

武术动作矫正项目简介

嵌入式人工智能

梁东 惠佳煊 韦东健

2024 年 5 月 28 日



① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

为什么选择武术动作检测项目

- 武术动作检测本质上是动作检测. 动作检测是是机器学习的重要研究内容. 因此在网络上存在大量的资料和项目可供参考.

为什么选择武术动作检测项目

- 武术动作检测本质上是动作检测. 动作检测是是机器学习的重要研究内容. 因此在网络上存在大量的资料和项目可供参考.
- 武术动作检测是一件有意义的事情.
首先, 动作检测在生活中存在广泛的应用空间. 就拿医学来说, 老年人健康检查, 行为评估, 疾病筛查都是很有发展前景的领域. 其次, 对于武术来说, 如果能开发一套智能的武术学习系统, 也能帮助大家更好接触学习传统武术.

为什么选择武术动作检测项目

- 武术动作检测本质上是动作检测. 动作检测是是机器学习的重要研究内容. 因此在网络上存在大量的资料和项目可供参考.
- 武术动作检测是一件有意义的事情.
首先, 动作检测在生活中存在广泛的应用空间. 就拿医学来说, 老年人健康检查, 行为评估, 疾病筛查都是很有发展前景的领域. 其次, 对于武术来说, 如果能开发一套智能的武术学习系统, 也能帮助大家更好接触学习传统武术.
- 我喜欢武术, 选择武术检测项目可以一箭双雕 (Doge)

1 项目背景

2 研究现状

目标检测研究现状
姿态估计研究现状
动作识别研究现状
动作评价研究现状

3 项目阶段性成果展示

4 项目流程

1 项目背景

2 研究现状

目标检测研究现状
姿态估计研究现状
动作识别研究现状
动作评价研究现状

3 项目阶段性成果展示

4 项目流程

- 传统目标检测算法

传统的目标检测算法通常使用机器学习和经典视觉图像处理技术对目标特征进行学习，精确定位目标在图像中的位置，并通过分类器得到目标类别。

- 传统目标检测算法

传统的目标检测算法通常使用机器学习和经典视觉图像处理技术对目标特征进行学习,精确定位目标在图像中的位置,并通过分类器得到目标类别。

- 基于深度学习的目标检测算法

基于候选区域的目标检测算法:首先获取目标的候选区域,其次由神经网络做分类和回归。

基于回归的目标检测算法:直接检测目标位置和类别。

1 项目背景

2 研究现状

目标检测研究现状
姿态估计研究现状
动作识别研究现状
动作评价研究现状

3 项目阶段性成果展示

4 项目流程

- 单人姿态估计单人姿态估计是对图像中独立的个体进行人体姿态估计, 通过算法输出该个体所有关键点。

- 单人姿态估计单人姿态估计是对图像中独立的个体进行人体姿态估计, 通过算法输出该个体所有关键点。
- 自顶向下的多人姿态估计
自顶向下的主要思路: 首先需要有一个目标检测模型, 用其检测出每个人体的边界框, 然后对每一个检测到的人体区域提取关键点来进行一次单人姿态估计, 最终实现多人姿态估计。

- 单人姿态估计单人姿态估计是对图像中独立的个体进行人体姿态估计, 通过算法输出该个体所有关键点。
- 自顶向下的多人姿态估计
自顶向下的主要思路: 首先需要有一个目标检测模型, 用其检测出每个人体的边界框, 然后对每一个检测到的人体区域提取关键点来进行一次单人姿态估计, 最终实现多人姿态估计。
- 自底向上的多人姿态估计
自底向上的主要思路: 与自顶向下恰恰相反, 首先检测图像中所有人体的关键点, 然后利用聚类的方法将这些人体关键点分配给不同的人体目标, 最终得到每个人的姿态信息。

1 项目背景

2 研究现状

目标检测研究现状
姿态估计研究现状
动作识别研究现状
动作评价研究现状

3 项目阶段性成果展示

4 项目流程

基于深度学习的动作识别算法

- 双流法

其本质是将采集到的视频划分为两种类型：空间流、时间流。其中，空间流是由视频中的任意一帧图像构成，时间流则是由该帧图像邻近帧之间的密集光流组成，分别构建两组二维卷积神经网络对上述空间流信息、时间流信息进行特征提取并在网络最后进行两种特征的融合，经过这种处理之后，模型学习到空间信息的同时也学习到空间信息的变化，实现动作识别。

基于深度学习的动作识别算法

- 双流法

其本质是将采集到的视频划分为两种类型：空间流、时间流。其中，空间流是由视频中的任意一帧图像构成，时间流则是由该帧图像邻近帧之间的密集光流组成，分别构建两组二维卷积神经网络对上述空间流信息、时间流信息进行特征提取并在网络最后进行两种特征的融合，经过这种处理之后，模型学习到空间信息的同时也学习到空间信息的变化，实现动作识别。

