

基于RK3588的老人智能监控系统设计

李志翔^{1,2}, 黄剑华¹, 甘仿², 蒋淦华², 程巧玲²

¹江西师范大学物理与通信电子学院, 江西 南昌

²江西软件职业技术大学智能科技学院, 江西 南昌

收稿日期: 2024年3月15日; 录用日期: 2024年4月19日; 发布日期: 2024年4月29日

摘要

随着人口老龄化的不断加剧, 老年人的居家安全问题变得日益突出, 尤其是跌倒和火灾等意外事件。传统的监控方式主要依赖于人工观察和分析, 难以实现对视频内容的智能识别、分析与处理, 无法及时有效地防范潜在的安全风险。该文研究并设计一种基于RK3588的老人智能监控系统, 内容包括系统平台设计、智能算法设计、智能识别等。该系统不仅实现了视频画面的线上实时监控与视频录制, 而且可以快速并准确地对姿态和火焰进行识别。与传统的家庭监控系统相比, 本研究设计的系统解决了监控存在的效率低下和便捷性不足的问题, 大幅提升了老人安全监控的效能。

关键词

智能监控系统, 姿态识别, 火焰识别, 加速推理

Design of Intelligent Monitoring System for the Elderly Based on RK3588

Zhixiang Li^{1,2}, Jianhua Huang¹, Fang Gan², Ganhua Jiang², Qiaoling Chen²

¹College of Physics and Communication Electronics, Jiangxi Normal University, Nanchang Jiangxi

²School of Intelligent Science and Technology, Jiangxi University of Software Professional Technology, Nanchang Jiangxi

Received: Mar. 15th, 2024; accepted: Apr. 19th, 2024; published: Apr. 29th, 2024

Abstract

With the escalating aging population, the issue of elderly home safety has become increasingly prominent, especially regarding accidents such as falls and fires. Traditional monitoring methods primarily rely on manual observation and analysis, which are inadequate for intelligent recognition, analysis, and processing of video content. Consequently, they fail to promptly and effectively prevent

potential safety risks. This paper studies and designs an intelligent elderly monitoring system based on RK3588, encompassing system platform design, intelligent algorithm design, and intelligent recognition. This system not only enables real-time online monitoring and video recording of video footage but also rapidly and accurately identifies posture and flame detection. Compared to traditional home monitoring systems, the system designed in this study addresses issues of inefficiency and inconvenience, significantly enhancing the effectiveness of elderly safety monitoring.

Keywords

Intelligent Monitoring System, Posture Recognition, Flame Detection, Accelerated Inference

Copyright © 2024 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

随着中国社会的快速发展和人口结构的转变，老龄化问题日益凸显。根据国家统计局发布的数据，截至 2023 年末，全国有 60 岁及以上人口 29697 万人，占全国总人口的 21.1%，其中 65 岁及以上人口 21676 万人，占全国总人口的 15.4% [1]。国家卫生健康委老龄司司长王海东表示，据测算，预计“十四五”时期，60 岁及以上老年人口总量将突破 3 亿，占比将超过 20%，进入中度老龄化阶段。庞大的老年人口基数使得老年人的居家安全问题成为了社会关注的焦点。据调查显示，老人在家庭中事故主要集中在火灾和跌倒两个方面，而这些事故往往因为缺乏及时的反应和救助而造成不可挽回的后果[2] [3]。

传统的家庭监控系统在老年人居家安全方面存在一定的局限性，如反应速度慢、误报漏报率高等问题，无法满足对老年人实时监控和及时救援的需求。特别是在老年人跌倒、突发疾病等紧急情况下，传统的监控系统往往难以及时做出有效反应，给老年人的生命安全带来潜在威胁。随着科技的进步和智能化的发展，智能监控系统将在未来为家庭安全保驾护航。

1.1. 视频监控系统的研究现状

家庭监控系统从早期的简单摄像头监控，到现在的智能化、网络化的监控系统，家庭监控技术在不断地发展和完善。主要分为以下几个时期。

1) 早期阶段(20 世纪 90 年代之前)

家庭监控系统的概念在这一时期开始萌芽。最初的家庭监控系统主要是基于模拟信号的电视系统。如图 1 所示，这些系统通常包括一个或多个摄像头、一台录像机和一些简单的控制设备。摄像头捕捉到的视频信号通过同轴电缆传输到录像机进行录制，用户可以通过观看录像或实时监控来确保家庭安全。视频录像使用 VCR 磁带对进行存储，消耗了大量资源。

2) 发展阶段(20 世纪 90 年代中期至 21 世纪初期)

随着网络技术的普及和数字化趋势的加速，出现了第二代视频监控系统——模数结合的视频监控系统(图 2)，它以硬盘录像机为代表的监控系统，DVR (digital video recorder)可以将模拟的视频信号进行数字化，并储存在计算机硬盘而不是盒式录像带上。这种数字化的存储大大提高了用户对录像信息的处理能力，使用户可以通过 DVR 来控制摄像机的开关，从而实现了移动侦测功能。此外，对于报警时间以及事前和事后报警信息的搜索也变得非常简单[4]。

