arXiv:2301.05277v1 [cs.HC] 12 Jan 2023

DriCon：面向意外驾驶事件的设备端即时情境表征

Debasree Das、Sandip Chakraborty、Bivas Mitra 计算机科学与工程系，印度理工学院卡拉格浦尔分校，印度 721302 电子邮箱：{debasreedas1994,sandipchkraborty, bivasmitra}@gmail.com

摘要——驾驶是一项复杂任务，受到各种空间对象及其时间交互的影响。因此，驾驶行为突然波动，可能是由于驾驶技能不足，也可能是行人移动、周围车辆的动作等各种道路空间因素所致。因此，要确保道路安全，就有必要即时理解驾驶行为退化背后的情境。本文开发了一种名为 DriCon 的系统，该系统利用从仪表盘安装的边缘设备获取的信息，从多种道路空间因素以及车内驾驶操作中提取微事件，以理解相关情境。DriCon 使用实时内部测试平台和目前公开可用的最大驾驶数据集，针对意外驾驶事件生成易于人类理解的解释。此外，与现有驾驶行为表征技术相比，DriCon 在超过 50 小时的驾驶数据上取得了 80% 的相似度，提供了更深入的洞察。

索引词——驾驶行为、空间事件、情境分析

一、引言

随着交通参与者数量增加，过去几年里道路交通事故显著上升。根据世界卫生组织（WHO）的数据 [1]，，损失不仅限于人员伤亡，也会影响国家的国内生产总值。官方报告的道路交通碰撞事故大多依据宏观情境进行调查，例如车辆速度、道路状况等。仔细审视这些宏观情境，可以发现一系列导致此类伤亡的微事件。例如，假设一名驾驶员在非拥堵道路上驾驶时撞上道路隔离带并受伤。从宏观角度看，我们可能会认为这是由于驾驶员驾驶技能生疏或车辆速度过快所致。然而，也有可能是某些意外障碍物（如穿越行人/动物）突然出现在视线之外。驾驶员为避免与它们相撞而减速时偏离了车道。因此，记录这些微事件对于查明此类事故背后的原因至关重要。这些情境信息或微事件因此能够帮助汽车保险公司或网约车公司等各类利益相关方分析其驾驶员的道路驾驶行为。有趣的是，网约车公司可以根据驾驶员对这类情境的处理方式，决定对其进行处罚或奖励，并采取应对措施以避免事故。

一种简单的提取上下文信息的方法是分析交通视频。值得注意的是，闭路电视摄像机 [2]只能捕捉与行驶中的车辆相关事件的静态快照。

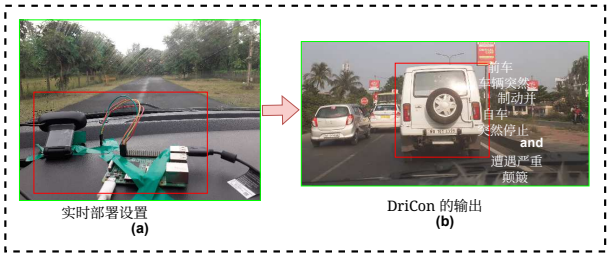


图 1：DriCon：硬件组件，以及车辆遭遇严重冲击时运行的实例

现有研究 [2], [3] 使用行车记录仪视频和IMU传感器数据，对事故进行人工或半自动调查。需要注意的是，人工干预容易出错且劳动强度大，成本也更高。当多个事件共同导致事故时，情况会更加复杂。例如，假设前车为避免与行人发生碰撞，或因遇到即将变红的交通信号灯而突然制动。因此，自车不得不突然减速，导致意外停车背后存在两步因果事件链。因此，识别交通对象之间的时空交互，对于确定此类事件的根本原因至关重要。

重要的是，实时理解驾驶行为退化背后的情境并非易事，还面临多重挑战。首先，这需要持续监测驾驶员的驾驶行为，并全面掌握各种道路空间微事件。高端车辆使用激光雷达、雷达等设备感知驾驶员和周围环境 [4]，[5]；然而，由于利润率较低，网约车公司不愿投资此类高端车辆。其次，根据所采取的驾驶操作，依据车辆与道路空间物体的交互，对微事件进行时间上的关联，是一项重大的研究挑战。例如，如果某天观测到恶劣雪天气候，其对交通流动的影响可能持续到次日。相比之下，鲁莽驾驶只会影响周围少数几辆车，几分钟后便不会再产生显著的时间影响。事件在时间上的影响会因事件的类型和发生空间而异。第三，周围物体的空间位置会影响驾驶操作。具体而言，除了时间依赖关系，自车与周围物体之间的距离也至关重要。

