

# 基于红外智能视频检测的自适应加速检测算法

祁海军,刘智嘉,赵金博,夏寅辉,高旭辉  
(北京波谱华光科技有限公司,北京 100015)

**摘要:**随着人工智能技术的快速发展,智能车辆辅助驾驶系统已经逐渐走进人们的生活。基于可见光成像的智能车辆辅助驾驶系统虽然日趋成熟,但是其受限于可见光成像条件的制约,在黑暗、雨雾等条件下,使用效果可能大打折扣甚至无法使用。红外智能车辆辅助驾驶系统因成像依靠红外成像设备,使其成为解决这一问题的重要手段而受到广泛关注。但是红外智能车辅系统在一些特殊情况下,如车辆突然提速、应用场景发生较大变化等,其检测及成像实时性可能会受到影响进而导致显示画面卡顿。上述情况的出现会严重影响驾驶员判断甚至行车安全。针对上述情况本文提出了一种自适应隔帧检测算法可以在红外智能车辅系统应用过程中针对不同情况出现的检测速度波动,自主调整检测间隔帧数,以保证检测和成像的实时性。实验证明,该算法对于视频检测速度提升有效,具有一定的实用性。

**关键词:**红外;智能检测;隔帧;自适应

中图分类号:TN219 文献标识码:A DOI:10.3969/j.issn.1001-5078.2020.10.010

## An adaptive accelerated detection algorithm based on infrared intelligent video detection

QI Hai-jun, LIU Zhi-jia, ZHAO Jin-bo, XIA Yin-hui, GAO Xu-hui  
(Beijing Bop Opto-Electronics Technology Co., Ltd., Beijing 100015, China)

**Abstract:** With the rapid development of artificial intelligence technology, intelligent vehicle assistant driving system has gradually entered people's life. Although the intelligent vehicle assistant driving system based on visible light imaging is becoming more and more mature, it is limited by the visible light imaging conditions. In the dark, rain and fog circumstances, the effect of utilization may be greatly reduced or even unable to use. Infrared intelligent vehicle assistant driving system is widely concerned, because it relies on infrared imaging equipment, which makes it an important means to solve this problem. However, in some special cases, such as sudden acceleration of the vehicle and great changes in the application scene, the real-time performance of the detection and imaging of the infrared intelligent vehicle auxiliary system may be affected, which may cause the display picture to jam. The above situation will seriously affect the driver's judgment and even driving safety. In view of the above situation, this paper proposes an adaptive frame detection algorithm, which can adjust the number of detection intervals to ensure the real-time performance of detection and imaging. Experimental results show that the algorithm is effective and practical for improving the speed of video detection.

**Keywords:** infrared; intelligent detection; frame skip; adaptive

## 1 引言

红外热像仪可在黑暗、雨雾以及烟云条件下对周围环境及探测目标成像,所以热像仪具有全天候及全时段工作的特点而广泛应用于国防及安全领域。由于早期的红外成像设备较为昂贵,使用红外成像设备的车辆辅助驾驶系统仅针对军用车辆使用而开发,以便在烟尘、无照明条件的野外以及各种可视条件不佳的情况下为军用车辆提供驾驶辅助信息。随着半导体技术的发展及红外焦平面探测器成本的不断下降,红外热像仪的应用已经从军用领域逐渐扩展至民用领域。

早期的红外图像目标检测算法主要基于传统的图像处理方法,如 HOG<sup>[1]</sup>,LBP<sup>[2-3]</sup>以及 SIFT<sup>[4-5]</sup>等方法。传统检测方法的基本原理均采用人工设计的样本轮廓特征提取方法,这样的做法通常不可避免如下缺点:一方面对设计人员要求较高,需要设计人员具有相当程度的专业知识和设计经验;另一方面由于仅使用简单的轮廓作为判断特征,识别准确率很难随算法改进而不断提高。基于传统图像算法的红外车辆辅助驾驶系统也没有消除上述缺点。

