

李晓雪, 李争平, 曹树林, 等. 基于视觉的航空医学应急救援心肺复苏操作姿态评价算法[J]. 安全与环境学报, 2026, 26(6): 2409–2416.

LI X X, LI Z P, CAO S L, et al. Visual-based CPR compression posture recognition system for aviation medical emergency rescue[J]. Journal of Safety and Environment, 2026, 26(6): 2409–2416.

基于视觉的航空医学应急救援心肺复苏操作姿态评价算法*

李晓雪^{1,2}, 李争平³, 曹树林³, 郝昱文^{1,2}, 王立军^{3,4}, 贾可豪³

(1 中国人民解放军总医院医学创新研究部, 北京 100853;

2 智能应急医学救援北京市重点实验室, 北京 100853;

3 北方工业大学人工智能与计算机学院, 北京 100144; 4 西安电子科技大学杭州研究院, 杭州 311231)

摘要:我国的航空医学应急救援人才短缺, 航空医学应急救援训练装备落后。针对航空医学应急救援模拟训练关键问题, 即心肺复苏 (Cardiopulmonary Resuscitation, CPR) 按压动作识别精度低, 提出一种基于视觉的训练操作姿态识别模型。通过人体姿态关键点提取, 并进行时域均值填充和平滑处理, 完成数据预处理, 之后提出训练姿态识别模型, 进行心肺复苏按压姿态识别。该模型在双向门控循环单元 (Bi-directional Gated Recurrent Unit, BiGRU) 网络基础上加入卷积神经网络 (Convolutional Neural Network, CNN), 并添加注意力机制, 使模型更关注动作本身, 减小动态环境等因素的干扰。试验结果表明, 此模型的姿态识别准确率达到 93%, 高于其他模型, 可以满足需求。

关键词:公共安全; 心肺复苏; 姿态识别; 注意力机制

中图分类号: X928; TP391.4

文献标志码: A

文章编号: 1009-6094(2026)06-2409-08

DOI: 10.13637/j.issn.1009-6094.2025.0392

0 引言

航空医学应急救援是指应对重大事故和灾害, 利用航空技术和技术装备, 以直升机、飞机等为主要载体, 实施紧急病人快速运送、空中应急救援、医疗物资运送和搜救飞行活动^[1]。航空医学应急救援被誉为反应最快、安全性最高、综合效率最突出的救援手段, 具备响应迅速、机动能力强、覆盖范围广等优势, 在其发展与应用过程中, 呈现政策支持持续加强、社会需求日益迫切、应急响应不断提速、能力建设稳步推进的鲜明特征^[2-3]。

我国的航空医学应急救援相较于发达国家经验还不足, 并且由于时间、空间、人力和物力等因素限制无法大规模开展直升机救援培训活动^[4-5]。在航空医学应急救援中, 缺乏对心肺复苏 (Cardiopulmonary Resuscitation, CPR) 按压姿势正确性的自动化评估体系^[6-7]。现有评估多依赖人工或仅评估按压吹气操作, 缺乏规范性, 且需大量人力资源, 导致应急救援能力提升缓慢^[8-9]。而对于姿态评价, 也有许多研究者进行了研究。2023 年, 张冰

冰等^[10]为提高散打运动辅助打分的公平公正性, 第一次提出了基于骨骼关键点的散打动作识别与评价的概念, 然后通过分析这些关键点的信息完成人体动作的识别与评价。2024 年, Wang 等^[11]首次提出通过三维卷积神经网络 (Convolutional Neural Network, CNN) 进行篮球技术动作识别, 旨在提高篮球比赛中各种动作的准确和自动化识别。在航空应急救援领域, 一些研究主要集中在救援场点布局或能力评估^[12-13], 而本文研究主要集中在航空医学应急救援模拟训练关键技术。

针对航空医学应急救援中的心肺复苏按压姿势规范性评价, 本文提出基于视觉的操作动作评价方法。通过工业级 RGB 相机采集视频数据, 利用人体姿态识别算法获取姿态信息, 并经预处理输入识别模型, 实现心肺复苏按压姿势的正确性评估。心搏骤停救治模拟训练计算机数据评估设计流程如图 1 所示。

1 人体动作估计算法及数据集

1.1 动作估计算法

基于视觉的人体动作估计是指利用图像中的人

* 收稿日期: 2025-04-09

作者简介: 李晓雪, 副主任医师, 博士, 从事于救援医学理论和技术研究, lixiaoxue@301hospital.com.cn; 李争平 (通信作者), 副教授, 博士, 从事信号处理、通信工程、机器学习和深度学习研究, lizp@ncut.edu.cn。

基金项目: 国家重点研发计划项目 (2020YFC0811004)

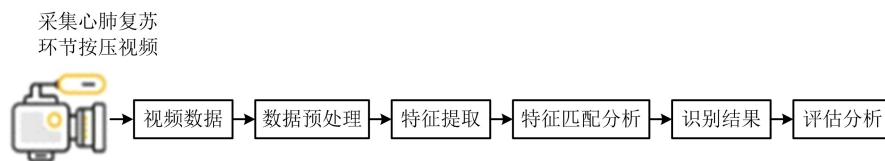


图1 心搏骤停救治模拟训练计算机数据评估流程图

Fig. 1 Flowchart of computer data evaluation for cardiac arrest rescue simulation training

