

基于OpenPose的脑卒中肢体康复训练评估辅助系统开发

刘肇川, 毛文婷, 陈 炫

重庆科技学院, 重庆

收稿日期: 2022年4月26日; 录用日期: 2022年8月8日; 发布日期: 2022年8月16日

摘 要

针对脑卒中患者的居家肢体康复训练缺乏指导、看护者的康复知识欠缺的问题, 开发一套针对脑卒中患者肢体康复训练的智能化辅助系统。通过摄像头获取患者康复训练过程中的图像, 利用OpenPose与MoveNet姿态检测库与提取患者训练过程的实时动作骨架坐标序列, 构建用户肢体运动评价模型, 对康复训练者的训练过程进行评价, 得出训练者的康复训练质量, 并通过文字提示的方式提醒用户不达标的动作以及改正方式, 指导和引导患者逐步达到康复训练的标准。

关键词

计算机视觉, 神经网络, OpenPose算法, MoveNet算法

Development of an OpenPose-Based Assessment Aid System for Stroke Physical Rehabilitation Training

Zhaochuan Liu, Wenting Mao, Xuan Chen

Chongqing University of Science and Technology, Chongqing

Received: Apr. 26th, 2022; accepted: Aug. 8th, 2022; published: Aug. 16th, 2022

Abstract

To address the problems of lack of guidance and caregiver's knowledge in home physical rehabilitation training for stroke patients, an intelligent assisted system for physical rehabilitation training of stroke patients was developed. The system uses the OpenPose and MoveNet posture detection

libraries to extract the real-time skeleton coordinate sequences of the patient's training process, construct a user's limb movement evaluation model, evaluate the training process of the trainer, and derive the quality of the trainer's rehabilitation training, and remind the user of the substandard movements and the correction methods through text prompts. It also reminds the user of the substandard movements and the ways to correct them by means of textual prompts, and guides the patient to gradually achieve the rehabilitation training standard.

Keywords

Computer Vision, Neural Networks, OpenPose Algorithm, MoveNet Algorithm

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑卒中又被称为中风，是脑组织和脑细胞坏死的一种疾病，此病主要是由于大脑中的血管突然破裂或因血管阻塞从而导致血液不能流入大脑进而引发的脑组织损伤，引起临床症状超过 24 小时将有可能致死，此病具有明显的季节性特征，在寒冷季节发病率更高。循证医学证实，在脑卒中存活的患者中，进行积极的康复治疗，可使 90% 的存活患者能重新恢复步行和生活自理能力，可使 30% 的存活患者能恢复一些较轻的工作。相反，不进行康复治疗，上述两方面恢复的百分率相应只有 6% 和 5%。大量临床康复实践表明[1] [2] [3] [4]，早期有效的康复训练有助于改善脑卒中患者受损的功能(如感觉、运动、语言、认知和心理等)，减轻残疾的程度，预防脑卒中后可能发生的并发症(如压疮、坠积性或吸入性肺炎、泌尿系感染、深静脉血栓形成等)，提高患者的日常生活活动能力和适应社会生活的能力，提高其生活质量。现阶段在康复训练领域对脑卒中患者进行早期有效的康复训练有两种解决方案，第一部分是依靠康复器械[5] [6]，如上肢康复机器人、下肢康复机器人、肢体康复治疗仪等，康复器械主要依赖进口，价格较高，并且现有的康复器械只能覆盖部分的康复训练，具有一定的局部性，其中肢体康复治疗仪对脑卒中急性期生命体片有危险倾向者不适用。另一个解决方案是康复机构包括公立康复医院和私立康复医院，我国公立康复专科医院数量严重不足，同时民营康复医院水平参差不齐，且同样存在治疗费用较高的情况，为了解决以上问题，让脑卒中患者不需要去支付昂贵的治疗费用购买昂贵的器械，居家就可以实现早期简单的康复训练，系统提出的解决方法是通过摄像头获取患者康复训练过程图像，利用计算机视觉技术，结合 OpenPose 和 MoveNet 姿态检测库提取患者训练过程的实时动作骨架坐标序列，构建用户肢体运动评价模型，对康复训练者的训练过程进行评价，得出训练者的康复训练质量，并通过语音的方式提醒用户不达标的动作以及改正方式，指导和引导患者逐步达到康复训练的标准。

