

# 基于小样本学习的老年人异常行为识别模型研究

王聪敏<sup>1</sup>, 周艳聪<sup>2</sup>

<sup>1</sup>天津商业大学理学院, 天津

<sup>2</sup>天津商业大学信息工程学院, 天津

收稿日期: 2023年2月18日; 录用日期: 2023年3月20日; 发布日期: 2023年3月28日

## 摘要

当今社会老龄化程度逐步加重, 老年人的看护问题日益凸显。基于视频识别的老年人异常行为识别旨在助力老年人看护问题。但该领域相关公开的老年人异常行为数据稀缺。针对这种情况, 本文收集了老年人异常行为数据, 并结合小样本学习设计了老年人异常行为识别模型(RAAE, Recognition model of Abnormal Action of the Elderly)。该网络通过patch级富集模块获取patch级空间信息, 后接通道注意力模块获取帧间时间语义信息, 最后通过CrossTransformer进行时间建模。实验表明该网络和以往模型相比, 对老年人异常行为的识别准确率得到了4.7%的提升。

## 关键词

异常行为识别, 小样本学习, 深度学习

# Research on Recognition Model of Abnormal Behavior of the Elderly Based on Few-Shot Learning

Congming Wang<sup>1</sup>, Yancong Zhou<sup>2</sup>

<sup>1</sup>School of Science, Tianjin University of Commerce, Tianjin

<sup>2</sup>School of Information Engineering, Tianjin University of Commerce, Tianjin

Received: Feb. 18<sup>th</sup>, 2023; accepted: Mar. 20<sup>th</sup>, 2023; published: Mar. 28<sup>th</sup>, 2023

## Abstract

Nowadays, the aging of society is gradually increasing, and the care of the elderly is increasingly

文章引用: 王聪敏, 周艳聪. 基于小样本学习的老年人异常行为识别模型研究[J]. 计算机科学与应用, 2023, 13(3): 465-471. DOI: 10.12677/csa.2023.133045

**prominent. The recognition of abnormal behaviors of the elderly based on video recognition aims to help the elderly care problems. However, relevant public data on abnormal behavior of the elderly in this field are scarce. In response to this situation, this paper collected the data of abnormal behavior of the elderly, and designed a recognition model of abnormal behavior of the elderly (RAAE, Recognition model of Abnormal Action of the Elderly) based on few-shot learning. The network obtains the patch-level spatial information through the patch-level enrichment module, and then obtains the inter-frame temporal semantic information through the channel attention module. Finally, time modeling is performed through Cross Transformer. The experiment shows that compared with previous models, the recognition accuracy of the network for abnormal behavior of the elderly has been improved by 4.7%.**

## Keywords

**Abnormal Behavior Recognition, Few-Shot Learning, Deep Learning**

Copyright © 2023 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

我国自 21 世纪进入老龄化社会以来, 社会生育率逐年下降。随着社会福利、医疗技术和卫生水平的提升, 老年人的平均寿命在不断延长。这也加剧了老龄化社会的问题, 年轻人忙于工作和学业, 不能随时照看老人。养老系统目前存在人力短缺的情况, 科技的发展给养老问题带来了新的解决思路。异常行为识别技术可以监护老年人的日常生活[1], 一旦识别出老年人有异常行为, 例如出现摔倒、发病等异常情况[2], 立即发出警报, 提醒监护人员快速采取有效措施, 避免不必要的伤害[3]。因此, 老年人异常行为识别成为当前研究热点。

