

基于YOLOv3的深度学习交通标志识别系统

张钟文 高宇 王静 曹登平 (中铁二院工程集团有限责任公司, 成都市 610031)

Deep Learning Traffic Sign Recognition System Based on YOLOv3

ZHANG Zhongwen GAO Yu WANG Jing CAO Dengping

(China Railway Eryuan Engineering Group Co., Ltd., Chengdu 610031, China)

Abstract: The difficulties in the study of target detection using deep learning algorithm in the identification of prohibited traffic signs and the related research results at home and abroad are introduced; the use of YOLOv3 target detection network model and its training method is expounded; and the research results are evaluated by measuring the average accuracy.

Key words: traffic sign recognition; neural network; deep learning; automatic drive; YOLOv3; target detection network model; data processing; model training

摘要: 介绍采用深度学习算法中的目标检测网络模型在禁令交通标志识别研究方面的难点以及国内外对此类问题的相关研究成果; 阐述基于YOLOv3目标检测网络模型以及其训练方法; 通过测定平均精度及平均召回率对研究结果进行评估。

关键词: 交通标志识别; 神经网络; 深度学习; 自动驾驶; YOLOv3; 目标检测网络模型; 数据处理; 模型训练

中图分类号: TP391

文献标识码: A

doi: 10.3969/j.issn.1003-8493.2020.07.011

0 引言

辅助驾驶系统是解决交通安全问题以及实现无人

驾驶的关键技术之一, 而禁令交通标志识别是其中的一个重要组成部分。目前对于这一问题的主要研究思路是通过安装在车辆上的摄像机获取自然场景图像, 进而通过图像处理与模式识别等技术对场景中的禁令交通标志进行实时检测和识别。由于真实的道路交通环境复杂多变, 而现实应用又要求其保证较高的准确率和实时性, 传统方法使用滑动窗口结合手工设计特征进行交通标志识别, 在精度和实时性上难以满足智能驾驶系统的需要。近年来, 卷积神经网络(Convolutional Neural Network, CNN)和候选区域网络(Region Proposal Network, RPN)等深度学习算法的发展为交通标志识别技术提供了新的可能。

1 研究难点及相关研究

采用深度学习算法中的目标检测思想对禁令交通标志识别有两个难点:

a. 由于现代的深度卷积神经网络拥有强大的特征提取和表达能力, 网络本身需要大量甚至海量数据来驱动模型训练, 否则可能会陷入模型训练过拟合。

b. 在不同的天气和光照条件下, 辅助驾驶系统都应能够很好地执行检测。这就意味着要保证各种条件下的训练图像数据充足且数量相近, 以便于神经网络模型正确学习不同天气和光照条件的特征。

德国交通标志检测基准(GTSDB)^[1]是2012年成立的挑战项目。GTSDB数据集包含有900张街景

作者信息

张钟文, 女, 中铁二院工程集团有限责任公司, 工程师。

高宇, 男, 中铁二院工程集团有限责任公司, 高级工程师。

王静, 男, 中铁二院工程集团有限责任公司, 高级工程师。

曹登平, 男, 中铁二院工程集团有限责任公司, 工程师。

图像,并在每张图片的不同位置上具有一个或多个交通标志。每年都会有计算机视觉的研究人员参与这项挑战,并建立不同的模型以检测图像中交通标志的位置和对它们的类型进行分类。

但是,GTSDb的数据集对于现代神经网络模型来说不够充足,容易导致模型过拟合。所以为了解决这个问题,有两种不同的方法可以用来扩展原始数据集的大小:

a. 第一种方法可以手动在“道路/车道检测评估(RLDE)”数据集图像上粘贴德国交通标志识别基准(GTSRB)数据集上的交通标志。具体可以参考Evan Peng等人在《Traffic Sign Detection with Convolutional Neural Networks》一文中提及的方法。

b. 第二种方法可以使用其他交通标志检测数据集替换或扩展原始训练数据。选择之一是BelgiumTS数据集^[2],其中包含5 905张图片和8 851个标签。另两种替代方法是北京交通大学开发的中国交通标志检测数据库(TSDD)^[3]和清华—腾讯联合实验室开发的TT100K数据集^[4],这其中包含高达10万张带有不同交通标志标签的图片。

这两种方法相比,第二种方法更为简便可靠。由于BelgiumTS数据集中只有10%左右的图片含有禁止交通标志,而TT100K数据集中含有最新的街景图像,并且涵盖了不同光照强度和天气条件。这在某种程度上也解决了在各种天气和光照条件下检测交通标志的第二个难点。所以本文选择使用TT100K数据集对神经网络进行训练。

