

算法描述文档

算法挑战赛—基于无人机的人体行为识别

算法说明书



队伍名称：魑魅魍魉

队员：赵佰亮、陈庭宇、张健

队伍编号：DIGIX2024TEAM130047

目录

**一、思路阐述1**

**二、实现描述1**

1、TE-GCN1

2、HDBN3

**三、创新点介绍4**

1、双分支网络+TEGCN5

2、基于单目标优化模型的参数搜索6

**四、参考文献资料及引用8**

1. **思路阐述**

基于官方提供的UAV-Human处理后的数据集，我们的思路是先复现现有优秀仓库中的训练结果【1】，然后通过模型的训练结果与验证测试集的准确率进行相应的创新优化，通过调整模型参数、优化模型实现对结果数据进行加权等方法尝试来找到能够更好的提高准确度的方法。

1. **实现描述**

我们对TE-GCN与HDBN两个仓库的模型进行了训练复现，测试了数据集达到的效果以及对应的准确度，验证了模型网络的思路并且对论文思路进行了学习解读。

1. **TE-GCN：**

《Temporal Graph Modeling for Skeleton-based Action Recognition》【1】【4】 提出了一种用于骨架基础动作识别的时序增强图卷积网络（Temporal Enhanced Graph Convolutional Network, TE-GCN）。该网络旨在通过构建时序关系图来捕捉骨架序列中的复杂时序动态，特别是非相邻时间步之间的时序关系。TE-GCN通过显式建立语义相关时间特征之间的连接来模拟相邻和非相邻时间步之间的时序关系。此外，为了进一步探索充分的时序动态，作者设计了多头机制来研究不同时间步之间的多种时序关系。在NTU-60 RGB+D和NTU-120 RGB+D两个大规模数据集上的广泛实验表明，所提出的模型通过时序建模为动作识别做出了贡献，具有优秀的性能。

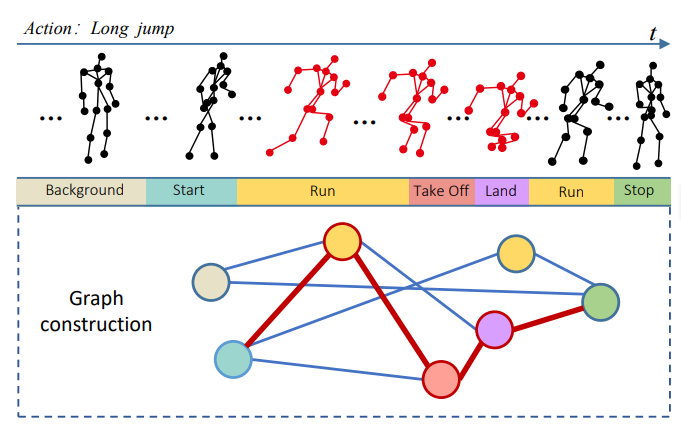
我们对技术架构进行了整理总结：

**1.时序关系图构建：** TE-GCN通过两种转换函数（Feature Calculated和Feature Learned）来计算时间特征之间的相关性，构建时序关系图。

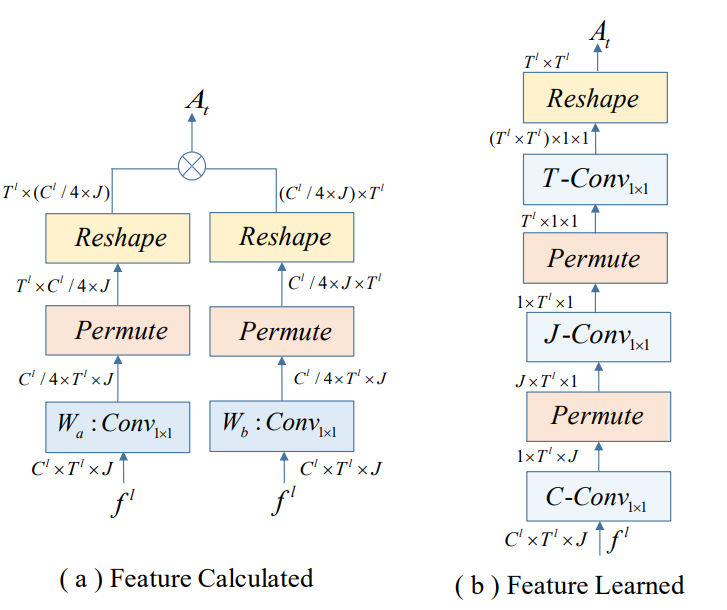
**2.多头时序图卷积：** 模型设计了多头机制，通过多个独立的相关函数获得不同的时间相邻矩阵，然后对每个时间图进行图卷积，最后通过元素求和聚合特征。