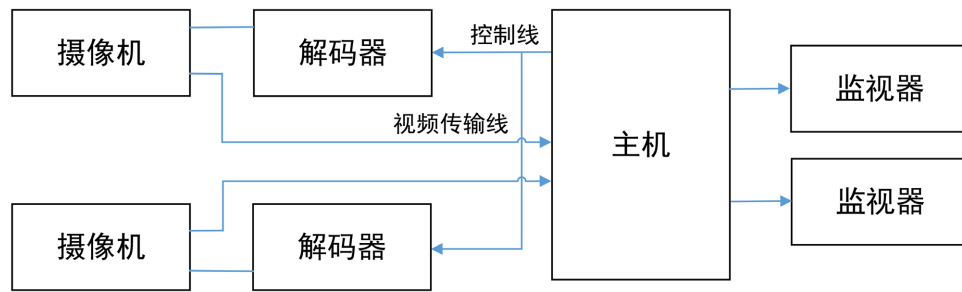


Figure 1. Analog signal monitoring system
图 1. 模拟信号监控系统

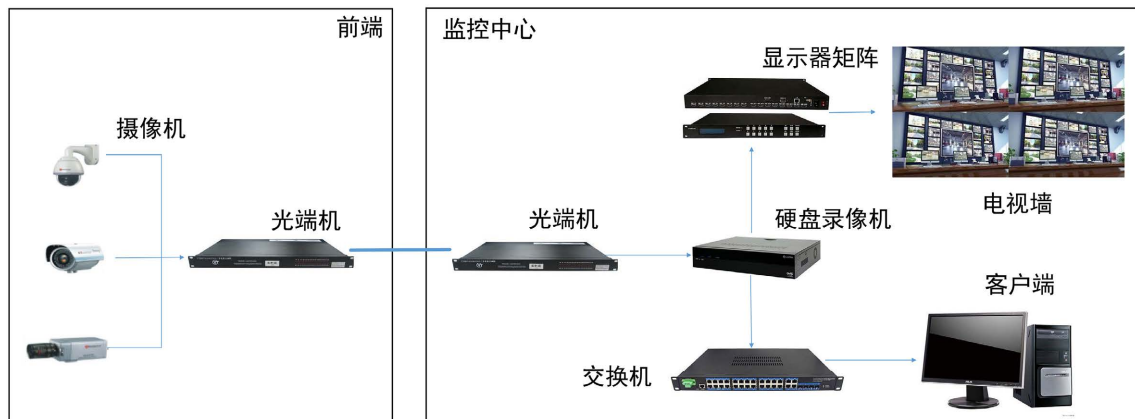


Figure 2. Hybrid analog-digital video monitoring system
图 2. 模数混合视频监控系统

3) 现阶段(21 世纪中期至今)

近年来，随着人工智能、物联网和大数据技术的飞速发展，家庭监控系统正逐步向智能化、集成化和个性化的方向发展。现代的家庭监控系统不再仅仅局限于安全监控，而是融入了智能家居、健康管理等多种应用场景。例如，一些家庭监控系统可以通过人脸识别技术，自动识别家庭成员并调整监控策略。通过集成智能家居设备，实现与灯光、窗帘等设备的联动控制，甚至可以通过大数据分析，提供健康管理和生活建议等服务。

随着深度学习算法的发展，家庭监控系统的智能化水平不断提高。研究者们致力于开发更加高效的目标检测、识别、跟踪等算法，以提高家庭监控系统的准确性和实时性。同时，还有一些研究关注于隐私保护问题，如何在实现智能化监控的同时，保护用户的隐私不被侵犯。

1.2. 跌倒检测的研究现状

跌倒检测与识别主要包含两种方式：一种是使用者穿戴具有惯性测量单元的设备，该设备收集姿态信息，通过对数据进行姿态解算获取当前使用者信息；另一种是使用视觉传感器对跌倒者进行状态检测与识别。本系统采用视觉传感器实现，因此主要讨论第二种方式的研究。

在视觉检测中，跌倒检测与识别的主要步骤分别是目标检测、特征提取和定性分类。在目标检测方面中，通过背景差分、帧间差分等方法实现对图像和视频中的被测目标进行定位[5]。特征提取部分则是根据算法结构和处理方式采取不同的方法获取特征。最后，根据不同属性判定数据的类别实现定性分类。研究人员主要围绕特征提取部分展开工作。在特征提取方面，对跌倒检测采用阈值分析法、机器学习和深度学习三类方法实现。

1) 阈值分析法是通过检测人体体态转变时的特征值来判断是否产生跌倒动作，常用于便携式设备中。如 Sun 等[6]使用足底倾斜角传感器系统来识别人体状态，通过分析双足移动时的 X 轴和 Y 轴倾斜角变化获取姿态，但传感器的测量数据波动幅度呈现两极分化。Yuan 等[7]采用 Kinect 传感器进行跌倒检测，通过传感器收集深度图像序列，再提取图像序列中的人体轮廓计算人体长宽比等相关参数，最后将参数与阈值作比较，获得目标的行为状态，但该算法容易出现混淆弯腰与跌倒的问题，需要进一步的改善阈值的设置。通过设定阈值分析人体动作是一种较为简单的算法，识别效果的差距也较大，容易出现识别不准确的问题。

