



(12) 发明专利

(10) 授权公告号 CN 113158970 B

(45) 授权公告日 2023.02.07

(21) 申请号 202110510781.9

G06N 3/04 (2006.01)

(22) 申请日 2021.05.11

G06N 3/08 (2006.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 113158970 A

(56) 对比文件

CN 110059598 A, 2019.07.26

WO 2017133009 A1, 2017.08.10

(43) 申请公布日 2021.07.23

CN 111860128 A, 2020.10.30

(73) 专利权人 清华大学
地址 100084 北京市海淀区清华园

审查员 刘梦佳

(72) 发明人 高跃 陈自强

(74) 专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201
专利代理师 王萌

(51) Int. Cl.

G06V 40/10 (2022.01)

G06V 10/764 (2022.01)

G06V 10/82 (2022.01)

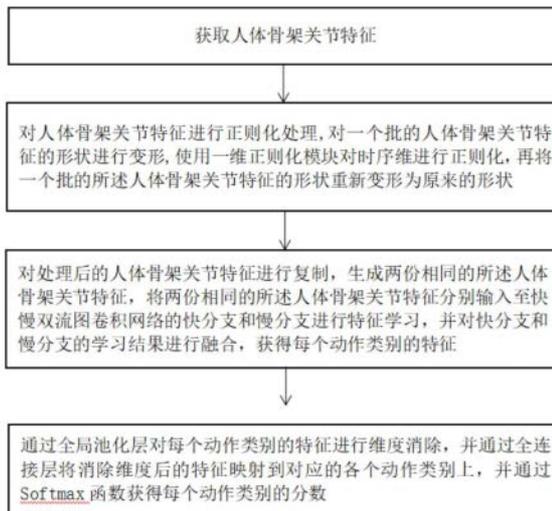
权利要求书3页 说明书8页 附图3页

(54) 发明名称

一种基于快慢双流图卷积神经网络的动作识别方法与系统

(57) 摘要

本发明提出一种基于快慢双流图卷积神经网络的动作识别方法与系统,其中,方法包括:获取人体骨架关节特征;对所述人体骨架关节特征进行正则化处理,对一个批的所述人体骨架关节特征的形状进行变形;对处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习;通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数。本申请解决了现有技术中对于时序信息建模较为薄弱的问题,是一个更好地捕捉时序信息和快慢动作信息的方法。



1. 一种基于快慢双流图卷积神经网络的动作识别方法,其特征在于,包括以下步骤:

步骤S10,获取人体骨架关节特征;

步骤S20,对所述人体骨架关节特征进行正则化处理,其中,对一个批的所述人体骨架关节特征的形状进行变形,使用一维正则化模块对时序维进行正则化,再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状;

步骤S30,对步骤S20处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习,并对所述快分支和慢分支的学习结果进行融合,以获得每个动作类别的特征,其中,快慢双流图卷积网络的快分支和慢分支具有相同的网络结构,具有不同的网络参数配置和输入特征;

步骤S40,通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数;

步骤S30中具体步骤包括:

在每个分支中,都包含连续叠加的多个图卷积区块,每个图卷积区块包括空间图卷积层和时序卷积层;时序卷积层是一个二维卷积模块,卷积核大小为 $(t, 1)$, t 即为卷积核的时序感受野;两个卷积层后都附有批正则化层和ReLU激活函数,保证各个通道的特征保持相同的分布;图卷积区块的计算使用如下公式描述:

$$f_{out} = \sum_k^{K_v} W_k f_{in}(A_k + B_k + C_k)$$

B_k 和 C_k 是在2s-AGCN中提出的自适应性的邻接矩阵,在网络训练的过程中会有改变,其中, B_k 初始化时被设置为 A_k ,用于学习任意两个节点潜在的关联; C_k 是根据样本特征计算得到的矩阵,用于描述样本特定的节点关联;

使用横向连接模块来共享快慢两个分支学习到的信息,从快分支融合到慢分支,由于 $f_{fastout}^i$ 和 $f_{slowout}^i$ 的特征形状分别为 $(B, \beta C, \alpha T, V, M)$ 和 (B, C, T, M, V) ,首先采用一个二维卷积层来进行特征形状转换,并在进行特征形状转换之后增加了批正则化层和ReLU函数,然后用拼接或相加的方式将两个特征融合。

