

基于深度学习的运动员动作识别算法研究

张娜 何建强 王园园

商洛学院 人工智能研究中心 陕西商洛 726000

【摘要】在当下，动作识别技术于体育训练、医疗康复等多个领域的应用呈现出愈发广泛的态势。本文给出了一种借助深度学习实现的运动员动作识别算法。首先采用改进的YOLOv5s-CBAM模型进行目标检测，再利用HRNet模型提取骨骼关键点信息，最后结合时空图卷积网络（ST-GCN）完成动作分类。实验结果表明，所提出的方法在目标检测、关键点检测和动作识别任务中均取得了优异的性能，具有较高的实用价值和应用潜力。

【关键词】深度学习；目标检测；动作识别；骨骼关键点检测

Research on athletes' motion recognition algorithm based on deep learning

Zhang Na He Jianqiang Wang Yuanyuan

Artificial Intelligence Research Center of Shangluo University, Shaanxi Shangluo 726000

【Abstract】At present, the application of action recognition technology in sports training, medical rehabilitation and other fields has shown an increasingly extensive trend. This paper presents a motion recognition algorithm with deep learning. Firstly, the improved YOLOv5s-CBAM model is used for target detection, and then the HRNet model is used to extract bone key bone information. Finally, the action classification is combined with the spatiotemporal graph convolutional network (ST-GCN). The experimental results show that the proposed method has achieved excellent performance in target detection, key point detection and action recognition tasks, with high practical value and application potential.

【Key words】deep learning; object detection; action recognition; bone key point detection

1 引言

最近几年，随着计算机视觉与深度学习技术突飞猛进，动作识别日益成为科研焦点，在体育训练、医疗康复以及安防监控等诸多领域，都呈现出极为广阔的应用前景。特别是在体育领域，运动员动作识别技术能够辅助教练和运动员分析技术动作，优化训练策略，从而提升竞技表现。动作识别方法不断演进，主要通过结合循环神经网络（RNN）与CNN、YOLO模型、HRNet网络、ST-GCN模型结合空间图卷积与时间卷积网络等方法实现高效的动作识别^[1-5]。

基于以上先进的技术，结合运动员动作的特点，提出了一种新的动作识别算法。首先，运用YOLOv5s-CBAM模型开展人体定位工作，极大地增强了目标检测的准确率与鲁棒性。其次，利用HRNet模型实施骨骼关键点检测，提升了关键点检测的精度。最后，结合时空图卷积网络（ST-GCN）完成动作分类。

2 基于深度学习的运动员动作识别

图1所呈现的是借助深度学习实现运动员动作识别的原理架构。此流程涵盖四个主要环节：①针对运动员动作数据展开采集工作，并完成相应预处理；②运用Yolov5s -

CBAM目标检测网络来确定人体所在位置；③依靠HRNet网络对人体骨骼的关键点予以检测；④借助ST-GCN网络达成对动作的识别以及评价任务。

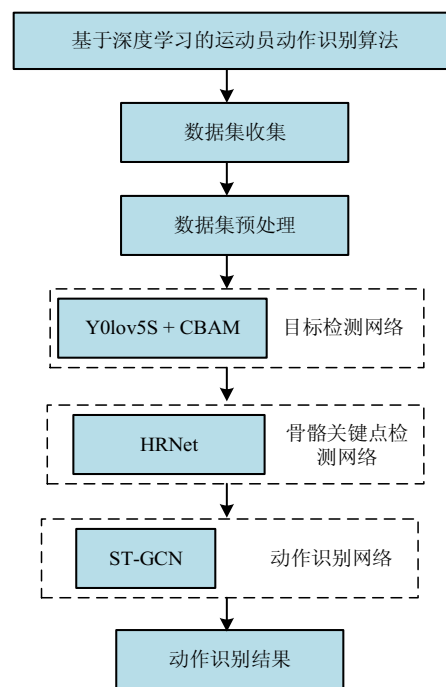


图1 原理框图

2.1 运动员动作数据集构建

构建了一套散打动作数据集，涵盖前摆拳、后摆拳、前

直拳、后直拳、前正蹬腿、后正蹬腿、前侧踹踢、后侧踹踢、前鞭腿、后鞭腿等动作。每组动作包含 800 张图像，总计 8000 张。为进一步扩充数据，从 Pascal VOC 数据集的 Person 类别中选取 2000 张图像，最终数据集规模达到 10000 张。