2) 采用机器学习算法进行跌倒检测是通过人工设计的特征算子的方式处理视觉数据，如 HOG (Histogram of Oriented Gradients)特征[8]、ORB (Oriented FAST and Rotated BRIEF)特征[9]和 LBP (Local Binary Pattern)特征[10]等，该类算法含有多种功能的分类器，具有显著的分类优势。由于真实跌倒的数据集量较少，并且机器学习算法在轻量级的数据集在性能与效果上要胜于重量级数据集，因此现阶段采用机器学习的跌倒检测算法仍在进一步研究中。

3) 深度学习是当前一项热门的技术，在图像处理领域取得了显著的成果，其中的 CNN (Convolutional Neural Network, CNN)在人体动作识别发挥着重要的作用。CNN 算法主要可分为 CNN、1D CNN (one-dimensional convolutional neural network)及 3D CNN (three-dimensional convolutional neural network) [10]。

阈值分析、机器学习和深度学习都在跌倒检测领域中发挥着重要作用，其中深度学习模型作为该领域的前沿算法，拥有较高的识别准确率，可以预估今后跌倒检测的发展趋势将更倾向与 CNN 结合及其拓展算法。

1.3. 火灾检测的研究现状

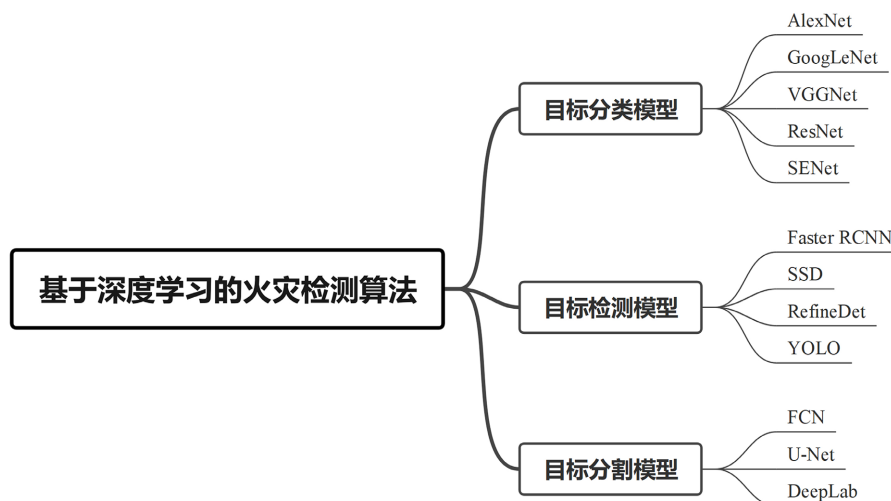


Figure 3. Common models for fire detection algorithms based on deep learning
图 3. 基于深度学习的火灾检测算法常用模型

对于基于视频图像的火灾检测算法，国内外研究学者采用机器学习和深度学习的实现火灾检测[11]。基于机器学习的火灾检测在预处理阶段存在大量冗余的计算开销，而在特征提取阶段仅能获取图像的低级特征，且提取的特征依赖于先验知识，缺乏普适性。深度学习是一种能够自动学习训练数据集中数据内部结构的多层级神经网络算法，可自动检测和学习特征，与基于机器学习的烟雾火灾检测算法相比，基于深度学习的烟雾火灾检测算法可抽取更抽象化、更高阶的特征，具有速度快、准确性强、鲁棒性强的优势[12]。

基于深度学习的烟雾火灾检测通常采用目标分类模型、目标检测模型、目标分割模型 3 大类方法实现，其常用模型如图 3 所示。

1.4. 本文的内容组织安排

本文以老人居家安全风险防范现实需求为背景，基于深度学习，研究设计一种老人智能监控系统，对系统整体方案进行设计，包括软硬件设计、目标检测算法设计，并对系统的功能进行了测试和分析。最后对全文做出总结。

2. 系统总体方案设计

老人智能监控系统如图 4 所示。

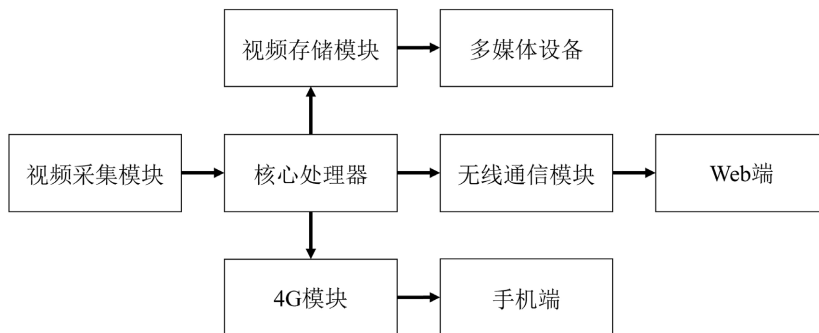


Figure 4. Block Diagram of an Elderly Smart Home Monitoring System
图 4. 老人智能家庭监控系统框图

