

基于暗通道去雾和深度学习的行人检测方法

田青¹, 袁瞳阳^{1*}, 杨丹¹, 魏运²

¹北方工业大学电子信息工程学院, 北京 100144;

²北京城建设计发展集团有限公司, 北京 100037

摘要 行人检测是实现智能交通与客流监控的关键技术,深度学习训练模型已经在行人检测领域取得了良好的效果。但是当训练样本质量不佳时,训练的模型往往不能得到令人满意的效果。为了提高雾霾天气与曝光较强环境下的行人检测效果,提出了将暗通道去雾算法应用于深度学习的样本预处理中,并使用快速深度卷积神经网络训练行人检测模型。在实验中,首先对 10000 张样本图片采用暗通道去雾算法进行预处理,之后分别使用有无暗通道去雾算法预处理的样本图片训练模型,最后比较这两种模型在不同场景下的模型检测准确率。实验结果表明,使用暗通道去雾预处理后的样本训练得到的深度模型具有更好的检测效果,在多种场景下的检测率都得到提升。

关键词 图像处理; 行人检测; 暗通道去雾; 深度学习; 神经网络

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP55.111007

A Pedestrian Detection Method Based on Dark Channel Defogging and Deep Learning

Tian Qing¹, Yuan Tongyang^{1*}, Yang Dan¹, Wei Yun²

¹ School of Electronic Information Engineering, North China University of Technology, Beijing 100144, China;

² Beijing Urban Construction Design & Development Group Co., Ltd., Beijing 100037, China

Abstract Pedestrian detection is the key technology to realize intelligent traffic and passenger flow monitoring. Currently, the training model of deep learning method has achieved good results in pedestrian detection. However, when the training samples are poor, the training model often fails to achieve good results. In order to improve the effect of pedestrian detection under hazy weather and strong exposure environment, the dark channel defogging algorithm is applied to pretreat deep learning samples. And pedestrian detection model is trained with fast deep convolutional neural network. In this experiment, the dark channel defogging algorithm is applied to preprocess the 10,000 sample images. After that, the sample images preprocessed by the defogging algorithm with and without dark channel are used to train model, respectively. Finally, detection accuracy of these two models under different scenarios are compared. The experimental results show that the depth model obtained by using the dark channel defogging pretreatment sample has a better detection effect and the detection rate increases under many scenarios.

Key words image processing; pedestrian detection; dark channel defogging; deep learning; neural networks

OCIS codes 100.2980; 100.4996; 100.5010

1 引言

行人检测技术具有广泛的应用,例如智能驾驶辅助、客流监测、行人行为分析等。行人检测具有更高的准确率也一直是科研人员的追求。随着机器视

觉与模式识别技术的进步,自 2005 年以来,行人检测技术进入一个快速发展的阶段^[1]。近年来基于深度学习的目标检测逐渐发展起来,比如,2013 年提出的基于区域的卷积神经网络(R-CNN)是一个可以将深度学习进行目标检测应用到工业领域的方

收稿日期: 2018-04-16; 修回日期: 2018-05-16; 录用日期: 2018-05-29

基金项目: 轨道交通突发事件应急抢险指挥辅助决策系统研发及示范(Z161100001016003)、国家重点研发计划资助(2016YFB1200402)

* E-mail: 2017311020137@mail.ncut.edu.cn

案,后续通过将很多传统的机器视觉方法与深度学习方法相结合,又产生了众多性能更好的基于区域的卷积神经网络,例如更快的区域卷积神经网络(Faster R-CNN)[2]。

深度学习相比于传统的机器学习,最主要的优势是不需要人工设计特征提取器,而是由机器自动学习而获得。而机器学习自动学得特征十分适用于变化多端的自然数据,具有很好的泛化能力与稳健性[3]。深度学习有着广泛的应用,在再检测与识别领域取得了很好的效果,最常见的有行人检测、遥感图像检测、人脸识别等[4]。但不论是使用传统的机器学习方法还是使用深度学习进行行人检测都存在着一些难点,一是因周围光照环境不同,例如在阴影较多或是雾霾较多的天气下,目标的检测会变得更加困难;二是分类器的性能受训练样本的影响较大。所以本研究想要通过对质量不佳、曝光度较强的行人样本使用暗通道去雾算法进行去雾增强处理,得到质量更优的图片,尝试训练更好的模型。暗通道去雾方法是一种实用性很强的图像处理方法,对于雾气天气的遥感图像检测发挥了很大作用[5],本文将其应用于行人检测的预处理中。图片具有雾感主要有如下三个原因,一是拍照现场存在雾,二是空气中灰尘等小颗粒过多引起的漫反射也会使图片存在雾感,三是曝光过度导致照片泛白。这三种情况在使用行人检测系统检测时是比较常见的。