体关键点信息来识别并定位人体的姿态,进而构建人体的近似骨架^[14]。MultiPose 是由中东科技大学提出的自底向上多人姿态估计方法^[15]。其核心是多任务模型,同步执行人体检测、关键点检测和人体分割。该方法通过同步检测与跟踪机制,显著提升多人场景(尤其是姿态重叠时)的人体姿态估计准确性、鲁棒性和个体动作分割精度,并增强对复杂运动细节的捕捉能力。同时,MultiPose 优化网络结构与算法,同步改进骨骼点连接精度和检测效率,综合性能超越传统方法。

MultiPose 结构如图2所示。其骨干网络(Backbone)采用 ResNet 作为主干,残差块 C2、C3、C4、C5 分别以步长 4、8、16、32 对图像进行特征提取,为后续关键点检测和行人检测网络提供多尺度特征。ResNet 后接两个特征金字塔网络 (Feature Pyramid Network, FPN), 分贝用于关键点检测和行人检测。FPN 通过自底向上提取多分辨率特征图,再

自顶向下上采样统一尺寸,横向连接形成多尺度特征金字塔,增强网络对细节与语义信息的捕捉,提升图像理解力。

MultiPose 结构的第二部分包含两个相互独立的分支网络:关键点检测网络与人体检测网络。关键点检测网络基于 FPN 对 CNN 特征进行融合处理,输出人体关键点热图与人体分割掩码。其具体流程为:对特征图进行下采样后,经两次 3×3 卷积得到中间特征图 D ;将多尺度特征整合为通道数为 512 的特征图,再通过 3×3 卷积与 ReLU 激活函数进行平滑,最终由 1×1 卷积输出 $K+1$ 通道热图,其中 K 代表人体关键点类别数,额外 1 个通道对应人体分割掩码。

人体检测网络采用 RetinaNet 结构,输出维度为 $N \times 5$ 的检测结果; N 为图像中检测到的人体目标数量,5 维向量分别对应人体边界框的 4 个坐标值和目标置信度。第三部分为伪随机噪声码 (Pseudo