例如，远处的行人可能会以较高速度过马路，同时保持安全距离；但如果与车辆的距离很近，后果就可能致命。现有研究 [6], [7] 曾尝试通过惯性测量单元和视频分析识别危险驾驶行为，例如车辆与行人交互；然而，这些研究未能捕捉此类时间尺度变化或周围物体之间的空间依赖关系。第四，在边缘设备（例如行车记录仪）上实时识别情境，对于提供即时反馈至关重要。但要在资源受限的边缘设备上，基于多模态数据部署用于情境表征与分析的系统，并非易事。

为应对这些挑战，我们提出了 DriCon。该系统开发了一款安装在车辆仪表板上的智能行车记录仪，用于表征微事件，从而为驾驶员及其他利益相关者（如出租车公司）提供即时情境反馈。它通过惯性测量单元和全球定位系统传感器感知自车所执行的驾驶操作。此外，设备本身搭载的前置摄像头用于分析各种道路微事件与所执行驾驶操作之间的关系。这使系统能够在每辆车上独立运行，并兼具低成本、轻量化的特点。图 1 展示了安装在车辆上的系统硬件组件，以及一个示例场景：DriCon 对车辆突然颠簸的原因生成实时情境解释。总之，本文的贡献如下。

(1) 通过先导研究推动微事件表征：我们在伯克利深度驾驶（BDD）数据集上开展了一系列先导研究。该数据集是互联网上规模最大的公开驾驶数据集（截至 2023年1月16日） [8], 。我们旨在考察驾驶行为如何随道路类型、一天中的时间、星期等因素而变化，并指出导致驾驶操作发生突变的时空微事件。

（2）设计一种可由人类解释的轻量级因果模型：DriCon 的开发依赖于（i）惯性测量单元和全球定位系统数据来推断驾驶操作，以及（ii）目标检测模型和透视变换 [9] 来检测周围物体及其动作，从而捕捉各种空间微事件。随后，我们识别行程中驾驶行为恶化时的时空情境。最后，我们采用自组织映射（SOM），这是一种轻量而有效的因果模型，用于捕捉特征之间的时空依赖关系，以学习情境并生成便于人类理解的解释。

（3）边缘端部署：我们将 DriCon 的完整架构部署在一台嵌入了前置摄像头、惯性测量单元和 GPS传感器的树莓派3设备上（图 1）。为此，我们使惯性测量单元和视觉数据处理都更加轻量，并满足低延迟要求。随后，预训练模型根据正在进行的驾驶行程生成建议，使其能够实时高效地运行，以进行即时因果推断。

（4）在实时系统部署和 BDD 数据集上评估 DriCon：我们在自有的实时部署系统以及 BDD 数据集 [8] （使用标注数据 [10]）上评估 DriCon，二者分别包含 33 小时和 17 小时的驾驶数据。对于模型返回的排名前 3 和前 5 的特征，在正确识别导致驾驶行为变化的微事件方面，推导出的因果特征与真实因果特征之间的平均相似度分别达到 70% 和 80%。值得注意的是，在大多数情况下，我们观察到推导出的特征与观测到的驾驶行为之间具有良好的因果关系（以平均处理效应衡量）。此外，我们在边缘设备上开展了不同的资源消耗基准测试研究，以更深入地了解所提出的模型。

II. 相关工作

已有多项研究致力于理解道路交通及其对道路事故死亡的影响。早期研究主要集中于基于交通监控的技术，以预防道路交通事故。例如，美国国家公路交通安全管理局（NHTSA） [3] 记录了致命事故案例的统计数据；TUAT [2] 自 2005 年起一直收集出租车视频记录和驾驶员面部图像，将受伤案例划分为多个类别，并进行驾驶行为估计。在印度，关于交通伤害成因的信息来源是当地交通警察 [11]。相比之下， [12], [13] 等研究分别从街景图像和原始轨迹流中即时学习犯罪类型和驾驶员移动模式。除了利用视频和众包信息外，还有多项研究 [14], [15] 通过利用惯性测量单元（IMU）和全球定位系统（GPS）数据来检测异常驾驶行为。为预防致命事故，作者 [16]–[18] 尝试在观察到危险驾驶特征时向驾驶员发出警报，例如车道偏离或表明拥堵的突然减速。然而，这些研究尚未探讨邻近车辆或其他周围因素对各种驾驶操作的影响。

在复杂的城市交通中，自车与行人等其他障碍物之间的交互，以及恶劣天气等因素，往往会影响车辆的运动，进而影响驾驶行为。现有研究 [19] 表明，道路类别、无信号灯人行横道和车速往往会导致行人对是否过马路意见不一，从而引发道路交通死亡事故。一项更深入的研究 [20], [21]聚焦于自动驾驶的因果关系分析，但难以实现实时部署。此外，这些研究仅使用了有限的一组驾驶操作，例如只考虑速度变化。尤其是，由于驾驶数据存在高度差异性，且交通对象与驾驶操作之间存在虚假相关 [22] ，因果推断颇具挑战。现有研究仅考虑静态道路属性，或依赖联网道路系统提供的单模态或多模态数据，因此研究范围有限。若未连接到该系统，这些方法便无法应用于单车的实时部署