本文通过对在北京地铁站收集到的行人样本使用去雾算法处理,得到更为清晰的、目标更加突出的样本图片集,再使用深度学习的方法训练行人检测的深度模型,从而获得比直接使用原始图片训练深度模型更好的检测效果。本文开展的研究工作共包括如下方面:1) 使用标签工具对采集到的 10000 张行人图片标标签;2) 根据暗通道去雾算法编写 Python 程序,并对这 10000 张图片进行去雾增强操作;3) 将原始样本图片与处理后的样本图片分别投入到深度神经网络框架下,设置相应的网络参数,并开始训练;4) 测试这两个深度模型在几组不同场景下的行人检测效果;5) 根据实验结果做出分析。

2 基本原理

2.1 暗通道去雾算法

根据文献[6],图片暗通道的含义为,在绝大多数非天空的局部区域里,有些像素至少有一个颜色通道的像素值很低,或者说这个区域的光强度值是一个很小的值,这个区域称为图片暗通道。对于一

幅输入图像 J 的暗通道可以表示为

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} \left[\min_{c \in \{r, g, b\}} J^c(y) \right], \quad (1)$$

式中 J^c 表示彩色图像的每个通道, $\Omega(x)$ 表示以像素 x 为中心的一个窗口, r, g, b 表示图像的三个颜色通道, y 表示某个颜色通道的像素值, x 表示暗通道的像素值。根据暗通道的先验理论得知:

$$J^{\text{dark}} \rightarrow 0, \quad (2)$$

在机器视觉中,通常使用下式所描述的雾感图形成模型,即:

$$I(x) = J(x)t(x) + [A(1 - t(x))], \quad (3)$$

式中 $I(x)$ 为需要进行暗通道去雾的图像, $J(x)$ 是需要恢复的无雾感图像,参数 A 是全球大气光的成分, $t(x)$ 为透射率。现在已知 $I(x)$, 需要求解出 $J(x)$ 。根据代数知识可知,这个方程有无数解,所以需要有一些先验基础知识才能达到定解,首先将上述方程变化为

$$\frac{I^c(x)}{A^c} = t(x) \frac{J^c(x)}{A^c} + 1 - t(x), \quad (4)$$

式中右上角的 c 表示 r, g, b 三个通道。假设每个窗口的内透射率 $t(x)$ 为常数,将其定义为 $\tilde{t}(x)$, 并且 A 值已经给定,对(4)式求两次最小值运算,得到:

$$\min_{y \in \Omega(x)} \left[\min_c \frac{I^c(y)}{A^c} \right] = \tilde{t} \min_{y \in \Omega(x)} \left[\min_c \frac{J^c(y)}{A^c} \right] + 1 - \tilde{t}(x). \quad (5)$$

并且根据暗原色的先验理论可以得到:

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} \left[\min_c J^c(y) \right] = 0. \quad (6)$$

因此,可以进一步推导出:

$$\min_{y \in \Omega(x)} \left[\min_c \frac{J^c(y)}{A^c} \right] = 0. \quad (7)$$

把(7)式结果代入(5)式中,就可以得到透射率 $\tilde{t}(x)$ 的预估值,即:

$$\tilde{t}(x) = 1 - \min_{y \in \Omega(x)} \left[\min_c \frac{J^c(y)}{A^c} \right]. \quad (8)$$

由于在现实生活中,雾的存在使人感到景深的存在,所以在去雾处理时通常需要保留一定的雾,这需要通过(8)式引入一个 $[0, 1]$ 之间的因子来实现,从而将(8)式修正为

$$\tilde{t}(x) = 1 - \varpi \min_{y \in \Omega(x)} \left[\min_c \frac{J^c(y)}{A^c} \right]. \quad (9)$$

在上述推论中假设全球大气的光值 A 值是已知的,而在实际中则可以使用暗通道图像从有雾感的图像中获得该值。根据文献[6]中暗通道去雾原理,首先从暗通道图像中根据亮度大小提取出最亮

