

文章编号: 2095-2163(2022)01-0065-04

中图分类号: V323.19

文献标志码: A

基于图片扰动的自动驾驶测试数据生成方法

郑雅婷, 刘亚男, 周亚辉, 韦晓梦, 房丽婷

(北京信息科技大学 计算机学院, 北京 100101)

摘要: 自动驾驶是一个复杂的、系统性的研究项目, 在实际的测试过程中, 为了确保安全, 虚拟仿真测试成为自动驾驶测试的关键, 在检测驾驶模型是否可以做出正确的判断, 减少安全事故的发生等方面有着不可替代的作用。本文在虚拟仿真测试的基础上, 提出有关图片扰动的自动驾驶测试的数据生成方法, 通过对检测数据的核算来测试模型是否可以做出缺陷判断。数据显示, 本文使用到的模型的缺陷检测会受到相应测试环境的影响。

关键词: 自动驾驶; 虚拟仿真; 图片扰动

Automatic driving test data generation method based on image disturbance

ZHENG Yating, LIU Yanan, ZHOU Yahui, WEI Xiaomeng, FANG Liting

(School of Computer Science, Beijing Information Science and Technology University, Beijing 100101, China)

[Abstract] Automatic driving is a complex and systematic research project. In the actual testing process, in order to ensure safety, virtual simulation test becomes the key of automatic driving test, which plays an irreplaceable role in detecting whether the driving model can make correct judgment and reduce the occurrence of safety accidents. On the basis of virtual simulation test, this paper proposes a data generation method for automatic driving test of image disturbance, and tests whether the model can make defect judgment by accounting the test data. The data show that the defect detection of the model used in this paper will be affected by the corresponding test environment.

[Key words] autonomous driving; virtual simulation; image disturbance

0 引言

自动驾驶技术在近年来得以迅速发展, 其应用需要严格的测试流程, 以保证可靠性。目前有许多自动驾驶的测试方式, 如: 基于用例的测试方式、基于场景的测试方式以及基于公共道路测试方式等^[1]。基于实际道路的测试除了需要解决测试材料以及相应场景构建的问题之外, 更重要的是要确保测试过程的安全, 驾驶模型的部件设计不够合理将会造成交通事故。如: 2016年谷歌的自动驾驶汽车在加州山景城地区的事故; 特斯拉2019年自动驾驶电动车测试中, 因电动车的摄像头设计不够合理, 导致其对车辆的识别出现问题; 又因雷达的布局不够合理, 无法准确识别周围车辆, 而做出错误的判断, 使得该电动车直接撞向卡车^[2]。

在这样的背景下, 虚拟仿真测试成为自动驾驶检测过程重要方法, 许多虚拟仿真技术相继出现, 其中关于检测在不同环境包括角落情况驾驶的测试技术, 如: DeepTest, DeepRoad 等都已经取得了重大的

成就。但目前的技术仍有许多不足和缺陷, 比如拟合不足, 许多图片需要手动标注, 无法检测到角落以及边界情况等。

本文在调研了目前的自动驾驶场景测试的情况和问题后, 提出了有关数据的处理方法, 围绕数据变化前后大小的计算和对变化前后数据的相关程度的计算, 对自动驾驶的模型的性能进行检测。对相应的驾驶场景进行不同天气的模拟, 通过比对的结果来验证自动驾驶模型的检测性能, 如评判模型是否可以检测到缺陷等。

1 深度神经网络(DNN)

深度神经网络除了标记训练数据的功能外, 还可以从输入的数据中提取相关特征。DNN 每层包含大量的神经元, 每层神经元之间相互连接, 每一层会将处理得到的信息进行反馈。DNN 通常包含3层, 即输入层、隐藏层和输出层, DNN 的结构如图1所示。

自动驾驶中通常都会应用到卷积神经网络

基金项目: 北京信息科技大学2021年大学生创新创业训练计划项目(5102110805)。

作者简介: 郑雅婷(2001-)女, 本科生, 主要研究方向: 计算机应用。

收稿日期: 2021-09-19

(CNN)和循环神经网络(RNN),前者是由卷积层、池化层和全连接层构成,通常用于图像的处理^[3]。本文中的图像处理涉及到卷积变换,而循环神经网络通常可以用于反馈相关的输入信息^[4]。另外,循环神经网络还可以将数据向下一层反馈和向上一层反馈^[5]。