系统。相比之下，利用车载传感器提供的多模态数据，可以高效地刻画意外驾驶行为波动背后的连续动态情境。DriCon 正是朝这一方向开发的系统。

III. 动机

在理想情况下，在相同驾驶环境下，两辆车辆往往会采取相似的驾驶操作；但现实并非如此。驾驶行为会因驾驶员独特的技能组合而有所不同，并受到各种道路事件影响，例如其他大型和小型车辆的移动、行人的移动、道路拥堵、前车采取的驾驶操作等，我们将这些事件称为空间微事件，简称微事件。本节将开展一组初步研究，以回答以下问题。(a) 驾驶员的驾驶行为是否表现出时空变化？(b) 一次行程中发生的所有微事件是否会以相似的方式影响驾驶行为？(c) 一系列相互依赖的微事件是否会共同影响驾驶行为？随后，我们分析名为伯克利深度驾驶数据集（BDD）的公开开源驾驶数据集 [8]，以回答这些问题，并阐明不同微事件对驾驶行为的影响。该数据集包含由10,000名志愿驾驶员从两国——美国和以色列——18个城市众包收集的100,000次行程。对于每次驾驶行程中的每5秒，数据集都按照李克特量表进行了驾驶评分，评分范围为1分（驾驶最差）至5分（驾驶最佳）。

A. 驾驶行为在空间和时间上的变化

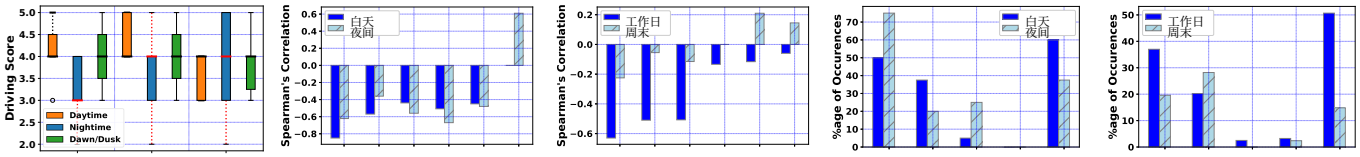
我们首先检验道路上的驾驶行为是否呈现时空变化。为此，我们在 BDD 数据集中改变两个参数——将道路类型作为空间参数（例如“高速公路”、“城市街道”、“住宅区道路”），将一天中的时间作为时间参数（例如“白天”、“夜间”、“黎明/黄昏”）。在这项先导研究中，我们构建了 9 个组，每组包含 30 次行程，总计 270 次；每个组中的行程均从 BDD 数据集中随机选取。我们绘制了每个组所有行程的驾驶评分分布。从图 2(a) 可以明显看出，评分分布会发生变化：(a) 对于同一种道路类型，在一天中的不同时间有所变化；(b) 对于不同道路类型，在一天中的任一给定时间也有所变化（ $p < 0.05$ 反映了其统计显著性）。下文将探讨各种微事件在驾驶行为变化背后所发挥的作用。

B. 空间微事件的作用

接下来，我们考察各种道路微事件——这些事件由其他空间对象的移动所表征，例如“汽车”“行人”“卡车”“公交车”“摩托车”“自行车”等——在一天中的不同时段是否会以相同方式影响驾驶员的驾驶行为。我们从 BDD 数据集中人工挑选了 30 次行程及其标注的驾驶评分，涵盖白天和夜间，并据此开展研究。我们使用现有的目标检测算法 [23] 从行程期间拍摄的视频中提取空间对象的数量（例如计数），并计算每个物体在 5 秒时间窗口内的平均计数。因此，对于白天和夜间，我们分别得到两种时间序列分布：(a) 每个时间窗口内，从行程视频中捕获的各类道路空间物体的数量；(b) 这些时间窗口内标注的驾驶评分。接下来，我们分别计算白天和夜间这两种分布之间的斯皮尔曼相关系数（SCC）。从图2(b)可以推断，几乎所有道路空间物体都会对驾驶行为产生负面影响（表现为负相关）。白天，汽车和行人对驾驶评分的影响最为显著；而夜间，卡车和公交车以及汽车都会影响驾驶行为，这是因为卡车等重型车辆主要在夜间行驶。不过，摩托车和自行车等轻型车辆的影响并不显著，这是由于它们行驶时有专用车道。我们还在图2(c)中进一步考察了这一现象，对比工作日与周末的情况。我们根据 BDD 数据集中已有的时间戳提取星期信息，并将星期一至星期五的 30 次行程归为工作日，将星期六至星期日的 30 次行程归为周末。从图2(c)可以看出，在一周的前几天，汽车、行人和卡车会对驾驶行为产生负面影响，而周末的影响较小。因此，我们得出结论：不同的道路物体会对驾驶行为产生不同的时间效应。