的前 0.1% 像素, 选取范围较小的亮度值, 更加接近最亮的像素。然后, 在原始有雾霾图像 I 中寻找对应位置上的具有最高亮度点的值, 最后将这个值作为 A 的值。并且考虑到当透射图像的 t 很小时, 会导致 J 的值偏于过大, 导致图像整体偏白场过度。所以一般还需要设置一个阈值 t_0 , 从而使得当 t 小于 t_0 时, 令 $t = t_0$ (例如 $t_0 = 0.1$)。最终得到图像恢复公式为

$$J(x) = \frac{I(x) - A}{\max[t(x), t_0]} + A. \quad (10)$$

根据(10)式便可以进行最终的无雾图像恢复。

2.2 更快的区域卷积神经网络

本文采用深度学习在目标检测领域的经典框架 Faster R-CNN 框架训练检测模型, 即更快的区域卷积神经网络框架训练行人检测模型, 此种方法被广泛应用到行人检测当中^[7]。相对于更新的 YOLO (you only look once)、单一精准多框检测网络 (SSD) 等深度学习框架, Faster R-CNN 采用候选区域提取网络, 最终可以获得 300 个候选区域, 多于 YOLO 的 98 个候选区域, 虽然速度比 YOLO 框架慢, 但可以具有更高的检测精度。SSD 框架与 YOLO 框架更为接近, 但 SSD 最终的输出结果显示为对应目标属于哪一类的概率, 而不是直接显示为类别, 因此 Faster R-CNN 的结果更为直观。Faster R-CNN 把目标检测的 4 个基本步骤统一到一个深度网络框架之内。训练得到检测模型的 4 个基本步骤为候选区域生成、特征提取、分类以及边框回归。

Faster R-CNN 使用 ZF (Zeiler-Fergus) 网络的前 5 层进行特征提取^[2], 使用区域生成网络 (RPN) 提取候选框区域。Faster R-CNN 框架训练的步骤为, 首先向预先训练好的卷积神经网络 ZF 中输入任意大小的图片, 然后经过 ZF 网络前向传播至最后共享的卷积层, 这样一方面可以得到共 RPN 输入的特征图, 另一方面继续前向传播至特有卷积层, 以产生更高维度的特征图。接着将供 RPN 输入的特征图使用 RPN 处理得到区域建议和区域得分, 为了能够应对尺寸不同的物体, Faster-RCNN 采用了三种不同类型的滑动窗口, 长宽比分别为 1:1、1:2、2:1, 并用三个尺度滑动缩放窗口, 最终采用 9 种类型的滑动窗口, 这些窗口经过卷积产生 256 维的向量, 每张图片最终通过分类挑选出 300 个得分最高的区域作为建议区域窗口。将得到的更高维特征图与产生的区域建议同时输入到感兴趣区域

(ROI) 池化层, 提取对应的区域建议部分尺寸固定的特征。将这些尺寸固定的特征通过全连接层, 得到该区域的分类得分与边框回归。

ZF 网络的结构^[8]如图 1 所示, Faster R-CNN 框架使用 ZF 网络全连接层之前的层用于特征提取。在 ZF 网络结构中, 首先输入大小为 $224 \times 224 \times 3$ 的图片, 其中 224 为图像大小, 3 是由于输入的为彩色图像所具有的三个通道, 之后第一层卷积核的维度为 $7 \times 7 \times 3$, 所以对原始图片卷积得到的卷积层 1 的结果为 $110 \times 110 \times 96$, 其中 96 是由于采用了 96 个不同类别的卷积核, 110 的计算方法为 $(224 - 7 + \text{填充像素}) / 2 + 1$, 在图片周围补充像素的目的是为了整除。然后是对卷积后的结果进行池化, 池化核的大小为 3×3 , 池化的步长为 2, 所以池化后的图片维度为 $55 \times 55 \times 96$, 其中 55 的计算方法为 $(100 - 3 + \text{填充像素}) / 2 + 1 = 55$ 。之后几层的计算方法与前面类似, 最后取卷积层 5 的结果也就是 $13 \times 13 \times 256$ 输送给 RPN。