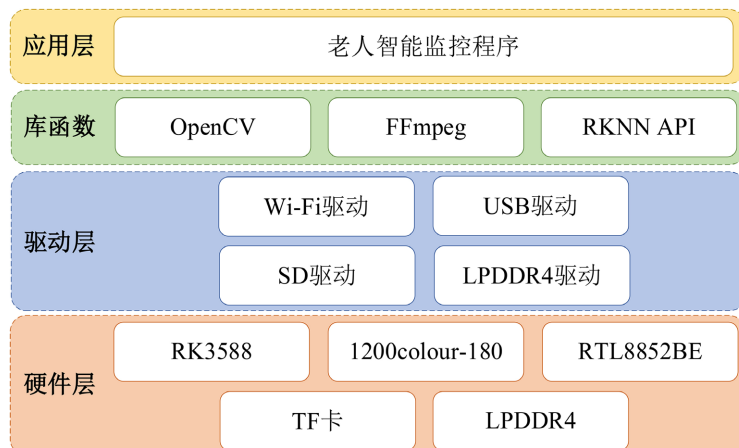


Figure 5. Software architecture of elderly smart monitoring system
图 5. 老人智能监控系统的软件架构

视频采集模块将采集到的视频数据传输到核心处理器进行处理和分析。核心处理器对视频数据进行压缩编码，以减小数据体积，便于存储和传输。同时，核心处理器还将处理后的视频数据通过无线通信模块发送到网络上。无线通信模块负责将视频数据转换为适合网络传输的格式，并通过无线网络将视频流推送到云端或互联网上的服务器。客户端设备(如手机、电脑等)可以通过网络连接访问这些服务器，对视频数据进行拉流操作，实现远程查看家庭的实时监控画面。在紧急情况下，如老人跌倒或火灾发生，核心处理器会检测到这些异常情况，并通过 4G 模块向手机发送紧急信息。4G 模块与核心处理器之间通过串口相连，它负责将核心处理器发送的紧急信息转换为无线传输的格式，并通过移动网络将信息发送到指定的手机号码上，以便及时通知用户。

系统软件架构如图 5 所示。在硬件层中，采用 RK3588 为主处理器、12colour-180 作为摄像头模块、

RTL8852BE 为 Wi-Fi 模块, TF 卡存储监控录制的视频。驱动层为对应硬件层的驱动。在应用层, 需要调用 OpenCV、FFmpeg 和 RKNN API 函数, 编写 C++ 代码实现该系统。

本系统使用的 RK3588 芯片性能强大, 可以运行深度学习算法。使用该芯片部署深度学习的开发流程如图 6 所示。

首先需要进行模型训练, 模型训练前根据本系统具体需求, 选择合适的模型, 采集并标注数据, 然后进行模型训练。本系统需要对老人进行跌倒检测和火焰识别, 因此在训练前需要准备老人跌倒和火焰相关的数据集, 然后对数据集进行标注, 并选择 YOLOv5s 作为本系统的目标检测模型, 在主机端进行模型训练, 训练完成后获得模型权重文件。然后进行模型转换, RK3588 芯片使用 rknn 模型进行推导, 中间需要进行模型转换, 将训练完成的深度学习模型转化为 rknn 模型。最后将转换的 rknn 模型部署到 RK3588 开发平台, 实现深度学习的模型在嵌入式系统中运行。