C. 导致突发驾驶操作的微事件：以急停为例

最后，我们探究多个相互依赖的微事件是否会导致某种特定的驾驶操作，从而降低驾驶行为表现。为此，我们选择急停作为驾驶操作，并从全球定位系统和惯性测量单元数据中提取急停事件（即停车时产生严重颠簸的情况 [24]）。我们针对白天、夜间、工作日和周末四种场景，各选取30次行程。对于每种场景，我们提取发生急停的实例，并记录这些实例中观察到的相应微事件。具体而言，我们采用成熟方法 [10], [23]，提取以下微事件是否出现：红色交通信号灯、行人移动、重型车辆（如卡车和公交车）、轻型车辆（如摩托车和自行车），以及前车的制动行为（作为同伴车辆操作）。我们统计上述四种场景中所有行程里各类微事件出现的累计计数，以及急停的次数。从图2(d)和(e)可以看出，夜间和工作日发生的急停大多由红色交通信号灯、同伴车辆操作和重型车辆引起。因此，我们认为，多个道路微事件（例如重型车辆在夜间鲁莽行驶）会迫使即使是优秀的驾驶员也猛踩制动器，采取不安全操作。



(a) (b) (c) (d) (e)

图 2：(a) 驾驶行为随道路类型和一天中的时间的变化；(b)-(c) 空间微事件在不同 (i) 一天中的时间、(ii) 星期对驾驶评分的影响；(d)-(e) 在不同 (i) 一天中的时间、(ii) 星期观察到的急停背后的影响因素

IV. 问题陈述与系统概述

A. 问题陈述

设 $\mathcal{F}_M$ 表示驾驶操作的集合， $\mathcal{F}_S$ 表示空间微事件的集合。 $\mathbb{F}^i$ 表示与行程 i期间采取的驾驶操作及遇到的道路空间微事件相对应的时间表示的特征变量。设 $\mathcal{R}_T^i$ 为行程 i期间时间 T 的驾驶评分。我们希望检查发生的事件，并在$|\mathcal{R}_T^i - \hat{}\mathcal{R}_{T-1}^i| > \epsilon$ 时表示特征值 $\mathbb{F}^i$（ϵ 为超参数，我们设定 $\epsilon = 1$），以反映驾驶行为的波动。这里，$\hat{}\mathcal{R}_{T-1}^i = \lceil \text{mean}([\mathcal{R}_1^i, \mathcal{R}_{T-1}^i]) \rceil$ 表示截至$T-1$的平均驾驶行为。系统的输出是对 $\{\mathcal{F}_M^i, \mathcal{F}_S^i\}$的刻画，用以判断驾驶行为的波动究竟仅由驾驶操作导致（$\mathcal{F}_M^i$），还是由具有空间因果关系的微事件所致（$\mathcal{F}_S^i$）。最后，我们旨在根据 $\{\mathcal{F}_M^i, \mathcal{F}_S^i\}$ 生成解释，为利益相关者提供反馈，以便进一步分析驾驶档案。

B. 特征选择

借鉴现有文献 [24], ，我们确定了一组时间戳 T 处的特征变量，用于表示自车的各种驾驶操作 $\mathcal{F}_M$ 。这些特征包括：蛇形行驶（$\mathcal{A}_T^W$）、急转向（$\mathcal{A}_T^S$）、侧滑（$\mathcal{A}_T^L$）、急停（$\mathcal{A}_T^Q$）、急转弯（$\mathcal{A}_T^U$）和严重颠簸（$\mathcal{A}_T^J$）。同样，对于以下空间微事件 $\mathcal{F}_S$，我们考虑相应的特征变量：自车与前车之间的相对速度（$\mathcal{S}_T$）和距离（$\mathcal{D}_T$）、前车的制动动作（$\mathcal{B}_T$）、自车前方周围车辆的数量（用于指示道路拥堵情况）（$\mathcal{C}_T$）、行人（$\mathcal{P}_T$）及其速度（$\mathcal{Q}_T$）、交通信号灯（$\mathcal{L}_T$）、重型车辆： {公交车和卡车}（$\mathcal{H}_T$）、道路类型（$\mathcal{G}_T$）以及天气状况（$\mathcal{W}_T$）。需要注意的是，我们根据现有文献和数据集中的观察结果，通过经验选取了这些特征；在不影响 DriCon 通用性的前提下，也可以纳入其他特征。

接下来，我们将概述系统架构。DriCon 从行车记录仪（例如，边缘设备）采集惯性测量单元、全球定位系统和视频数据，并实时刻画驾驶行为改善/恶化背后的情境。系统由三个组件构成：(a) 数据预处理和特征提取，(b) 驾驶行为改善/恶化检测，以及 (c) 可能情境识别（见图 3）。

