

文章编号: 2095-2163(2020)04-0253-05

中图分类号: TP391.4

文献标志码: A

复杂背景下目标识别算法分析与改进

王伟¹, 何姣¹, 石强²

(1 贵州电子科技职业学院, 电气工程系, 贵阳 550025; 2 中国核动力研究设计院 第一研究所, 四川 乐山 614100)

摘要: 为解决复杂背景下目标不确定性和环境复杂性带来的算法识别速度慢、检测效率低等问题, 系统采用多线程编程技术, 在嵌入式 Samsung SC6410 和阿里云等平台, 引入开源机器视觉库 OpenCV、以及 DarkNet、keras 第三方深度学习框架, 重点研究了现有的主流目标检测算法, 针对现有算法存在的优势和不足进行剖析, 提出了一种基于密集卷积神经网络目标检测算法改进思路, 即优先采用帧间差分法对样本进行分割, 再调用改进的 YOLO v3 算法实现对目标的分类和检测。实验结果表明, 改进后的算法不仅能够大幅度提升检测速度, 而且对复杂背景下动态目标检测具有优良的效果。

关键词: 复杂背景; DarkNet; 密集卷积网络; YOLO v3

Analysis and improvement of target recognition algorithm in complex background

WANG Wei, HE Jiao, SHI Qiang

(1 Guizhou Electronic Technology Vocational College, Department of electrical engineering, Guiyang, Guizhou, 550025, China;
2 The first Institute of China Nuclear Power Research and Design Institute, Leshan, Sichuan 614100, China)

[Abstract] In order to solve the problems of slow recognition speed and low detection efficiency caused by the uncertainty of target and the complexity of environment in the complex background, this paper introduces opencv, Darknet, keras and other third-party deep learning frameworks in embedded Samsung sc6410 and alicloud platforms by using multithreaded programming technology, focuses on the existing mainstream target detection algorithms, analyzes their advantages and disadvantages, and proposes an improved algorithm based on dense convolution network, That is to say, the priority is to use the frame difference method to segment the samples, and then call the improved Yolo V3 algorithm to achieve the target classification and detection. The experimental results show that the improved algorithm can not only greatly improve the detection speed, but also has excellent results for dynamic target detection in complex background.

[Key words] complex background; Darknet; DenseNet; Yolo V3

0 引言

随着物联网的发展, 大数据、云计算、人工智能等新兴的科学技术开始走进生活, 彻底颠覆了几千年来固有的生产方式和思维习惯, 尤其是在安防行业, 在新时代科学技术日新月异的浪潮下, 传统安防系统发生了翻天覆地的变化。比如智能家居系统让远在异地的人们随时随地掌控家中的情况; 智慧农业系统可以结合天气的变化自适应调节环境参数; 智能交通系统运用摄像机, 结合先进的图像处理算法, 成功的在道路、车辆和驾驶员三者之间建立了一条监控屏障等。虽然安防系统在各行各业发展相对成熟, 但由于存在网络延迟、软硬件成本过高、图像识别率低和安全性差等问题, 导致了安防系统的普及受到限制。据相关数据统计, 我国的民用智能安

防产品普及率只有 10% 左右^[1]。如何采取最优的办法, 解决上述问题成为打开安防市场的关键。

目前, 大多数图像算法多着眼于提高图像处理速度和识别精度, 针对复杂背景、目标遮挡、图像模糊等问题研究甚少。本文就以复杂背景下目标识别作为切入点, 借助阿里云和嵌入式 Linux 平台, 以及 MATLAB 等优秀的仿真软件, 剖析多种经典目标识别算法, 通过比较和分析, 最终得出 YOLO v3 算法具有检测速度快、精度高等优点, 并且在其基础上, 改进算法处理结构^[2], 使得 YOLO v3 算法在捕捉和处理异常目标上更具优势。

1 系统总体架构

在设计整体框架之前, 先后对当地企业食安云公司和润可信息科技有限公司进行了实地调研, 调

作者简介: 王伟(1991-), 男, 硕士, 助理讲师, 主要研究方向: 计算机图像处理技术、物联网、计算机控制技术; 何姣(1991-), 女, 硕士, 助理讲师, 主要研究方向: 嵌入式系统设计与研究、物联网; 石强(1994-), 男, 硕士, 助理工程师, 主要研究方向: 图像处理、故障诊断、目标检测与追踪。