2. 如权利要求1所述的方法,其特征在于,所述步骤S10包括以下步骤:

从数据集中获取人体骨架关节特征,每个样本的特征形状为:

$$(C, T, M, V)$$

其中, C 是特征通道数,值为3,表示关节的三维坐标 (x, y, z) ; T 表示动作的帧数; M 表示表演动作的人数; V 表示人体关节的数量。

3. 如权利要求1所述的方法,其特征在于,所述步骤S20包括以下步骤:

对数据进行正则化处理,在训练过程中使用批训练,所述一个批的张量的特征形状为:

$$(B, C, T, M, V)$$

首先对一个一维批的张量进行变形,变形为:

$$(B, M*V*C, T)$$

然后使用一个所述一维正则化模块对时序 T 维进行正则化,并将特征重新变形为原来

的形状(B,C,T,M,V)。

4. 如权利要求1所述的方法,其特征在于,下面两个公式分别描述了同一个阶段的图卷积区块的输入特征的特征形状:

$$f_{fastin} = (B, \beta C, \alpha T, V, M)$$

$$f_{slowin} = (B, C, T, V, M)$$

快分支的时序维度始终为 αT , α 是一个正整数,表示在初始输入特征中,快分支的输入帧率与慢分支的帧率的比值,在快分支中,通道数量 $\beta_i C_i$ 显著小于同阶段慢分支图卷积区块通道数 C_i ,其中 i 是区块编号, β_i 是一个小于1的值,两个分支的 V 是一致的,都是图节点的数量。

5. 如权利要求1所述的方法,其特征在于,步骤S40中,将步骤S30得到的最终特征通过一个全局池化层消去时序 T 、图节点 V 、人数 M 三个维度,并通过一个全连接层将特征映射到各个动作类别上,最终,通过Softmax函数得到每个动作类别的分数。

6. 一种基于快慢双流图卷积神经网络的动作识别系统,其特征在于,包括:

获取模块,用于获取人体骨架关节特征;

处理模块,用于对所述人体骨架关节特征进行正则化处理,其中,对一个批的所述人体骨架关节特征的形状进行变形,使用一维正则化模块对时序维进行正则化,再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状;

生成模块,用于对所述处理模块处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习,并对所述快分支和慢分支的学习结果进行融合,以获得每个动作类别的特征,其中,快慢双流图卷积网络的快分支和慢分支具有相同的网络结构,具有不同的网络参数配置和输入特征;

确定模块,用于通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数;

所示生成模块,还用于:

在每个分支中,都包含连续叠加的多个图卷积区块,每个图卷积区块包括空间图卷积层和时序卷积层;时序卷积层是一个二维卷积模块,卷积核大小为 $(t, 1)$, t 即为卷积核的时序感受野;两个卷积层后都附有批正则化层和ReLU激活函数,保证各个通道的特征保持相同的分布;图卷积区块的计算使用如下公式描述:

$$f_{out} = \sum_k^{K_k} W_k f_{in} (A_k + B_k + C_k)$$

B_k 和 C_k 是在2s-AGCN中提出的自适应性的邻接矩阵,在网络训练的过程中会有改变,其中, B_k 初始化时被设置为 A_k ,用于学习任意两个节点潜在的关联; C_k 是根据样本特征计算得到的矩阵,用于描述样本特定的节点关联;

使用横向连接模块来共享快慢两个分支学习到的信息,从快分支融合到慢分支,由于 $f_{fastout}^i$ 和 $f_{slowout}^i$ 的特征形状分别为 $(B, \beta C, \alpha T, V, M)$ 和 (B, C, T, M, V) ,首先采用一个二维卷积层来进行特征形状转换,并在进行特征形状转换之后增加了批正则化层和ReLU函数,然

后用拼接或相加的方式将两个特征融合。

7. 一种计算机设备,其特征在於,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时,实现如权利要求1-5中任一所述的方法。