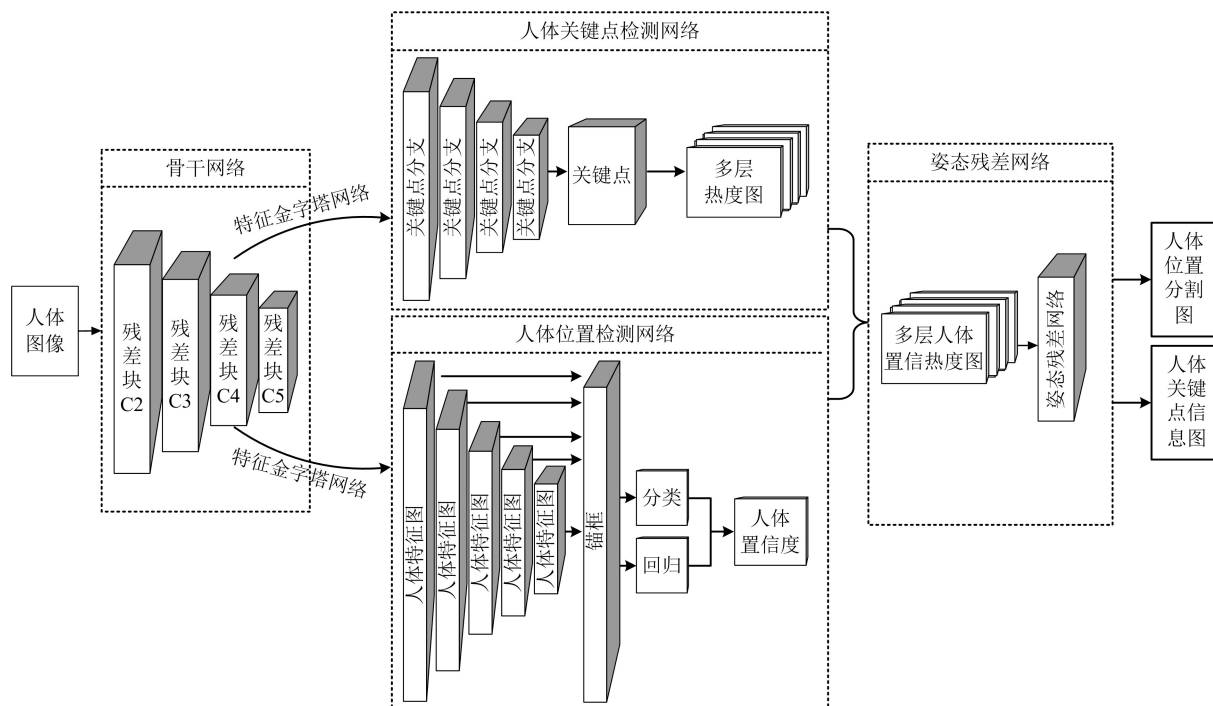


图2 MultiPose 结构图

Fig. 2 Schematic diagram of MultiPose

Random Noise code, PRN) 网络, 该网络用于对检测得到的关键点和人体目标进行关联聚类, 并基于感兴趣区域 (Region of Interest, RoI) 对关键点检测结果进行聚合, 随后通过残差型多层感知机进行特征处理, 以学习完整的人体关节姿态配置。具体过程为关键点检测网络的输出被裁剪至人体位置检测网络检测的人体框内, X 表示 PRN 输入, $X = (x_1, x_2, \dots, x_k)$, $x_k \in \mathbf{R}^{W \times H}$, k 是关键点的数量, W 为热度图的宽度, H 为热度图的高度。通过式 (1), PRN 输出 $Y = (y_1, y_2, \dots, y_k)$, 其中 $y_k \in \mathbf{R}^{W \times H}$, 特征图尺寸与输入 x_k 保持一致。

$$y_k = \varphi_k(X) + x_k \quad (1)$$

函数 $\varphi_k()$ 对 X 中的姿态应用残差修正, 网络输出层采用 Softmax 归一化函数对关键点热图进行概率归一化处理。

1.2 数据集

本文通过深度学习方法评估航空医学心搏骤停救援中的关键动作, 鉴于尚无相关心肺复苏动作评价数据集, 需构建此类数据集以支持研究。

依据美国心脏协会 (American Heart Association, AHA)《心肺复苏指南》等规范文件, 正确的心肺复苏按压姿势须确保手、肘、肩三点成一线且双肩水平, 即肘关节伸直角度为 $180^\circ \pm 15^\circ$, 肩部水平夹角 $< 10^\circ$, 常见错误包括曲肘和斜肩, 是直接影响深度与位置的核心错误。

本文使用单目相机从训练者正前方录制心肺复苏按压动作。邀请 10 名志愿者对模拟人执行心肺复苏, 录制标准与非标准动作 (斜肩、曲肘等), 每人 3 组, 每组 30 次。专业人员评估视频质量, 确保视频适用于训练数据集。在本数据集中, 采集的视频数据帧率均为 30 帧/s, 分辨率为 1 280 像素 $\times$ 720 像素。

构建的数据集包含 900 个心肺复苏动作样本 (标准、曲肘、斜肩各 300 个)、600 个无关负样本 (走路、摔倒等), 总长 1 500 s (45 000 帧, 30 帧/s), 均为分辨率 1 280 像素 $\times$ 720 像素的 MP4 格式文件。数据集附带 2 个 TXT 文件: 一个存储 13 个关键点的归一化坐标, 另一个存储对应分类标签。

2 航空医学应急救援动作识别与评价

2.1 卷积神经网络与循环神经网络

深度学习在高精度连续动作识别中表现优异。CNN 擅长提取空间特征, 但时序处理能力不足。双流法利用光流信息提升效果, 但依赖高质量光流数

据^[16]。鉴于心肺复苏按压姿势识别具有时序依赖性且易受光照限制, 本文主体采用基于时序信息的分类算法, 并融合 CNN 提取的空间关系, 以综合时空特征并提升模型性能。

CNN 通过卷积操作和局部感受野, 能高效提取图像和音频数据特征, 具有权值共享的特性^[17-18]。循环神经网络 (Recurrent Neural Network, RNN) 能处理序列特征并记忆历史信息, 但存在梯度消失或爆炸问题。长短期记忆 (Long Short-Term Memory, LSTM) 网络通过门机制和记忆细胞解决 RNN 的梯度问题, 擅长长期依赖^[19]。门控循环单元 (Gate Recurrent Unit, GRU) 作为 LSTM 的简化结构, 通过更新门和重置门也能有效缓解梯度问题, 且在小型数据集上表现更优^[20]。GRU 结构示意图如图 3 所示。

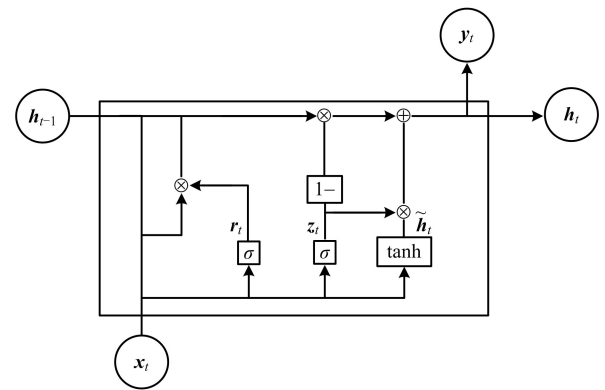


图 3 GRU 结构图

Fig. 3 Schematic diagram of GRU

GRU 的计算公式为

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \quad (2)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \quad (3)$$

$$\tilde{h}_t = \tanh(W[r_t h_{t-1}, x_t] + b) \quad (4)$$

$$h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t \quad (5)$$

式中 z_t 是更新门的状态; r_t 是重置门的状态; x_t 是当前输入; W_z 、 W_r 和 W 分别代表各门及状态的权重参数; b_z 、 b_r 和 b 分别表示各门及状态的偏置项; σ 是 sigmoid 激活函数; 隐藏状态 h_t 是更新门和候选隐藏状态的加权和, h_{t-1} 是上一级更新门和候选隐藏状态的加权和; $\tilde{h}_t$ 是候选隐藏状态, 用于提供新的信息。