**3.网络结构：** TE-GCN基于ST-GCN骨架构建，包含输入层和多个图卷积层，其中常规的1D时序卷积被替换为提出的时序增强图卷积。

**4.多流网络与多模态：** 模型还探索了关节、骨骼及其相应的运动信息等多种模态，通过加权求和融合不同模态的softmax分数。



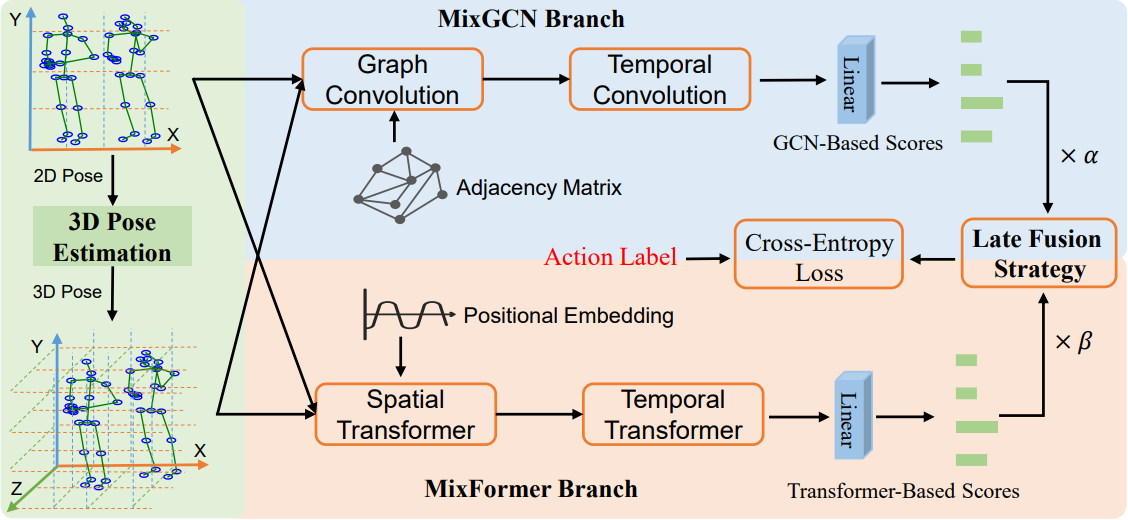
**Figure1：时序建模过程图**



**Figure 2: 计算过程**

1. **HDBN：**

《HDBN: A NOVEL HYBRID DUAL-BRANCH NETWORK FOR ROBUST SKELETON-BASED ACTION RECOGNITION》【3】【5】 提出了一种名为**Hybrid Dual-Branch Network (HDBN)**的新型双分支框架，用于鲁棒的基于骨架的动作识别。该框架结合了**图卷积网络 (GCNs)** 和 **Transformers** 来处理2D和3D骨架数据，旨在充分利用这两种网络架构的互补特性。

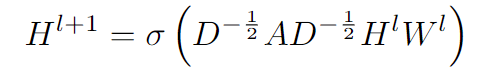


**Figure3: HDBN结构示意图**

下面我们对技术架构进行了整理总结：

**双分支架构**：

**MixGCN分支**：使用GCNs来处理图结构数据，特别是骨架数据。该分支采用了三种动态邻接矩阵（TD-GCN、CTR-GCN和MST-GCN）来处理2D和3D骨架数据。一般的图卷积过程可以用公式表示：



其中是第层的特征，是激活函数，是n个节点的度矩阵，是第 层的可学习参数，是邻接矩阵。

**MixFormer分支**：利用Transformers强大的全局信息建模能力，通过Transformer架构来编码骨架数据。

编码特征的简化表示如下：



其中，、、 和分别表示查询、键、值和输出权重矩阵。

HDBN通过结合GCNs和Transformers，有效地提高了基于骨架的动作识别的鲁棒性。论文实验结果表明，该方法在UAV-Human数据集上表现优异，优于大多数现有方法。

1. **创新点介绍**

我们基于论文提出了一种名为Hybrid Dual-Branch Plus Network（HDBPN）的新型双分支网络，用于鲁棒的基于骨架的动作识别。HDBPN在论文的原有的基础上，保留了原有的双分支网络这一创新结构【7】，并创新性引入了TEGCN这一改进后的GCN分类模型【6】，建立了单目标优化模型来搜寻融合权重。HDBPN结合了图卷积网络（GCN）和Transformer，旨在充分利用这两种网络架构的互补特性，以提高动作识别的准确性和鲁棒性。