本文主要实验并比较了 OpenPose 和 MoveNet 两种姿态识别的算法，采用了在移动端效果更好的 MoveNet 神经网络算法进行设计。

2. 系统设计

系统实验环境为机械革命 X8Ti Plus 笔记本电脑，CPU 为 Inter i7-8750H，Windows 10 64 位操作系统，内存为 16 G，显卡为 NVIDIA GEFORCE 2060，软件平台为 Visual Studio Code。系统安装简单，所需要的设备数量少，成本低。系统流程图如图 1 所示。

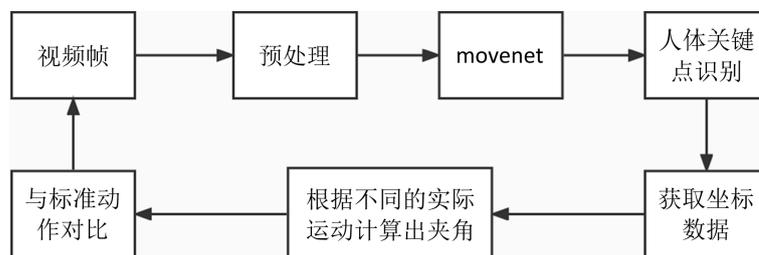


Figure 1. System flow chart
图 1. 系统流程图

由于成像设备和外部环境等因素会在图像运输和转换的过程中造成影响，图片的大小在不同时段色彩具有差异，这些因素会干扰结果的正确性，为了最大程度减少环境因素对图像的影响，则需要先对图像的数据进行宽高固定化，将图片等比例缩放到一样的大小，再对图片进行灰度化[7]处理。灰度化处理就是将彩色图片转化为灰色图像，由于彩色图像的所含信息量太多，会对后续工作造成影响，所以选择使用图像灰度化处理从而达到提高运算速度的目的。经过测试，预处理操作对结果有一定的积极作用。如图 2 所示。



Figure 2. Preprocessing effect
图 2. 预处理效果

3. 姿态估计

3.1. 基于 OpenPose 的人体关键点检测

OpenPose 算法分析

OpenPose [8]人体姿态识别项目是依赖于卷积神经网络和监督学习并以 Caffe 为框架开发的开源库。它支持在单人或多人情况下对人的面部表情和肢体动作进行估计。它的主要优点是适用于单人或多人，以及具有良好的鲁棒性，可以说是世界上第一个基于深度学习的实时 2D 姿态估计。OpenPose 采用的是自下向上的方式，即先将所有关节整合出来，再通过仿射场将其划分到个人。OpenPose 的网络结构如图 3 所示。

在该模型中输入了一幅图像，并通过 VGG19 [9]卷积网络提取特征，得到了一组特征图之后将其分成两个岔路 Branch 1 & 2，分别使用 CNN 网络提取 Part Confidence Maps(置信度)和 Part Affinity Fields(关联度) [10]，使用图论中的 Bipartite Matching(偶匹配)求出 Part Association [11]，并且连接同一个人的关节点，由于 PAF 自身的矢量性，使得生成的偶匹配正确，最终合并为一个完整的人体骨架；最后基于 PAFs 求 Multi-Person Parsing→把 Multi-Person Parsing 问题转变为 graphs 问题→Hungarian Algorithm(匈牙利算法)。