DNN的局限性也是非常明显,对数据的标记就是限制DNN发展的一个重要因素,主要表现在标记的不够完善,无法对大量的数据进行手动的一一标注;另外,有些标注,因为技术性要求过高,对相关专家来说都难以完成^[6]。另外一个局限主要表现在神经元覆盖率低,许多的测试输入常常得到错误的结果。

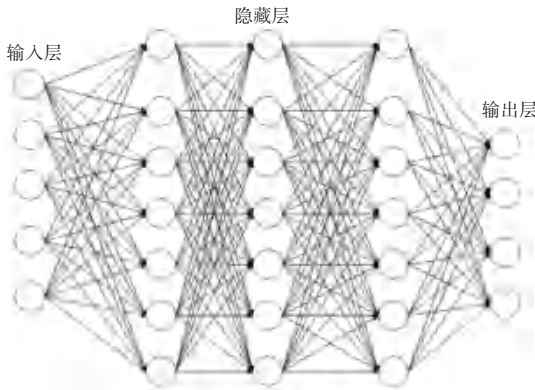


图1 DNN基本结构

Fig. 1 Basic structure of DNN

2 实验流程及关键步骤

2.1 自动驾驶测试场景构建

自动驾驶基于场景的测试中除了对环境的测试,还包含对相关车辆模型的测试^[7]。自动驾驶测试中包含的环境因素是变化的,从某种意义上说,相关场景的选取具有无限性。自动驾驶场景的数据通常来源于测试数据、事故数据以及专家经验等^[8]。基于相关数据以及场景要素设计相应的驾驶测试,测试主要依据真实度、成本、效率、准确性等方面做出评估。

2.2 实验流程

目前的虚拟测试方式主要包括模型在环测试、硬件在环测试和车辆在环测试等^[1]。相关的场景构建主要利用数据,通过统计模型分析出相应的场景案例。因为在现实生活中,危险事故的发生概率较小,因而数据的统计困难。另外,驾驶时是否会发生危险也难以确定。

首先从相应数据集中截取典型的驾驶场景,如:

右转弯的驾驶情况,为不同的场景添加不同的天气背景,基于驾驶模型测试和记录数据,对数据进行核算,分析和对比核算结果。实验过程流程图如图2所示。

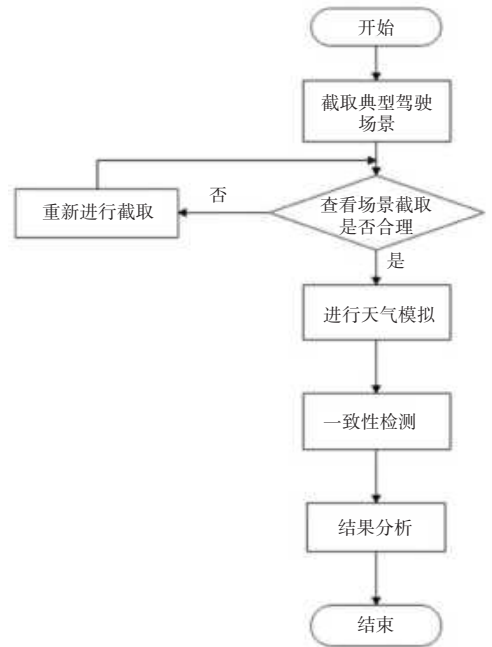


图2 实验过程流程图

Fig. 2 The Experimental process of flow chart

2.3 关键步骤

2.3.1 图像合成

本文使用Photoshop实现对不同驾驶场景下的天气情况模拟,当然也可以利用OpenCV编写相应的天气脚本来实现批处理,但是后者实现的图像真实性不及前者,因此本文使用的测试场景是手动合成的,以此来提升检测到的数据的可靠性。

2.3.2 一致性检测

模型需要检测模拟前后的驾驶场景来获得相应的数据,本文采用求两组数据的差值绝对值和相关系数的方法来评估原数据与变换后的数据之间的差距和相关程度,通过这两种方法的组合来对结果进行分析。

由于每组驾驶场景对应4种不同的天气,包括无天气模拟的一组和有天气模拟的3组,为了便于叙述,将每组驾驶场景的4组数据分别命名为0、1、2、3。为了评估模拟前后检测到的数据之间的差距,每个驾驶场景下的数据组1、2、3都需要利用公式(1)求得与数据组0的差值绝对值 f , f 越小,说明天气变化后对模型决策的影响越小,模型对相应驾驶场景下的路况识别程度越好,但这个绝对值又不希望为0,即希望看到模型可以检测到缺陷和变化。