老年人异常行为识别根据关键技术不同, 可分为三类[4]: 基于穿戴式设备[5]、基于环境传感器[6] [7] [8]和基于计算机视觉[9] [10]。前两种存在影响老年人正常生活、设备条件要求高的问题, 基于计算机视觉的老年人异常行为识别不会影响老年人的正常生活, 并且对环境和设备要求较低, 不用另外安装传感器等设备。但是基于计算机视觉的老年人异常行为识别领域存在相关公开数据集稀缺的情况。由于视频涉及隐私、收集成本高等原因, 无法像其他行为识别一样用大量的数据训练模型。此外, 由于对异常行为的界定及判别方法与场景因素紧密相关[11], 因此需要收集针对老年人日常生活场景相关的异常行为数据, 才能更好的识别。目前的老年人异常行为识别技术大多是基于统计模型的方法, 即人体序列的排列组合情况及各个时间点信息进行训练和分类。常见的方法有模板匹配法、距离分类法、状态空间法和状态统计法[12], 但是这些方法都是建立在大量数据基础上, 针对识别数据稀缺的异常行为类别的情况表现不佳。小样本学习旨在解决样本量小的识别任务, 它的提出基于对人类学习思维的模仿[13], 主要通过训练模型识别事物的异同来将其归类[14], 用在老年人异常行为领域, 可以提高识别效率。针对以上问题, 本文首先收集老年人异常行为数据, 建立需要关注的老年人异常行为数据集, 其次引入小样本学习来对老年人异常行为进行识别。

## 2. 相关工作

### 2.1. 老年人异常行为数据集

老年人身体状况与年轻人有较大差异, 需要重点关注的行为类别也与平时的行为识别数据集所包含

的不同。目前, 针对家居场景下的老年人异常行为的数据集还未发现。相关公开数据集有, 但相关系数较小。例如跌倒检测数据集、异常行为数据集。跌倒检测数据集包括 the UR fall detection dataset (URFD) [15]、the fall detection dataset (FDD) [16]等, 但是只包含跌倒一个动作。异常行为数据集如 UMN (University of Minnesota)数据库[17], 所包含的动作类与老年人无关。

## 2.2. 老年人异常行为识别

老年人异常行为识别受到计算机视觉领域研究人员的关注和探索。利用老人房间内的摄像头采集的视频数据, 从图像序列中提取行为特征, 实现对行为的分类并判断是否有异常行为发生。而在研究中收集各种可能异常的数据或者大量带标签的异常实例是极困难的。针对这种情况, 无监督设置假设只有正常的训练数据, Luo [18]等人提出了一种基于时间相干稀疏编码的方法, 该方法可以映射到堆叠的 RNN 框架; 弱监督设置可以访问带有视频级标签的视频。Hui Lv [19]等人针对异常检测中的噪声干扰和定位引导缺失提出了一种弱监督异常定位方法, 提出的一种高阶上下文编码模型, 不仅可以提取语义表示, 还可以测量动态变化, 从而有效地利用时间上下文。相对无监督学习及弱监督学习, 对于老年人异常行为存在的数据少、个性化强等特点, 本文采用小样本学习方法进行识别。