2 神经网络模型选择及训练方法

2.1 神经网络模型选择

有许多潜在的神经网络模型适合此任务,例如,目标检测模型(如Faster R-CNN、YOLO、SSD)或图像语义分割模型(例如Mask R-CNN、DeepLab等)。其中清华—腾讯联合实验室提出的一个高效检测模型,其在精确度和召回率测量方面具有比Fast R-CNN更好的性能^[5]。

该模型是从AlexNet^[6](AlexNet是2012年

ImageNet竞赛冠军获得者Hinton和他的学生Alex Krizhevsky设计)派生而来,只是它在第6卷积层上分支成3个不同的流(如图1所示)。该模型还使用了与YOLO模型非常相似的检测输出,可以一次预测对象的类别和边界框。3个分支流将输出:检测对象存在于某个特定网格的概率(conv8-pixel);检测目标的实际边界框与上述特定网格的相对位置(conv8-bbox);目标对象所属的类标签(conv8-label)。

本文之所以选择YOLOv3模型^[7]来完成此研究,主要是因为它使用了基于残差卷积神经网络的darknet-53主干卷积网络模型。此类模型比TT100K使用的AlexNet模型更深、更复杂,更容易提炼出检测目标的深层特征。YOLOv3会将图片分成3个大小不同的区域来检测各种大小的目标对象(如图2所示)。

YOLOv3模型的最终输出信息里包含:

- 边界框中心的 X, Y 坐标以及边界框的宽度和高度;
- IOU(交并比):代表预测的边界框和地面真值边界框之间的重叠区域的百分比;
- 目标对象所属的类标签。

选择YOLOv3的另一个原因是它的训练操作相对

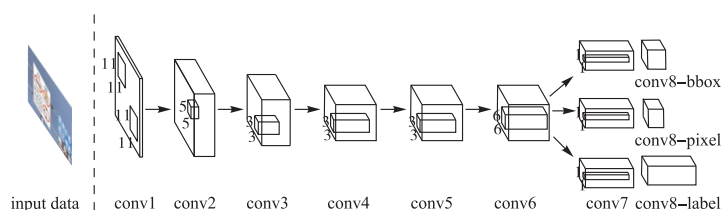


图1 TT100K目标检测模型

Fig. 1 TT100K target detection model

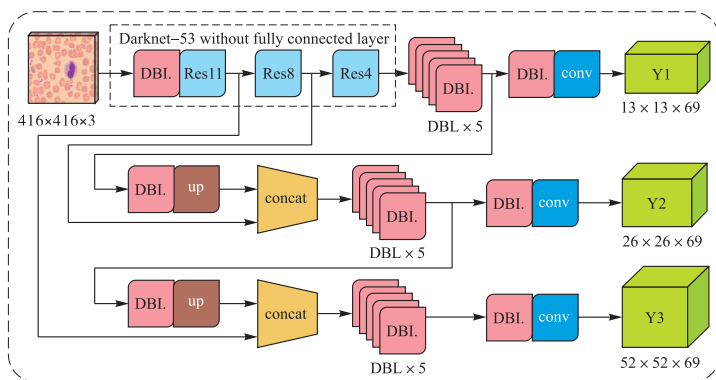


图2 YOLOv3目标检测模型

Fig. 2 YOLOv3 target detection model

简便。Darknet 深度学习框架^[8]允许使用者在数据训练之前,只需对其原始配置文件进行一些更改,并将检测目标的标签格式转换为 YOLOv3 所使用的格式即可开始训练。

2.2 数据预处理

本文首先对原始的 TT100K 进行数据扩充。有效的数据扩充不仅能扩充训练样本的数量,增加样本的多样性,而且还能避免模型过拟合的发生。本文主要使用伽马校正原始数据集中增添不同光照强度的街景图像(如图3所示)。

在 TT100K 数据集中,除了“ps”(“停止”交通标识)和“pg”(三角形交通标识)外,所有标签以“p”开头的都为禁令交通标识(如图4所示)。准备训练数据的标签时,本文将所有禁令交通标识都视为同一类型,并将其他不包含红色圆圈符号的图像都



图3 使用伽玛校正改变街景的光照强度
Fig. 3 Using Gamma correction to change the lighting intensity of the street view