近些年随着人工智能技术的发展,智能车辆辅助驾驶系统已经越来越成熟,随着智能检测算法在很多领域取得接近甚至超过人类对相同目标的识别准确率,一些厂商已经推出具备自动驾驶功能的车辆。不过,当前的智能车辅几乎都是基于可见光成像设备,工作环境受到可见光成像条件的制约。将红外成像和智能识别的优势相结合就成为解决上述一系列问题的重要方法之一。作为智能检测算法种非常重要的组成部分——深度学习检测算法利用大量卷积运算对图像中样本特征进行全面提取并自主抽象高级别特征,具有传统识别算法无法比拟的检测准确率,常用的方法如 R-CNN<sup>[6]</sup>、Fast R-CNN<sup>[7]</sup>、Faster R-CNN<sup>[8]</sup>、SSD<sup>[9-10]</sup>及 YOLO<sup>[11-12]</sup>等。由于深度学习检测算法对硬件的算力要求较高,所以智能检测算法虽然在检测准确率的表现上远超传统检测算法,但其检测速度却相对缓慢得多,如何在保证检测准确率的前提下提升检测速度一直相关研究的重点和难点。特别是红外智能车辅系统关系到目视条件较差的行车安全问题,所以应用中的实时性需要绝对保证。

红外智能车辆辅助系统硬件部分均为可部署于车辆的移动终端,通常使用嵌入式硬件平台或 FP-

GA。尽管近些年移动终端在小型化、低功耗、高性能方面取得了不小的进步,但相对于使用专业计算卡的计算机性能仍相差甚远,所以大部分嵌入式硬件平台即使可以部署深度学习算法,也要在检测准确率和检测速度之间进行取舍。目前基于深度学习的红外目标检测方法的终端应用优化主要利用对卷积网络中卷积层数量进行压缩以减少卷积运算,提高整体运行速度。但是这样的结果会导致检测准确率下降和目标漏检及误检。为了尽可能不降低算法模型的检测准确率,优化策略往往选择在满足实时检测的最低条件下尽可能少的降低检测准确率。但是,在实际应用中由于车速突然改变或者检测背景发生较大变化时,可能会出现检测速度突然明显下降的情况,导致红外智能车辆辅助驾驶系统的显示画面出现卡顿。上述情况的出现不仅影响车辆辅助驾驶系统对成像视野内目标的检测,也会严重影响对外界环境成像,妨碍驾驶者判断,甚至导致恶性交通事故发生。

为了解决上述问题,本文提出了一种基于红外智能目标检测算法的自适应隔帧检测算法,本算法可以根据智能检测算法对红外视频的即时检测速度调整检测算法对图像序列的检测模式,使红外智能检测算法具备在实际应用中自动调整整体检测速度的功能以确保检测及成像的实时性。

## 2 智能算法视频检测原理分析

红外智能识别算法检测视频的本质是先将视频分解为逐帧图像,然后对分解图像依照顺序进行逐帧检测,再将带有检测信息的图像合成视频输出至显示设备。每帧图像的检测信息包括图片内所有被检测出目标的检测框坐标信息和目标的类别信息,并根据当前图像检测速度计算出即时视频检测速度,即每秒检测帧数(FPS):

$$FPS = \frac{1}{t} \quad (1)$$

式中, $t$ 为检测模型检测单张图像所用时间,FPS也称即时检测帧率,由智能识别算法实时获得并输出。

深度学习算法网络模型逐帧检测流程如下(如图1所示):

- (1) 视频信号输入;
- (2) 视频被拆分为逐帧图像;
- (3) 逐帧图像送入智能检测模型;
- (4) 对图像进行检测并对识别目标进行标记;

- (5) 将标记后图像序列合成视频;  
 (6) 合成视频输出至显示设备。

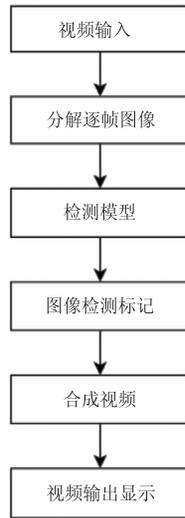


图1 逐帧智能检测流程图

Fig. 1 Flow chart of frame-by-frame intelligent detection

通过上述检测流程,输入视频经过智能检测算法模型检测标记后再输出显示,为驾驶员提供驾驶辅助信息。当每帧图像检测时间小于 40 ms 时,即 FPS 不少于 25 时实现实时检测。若使用中 FPS 小于 25,则显示画面会出现卡顿,驾驶者不能实时获得检测及成像信息。

### 3 自适应隔帧检测算法

自适应隔帧检测算法原理说明:

为了方便叙述定义变量如下:IB 为存储检测模型对被检图像检测过程中生成的最新目标检测信息及检测帧率(FPS)的临时变量;JF 为隔帧总数,表示检测模式为每隔 JF 帧图像检测一帧图像,缺省值为 0;RJF 为当前隔帧数,若  $RJF > 0$ ,表示当前图像为被隔图像,若  $RJF = 0$ ,则当前图像需进行目标检测,RJF 缺省值为 0;LT 为阈值下限,预设值为 25;HT 为阈值上限,预设值为 32,阈值可根据具体使用场景设置;JFPS 为整体输出帧率即每秒输出图像数量,由于检测过程输出的即时速度不包括隔帧帧数,所以应重新计算全部图像的输出帧率,计算公式为:

$$JFPS = (JF + 1) \times FPS \quad (2)$$

当逐帧检测时,JF 为 0,JFPS 即为 FPS。若隔帧总数为 1,则速度加快一倍。

视频信号输入后被拆分为逐帧图像,图像依次进入检测系统前通过 RJF 数值对当前图像是否需要检测进行判断。由于参数设置,首帧图像默认为

被检图像,检测所获得的目标检测信息及速度信息保存于 IB 中并根据图像检测速度对 JF 进行操作:若检测速度高于实时检测速度,则保持目前检测状态不变;若检测速度低于实时检测速度,则 JF 值增加并将更新后的 JF 值赋予 RJF,从检测下一帧图像开始增加隔帧数量,以提升检测速度,直至满足实时检测及显示的速度需求。被隔图像跳过检测流程直接将 IB 中的检测信息复制至图像并显示检测框及类别信息,并根据 IB 中保存的速度信息计算根据公式(2)计算即使速度 JFPS,并根据 JFPS 进行判断以选择后续隔帧检测模式。

当 JFPS 超过 HT 则说明检测条件好转,可减少隔帧数量,则 JF 减少并赋值于 RJF,立即增加检测图像数量,随后根据 JFPS 继续判断是否需要继续减少 JF,直至恢复逐帧检测。

由于被隔图像不进行目标检测,图像中缺少检测信息,所以将 IB 中的检测信息直接复制至被隔图像,以避免因合成视频中仅被检图像具有检测框而出现显示画面闪烁的问题,方便使用者观察。并利用保存在 IB 中的检测速度信息计算不同隔帧模式下的整体即时检测速度。结合智能检测算法的整体流程如图 2。

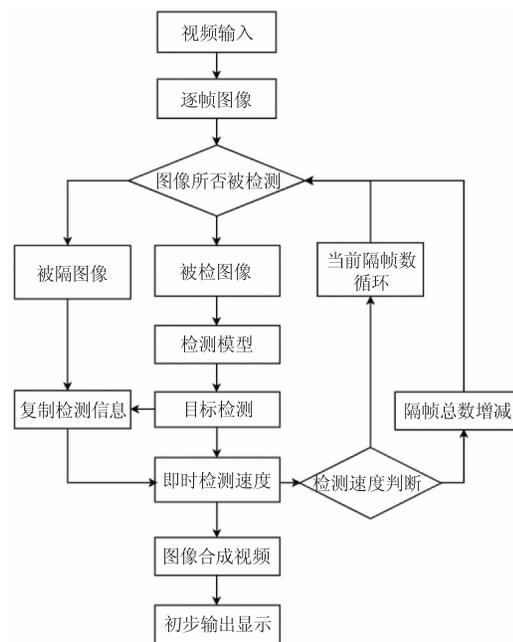


图2 自适应隔帧智能检测流程图

Fig. 2 Adaptive frame skip intelligent detection detection flowchart

整体检测流程如下:

- (1) 采集视频或者已录制视频作为输入读入至识别网络;

(2)待检测视频拆分为顺序逐帧图像;进行图像是否被检测判断:如果  $RJF > 0$ ;复制检测信息至被隔图像,并将检测框标注至被隔图像相应位置;转至流程(6);否则:进行下一步;

(3)被检图像送至检测模型;

(4)检测模型对送入图像进行检测并标记检测信息,并将检测信息及当前图像目标检测的速度信息更新至变量  $IB$ ;

(6)利用公式(2)计算当前整体即时检测速度;

(7)检测速度判断:如果  $JFPS < LT$ : $JF = JF + 1$ ;  $RJF = JF$ ;如果  $JFPS > 30$ :如果  $JF > 0$ , $JF = JF - 1$ ;  $RJF = JF$ ; 否则:  $RJF = RJF - 1$ ; 如果  $RJF < 0$ :  $RJF = JF$ ;

(8)依照拆分顺序合成视频;