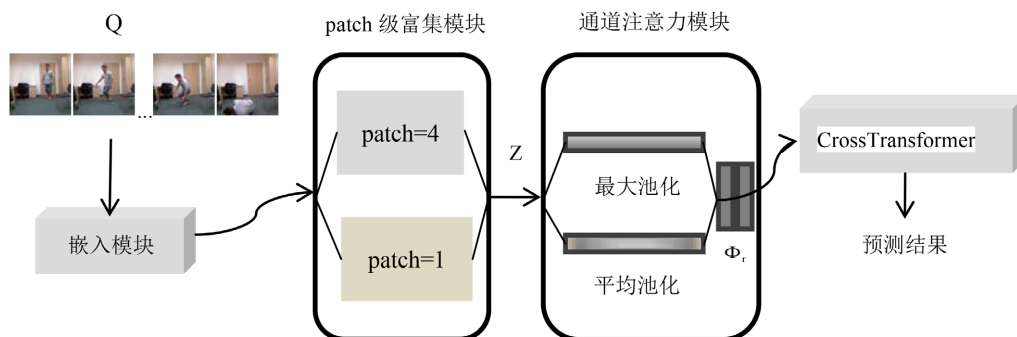


Figure 1. RAAE architecture  
图 1. RAAE 框架

## 3. 本文方法

### 3.1. 问题定义

在小样本动作识别中, 需要分类的视频称为查询视频, 小样本学习的一个任务就是将一个不带标签的查询视频划分为几个类中的具体类别, 每个类都由少量在训练中未见过的有标签的样本组成, 称为支持集。每个任务视为一个 N-way K-shot 分类问题。N-way K-shot 是指支持集包含 N 个类别, 每个类别包含 K 个有标签的支持集样本。每个任务的数据由支持集 S 和查询集 Q 组成, 大多数文献中,  $Q = \{q_1, \dots, q_F\}$  是一个  $F = 8$  均匀采样的帧样本集合。  $S = \{S_c\}$ ,  $c = 1, \dots, N$ , 是包含 N 个类的支持集, 第 c 类第 i 个视频表示为  $S^c = \{S_{i1}^c, \dots, S_{iF}^c\}$ 。为了方便比较, N 设置为 5, K 设置为 5。

### 3.2. 异常行为识别网络

TRX [20]是一种先进的小样本行为识别方法, 在支持集中的查询和视频之间寻找时间对应的帧元组。它使用 ResNet [21]网络作为嵌入网络提取特征, 该网络所包含的 Residual Network 结构, 能使网络的学习能力随着网络深度的增加而增加; 使用 CrossTransformers [22]注意机制来构建类原型来观察所有支持视频的相关子序列, 而不是使用类平均值或单个最佳匹配。TRX 结合 Resnet50 和 CrossTransformers 实现

了高效的小样本行为识别, 但是 TRX 中的 CrossTransformers 对于时间的建模仅限于采样帧的重排, 而没有提取帧维度上新的语义信息, 本文在 TRX 的基础上进行改进, 引入通道注意力[23]学习帧维度上的时间语义信息, 以及 patch 级富集模块学习 patch 级空间信息。

本文使用 Resnet50 作为嵌入网络, 结合 patch 级特征富集模块和通道注意力模块提取空间和时间信息进行老年人异常行为的识别。具体模型及流程如图 1, 将稀疏采样的视频帧  $Q = \{q_1, \dots, q_F\}$  输入嵌入网络后, 经过 patch 级富集模块获取不同 patch 级的空间上下文信息, 输出 patch 级空间特征, 提升帧中的动作相关场景人物等上下文的聚合。patch 级富集模块包含两个不同 patch 大小的提取器, 分别提取不同 patch 级特征  $I_1^2 * D, I_2^2 * D$ , 然后将其堆叠展平得到  $M_i \in \mathfrak{R}^{(I_1^2 + I_2^2) * D}$ 。将  $M_i$  进行空间平均得到  $Z$  以聚合 patch 级信息。随后将 patch 级富集特征输入通道注意力模块, 分别经过平均池化和最大池化, 进行视频级特征压缩, 提取通道注意力权重。再利用一个双层 MLP  $\Phi_r$  对中间特征进行细化, 公式如下