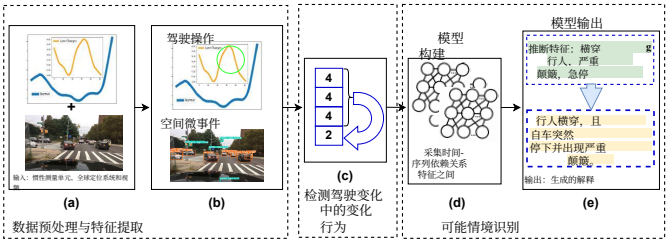


图 3：DriCon 系统流程与建模管线

C. 数据预处理与特征提取

采集到的惯性测量单元和全球定位系统传感器数据容易受到地球引力、信号衰减和大气干扰的影响而产生噪声。因此，我们采用低通滤波器消除惯性测量单元和全球定位系统数据中的此类噪声，以计算用于提取驾驶操作的惯性特征（$\mathcal{F}_M$）。接下来，在提取道路空间微事件及其动作之前，我们会对视频数据进行预处理（$\mathcal{F}_S$）。我们将采集的视频上采样或下采样至960×540p分辨率，同时确保信噪比高于 20 dB。

1）驾驶操作 - $\mathcal{F}_M$：为生成与不同驾驶操作对应的特征（$\mathcal{F}_M$），我们采用标准的加速度测量分析 [10], [24]，从惯性测量单元数据中提取蛇形行驶（$\mathcal{A}_T^W$）、急转向（$\mathcal{A}_T^S$）、侧滑（$\mathcal{A}_T^L$）、急停（$\mathcal{A}_T^Q$）、急转弯（$\mathcal{A}_T^U$）和严重颠簸（$\mathcal{A}_T^J$）的实例。

2) 空间微事件 - $\mathcal{F}_S$：接下来，我们采用最先进的、基于视频数据的目标检测算法，并根据需求进一步微调这些算法，因为开发基于视觉的算法超出了本研究的范围。我们利用在 COCO数据集 [25] 上训练的 YOLO-V3 [23] 算法，检测一部分交通对象，例如行人、汽车、公交车、卡车和交通信号灯（如 $\mathcal{F}_S$所示）。接下来，我们估算行人之间的互动、重型车辆（公交车和卡车）的出现、交通信号灯状态变化（红灯、黄灯和绿灯）以及汽车对自车驾驶行为的影响。随后，我们舍弃检测结果中置信度得分

$< 50\%$ 且边界框面积为 $< 10k$的物体，以体现远处交通对象相比近处交通对象对自车的影响微乎其微。此外，位于道路中间、从驾驶员仪表板视角清晰可见的交通对象，比左侧或右侧车道上的交通对象影响更大，因为自车会立即跟随它们。因此，我们沿水平方向将每一帧按 0.2:0.6:0.2 的比例划分为左侧、中间和右侧车道。因此，我们保留边界框坐标位于中间车道范围内的行人 $\mathcal{P}_T$、汽车和重型车辆（即 {公交车和卡车} $\mathcal{H}_T$），以及不包含车道信息的交通信号灯状态变化 $\mathcal{L}_T$（红灯、黄灯和绿灯），因为交通信号灯通常位于左侧或右侧车道。先导研究表明，行人和周围车辆的动作会显著影响自车的驾驶操作，因此，(a) 我们提取行人速度（$\mathcal{Q}_T$），并识别道路中间的横穿行人；(b) 我们计算前车的制动动作（$\mathcal{B}_T$）和拥堵情况（$\mathcal{C}_T$），并检测自车与前车之间相对速度（$\mathcal{S}_T$）和距离（$\mathcal{D}_T$）的变化。我们采用透视变换和深度学习方法 [9], [26]来推断上述信息。最后，上述流程会在每一帧上运行，视频则重采样为每秒 15 帧。

D. 驾驶行为波动的检测

DriCon 的核心是在行程中观察到驾驶行为发生变化时，捕捉各种驾驶操作与空间微事件之间的时间依赖关系。为了对驾驶行为进行运行时标注，我们采用一项已有研究 [10]，该研究通过分析驾驶操作及其他周边因素，按照李克特量表 $[1-5]$ 给出驾驶行为评分。我们将行程划分为大小为 δ 的连续且互不重叠的时间窗口，并在每个窗口结束时 $\mathcal{U}$，利用该窗口 [10]期间采集的特征值计算驾驶评分（记为 $\mathcal{R}_{\mathcal{U}}^{P}$）。为了定量监测窗口 $\mathcal{U}$内驾驶行为是否发生变化，我们比较 $\mathcal{R}_{\mathcal{U}}^{P}$ 与 $\hat{}\mathcal{R}_{\mathcal{U}}^{P} = \frac{1}{\mathcal{U}-1}\sum^{\mathcal{U}-1}\mathcal{R}_{i}^{P}$（前 $\mathcal{U}-1_{i=1}$个时间窗口内的平均驾驶评分）。假设该差异显著（大于某个预定义阈值 ϵ），那么 DriCon 就会进一步分析不同时间窗口中特征向量之间的时间依赖关系，以了解这一差异背后的原因。