通讯作者: 何姣 Email: 328079904@qq.com

收稿日期: 2020-01-11

研的主要内容包括方案可行性分析、技术开发要点和实用价值等,并且专访了相关专业的教授,请他们提出宝贵意见,结合目前市场短板,最终设计了基于云平台的目标识别系统,整体框架如下图1所示。

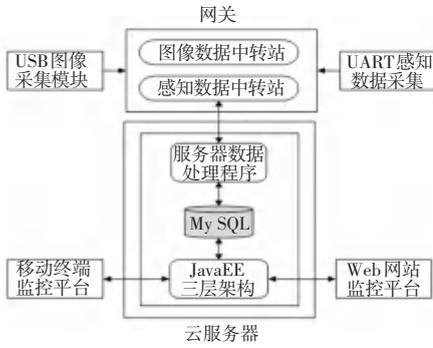


图1 系统总体框架

Fig. 1 Overall framework of the system

主要设计思路:通过网关,将原始的图像数据上传到阿里云服务器,然后借助部署的服务器处理程序进行图像分析和数据保存,移动终端、web平台负责与JavaEE三层架构建立高速信息交互通道,用于异常数据检索和查询。

2 系统结构设计

为解决复杂背景下目标识别过程中遇到的关键问题,首先要研究传统算法在目标识别中存在的优势和劣势,进一步改进已有的算法结构,提升程序运行效率,本文分别从硬件结构、软件结构和界面设计三方面进行功能设计。

2.1 硬件结构设计

网关是系统硬件结构的核心,本文采用Samsung S3C6410作为主控。针对所需的模块,在原有的功能板基础上进行了接口扩展。为了充分利用CPU多核资源,提高数据的吞吐率,满足图像采集和网络上传过程中数据并行处理的需求,本文提出了一种基于嵌入式Linux系统下的网关设计方案,即在硬件系统上移植Linux操作系统^[3]。经过论证与检验,该方案提供的多线程、图像化编程技术可以稳定的为云服务器输送大量的数据资源,为数据分析提供了保障。硬件结构模块化设计如图2所示。

2.2 软件结构设计

复杂背景下目标识别系统软件架构是一种基于B/S架构的数据管理方案,如图3所示。该方案由客户端、云平台服务器和手持移动终端三部分构成。在客户端用户可以实现信息注册、网关登录、数据信息采集和网络传送等功能。云服务器端主要负

责图像算法分析,通过搭建不同的算法,在多种复合场景下,从识别率、处理速度、模糊等角度进行性能比较,结合JavaEE三层架构建立信息反馈通道,查找不足并适当改进程序结构,最终得到最优算法部署方案。移动终端和web平台依靠HTTP请求,建立与云服务器信息交互通道,针对MySQL数据库收集的异常信息,使用户可以在大众平台上实时查阅^[4]。

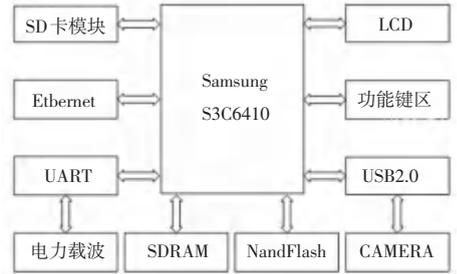


图2 硬件结构设计

Fig. 2 Hardware structure design

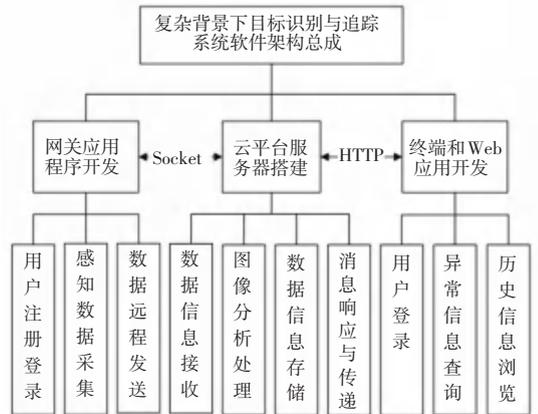


图3 软件架构图

Fig. 3 Software architecture

3 算法分析与优化

在复杂背景下,处理图像的步骤一般是将整张图像送到准备好的训练集进行特征分析和提取,再借助分类器归类,由于检测目标往往只占实际图像的很小一部分,如果采用这种方式作为检测手段,势必造成大量计算机资源浪费,图像检测速度也随之降低。为了提高检测效率,本文提出了一种“先分割,再检测”的处理方案。