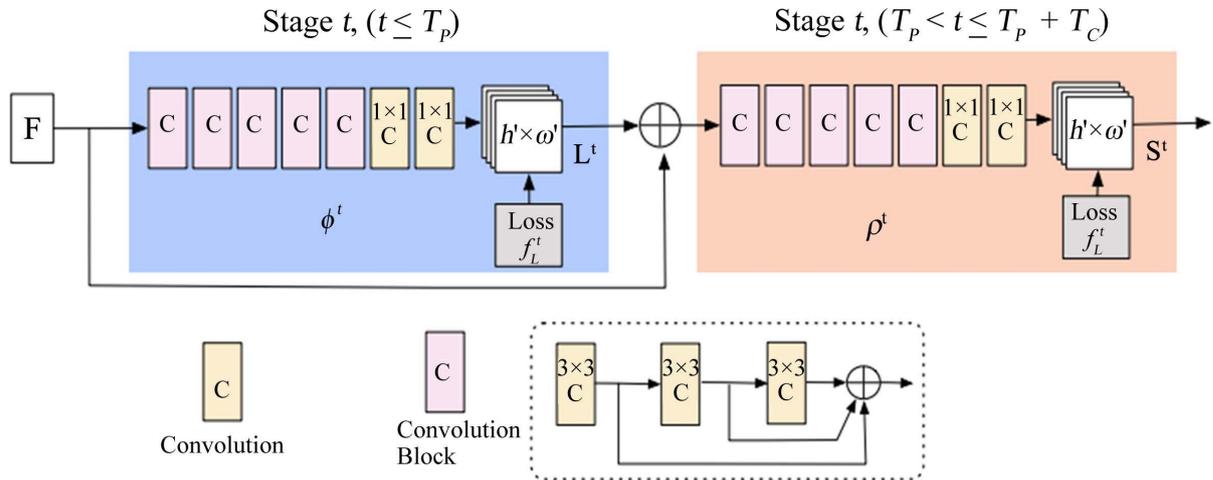


Figure 3. Network structure of OpenPose
图 3. OpenPose 的网络结构

3.2. 基于 MoveNet 的人体姿态检测

MoveNet 是由 google 公司 2021 年推出的人体姿态检测，MoveNet 是自下而上的估计模型如图 4，能够检测出身体的 17 个关键点，利用热图来精确定位人体的关键点。该架构是由两个部分组成：一个特征提取器和一组预测头。预测方案大致遵循 CenterNet，但是相较该架构而言，它大幅度提升了速度和准确性。所有模型均使用 TensorFlow 对象检测 API 进行训练。