这样,通过公式(1)会得到每个驾驶场景的 3 组差值绝对值,计算每组差值绝对值的平均值,如此得到 30 个相应的平均值。

$$f = |\theta_1 - \theta_2| \quad (1)$$

其中, θ_1 表示模拟后的数据组中的一个转向角, θ_2 表示对应的原驾驶场景数据组 0 中与 θ_1 对应的图片的转向角。

为了评估相关程度以及变化趋势,本文使用了相关系数 r , 公式(2):

$$r = \frac{\sum(\theta_{1i} - \bar{\theta}_1)(\theta_{2i} - \bar{\theta}_2)}{\sqrt{\sum(\theta_{1i} - \bar{\theta}_1)^2} \sqrt{\sum(\theta_{2i} - \bar{\theta}_2)^2}} \quad (2)$$

其中, θ_{1i} 为模拟后的数据组中对应的第 i 个转向角; $\bar{\theta}_1$ 为模拟后的数据组所对应的平均值; θ_{2i} 为数据组 0 中与 θ_{1i} 对应的第 i 个转向角; $\bar{\theta}_2$ 为数据组 0 所对应的平均值。

相关系数 r 的取值范围在 $[-1, 1]$ 之间, r 为正数表示正相关, r 为负数表示负相关, 值越大表示相关程度越高。本文使用 $|r|$ 来表示相关程度的大小, 即 $|r|$ 值越大相关程度就越高, 两组数据的联系紧密。利用同样方法可以获得每个驾驶场景下的数据组 1, 2, 3 与数据组 0 之间的相关系数。

3 实验及评估

3.1 测试数据及模型

本文使用 Udacity 数据集, 该数据集提供了上千张连续帧图片, 由于数据集中的图片地点在公路, 很多驾驶路况相似, 因此, 本文截取了一部分典型的驾驶场景, 分别为右转、左转、转弯前方有一颗树、直行、右转之后再左转、右转的时候左侧出现了行人、左转前方有障碍物、前方有车辆驶入主路、左转提示、分岔路口, 共 10 种场景。另外, Udacity 提供 5 个模型, 其中 Autumn、Chauffeur 和 Rwrightman 的训练模型是公开的, 可以直接评估图片对应的数据, 本文选用 Rwrightman 模型进行评估。

本文通过模型检测截取的场景图片来获取转向角数据。在实验的过程中不断增加对应驾驶场景所包含的图片数量, 同时手动核对转向角的情况, 以确保模型做出的决策是正确的。如左转驾驶场景第一次截取后包含的图片在 30 张左右, 此时场景的检测数据就会出现错误, 这是由于该场景包含的图片数量不足, 此时模型只能识别左转的一部分路况, 而模型每个数据的产生都受上一个数据的影响, 不完整的驾驶场景的检测结果就会出现错误, 当左转场景

对应的图片被逐个添加完整时, 数据才变得符合实际。

3.2 场景模拟

本文合成的图片主要是日常生活中常见的沙尘、雾以及下雨等天气。由于需要处理的图片数量多, 处理单个图片需要的时间较长, 因此在模拟单个图片的天气效果后, 利用 Adobe Photoshop 录制相应的处理流程, 利用该流程处理本驾驶场景下的所有图片。由于本文使用的图片是连续帧的, 截取的驾驶场景中的图片连续显示时效果类似于视频, 为使图片联合起来有动感, 需要在图片对应的部位依据相应的环境背景适时做出调整, 天气效果图如图 3 所示。



图 3 原图及添加沙尘、下雨及大雾天气后的图片

Fig. 3 Original picture and pictures after adding dust weather, rain and fog weather

3.3 数据分析

统计 30 组差值绝对值对应的平均值, 都在 $[0, 0.32]$ 之内, 因此阈值 ε 的设置可参考每组差值绝对值的平均值, 设置为 $[0.01, 0.32]$ 内, 初值为 0.01, 每次增长 0.01, 用来查看每组驾驶场景的单个数据 f 是否小于等于一个阈值 ε , 见表 1。规定: 满足 $f \leq \varepsilon$ 为正确行为; 否则, 为错误行为。通过表 1 可知, 模型对变化后的图片的识别存在一定差异, 在 ε 非常小的时候, 相关的错误行为较多; 随着 ε 不断增大, 正确行为的数量不断上升, 在 0.3 左右时几乎没有错误行为产生。另外, 在 ε 一定时, 不同的天气情况下错误行为的数量不同, 沙尘天气的错误行为数量最高, 大雾天气的最低, 这也说明不同的天气对模型的影响不相同, 与实际相符, 即在不同的天气驾驶相同的车辆在相同的地点所作出的反应是有差异

