

1、一种动作识别方法，包括：

对原始压缩视频数据进行分组处理，得到分组视频数据；

将所述分组视频数据输入至第一预设模型中，并根据所述第一预设模型的输出结果确定包含动作的目标分组视频数据；

对所述目标分组视频数据进行解码，得到待识别分组视频数据；

将所述待识别分组视频数据输入至第二预设模型中，并根据所述第二预设模型的输出结果确定所述待识别分组视频数据中包含的动作类型。

2、根据权利要求1所述的方法，其中，所述对原始压缩视频数据进行分组处理，得到分组视频数据，包括：

基于预设分组规则对原始压缩视频数据进行区间划分，得到区间压缩视频；

基于预设提取策略提取所述区间压缩视频中的关键帧 I 帧数据和前向预测编码帧 P 帧数据，得到分组视频数据，其中，所述 P 帧数据包含 P 帧对应的以下至少一种信息：运动矢量信息，色素变化残差信息。

3、根据权利要求2所述的方法，其中，所述第一预设模型中包含第一 2D 残差网络、第一拼接层和第一全连接层；

所述分组视频数据被输入至所述第一预设模型中后，经由所述第一 2D 残差网络得到对应的维度相同的特征图；

所述特征图经由所述第一拼接层得到按照帧的先后顺序进行拼接操作后的拼接特征图；

所述拼接特征图经由所述第一全连接层得到是否包含动作的分类结果。

4、根据权利要求3所述的方法，其中，基于预设提取策略提取所述区间压缩视频中的 I 帧数据和 P 帧数据，得到分组视频数据，包括：

基于预设提取策略提取所述区间压缩视频中的 I 帧数据和 P 帧数据；

对所述 P 帧数据进行累加变换，以使得变换后的 P 帧数据依赖于前向相邻的所述 I 帧数据；

根据所述 I 帧数据和变换后的 P 帧数据确定分组视频数据；

所述第一预设模型中还包括位于所述拼接层之前的相加层，所述特征图中对应 P 帧数据的特征图记为 P 帧特征图，所述特征图中对应 I 帧数据的特征图记为 I 帧特征图；

所述 P 帧特征图和所述 I 帧特征图经由所述相加层得到在所述 I 帧特征图基

础上经过相加操作后的P帧特征图；

所述I帧特征图和所述经过相加操作后的P帧特征图经由所述第一拼接层得到按照帧的先后顺序进行拼接操作后的拼接特征图。

5、根据权利要求3所述的方法，其中，在所述第一2D残差网络的残差结构前包括特征变换层；

所述分组视频数据在进入所述残差结构前，经由所述特征变换层得到以下至少一种分组视频数据：经过向上特征变换的分组视频数据，经过向下特征变换的分组视频数据。

6、根据权利要求2所述的方法，其中，基于预设提取策略提取所述区间压缩视频中的P帧数据，包括：

采用等间隔方式提取所述区间压缩视频中的预设数量的P帧数据；

其中，在所述第一预设模型的训练阶段，采用随机间隔方式提取区间压缩视频中的预设数量的P帧数据。

7、根据权利要求1所述的方法，其中，所述对所述目标分组视频数据进行解码，得到待识别分组视频数据，包括：

对所述目标分组视频数据进行解码，得到待识别分段视频图像；

获取所述待识别分段视频图像中的频域信息，根据所述频域信息生成对应的频域图；

将所述待识别分段视频图像和对应的频域图作为待识别分组视频数据。

8、根据权利要求7所述的方法，其中，所述第二预设模型包括基于用于在线视频理解的高效卷积网络ECO架构的模型。

9、根据权利要求8所述的方法，其中，所述第二预设模型包括第二拼接层、第二2D残差网络、3D残差网络、第三拼接层和第二全连接层；

所述待识别分组视频数据被输入至第二预设模型中后，经由所述第二拼接层得到对所述待识别分段视频图像和对应的频域图进行拼接后的拼接图像数据；

所述拼接图像数据经由所述第二2D残差网络得到2D特征图；

将所述第二2D残差网络的中间层输出结果作为所述3D残差网络的输入，并经由所述3D残差网络得到3D特征图；

所述2D特征图和所述3D特征图经由所述第三拼接层得到拼接后的待识别特征图；

所述待识别特征图经由所述第二全连接层得到对应的动作类型标签。

10、根据权利要求 9 所述的方法，其中，所述第二预设模型还包括第一池化层和第二池化层；

所述 2D 特征图在被输入至所述第三拼接层之前，经由所述第一池化层得到对应的包含第一元素数量的一维 2D 特征向量；

所述 3D 特征图在被输入至所述第三拼接层之前，经由所述第二池化层得到对应的包含第二元素数量的一维 3D 特征向量；

所述 2D 特征图和所述 3D 特征图经由所述第三拼接层得到拼接后的待识别特征图，包括：

所述一维 2D 特征向量和所述一维 3D 特征向量经由所述第三拼接层得到拼接后的待识别向量。

11、根据权利要求 10 所述的方法，其中，所述第一池化层包括多感受野池化层。

12、根据权利要求 11 所述的方法，其中，所述第一池化层包括一级局部池化层、二级全局池化层和向量合并层，所述一级局部池化层中包含至少两个尺寸不相同的池化核；

所述 2D 特征图经由所述一级局部池化层得到对应的至少两组不同尺度的 2D 池化特征图；

所述至少两组不同尺度的 2D 池化特征图经由所述二级全局池化层得到至少两组特征向量；

所述至少两组特征向量经由所述向量合并层得到对应的包含第一元素数量的一维 2D 特征向量。

13、一种动作识别装置，包括：

视频分组模块，设置为对原始压缩视频数据进行分组处理，得到分组视频数据；

目标分组视频确定模块，设置为将所述分组视频数据输入至第一预设模型中，并根据所述第一预设模型的输出结果确定包含动作的目标分组视频数据；

视频解码模块，设置为对所述目标分组视频数据进行解码，得到待识别分组视频数据；

动作类型识别模块，设置为将所述待识别分组视频数据输入至第二预设模型中，并根据所述第二预设模型的输出结果确定所述待识别分组视频数据中包含的动作类型。

14、一种计算机设备，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，所述处理器执行所述计算机程序时实现如权利要求 1-12 任一项所述的方法。

15、一种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如权利要求 1-12 中任一所述的方法。