基于视频的体育动作标准度评估算法

摘要： 本研究旨在开发一种基于视频的体育动作标准度评估算法，以解决传统姿态检测方法在设备依赖性和实时反馈方面的不足。首先本人对当前体育训练中的问题和需求进行了分析，发现传统评估方法在设备成本和实时性上存在限制。随后我通过研究现有的计算机视觉和机器学习技术，选择了MoveNet人体骨骼点检测模型作为基础框架，并结合姿态匹配技术，实现了对深蹲、引体向上动作的识别和评估。在算法实现过程中，我逐步克服了姿态匹配精度和动作实时性滞后等困难，通过不断优化算法和调整模型参数提高了计算效率，有效改善了上述问题。最终我搭建了一套简要的实时评分系统，其能够在动作进行过程中及时给出评分和反馈。测试结果表明，新算法在准确性和实时性上均达到了较高水平，为体育训练提供了一种全新的高效评估方法。本研究的成果将为体育教练员、运动员和普通用户提供更加便捷有效的训练辅助工具。

关键词：计算机视觉；TensorFlow；神经网络；MoveNet；姿态检测；

Algorithm for Evaluating Sports Action Standardization Based on Video Streams

Abstract: This study aims to develop a video-based algorithm for assessing the standardization of sports movements, addressing the shortcomings of traditional methods in device dependency and real-time feedback. Firstly, an analysis of the current problems and demands in sports training was conducted, revealing limitations of traditional assessment methods in terms of device costs and real-time constraints. Subsequently, by researching existing computer vision and machine learning technologies, the MoveNet human body skeleton key point detection model was selected as the foundational framework. Combined with pose matching techniques, recognition and assessment of squatting and pull-up movements were achieved. Throughout the algorithm implementation process, challenges such as pose matching accuracy and real-time lag in movement were gradually overcome. This was achieved through continuous algorithm optimization and adjustment of model parameters, thereby enhancing computational efficiency and effectively addressing the aforementioned issues. Ultimately, a concise real-time scoring system was constructed, providing timely scoring and feedback during movement execution. Test results indicate that the new algorithm achieves high levels of accuracy and real-time performance, offering a novel and efficient assessment method for sports training. The outcomes of this study will provide coaches, athletes, and general users with a more convenient and effective training aid.

Keywords: Computer Vision; TensorFlow; Neural Networks; MoveNet; Pose Detection;

目录

摘要 i

Abstract i

[目录 iii](#_Toc163050261)

[1 绪论 1](#_Toc163050262)

[1.1 课题的目的和意义 1](#_Toc163050263)

[1.2 国内外研究现状与水平 2](#_Toc163050264)

[1.3 项目创新点 2](#_Toc163050265)

[1.4 论文要求 3](#_Toc163050266)

[1.4.1 论文主要研究内容 5](#_Toc163050267)

[1.4.2 论文章节安排 5](#_Toc163050268)

[2 技术与工具 6](#_Toc163050269)

[2.1 采取技术 6](#_Toc163050270)

[2.1.1 Anaconda3 6](#_Toc163050271)

[2.1.2 TensorFlow 6](#_Toc163050272)

[2.1.3 OpenCV 6](#_Toc163050273)

[2.1.4 MoveNet 7](#_Toc163050274)

[2.2 开发工具 7](#_Toc163050275)

[2.2.1 硬件环境 7](#_Toc163050276)

[2.2.2 软件环境 7](#_Toc163050277)

[3 算法设计 8](#_Toc163050278)

[3.1 系统架构 8](#_Toc163050279)

[3.2 测量标准 8](#_Toc163050280)

[3.2.1 深蹲动作 8](#_Toc163050281)

[3.2.2 引体向上动作 9](#_Toc163050282)

[3.3 姿态检测 10](#_Toc163050283)

[3.3.1 模型初始化 10](#_Toc163050284)

[3.3.2 视频帧读取 11](#_Toc163050285)

[3.3.3 骨骼点识别 11](#_Toc163050286)

[3.3.4 骨骼点归一化 13](#_Toc163050287)

[3.4 匹配算法 14](#_Toc163050288)

[3.4.1 相似度计算 14](#_Toc163050289)

[3.4.2 相似度阈值过滤 15](#_Toc163050290)

[3.4.3 视频帧抖动过滤 15](#_Toc163050291)

[3.4.4 计数匹配 16](#_Toc163050292)

[3.5 动作评分 16](#_Toc163050293)

[3.6 优化措施 16](#_Toc163050294)

[4 函数简介 17](#_Toc163050295)

[4.1 视频帧读取 17](#_Toc163050296)

[4.2 姿态检测 17](#_Toc163050297)

[4.3 基准数据构建 17](#_Toc163050298)

[4.4 匹配与评分 18](#_Toc163050299)

[4.5 用户界面 18](#_Toc163050300)

[4.6 系统主控制 19](#_Toc163050301)

[5 系统演示 20](#_Toc163050302)

[5.1 环境配置 20](#_Toc163050303)

[5.2 运行实例 20](#_Toc163050304)

[5.2.1 本地视频模式 20](#_Toc163050305)

[5.2.2 摄像录制模式 20](#_Toc163050306)

[6 结论 21](#_Toc163050307)

[7 致谢 22](#_Toc163050308)

[8 参考文献 23](#_Toc163050309)

[附录 25](#_Toc163050310)

# 绪论

开发与研究基于视频的体育动作标准度评估算法是一项具有重要实践意义和广泛应用前景的任务。在现代体育领域，准确评估运动员的动作标准度对于提高竞技水平、预防运动伤害以及优化训练方法至关重要。

## 课题的目的和意义

随着体育科技的不断发展，运动员的训练和比赛姿态对于他们的表现和健康来说变得越来越关键。对运动动作姿态进行科学评估不仅有助于优化训练方法，预防运动伤害，也对提高运动效率和成绩至关重要。准确的姿态评估可以指导运动员进行更有效的训练，从而在竞技体育中取得更好的成绩。

同时，深度学习算法和姿态检测模型的快速发展与应用为体育运动姿态的评估提供了新的可能性。姿态检测模型能够在短时间内处理海量的运动数据，提供更加精准和全面的分析，这对于指导运动员的训练和比赛策略具有重大意义。例如，通过分析大量的运动数据，可以精确地识别出运动员在特定动作中的不足，从而提供针对性的训练建议。此外，利用姿态检测模型，配合机器学习和深度学习，可以自动化地分析运动姿态，为教练员和运动员提供即时反馈，极大地提高了训练的效率和效果。

除了在特定的竞技体育领域的应用，该课题研究对于普及体育运动和提高公众的身体健康也有积极作用：分析普通人的运动动作，不仅可以帮助他们进行更加安全有效的锻炼，还能够预防运动相关的伤害。这对于提高公众的整体健康水平和生活质量具有重要意义。

体育运动始终是现代生活的重要组成部分。开发基于视频的体育动作标准度评估应用算法，不仅对运动员和教练在高水平竞技体育领域至关重要，也对推广健康生活方式和提高公共健康水平具有深远影响。随着技术的不断发展，这一领域展现出广阔的研究和应用前景，预期将在未来带来更多的创新和突破。

从基础的训练到专业运动领域，姿态评估技术的进步不仅改善了运动员的表现和健康状况，还在更广泛的社会层面产生积极影响。通过对运动员的动作进行细致分析和调整，不仅能够提高他们的竞技水平，还可以减少运动相关的潜在伤害风险。这些技术的普及应用也有助于普通民众更好地理解正确的运动姿势，并在日常生活中保持健康的身体素质。因此，姿态评估技术的发展不仅在专业体育领域发挥重要作用，也在促进社会健康和生活质量方面发挥着不可或缺的作用。

## 国内外研究现状与水平

在国际层面上，体育动作姿态评估的研究已经取得了显著进展，特别是在运动生物力学、计算机视觉和人工智能等领域。多项研究表明，利用先进的传感器技术和高分辨率摄像头捕捉运动员的动作数据，结合人工智能，尤其是机器学习和深度学习技术，可以有效地对运动姿态进行分析和评估。

一些研究使用穿戴式传感器，如加速度计和陀螺仪，实时捕捉运动员的动作数据。这些传感器能够提供关于运动员运动速度、方向和加速度等详细信息。通过计算机视觉的应用，如采用高速摄像头捕捉运动员的动作，利用计算机视觉技术提取运动特征。这些技术能够生成运动员动作的三维重建，提供更加全面的运动分析。利用深度学习技术，特别是卷积神经网络（CNN）和循环神经网络（RNN），被广泛应用于运动数据的分析，例如对篮球、足球等运动中的动作进行深度学习分析，不仅有助于优化运动员的表现，还能有效减少受伤风险。