3.1 复杂背景下目标分割技术

目前,能够实现目标与背景分离的算法很多,比如:背景差分法、光流场法、ViBe算法等。背景差分法鲁棒性差,自然环境下场景参数的变换对检测结果有很大影响;光流场法算法复杂度高,执行效率低;ViBe算法对静止物体或者局部移动的物体,容

易产生误判和孤立噪声, 几种分割算法的比较如表 1 所示。相比之下, 帧间差分法在抗噪声、运算速度、算法复杂度上具有很大优势。

表 1 几种背景分割算法比较

Tab. 1 Comparison of several background segmentation algorithms

算法	抗噪声能力	执行速度	复杂度
背景差分法	鲁棒性较差	快	简单
光流场法	很差	耗时	复杂
ViBe 算法	一般	较快	较复杂
帧间差分法	一般	快	简单

3.1.1 算法原理

帧间差分法原理上又分为二帧差分法和三帧差分法。由于二帧差分法在目标移动过快的情况下, 经常出现重影和噪声, 所以本文采用三帧差分法对目标进行分割, 原理如图 4 所示。

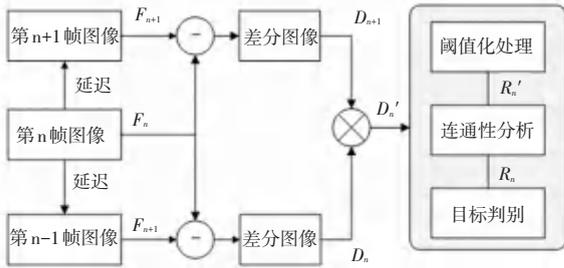


图 4 三帧差分法原理框图

Fig. 4 Principle block diagram of three frame difference method

随机选取摄像头实时采集的三张连续的图像, 分别记为: $F_n(x, y)$ 、 $F_{n-1}(x, y)$ 和 $F_{n+1}(x, y)$, 将原始图像灰度化处理, 得到灰度图像 $f_n(x, y)$ 、 $f_{n-1}(x, y)$ 和 $f_{n+1}(x, y)$, 再对三帧灰度图做差分处理, 运算公式如下:

$$\begin{cases} D_n(x, y) = |f_n(x, y) - f_{n-1}(x, y)|, \\ D_{n+1}(x, y) = |f_{n+1}(x, y) - f_n(x, y)|. \end{cases} \quad (1)$$

根据得到的差分图像, 选取合适的灰度阈值 T , 再进行阈值化处理, 为了能够让阈值自适应环境光照的变化, 特提出新的约束条件 T' , T' 在 T 的基础上, 考虑了图像整体光线变化参数, 当光线变化较弱时, $T' = T$, 反之光线突变时, $T' \gg T$, 公式(2)如下:

$$\begin{cases} R_n(x, y) = \begin{cases} B(x, y), & \text{IF } |D_n(x, y)| < T', \\ M(x, y), & \text{IF } |D_n(x, y)| \geq T'; \end{cases} \\ R_{n+1}(x, y) = \begin{cases} B(x, y), & \text{IF } |D_{n+1}(x, y)| < T', \\ M(x, y), & \text{IF } |D_{n+1}(x, y)| \geq T'; \end{cases} \\ T' = T + \frac{1}{N} \sum_1^N |D_n(x, y)|. \end{cases} \quad (2)$$

其中 $B(x, y)$ 为阈值化后的背景像素点, $M(x, y)$ 为阈值化后的目标变化像素点, $R_n(x, y)$ 是阈值化图像。使用二值化图像 $R_n(x, y)$ 和 $R_{n+1}(x, y)$ 进行与运算和形态学处理, 得到 $P_n(x, y)$, 公式(3)如下所示。记录 $P_n(x, y)$ 的像素点的突变位置, 返回原图像 $F_n(x, y)$, 再按照定位点对图像分割和边沿扩充, 最终得到分割后的目标图像。

$$P_n(x, y) = \begin{cases} B(x, y), & R_n(x, y) \cap R_{n+1}(x, y) = 1, \\ M(x, y), & R_n(x, y) \cap R_{n+1}(x, y) = 0. \end{cases} \quad (3)$$

