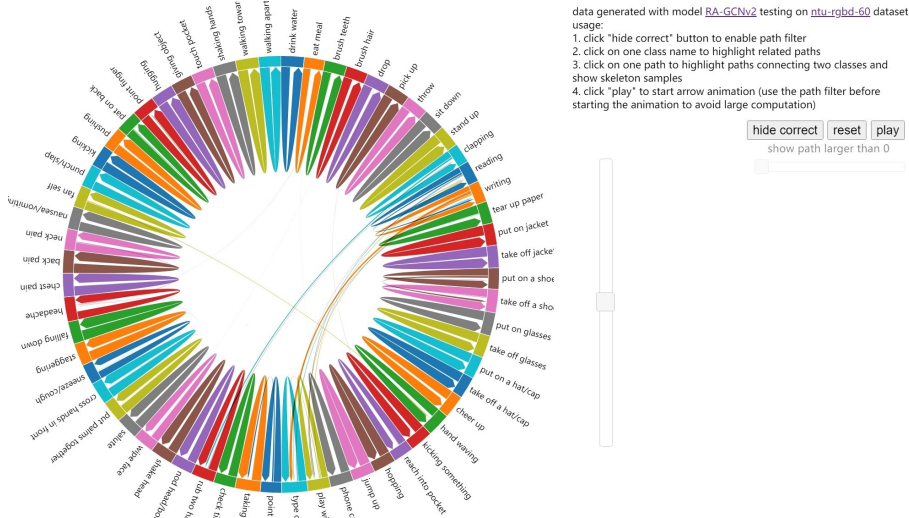


图 4 初始界面

本文设计的可视化界面如图4所示，初始时和弦图的数据与混淆矩阵一致，图中的路径是带方向的，起点为真实类别，终点为预测结果。由于混淆矩阵对角线元素值相对更大，所以在和弦图中每一个类指向自身的路径占据了很大的空间，使非对角线元素对应的路径难以被观察到。所以此处添加了隐藏对角线元素

的功能，从而隐藏类指向自身的路径。通过点击“hide correct”按钮激活此功能，然后调节按钮下方的水平滑动条，可以进一步隐藏错误样本数小于选定阈值的路径，得到结果如图5。

