基于深度学习的人体行为识别算法综述

人体行为识别在机器人交互、虚拟现实、家庭和公共安全等领域的广泛应用， 使其正逐渐成为计算机视觉最活跃的研究领域之一。基于机器视觉的人体行为识别是将包含人体动作的视频添加上动作类型的标签。

人体行为识别工作主要分为两个过程：特征表征和动作的识别及理解。目前的识别算法和模型可以大概地分为两类，一类是基于传统手选特征的识别算法， 一类是基于深度学习的识别算法和模型。其中，基于深度学习的方法在各类具有挑战性的视频数据集上展现出了优于传统方法的较大优势。尽管如此，如何准确地区分不同类别的行为动作仍然存在巨大的挑战性。 比如光照或遮挡等视频环境因素、动作类别的类间和类内差异、视频数据集较少，这些问题都对鲁棒特征提取和动作分类构成了巨大挑战。

深度学习算法在近些年来，特别是模式识别与计算机视觉领域，大大改善了语音识别、视觉对象识别、对象检测以及许多其他领域（如药物发现、基因组学等）的最新技术。特别是深度卷积神经网络在处理图像、视频、语音音频等方面的相关问题带来了突破性的进展。因此，研究基于深度学习的人体动作识别问题具有十分重要的理论价值与现实意义。

**人体动作识别研究概述**

**Ⅰ 传统的特征提取方法**

动作识别近几年蓬勃发展，不断涌现各种算法、理论、技术。基于非深度学

习算法的特征提取有三类：低层特征、中层和高层特征、剪影。

传统特征提取方法一般是通过人工观察和设计，手动设计出能够表征动作特征的特征提取方法。 人体行为识别特征提取方法主要分为两部分： 基于人体几何或者运动信息的特征提取方法和基于时空兴趣点的特征提取方法。主要包括基于人体几何特征或运动信息的特征提取方法，基于人体几何特征或运动信息的特征提取方法。

**Ⅱ 基于深度学习的特征提取方法**

由于深度网络 可以无监督地从数据中学习到特征， 而这种学习方式也符合人类感知世界的机理， 因此当训练样本足够多的时候通过深度网络学习到的特征往往具有一定的语义特征， 并且更适合目标和行为的识别。

基于3D卷积神经网络的行为识别

卷积神经网络（Convolution neural network，CNN）是基于深度学习理论的一种人工神经网络，它主要利用权值共享来减小普通神经网络中的参数膨胀问题并在前向计算过程中使用卷积核对输入数据进行卷积操作，将得到的结果通过一个非线性函数作为该层的输出，这样的层称为卷积层，卷积层和卷积层之间会出现下采样层，下采样层主要用于获取局部特征的不变性，同时降低特征空间的尺度。一般在卷积层和下采样层之后是一个全连接的神经网络用于最终的识别。

递归神经网络及其扩展模型

在深度学习领域，传统的前馈神经网络（Feedforward neural net，FNN）取得了显著的成就。 但近年来随着研究的深入， FNN 模型对声音、 文本、视频等信息表征时， 无法学习到信息的逻辑顺序。 为解决这一问题， 能够反映序列前后关联信息的递归神经网络 （Recurrent neural networks， RNN）［发展迅速。RNN将上几个时刻的隐含层数据作为当前时刻的输入，从而允许时间维度上的信息得以保留。

结论

总体来说，非深度学习算法的特征提取大多是基于人工先验知识的设计，且

对视频的环境、拍摄条件有较高的要求。深度学习算法无需手动选取特征就可在

视频数据上学习得到最有效的表征，对数据具有很强的适应性。

通过以上工作,可看出人体动作行为识别的研究取得卓越的成果和进展。

但同时需要注意的是很多方法仍处于实验室研究阶段，能够应用在实际的监控视频中，如何能够更加有效地应用于社会生活中还需要进一步的实验研究。此外应该研究算法如何在进一步提高识别准确率的同时，提高时效性与即时性。

人体动作行为识别是一个综合性课题,模式分类的数学方法、图像传感技术以及人体运动学研究中的创新理论在动作识别领域的应用是推动其发展的重要动力。

参考文献

[1]惠通. 基于轨迹和卷积神经网络的人体行为识别方法[D].西安电子科技大学,2017.

吴军, 肖克聪. 基于深度卷积神经网络的人体动作识别[J]. 华中科技大学学报(自然科学版), 2016, 44(s1).