- 三维卷积法

三维卷积法在二维卷积的基础上，三维卷积法引入了时间维度，可以从连续视频帧中同时获取到时、空两种特征信息，以此实现动作识别。

基于深度学习的动作识别算法

- 基于人体骨骼的动作识别

基于人体骨骼的动作识别方法的基本原理是先由关键点检测网络获得关键点的位置信息,然后在时间、空间两个维度上分别提取骨骼关键点的动态变化信息,最终实现动作识别。

1 项目背景

2 研究现状

目标检测研究现状
姿态估计研究现状
动作识别研究现状
动作评价研究现状

3 项目阶段性成果展示

4 项目流程

- 采用什么方法评价动作？[1]
 1. 根据关节在某个动作中的重要性
 2. 根据大量专家打分的大数据，设计专家系统
 3. 定义规则，设置错误动作

- 采用什么方法评价动作？[1]
 1. 根据关节在某个动作中的重要性
 2. 根据大量专家打分的大数据，设计专家系统
 3. 定义规则，设置错误动作
- 目前大量使用定量数据，如角度、加速度。但是对于不容易定量的优美性和流畅度评价不足。

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

项目成果

- ESP32-CAM 网络摄像头

项目成果

- ESP32-CAM 网络摄像头
- Mediapipe 关节点识别

项目成果

- ESP32-CAM 网络摄像头
- Mediapipe 关节点识别
- S-ET-GCN 网络模型识别动作

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

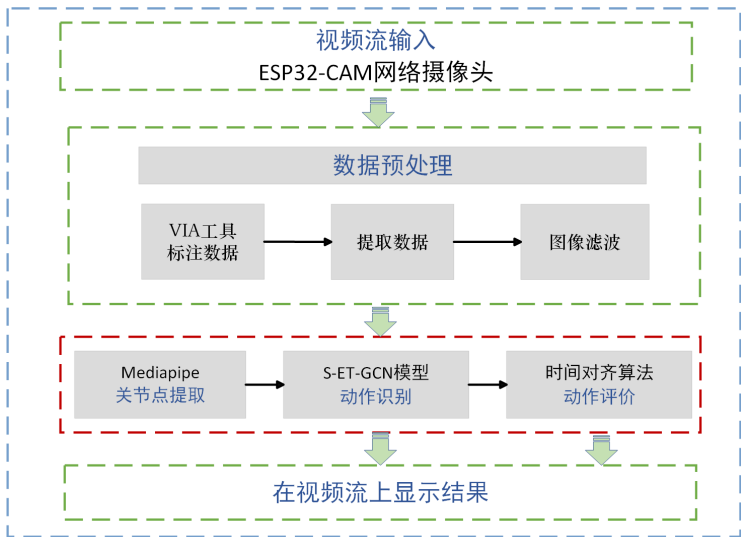
关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

流程图



① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

数据预处理

- 使用 VIA 工具标注视频片段中的各种动作.
- VIA 下载地址

`https://www.robots.ox.ac.uk/~vgg/software/via/`

注意:VIA3 才有标注视频的功能.

提取关键帧和图像滤波

- 我们把包含需要动作的图片叫做关键帧。
所谓提取关键帧就是将在 VIA 中标注的视频片段提取出来。
把他们的每一帧数据都放入一个列表中。
- 图像滤波, 过滤图像中的噪声。

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

利用 Mediapipe 检测关节点

- 利用 Mediapipe 检测关节点, 并将关节点导出. Mediapipe 是一个关节点检测模型, Mediapipe 使用简单, 识别迅速.
- 保存关节点数据为 JSON 数据文件, 准备提供给下一步 S-ET-GCN 模型训练.

① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

S-ET-GCN 模型简介

- GCN 模型
GCN(Graph Convolutional Network,GCN) 模型, 其本质是利用节点间的连接关系对节点的特征进行更新, 起到特征提取器的作用。
- ST-GCN 模型
ST-GCN 模型的输入为人体骨骼数据构建的骨架时空图, 利用空间图卷积网络 (Spatial Graph Convolutional Network,SGCN) 和时间卷积网络 (Temporal Convolutional Network,TCN) 聚合骨架时空图的空间和时间特征, 逐步生成更高层次的特征图, 实现散打动作识别。

S-ET-GCN 模型简介

- S-ET-GCN 模型: 扩展 TCN 层次并加入残差机制的 ST-GCN 动作识别网络 [3]