鉴于心肺复苏按压姿势的连续性, 采用双向门控循环单元 (Bi-directional Gated Recurrent Unit, BiGRU) 提取心肺复苏按压关键动作序列关键

信息,有助于提取关键动作数据的前后特征依赖,提高模型理解力和泛化能力。

BiGRU 结构如图 4 所示,BiGRU 由两个传递方向相反的单向 GRU 组成,BiGRU 分别从正向和反向两个方向处理输入序列,以捕捉序列中的双向信息,将不同时刻的 GRU 输出相加,最终成为 BiGRU 的输出层。BiGRU 计算公式为式(6)~(13),可知双向 GRU 在单向的信息传递与遗忘中与单向 GRU 类似。

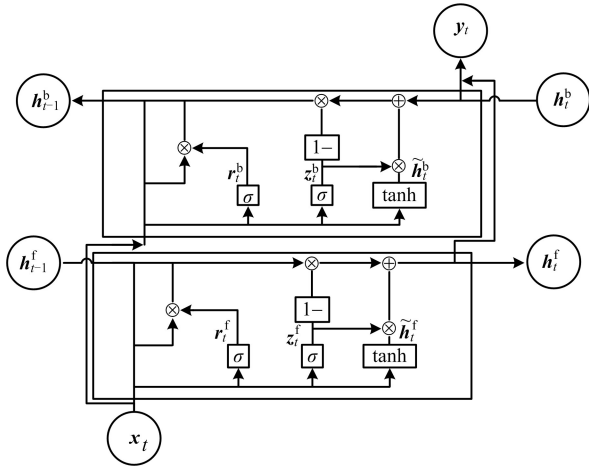


图 4 BiGRU 结构图

Fig. 4 Schematic diagram of BiGRU

式(6)~(9)为正向 GRU 公式。在正向 GRU 中,隐藏状态根据当前输入和上一个时间步的隐藏状态进行计算。正向 GRU 能够捕获当前时刻及之前的上下文信息。

$$z_t^f = \sigma(W_z^f[h_{t-1}^f, x_t] + b_z^f) \quad (6)$$

$$r_t^f = \sigma(W_r^f[h_{t-1}^f, x_t] + b_r^f) \quad (7)$$

$$\tilde{h}_t^f = \tanh(W_h^f[r_t^f h_{t-1}^f, x_t] + b_h^f) \quad (8)$$

$$h_t^f = (1 - z_t^f) h_{t-1}^f + z_t^f \tilde{h}_t^f \quad (9)$$

式中 z_t^f 是正向更新门的状态; r_t^f 是正向重置门的状态; W_z^f, W_r^f 与 W_h^f 分别代表正向各门及状态的权重参数; b_z^f, b_r^f 与 b_h^f 分别是正向各门及状态的偏置项; 隐藏状态 h_t^f 是正向更新门和候选隐藏状态的加权和, h_{t-1}^f 是正向上一级更新门和候选隐藏状态的加权和; $\tilde{h}_t^f$ 是正向候选隐藏状态, 用于提供新的信息。

式(10)~(13)为反向 GRU 公式。在反向 GRU 中,隐藏状态会根据当前输入和下一个时间步的隐藏状态进行计算。反向 GRU 则能够捕获当前时刻及之后的上下文信息。

$$z_t^b = \sigma(W_z^b[h_{t-1}^b, x_t] + b_z^b) \quad (10)$$

$$r_t^b = \sigma(W_r^b[h_{t-1}^b, x_t] + b_r^b) \quad (11)$$

$$\tilde{h}_t^b = \tanh(W_h^b[r_t^b h_{t-1}^b, x_t] + b_h^b) \quad (12)$$

$$h_t^b = (1 - z_t^b) h_{t-1}^b + z_t^b \tilde{h}_t^b \quad (13)$$

式中 z_t^b 是反向更新门的状态; r_t^b 是反向重置门的状态; W_z^b, W_r^b 与 W_h^b 代表反向各门及状态的权重参数; b_z^b, b_r^b 与 b_h^b 是反向各门及状态的偏置项; 隐藏状态 h_t^b 是反向更新门和候选隐藏状态的加权和, h_{t-1}^b 是反向上一级更新门和候选隐藏状态的加权和; $\tilde{h}_t^b$ 是反向候选隐藏状态, 用于提供新的信息。

双向 GRU 隐藏状态为

$$h_t = [h_t^f; h_t^b] \quad (14)$$

式中 h_t 为当前时刻的最终隐藏状态, 由正向和反向隐藏状态拼接而成。当正向和反向 GRU 计算完隐藏状态后, 将两者拼接形成一个综合的隐藏表示, 然后输出。这个综合的隐藏表示包含输入序列中每个时间步的信息及上下文关系, 因此可以更全面地表达输入序列的特征。

2.2 注意力机制

注意力机制由 Bengio 等提出^[21], 它通过为信息分配权重来突出关键部分, 帮助模型更好地理解 and 处理复杂数据。本文使用注意力机制来融合多层双向 GRU 输出, 动态地给予不同时间步的隐藏状态不同的权重, 能更好地捕捉输入序列中的重要信息。假设有 l 层双向 GRU, 每一层的正向和反向隐藏状态分别为

$$[h_t^{f(1)}, h_t^{b(1)}], [h_t^{f(2)}, h_t^{b(2)}], \dots, [h_t^{f(l)}, h_t^{b(l)}] \quad (15)$$