的。

计算所得相关系数 $|r|$,见表2,沙尘、雾,下雨天气情况下, $|r|$ 的值在区间 $[0.9,1]$ 的驾驶场景个数分别达到对应天气模拟效果总场景个数的50%,80%,50%, $|r|$ 的值在 $[0.8,1]$ 达到70%,90%和70%,说明模拟前后检测到的数据之间的相关性非常高,变化趋势也非常的相似,即模型对不同环境下的交通场景的预测基本与原场景一致。另外,通过分析表2数据也可以发现,不同天气情况对模型的检测也有不同程度的影响。

表1 对应天气模拟正确与错误行为数量统计

Tab. 1 Statistics of correct and wrong behaviors under corresponding weather simulation effect

ε	沙尘天气		大雾天气		下雨天气	
	正确行为	错误行为	正确行为	错误行为	正确行为	错误行为
0.01	37	402	95	344	63	376
0.02	68	371	176	263	108	331
0.03	98	341	241	198	159	280
0.04	138	301	291	148	196	243
0.05	178	261	313	126	221	218
0.06	207	232	334	105	241	198
0.07	227	212	353	86	253	186
0.08	249	190	368	71	259	180
0.09	274	165	378	61	61	391
0.1	305	134	382	57	57	392
0.11	316	123	390	49	49	392
0.12	331	108	395	44	44	393
0.13	336	103	402	37	37	393
0.14	347	92	407	32	32	393
0.15	352	87	411	28	28	394
0.16	362	77	414	25	25	394
0.17	371	68	417	22	329	110
0.18	379	60	420	19	332	107
0.19	387	52	420	19	339	100
0.20	394	45	421	18	340	99
0.21	401	38	421	18	346	93
0.22	408	31	428	11	354	85
0.23	412	27	434	5	358	81
0.24	415	24	439	0	363	76
0.25	419	20	437	2	371	68
0.26	422	17	437	2	375	64
0.27	427	12	437	2	382	57
0.28	431	8	437	2	393	46
0.29	436	3	437	2	400	39
0.30	436	3	437	2	401	38
0.31	437	2	437	2	403	36
0.32	437	2	437	2	390	49

表2 利用相关系数绝对值进行场景分类

Tab. 2 Scene classification using absolute value of correlation coefficient

$ r $	统计相关的场景数量		
	沙尘天气	大雾天气	下雨天气
$0.3 \leq r < 0.7$	3	0	3
$0.7 \leq r < 0.8$	0	1	0
$0.8 \leq r < 0.9$	2	1	2
$0.9 \leq r \leq 1$	5	8	5
合计	10	10	10

4 结束语

本文选用数据集 Udacity,截取其中一些典型的驾驶场景,对不同天气效果的模拟,提出一种基于图片扰动的自动驾驶测试数据生成方法。通过驾驶模型 Rwrightman 对模拟前后的驾驶场景进行测试,对不同天气效果模拟前后获得的数据进行核算和分析,发现不同天气背景对模型的检测有着不同的影响,能够在不同的环境背景中检测到变化,做出缺陷判断。

参考文献

- [1] 余卓平,邢星宇,陈君毅. 自动驾驶汽车测试技术与应用进展[J]. 同济大学学报(自然科学版),2019,47(4):540-547.
- [2] 冯洋,夏志龙,郭安,等. 自动驾驶软件测试技术研究综述[J]. 中国图象图形学报,2021,26(1):13-27.
- [3] Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks[C]//In Advances in neural information processing systems, 2012:1097-1105.
- [4] TIAN Y, PEI K, JANA S, et al. DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars[J]. 2017.
- [5] SAK H, SENIOR A, BEAUFAYS F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling[J]. computer science, 2014.
- [6] PEI K, CAO Y, YANG J, et al. DEEPPXPLOR: Automated Whitebox Testing of Deep Learning Systems[J]. Mobile Computing and Communications Review, 2018, 22(3):36-38.
- [7] 任秉韬,邓伟文,白雪松,等. 面向智能驾驶测试的仿真场景构建技术综述[J]. 中国图象图形学报,2021,26(1):1-12
- [8] 朱冰,张培兴,赵健,等. 基于场景的自动驾驶汽车虚拟测试研究进展[J]. 中国公路学报,2019,32(6):1-19.