视为单纯的背景。

2.3 神经网络模型训练

由于 TT100K 数据集使用检测目标边界框的边线坐标来表示目标位置(X_{\min} 表示左边线, X_{\max} 表示右边线, Y_{\min} 表示下边线, Y_{\max} 表示上边线),所以训练 YOLOv3 模型之前,需要用下面的公式把 TT100K 的标签格式转换为 YOLOv3 模型所要求的“目标边界框中心点坐标(X_{center} , Y_{center}) + 高度($Height_{\text{bbox}}$) + 宽度($Width_{\text{bbox}}$)”格式:

$$X_{\text{center}} = \frac{X_{\max} + X_{\min}}{2} \times Width_{\text{image}} \quad (1)$$

$$Y_{\text{center}} = \frac{Y_{\max} + Y_{\min}}{2} \times Height_{\text{image}} \quad (2)$$

$$Width_{\text{bbox}} = (X_{\max} - X_{\min}) \times Width_{\text{image}} \quad (3)$$

$$Height_{\text{bbox}} = (Y_{\max} - Y_{\min}) \times Height_{\text{image}} \quad (4)$$

式中: X_{center} , Y_{center} ——目标边界框中心点 X , Y 坐标;

$Height_{\text{bbox}}$ ——目标边界框高度;

$Width_{\text{bbox}}$ ——目标边界框宽度。

除此之外,在训练过程中,有3个配置文件需要输入到“Darknet”的框架中,它们分别是:

a. “red_sign_obj. names”文件:包含所有对象类别的名称;

b. “red_sign_obj. data”文件:包含对象类别的数量、训练数据集和测试数据集的文件路径、“red_sign_obj. names”文件的路径,以及用于保存模型权重文件的路径;

c. “red_sign_yolov3. cfg”文件:包含了修改后的 YOLOv3 结构配置以及各类训练算法的超参数。

因为原始 YOLOv3 的结构并不适用于检测图片中的较小物体,所以并不适合检测单个交通标志。为了解决这个问题,笔者修改了 YOLOv3 的模型结构。由于原始训练数据是尺寸为 2048×2048 的高分辨率图像,把 YOLOv3 输入层的尺寸从 416×416 放大至 608×608 ,并把训练图像向下采样,从 2048×2048 缩小到 608×608 。这样可以极大地减少高分辨率图片在训练中特征信息的损失,且也不会由于模型增大而过多增加训练所需要的时长。

由于目标类型的数量从原始默认的 80 个更改为了只有 1 个类型（禁止交通标识类型），在 YOLOv3 输出层“Y1, Y2, Y3”（如图 2 所示）之前的 3 个卷积层中，每层卷积核的数量也需要从 255 个修改为 18 个。修改完成后，将模型训练的输入批处理大小设定为 64 张图片，学习率（learning rate）设置为 0.001。使用 Darknet 工具的“detector train”命令对模型进行训练。YOLOv3 模型将根据预测结果输出训练的交叉熵（cross entropy loss）损失，并使用 ADAM 优化器对其神经网络的权重进行优化。

3 训练结果评估

该 YOLOv3 模型的损失在 50 000 个迭代后收敛，平均损失为 0.342。使用 20 000 张独立于训练数据集的测试图片给 YOLOv3 模型进行禁止交通标志检测，并计算出平均精度（mean Average Precision, mAP），对其预测结果进行评估。mAP 是评估检测效果的重要指标，AP 指的是在 11 个召回率等分候选点上的精度，而 mAP 是所有目标类别的 AP 平均值。

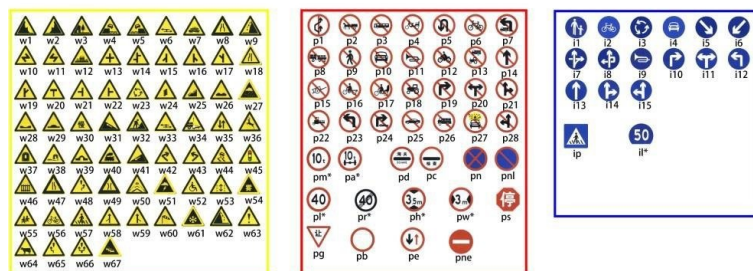


图 4 TT100K 数据集中的交通标志标签示例

Fig. 4 Examples of traffic sign and label in TT100K data set

在预测阈值设为 0.25 时，本文所使用模型的平均精确率以及平均召回率可以分别达到 80% 和 85 %。图 5 为部分 TT100K 测试数据集的检测输出示例。

4 总结和展望

尽管禁令交通标志的目标相对于高清训练图像本身较小，但是 YOLOv3 模型还是达到了良好的 mAP。未来通过以下修改，YOLOv3 模型的结果可能会得到极大改善：

- 由于要检测的所有对象都相对较小，在计算 YOLOv3 检测框的锚点时，可以不必使用大尺寸的锚点，而全部使用中尺寸以及小尺寸的锚点；
- 可以将模型的输入分辨率从 608×608 增加到 1088×1088 。尽管这将导致更长的训练时间，但将有助于提高高分辨率图像中小物体的检测精度。