E. 可能情境识别

在最后一个模块中，我们利用不同时间窗口的特征向量来构建模型，识别在窗口 $\mathcal{U}$ 期间，哪些特征（$\mathcal{F}_{GEN}$）导致了驾驶行为的变化。该模型通过分析过去窗口 $[1,\cdots,(\mathcal{U}-1)]$以及当前窗口 $\mathcal{U}$中发生的微事件所产生的影响，主动探寻此类波动背后的解释。最后，系统会生成基于自然语言的人类可解释说明，并反馈给利益相关者，以供进一步分析。

V. 模型开发

为了开发 DriCon 的核心模型，我们利用已提取的特征（ $\mathcal{F} \in \{\mathcal{F}_M \bigcup \mathcal{F}_S\}$，详见 §IV-C）来捕捉过去事件和当前事件的时间依赖关系。此外，DriCon 通过解释性特征（ $\mathcal{F}_{GEN}$）推导出对检测到的事件的解释。为此，我们需要一个自解释模型，用于捕捉不同驾驶操作与相关微事件之间的时空依赖关系，这些操作和微事件与道路上的驾驶行为相关。我们选择自组织映射（SOM）（ [27] ）来构建模型，使其能够在数据可用性有限的情况下利用此类时空依赖关系。经典深度学习模型（如卷积神经网络或循环神经网络）的主要局限在于：（i）深度网络消耗大量资源（例如内存），且严重依赖海量数据；（ii）它们如同黑箱，因此无法对某些预测（ [28]）生成人类可理解的解释。另一方面，SOM 能够在运行时利用特征变化和未标注数据来刻画微事件。

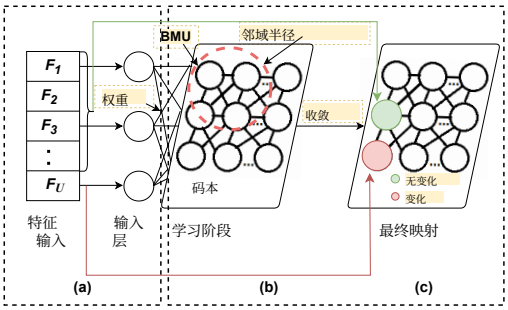


图4：SOM 的工作原理

A. 使用 SOM 推断解释性特征

获取解释性特征的关键在于首先发现时空特征依赖关系。在 DriCon 中，我们使用 Kohonen自组织映射（见图4）来推导这种关系，因为它是一种利用竞争学习方法的基于人工神经网络的无监督技术。由于 DriCon 在边缘设备上运行，我们采用尽可能少的模型参数，以加快处理速度，同时不影响性能。具体而言，我们使用 147 个神经元构建码本，并将它们分布在大小为 7×21 的二维阵列上（其中 7 是一个超参数，取决于一次行程中过去窗口的最大影响；21 对应特征空间中的特征数量）。这些神经元以随机权重初始化（见图4(a)），权重向量的长度与特征向量相同（均为 21）。接下来，我们用大小为 8×21 的二维网格表示每次行程（考虑行程中的 8 个连续窗口），以捕捉过去窗口 $[1,\cdots,(\mathcal{U}-1)]$ 和当前窗口 $\mathcal{U}$ 的影响。原则上，当驾驶行为没有变化时，SOM 固有的拓扑排序会将相似的特征空间（在窗口 $[1,\cdots,(\mathcal{U}-1)]$ 中）归为一组。相反，不相似的特征空间

（例如，在窗口 $\mathcal{U}$ 期间），当驾驶行为发生变化时，会被映射到不同的组中，如图4(b,c)所示。

例如，假设在某次行程中，交通信号灯突然变化后，前车采取制动动作，导致自车急停。因此，窗口 $[1,\cdots,(\mathcal{U}-1)]$ 中的特征空间呈现出相似的特征（直到发生急停），随后会映射到单个神经元。然而，在急停期间，特征空间会发生变化（例如驾驶操作和其他空间事件）。这些特征空间的变化会使其被分配到另一个神经元，并根据窗口 $[1,\cdots,(\mathcal{U}-1)]$ 与窗口 $\mathcal{U}$之间特征空间的变化，自动调整其他神经元的权重。此过程使 SOM 能够以无监督方式利用空间事件之间的时间依赖关系，而无需显式使用驾驶评分。

