

UAV-GCN 算法说明书

队伍名称	骼固鼎芯
参赛院校	东北大学秦皇岛分校
参赛队员	魏祥睿
	欧阳广贤
	陈少勇
指导老师	吕艳霞

目录

1 数据处理	2
2 CTR-GCN	2
3 BlockGCN	3
4 CD-GCN	4
5 InfoGCN	4
6 分数融合	5
7 总结	6

1 数据处理

为确保数据质量和处理效率，我们对原始数据实施了一系列系统化的预处理步骤：

- 正则化处理: 对原始数据进行正则化处理，确保数据格式一致。
- 丢失帧检查: 检查数据中是否存在丢失帧，若存在则进行补帧或删除。
- 基于帧长度的降噪: 根据帧长度对数据进行降噪处理，去除异常帧。
- 基于运动量的降噪: 计算骨架运动量，过滤掉运动量过大 (> 3) 或过小 (< 0.089) 的序列。
- 去除 NaN 值帧: 检查数据中是否存在 NaN 值帧，若存在则进行删除或替换。
- 序列对齐: 将所有序列对齐到相同帧长，确保数据格式一致。

2 CTR-GCN

在国赛中，我们首先沿用了省赛中使用的 CTR-GCN [CZY+21] 模型，模型架构如下图所示：

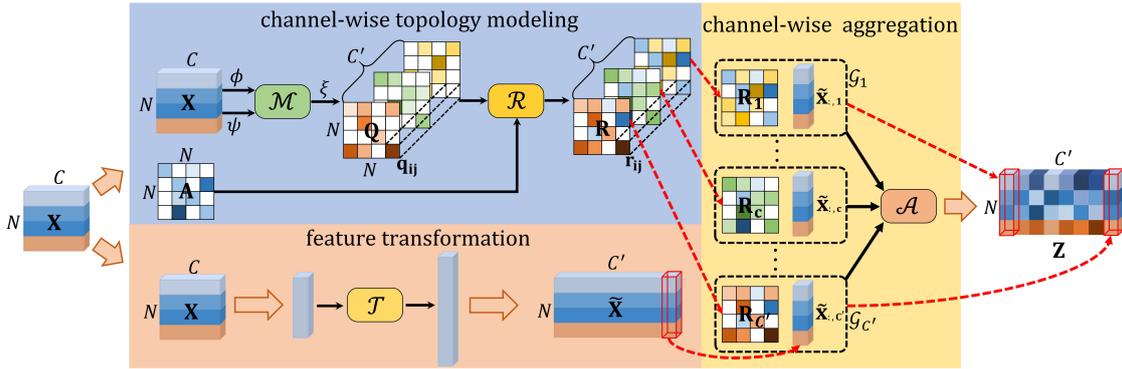


图 1: Framework of CTRGCN

CTR-GCN 是一种创新的骨架动作识别图卷积网络模型，其主要创新点在于提出了通道级拓扑细化图卷积 (CTR-GC)。该方法首先学习一个所有通道共享的通用拓扑结构作为先验知识，然后结合每个通道特定的动态相关性进行细化，从而实现了高效的通道级拓扑建模。模型架构由 10 个基本块构成，每个块包含空间建模和时序建模两个关键模块。其中，空间建模采用 3 个并行的 CTR-GC 来捕获人体关节间的相关性；时序建模则使用多尺度时序卷积来处理不同长度的动作特征。在本项目中，我们针对 joint、bone、joint motion 和 bone motion 四种模态。

针对 joint 模态，我们创新性地引入了基于 BatchNorm 层的测试时适应 (TTA) 算法。该算法通过选择性地更新 BatchNorm 层参数，同时保持其他网络参数不变，使模型能够在测试阶段自适应地调整特征分布。这种方法有效缓解了域偏移问题，显著提升了模型的泛化性能。

对于 bone 模态，我们采用了 Logit Adjustment 算法来解决数据长尾分布问题。该算法通过智能调整预测 logits，为样本数量较少的类别提供补偿机制。通过为少数类别分配更大的补偿权重，算法有效提升了模型对这些类别的识别准确率，同时保持了整体性能的稳定。