国内在体育动作姿态评估算法的研究方面也在逐步发展。随着科技的进步和人工智能技术的普及，越来越多的高校和研究机构开始关注运动姿态分析的重要性，并在此领域进行探索和实践。国内许多高校和研究机构已经开始布局运动生物力学和运动数据分析领域，投入资源进行相关研究。国内研究者们也在尝试将机器学习、深度学习等先进技术应用于运动姿态的分析，例如，一些研究通过分析羽毛球、乒乓球等运动的运动员姿态，以帮助提升运动员的技术水平和比赛表现，国内研究者正努力缩小与国际先进水平之间的差距。在国际和国内双重推动下，基于视频的体育动作标准度评估算法的研究和应用前景更加广阔，将为运动员、教练员以及整体社会的健康与发展带来重要的推动力。

## 项目创新点

本项目结合了先进的技术和实用的功能，为体育训练和评估提供了一种全新的方法，具有较高的创新性和实用性。本项目有多个方面的创新之处：

1. 基于视频的体育动作标准度评估：采用视频流作为输入，可同时接受视频与摄像头采集的画面，而不仅仅是传统的基于传感器的方法，有效提高可用性，因为许多人可能没有专业的传感器设备，但他们可能有摄像头或智能手机。
2. 采用新一代人体骨骼点检测模型和姿态匹配技术：该算法利用了先进的机器学习技术，使用MoveNet取代了传统的PoseNet模型，能够更加准确地检测和跟踪人体骨骼点，从而实现对深蹲和引体向上动作的准确识别和评估。
3. 实时评分和反馈：系统能够在动作进行过程中实时给出评分，而非是等到动作完成后才提供反馈。这种实时性可以有效帮助用户更快地调整姿势和动作，从而提高体育运动的训练效果。
4. 支持多科目切换：系统能够根据用户的需求切换识别的体育动作，包括深蹲和引体向上。该功能这增加了算法系统的灵活性，使其可以应用于不同类型的体育训练和评估场景。
5. 动作评分的定制化：系统提供了可调整的评分阈值参数，用户可以根据自己的需求和标准来定制评分规则。这种灵活性使系统可以适应不同用户和不同训练阶段的需求。

## 论文要求

在论文中图的用法和格式参考如下。

图的标题可以在插入图后“插入”菜单“引用”“题注”，标签选择“图”即可（没有“图”标签新建即可）。为了方便操作，图标题也可以复制后再修改。图片的样式选择“插图”，图标题的样式可以选择“图标题”。如果文中需要引用图片，可以在“插入”菜单“引用”“交叉引用”，引用类型选择“图”，引用内容“只有标签和编号”。例如：如图 ‎1.1所示。

注意如果在论文中粘贴屏幕截图，应选用合适的工具选取有说明性的部分，一般不宜截取整个屏幕。截图后用工具保存为JPG或者PNG格式，作为备份。然后根据论文中的位置修改到合适的尺寸，再插入到Word中。



图 ‎1.1 例图

（前一个1表示第1章，后一个1表示这一章的第1个图）

下面这段Word宏代码可以将全部图片设置为样式“插图”。

Sub SetAllImageStyle()

Dim doc As Document

Dim img As InlineShape

Set doc = ActiveDocument

For Each img In doc.InlineShapes

img.Select

Selection.Style = "插图"

Next

End Sub

论文中表格的用法参考如下。

表的标题可以在插入表格后选中整个表格后“插入”菜单“引用”“题注”，标签选择“表”即可。为了方便操作，表格标题也可以复制后再修改。表格标题的样式可以选择“表标题”。

表格标题的引用同图标题的引用。可以在“插入”菜单选择“交叉引用”，引用类型选择“表”，引用内容“只有标签和编号”。例如：数据如表 ‎1.1所列。

表格的居中和表格后空白设置：右键点击表格后选择“表格属性”，对齐方式选择“居中”，文字环绕选择“环绕”，定位，距正文，下：0.5厘米，左、右：10厘米即可。

下面这段Word宏代码可以将全部表格的对齐方式设置为“居中”。

Sub CenterAllTable()

Dim doc As Document

Dim tbl As Table

Set doc = ActiveDocument

For Each tbl In doc.Tables

tbl.Rows.Alignment = wdAlignRowCenter

Next

End Sub

表 ‎1.1 MsgBox 函数返回值

|  |  |  |
| --- | --- | --- |
| **常数** | **值** | **按钮** |
| vbOK | 1 | 确定 |
| vbCancel | 2 | 取消 |
| vbAbort | 3 | 放弃 |
| vbRetry | 4 | 重试 |
| vbIgnore | 5 | 忽略 |
| vbYes | 6 | 是 |
| vbNo | 7 | 否 |

论文字数要求15000字以上，是指正文部分。摘要、目录、参考文献、附录等不计在内，统计时请注意。



图 ‎1.2 全垒打

### 论文主要研究内容

### 论文章节安排

# 技术与工具

以下为本项目所使用到的技术与工具：

## 采取技术

### Anaconda3

Anaconda是一个用于科学计算的开源软件包管理系统和环境管理系统，由Anaconda, Inc.开发并维护。Anaconda集成了常用的Python数据科学工具和库，使得本项目中科学计算、数据分析、机器学习和大数据处理等任务运行调试更加便捷。

Anaconda的核心组件是Anaconda Distribution，它是一个包含了最新Python解释器、一系列常用科学计算库以及其他工具的集合。其中包括了诸如NumPy、SciPy、Pandas、Matplotlib等常用的科学计算和数据分析库，还囊括了用于机器学习的工具包如scikit-learn和TensorFlow等。

Anaconda Distribution的另一个重要特性是它自带了名为conda的包管理工具，它可以用于安装、升级、删除和管理软件包，以及创建和管理不同的Python环境。这使得用户可以轻松地管理不同项目的依赖关系，并在不同项目中使用不同版本的软件包，而不会相互干扰。

### TensorFlow

TensorFlow是由Google团队开发的深度学习框架之一，它是一个完全基于 Python 语言设计的开源的软件。TensorFlow 的初衷是以最简单的方式实现机器学习和深度学习的概念，它结合了计算代数的优化技术，使它便计算许多数学表达式。

TensorFlow 可以训练和运行深度神经网络，它能应用在许多场景下，比如，图像识别、手写数字分类、递归神经网络、单词嵌入、自然语言处理、视频检测等等。TensorFlow 可以运行在多个 CPU 或 GPU 上，同时它也可以运行在移动端操作系统上（如Android、iOS 等），它的架构灵活，具有良好的可扩展性，能够支持各种网络模型（如OSI七层和TCP/IP四层）。

### OpenCV

OpenCV (Open Source Computer Vision Library) 是一款由英特尔公司发起并维护的开源跨平台计算机视觉库。其提供了丰富的图像处理和计算机视觉功能，可广泛应用于图像处理、对象检测与跟踪、人脸识别、手势识别以及机器学习等各种视觉应用程序开发领域。OpenCV实现了多种通用算法，成为计算机视觉领域最重要的研究工具之一。

### MoveNet

MoveNet 是由 Google 开发的实时多人姿势估计模型。该模型旨在从视频流中检测和跟踪多个人体的姿势，包括身体的关键点位置和姿态。MoveNet 基于深度学习技术，采用了轻量级的神经网络架构，可以在移动设备和嵌入式系统上实现实时的姿势估计。

MoveNet 采用了深度可分离卷积（DepthwiseSeparable Convolution）和逐步扩张卷(Dilated Convolution)等操作，以减少参数数量和计算量，从而实现轻量级和高效的模型。MoveNet 模型的整体结构包括两个主要组件：移动网络（MobileNet）和动作分支（Pose orAction Branch）。MoveNet 整体的结构如上所示，Backbone 部分是比较经典的带三层 deconv 的 MobileNetv2，并且使用了残差连接来获取浅层特征。