8. 一种非临时性计算机可读存储介质,其上存储有计算机程序,其特征在於,所述计算机程序被处理器执行时实现如权利要求1-5中任一所述的方法。

一种基于快慢双流图卷积神经网络的动作识别方法与系统

技术领域

[0001] 本发明涉及基于骨架信息的动作识别的技术领域,尤其涉及一种基于骨架信息的动作识别的技术领域。

背景技术

[0002] 在基于骨架信息的动作识别这一任务中,基于图卷积神经网络的方法是当前的主流方法。图卷积神经网络是针对单个静态图结构进行特征提取而设计的网络,对于时序信息的提取较为薄弱。而人体骨架信息是一种时序连续的图结构数据,也可以视为一种动态图数据。对于动作识别这一任务而言,仅仅捕捉静态图空间结构信息(单帧骨架信息)而忽略时序信息并不能取得令人满意的性能。一般认为,对于仅需要单帧静态信息即可区分的动作,则基于图卷积神经网络的方法能取得较好的性能;而有些动作由于静态帧与其他动作很相似,需要辅以时序动作信息才能区分,此时需要模型具有较好的时序信息的建模能力。

[0003] 当前许多基于图卷积神经网络的方法的设计重心在捕捉空间结构信息上,通过定义自适应性邻接矩阵、新的图结构建模方法、新的节点连接等方式,提高了模型的性能。这些方法相较首次将GCN应用在人体骨架动作识别这一任务上的方法ST-GCN,都取得了一定的性能提升。但是在时序信息的建模上,这些方法都仅仅是简单地沿用了ST-GCN所使用的二维卷积来建模时序信息,并没有做较大的改进。

[0004] 在基于RGB视频的方法中,建模时序信息与建模时空信息的交互一直是重要的课题,研究者们使用光流模态来建模运动信息或者使用3D卷积网络来同时建模时间和空间信息。近年来,在基于RGB视频的动作识别方法中,基于卷积神经网络的方法Slowfast取得了极大的成功。

发明内容

[0005] 本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

[0006] 为此,本发明的第一个目的在于提出一种基于快慢双流图卷积神经网络的动作识别方法,是基于图卷积神经网络的方法上设计的一种使用快慢双流图卷积神经网络来更好地捕捉时序信息和快慢动作信息的方法,以提高动作识别的准确性。

[0007] 本发明的第二个目的在于提出一种基于快慢双流图卷积神经网络的动作识别系统。

[0008] 本发明的第三个目的在于提出一种计算机设备。

[0009] 本发明的第四个目的在于提出一种非临时性计算机可读存储介质。

[0010] 为达上述目的,本发明第一方面实施例提出了一种基于快慢双流图卷积神经网络的动作识别方法,包括以下步骤:

[0011] 步骤S10,获取人体骨架关节特征;

[0012] 步骤S20,对所述人体骨架关节特征进行正则化处理,其中,对一个批的所述人体

骨架关节特征的形状进行变形,使用一维正则化模块对时序维进行正则化,再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状;

[0013] 步骤S30,对步骤S20处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习,并对所述快分支和慢分支的学习结果进行融合,以获得每个动作类别的特征,其中,快慢双流图卷积网络的快分支和慢分支具有相同的网络结构,具有不同的网络参数配置和输入特征;

[0014] 步骤S40,通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数。

[0015] 可选地,在本申请的一个实施例中,所述步骤S10包括以下步骤:

[0016] 从数据集中获取人体骨架关节特征,每个样本的特征形状为:

[0017] (C, T, M, V)

[0018] 其中,C是特征通道数,值为3,表示关节点的三维坐标(x, y, z);T表示动作的帧数;M表示表演动作的人数;V表示人体关节点的数量。

[0019] 可选地,在本申请的一个实施例中,所述步骤S20包括以下步骤:

[0020] 对数据进行正则化处理,在训练过程中使用批训练,所述一个批的张量的特征形状为:

[0021] (B, C, T, M, V)

[0022] 首先对所述一个一维批张量进行变形,变形为:

[0023] $(B, M*V*C, T)$

[0024] 然后使用一个所述一维批正则化模块对时序T维进行正则化,并将特征重新变形为原来的形状 (B, C, T, M, V) 。

[0025] 可选地,在本申请的一个实施例中,步骤S30中具体步骤包括:

[0026] 在每个分支中,都包含连续叠加的多个图卷积区块,每个图卷积区块包括空间图卷积层和时序卷积层;时序卷积层是一个二维卷积模块,卷积核大小为(t, 1),t即为卷积核的时序感受野;两个卷积层后都附有批正则化层和ReLU激活函数,保证各个通道的特征保持相同的分布;图卷积区块的计算使用如下公式描述:

[0027]
$$f_{out} = \sum_k^{K_v} W_k f_{in} (A_k + B_k + C_k)$$

[0028] B_k 和 C_k 是在2s-AGCN中提出的自适应性的邻接矩阵,在网络训练的过程中会有改变,其中, B_k 初始化时被设置为 A_k ,用于学习任意两个节点潜在的关联; C_k 是根据样本特征计算得到的矩阵,用于描述样本特定的节点关联。

[0029] 可选地,在本申请的一个实施例中,下面两个公式分别描述了同一个阶段的图卷积区块的输入特征的特征形状:

[0030] $f_{fastin} = (B, \beta C, \alpha T, V, M)$

[0031] $f_{slowin} = (B, C, T, V, M)$

[0032] 快分支的时序维度始终为 αT_1 , α 是一个正整数,表示在初始输入特征中,快分支的输入帧率与慢分支的帧率的比值,在快分支中,通道数量 $\beta_i C_i$ 显著小于同阶段慢分支图卷积

区块通道数 C_i ,其中 i 是区块编号, β_i 是一个小于1的值,如1/3,两个分支的 V 是一致的,都是图节点的数量。

[0033] 可选地,在本申请的一个实施例中,使用横向连接模块来共享快慢两个分支学习到的信息,从快分支融合到慢分支,由于 $f_{fastout}^i$ 和 $f_{slowout}^i$ 的特征形状分别为 $(B, \beta C, \alpha T, V, M)$ 和 (B, C, T, M, V) ,首先采用一个二维卷积层来进行特征形状转换,并在进行特征形状转换之后增加了批正则化层和ReLU函数,然后用拼接或相加的方式将两个特征融合。

[0034] 可选地,在本申请的一个实施例中,步骤S40中,将步骤S30得到的最终特征通过一个全局池化层消去时序 T 、图节点 V 、人数 M 三个维度,并通过一个全连接层将特征映射到各个动作类别上,最终,通过Softmax函数得到每个动作类别的分数。

[0035] 为达上述目的,本申请第二方面实施例提出本发明一种基于快慢双流图卷积神经网络的动作识别系统,包括以下模块:

[0036] 获取模块,用于获取人体骨架关节特征;

[0037] 处理模块,用于对所述人体骨架关节特征进行正则化处理,其中,对一个批的所述人体骨架关节特征的形状进行变形,使用一维正则化模块对时序维进行正则化,再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状;

[0038] 生成模块,用于对所述处理模块处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习,并对所述快分支和慢分支的学习结果进行融合,以获得每个动作类别的特征,其中,快慢双流图卷积网络的快分支和慢分支具有相同的网络结构,具有不同的网络参数配置和输入特征;

[0039] 确定模块,用于通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数。

[0040] 为达上述目的,本申请第三方面实施例提出了一种计算机设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时,实现本申请第一方面实施例所述的基于关节组间关联建模的动作识别方法。

[0041] 为达上述目的,本申请第四方面实施例提出了一种非临时性计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现本申请第一方面实施例所述的基于关节组间关联建模的动作识别方法。

[0042] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0043] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0044] 图1为本申请实施例的一种基于快慢双流图卷积神经网络的动作识别方法的流程图。

[0045] 图2是本申请实施例的快慢双流图卷积神经网络结构示意图;

[0046] 图3是本申请实施例的快慢两个分支的输入特征的特征形状随着图卷积块编号增加而改变情况的示意图；

[0047] 图4是本申请实施例的横向连接模块示意图。

[0048] 图5为本申请实施例的一种基于快慢双流图卷积神经网络的动作识别系统的结构示意图。

具体实施方式

[0049] 下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

[0050] 下面参考附图描述本发明实施例的一所提供的一种基于快慢双流图卷积神经网络的动作识别方法。

[0051] 如图1所示，为达上述目的，本发明第一方面实施例提出了一种基于快慢双流图卷积神经网络的动作识别方法，包括以下步骤：

[0052] 步骤S10，获取人体骨架关节特征；

[0053] 步骤S20，对所述人体骨架关节特征进行正则化处理，其中，对一个批的所述人体骨架关节特征的形状进行变形，使用一维正则化模块对时序维进行正则化，再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状；