经过在官方 train_joint 数据集上的训练和优化，我们的 CTR-GCN 模型在 val_joint 验证集上取得了显著的性能表现，具体结果如下表所示。

模态	结果
joint	43.90
bone	43.30
joint_motion	36.85
bone_motion	37.80
joint_tta	44.45
bone_longtail	43.55

Table 1: CTR-GCN 模型在验证集上的性能表现

3 BlockGCN

BlockGCN [ZYC+24] 模型架构如下图所示：

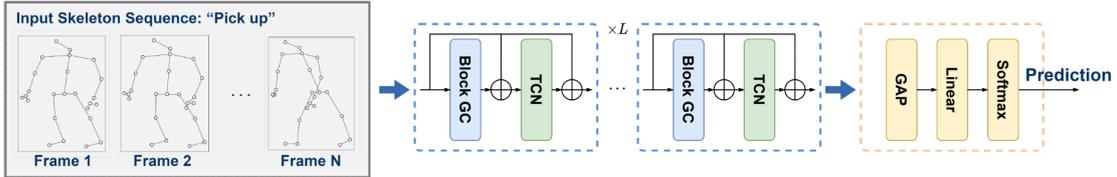


图 2: BlockGCN 模型架构

BlockGCN 是一种突破性的骨架动作识别模型，针对性地解决了传统图卷积网络在骨骼拓扑建模中的局限性。该模型的创新主要体现在两个关键设计上：一方面，通过引入基于图距离的拓扑编码机制，结合持久同调分析技术，BlockGCN 能够精确捕获和保留动作序列中的关键拓扑特征，这些特征在传统 GCNs 中往往会被忽略。另一方面，模型创新性地提出了 BlockGC 图卷积改进方案，该方案在显著降低参数数量的同时，仍能高效地建模关节间的潜在关联关系。

在官方 train_joint 数据集上的实验验证中，BlockGCN 模型展现出了卓越的性能表现，其在 val_joint 验证集上的具体评估结果如下表所示：

模态	结果
joint	43.45
bone	44.55

Table 2: BlockGCN 模型在验证集上的性能表现

4 CD-GCN

CD-GCN [MKZ⁺23] 为骨架动作识别任务提供了一种创新的图卷积解决方案。其核心是新提出的中心差分图卷积（CDGC）操作，该操作不仅能有效聚合节点特征，还能精确捕获节点间的梯度变化，从而更全面地刻画动作的局部运动特征。CDGC 的一大优势在于其可以直接替代传统图卷积网络中的标准卷积操作，无需引入额外参数，同时通过优化的加速版本显著提升了模型训练效率。这种基于中心差分的设计使 CD-GCN 在处理复杂动作时表现出色，尤其善于捕捉关节间的动态运动关系。