尽管在现阶段，使用候选区域网络（RPN）模型、如 R-FCN、R-CNN 等，在检测图片中的小物体时可能可以比 YOLOv3 达到更高的精度，但是 YOLOv3 模型具有体积更小、检测速度更快的特点。随着此类模型的不断改善，在未来的智能辅助驾驶系统中，YOLO 模型可以更好应用于实时的交通标志以及其他目标对象的检测。

参考文献

- [1] S. Houben, J. Stallkamp, J. Salmen, et al. Detection of Traffic Signs in



图 5 模型测试输出示例

Fig. 5 Examples of model test output

Real - World Images: The German Traffic Sign Detection Benchmark[C] // 2013 Int. Jt. Conf. Neural Networks: 1 - 8.

[2] M. Mathias, R. Timofte, R. Benenson, et al. Traffic sign recognition 2014 — How far are we from the solution[C] // 2013 Int. Jt. Conf. Neural Networks: 1 - 8.

[3] L. Huang. Chinese Traffic Sign Database [EB / OL]. <https://www.nlpr.ia.ac.cn/pal/trafficedata/detection.html>, 2015.

[4] S. Zhang, Z. Zhu, X. Huang, et al. Traffic - Sign Detection and Classification in the Wild[C] // Proceedings of the IEEE conference on computer vision and pattern recognition: 2110 - 2118.

[5] R. Girshick. Fast R - CNN [C] // International

Conference on Computer Vision, ICCV 2015: 1440 - 1448.

[6] A. Krizhevsky, I. Sutskever, G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks[J]. Adv. Neural Inf. Process. Syst., 2012 (12): 1 - 9.

[7] J. Redmon, A. Farhadi. YOLOv3: An Incremental Improvement [EB / OL]. <http://pjreddie.com/media/files/papers/YOLOv3.pdf>.

[8] J. Redmon. Darknet: Open Source Neural Networks in C [EB / OL]. <https://pjreddie.com/darknet>, 2013.

2019 - 12 - 09 来稿

2020 - 06 - 28 修回

新书推荐

——《民用建筑电气设计标准实施指南》出版发行

根据中华人民共和国住房和城乡建设部 2019 年第 314 号公告, 由主编单位中国建筑东北设计研究院有限公司会同中国建筑标准设计研究院有限公司、北京市建筑设计研究院有限公司、中国建筑标准设计研究院有限公司、华东建筑设计研究院有限公司、上海建筑设计研究院有限公司、天津市建筑设计院、中国建筑西南设计研究院有限公司、中国建筑西北设计研究院有限公司、中南建筑设计院股份有限公司、悉地(北京)国际建筑设计顾问有限公司、哈尔滨工业大学、广东省建筑设计研究院、福建省建筑设计研究院有限公司等 23 家单位共同编制的国家标准 GB 51348 - 2019《民用建筑电气设计标准》将于 2020 年 8 月 1 日实施。

国标 GB 51348 - 2019 是根据住建部《关于印发

《2013 年工程建设标准规范制订、修订计划》的通知》(建标[2013] 6 号)的要求编制。标准编制组经广泛调查研究, 认真总结实践经验, 参考有关国际标准和国外先进标准, 并在广泛征求意见的基础上, 对《民用建筑电气设计规范》(JGJ 16 - 2008)进行了修订。新修订的国标 GB 51348 - 2019 与修订前相比较, 在架构和主要技术内容方面均有较大变化。

为配合国家标准 GB 51348 - 2019 宣贯实施, 标准主编单位和编制组组织编写单位人员历时数年编著了《民用建筑电气设计标准实施指南》。全书共 26 章, 对应国标 GB 51348 - 2019 各章的重点条文。为方便和引导读者阅读, 每章的开始增加了修订简介, 介绍本章主要修订内容。在注释中对条文进行深入的分析, 阐明立条依据或对条文说明进行引申和扩展, 同时配有大量的设计举例, 使读者能更好地理解、掌握和执行国家标准。

该书是一部配合国标宣贯的良好教材, 其技术内容涵盖民用建筑电气的各个领域, 吸取了国内外在该领域里的新技术、新设备、新材料, 反映了该领域技术新进展, 是从事民用建筑电气设计必备的指导用书, 也可作为大专院校师生、施工人员和房地产开发等人员的参考书。

该书由四川科学技术出版社出版, 建筑电气杂志社独家发行。正 16 开本精装, 字数 1 034 千字, 定价 236 元/本。



《民用建筑电气设计标准实施指南》出版发行

中国建筑东北设计研究院有限公司
《民用建筑电气设计标准》编制组
建筑电气杂志社