由文献[2]可知 RPN 的实现方式, 如图 2 所示, 对前边 ZF 网络卷积层 5 得到的 $13 \times 13 \times 256$ 特征图进行处理, 通过增加滑动窗口操作以及使用两个卷积层实现区域建议的功能。首先对卷积层 5 得到的特征图用 $n \times n$ 的滑动窗口在特征图上滑动扫描, 每个滑窗区域通过第一层卷积层映射到一个 256 维的低维特征向量, 由于每个滑动窗口考虑 K 种可能的参考窗口, 本文采用 9 种, 使得 RPN 在对特征图进行滑窗时, 可以对滑窗进行多尺度多长宽比的采样, 并对多尺度和多长宽比的候选区域进行回归和分类, 对于一个 $W \times H$ 的特征图, 可以产生 $W \times H \times K$ 个区域建议。之后将得到的 256 维低维特征向量输入两个并行连接的卷积层, 即窗口回归层和窗口分类层, 窗口回归层可以对回归区域产生建议, 窗口分类层可以为框内物体打分, 判断是前景还是背景, 由于每个滑动窗口共产生 K 个区域建议, 所以窗口回归层得到 $4K$ 个参数, 分别为建议框中心锚点对应的坐标 x 、 y 和建议框的宽高 w 、 h 。窗口分类层共产生 $2K$ 个得分, 分别给出 K 个区域建议为目标还是背景的概率。