**1.双分支网络+TEGCN：**

基于论文的启发，我们在HDBNP中构建了MixGCN Branch和Mixformer Branch这两大分支。这两个分支的构建旨在通过结合不同的模型和技术，增强我们对复杂人体行为数据的处理和识别能力。

在MixGCN Branch的设计中，我们采用了三种动态邻接矩阵来构建图卷积网络（GCN），分别是TD-GCN、CTR-GCN和MST-GCN。这三种模型各自具有独特的优势，能够从不同的角度捕捉人体行为的特征。通过将它们集成到一个分支中，我们能够充分利用这些模型之间的互补性，从而提高整体的识别性能。

同时考虑到TE-GCN这类新的模型在人体行为识别上的卓越性能，我们一并增加了这一新的模型。通过对这些模型进行独立的训练，我们得到了各自的置信度文件。为了进一步优化模型的融合效果，我们引入了一个模型参数α，用于调整不同GCN模型在最终决策中的权重，以实现最佳的性能平衡。

在Mixformer Branch的设计中，我们认识到Transformer模型在处理全局信息和抽象特征方面的强大能力。因此，我们将Transformer作为该分支的核心骨干网络。通过利用丰富的行为数据对Transformer模型进行训练，我们同样得到了置信度文件。为了进一步调整和优化模型的表现，我们引入了模型参数β，用于控制Transformer模型在整体框架中的影响力。

通过这种双分支的设计，我们的模型不仅能够捕捉到局部的图结构特征，还能够有效地处理全局的信息和上下文。这种多模型融合的方法，使得我们的HDBPN框架在处理复杂多变的人体行为数据时，能够提供更加准确和鲁棒的识别结果。参数α和β的引入，为模型的调整和优化提供了灵活性，使得我们的系统能够根据不同的应用场景和数据特性进行自适应调整，以达到最佳的识别效果。

**2.基于单目标优化模型的参数搜索**

**参数优化策略：**

在模型融合的过程中，我们面临着一个关键的挑战：如何确定各个模型在融合过程中的最优权重。为了解决这一问题，我们建立了一个单目标优化模型，其核心目标是最大化训练集验证集的准确率。这一策略的制定基于这样一个假设：如果模型在验证集上表现良好，那么它在未知数据上的表现也很可能同样出色。

**Powell方法的应用 ；**

为了寻找最优参数，我们采用了Powell方法，这是一种无约束优化算法，特别适合于处理维度n小于20的问题。Powell方法通过共轭方向法加速收敛，这种方法不需要导数信息，因此对于非线性或不可微的函数同样适用。在我们的案例中，Powell方法帮助我们在参数空间中有效地搜索，以找到最大化验证集准确率的参数组合。

**Python实现与优化过程**

我们利用Python的scipy.optimize模块中的minimize函数来实现Powell方法。由于minimize函数是寻找最小值的，我们将准确率的负值作为目标函数，通过最小化负准确率来实现准确率的最大化。在实际应用中，我们将模型的准确率作为目标函数，然后通过minimize函数的method='Powell'参数来指定使用Powell方法。

**结果获取与应用**

优化完成后，我们通过minimize函数返回的对象result获取最优参数。result.x包含了优化后的权重比例，而-f1(result.x)给出了使用这些权重比例计算出的最大准确率。这些参数随后被应用到测试集置信度文件的融合中，以同样的比例进行权重融合，最终得到测试集融合后的置信度文件。

1. **参考文献资料及引用：**

1. [UAV-Human Benchmark (Skeleton Based Action Recognition) | Papers With Code](https://paperswithcode.com/sota/skeleton-based-action-recognition-on-uav?metric=CSv2(%25))
2. <https://doi.org/10.48550/arXiv.2012.08804>
3. <https://doi.org/10.48550/arXiv.2404.15719>
4. [xieyulai/TE-GCN: skeleton-based action recognition (github.com)](https://github.com/xieyulai/TE-GCN)
5. [liujf69/ICMEW2024-Track10: [ICMEW 2024] Implementation of the paper “HDBN: A Novel Hybrid Dual-branch Network for Robust Skeleton-based Action Recognition“. (github.com)](https://github.com/liujf69/ICMEW2024-Track10)
6. [lansijian/my\_tegcn: in master (github.com)](https://github.com/lansijian/my_tegcn)
7. [lansijian/reproduce\_HDBN: in master (github.com)](https://github.com/lansijian/reproduce_HDBN)