2.2 运动员动作图像预处理

由于拍摄过程中光照、角度等因素会影响图像质量，且采集到的图像尺寸不一致，所以在模型训练前，需要针对图像开展一系列预处理操作。

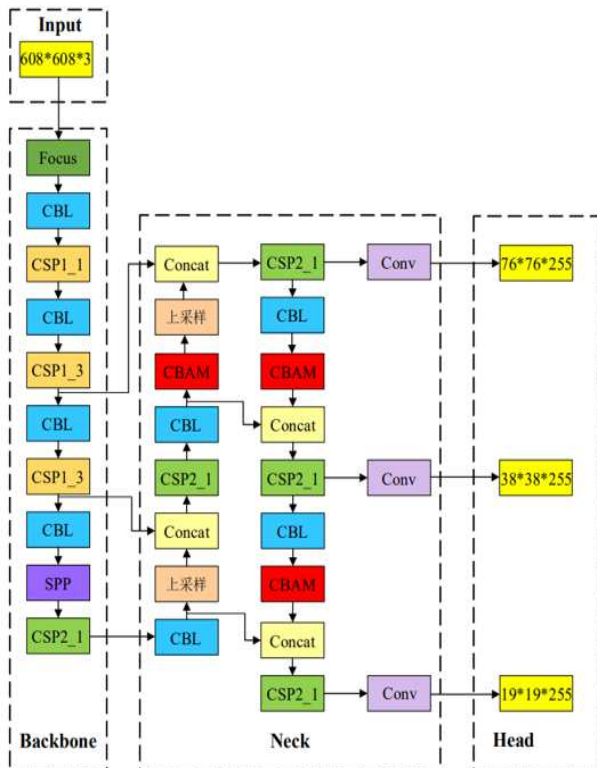


图 2 Yolov5s-CBAM 网络结构图

3 YOLOv5S-CBAM 目标检测网络模型

表 1 目标检测模型性能对比

网络模型	P (精确率)	R (召回率)	mAP (平均精度均值)	FPS (帧速)
SSD	0.889	0.866	0.931	26
Yolov5s	0.896	0.893	0.982	42
Yolov5s-CBAM	0.934	0.932	0.995	35

实验数据表明，Yolov5s - CBAM 模型在精确度、mAP 值等方面均有着优异表现。与 Yolov5s 模型相比，Yolov5s - CBAM 模型的精确度提高了 3.8%，和 SSD 模型相比，其精确度更是提升了 4.5%。尽管 Yolov5s-CBAM 的帧率 (FPS) 比 Yolov5s 低 7 帧，但比 SSD 高 9 帧，这意味着它在维持高水平检测精准度的状况下，依旧具备一定程度的

3.1 YOLOv5S-CBAM 网络模型

YOLOv5s 模型由四个核心组件构成：①Backbone，是一个基于卷积的网络，从图像中提取特征；②Head，用于预测目标的位置和类别；③Neck，位于 Backbone 和 Head 之间，用于特征融合和调整；④预测层，负责输出目标检测框和相应的标签^[6]。

在 YOLOv5s 模型中，通过在每个检测尺度前集成 CBAM 模块，能够更有效地集中注意力于目标的关键区域。CBAM 模块由两个子模块组成：通道注意力模块 (CAM) 和空间注意力模块 (SAM)。CAM 通过关注输入特征图中的重要通道来增强特征的显著性，而 SAM 则专注于在空间维度上识别关键区域。Yolov5s-CBAM 网络结构如图 2 所示。

3.2 模型评价指标

1.精确率 (P)、召回率 (R) 如式 (1) 所示：

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN} \quad (1)$$

其中，TP 为正确检测数量；FP 为误检数量；FN 为未被检测出数量。

2.平均精度 (AP)、平均精度均值 (mAP) 如式 (2) 所示：

$$\begin{cases} AP = \int_0^1 P(R) dR \\ mAP = \frac{\sum P_A}{N_c} \end{cases} \quad (2)$$