3.1.2 算法实现

该算法的实现依托于开源机器视觉库 OpenCV 3.4.8, 结合 QT 界面, 最终在 UBUNTU 环境下实现了基于三帧差分法背景目标分离。通过对比发现, 两张连续的图片差分操作以后, 运动物体的轮廓出现了重影和噪声, 随着物体移动速度的递增, 重影的距离也在变化。为了消除重影, 本文在微调阈值的基础上, 将得到的两张差分后的二值化图再次与运算, 最大限度消除重影。最后根据处理结果, 采用连通性分析函数查找物体边界, 并对边界进行与运算 ($\text{maxrect} = \text{maxrect} | \text{rect}$) 保留最大边界, 适当扩充框选范围后, 分割目标如图 5 所示。



(a) 二值化图 (b) 与操作图

图 5 帧间差分法目标分割图

Fig. 5 Target segmentation by frame difference method

3.2 目标识别

复杂场景下的目标识别是安防监控系统部署的一项重要内容, 检测的准确性和运算速度直接影响着系统的其他环节。如何解决复杂场景、遮挡、光照参数变化导致的系统误判, 以及不同距离下目标大小变化问题, 对于安防监控行业的发展具有深远的意义^[5]。本文从深度学习模型出发, 研究了多种算法, 通过性能对比, 最终得出 YOLO V3 在目标检测领域具有卓越的优势, 如表 2 所示。

3.2.1 YOLO V3 算法

YOLO V3 是 Joseph Redmon 等在 YOLO V1 和 YOLO V2 版本上开发的基于回归的目标识别算

法^[6]。它沿用 Darknet 框架,并在此基础上将网络层升级,由原来 Imagenet 训练的 53 层网络,增加到 106 层底层框架,主要针对检测任务堆叠了 53 层网络^[7]。该算法的基本思想借鉴了残差网络 Resnet 的做法,在 Darknet-53 的某些卷积网络中设置了 shortcut connections,通过采样卷积方向变换和特征图融合,最终得到 3 种不同尺度特征图(13 * 13, 26 * 26, 52 * 52),再次采用 9 种尺度的先验框对特征图进行检测,13 * 13 特征图使用较大的先验框,用于检测大 object;26 * 26 特征图使用中等先验框,检测中等 object;52 * 52 特征图使用较小先验框,用于检测小 object^[8],结果表明该算法尤其适用视距较远的小目标。

表 2 基于深度学习部分模型性能对比

Tab. 2 Performance comparison of some models based on deep learning

检测器	训练集	Map/%	AP/%	FPS/s
RCNN	Voc2007	52.3	53.2	—
FasterRCNN	Voc2007	73.2	79.6	7
YOLO	Voc2007	63.4	63.5	48
SSD300	Voc2007	74.3	65.0	46
YOLO V3	coco	76.0	73.0	43

3.2.2 算法改进思想

在众多目标检测的算法中,并不是层数越深的卷积网络提取特征能力越强,网络的加深则会导致过拟合、梯度弥散和梯度爆炸等问题的产生^[9],这对于检测模型训练会产生一定的影响。