以下图2.1为标准MoveNet模型结构图：

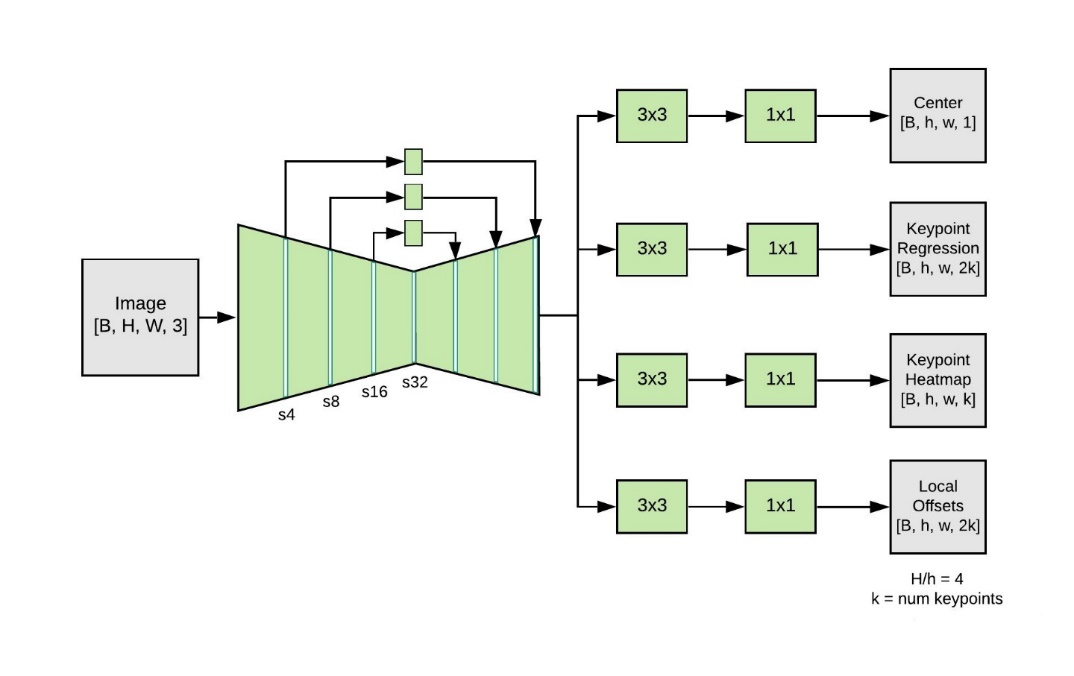


图 ‎2.1 MoveNet模型结构图

其中：

* Center Heatmap[B, h, w, 1]：预测人体的几何中心，主要用于存在性检测，用Heatmap上一个锚点来代替目标检测的Box。
* Keypoint Regression[B, h, w, 2k]: 基于中心点来回归17个关节点坐标值。
* Keypoint Heatmap[B, h, w, k]：每种类型的关键点使用一张Heatmap进行检测，这意味着多人场景下这张Heatmap中会出现多个高斯核。
* Local Offsets[B, h, w, 2k]：回归Keypoint Heatmap中各高斯核中心跟真实坐标的偏移值，用于消除Heatmap方法的量化误差。

MoveNet模型主要分为两个版本，可适用于不同的识别场景，在本项目中均会采用：

1. MoveNet Thunder

MoveNet Thunder 是 Google 开发的高精度版本的 MoveNet 模型。其在保持实时性的同时，专注于提高姿势估计的准确性和稳定性。这个版本通常用于对姿势估计精度要求较高的应用场景，如医疗诊断、运动分析等。

1. MoveNet Lightning

MoveNet Lightning 则是Google 开发的轻量级版本的 MoveNet 模型。其注重模型的小型化和性能的优化，以实现在资源受限的设备上实时运行。这个版本通常用于移动端和嵌入式设备，如智能手机、智能眼镜等，以提供实时的姿势估计功能。

## 开发工具

以下为本项目设计所使用的硬件环境与软件环境：

### 硬件环境

表格2-1为本次项目设计搭建所用到的设备硬件规格：

表格 ‎2‑1 设备硬件规格表

|  |  |
| --- | --- |
| **项目** | **规格** |
| CPU | 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz |
| GPU | Nvidia GeForce RTX 3070 Laptop |
| 硬盘 | 3TB Samsung PM9A1 |
| 内存 | 32GB DDR4 3200MHz |
| 摄像头\_0 | 640\*480 30fps内置摄像头，用于摄像测试 |
| 摄像头\_1 | 1920\*1080 30fps外置摄像头，用于采集图像数据 |

### 软件环境

操作系统：Windows11 Pro 22631.3296 x64

开发环境：Anaconda3 (Python 3.8.19、Python 3.10.11)