式中 t 表示时间步; $h_t^{f(l)}$ 和 $h_t^{b(l)}$ 分别表示第 l 层 GRU 的正向和反向隐藏状态。注意力机制通过对每个时间步的隐藏状态加权并求和, 得到最终表示。式(16)和(17)为注意力权重计算公式。

$$\alpha_t^{(l)} = \frac{\exp(e_t^{(l)})}{\sum_{k=1}^T \exp(e_k^{(l)})} \quad (16)$$

$$e_t^{(l)} = v^T \tanh(W_1 h_t^{f(l)} + W_2 h_t^{b(l)}) \quad (17)$$

式中 $\alpha_t^{(l)}$ 是在第 l 层 GRU 中第 t 个时间步的注意力权重; $e_t^{(l)}$ 是计算注意力权重的中间变量; v^T 是维度可学习的注意力向量 v 的转置; W_1, W_2 是注意力权重计算中的可学习参数。

根据所得注意力权重, 对特征进行加权融合, 融合后的最终特征表达式为

$$\mathbf{h}_t^{(l)} = \sum_{k=1}^T \mathbf{z}_k^{(l)} \quad (18)$$

式中 $\mathbf{z}_k^{(l)}$ 是加权后的隐藏状态,表示第 l 层 GRU 中第 t 个时间步的融合表示; $\mathbf{h}_t^{(l)}$ 是最终的融合表示,是所有时间步融合后的结果。

通过注意力机制动态融合多层双向 GRU 隐藏状态,根据时间步的重要性调整权重,可以更好地捕捉序列特征。

2.3 医学救援动作识别网络设计

对航空医学救援模拟训练中的关键动作进行识别,设计的航空救援动作识别网络预测模型如图 5 所示。首先,人体姿态关键点信息被输入卷积神经网络的输入层。然后,在卷积层对输入信息进行初步特征提取。卷积层使用 32 个尺寸为 2、步长为 1 的卷积核。同时,采用 ReLU 激活函数来引入非线性因素。ReLU 激活函数公式为

$$\text{ReLU}(x) = \max(x, 0) \quad (19)$$

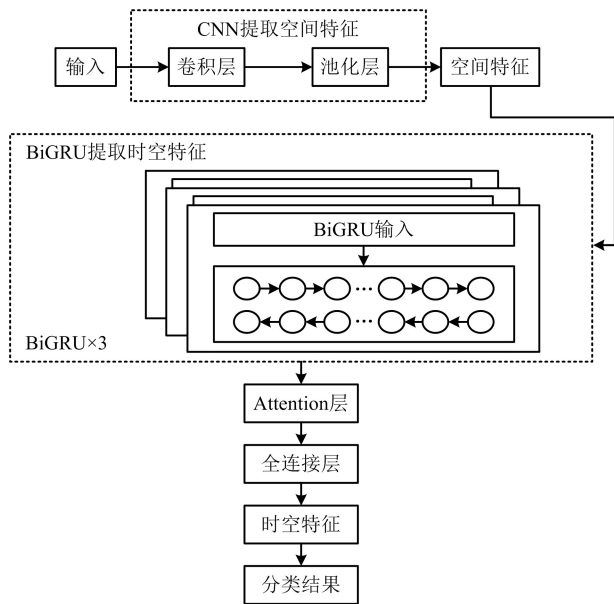


图 5 航空医学动作识别模型结构图

Fig. 5 Architecture diagram of the recognition model for aviation medicine action

池化层采用最大池化进行降维,池化大小为 2,步长为 1。CNN 通过提取空间特征并减少数据尺寸降低复杂性,保留关键特征。然后将 CNN 提取的信息送入 3 层堆砌的 BiGRU(每层 50 个神经元),用于提取浅层、中层和深层时序特征,增强模型的抽象表达能力。

鉴于模拟舱内应用场景,引入注意力机制对 BiGRU 输出加权,以聚焦动作本身。最后通过全连接层与 Softmax 分类器输出分类结果。假设分类数

为 C , Softmax 函数会将每个节点的原始输出转换为一个 C 维的概率向量,这些元素代表该分类的概率,总的概率之和为 1。Softmax 函数公式为

$$S_i = \frac{e^i}{\sum_j e^j} \quad (20)$$

式中 e^i 代表第 i 节点的输出值; j 表示分类数; e^j 代表第 j 节点的分类值; S_i 表示分类后的概率。

2.4 网络性能评价指标

以分类准确率为评价指标,将心肺复苏按压姿势分为 3 类,即标准、斜肩和曲肘,分类准确率定义为

$$A = \frac{T_{B_z} + T_{X_j} + T_{Q_z}}{B_z + X_j + Q_z} \quad (21)$$

式中 B_z 、 X_j 和 Q_z 分别表示标准姿势、斜肩姿势和曲肘姿势的总样本数, T_{B_z} 、 T_{X_j} 和 T_{Q_z} 分别代表这三种按压姿势被正确分类的样本数。