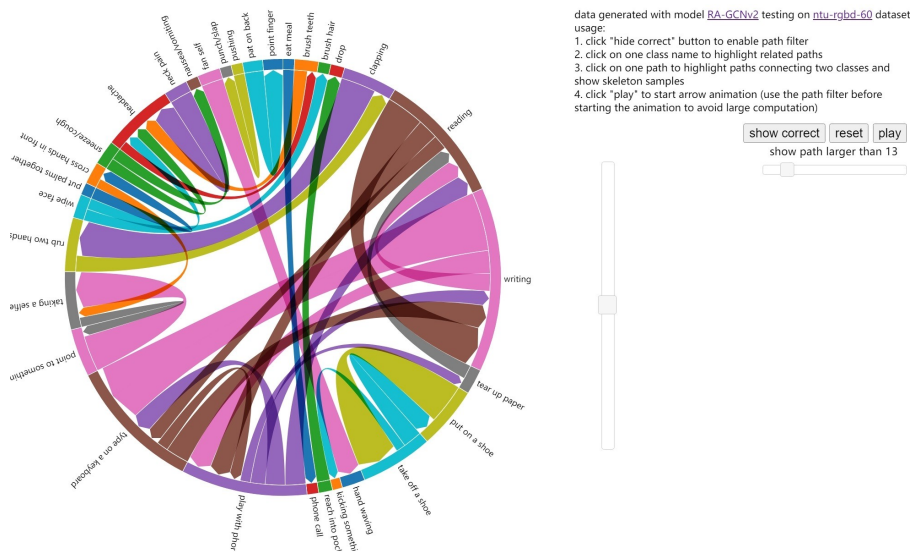


图 5 隐藏混淆矩阵对角线元素和小于选定阈值的元素

如图6，通过点击类别标签可以突出显示与其相关的路径，包括以该类为起点和终点的路径。

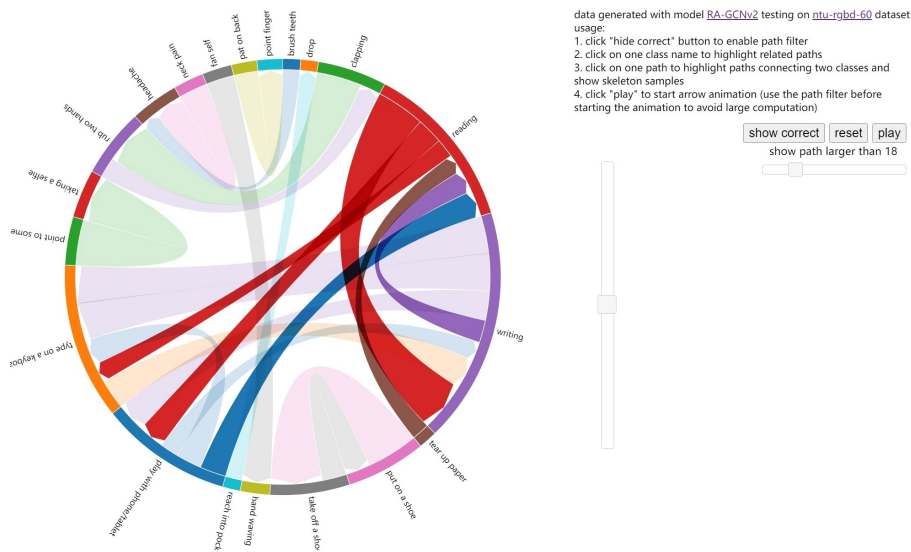


图 6 点击类别以标签突出显示相关路径

由于预测错误通常是因为两个类别的样本比较相似，因此可能互相识别错，所以添加了突出显示关于对角线对称的元素对应路径的功能。如图7，点击路径以激活此功能，同时界面右侧会展示相关的样本的骨架序列。骨架点的颜色与激活值对应，从蓝色到红色，激活值增大。

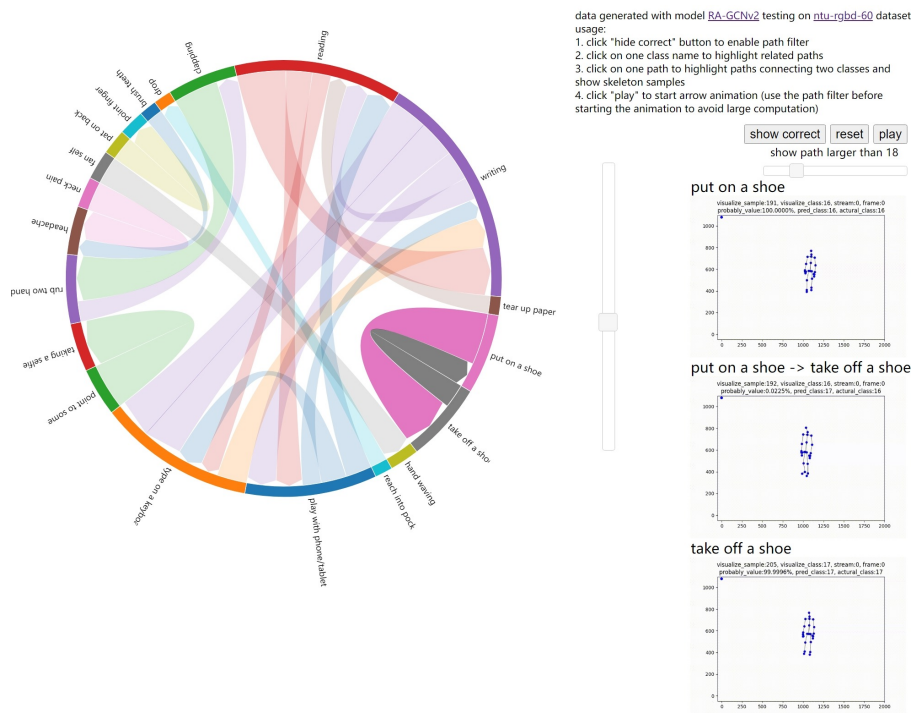


图 7 点击路径以突出显示该路径和反向的路径（若存在）

Min Lu^[4] 等人提出可以使用动画来增强静态图像，利用动画对数据中的属性进行编码，从而予以强调，加快感知理解。作者归纳了三种动画效果，其中与和弦图可以自然融合的是“行军蚂蚁（Marching Ants）”效果。如图8，设计此动画需要指定 Visual Proxy（移动的元素是什么），Dynamic Function（移动的路径、边界、元素间隔、速度）。此效果可以对和弦图中路径的方向和数值进行编码。