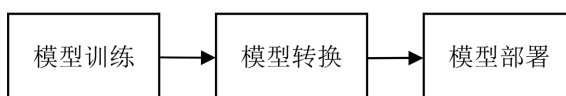


Figure 6. RKNN development process
图 6. RKNN 开发流程

3. 模型训练

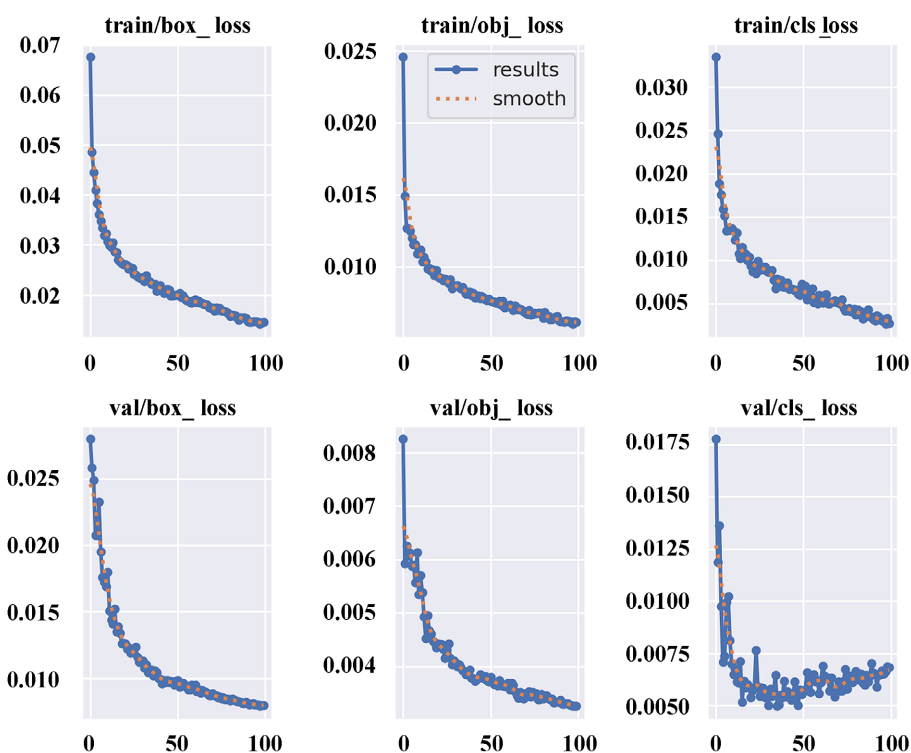


Figure 7. Loss of the training set and validation set during the training process
图 7. 训练过程中训练集和验证集的损失

训练模型的数据由两类组成, 一类是家庭监控拍摄下的一些人体活动的数据集, 另一类是火灾烟雾的数据集。人体活动数据集一共 4000 余张, 火灾烟雾数据集一共 12,000 余张。在人体活动数据集中, 将人体活动分成了“跌倒”“坐立”“直立”三类, 并分别对三类数据进行标注。在火灾烟雾数据集中

将火焰和烟雾分别进行标注。将总数据集按照训练集、测试集和验证集进行分类。

本系统训练轮数为 100，学习率为 1×10^{-5} ，每次输入 8 张图像。训练完成后，训练集和验证集的损失如图 7 所示。从图中可知，随着训练迭代次数的提升，损失呈下降趋势。

4. 模型转换

由于 RK3588 推理框架使用 rknn 模型进行推理，而在 PC 端是采用 Pytorch 模型实现，因此需要对训练后的模型进行模型转换。本系统使用 onnx 作为中间的转换的模型，实现 rknn 模型和 Pytorch 模型之间的转换。onnx 是一种开放的模型格式，可以用来表示深度学习模型，目前已经成为了深度学习模型的标准格式，可用于各种框架模型转换的一种桥梁。因此转换流程为先将 Pytorch 模型转换为 onnx 模型，然后再将 onnx 模型转化为 rknn 模型。

将 Pytorch 模型转换为 onnx 模型需要使用 YOLOv5 文件夹的 export.py 文件。首先定义命令行参数的解析器，并解析命令行参数。这些参数包括模型权重路径、图像尺寸、批量大小、设备选择等；然后加载 Pytorch 模型。根据参数使用 select_device 函数选择运行模型的设备。使用 attempt_load 函数加载指定路径的模型权重，从加载的模型中获取类别标签，通过计算模型的最大步长，确定图像尺寸是否是步长的倍数并进行验证；接着创建的全零张量，用于模型推理的输入；更新模型的处理方式后，使用 TorchScript 转换为 TorchScript 模型。创建一个文件路径 f，将模型以 TorchScript 格式保存在该路径下。转换的模型可以在没有 Python 环境的设备上运行和部署。导出 onnx 模型并放在指定目录下。

将 onnx 模型转换为 rknn 模型的过程中，利用 rknn 神经网络前向推理框架提供的 API 进行操作。整个转换流程如所示。首先，创建一个 rknn 对象作为转换的起点。接着，通过调用 rknn.config 函数进行一系列预处理配置，这包括设置输入图像的归一化参数、调整 RGB 通道的顺序、确定优化等级以及指定目标平台为 RK3588。在归一化处理时，需要准确设置均值和方差以确保数据的正确性。随后，我们提供 onnx 模型的地址并加载模型数据。紧接着，基于加载的模型和数据集，我们创建 rknn 模型，并启用模型的量化功能以提高运行效率。最后，将生成的 rknn 模型导出至指定的保存地址，完成整个模型的转换过程。

5. 模型部署

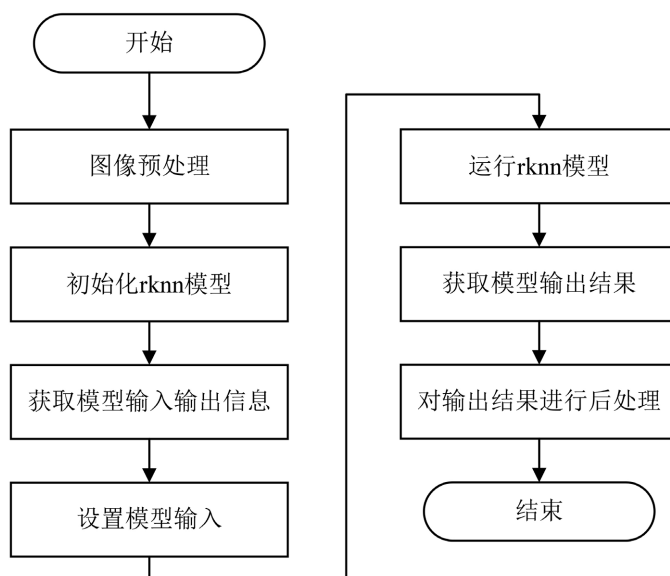


Figure 8. Execution flow chart of the application layer

图 8. 应用层执行流程图

在应用层中，主要是实现通过调用库函数实现该系统功能，如图 8。首先，会对输入图像进行预处理。然后对图像进行缩放操作，以满足 rknn 模型中对图像大小的要求。预处理阶段的图像处理采用 RGA 实现。然后初始化 rknn 模型、获取模型输入输出信息、设置模型输入。接着 rknn_run 函数运行 rknn 模型，得到模型输出结果，对模型输出结果进行后处理。最后释放系统资源。