$$g(H) = [\Phi_r(AvgPool(Z)) + \Phi_r(MaxPool(Z))] * Z \tag{1}$$

最后使用 CrossTransformers 注意机制来构建类原型, 它通过时间建模观察所有支持视频的相关子序列, 以此输出异常行为预测结果。

## 4. 实验结果与分析

### 4.1. 数据集

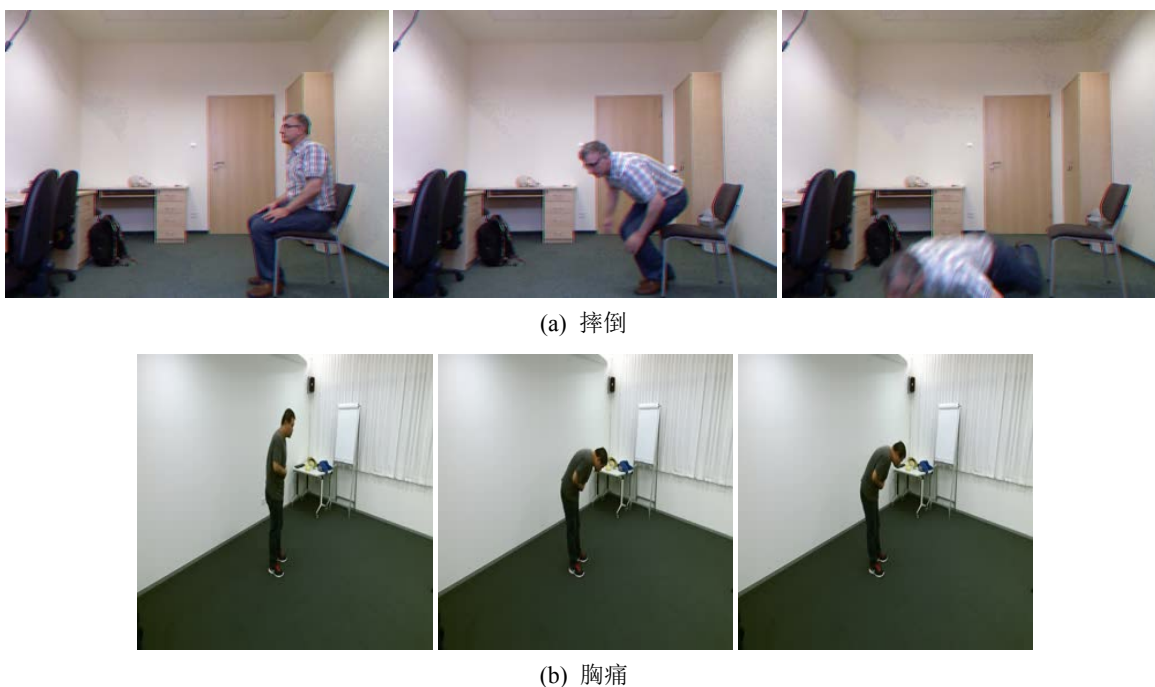


Figure 2. Example of abnormal behavior video  
图 2. 异常行为视频示意图

本文针对老年人需要关注异常行为类别, 确定了 12 个需要关注的老年人异常行为类别。并收集了相关的视频数据。其中包括打喷嚏/咳嗽、走路摇晃、摔倒、头痛、胸痛、背痛、颈部疼痛、恶心/呕吐、给自己扇风、打哈欠、伸展身体、擦鼻子。数据主要来源于[24] [15]。所有视频都是室内场景。每个异常行为类别的视频数量均在 900~1000 范围内。图 2 显示了其中异常行为的例子, a 是摔倒, b 是胸痛。