(9)终端显示。

视频的检测过程为上述过程的持续循环,直至视频检测结束。

由于实时检测最低检测速度为  $25 \text{ f/s}$ ,故  $LT$  设定为  $25$ 。根据红外视频的智能检测在实际应用中检测速度会在较小范围内波动的现象,将  $HT$  设定为  $32 \text{ f/s}$ ,该值可根据具体检测算法和检测环境自行设定。通过阈值判断进行增减隔帧数量实现对序列图像检测模式的改变,保证红外智能车辆辅助驾驶系统检测及成像的实时性。

#### 4 实验验证及结果分析

本文对同一智能检测算法模型分别使用逐帧和隔两帧检测模式进行速度对比实验,通过实验结果分析,验证自适应隔帧算法对智能检测算法模型检测速度提升的有效性。为使实验平台更接近实际应用终端硬件性能,本次实验选择 Nvidia Jetson TX2 嵌入式平台,算法模型使用相同的 Darknet 深度学习框架下的 YOLOv3 改进模型。由于使用相同软硬件进行测试,实验结果具有可对比性。

表 1 Nvidia Jetson TX2 主要参数

Tab. 1 Parameters of Nivida Jetson TX2

操作系统	Ubuntu 16.04
GPU	Pascal 256 CUDA 核心
CPU	Denver 2 & ARM A57
视频输出	HDMI 输出:最高可支持 4k 视频
内存	8GB
接口	USB 2.0、USB 3.0、PCIE、SATA 等



图 3 Nvidia Jetson TX2

Fig. 3 Nvidia Jetson TX2

实验对同一视频分别利用逐帧检测和隔两帧检测模式进行检测,结果如图 4 所示。

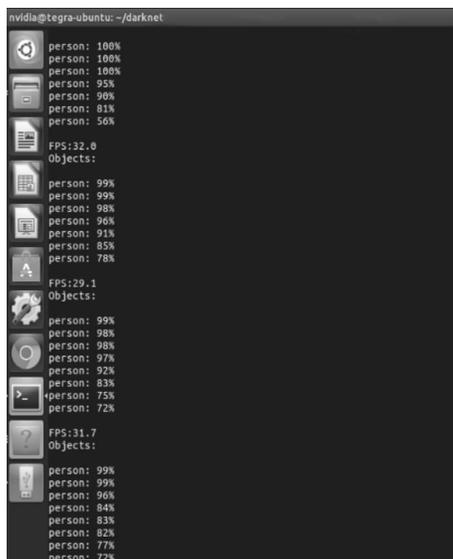
图 4(a)为使用逐帧检测模式检测帧率,最低检测帧率及最高检测帧率分别为  $29.1 \text{ f/s}$  和  $32 \text{ f/s}$ 。图 4(b)所示为使用隔两帧检测模式下整体检测帧率,检测最低帧率和检测最高帧率分别为  $55 \text{ f/s}$  和  $77.2 \text{ f/s}$ 。通过图中所示实验结果对比可以看出使用隔两帧检测模式相对于逐帧检测模式的检测速度提升效果显著。

相同图像使用逐帧与隔两帧检测模式效果对比如图 5 所示,图 5(a)为逐帧检测模式效果图像,图 5(b)所示为隔两帧检测同一图像效果。图像中目标行人从摄像设备左侧向右侧步行通过,距离适中。从图像中可以看到,使用相同模型对同一图像进行检测,但是由于采用不同隔帧检测模式所以检测效果有较明显区别。逐帧检测图像检测框紧贴行人外轮廓,定位准确。隔两帧检测由于检测信息为前两帧图像检测信息直接复制获得,检测框相对检测样本有轻微偏移现象,行人身体有很小部分在检测框外,但被检目标绝大部分仍在检测框包围面积内。

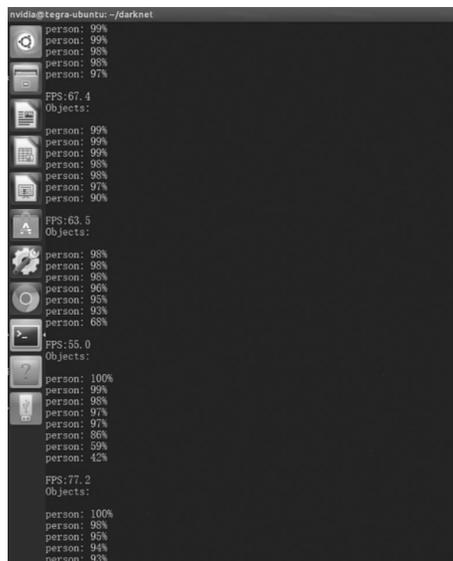
行人径向移动是逐帧和隔两帧模式下检测效果对比如图 6 所示。图中行人以径向远离摄像设备方向步行移动,距离摄像设备距离适中。其中图 6(a)为逐帧检测图像,图像中检测框紧贴行人外轮廓,检测效果良好。图 6(b)为相同图像使用隔两帧检测方式检测效果图像,可以看到检测框的位置和形态