6. 加速推理实现

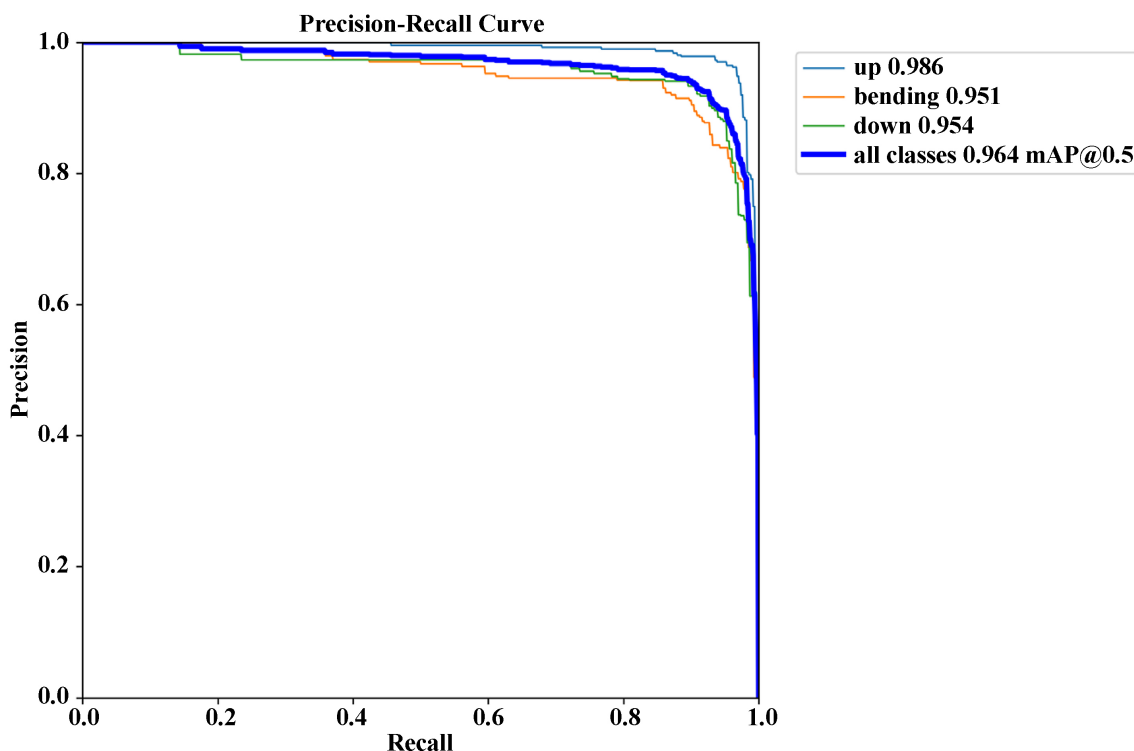
在加速处理部分，首先使用 RKNN 内置的量化方式，目前支持的类型是：asymmetric_quantized-8(非对称量化 8)。量化操作使用在 onnx 转 rknn 阶段。首先初始化模型对象，并调用 rknn.config 配置模型的通道均值、归一化值、量化图片 RGB2BGR 转换、目标平台类型。接着调用 rknn.load_onnx 加载已经转换完成的 onnx 模型，通过将 rknn.build 函数的 do_quantization 设置为 True 开启量化功能。RKNN-Toolkit2 在量化过程中，会根据量化数据集，计算推理结果所需要的量化参数。基于该原因，校准数据集里的数据选择训练集中的一部分，此处从训练集中选择 100 张。最后导出 RKNN 模型。