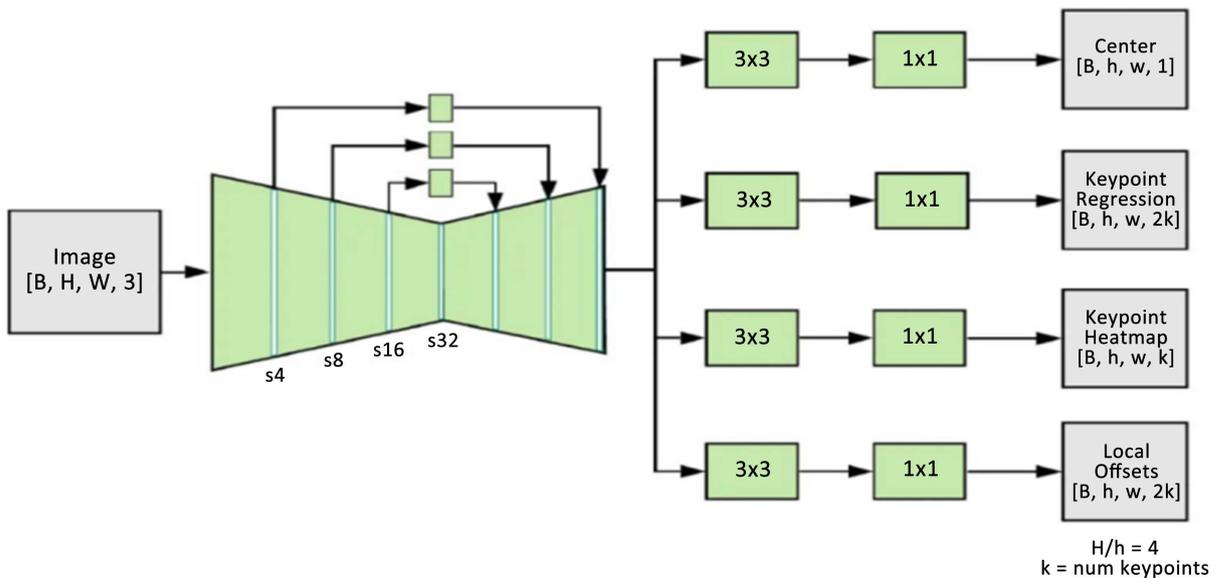


Figure 4. Structure of MoveNet
图 4. MoveNet 的结构

MoveNet 中的特征提取器是 MobileNetV2 [12]，带有附加的特征金字塔网络(FPN)，能够实现高分辨率(输出步长为 4)且语义丰富的特征图输出。特征提取器上附带了四个预测头，负责密集预测：人体中心热图是用来预测人体实例的几何中心；关键点回归场是用来预测人体的完整关键点集，用于将关键点分组到实例中；人体关键点热图是用来预测所有关键点的位置，使其不依赖人体实例。每个关键点的二维

偏移场是用来预测从每个输出特征图像素到每个关键点的精确子像素位置的局部偏移量。其步骤如下：

第 1 步：人体中心热图用于识别框架中所有个人的中心，定义为属于个人的所有关键点的算术平均值。选择得分最高的位置(通过与框架中心的反距离加权)。

第 2 步：通过对象中心对应的像素分割关键点回归输出来生成该人体的初始关键点集。由于这是中心向外的预测(必须在不同的尺度上操作)，所以回归关键点的质量不会特别准确。

第 3 步：关键点热图中的每个像素都乘以一个权重，这个权重与对应的回归关键点的距离成反比。这可以确保我们不接受来自背景人物的关键点，因为他们通常不会靠近回归的关键点，因此得分较低。

第 4 步：通过获得的每个关键点通道中热图值的最大坐标来选择关键点预测的最终集合，然后将局部 2D 偏移量预测添加到这些坐标以给出精确的估计。关于这四个步骤的详细信息，请参见下图 5。