其中，Nc 为类别个数；AP 为不同类别的平均精度。得到每一类的值后，将其平均就得到 mAP。

3.帧率 (FPS) 常被用于评估目标检测模型的实时性能，它体现的是神经网络在一秒内处理图片的数量。

3.3 实验结果分析

对于所采集到的数据将训练集与测试集按照 8:2 进行数据划分。对 SSD、YOLOv5s 以及 Yolov5s-CBAM 网络模型进行训练测试，实验结果如表 1 所示。

实时处理能力。

4 骨骼关键点检测及动作识别算法

4.1 HRNet 骨骼关键点检测网络模型

在利用目标检测网络完成人体目标区域的提取后，紧接

着的步骤是把提取出的人体目标框图坐标信息,输入至骨骼关键点检测网络,以此来检测人体骨骼关键点的位置。本次选取了 17 个骨骼关键点,具体位置分布如图 3 所示。

4.2 ST-GCN 动作识别算法

随着 AI 技术的进步,ST-GCN 模型应运而生,它结合了图卷积和动作识别的优势。该模型以构建骨架时空图为基础,借助空间图卷积网络 (SGCN) 与时间卷积网络 (TCN),对人体动作的空间特性和时间特性加以捕捉^[9],从而达成运动员动作识别的目的,具体流程可见图 4。

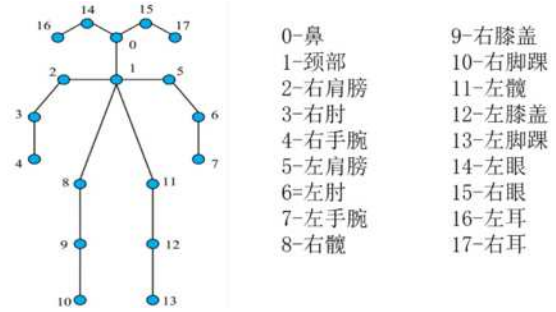


图 3 人体姿态模型图

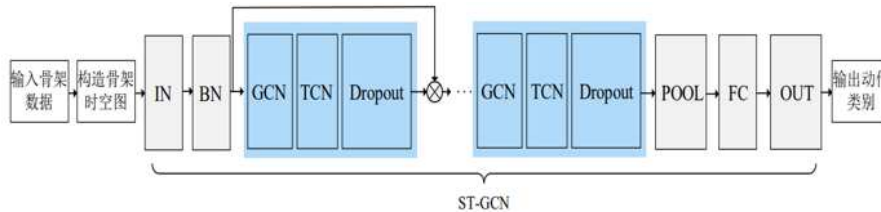


图 4 ST-GCN 动作识别原理图

4.3 实验结果分析

将骨骼数据转化为骨架时空图,作为动作识别研究的材

料,运用 ST-GCN 网络模型进行训练与测试,其识别准确率展如表 2。

表 2 动作识别准确率

动作名称	前直拳	后直拳	前摆拳	后摆拳	前正蹬腿	后正蹬腿	前鞭腿	后鞭腿	前侧踹退	后侧踹退
ST-GCN	94%	94%	94%	95%	90%	91%	95%	95%	90%	90%

实验数据表明,ST-GCN 网络模型在处理运动员动作的识别与分类任务时,展现出了优越的性能。

法,通过优化目标检测、骨骼关键点检测及动作识别模型,改善了散打动作识别中存在的遮挡和姿态多样性等挑战。实验证明,改进后的 YOLOv5s-CBAM、HRNet 和 ST-GCN 模型在各项指标上兼具高精度与实时性。对动作识别具有实用价值。

5 结论

本研究设计了一种基于深度学习的运动员动作识别算

参考文献

- [1]Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. "Large-scale video classification with convolutional neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp.1725-1732).2014.
 - [2]Donahue, J., McAuley, J., & Ponce, J. "Long-term recurrent convolutional networks for visual recognition and description." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp.2625-2634).2015.
 - [3]Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition.2016.
 - [4]Sun, Ke, et al. "Deep high-resolution representation learning for human pose estimation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.2019.
 - [5]Chen, L., et al. "Spatio-temporal graph convolutional networks: A deep learning framework for time-series forecasting." arXiv preprint arXiv: 1801.07865, 2018.
 - [6]武历展.基于深度学习的实时动作识别方法研究[D].长安大学, 2023.
- 项目基金: 陕西省体育科研常规课题 (20240360)