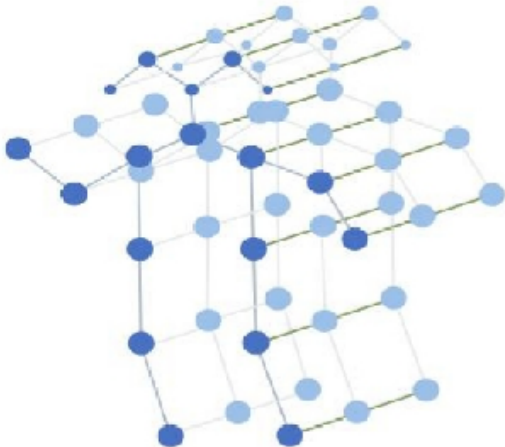
ST-GCN 模型的时间卷积网络中只用了一个标准卷积层不能充分提取时间特征。由于扩展 TCN 网络层次可以提高网络的表征能力, 使得网络对于一些细微的变化具有更好的鲁棒性; 残差模块可以通过跨层连接将输入数据与输出数据相加, 使得网络可以直接传递前一层的信息到后一层, 从而减少信息的损失。

S-ET-GCN 动作识别网络模型可以较好地捕捉到动作之间短期与长期的时间依赖关系, 提高动作识别准确率。



S-ET-GCN 模型简介

ST-GCN 骨架时空图



① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

时间规整算法

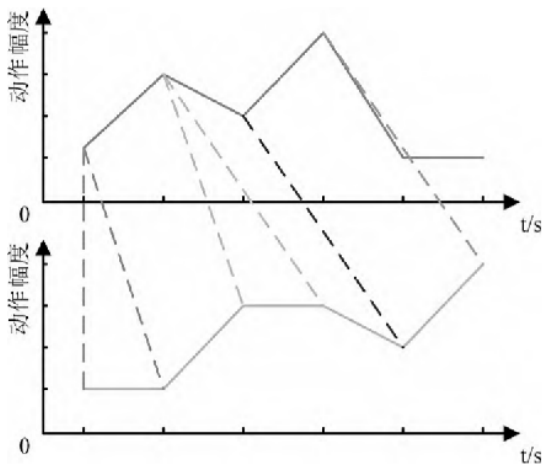
- 时间规整算法

大多武术爱好者完成武术动作的时间都会与标准时间存在差异, 所以需要对采集到的动作进行对齐。

动态时间规整算法能够在时序上对动作进行对齐和匹配, 从而实现动作的准确分析和比较。[2]

时间规整算法

时间规整算法示意图如下所示



① 项目背景

② 研究现状

③ 项目阶段性成果展示

④ 项目流程

流程图

数据预处理

关节点检测

S-ET-GCN 模型

时间规整算法 (未实现)

ESP32-CAM 网络摄像头

ESP32-CAM 网络摄像头

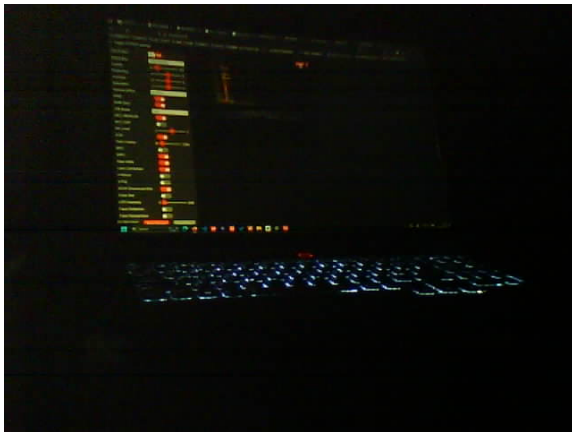
网络摄像头使用的硬件设备是 ESP32-CAM.

配置教程可以参考https://www.cnblogs.com/qiaoyurensheng/articles/esp32cam_001.html

1. 我们将 ESP32 配置为网络摄像头, 通过 Arduino 的串口监视器获取 ESP32-CAM 传输视频的地址. 然后打开地址, 选择适合的分辨率, 开始录制.
2. 关闭网页上的图像 (注意: 只点图上的小 x, 而不是关闭网页.)
3. 在程序中的视频接口处, 输入网址:
`192.168.xx.xxx:80/stream`

ESP32-CAM 网络摄像头

ESP32-CAM 拍到的图片.



- [1] 宋震, 张宇姝, and 杨刚.
人体动作识别与评价综述.
中国传媒大学学报 (自然科学版), 28(03):58–65, 2021.
- [2] 常琪.
基于动态时间规整算法的健美操动作分析及评价系统设计研究.
自动化与仪器仪表, pages 136–139+144, 2024.
- [3] 张冰冰.
基于深度学习的运动员动作识别方法研究.
Master's thesis, 沈阳工业大学, 2023.

致谢

谢谢卢红旺老师!

Thanks!