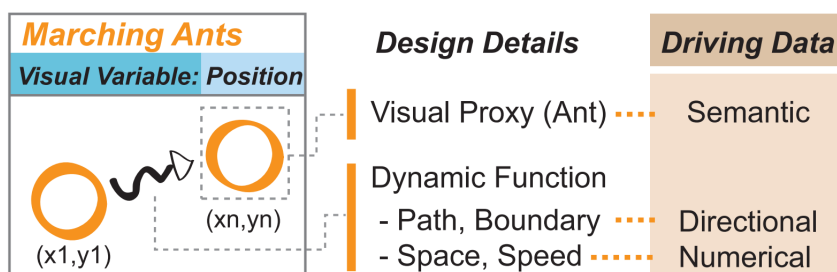


图 8 使用“行军蚂蚁”效果对数值、方向等信息进行编码

由于论文提供的代码在与和弦图结合时卡顿比较严重，以及和弦图中路径的边界（Boundary）较难定义，所以此处通过其他方式复现了此效果，如图9。通过定时更新每一个箭头图形的位置和方向实现动画效果，数值越大的路径中箭头的移动速度会越快，但是在路径数较多的时候仍会有些卡顿。

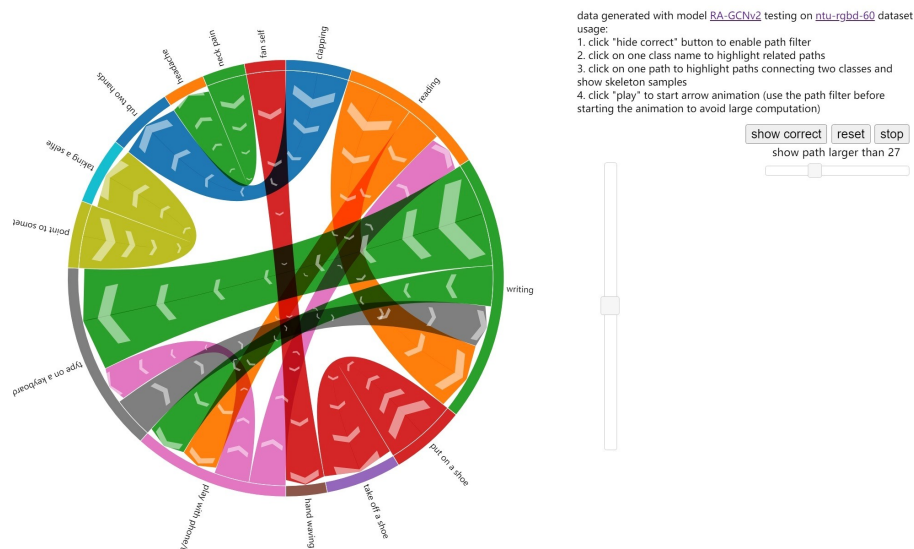


图 9 “行军蚂蚁”效果复现

通过使用本文的可可视化工具，可以比较方便地展示误判较多的情况是哪些，通过查看样本可以发现有的样本确实单看骨架信息不太容易识别准确，存在整体动作很相似的类别，例如穿鞋和拖鞋，写字和打字，若要减少这些误判，需要增强对手部动作细节的判断。但是此工具也有一些缺点，由于和弦图中路径的显示占用空间较大，在避免杂乱的情况下能展示的数据有限，所以还要结合混淆矩阵来使用，可以先从混淆矩阵中获取整体的感知，再使用本文的工具做下一步分析。

参考文献

- [1] SHAHROUDY A, LIU J, NG T T, et al. Ntu rgb+ d: A large scale dataset for 3d human activity analysis[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2016: 1010-1019.
- [2] SONG Y F, ZHANG Z, SHAN C, et al. Richly activated graph convolutional network for robust skeleton-based action recognition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(5):1915-1925.
- [3] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2016: 2921-2929.
- [4] LU M, FISH N, WANG S, et al. Enhancing static charts with data-driven animations[J]. IEEE Transactions on Visualization and Computer Graphics, 2020.