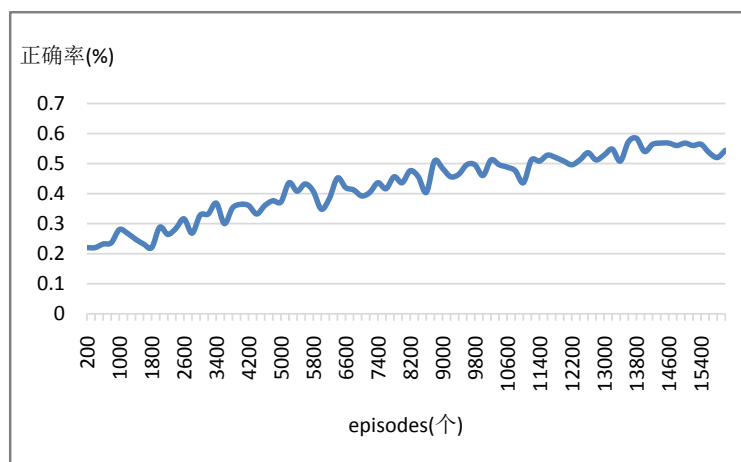


Figure 3. TRX training accuracy line chart

图 3. TRX 训练准确率折线图

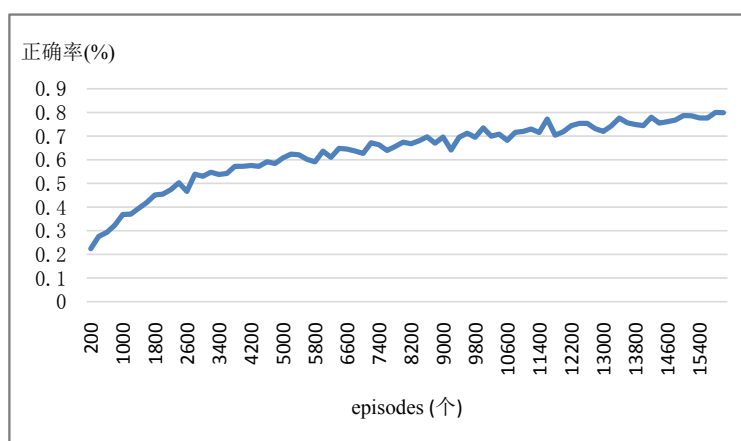


Figure 4. RAAE training accuracy line chart

图 4. RAAE 模型训练准确率折线图

## 4.2. 实验设置

本文采用小样本学习的 episodes 训练策略来训练模型。在 5-way 5-shot 的基准上评估本文方法, 并且记录了在测试集上 1000 个随机任务的平均结果。视频在输入网络之前先进行提取关键帧和缩放的处理。提取关键帧采用  $F = 8$  的均匀采样方式。缩放大小为  $256 \times 256$ 。模型训练时, 对数据进行随机水平翻转和  $224 \times 224$  的中心裁剪的数据增广处理。测试时, 只有中心裁剪。使用学习率为 0.0001 的 SGD 优化器, 训练至少 15,000 个任务。实验环境为: 在一个 NVIDIA RTX A40GPU 上训练。

## 4.3. 实验结果

首先通过 9 个异常行为类别的数据来对原始的 TRX 模型进行训练, 训练了 16,000 个 episodes, 然后用 1000 个测试样本集对网络模型的识别率进行了验证。最终的准确率变化曲线如图 3 所示。

TRX 模型准确率的折线图显示, 随着训练过程中 episodes 数的不断增加, 准确率整体呈现不断上升的趋势, 在 14,000 个 episodes 时准确率为 55% 并不再上升。episodes 数在 0~10,000 之间准确率上升的速度较快, 在 10,000~14,000 个 episodes 之间准确率上升曲线变得平缓。最终 TRX 的测试准确率为 55.7%。

使用相同的训练设置和数据对改进之后的 RAAE 模型进行训练和测试, 最终的准确率折线图如图 4 所示。

根据图 4 模型训练准确率折线图可以看出, 随着 episodes 个数的不断增加, 异常行为识别训练准确率不断上升, 准确率折线图形似指数曲线。在 5000 个 episodes 之前训练准确率上升比较快, 从约 20% 上升到约 60%, 5000 个 episodes 之后, 训练准确率上升趋势逐渐趋于平稳。15000 个 episodes 之后, 训练准确率基本平稳, 不再上升。最终的训练准确率为 80%, 此时的测试准确率达到 60.4%。与 TRX 相比, 准确率有了进一步的提高。此外, 观察两图, RAAE 训练准确率变化更平缓, 波动更小, 这说明本文模型更稳定, 训练前期准确率的快速增长也说明 RAAE 模型学习效率更高, 能在短时间内学习更多特征。

**Table 1.** Comparison table of detection accuracy of different models

**表 1.** 不同模型检测准确率对比表

模型	TRX	RAAE
准确率(%)	55.7	60.4

TRX 与 RAAE 之间的测试准确率对比如表 1 所示, 可以看出, 本文模型 RAAE 对 TRX 的识别率进行的改进非常有效, 这都归功于 RAAE 对于空间信息和帧间信息的利用, 使其得到了 4.7% 的提升。

## 5. 结束语

针对老年人异常行为数据稀缺的现状, 本文先是收集了老年人异常行为相关数据集, 其中包括头痛、胸痛、摔倒等亟需关注的老年人异常行为。随后介绍了 TRX 模型, 针对 TRX 中时间建模中没有产生新的时间语义信息的问题, 提出了使用通道注意力模块学习帧间时间语义信息的方法, 并且结合 patch 级富集模块学习 patch 级空间信息, 设计了老年人异常行为识别网络 RAAE。实验结果表明, 本文模型 RAAE 对老人异常行为识别的准确率高达 60.4%, 准确率相对 TRX 得到了明显的提升, 能有效的识别老年人异常行为。