深度学习框架：TensorFlow、Torch 2.1.2 (CUDA 11.8)、xFormers 0.0.23

编程工具：JetBrains PyCharm Professional 2023.3.5、Microsoft VSCode

数据分析工具：Pandas、NumPy等用于数据处理和分析的库

可视化工具：Jupyter Notebook，用于数据和结果的可视化

# 系统流程

## 系统主流程

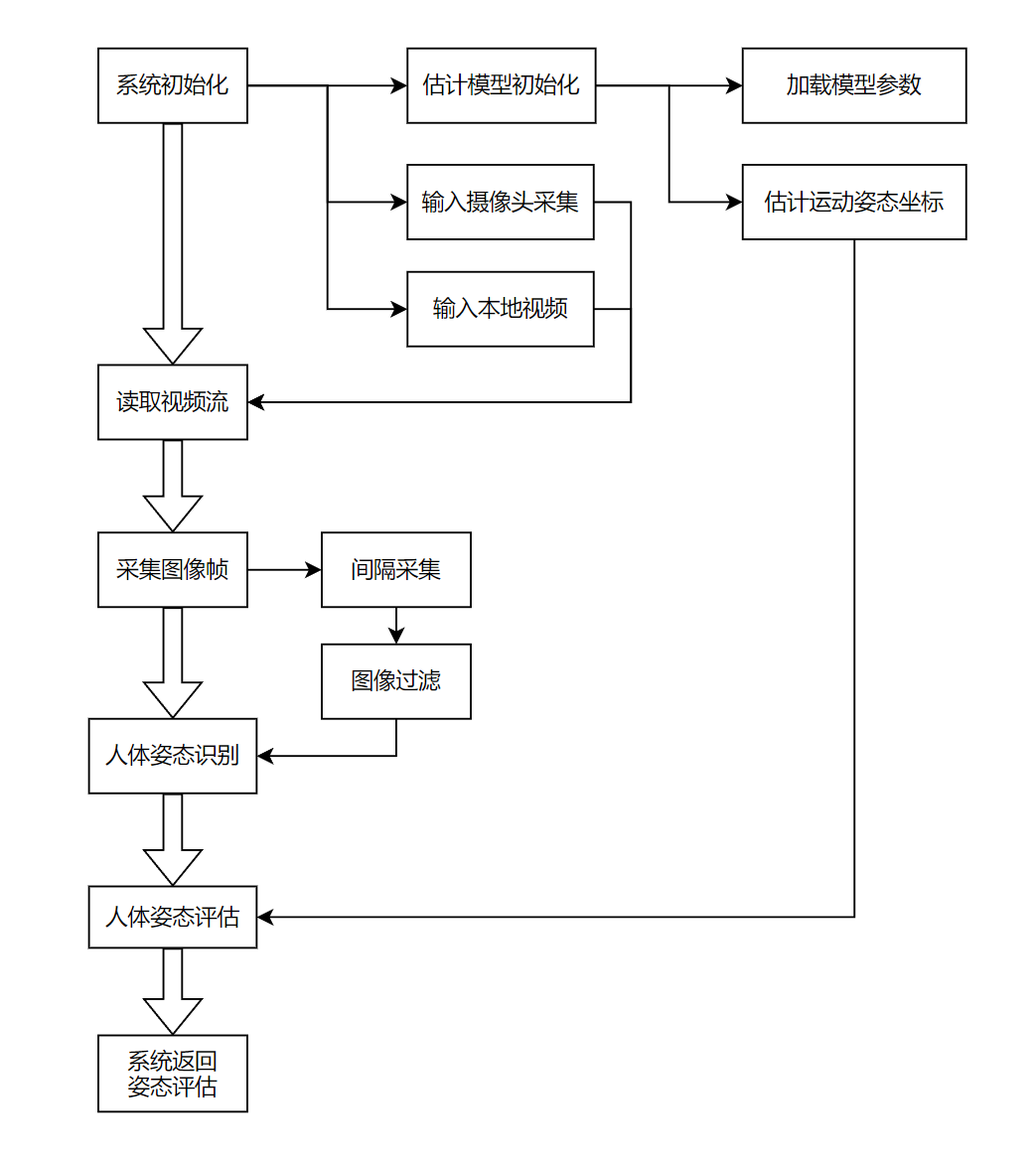


图 ‎3.1 系统主流程图

## 人体姿态识别流程

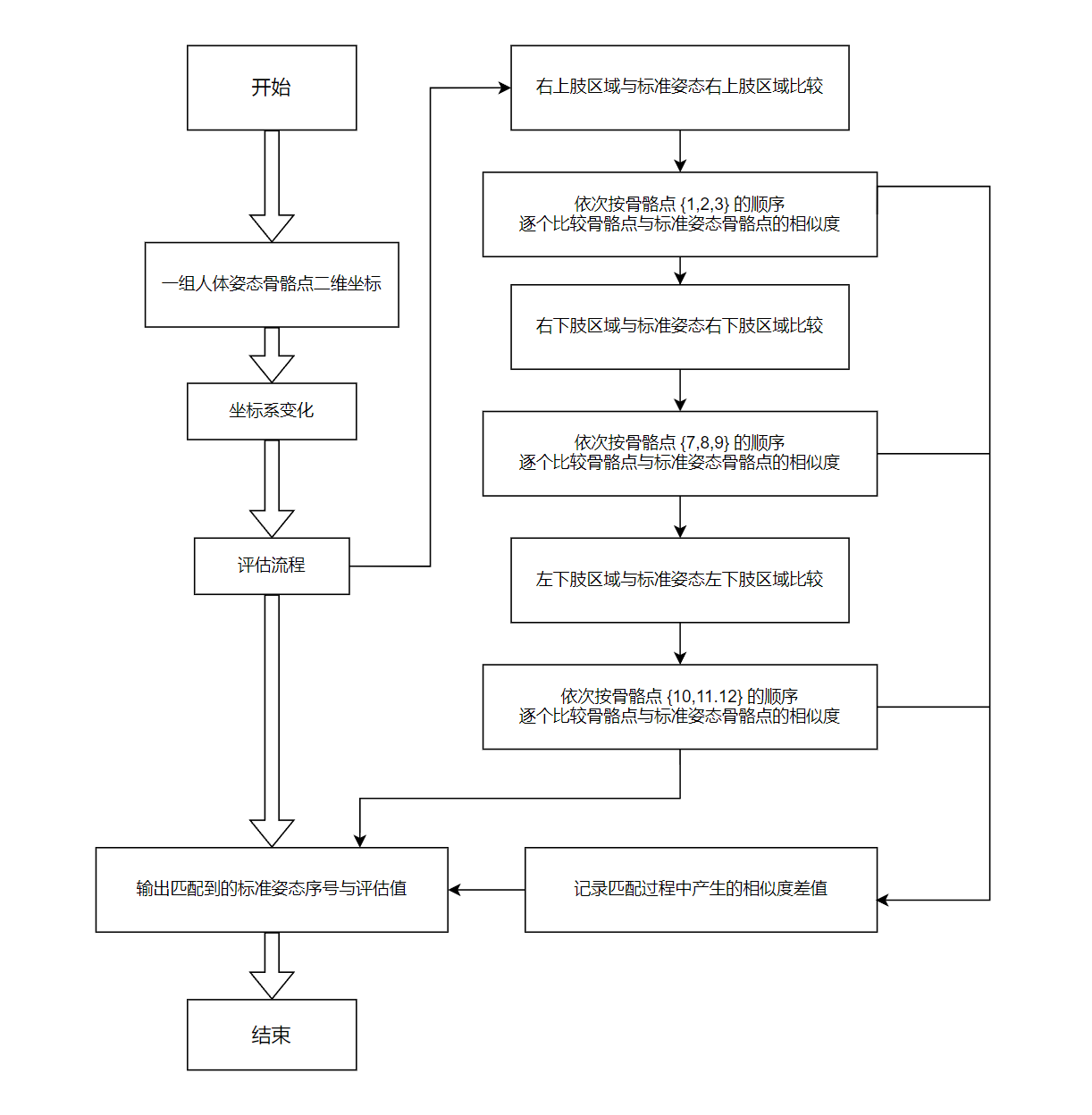


图 ‎3.2 人体姿态识别流程图

## 相似度估计流程

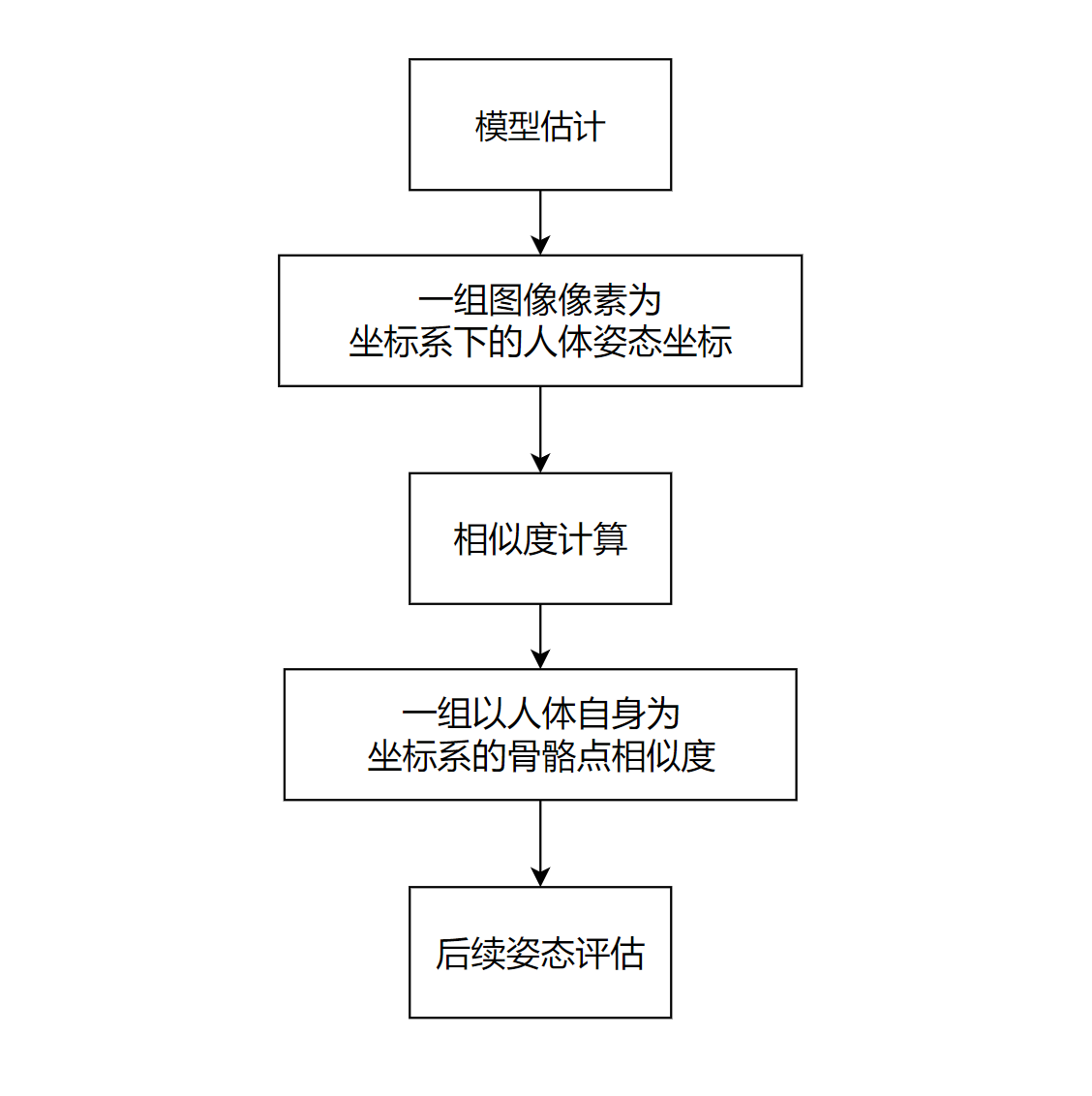


图 ‎3.3 相似度估计流程图

## 人体姿态评估流程

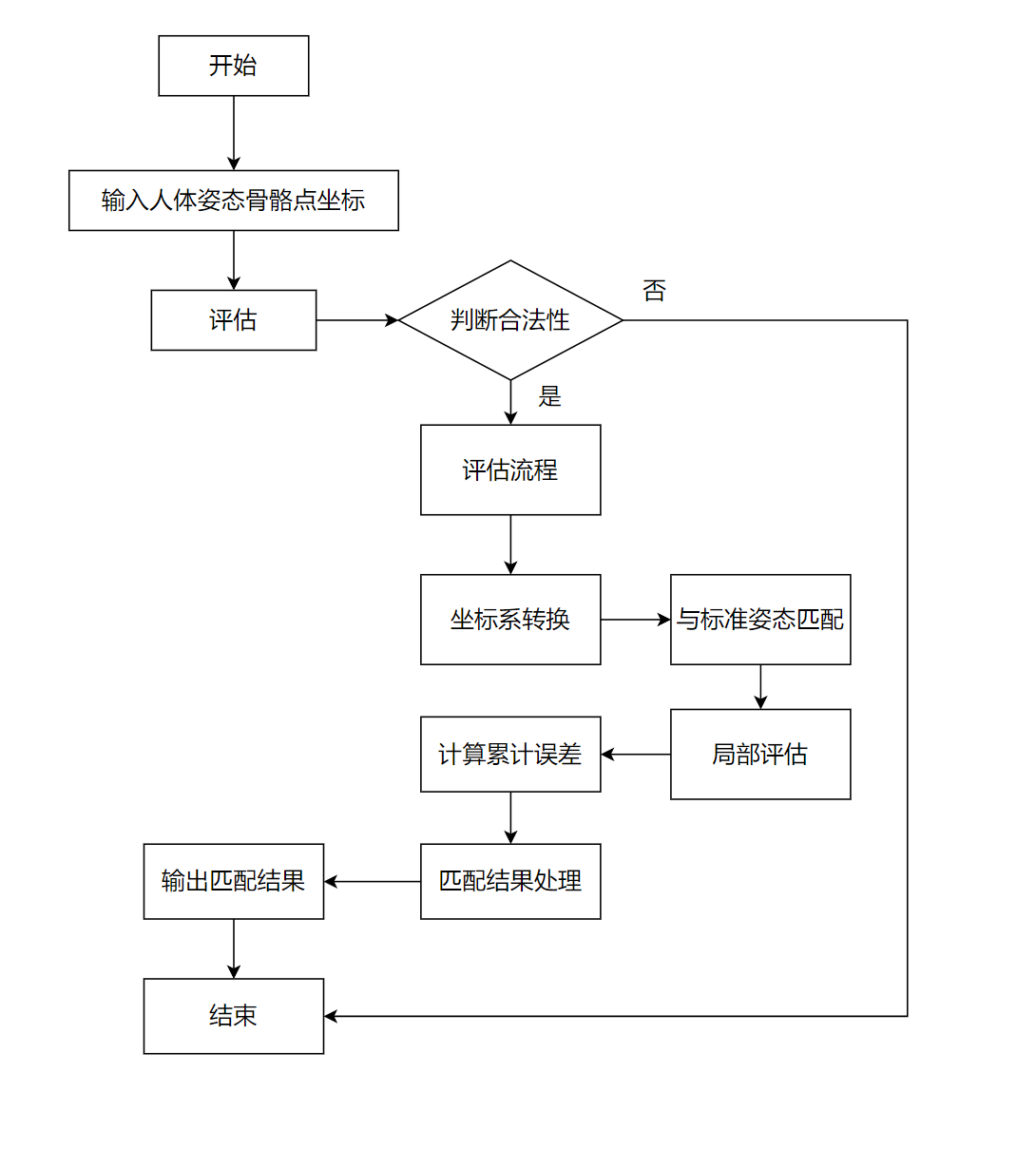


图 ‎3.4 人体姿态评估流程图

# 算法设计

基本检测思路：抓取当前帧的骨骼与标准库图像（base\_data）进行相似度计算，通过比对识别运动中up和down两种姿态以及关节活动幅度，给出该次运动姿态的标准度评价分数。

## 测量标准

深蹲与引体向上两种运动都有up和down两种姿态，初始状态默认是up，然后先后完成一次down和up算识别到一次，当运动者做完一次时给出该次评价分数。

### 深蹲动作

深蹲动作标准姿势：在正面拍摄的视角下，运动者的足关节应与肩关节保持同宽的位置，同时髋关节应低于膝关节，形成一个标准的下蹲姿势。

深蹲动作评估标准：根据运动者臀部下沉的程度来确定，臀部下沉得越低，表示姿势越正确，得分也就越高。以下为深蹲动作的up与down两种姿态的模型参考标准图：



图 ‎4.1 深蹲动作“up姿态”参考图



图 ‎4.2 深蹲动作“down姿态”参考图

### 引体向上动作

引体向上动作标准姿势：在正面拍摄的视角下，在开始引体向上动作时，运动者的两手应略微宽于肩宽，手臂自然垂直下垂。在上拉的过程中，下巴的位置应高于手握处。

引体向上动作评估标准：引体向上动作的评估标准是根据运动者脸部点相对于小臂位置的高度来确定，脸部点越高，表示动作完成得越好，得分也就越高。以下为引体相上动作的up与down两种姿态的模型参考标准图：



图 ‎4.3 引体向上动作“up姿态”参考图



图 ‎4.4 引体向上动作“down姿态”参考图

## 模型初始化

预设TensorFlow SavedModel文件：该文件存储了一个已经训练好的 MoveNet 模型，包括模型本体与权重参数，以本项目中所用到的thunder模型为例：

* saved\_model.pb：存储了模型的计算图结构和操作定义。该文件描述了模型的网络架构，包括各层之间的连接方式和操作的定义。在模型部署和加载过程中，TensorFlow 可根据这个文件来构建模型的计算图。