3 实验过程

3.1 对样本图片标签

本文基于正在进行的地铁项目, 将由北京地铁站的监控摄像头所采集到的 10000 张客流图片用于

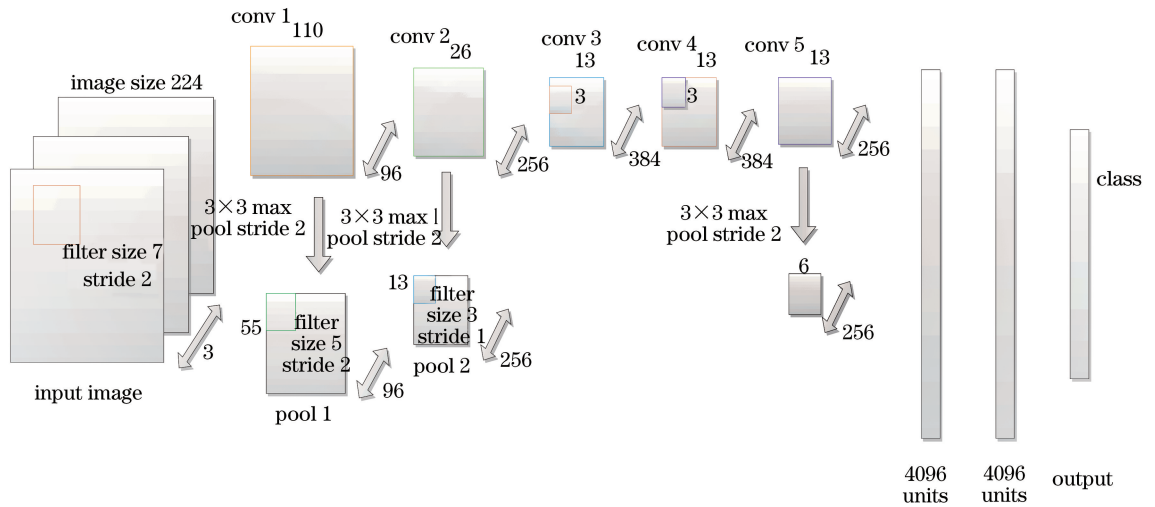


图 1 ZF 网络结构

Fig. 1 Structure of ZF network

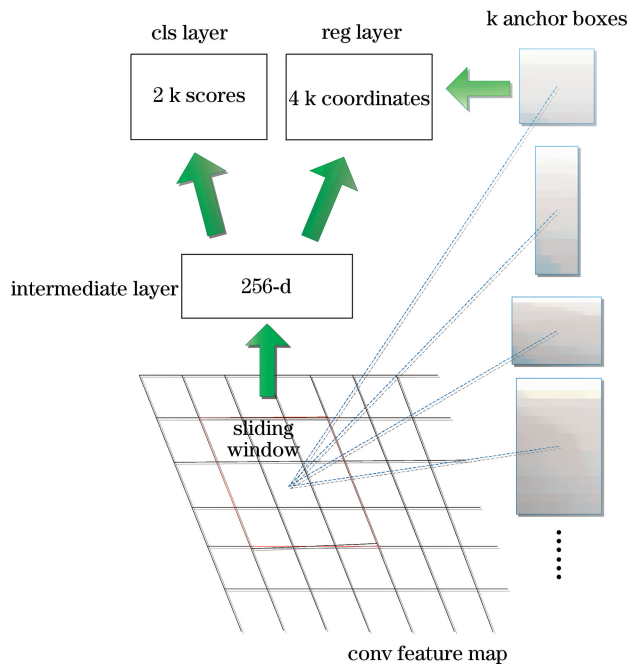


图 2 RPN 结构图

Fig. 2 Structure of RPN

训练模型,随机选取 9000 张作为训练集,1000 张作为测试集,并且选择行人的头肩作为检测行人目标的依据。首先使用标签工具对这 10000 张图片中的所有行人进行标注,标签分为两类,一类标注为头肩,另一类标注忽略,“忽略”即为被帽子遮挡难以识别的行人。样本中所标注的每一个头肩框都是手动框选的,使用鼠标框选目标的“最小外接矩形”,该最小外接矩形尽量覆盖目标完整的边缘,而尽可能少的包含背景。从而使得网络进行学习时尽可能多于完整的学习目标的特征信息,减少对无用背景信息

的学习。标注示例如图 3 所示,图 3(a)为使用标签工具框选目标的过程,图 3(b)为打框标注完成后所得到的图片标签,包括每一个标注框的横纵坐标、大小以及所标注的类别等信息。

3.2 暗通道去雾算法处理图片

根据文献[9],暗通道去雾算法对夜晚图像具有较好的增强效果,将其应用在雾霾天、地铁站等光线不佳的环境中,从而实现样本质量增强。标注完成后使用暗通道去雾算法对地铁中曝光不佳的行人样本与室外雾霾天气下的行人样本进行处理。使用



图 3 (a) 示例标签图片; (b) XML 格式的标签文件

Fig. 3 (a) Sample label image; (b) label file in XML

python 软件载入 opencv 模块编写暗通道去雾程序处理图片样本集。图片大小为 640×480 。图 4 为暗通道去雾的算法流程图, 描述图片的处理过程。

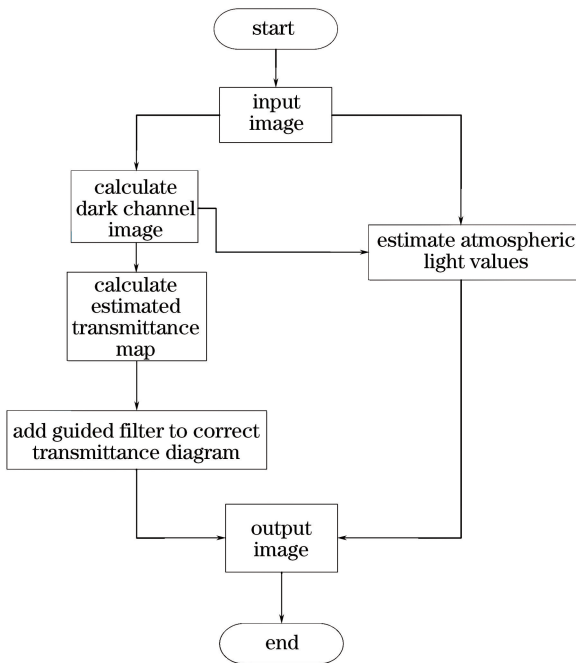


图 4 暗通道去雾算法流程图

Fig. 4 Flow chart of dark channel defogging algorithm

3.3 使用 Faster R-CNN 框架训练检测模型

根据文献[10], Faster R-CNN 被用于人脸检测的系统中并取得较好的效果; 根据文献[11], Faster R-CNN 对于小目标的检测结果较优, 检测率优于已有的多数检测模型。因此本文的行人检测方法为在 640×480 大小的地铁客流数据集上进行仿真实验。实验数据来自地铁客流图片样本集。实验环境为 k80 GPU 服务器、ubuntu14.04 操作系统、网络框架为 caffe, 采用检测率(DR)、虚警率(FAR)和平均运行时间作为实验评价指标。根据文献[2], 将图像测试时交并比(IoU)阈值设为 0.5, IoU 区域即为手动框选区域与使用算法训练所得模型的预测到区域的