## 参考文献

- [1] Lentzas, A. and Vrakas, D. (2020) Non-Intrusive Human Activity Recognition and Abnormal Behavior Detection on Elderly People: A Review. *Artificial Intelligence Review*, **53**, 1975-2021. <https://doi.org/10.1007/s10462-019-09724-5>
- [2] 于乃功, 柏德国. 基于姿态估计的实时跌倒检测算法[J]. 控制与决策, 2020, 35(11): 2761-2766.
- [3] 黄加成. 中国老龄化社会独居老人研究进展[J]. 中国老年学杂志, 2015, 35(23): 6954-6956.
- [4] Mubashir, M., Shao, L. and Seed, L. (2013) A Survey on Fall Detection: Principles and Approaches. *Neurocomputing*, **100**, 144-152. <https://doi.org/10.1016/j.neucom.2011.09.037>
- [5] 李饶. 老年人群居家异常行为检测技术研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2021.
- [6] 王宇, 周迎, 王灵灵, 等. 室内环境下老年人异常行为检测[J]. 土木工程与管理学报, 2022, 39(4): 145-152.
- [7] 李一凡. 基于 Wi-Fi 的非接触式老年人异常行为识别研究[D]: [硕士学位论文]. 兰州: 西北民族大学, 2022.
- [8] 彭小英. 室内环境下老年人异常行为检测系统设计与研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2020.
- [9] 温佳昆. 基于机器学习的老年人异常行为检测方法研究[D]: [硕士学位论文]. 沈阳: 沈阳理工大学, 2021.
- [10] 孟祥泽, 胡啸峰, 沈兵. 社区老年人空间行为轨迹异常分析方法[J]. 科学技术与工程, 2021, 21(9): 3676-3681.
- [11] 张晓平, 纪佳慧, 王力, 等. 基于视频的人体异常行为识别与检测方法综述[J]. 控制与决策, 2022, 37(1): 14-27.
- [12] 杨鹏, 李潇婧, 张慧君, 等. 智能监控系统下的老年人异常行为研究进展[J]. 中国康复医学杂志, 2019, 34(1): 112-115.
- [13] 赵凯琳, 靳小龙, 王元卓. 小样本学习研究综述[J]. 软件学报, 2021, 32(2): 349-369.
- [14] 韦世红, 刘红梅, 唐宏, 朱龙娇. 多级度量网络的小样本学习[J]. 计算机工程与应用, 2023, 59(2): 94-101.

- 
- [15] <http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html>
- [16] <http://falldataset.com>
- [17] <http://mha.cs.umn.edu/projevents.shtml#crowd>
- [18] Luo, W.X., Liu, L. and Gao, S.H. (2017) A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 341-349. <https://doi.org/10.1109/ICCV.2017.45>
- [19] Lv, H., Zhou, C.W., Xu, C.Y., *et al.* (2021) Localizing Anomalies from Weakly-Labeled Videos. *IEEE Transactions on Image Processing*, **30**, 4505-4515.
- [20] Perrett, T., Masullo, A., Burghardt, T., Mirmehdi, M. and Damen, D. (2021) Temporal-Relational CrossTransformers for Few-Shot Action Recognition. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 475-484. <https://doi.org/10.1109/CVPR46437.2021.00054>
- [21] He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778.
- [22] Doersch, C., Gupta, A. and Zisserman, A. (2020) CrossTransformers: Spatially-Aware Few-Shot Transfer. 34th *Conference on Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, 6-12 December 2020, 1-3+5.
- [23] Woo, S., Park, J., Lee, J.-Y. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. *Computer Vision—ECCV 2018*, Munich, 8-14 September 2018, 3-19.
- [24] Liu, J., Shahroudy, A., Perez, M., Wang, G., Duan, L.-Y. and Kot, A.C. (2019) NTU RGB+D 120: A Large-Scale Benchmark for 3D Human Activity Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, **42**, 2684-2701. <https://doi.org/10.1109/TPAMI.2019.2916873>