* variables 目录存储了模型的权重参数。这些参数是模型在训练过程中学到的数值，在推理过程中用于计算模型的输出。权重参数的保存格式包括两个文件：.data 文件与 .index 文件。其中，.data 文件存储了模型参数的具体数值，而 .index 文件则存储了参数的索引信息，包括参数的名称和形状等。

以上三个文件和一个目录组合起来构成了一个完整的TensorFlow SavedModel文件。

在姿态识别中，通过加载saved\_model.pb文件和variables目录中的权重文件来恢复模型并进行推理。

设置一个MoveNet模型的Python类：用于加载和使用预训练的MoveNet模型进行单人姿势估计。

1. 定义初始化方法（\_\_init\_\_）：用于创建Movenet 类的实例时。

在初始化方法中，根据传入的模型名称，选择加载对应的MoveNet模型。如果模型名称中包含movenet\_lightning，则加载Lightning版本的模型；如果包含 movenet\_thunder，则加载 Thunder 版本的模型。

1. 定义预测方法 (predict)：用于对输入的图像进行姿势估计。

* 首先对输入的图像进行预处理，包括颜色空间转换和图像大小调整。
* 接着调用加载的模型进行推理，获取图像中人体的关键点及其置信度。
* 最后对输出进行后处理，将关键点坐标映射回原始图像坐标空间。

1. 创建main程序加载模型。

加载模型实例，并使用OpenCV库读取一张图像。并利用循环对图像进行多次姿势估计，并输出每次推理的时间。

## 姿态检测

使用Movenet模型姿态检测模型，对输入的视频帧进行姿态检测，获取人体关键点的位置信息。接着，通过对于人体骨骼点的二维坐标估计，从而获取骨骼坐标点以评估当前目标的运动动作姿态。

### 视频帧读取

在运动标准度检测方法中，帧读取主要负责从视频文件或摄像头中读取视频帧，并将每一帧作为输入送入Movenet模型进行姿势估计。通过以下步骤来实现：

1. 视频帧读取：使用OpenCV库中的VideoCapture模块进行视频帧的读取。首先，输入视频文件或从摄像头中捕获视频流，然后利用VideoCapture对象的read()方法逐帧读取视频。
2. 帧处理：读取到的每一帧图像会被送入Movenet模型进行姿势估计，得到人体关键点的坐标信息，用于后续相似度估计判断。
3. 循环处理：在视频流未结束的情况下，持续读取并处理每一帧图像。
4. 资源释放：循环结束后，释放视频流相关的资源。

### 骨骼点识别

1. 确定骨骼点

在人体姿态标准度识别中，肢体的协调动作可以通过人体在空间中各个关节的坐标来量化。这些坐标点（骨骼点）的集合构成了所谓的人体姿态坐标集。人体姿态估计模型的主要功能是确定某一帧图像中人物姿态的关键点在图像像素坐标系中的二维位置。

使用传统COCO人体关节点中（用于MoveNet和PoseNet），包括以下17个骨骼点，它们分别是：0: 鼻子、1: 左眼、2: 右眼、3: 左耳、4: 右耳、5: 左肩、6: 右肩、7: 左肘、8: 右肘、9: 左腕、10: 右腕、11: 左胯、12: 右胯、13: 左膝、14: 右膝、15: 左踝、16: 右踝。具体实例如图3.5所示：