2.5 航空医学应急救援动作评价

动作识别模型将心肺复苏按压姿势分类为标准、斜肩和曲肘。标准动作定义为单侧肘部夹角为 $180^\circ \pm 15^\circ$, 双肩水平夹角 $< 10^\circ$ 。斜肩和曲肘被视为错误姿势。在实际应用中,通过统计各姿势次数及计算标准姿势比重来评价训练效果,计算公式为

$$S = \frac{B_z}{B_z + X_j + Q_z} \quad (22)$$

式中 S 代表训练得分。表 1 为航空医学应急救援动作对应的分级评价标准。

表 1 航空医学应急救援动作评价表

Table 1 Evaluation form for the actions of aviation medical emergency rescue

得分	$S = 100$	$85 \leq S < 100$	$75 \leq S < 85$	$S < 75$
评估结果	优秀	良好	合格	差

2.6 训练参数

基于 Tensorflow 深度学习框架进行试验,训练采用 Adam 优化器,学习率设置为 0.001。每次模型训练迭代中使用的样本数量为 32,训练中随机丢弃神经元的比例为 0.2。选择交叉熵损失作为损失函数指导训练过程,试验轮数(epoch)设置为 50。采用 10 折交叉验证进行模型评估,训练集与测试集比例为 9:1。

模型训练过程如图 6 所示。训练过程占用 GPU 为 0.9 GB,模型参数数量为 164 548。

3 试验结果与分析

将原始的特征数据及经过筛选的上半身特征数

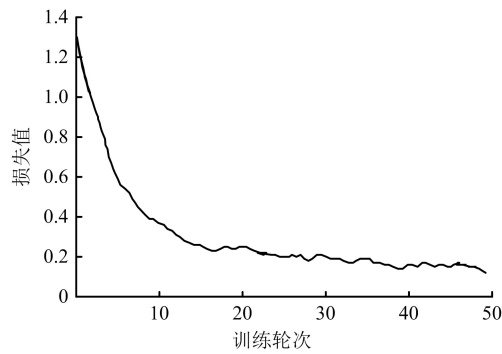


图6 网络模型训练情况

Fig. 6 Training status of the network model

据输入航空医学动作识别网络中进行分类,结果如图7所示,使用上半身关键点信息的识别效果优于全身关键点。原因是在实际胸部按压操作中,腿部无操作要求,且在相机正前方视角下,担架遮挡腿部,不易处理,因此上半身关键点更适合作为输入特征。

在后续试验中,输入模型的关键点为半身关键点;将经过第二步预处理(对关键点信息进行填充、平滑操作)与未经过该预处理的两类数据,分别输入识别网络进行分类,结果如图8所示。

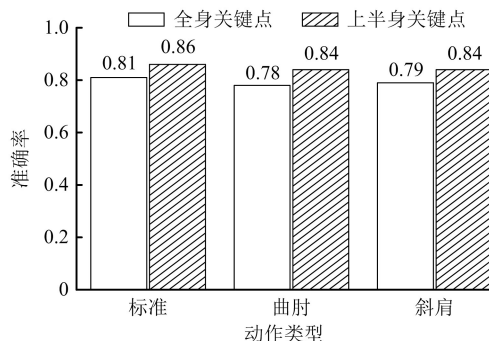


图7 全身关键点和半身关键点试验对比

Fig. 7 Comparison between full-body and half-body key point experiments

经过时域填充、平滑等预处理操作后,使用本文航空医学应急救援动作识别模型的准确率明显高于未处理的关键点数据,针对“标准”“曲肘”“斜肩”动作的识别正确率分别为0.95、0.92和0.93。因此可以说明在航空医学应急救援过程中,外界不利因素及自身的遮挡等问题对试验结果有影响。将关键点信息输入分类网络前需要进行数据预处理。

对时序网络进行了试验对比,试验网络分别为1层GRU+注意力、1层双向GRU+注意力、3层双向GRU、3层双向GRU+注意力及本文的航空救援

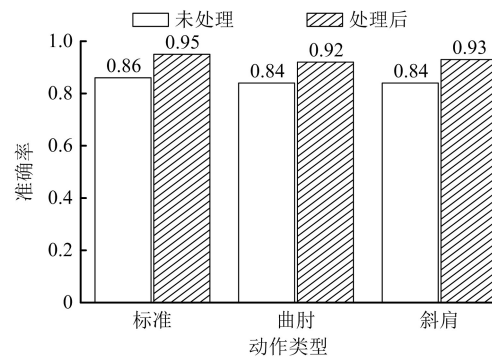


图8 填充与平滑预处理前后效果对比

Fig. 8 Comparison of results with and without filling and smoothing preprocessing