1) 模型训练：将输入的行程数据表示为 $2D$ 网格格式，以学习最佳匹配神经元，并优化特征空间与相应神经元权重向量之间的欧氏距离。为确保最佳拟合，最佳匹配神经元会尽可能学习特征空间的权重向量。此外，邻域中的神经元也会尝试将其权重调整得尽可能接近最佳匹配神经元。我们对该模型进行 500 个训练轮次的训练，每个神经元都会与最佳匹配的行程实例建立映射，并收敛到其在码本中的坐标位置。我们采用气泡邻域函数 [29] 更新邻域神经元的权重，直至邻域半径收敛到 ≈ 0。我们确保距离函数和邻域函数都具有较快的计算速度，从而实现准确学习并加速收敛。完成全部训练轮次后，我们得到收敛的码本，称为映射；每个行程实例都会分配给最佳匹配神经元，即最佳匹配单元（BMU）。与 BMU 坐标相对应的权重向量揭示了解释性特征$\mathcal{F}_{GEN}$。

2）模型执行：我们利用构建好的映射进行运行时推理。首先，对当前正在进行的行程进行特征处理（参见 §IV-C），同时将提取出的特征空间作为输入馈入构建好的映射。最终，我们获得最佳匹配单元的坐标，并提取其对应的权重向量以及给定行程实例的特征编码。我们从权重向量中提取前 k 个权重及其对应的特征名称（例如天气类型）和编码值（例如天气类型：雨天）。最后，我们将这些内容填入 $\mathcal{F}_{GEN}$ （称为生成式微事件），以便进一步生成便于人类理解的解释。

B. 生成文本解释

DriCon 旨在利用输出特征 $\mathcal{F}_{GEN}$ 生成文本形式的解释，以提高可读性并便于人类理解。由于这些特征 $f \in \mathcal{F}_{GEN}$ 已经与一些关键词（例如“严重颠簸”）相关联，我们需要将它们生成为句子形式，并根据特征 $f \in \mathcal{F}_M$ 或 $f \in \mathcal{F}_S$分别将其视为“动作”或“主语”。例如，如果特征是动作，我们就将自车设为主语，用描述性关键词替换相应的输出特征 f，最后将它们拼接起来，得到句子形式。例如，对于严重颠簸，构造的句子为“自车发生严重冲击”。然而，如果输出特征 f 表示主语，那么一个主语可以生成许多可能的句子。因此，我们挖掘若干交通指南 [30]，并使用 TF-IDF 向量化器计算特征与现有指南之间的余弦相似度。提取出最相关的指南后，我们获取句子中关联的物体，并为每个输出特征构造一个句子（例如，“行人过街” $\rightarrow$ “行人穿过交叉路口”）。接下来，对于所有生成的句子，我们使用 GloVe [31] 将每个特征对应的描述性关键词转换为形容词或副词，以使句子结构更流畅（例如，“自车发生严重冲击” $\rightarrow$ “自车发生了严重颠簸”）。最后，使用“和”连接词将各句拼接起来，并通过字符串操作将重复的主语替换为相应的代词形式，从而生成完整的解释，如图 3(e) 所示。

VI. 性能评估

本节介绍 DriCon 在实时环境和 BDD 数据集上的实现细节。我们报告 SOM 模型的性能，并将其与一个成熟的基线方法进行比较。此外，我们展示系统生成文本解释的效果，并通过敏感性分析辨别 DriCon 的出错倾向。以下从实验设置的细节开始介绍。

A. 实验设置

DriCon 部署在一套 Raspberry Pi 3 Model B 微处理器套件上，该套件运行 Raspbian 操作系统和 Linux 内核版本 $5.15.65-v7+$，配备 1 GB 主存和 ARMv7 处理器。我们主要使用惯性测量单元、全球定位系统，以及通过前置摄像头（朝向前挡风玻璃）采集的视频数据作为不同模态。为此，我们在 Raspberry Pi 开发板上安装了一个 MPU-9250 IMU 传感器、一个 u-blox NEO$-6M$ GPS 模块和一台罗技 USB 摄像头，如图 1(a) 所示。我们在三种不同类型的车辆（例如，运动型多用途汽车、轿车和掀背车）上部署了 DriCon。我们聘请了 6 名年龄在 $[20-45]$、平时经常开车的驾驶员。因此，整个实验历时两个多月，在三个城市开展，累计驾驶约 33 小时，行驶距离超过 1000 公里。驾驶员可以自由驾驶，没有收到任何具体指示；每次行程时长约为 20 分钟至 2 小时。此外，每位驾驶员都在五种不同类型的道路（城市街道、高速公路、住宅区道路、停车场和校园道路）上，于一天中的三个不同时段（白天、黄昏和夜间）驾驶。我们通过分析所提出的模型提取生成式

微事件的效果来评估 DriCon $\mathcal{F}_{GEN}$（见 §V-B）。在实现 DriCon 时，我们取 $\delta = 5$ 秒、 $\epsilon = 1$。其他超参数和资源消耗的影响将在后续分析中讨论。接下来，我们将介绍用于评估 DriCon 的真实标注流程。