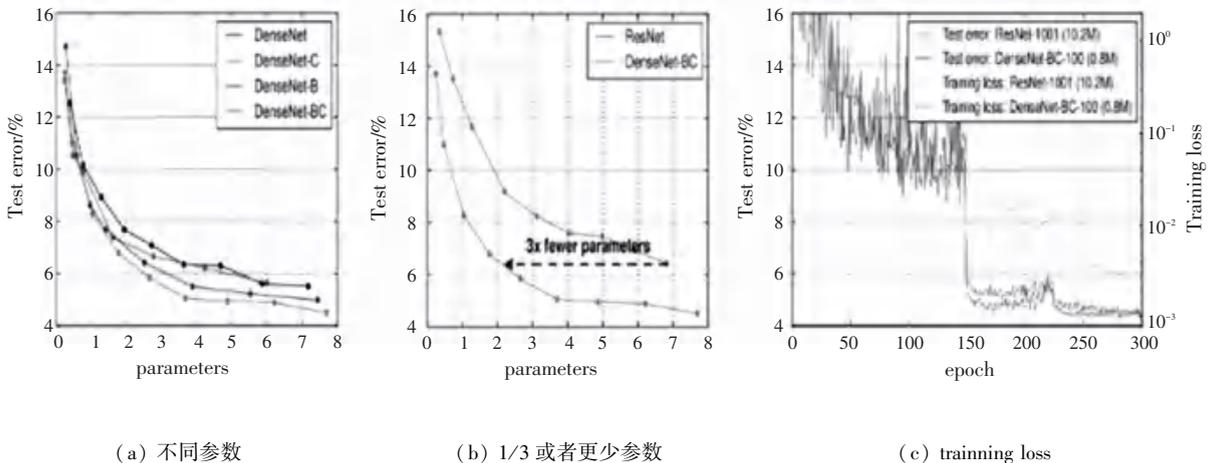


图 7 参数和计算复杂度对比结果

Fig. 7 Comparison results of parameters and calculation complexity

3.2.3 实验分析

YOLO V3 算法对计算机硬件性能要求极高,由

YOLO V3 采用 Darknet-53 框架提取特征图,内含 5 组残差网络(Resnet),层与层之间使用顺序连接的方式。2017 年稠密卷积神经网络 DenseNet,打破了 Resnet 的顺序连接方式,通过前馈神经网络建立跨层连接,解决了随着检测网络深度的增加而产生的退化问题,残差模块和密集卷积网络结构对比如图 6 所示。

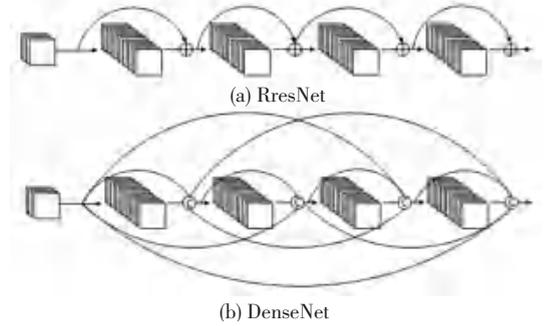


图 6 残差模块和密集卷积网络结构对比

Fig. 6 Structure comparison between residual module and dense convolution network

比较发现,密集卷积网络的引入,减少了实验参数,提升整体运算效率,如图 7 所示。图 7(a)是稠密卷积神经网络变种在不同参数下的 test error 统计曲线;图 7(b)的曲线图则得出稠密卷积神经网络只需要 1/3 或者更少参数,所得的 test error 统计几乎和残差网络相同;图 7(c)显示 100 层密集卷积网络的 training loss,以及 test error 曲线几乎等价于 1001 层的残差网络^[10]。

于条件有限,这里主要采用 Ubuntu 64 位系统、开源机器视觉库、Darknet 和 keras 深度学习框架等进行

模拟测试^[11],实验结果表明:改进后的 YOLO V3 算法不仅对运动目标检测性能优越,对于多目标种类的识别、目标遮挡和不同视距下目标大小变化的识别具有较好的检测效果,如图 8 所示。



(a) 编译结果 (b) 识别效果

图 8 YOLO V3 算法编译识别结果

Fig. 8 Results of Yolo V3 algorithm compilation and recognition

3.3 程序结构改进与结果分析

通过对上述算法分析可得:帧间差分法能够较为简单迅速的检测运动目标,但对于复杂背景环境参数较为敏感,容易造成误判;改进后的 YOLO v3 算法综合检测性能优越,唯一缺点是程序结构复杂。结合两种算法的优势与劣势,本文提出一种基于目标分割检测的处理思路,如下图 9 所示。第一步,将标定后的摄像头,放在检测目标出没的位置,实时采集图像以待后台处理;第二步,对图像数据抽帧检测,被抽中的图像利用帧间差分法进行过滤,如果检测到目标存在,则调用 src 函数裁剪目标区域;第三步,针对裁剪的图像再次调用 YOLO v3 算法进行目标分类检测,根据扫描的结果进行归类,从而判定物体的属性。