在官方 train_joint 数据集上的实验中，CD-GCN 模型展现出了优异的性能，其在 val_joint 验证集上的详细评估结果如下表所示：

模态	结果
joint	43.85
bone	44.25

Table 3: CDGCN 模型在验证集上的性能表现

5 InfoGCN

InfoGCN [CHC⁺22] 模型架构如下图所示：

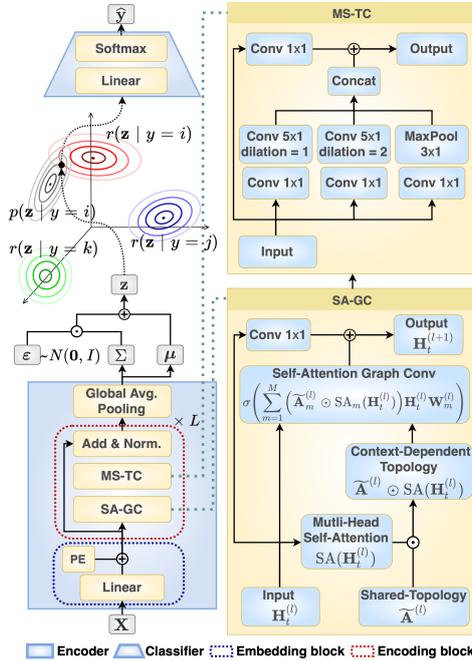


图 3: InfoGCN 模型架构

InfoGCN 是一种新型的骨架动作识别模型, 其核心创新在于将信息瓶颈 (IB) 理论与自注意力图卷积 (SA-GC) 模块进行了有机结合。该模型通过信息瓶颈原理构建学习目标, 实现了对动作特征的高效压缩, 确保提取的特征既紧凑又包含充分的目标相关信息。模型中引入的自注意力机制能够动态捕获动作序列中的上下文依赖关系, 有效建模了人体关节间的内在拓扑结构。为进一步提升性能, InfoGCN 还创新性地采用了基于关节相对位置的多模态骨架表示方法, 为动作识别提供了更丰富的空间信息特征。

在官方 train_joint 数据集上的实验表明, InfoGCN 模型在 val_joint 验证集上展现出了优异的性能, 具体评估结果如下表所示:

模态	结果
k_1	43.20
k_2	43.10
k_6	43.90
32Frame_1	40.85
32Frame_2	40.75
32Frame_6	41.70
128Frame_1	43.15
128Frame_2	43.05
128Frame_6	43.80

Table 4: InfoGCN 模型在 val_joint 验证集上的性能评估结果

6 分数融合

为了有效融合多个模型的预测结果, 我们设计了一种基于贝叶斯优化的自适应权重分配方法。该方法以 7 个不同 GCN 变体模型 (包括不同的网络架构和输入模态) 在验证集上的预测分数作为输入, 利用高斯过程回归作为代理模型, 在 $[0.2, 1.2]$ 区间内精确搜索最优权重组合。算法采用 0.01 的离散化精度, 通过 150 轮迭代优化来最小化负准确率损失。在优化过程中, 系统不断计算加权集成的预测准确率, 并基于贝叶斯优化策略动态调整搜索方向, 最终得到一组能使集成模型达到最佳验证集性能的权重参数。以下是我们选择的 7 个预测分数:

模型 _ 模态	结果
ctrfcn_joint	43.90
ctrfcn_joint_tta	44.45
blockfcn_bone	44.55
cdfcn_joint	43.85
cdfcn_bone	44.25
infofcn_k_6	43.90
infofcn_128Frame_6	43.80

Table 5: 最终选择用于融合的 7 个预测分数

7 总结

本项目构建了一个完整的骨架动作识别解决方案，通过多模型、多模态的协同作用实现了较好的识别效果。在数据处理方面，建立了包含正则化、降噪、序列对齐等在内的全面预处理流程；在模型架构方面，分别采用了 CTR-GCN（引入通道级拓扑细化）、InfoGCN（结合信息瓶颈理论与自注意力机制）、BlockGCN（创新性地引入基于图距离的拓扑编码）和 CD-GCN（提出中心差分图卷积操作）四种先进的图卷积网络模型；在优化策略方面，实现了多模态训练，并针对性解决了域偏移、长尾分布等问题；最终通过基于贝叶斯优化的模型融合方案，将各个模型的优势有机结合。

参考文献

- [CHC⁺22] Hyung-gun Chi, Myoung Hoon Ha, Seunggeun Chi, Sang Wan Lee, Qixing Huang, and Karthik Ramani. Infogcn: Representation learning for human skeleton-based action recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20186–20196, 2022.
- [CZY⁺21] Yuxin Chen, Ziqi Zhang, Chunfeng Yuan, Bing Li, Ying Deng, and Weiming Hu. Channel-wise topology refinement graph convolution for skeleton-based action recognition. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 13359–13368, 2021.
- [MKZ⁺23] Yulong Ma, Bing Kong, Lihua Zhou, Hongmei Chen, and Chongming Bao. CdgcN: An effective and efficient algorithm based on community detection for training deep and large graph convolutional networks. In *International Conference on Spatial Data and Intelligence*, pages 128–139. Springer, 2023.
- [ZYC⁺24] Yuxuan Zhou, Xudong Yan, Zhi-Qi Cheng, Yan Yan, Qi Dai, and Xian-Sheng Hua. BlockgcN: Redefine topology awareness for skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2049–2058, 2024.