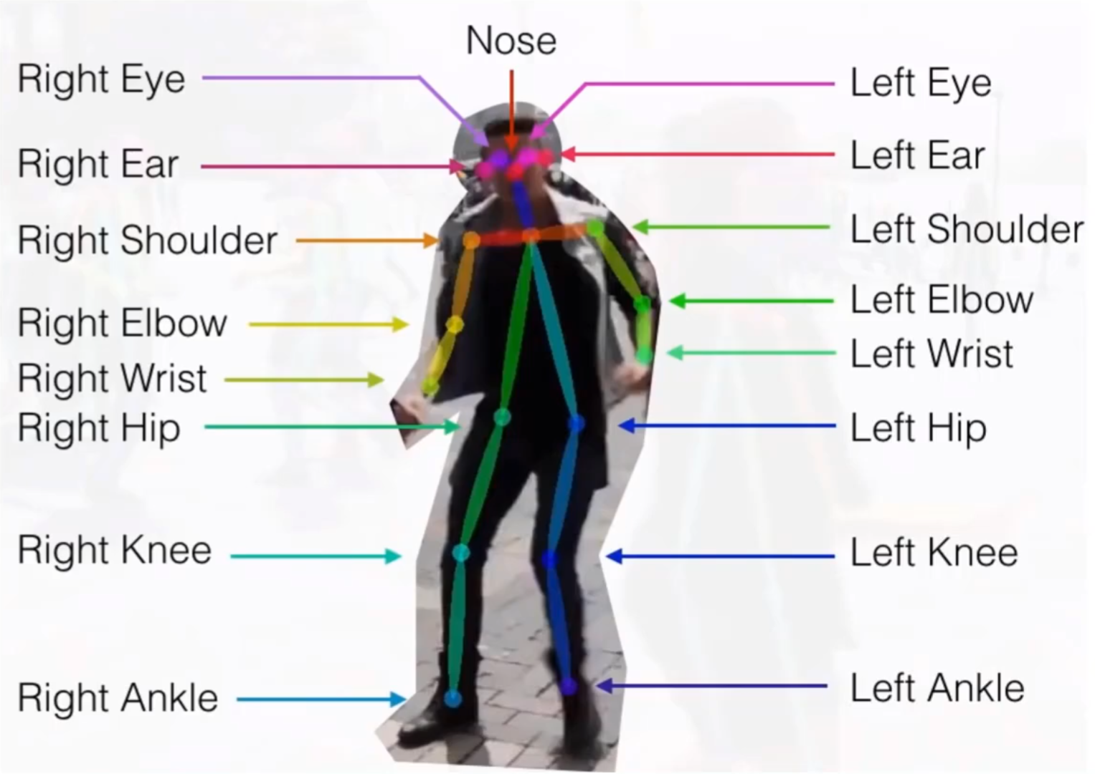


图 ‎4.5 标准17点人体姿态模型

1. 绘制骨骼点

创建姿态检测器模块，用于在图像上绘制姿势估计结果中的骨骼点和连接线的函数。通过将上述的人体关键点和骨骼连接线绘制在图像上，可以直观地展示出人体的姿势以及关键点之间的关联。此外，通过计算关键点之间的夹角，可以量化和评估特定姿势或动作的准确性和标准度。该模块主要包含以下两个主要函数：

首先，创建plot\_skeleton函数用于绘制骨骼点和连接线，以可视化姿势估计结果。参数说明：此函数接收两个参数：image（图像）和points（姿势估计结果中的关键点）。

* 将points转换为NumPy数组，并将其类型设置为整数类型。然后，它遍历每个关键点，并在图像上绘制一个小圆圈来表示关键点的位置。
* 遍历预定义的连接线索引（skeleton\_lines），并使用OpenCV的line函数绘制连接线。
* 返回处理后的图像。

其次，创建calculate\_angle函数用于计算三个给定点（关键点）之间的夹角，判定姿态类型。参数说明：此函数接收三个坐标点作为参数，并计算由这三个点形成的夹角。

* 计算三个点之间的距离，并使用余弦定理计算出夹角。为了避免除零错误，需要在分母中添加了一个很小的值（1e-6）。
* 将弧度转换为角度，并返回角度值。
* 函数中的三个点是按照顺序传递的，因此假设这三个点是按照特定的顺序排列的（例如，点1、点2、点3）。

### 骨骼点归一化

对从姿态检测器得到的人体关键点坐标进行归一化处理，使得大小位置不同的人体在图像上的关键点位置具有相同的尺度和位置参考，便于模型识别的后续处理和比较。这一步骤包括将关键点坐标进行偏移和缩放，以及计算关键点的置信度和总置信度。

详细的逻辑步骤如图4.6所示：

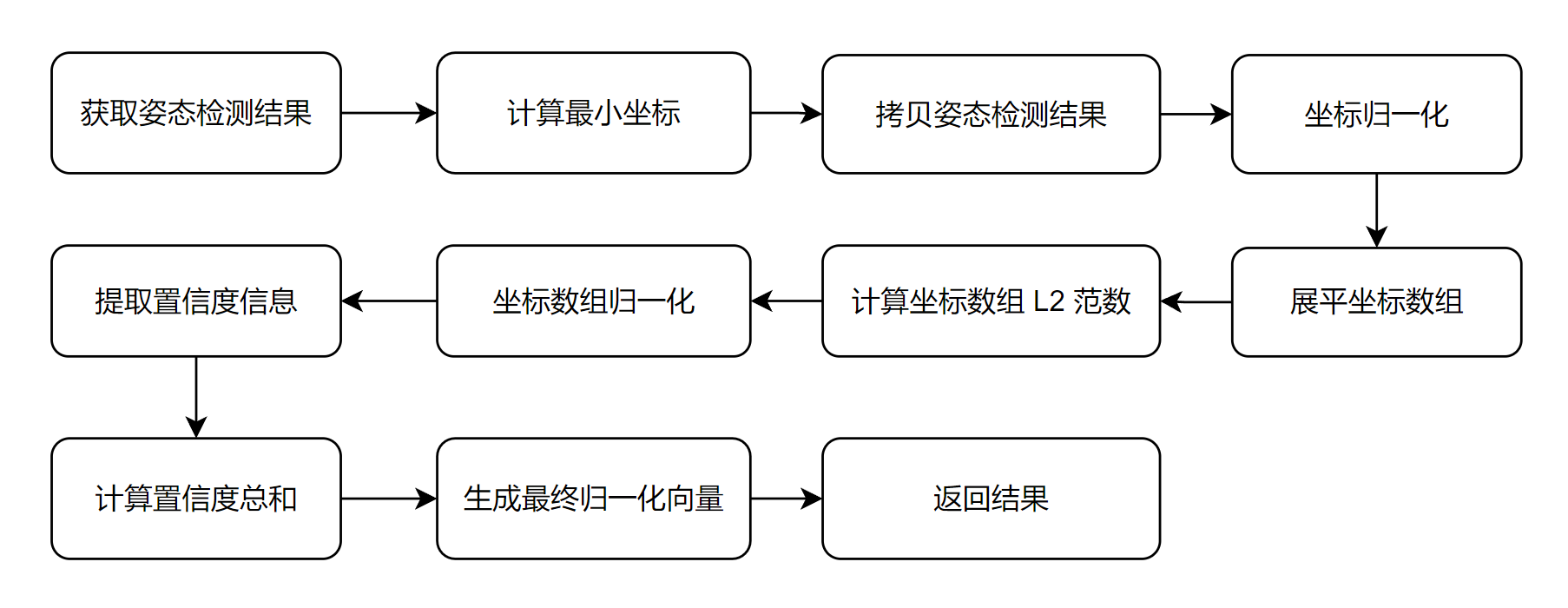


图 ‎4.6 骨骼点归一化步骤

具体操作说明如下：

首先从姿态检测器获取的结果中，获得每个关键点的坐标和置信度。通过计算所有关键点的最小x和y坐标，以确定一个参考原点，相当于将所有关键点整体移动到图像的左上角。为避免修改原始数据，对姿态检测结果进行深度拷贝，并将所有关键点的坐标减去最小坐标，以实现将人体关键点移动到图像的左上角，并保持了关键点之间的相对位置关系。同时，为了便于后续处理，将关键点坐标数组展平为一维数组，计算坐标数组的L2范数以用于归一化处理，确定关键点坐标数组的尺度，使得所有关键点都处于同一尺度上，从而解决近大远小的问题。

此外，还需从姿态检测结果中提取出了关键点的置信度信息，并计算了所有关键点置信度的总和，以备后续处理使用。

最后，将归一化的关键点坐标、置信度信息以及总分信息连接在一起，形成最终的归一化向量，并将原始的姿态检测结果和归一化后的关键点向量返回，以供后续处理使用。

## 匹配算法

在运动标准度检测方法中，匹配算法是用来比较当前帧图像中的人体关键点与预先录制的标准动作关键点之间的相似程度，并找到最匹配的动作作为当前动作的评估结果。在本项目中，匹配算法的实现包括以下关键步骤：

1. 相似度计算：使用相似度计算函数来比较当前帧图像中的关键点向量与标准动作的关键点向量之间的相似度。在项目中，采用了余弦相似度作为相似度计算的方法。这个计算过程在cosin\_distance\_matching函数中实现。余弦相似度的计算方法是将两个向量的点积除以它们的范数的乘积。
2. 多重匹配过滤：在计算得到相似度之后，会将匹配结果与相似度阈值进行比较。如果匹配结果的相似度低于阈值，则该匹配结果可能不够可靠，会被过滤掉。
3. 最佳匹配选择：经过相似度计算和过滤之后，会选择具有最高相似度的标准动作作为当前帧图像的最佳匹配。这个过程在find\_best\_match函数中实现。
4. 匹配结果返回：最终的匹配结果将包括匹配到的标准动作的序号和相似度值。这个结果可以被用来评估当前动作的标准度，并根据需要进行相应的处理。