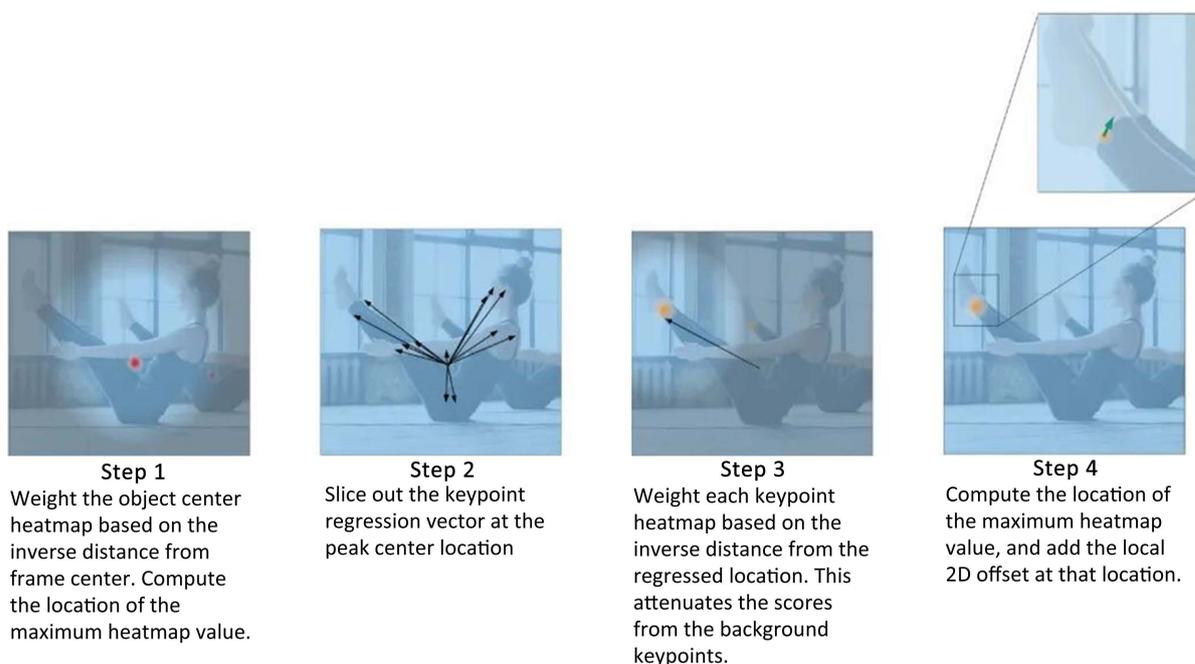


Figure 5. MoveNet flowchart

图 5. MoveNet 流程图

为了将 MoveNet 打造成高质量的检测器，官方团队选择在架构设计、后处理逻辑和数据上投入了大量精力。同时，团队也很重视推理速度。选择了将 MobileNetV2 的瓶颈层用于 FPN 横向连接。其次，为了加快输出特征图的执行速度，选择大幅减少每个预测头中的卷积滤波器数量。除 MobileNetV2 第一层外，整个网络都使用了深度可分离卷积[13]。

1) 优化

对 MoveNet 进行了反复的分析，比较慢的操作被找到并且删除。例如，将 `tf.math.top_k` 替换为 `tf.math.argmax`，因为其执行速度明显更快，而且适用于单人设置。

为了确保能够使用 TensorFlow.js 实现快速执行，所有模型的输出都将打包到一个输出张量中，因此从 GPU 到 CPU 的下载只有一次。或许，最显著的提速是针对模型使用 192×192 的图像输入(对于 Thunder，则为 256×256)。为抵消低分辨率，基于前一帧的检测结果应用了智能裁剪[14]。这使得模型能够把注意力和资源投入到主体上，而不是背景上。

2) 时间滤波

在高 FPS 摄像头视频流上操作能够从容地对关键点估计应用平滑处理。Lightning 和 Thunder 都对传入的关键点预测流应用了强大的非线性滤波器[15]。团队对该滤波器进行了调整,以同时抑制模型中的高频噪声(即抖动)和离群值,同时还可以在快速运动期间维持高带宽吞吐量。这使得在各种情况下都能以最小的延迟实现流畅的关键点可视化。

3) MoveNet 浏览器性能

为了量化 MoveNet 的推理速度,官方团队在多个设备上对该模型进行了基准测试如图 6。模型延迟时间(以 FPS 表示)在使用 WebGL 的 GPU 和 WebAssembly (WASM)上进行测量,后者是使用低端 GPU 或无 GPU 的设备的典型后端。

	MacBook Pro 15" 2019。 Intel core i9; AMD Radeon Pro Vega 20显卡。 (FPS)	iPhone 12 (FPS)	Pixel 5 (FPS)	桌面设备 Intel i9-10900K; Nvidia GTX 1070 GPU。 (FPS)
WebGL	104 77	51 43	34 12	87 82
WASM 与 SIMD +多线程	42 21	不适用	不适用	71 30

Figure 6. Inference speed of MoveNet across different devices and TF.js backends

图 6. MoveNet 在不同设备和 TF.js 后端之间的推理速度

3.3. 算法比较

由于 OpenPose 主要是针对于多人运动进行姿态估计[16],并且需要计算的数据量很大,为了能够达到实时的目的,使用了高并行的策略。基于 cuda 加速,需要很高的显存,相较而言,MoveNet 是一个快速并且轻量的模型,更加适用于本系统中。