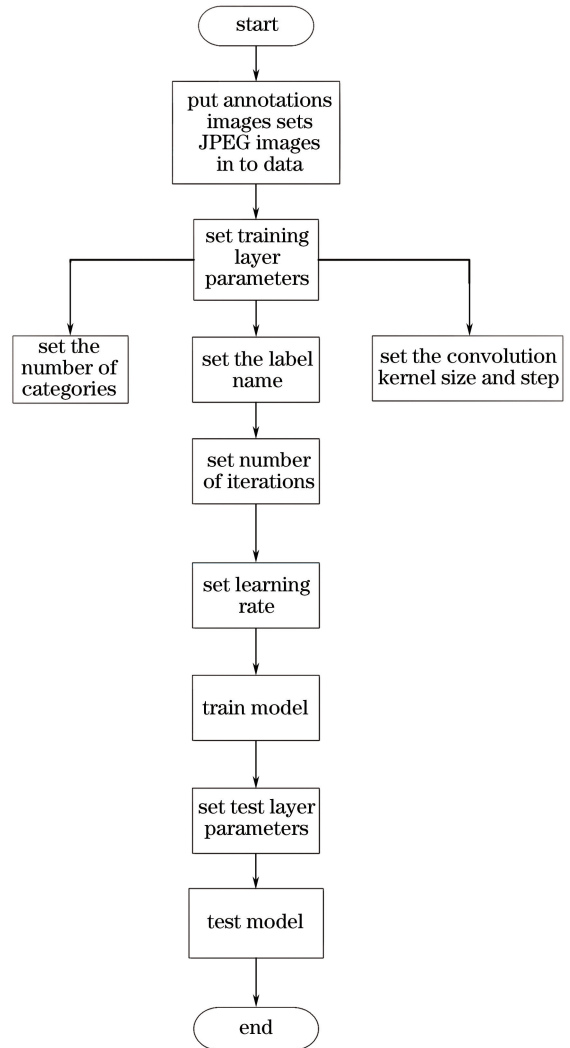


图 5 Faster R-CNN 训练算法流程图

Fig. 5 Flow chart of Faster R-CNN training algorithm

重合程度, 其值等于两个 box 的交集除以两个 box 的并集, 常用阈值范围为 0.3~0.5。本文将阈值设为 0.5, 如果某候选区域与任一目标区域的 IoU 最大, 则该候选区域判定为有目标; 如果某候选区域与任一目标区域的 $\text{IoU} > 0.5$, 则判定为有目标; 如果某候选区域与任一目标区域的 $\text{IoU} < 0.5$, 则判定为

背景。在 Faster R-CNN 论文中,使用 Faster R-CNN 做目标多分类任务,检测目标较多,所以将 IoU 值设为 0.3,本文只需检测行人目标,设置两类标签,所以将 IoU 值设为 0.5,减少网络训练量,同时在一定程度上降低过拟合。首先将原始图片样本集 JPEGImages、样本标签 Annotations、图片名称集 ImageSets 投入到基于 ZF 网络的 Faster R-CNN 框架下进行训练,采用端到端的训练方式,并且将迭代次数设置为 80000,使得神经网络能够更加充分地学习样本图片的特征。将前 60000 次迭代的学习率设置为 0.001,后 20000 次迭代的学习率设置为 0.0001。训练完成后得到模型 1。之后将使用去雾算法处理后的样本集 2 设置完全相同的参数,在相同环境下训练得到模型 2。最后选取多种场景下的行人测试集对模型进行测试,可以观察到,使用增强

图像训练得到的模型 2 在两个方面显示出优势,首先,模型 2 的检测结果与模型 1 在具备相同的正确检测效果的同时,降低了对行人的虚警率。其次,模型 2 在相同场景下能够降低模型的漏检率,对于图片中与背景对比度不高或是被遮挡的行人,实现了更好的检测效果,提高了模型的检测率。图 5 为使用 Faster R-CNN 训练深度模型并进行检测的流程图。

4 实验结果与分析

使用去雾算法处理后的图片与未处理图片的对比结果如图 6 所示,图 6(a)和(c)为室外雾霾天气获取到的行人图片与暗通道算法处理后的结果对比,图 6(b)和(d)为室内曝光不佳的行人图片与暗通道算法处理后的结果对比。



图 6 (a)(b)原始图片;(c)(d)暗通道去雾算法处理后图片

Fig. 6 (a) (b) Original images; (c)(d) image processed by dark channel defogging algorithm

从图 6 中可以观察到,使用暗通道去雾算法处理后的行人样本图片,雾霾天被雾气遮挡的行人能够凸显,目标更加清晰。对于室内拍摄的曝光度不足、环境光线较暗的样本图片,图片质量和对比度得到增强,亮度得到提升。

本次实验训练得到两个模型,使用原始图片样本集 1 训练得到的模型一,和使用暗通道去雾算法处理后的图片样本训练得到的模型二。训练得到模型后采用多种场景下的图片来检测效果,实验结果如下。