匹配算法能够快速而准确地找到当前帧图像中与标准动作最相似的动作，并将其作为当前动作的评估结果，从而实现了对运动标准度的检测与评估。

### 相似度计算

在本运动标准度检测方法中，相似度计算是一个重要步骤，其用于比较当前帧图像中检测到的人体关键点与预先录制的标准动作关键点之间的相似程度。在本项目中，该相似度计算采用了余弦相似度作为度量标准。下面是详细描述：

1. 采用余弦相似度： 余弦相似度是用来衡量两个向量方向的差异性的度量指标。它的取值范围在[-1, 1]之间，值越接近1表示两个向量越接近，越接近-1表示两个向量越远离，0表示两个向量之间的夹角为90度。
2. 关键点向量表示：在本项目中人体关键点被表示为一个向量，其中包含了关键点的坐标信息。该向量会忽略置信度和其他附加信息，只关注关键点的位置信息。
3. 计算过程：

* 从当前帧图像中提取的关键点向量中截取出关键点的坐标部分，作为输入的姿势向量。
* 将标准动作的关键点向量与当前帧图像中提取的关键点向量进行余弦相似度计算。
* 通过余弦相似度计算得到的相似度值，转换为欧式距离作为最终的相似度度量值。这里采用欧式距离的平方根作为最终的相似度值，用于衡量当前姿势与标准动作之间的差异程度。

计算得到的相似度值将被用于后续的匹配，用于找到当前帧图像中与标准动作最相似的动作。这样就能够通过比较当前动作与标准动作之间的相似度，来评估当前动作的标准度。

### 处理异常帧

#### 相似度阈值过滤

为了提高匹配的准确性，本项目中引入了相似度阈值参数，即match\_dist\_thresh。当匹配算法得到的相似度值低于该阈值时，系统将排除当前帧图像中动作与标准动作不相似的匹配结果。这个阈值可根据实际情况进行调整，以满足运动标准度检测的需求。在本项目中，相似度阈值被设置为0.3，表示只有当计算得到的相似度高于0.3时，系统才认为匹配结果有效，否则将被过滤。

经过测试，发现相似度的极限值为0.15。在此阈值下，视频识别表现正常，但摄像头录制的识别率显著降低。经过多次测试，最终将匹配相似度阈值设置为了0.3。

#### 视频帧抖动过滤

在视频帧中，可能由于人体运动或摄像机移动等原因导致抖动。为了应对这种情况，项目采用了连续帧匹配策略。具体方法是对连续帧进行匹配，并对匹配结果进行累积和平均，以减少抖动对匹配结果的影响，从而提高匹配的稳定性和准确性。

例如，考虑一个姿态类别信息队列（其中key表示姿态类别，value表示检测到的次数）：[[up,3], [down,15], [up,15], [down,1], [up,10], [down,10], [up,20]]。如果将视频帧抖动过滤阈值设为10（即value值必须大于等于10），则会过滤掉小于等于10的元素，使得队列变为[[down,15], [up,15], [up,10], [down,10], [up,20]]。

### 计数匹配

### 衡量待测姿态与标准姿态

人体姿态估计模型输出的姿态坐标集通常基于图像像素坐标系。将其转换到世界坐标系的步骤包括：

对图像进行去畸变，将理想图像映射到真实图像；

将图像像素坐标转换为物理坐标，进行单位变换；

利用相机内参矩阵进行光学投影，将物理坐标转换到相机坐标系下；

利用相机外参矩阵将相机坐标系转换到世界坐标系。

## 动作评分

## 优化措施

为了提高算法效率，本项目中采用了以下优化措施：

1. 数据结构优化：为了快速检索标准动作关键点数据，项目中使用了基于Ball树的数据结构。Ball树是一种高效的数据结构，可以快速地在多维空间中进行最近邻搜索。通过构建 Ball 树，可以将标准动作关键点数据进行组织和索引，以便快速地在实时帧图像中进行匹配。
2. 并行计算与多线程：在帧读取和关键点检测阶段，项目采用了多线程的方式来提高效率。通过多线程同时处理视频帧和关键点检测，可以减少处理时间，从而提高系统的整体性能。这种并行计算的方式可以充分利用多核处理器和显卡的计算资源，加快算法的执行速度，提高运动标准度检测的实时性。
3. 算法效率考虑：在设计匹配算法时，项目考虑了算法的效率和实时性。采用了余弦相似度作为相似度计算的方法，因其计算简单快速而被选择。此外，通过设置相似度阈值来过滤低相似度的匹配结果，可以减少不必要的计算，提高算法的执行效率。

通过以上优化措施，项目实现了对匹配算法的高效处理，保证了运动标准度检测方法的实时性和准确性。同时，采用了并行计算和多线程技术，充分利用了硬件资源，提高了系统的整体性能。

# 系统实现

本项目为基于视频的运动动作检测系统，可以通过摄像头或者本地视频文件输入，实时识别人体姿态，并根据预定义的基准数据进行匹配，最终给出运动动作的评分。

## 视频帧读取

模块功能：从视频源中读取视频帧并存入队列，供后续模块处理。

接口说明：

* init\_video\_capture(source, frame\_rate): 初始化视频流，设置视频源和帧读取频率。
* start\_video\_stream(): 开始视频流读取。
* read\_frame(): 读取视频帧并存入队列。

调用流程：

1. 初始化：设置视频源（本地视频文件路径或摄像头设备索引）、帧读取频率等参数。
2. 开启视频流：根据设置，打开本地视频文件或连接摄像头设备。
3. 循环读取：通过VideoCapture类的read()方法循环读取视频帧。
4. 队列存储：将读取到的视频帧存入队列，等待后续模块处理。

## 姿态检测

模块功能：对视频帧中的人体进行姿态检测，并提取关键点信息。

接口说明：

* load\_pose\_detection\_model(model\_path): 加载姿态检测模型。
* detect\_poses(frame): 对视频帧进行姿态检测，返回关键点位置信息。

调用流程：

1. 加载模型：使用预训练的MoveNet姿态检测模型进行模型加载。
2. 循环检测：从视频帧队列中获取视频帧，然后对每一帧进行姿态检测。
3. 提取关键点：从姿态检测模型输出中提取人体关键点的位置信息。
4. 姿态向量化：将关键点的位置信息转换成归一化的姿态向量，用于后续匹配。
5. 结果输出：将姿态向量传递给匹配与评分模块进行后续处理。

## 基准数据构建

模块功能：收集运动动作数据，构建基准数据的搜索树。

接口说明：

* collect\_motion\_data(): 收集运动动作数据并进行姿态向量化。
* build\_search\_tree(data): 构建基准数据的搜索树结构。

调用流程：

1. 运动动作数据收集：收集一系列不同人进行各种运动动作的视频数据。
2. 姿态检测与向量化：对收集到的视频数据进行姿态检测，并将关键点位置信息转换成姿态向量。
3. 搜索树构建：利用姿态向量构建搜索树结构，以便后续匹配与评分模块快速检索。

## 匹配与评分

模块功能：对姿态向量进行匹配，计算评分并输出结果。

接口说明：

* load\_search\_tree(tree\_path): 加载基准数据的搜索树。
* match\_and\_score(pose\_vector): 匹配姿态向量并计算评分。

调用流程：

1. 搜索树加载：加载预先构建好的基准数据的搜索树结构。
2. 姿态向量匹配：对于每一帧的姿态向量，利用搜索树进行快速匹配，找到最相似的基准数据。
3. 相似度计算：通过计算姿态向量之间的相似度，得出匹配结果。
4. 评分输出：根据相似度结果给出运动动作的评分，并将结果输出

## 用户界面

模块功能：提供用户界面，实时显示视频流和评分结果。

接口说明：

* initialize\_gui(): 初始化用户界面。
* update\_gui(frame, score): 更新用户界面显示视频帧和评分结果。

调用流程：

1. 显示设置：初始化用户界面，包括视频显示区域、评分区域、运动类型选择等。
2. 实时更新：在界面上实时显示视频帧、运动动作评分及相关信息。
3. 交互操作：根据用户操作，实现退出系统、切换运动类型等功能。

## 系统主控制

模块功能：控制系统整体运行流程，协调各模块之间的数据交互。

接口说明：

* initialize\_system(): 初始化系统参数，加载模型和基准数据。
* run\_system(): 运行系统主循环，依次处理视频帧，进行姿态检测、匹配与评分等操作。
* stop\_system(): 停止系统运行，释放资源。