与上图完全一样,这是因为6(b)为6(a)隔1帧后图像,图像中检测框为6(a)图检测信息复制获得,所以行人有部分身体在检测框外,其中以左侧行人更加明显。图6(c)为(b)图像逐帧检测效果,可见检测框紧贴行人,行人身体完全在检测框内,检测效果良好。

测框覆盖,对于驾驶者来说仍然具有一定程度的提示和辅助意义。并且使用自适应隔帧算法的红外智能车辆辅助系统即使在完全不能进行目标检测的极端情况,也可以保证红外智能车辆辅助系统对环境的实时探测成像,这对于驾驶者在可视条件不佳的状况下安全驾驶具有重要意义。



(a)



(b)

图4 逐帧检测和隔两帧检测帧率对比图

Fig. 4 Comparison of frame rate for frame-by-frame detection and two-frame-skipping detection

通过对使用隔帧检测的检测速度进行定量对比实验可以看到隔帧算法对检测提速具有明显效果,证明了算法的有效性。而后通过检测效果实验可以看到,使用了隔帧检测算法后虽然检测框相对于检测目标稍有偏移,但样本的绝大部分面积仍然被检



(a)



(b)

图5 横向移动行人逐帧与隔两帧检测效果对比图

Fig. 5 Comparison of detection results of frame-by-frame and two-frame-skipping detection of pedestrians moving horizontally



(a)



(b)

图6 径向移动行人逐帧与隔两帧检测效果对比图

Fig.6 Comparison of detection effect of frame-by-frame and two-frame-skipping detection for pedestrians moving radially

## 5 结论

本文针对红外智能车辅在应用过程中可能出现的检测速度下降而导致的画面卡顿及不能实时成像的问题,通过对常规智能算法模型检测模式流程的分析,提出了自适应隔帧检测算法。该算法可以在红外智能车辅使用过程中根据当前检测速度对检测策略进行自动调整,自主平衡检测精度和速度的关系,确保红外智能车辅检测及成像的实时性。通过实验及实验结果分析,自适应隔帧算法对检测速度提升的有效性得到验证。根据图像检测效果实验的结果,证明使用隔帧检测模式的检测图像仍具有一定的辅助提示作用,实现了算法提出的预期目的。与此同时,在算法的改进提出的过程中仍有一些不足,如是否可以简单预测被隔图像的检测框信息,使隔帧图像的检测信息更为准确,可以作为下一步改进的方向。

## 参考文献:

- [1] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 05), IEEE, 2005.
- [2] T Ojala, M Pietikäinen, D Harwood. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions[C]//Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), 1994:582-585.
- [3] Ojala T, Pietik Inen M, Harwood D. A comparative study of texture measures with classification based on featured distributions[J]. Pattern Recognition: The Journal of the Pattern Recognition Society, 1996, 29(1):51-59.
- [4] Lowe D G. Local feature view clustering for 3D object recognition[J]. Proc. IEEE Conf. on Computer Vision & Pattern Recognition Maui Hi, 2001, (1):682-688.
- [5] Lowe D G. Object recognition from local Scale-Invariant features[C]// Iccv. IEEE Computer Society, 1999.
- [6] R Girshick, J Donahue, T Darrell, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, Columbus, OH, 2014:580-587.
- [7] R Girshick. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision, 2015:1440-1448.
- [8] S Ren, K He, R Girshick, et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6):1137-1149.
- [9] W Liu, D Anguelov, D Erhan, et al. SSD: single shot multi box detector[C]//European Conference on Computer Vision, Amsterdam, 2016:21-37.
- [10] C-Y Fu, W Liu, A Ranga, et al. DSSD: deconvolutional single shot detector[J]. arXiv preprint arXiv, 2017, 1701:06659.
- [11] J Redmon, A Farhadi. YOLO9000: better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 2017:6517-6525.
- [12] J Redmon, S Divvala, R Girshick, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016:779-788.