[0054] 步骤S30，对步骤S20处理后的所述人体骨架关节特征进行复制，生成两份相同的所述人体骨架关节特征，将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习，并对所述快分支和慢分支的学习结果进行融合，以获得每个动作类别的特征，其中，快慢双流图卷积网络的快分支和慢分支具有相同的网络结构，具有不同的网络参数配置和输入特征；

[0055] 步骤S40，通过全局池化层对每个动作类别的特征进行维度消除，并通过全连接层将消除维度后的特征映射到对应的各个动作类别上，并通过Softmax函数获得每个动作类别的分数。

[0056] 在本申请的一个实施例中，进一步而言，所述步骤S10包括以下步骤：

[0057] 从诸如NTU RGB+D等公开数据集中获取人体骨架关节特征，每个样本的特征形状为：

[0058] (C, T, M, V)

[0059] 其中，C是特征通道数，值为3，表示关节点的三维坐标(x, y, z)；T表示动作的帧数；M表示表演动作的人数；V表示人体关节点的数量。

[0060] 在本申请的一个实施例中，进一步而言，所述步骤S20包括以下步骤：

[0061] 对数据进行正则化处理，在训练过程中使用批训练，所述一个批的张量的特征形状为：

[0062] (B, C, T, M, V)

[0063] 首先对所述一个一维批张量进行变形，变形为：

[0064] (B, M*V*C, T)

[0065] 然后使用一个所述一维批正则化模块对时序T维进行正则化，并将特征重新变形

为原来的形状 (B, C, T, M, V) 。

[0066] 如图2所示,我们的网络结构包含两个分支,我们分别称为快分支和慢分支。

[0067] 在本申请的一个实施例中,进一步而言,步骤S30中具体步骤包括:

[0068] 在每个分支中,都包含连续叠加的多个图卷积区块,每个图卷积区块包括空间图卷积层和时序卷积层;时序卷积层是一个二维卷积模块,卷积核大小为 $(t, 1)$, t 即为卷积核的时序感受;两个卷积层后都附有批正则化层和ReLU激活函数,保证各个通道的特征保持相同的分布;图卷积区块的计算使用如下公式描述:

$$[0069] \quad f_{out} = \sum_k^{K_v} W_k f_{in}(A_k + B_k + C_k)$$

[0070] B_k 和 C_k 是在 2s-AGCN 中提出的自适应性的邻接矩阵,在网络训练的过程中会有改变,其中, B_k 初始化时被设置为 A_k , 用于学习任意两个节点潜在的关联; C_k 是根据样本特征计算得到的矩阵,用于描述样本特定的节点关联。

[0071] 可选地,在本申请的一个实施例中,下面两个公式分别描述了同一个阶段的图卷积区块的输入特征的特征形状:

$$[0072] \quad f_{fastin} = (B, \beta C, \alpha T, V, M)$$

$$[0073] \quad f_{slowin} = (B, C, T, V, M)$$

[0074] 快分支的时序维度始终为 αT_1 , α 是一个正整数,表示在初始输入特征中,快分支的输入帧率与慢分支的帧率的比值,在快分支中,通道数量 $\beta_i C_i$ 显著小于同阶段慢分支图卷积区块通道数 C_i , 其中 i 是区块编号, β_i 是一个小于 1 的值,如 1/3, 两个分支的 V 是一致的,都是图节点的数量。

[0075] 如图3所示,在本申请的一个实施例中,进一步而言,假设网络结构中有 N 个卷积区块,在慢分支中,在图卷积区块中的时序卷积层中,我们将通过时序卷积层的步长降低帧率,因此有 $T_1 \geq T_2 \geq \dots \geq T_N$; 另一方面,在每个图卷积区块中,输出通道数会随着图卷积区块的增加逐渐增加,以提高慢分支对于图空间结构信息的捕捉能力,因此有 $C_1 \leq C_2 \leq \dots \leq C_N$ 。在快分支中,在所有图卷积区块的时序卷积层中,卷积核的步长都设置为 1, 以保证帧率不会下降,因此,快分支的时序维度始终为 αT_1 , α 是一个正整数,表示在初始输入特征中,快分支的输入帧率与慢分支的帧率的比值。在快分支中,通道数量 $\beta_i C_i$ 显著小于同阶段慢分支图卷积区块通道数 C_i , 其中 i 是区块编号, β_i 是一个小于 1 的值,如 1/3。两个分支的 V 是一致的,都是图节点的数量。