实验结果一:使用未进行暗通道去雾增强的图片测试模型 1 与模型 2,可以观察到两种模型都能够较好地检测出行人。但使用暗通道去雾增强后的图片训练得到的模型在两个方面展现出优势,首先在保证能检测出相同数目行人的情况下减少了虚警的情况,即将一些背包、垃圾桶等误判为行人的情况;其次对于原始模型不能检出的一些与背景颜色区分不明显的行人,检测效果也有了一些提升。实验效果如图 6 与图 7 所示。

图 7(a)和(c)为使用原始图片训练得到的模型 1



图 7 (a)(c)模型 1 检测结果;(b)(d)模型 2 检测结果

Fig. 7 (a)(c) Test results of model 1; (b)(d) test results of model 2

检测行人的结果示例,图 7(b)和(d)为使用去雾算法增强后图片训练得到的模型 2 检测行人的结果示例。图 7(a)左下角出现将书包检测为人的情况,图 7(c)右下角出现将手提袋检测为人的情况,而这些情况在图 7(b)和(d)中却不存在,从而可以得知增强后的模型减少了检测时的虚假检测情况。

图 8(a)(c)(e)为使用原始图片训练得到的模型 1 检测行人的结果示例,图 8(b)(d)(f)为使用去雾算法增强后图片训练得到的模型 2 检测行人

的结果示例。图 8(a)左侧有一名女士被漏检,而在图 8(b)中被检测出,图 8(c)有一名缠着纱带的老人未被检出,而在图 8(d)中被检测出,图 8(e)左侧有两名情侣未被检出,而在图 8(f)中被检出。从而可以得知增强后的模型减少了漏检,提升了检测效果。

实验结果二:使用暗通道去雾处理后的测试图片分别测试模型 1 与模型 2。测试结果如图 9 和 10 所示。



图 8 (a)(c)(e)模型 1 检测结果;(b)(d)(f)模型 2 检测结果

Fig. 8 (a)(c)(e) Test results of model 1; (b)(d)(f) test results of model 2

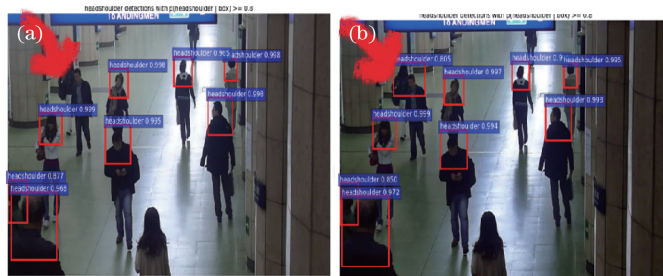


图 9 (a)模型 1 检测结果 1;(b)模型 1 检测结果 2

Fig. 9 (a) Test results 1 of model 1; (b) test results 2 of model 1



图 10 (a)模型 2 检测结果 1;(b)模型 2 检测结果 2

Fig. 10 (a) Test results 1 of model 2; (b) test results 2 of model 2

图 9 为模型 1 的测试结果,图 9(a)使用未去雾增强的测试图片,图 9(b)使用去雾增强后的测试图片,可以观察到使用增强图片检测到的行人人数增加(左上角)。图 10 为模型 2 的测试结果,图 10(a)使用未去雾增强的测试图片,图 10(b)使用增强后

的测试图片,可以观察到使用增强图片检测到的行人人数增加(左下角)。

根据表 1 中的数据可以得到模型 1 和模型 2 的检测率和虚警率。第一列数据采用室内地铁站曝光度不好的图片作为测试图片的结果,第二列数据为室外场

景中选取的有雾感测试图片进行行人检测的结果,第三列为这些图片混合在一起的平均检测结果,第四列数据为对这些测试图片做图像去雾前的测试结果。经

过检测发现,模型 2 的检测结果优于模型 1,检测率平均提高 2%,虚警率降低 3%。另外发现如果增强测试图像,测试结果同样可以得到一定程度的优化。

表 1 模型 1 与模型 2 的检测率与虚警率

Table 1 Detection rate and false alarm rate of model 1 and model 2

%

Evaluation index	Test sample			
	Poor quality test picture	Fog test picture	Average results with test image enhancement	Average results without test image enhancement
Detection rate of model 1	90	88	90	89
Detection rate of model 2	93	90	92.5	91.5
False alarm rate of model 1	5	8	5.5	7
False alarm rate of model 2	2	5	3	4