调用流程：

1. 初始化参数和模型。
2. 创建搜索树。
3. 启动帧读取线程。
4. 主循环中处理视频帧：
5. 进行姿势检测。
6. 进行相似度匹配。
7. 进行计数和评分更新。
8. 显示视频帧和信息。
9. 处理按键事件。
10. 释放资源。

# 系统演示

## 环境配置

命令行模式：

1. 环境安装：将项目中movenet.tar.gz加压放到miniconda3(或者conda3)/envs下。
2. 启用终端：打开自带的miniconda3 prompt命令行。
3. 环境激活：输入conda activate movenet。
4. 切换路径：使用cd命令进入项目代码所在路径。
5. 选择模式：在main函数开头处，选择是否启用摄像头及运动类别。
6. 命令输入：输入python main.py，即可开始运行代码。代码执行结果保存在result文件夹中。

## 运行实例

### 本地视频模式

### 摄像录制模式

# 结论

本次研究旨在开发一种基于视频的体育动作标准度评估算法，以解决传统方法在设备依赖性和实时反馈方面的不足。通过对现有技术的研究和分析，结合姿态匹配技术和机器学习算法，我们成功实现了对深蹲、引体向上等体育动作的识别和评估。在完成本研究的过程中，取得了以下主要成果和结论：

本项目成功开发了基于Movenet人体骨骼点检测模型的体育动作识别算法。该算法能够准确地识别视频中的人体骨骼点，并通过姿态匹配技术对深蹲、引体向上等动作进行识别和评估。实验结果表明，新开发的算法在准确性和鲁棒性方面表现出色，为体育动作标准度评估提供了可靠的技术支持。

同时，本项目搭建并实现了一个可用的实时评分体系，能够在动作进行过程中即时给出评分和反馈。这个系统结合了新开发的识别算法，结合多运动姿态的不同标准，能够实现对不同体系动作执行过程的实时监测和评估，直观简洁，为用户提供了及时的指导和反馈，有效提高了训练效果。

此外，项目中还对算法本身进行了深入的性能评估和优化。通过对多种估计模型以及检测算法的调试和参数优化，有效地提高了检测过程的计算效率和实时性，基本解决了人体在姿态匹配精度和实时性方面遇到的困难，使得算法能够在实际的姿态检测应用中具有较好的性能表现与资源开销。

综上，本研究的成果为体育训练和评估提供了一种全新的、高效的评估方法。新开发的算法以及新一代姿态检测模型的应用，将为体育教练员、运动员乃至普通用户提供更加便捷、有效的训练辅助工具，提高了检测效率的同时降低了设备要求的门槛，有望在体育训练和运动健康管理领域产生重要的应用和推广价值。

本研究的成果为体育动作标准度评估领域的研究和实践提供了重要的技术支持和理论基础，具有一定的创新性和实用性，对相关领域的发展具有积极的推动作用。在未来，我将继续深入研究和优化算法，进一步提升其性能和适用性，为体育训练和运动健康管理提供更加完善的解决方案。

# 致谢

本次毕业设计项目的顺利完成，离不开我的导师姜丽老师的悉心指导。从选题的初步确定，到开题报告、初稿的修改，再到定稿的细节调整以及论文格式的问题，导师都给予了我极大的支持和指导。她不仅是一位优秀的导师，更是一位负责任、耐心细致的教育者，她协助我发现问题，提出了许多宝贵的意见和建议，帮我理清了设计思路以及操作方法，为课题提出了有效的改进方案，让我受益匪浅。我深知自己在这段时间里成长了许多，这都离不开老师的倾囊相授。在此，我衷心感谢导师的教诲，我会永远铭记在心。

同时，我也感谢导师将引领我进入计算机视觉领域。本论文研究的方向是基于视频的体育动作标准度评估算法实现。作为计算机科学技术的一个分支，我有幸深入探索神经网络模型和理论框架，体验当前信息科学与技术中的源源不断的创新，并领悟图像处理算法的深刻内涵。尽管我仍然是计算机视觉学习道路上的一个新人，但我会继续努力前行，坚持脚踏实地，不断学习和成长。我相信，在姜丽老师的指导下，我会在这个领域取得更多的进步和成就。

其次，我要感谢大学四年中所有任课老师以及辅导员对我的严格要求。他们在我学习期间给予了我学术与生活上的指点和帮助，不仅让我学到了专业知识，同时也了解了为人处世的道理，使我能够在今后的求学道路上继续奋斗。他们的言传身教，将成为我人生道路上宝贵的财富，我将倍加珍惜。

在即将毕业的最后时刻，我非常感激在大学生涯中能够遇到一群志同道合的朋友们。他们不仅在学术上给予我不同的视角与专业的指导，同时也是与我共同成长的伙伴。在交往过程中，我们建立了信任，彼此鼓励、支持和帮助，使我的大学生活充实而满足，留下了珍贵难忘的回忆。我相信，我们在未来的人生道路上将会携手共进，共同追求更加美好的未来。

最后，再次感谢所有曾经帮助过我的良师益友们，以及在毕业论文设计中被我引用或参考的论著的作者们。没有你们的支持和帮助，我无法完成这篇毕业论文。愿我们共同的努力和付出，成就更加美好的未来。

# 参考文献

1. 王芫. 基于深度神经网络的人体运动姿态估计与识别.电子科技大学. 2020
2. 蔡敏敏. 基于人体姿态估计的视频动作对比分析研究. 2022
3. Mihai Fieraru, Anna Khoreva, Leonid Pishchulin, Bernt Schiele. Learning Human Pose Estimation Features with Convolutional Networks. 2018.

https://arxiv.org/abs/1804.07909

1. Arjun Jain, Jonathan Tompson, Mykhaylo Andriluka, Graham W. Taylor, Christoph Bregler. Learning human pose estimation features with convolutional networks. 2013. https://arxiv.org/abs/1312.7302
2. Rishabh Bajpai, Deepak Joshi. MoveNet: A Deep Neural Network for Joint Profile Prediction Across Variable Walking Speeds and Slopes. 2021. IEEE.

https://ieeexplore.ieee.org/document/9406043

1. Pramod Murthy, Bertram Taetz, Arpit Lekhra. DiveNet: Dive Action Localization and Physical Pose Parameter Extraction for High Performance Training. 2023. IEEE.

https://ieeexplore.ieee.org/document/10097490

1. Bernhard Hollaus, Bernhard Reiter, Jasper C. Volmer. Catch Recognition in Automated American Football Training Using Machine Learning. Sensors-Basel, January 2023. https://www.mdpi.com/1424-8220/23/2/840
2. Wenbo Li, Zhicheng Wang, Binyi Yin, Qixiang Peng, Yuming Du, Tianzi Xiao, Gang Yu, Hongtao Lu, Yichen Wei, Jian Sun. Rethinking on Multi-Stage Networks for Human Pose Estimation. May 2019.

https://arxiv.org/abs/1901.00148

1. Wei Tang, Pei Yu, Ying Wu. Deeply Learned Compositional Models for Human Pose Estimation. 2018.

https://openaccess.thecvf.com/content\_ECCV\_2018/html/Wei\_Tang\_Deeply\_Learned\_Compositional\_ECCV\_2018\_paper.html

1. Bin Xiao, Haiping Wu, Yichen Wei. Simple Baselines for Human Pose Estimation and Tracking. 2018.

https://openaccess.thecvf.com/content\_ECCV\_2018/html/Bin\_Xiao\_Simple\_Baselines\_for\_ECCV\_2018\_paper.html

1. Feng Zhang, Xiatian Zhu, Mao Ye. Fast Human Pose Estimation. 2019.

https://openaccess.thecvf.com/content\_CVPR\_2019/html/Zhang\_Fast\_Human\_Pose\_Estimation\_CVPR\_2019\_paper.html

1. Yanrui Bin, Xuan Cao, Xinya Chen, Yanhao Ge, Ying Tai, Chengjie Wang, Jilin Li, Feiyue Huang, Changxin Gao & Nong Sang. Adversarial Semantic Data Augmentation for Human Pose Estimation. November 2020.

https://link.springer.com/chapter/10.1007/978-3-030-58529-7\_36

1. Ke Sun, Bin Xiao, Dong Liu, Jingdong Wang. Deep High-Resolution Representation Learning for Human Pose Estimation. 2019.

https://arxiv.org/abs/1902.09212

1. TFBlog: Next-Generation Pose Detection with MoveNet and TensorFlow.js:

https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html

1. TFHub: movenet/singlepose/lightning:

https://www.kaggle.com/models/google/movenet/frameworks/tensorFlow2/variations/singlepose-lightning/versions/4?tfhub-redirect=true

附录

附录说明文字。附录篇幅不宜过大。如果文档或程序过多，可以写明附录见所附软盘或光盘。没有请删去此部分内容。