4. 角度计算

关节和骨骼之间的角度可以清晰地反映一个人的整体位置和移动方向,人与人之间存在的差异(身高、肤色、体型等)对计算机判断的结果没有影响。系统设定的骨骼点中有头(包括眼睛、鼻子和耳朵,但不会被分配权重)、脊柱中心、左右手等。如图 6 所示,将全部的关节分到主体躯干和四肢两个大组中,分组为:1) 主体躯干: 0, 1, 2, 5, 8, 11。2) 四肢: 2, 3, 4, 8, 9, 10(左); 5, 6, 7, 11, 12, 13(右)。

从一个关节到另一个关节之间的距离可用公式(1)表示

$$d_{x,y,z} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \quad (1)$$

式中 $d_{x,y,z}$ 为关节与关节之间的距离, (x_i, x_j) , (y_i, y_j) , (z_i, z_j) 为关键点的坐标,因为图像中每个人的关节点被 OpenPose 生成 2D 位置作为输出,所以将式(1)修改为式(2), (x_i, x_j) , (y_i, y_j) , 为二维像素坐标系上的关键点坐标。

$$d_{x,y} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2)$$

通过这个公式,我们可以先计算出一组里面两个关节之间的距离,然后再计算出关节与关节之间的

角度。角度计算如图7。

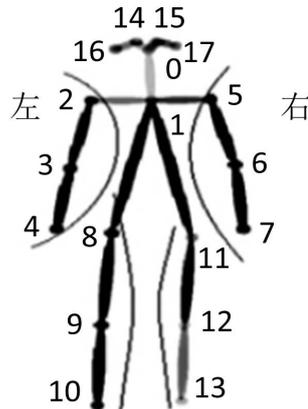


Figure 7. Joint grouping
图7. 关节分组

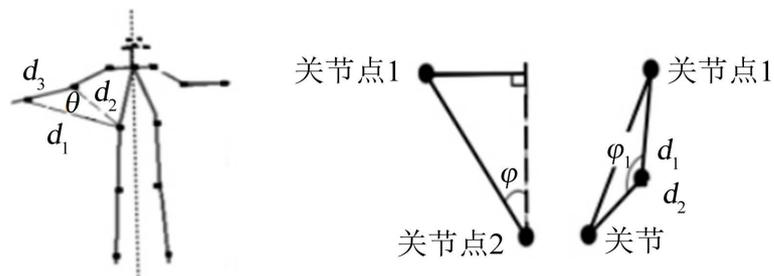


Figure 8. Angle calculation
图8. 角度计算

图8中 d_1 , d_2 是一个组里面相邻的两个关节点之间的距离。对于主体躯干, 还需要计算它们和垂直方向之间的偏移角 φ 。

$$\theta = \arccos \frac{d_1^2 + d_2^2 - d_3^2}{2d_1 \cdot d_2} \quad (3)$$

$$\varphi = \frac{\arccos |y_1 - y_2|}{d} \quad (4)$$

式中 θ 为一组中任意一个关节点与编号相邻的两个关节点连线之间的夹角, d_1 , d_2 , d_3 分别为三个关节点之间的距离, φ 为主干与垂线之间的偏移角(°)。式(5)可用于计算实际动作与标准动作之间角度偏差的百分比 P

$$P = 1 - \frac{|\theta_N - \theta_{\text{standrd}}|}{\theta_{\text{standrd}}} \times 100\% \quad (5)$$