模型 1 的大小为 230×185 kb;模型 2 的大小为 230×181 kb;模型 2 的大小略小于模型 1,但差别很小,所以两个模型对不同像素大小图片的检测时间

如表 2 所示,可以看出两者的检测时间几乎没有差别。所以根据表 1 与表 2 可知,模型 1 与模型 2 检测时间相近,但模型 2 的检测率更高。

表 2 模型 1 与模型 2 的检测时间

Table 2 Detection time of model 1 and model 2

Model	Pixel size			
	500×375	500×333	640×480	1280×720
Detection time of model 1 /s	2.38091	2.45501	2.47809	2.490890
Detection time of model 2 /s	2.37017	2.45496	2.47091	2.480745

这是因为使用暗通道去雾算法对图片做处理后,原始曝光效果不好或是由于拍摄场景存在雾感造成图片质量不佳的样本图片,图片对比度得到了改善,物体的目标与背景区域之间的差异更加明显。目标的轮廓和边缘也更清晰。因此,将这些图片自动添加到更快的 R-CNN 和卷积神经网络提取特征中,网络可以更好地获取目标的颜色和轮廓信息。即使质量差的样本图片也可以更好地提取目标特征。因此,通过增强图像训练的模型与使用未增强图像获得的模型相比具有更低的错误率,模型拟合效果更好。并且增强测试样本也可以得到更好的结果。总之,暗通道去雾算法可以用来增强样本图像或测试图像,然后再使用深度学习的方法训练,得到的检测模型效果更好。

质量不佳或是有雾感的场景中,更好地发挥深度学习训练行人检测模型的优势。

参 考 文 献

- [1] Sermanet P, Kavukcuoglu K, Chintala S, *et al.* Pedestrian detection with unsupervised multi-stage feature learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013: 3626-3633.
- [2] Ren S Q, He K M, Girshick R, *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.
- [3] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] Chen Y, Fan R S, Wang J X, *et al.* Cloud detection of ZY-3 satellite remote sensing images based on deep learning [J]. Acta Optica Sinica, 2018, 38 (1): 0128005.
陈洋, 范荣双, 王竞雪, 等. 基于深度学习的资源三号卫星遥感影像云检测方法[J]. 光学学报, 2018, 38(1): 0128005.
- [5] Guo H, Xu X T, Li B. Study on image dehazing methods based on dark channel prior[J]. Acta Optica Sinica, 2018, 38(4): 0410002.

5 结 论

基于暗通道去雾与深度学习的方法,利用 Faster R-CNN 框架训练了行人检测模型,提高了检测精度。测试结果表明该模型的平均识别率为 93%,检测效果更加准确,错误检测率远远低于没有图像增强的模型。因此,利用暗通道去雾图像增强方法对地铁客流图片进行预处理可以使 Faster R-CNN 网络更好地学习样本特征,训练得到检测精度更高的模型。这种方法可以应用于获取的样本图片

- 郭翰, 徐晓婷, 李博. 基于暗原色先验的图像去雾方法研究[J]. 光学学报, 2018, 38(4): 0410002.
- [6] He K M, Sun J, Tang X O. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [7] Ye G L, Sun S Y, Gao K J, *et al.* Nighttime pedestrian detection based on faster region convolution neural network [J]. Laser & Optoelectronics Progress, 2017, 54(8): 081003.
叶国林, 孙韶媛, 高凯珺, 等. 基于加速区域卷积神经网络的夜间行人检测研究[J]. 激光与光电子学进展, 2017, 54(8): 081003.
- [8] Gupta S, Girshick R, Arbeláez P, *et al.* Learning rich features from RGB-D images for object detection and segmentation [C] // European Conference on Computer Vision, 2014: 345-360.
- [9] Yang A P, Bai H H. Nighttime image defogging based on the theory of Retinex and dark channel prior [J]. Laser & Optoelectronics Progress, 2017, 54(4): 041002.
杨爱萍, 白煌煌. 基于 Retinex 理论和暗通道先验的夜间图像去雾算法[J]. 激光与光电子学进展, 2017, 54(4): 041002.
- [10] Wang K, Dong Y, Bai H L, *et al.* Use fast R-CNN and cascade structure for face detection [C] // Visual Communications and Image Processing (VCIP), 2016: 1-4.
- [11] Eggert C, Brehm S, Winschel A, *et al.* A closer look: small object detection in faster R-CNN [C] // IEEE International Conference on Multimedia and Expo (ICME), 2017: 421-426.