[0076] 在本申请的一个实施例中,进一步而言,如图4所示,使用横向连接模块来共享快慢两个分支学习到的信息,从快分支融合到慢分支,由于 $f_{fastout}^i$ 和 $f_{slowout}^i$ 的特征形状分别为 $(B, \beta C, \alpha T, V, M)$ 和 (B, C, T, M, V) , 首先采用一个二维卷积层来进行特征形状转换,并在进行特征形状转换之后增加了批正则化层和ReLU函数,然后用拼接或相加的方式将两个特征融合。

[0077] 我们首先采用了一个二维卷积层来进行特征形状转换,并在这之后增加了批正则化层和ReLU函数,然后用拼接或者相加的方式将两个特征融合。以上过程可以用如下公式描述。

$$[0078] \quad f_{fast_in}^{i+1} = f_{fast_out}^i$$

$$[0079] \quad f_{inter}^{i+1} = \text{ReLU}(\text{BN}(\text{Conv2D}\{f_{fast_out}^i\}))$$

$$[0080] \quad f_{slow_in}^{i+1} = \text{Fuse}(f_{inter}^{i+1}, \{f_{slow_out}^i\})$$

[0081] 其中Conv2D是二维卷积层,BN是批正则化层,ReLU是激活函数,Fuse是融合函数,融合的方式可以采用加和(Sum)、拼接(Concatenation)等方式,这两种方式性能接近。

[0082] 进一步地,在两个分支之间本实施例采用插入了横向连接模块用以共享两个模块之间的信息。在本实施例的实验中,你们采用了10个卷积区块,每个卷积区块中快分支和慢分支的输入通道数分别为3、128、128、128、128、256、256、256、512、512和3、32、32、64、64、64、64、64、128、128。

[0083] 在本申请的一个实施例中,进一步而言,步骤S40中,将步骤S30得到的最终特征通过一个全局池化层消去时序T、图节点V、人数M三个维度,并通过一个全连接层将特征映射到各个动作类别上,最终,通过Softmax函数得到每个动作类别的分数。

[0084] 如图5所示,为达上述目的,本申请第二方面实施例提出本发明一种基于快慢双流图卷积神经网络的动作识别系统,包括以下模块:

[0085] 获取模块,用于获取人体骨架关节特征;

[0086] 处理模块,用于对所述人体骨架关节特征进行正则化处理,其中,对一个批的所述人体骨架关节特征的形状进行变形,使用一维正则化模块对时序维进行正则化,再将一个批的所述人体骨架关节特征的形状重新变形为原来的形状;

[0087] 生成模块,用于对所述处理模块处理后的所述人体骨架关节特征进行复制,生成两份相同的所述人体骨架关节特征,将两份相同的所述人体骨架关节特征分别输入至快慢双流图卷积网络的快分支和慢分支进行特征学习,并对所述快分支和慢分支的学习结果进行融合,以获得每个动作类别的特征,其中,快慢双流图卷积网络的快分支和慢分支具有相同的网络结构,具有不同的网络参数配置和输入特征;

[0088] 确定模块,用于通过全局池化层对每个动作类别的特征进行维度消除,并通过全连接层将消除维度后的特征映射到对应的各个动作类别上,并通过Softmax函数获得每个动作类别的分数。

[0089] 为了实现上述实施例,本发明还提出一种计算机设备,该计算机设备包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,处理器执行计算机程序时,实现本申请实施例的基于快慢双流图卷积神经网络的动作识别方法。

[0090] 为了实现上述实施例,本发明还提出一种非临时性计算机可读存储介质,该计算机存储有计算机程序,所述计算机程序被处理器执行时,实现本申请实施例的基于快慢双流图卷积神经网络的动作识别方法。