其次是开启多线程实现神经网络的加速，本系统将视频流的识别分成三个线程：线程一读取视频帧；线程二推理模型，使用线程池处理；线程三根据推理的结果，绘制图形。

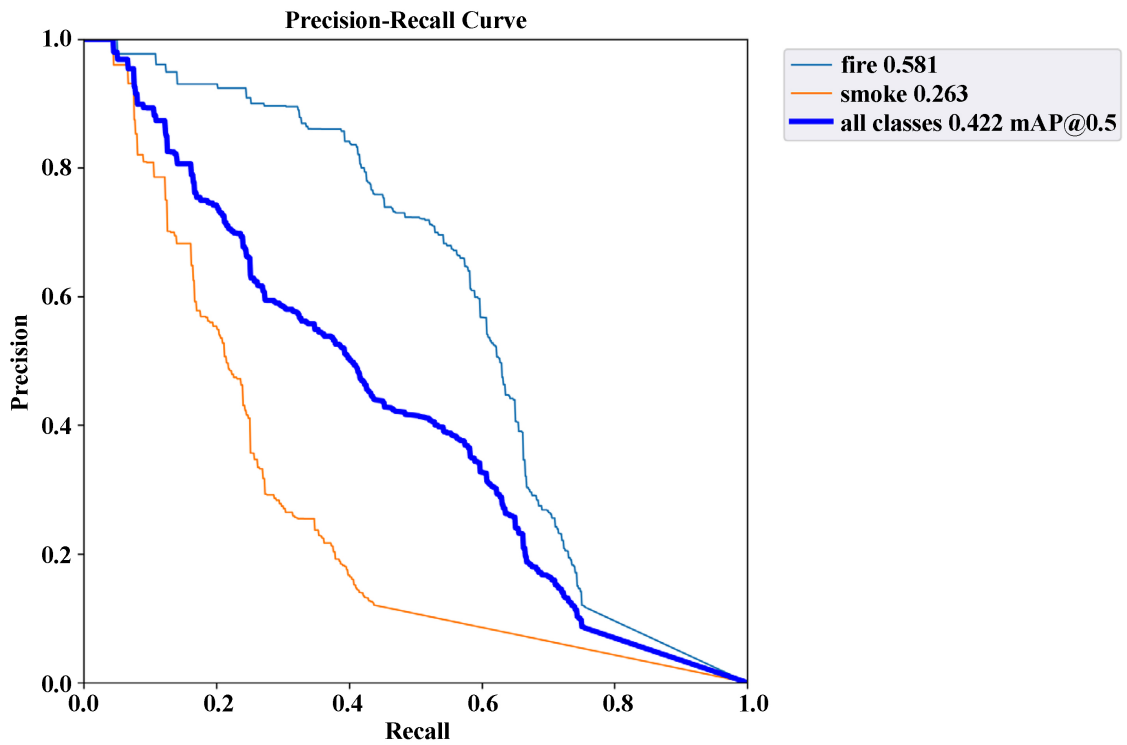
最后一步是修改激活函数，将原本的 SiLU 函数修改为 ReLU 函数。找到 yolov5/models/common.py 文件，将 default_act = nn.SiLU() 修改为 default_act = nn.ReLU() 即可实现激活函数的修改。

7. 系统测试

系统的性能测试，主要从识别准确性和执行速度两个方面展开，测试内容包含了人体姿态和火焰烟雾两个部分。精确率(Precision, P)表示为预测为正样本在所有正样本所占的比例，召回率(Recall, R)表示正样本中被预测为正样本所占的比例。其具体公式定义如公式所示：