式中 θ_N 为患者所做动作的实际角度, θ_{standrd} 为标准角度。式(6)能够计算出 P_{total} 人体动作的总体拟合百分比。

$$P_{\text{total}} = \frac{1}{n} \sum_{i=1}^n P_i \quad (6)$$

5. 效果

采用了广东工伤康复医院的居家康复训练指导动作进行测试如图9, 与实际做的动作进行对比, 系

统可以完整的检测到人体姿态，并给出提示，可以完成整个动作的无人式指导。



Figure 9. Rendering
图9. 效果图

6. 结论

基于 OpenPose 的脑卒中康复训练评估辅助系统可以较准确地识别出人体动作并进行比较提示，其稳定性好。本系统在现有技术和方法的基础上进行了分析研究，针对于现在康复领域所存在的设备器材及康复医院价格昂贵等特点，提出了第三种解决方案：在线康复训练，完成了康复训练动作的居家动作指导，系统设置了语音提示、采用得分制的方式对训练进行指导，使得整个训练过程更富有趣味。实验结果表明：该系统判断的结果与人眼识别的结果一致，并且安装简单，需要的设备较少，互动性较强，设备成本较低，涉及的数据较少。

基金项目

国家级大学生科技创新项目 202111551007。

参考文献

- [1] 王艳杰, 张丽娟. 近五年脑卒中康复训练领域研究热点的文献计量学分析[J]. 循证护理, 2022(4): 500-505.
- [2] 丁杏, 赵小玉. 成都市社区脑卒中患者自我护理能力与护理需求的调查研究[J]. 护士进修杂志, 2022, 37(5): 424-429.
- [3] 吴晓云. 脑卒中运动障碍康复护理的研究进展[J]. 中国城乡企业卫生, 2022, 37(2): 41-43.
- [4] 张小华, 曾媛媛, 王玉龙. 上肢康复机器人训练对脑卒中患者上肢功能的疗效观察[J]. 中国医学创新, 2017, 14(22): 29-31.
- [5] 王越, 张春青, 王悦, 甘宁, 戎善奎, 江潇, 计雄飞, 余新华. 国内外医用康复器械分类与管理现状[J]. 中国医疗设备, 2022, 37(3): 163-166.
- [6] 李夏楠. 中国农村脑卒中患者康复产品服务设计研究[D]: [硕士学位论文]. 郑州: 郑州轻工业大学, 2020: 77.
- [7] 李春阳. 彩色图像灰度化及其效果的客观评价方法研究[J]. 流行色, 2021(12): 93-95.
- [8] Cao, Z., et al. (2021) OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43, 172-186. <https://doi.org/10.1109/TPAMI.2019.2929257>
- [9] Karen, S. and Andrew, Z. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. *iComputer*

-
- Science*, **15**, 75-80.
- [10] Cao, Z., Simon, T., Wei, S.E., *et al.* (2017) Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 7291-7299. <https://doi.org/10.1109/CVPR.2017.143>
 - [11] Bulat, A. and Tzimiropoulos, G. (2016) Human Pose Estimation via Convolutional Part Heatmap Regression. *14th European Conference*, Amsterdam, 11-14 October 2016, 717-732. https://doi.org/10.1007/978-3-319-46478-7_44
 - [12] Sandler, M., Howard, A., Zhu, M.L., Zhmoginov, A. and Chen, L.-C. (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
<https://www.computer.org/csdl/proceedings-article/cvpr/2018/642000e510/17D45Wuc32W>
 - [13] Howard, A.G., *et al.* (2017) MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. <https://arxiv.org/abs/1704.04861>
 - [14] 王皓文, 桑农. 基于序列对抗域适应的智能裁剪算法[J]. 模式识别与人工智能, 2021, 34(8): 677-688.
 - [15] 刘金钢. 基于逆协方差交叉融合的多传感器非线性滤波器[J]. 黑龙江大学学报, 2021, 12(4): 59-65.
 - [16] 肖文福, 孙焱. 一种基于深度可分离卷积的实时多人姿态估计[J]. 计算机应用研究, 2020, 37(S2): 348-350.