动作识别网络(CNN-BiGRU-Attention)。试验结果如图9所示。

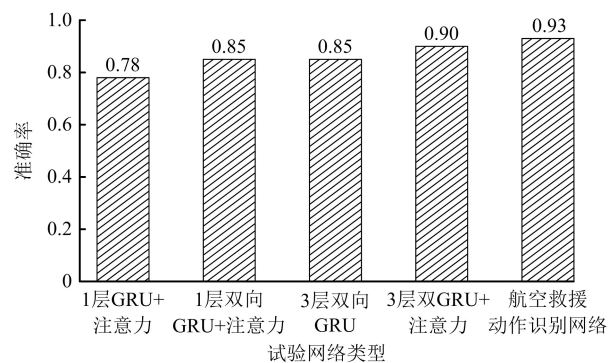


图9 时序网络对比

Fig. 9 Comparison of temporal networks

由图9可知,使用双向GRU的准确率比单向GRU有提高,添加注意力机制后3层双向GRU准确率提高,但效果不够理想,本文提出的双向GRU网络结合CNN并添加注意力机制(CNN-BiGRU-Attention)的航空救援动作识别网络对比其余网络在航空医学应急救援心搏骤停关键动作数据集上效果更好。

除此之外,还将本文提出的适用于航空医学应急救援动作识别网络与时间敏感型网络(Temporal Segment Network, TSN)^[22]、时间卷积网络(Temporal Convolutional Network, TCN)进行试验对比,结果如图10所示。

由图10可见:采用基于光流特征的TSN网络在本文心肺复苏按压姿势数据集上的识别准确率较低;TCN网络的识别准确率相较TSN有明显提升,但仍低于本文所提的航空应急救援动作识别网络。在本文数据集场景下,基于光流特征的识别方案效

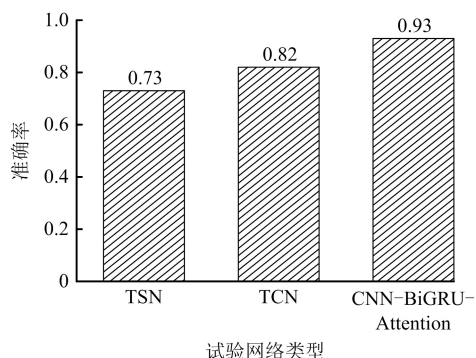


图 10 TSN、TCN 与 CNN-BiGRU-Attention 模型对比

Fig. 10 Comparison among TSN, TCN, and CNN-BiGRU-Attention models

果较差,这一结果与前文对识别算法的分析相一致。其原因主要在于:心肺复苏模拟训练场景光线条件复杂多变,且数据集规模有限,导致光流特征易受干扰、稳定性不足。因此,依赖光流图的动作识别算法在本研究中适用性较差,而基于人体骨骼关键点坐标的时序特征网络模型更适配本任务场景。经对比验证,本文提出的航空应急救援动作识别算法性能最优。

4 结 论

本文采用 MultiPose 人体姿态估计算法构建了心肺复苏按压关键动作数据集,阐述了针对人体姿态关键点信息的数据预处理方法;结合本文对按压姿势的评估需求,对比分析了现有典型动作识别算法的适用性,提出了一种融合 CNN 与双向 GRU 并引入注意力机制的分类网络,基于分类结果实现按压姿势的自动评价,进一步完善了基于计算机视觉的航空医学应急救援模拟训练评估方法。

参考文献 (References):

- [1] 娄玉峰. 加强航空应急救援体系研究:以河南省为例[J]. 中国应急救援, 2021(3): 64-68.
- [2] WANG L F, ZHANG Z H, SHEN Z N, et al. Aviation medical rescue research and equipment development[J]. Science and Technology & Innovation, 2022, 7: 49-51, 58.
- [3] GUO A B, LIU B, FU L, et al. Research on key technologies of air-ground collaboration in aviation emergency rescue[J]. Journal of Natural Disasters, 2022, 31(1): 157-167.
- [4] 高健, 张兵. 国外航空应急救援现状与启示[J]. 中国

民用航空, 2010(10): 22-24.

GAO J, ZHANG B. The status quo of aviation emergency rescue in foreign countries and the revelations[J]. China Civil Aviation, 2010(10): 22-24.

- [5] 中华人民共和国国务院. “十四五”国家应急体系规划[S]. 北京: 中国标准出版社, 2022: 4.
- [6] LI Y, LIU Z, CHEN K, et al. Research on modular scheme of aviation emergency rescue for natural disasters[J]. Journal of Physics: Conference Series, 2022, 2364(1): 012030.
- [7] CARSHAGEN U, VLOK N. An air medical approach to in-flight cardiac arrest management[J]. Air Medical Journal, 2024, 43(2): 100-105.
- [8] VÖGELE A, VAN VEELEN M J, DAL CAPPELLO T, et al. Effect of acute exposure to altitude on the quality of chest compression-only cardiopulmonary resuscitation in helicopter emergency medical services personnel: a randomized, controlled, single-blind crossover trial[J]. Journal of the American Heart Association, 2021, 10(23): e021090.
- [9] FRASCONI R, PASQUARELLA J, HARTIGAN M, et al. Automated versus manual cardiopulmonary resuscitation in flight: Are we being safe? [J]. Air Medical Journal, 2022, 41(3): 303-307.
- [10] 张冰冰, 赵柏山. 基于骨骼关键点的散打动作识别与评价方法[J]. 微处理机, 2023, 44(6): 36-39.
- [11] WANG J F, ZUO L, CORDENTE MARTÍNEZ C. Basketball technique action recognition using 3D convolutional neural networks[J]. Scientific Reports, 2024, 14(1): 13156.
- [12] 郭昕曜, 景国勋, 王远声, 等. 中国通用航空应急救援能力的测度评估方法及其应用研究[J]. 安全与环境学报, 2023, 23(5): 1383-1389.
- [13] 齐福强, 陈姝宁, 孟明源, 等. 山区城市航空应急救援场点布局研究[J]. 安全与环境学报, 2025, 25(3): 1067-1077.
- [14] 王珂, 陈启腾, 陈伟, 等. 基于深度学习的二维人体姿态估计综述[J]. 郑州大学学报(理学版), 2024, 56(1): 1-10.