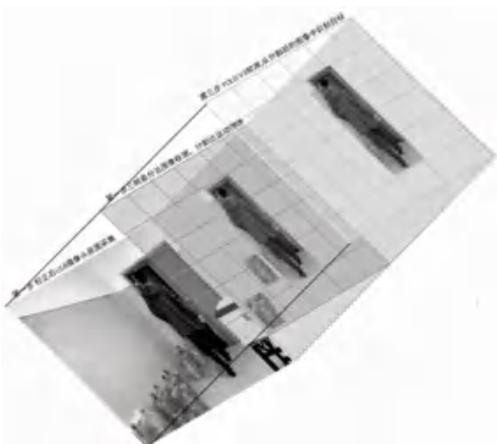


图 9 目标分割检测处理流程

Fig. 9 Target segmentation detection processing flow

为了验证改进前后算法性能变化,本文针对不同分辨率的图像进行了大量的实验,最终得出图 10

所示曲线,通过对曲线分析得出如下结论:

- 1) 图像越大,算法执行时间越长。
- 2) 改进算法结构后,节省了无效背景分析所占用的时间,程序整体的运行效率大幅度提升。

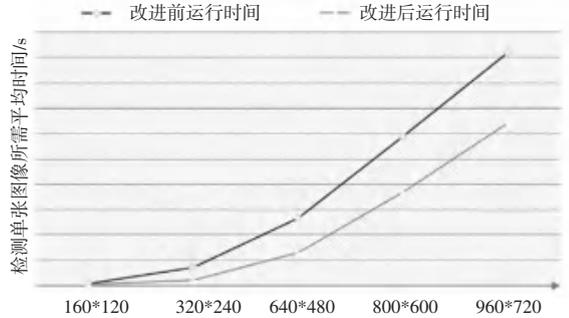


图 10 改进前后运行时间对比

Fig. 10 Comparison of operation time before and after improvement

4 结束语

本文对复杂背景下目标识别系统进行了详细研究和设计,说明了组成该系统的硬件架构和软件功能,重点分析了目标检测算法实现的过程,以及算法改进的思路。在保证精度的前提下,通过优化算法结构,降低数据样本的复杂度,为提升图像处理速度提供了一种可靠的思路,希望在未来的研究过程中,能够创造更大的价值。

参考文献

- [1] 田铭铭. 云计算在安防视频监控领域的应用[J]. 计算机产品与流通, 2019(8):104.
- [2] 毛宇丰,陈朝晖,鲍淑红,等. 基于综合监控系统软件平台的综合安防系统研究[J]. 城市轨道交通研究, 2018, 21(S2):1-8.
- [3] 穆健,于泽涛,刘浩,等. 智慧云平台系统在安全管理中的应用[J]. 山西建筑, 2019, 45(20):189-191.
- [4] 毛捷. 基于 S3C6410 的嵌入式远程视频监控系统设计研究[J]. 中国设备工程, 2018(10):103-105.
- [5] 蔡颖,陈思杰,张珍,等. 基于机器视觉与多旋翼无人飞行器的火场救援引导系统设计[J]. 自动化应用, 2019(06):5-6+12.
- [6] 张迪,樊绍胜. 基于 YOLO V3 的输电线路故障检测方法[J]. 自动化技术与应用, 2019, 38(7):125-129.
- [7] 刘生智,李春蓉,刘同金,等. 基于 YOLO V3 模型的奶牛目标检测[J]. 塔里木大学学报, 2019, 31(2):85-90.
- [8] 段慧芳,刘娟. 密集卷积神经网络和辅助特征相结合的乳腺组织病理图像有丝分裂检测方法[J]. 武汉大学学报(理学版), 2019, 65(5):434-440.
- [9] 孙颖. 基于 3D 残差密集网络的视频烟雾检测研究[D]. 东北师范大学, 2019.
- [10] 于丽,刘坤,于晨焱. 基于密集卷积神经网络的遥感飞机识别[J]. 计算机工程与应用, 2018, 54(19):179-185+203.
- [11] 郭进祥,刘立波,徐峰,等. 基于 YOLO v3 的机场场面飞机检测方法[J]. 激光与光电子学进展, 2019, 56(19):111-119.