(a) 人体姿态检测的 PR 曲线



(b) 火焰烟雾的 PR 曲线

Figure 9. Precision-Recall curve of detection results

图 9. 检测结果的 PR 曲线

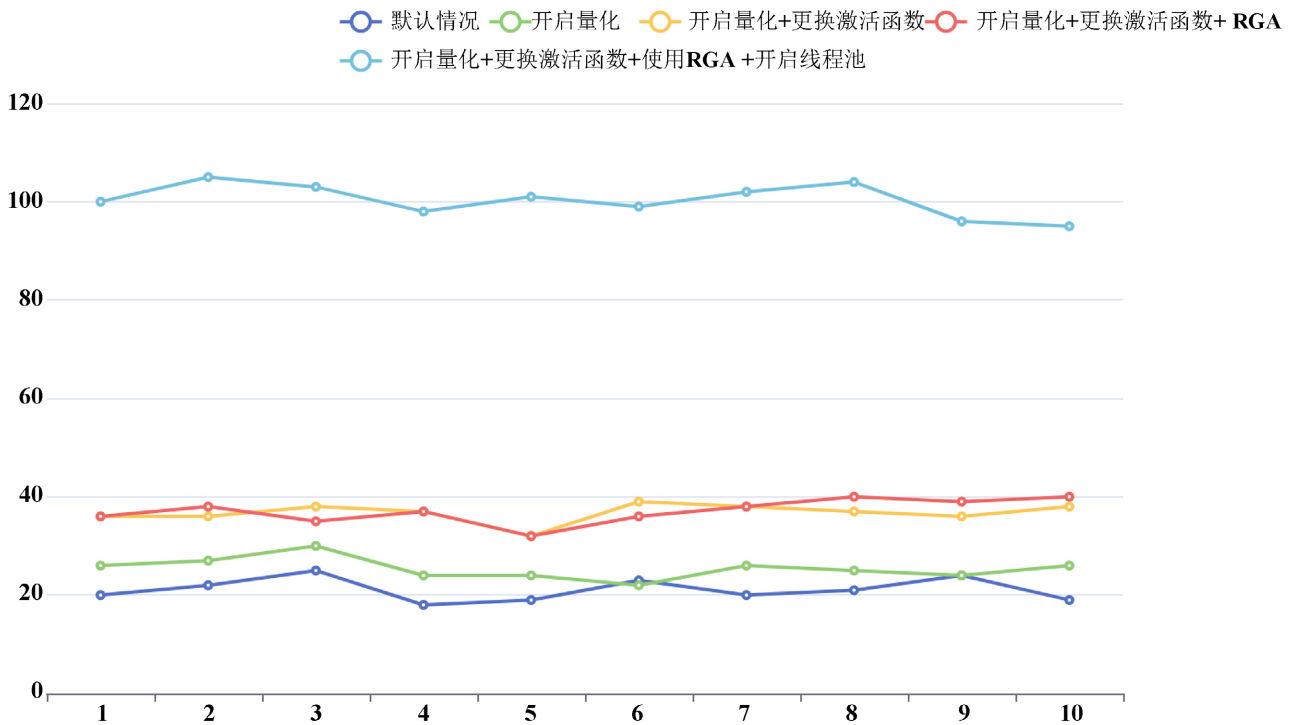


Figure 10. Comparison after enabling acceleration function

图 10. 开启加速功能后的对比

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

人体检测的 P 曲线、R 曲线以及 PR 曲线如图 9(a)所示。从中可以看出, 在 IoU 为 0.5 时, 识别“跌倒”的平均精确度均值能达到 95.4%。火焰烟雾识别的 PR 曲线如图 9(b)所示。识别“火焰”的平均精确度均值能达到 58.1%; 识别“烟雾”的平均精确度均值达到 26.3%。

除了对准确率有要求, 还需要达到较高的检测速度, 对各个加速方式后的帧率进行统计, 结果见图 10。开启 RK3588 内置的 RGA 硬件加速后, 性能并不会太大的提升, 主要原因是输入图像较小, 因此使用软件算法实现和使用硬件实现在效果上差别不会太大, 而当开启多线程功能后, 检测帧率能够得到显著地提升。

8. 结束语

本文结合嵌入式系统和人工智能技术, 实现了一种基于嵌入式系统的老人异常行为智能图像监控系统。相比于传统的线上监控和视频录制, 本系统引入深度学习技术实现了智能监控。系统采用 YOLOv5 实现老人的跌倒检测和火灾识别检测, 并完成该模型在 RK3588 上的移植。经测试, YOLOv5 在识别老人跌倒的准确率可以达到 95.4%。

基金项目

本文得到江西省教育厅科学研究项目(GJJ20212003)的资助。

参考文献

- [1] 浙江省经济信息中心. 热词里的中国活力 | 这个市场有多大? 数据带你看看老年人“新消费” [EB/OL]. https://zjic.zj.gov.cn/zkfw/sjxfw/202402/t20240219_21684944.shtml, 2024-02-22.
- [2] 国家卫生健康委疾病预防控制局. 预防老年人跌倒健康教育核心信息[EB/OL]. https://www.chinacdc.cn/jkzt/mxfcrjbhsh/jswj/202110/t20211012_250181.html, 2021-12-10.
- [3] 中国消防. 近 10 年全国居住场所火灾造成 11634 人遇难[EB/OL]. <https://www.119.gov.cn/gk/sjtj/2022/27328.shtml>, 2022-02-18.
- [4] 玄武纪. 安防视频监控技术的发展过程[EB/OL]. <https://zhuanlan.zhihu.com/p/139642793>, 2020-05-20.
- [5] 赵珍珍, 董彦如, 曹慧, 等. 老年人跌倒检测算法的研究现状[J]. 计算机工程与应用, 2022, 58(5): 50-65.
- [6] Sun, J., Wang, Z., Pei, B., et al. (2016) Fall Detection Using Plantar Inclinometer Sensor. 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), Beijing, 10-14 August 2015, 1692-1697.
- [7] Yuan, L., Nan, W., Lv, C., et al. (2015) Human Body Fall Detection Based on the Kinect Sensor. 2015 8th International Congress on Image and Signal Processing (CISP), Shenyang, 14-16 October 2015, 367-371.
- [8] Dalal, N., Triggs, B. (2005) Histograms of Oriented Gradients for Human Detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, 20-25 June 2005, 886-893.
- [9] Rublee, E., Rabaud, V., Konolige, K., et al. (2011) ORB: An Efficient Alternative to SIFT or SURF. 2011 International Conference on Computer Vision, Barcelona, 6-13 November 2011, 6-13. <https://doi.org/10.1109/ICCV.2011.6126544>
- [10] Wang, X., Han, T.X., Yan, S. (2009) An HOG-LBP Human Detector with Partial Occlusion Handling. 2009 IEEE 12th International Conference on Computer Vision, Kyoto, 29 September-2 October 2009, 32-39. <https://doi.org/10.1109/ICCV.2009.5459207>
- [11] Islam, M.M., Tayan, O., Islam, M.R., et al. (2020) Deep Learning Based Systems Developed for Fall Detection: A Review. IEEE Access, 8, 166117-166137. <https://doi.org/10.1109/ACCESS.2020.3021943>
- [12] 祝玉华, 司艺艺, 李智慧. 基于深度学习的烟雾与火灾检测算法综述[J]. 计算机工程与应用, 2022, 58(23): 1-11.