WANG K, CHEN Q T, CHEN W, et al. Review of 2D

- human pose estimation based on deep learning [J]. Journal of Zhengzhou University (Natural Science Edition), 2024, 56(1): 1–10.
- [15] KOCABAS M, KARAGOZ S, AKBAS E. MultiPoseNet: fast multi-person pose estimation using pose residual network [C]//15th European Conference on Computer Vision (ECCV), September 8–14, 2018, Munich, Germany. Cham, Switzerland: Springer, 2018: 417–433.
- [16] CARREIRA J, ZISSERMAN A. Quo vadis, action recognition? A new model and the kinetics dataset [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21–26, 2017, Honolulu, HI, USA. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, 2017: 6299–6308.
- [17] HU R W, XIANG S J. CNN prediction based reversible data hiding [J]. IEEE Signal Processing Letters, 2021, 28: 464–468.
- [18] GU J X, WANG Z H, LI T, et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77: 354–377.
- [19] ZHANG T M, ZHANG P B, XU P, et al. Multi-stream CNN–LSTM network with partition strategy for human action recognition [C]//2021 International Conference on Bioinformatics and Intelligent Computing, January 15–17, 2021, Virtual Conference. New York, NY, USA: Association for Computing Machinery, 2021: 431–435.
- [20] SUN S Q, WANG Y Y. A novel deep learning automatic modulation classifier with fusion of multichannel information using GRU [J]. EURASIP Journal on Wireless Communications and Networking, 2023, 2023(1): 66.
- [21] 高广尚. 深度学习推荐模型中注意力机制研究综述 [J]. 计算机工程与应用, 2022, 58(9): 9–18.
- GAO G S. Survey on attention mechanisms in deep learning recommendation models [J]. Computer Engineering and Applications, 2022, 58(9): 9–18.
- [22] DING C Y, WEN S, DING W W, et al. Temporal segment graph convolutional networks for skeleton-based action recognition [J]. Engineering Applications of Artificial Intelligence, 2022, 110: 104675.

Visual-based CPR compression posture recognition system for aviation medical emergency rescue

LI Xiaoxue^{1,2}, LI Zhengping³, CAO Shulin³, HAO Yuwen^{1,2}, WANG Lijun^{3,4}, JIA Kehao³

(1 Department of Medical Innovation Research, Chinese PLA General Hospital, Beijing 100853, China;

2 Beijing Key Laboratory of Intelligent Emergency Medical Rescue, Beijing 100853, China;

3 School of Artificial Intelligence and Computer Science, North China University of Technology, Beijing 100144, China;

4 Hangzhou Research Institute, Xidian University, Hangzhou 311231, China)

Abstract: China is facing a shortage of aviation medical emergency rescue personnel and relies on outdated training equipment. This paper addresses a critical challenge in aviation medical emergency rescue simulation training—the low recognition accuracy of Cardiopulmonary Resuscitation (CPR) compression actions—by proposing a vision-based training posture recognition model. The proposed solution is built upon the MultiPose network architecture, incorporating both human keypoint detection and human position detection networks. After processing through a pose residual network, the system generates human position segmentation maps and keypoint information maps, enabling multi-scale feature capture. Temporal mean filling and smoothing techniques are applied to eliminate noise, completing the data preprocessing phase. Subsequently, the proposed training posture recognition model is utilized for identifying CPR compression postures. This model integrates a Convolutional Neural Network (CNN) with a Bidirectional Gated Recurrent Unit (BiGRU) network. To address the common issue of lower-body keypoint occlusion caused by stretcher positioning and compression postures during CPR procedures, the method adaptively extracts spatial correlation features from 13 upper-body keypoints. An attention mechanism is incorporated to enhance the model's focus on essential action characteristics, thereby reducing interference from dynamic environments and visual obstructions. Additionally, recognizing the lack of relevant datasets for evaluating CPR actions, this study constructs a specialized dataset comprising key actions in aviation medical cardiac arrest rescue. Developed in accordance with American Heart Association guidelines for CPR posture compliance, the dataset categorizes actions into three classes: standard, bent elbows, and slanted shoulders, with the latter two representing the primary erroneous postures targeted for recognition. Irrelevant samples are included to improve model robustness. Experimental results demonstrate that by implementing upper-body keypoint recognition, processing keypoint data, and integrating an attention mechanism, the proposed model achieves a posture recognition accuracy of 93%. This performance surpasses that of other models and meets the operational requirements for aviation medical cardiac arrest rescue scenarios.

Key words: public safety; cardiopulmonary resuscitation; posture recognition; attention mechanism