[0091] 尽管参考附图详细地公开了本申请,但应理解的是,这些描述仅仅是示例性的,并非用来限制本申请的应用。本申请的保护范围由附加权利要求限定,并可包括在不脱离本申请保护范围和精神的情况下针对发明所作的各种变型、改型及等效方案。

[0092] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特

点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不必针对的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0093] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中,“多个”的含义是至少两个,例如两个,三个等,除非另有明确具体的限定。

[0094] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本发明的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本发明的实施例所属技术领域的技术人员所理解。

[0095] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为是用于实现逻辑功能的可执行指令的定序列列表,可以具体实现在任何计算机可读介质中,以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用,或结合这些指令执行系统、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPROM或闪速存储器),光纤装置,以及便携式光盘只读存储器(CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0096] 应当理解,本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如,如果用硬件来实现和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0097] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0098] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机

可读取存储介质中。

[0099] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

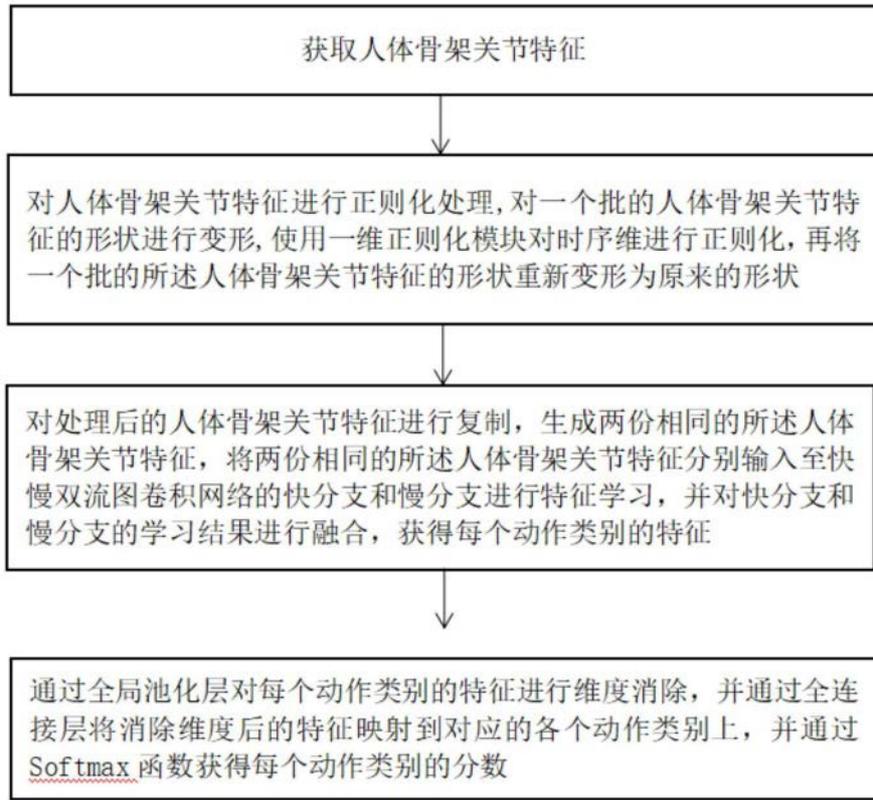


图1

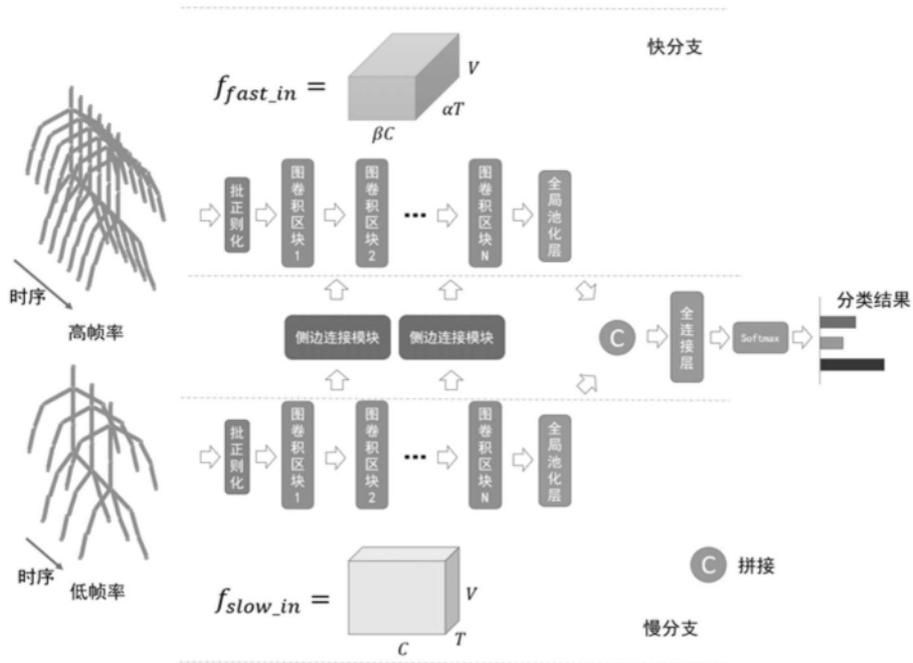


图2

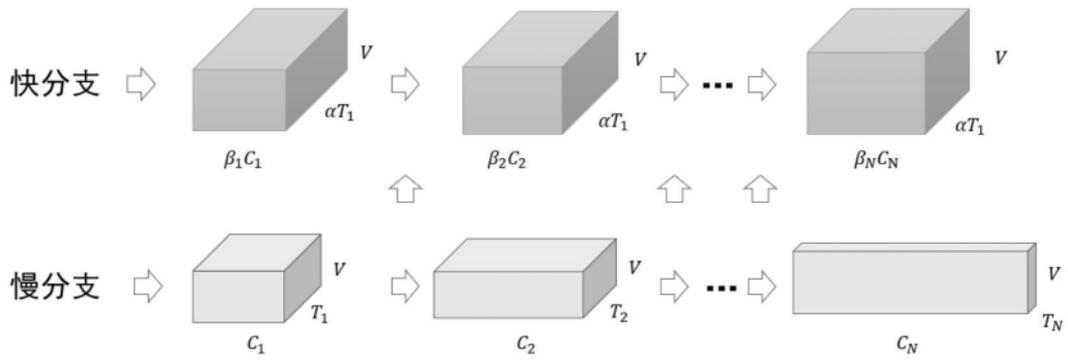


图3

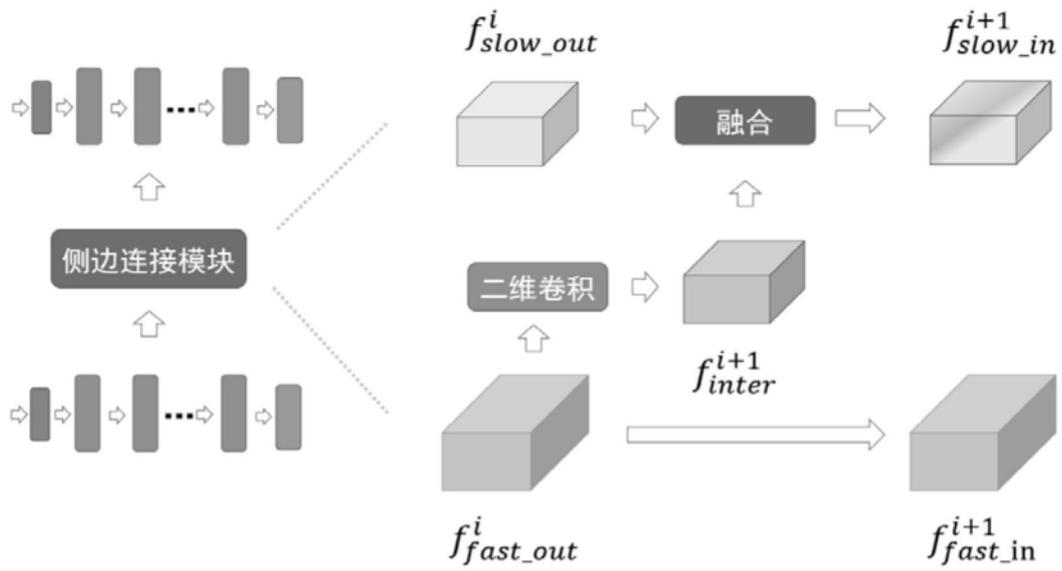


图4



图5