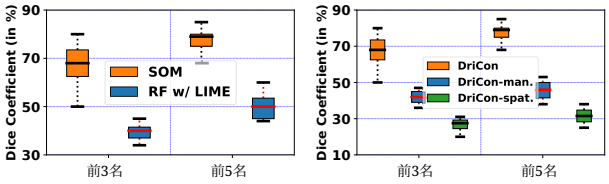
B. 微事件标注

我们向一组招募来的标注人员发放 Google 表单，启动了标注工作。标注人员需要观看一段不超过 10 秒的视频，并选出对驾驶行为影响最大的 3 个因素。我们针对内部数据（实时实验期间采集的视频数据）以及 BDD 数据集中的视频进行标注。对于表单中提供的两个数据集中的每段视频，我们只展示分数波动发生时对应的剪辑片段。接下来，标注人员需要从列表给出的 15 个因素（包括驾驶操作和空间微事件）中，依据视觉感知选出导致驾驶行为不佳的 3 个最重要因素。此外，我们还提供了模型生成的句子（§V-B），并请标注人员对这些句子在解释驾驶行为变化方面的相关性和结构完整性进行评分（评分范围为 $[1-5]$）。如果标注人员认为驾驶行为变化背后有更合理的原因，也可以自行撰写解释。由于行程数量相当多，我们需要设计一组 Google 表单（示例表单[1]），每份表单最多包含 20 段视频，以尽可能降低标注人员的认知负荷。我们还收集了标注人员的人口统计信息，例如年龄、性别、城市等。我们发现，大多数参与者（$>67\%$）此前具备驾驶技能。每个实例至少由三名独立标注人员进行标注。收到标注的因素后，我们需要计算标注人员之间的一致性，以确保获得的真实标注不存在偏差或随机性。由于标准的标注者间一致性指标（例如科恩卡帕系数）适用于定量分析或一对一映射，因此无法使用这类指标。于是，我们采用多数投票法，为列表中的每个因素分配一个百分比，表示标注人员选择该因素的频率。得票率至少为 60% 的因素将保留在$\mathcal{F}_{GT}$中。我们观察到， $\mathcal{F}_{GT}$ 的最小基数为 3，最大基数为 5。这也表明，标注人员在选择影响驾驶行为的因素方面达成了一致。 $\mathcal{F}_{GT}$ 包含经过标注的微事件，$\mathcal{F}_{GEN}$ 将据此进行评估。

C. 性能指标

我们使用戴斯相似系数评分 [32] ($\mathcal{N}$) 来计算 $\mathcal{F}_{GT}$与 $\mathcal{F}_{GEN}$之间的相似度，计算方式如下： $\mathcal{N} = \frac{2\times|\mathcal{F}_{GT}\cap\mathcal{F}_{GEN}|}{|\mathcal{F}_{GT}|+|\mathcal{F}_{GEN}|}$。我们报告所有行程的平均 $\mathcal{N}$，以衡量 DriCon 的准确性。接下来，我们还使用平均处理效应 [33] (ATE)，以报告模型识别出的特征中因果效应相对较高的特征。最后，我们将误差百分比定义如下。首先，我们

[1]https://forms.gle/97N6uk4ujRaZSWbj8（访问日期：2023年1月16日）



(a) (b)

图 5：（a）人工标注特征与模型生成特征之间的 Dice 系数相似度（%）；（b）消融研究

计算集合差 $\{\mathcal{F}_{GT} \setminus \mathcal{F}_{GEN}\}$，并提取相应的特征类别（例如， $\mathcal{F}_M, \mathcal{F}_S$）。统计每个特征类别的数量后，我们计算其占行程总数的百分比，作为误差百分比。

D. 基线实现

作为提取 $\mathcal{F}_{GEN}$ 的基线方法，我们实现了一种有监督的、基于规则的随机森林（RF）算法，其中包含 20 棵决策树，每棵树都在训练数据上扩展至无限深度。我们基于这样的直觉优化标签 $\mathcal{R}_u^p$： 不同特征对各个预测分数的贡献不同。尽管基于 RF 的模型具有特征重要性分数，用于表示每个特征在构建模型时的贡献，但我们还需要解释每个特征如何在单次行程实例的层面上为预测驾驶评分作出贡献。因此，我们在 RF 模型后台使用 LIME [34] 来生成解释性特征。由于 LIME 是一种模型无关方法，它通过调整特征值，尝试映射输入特征与输出分数之间的关系。因此，它会解释各特征对预测分数的贡献范围及其概率。我们从生成的解释中提取贡献特征 $\mathcal{F}_{GEN}$ 及其对应值，以便进一步生成文本解释。该流程以与 §VI-A 中所述类似的方式执行。

E. 表征上下文的准确性

我们使用 Dice 系数相似度 $\mathcal{N}$，在内部数据集上评估 DriCon 结合 SOM 和 RF+LIME 模型的准确性。我们从 $\mathcal{F}_{GEN}$ 中提取前 k 个特征，其中 $k \in \{3,5\}$，并计算两组特征之间的 $\mathcal{N}$—— $\mathcal{F}_{GEN}$ 和 $\mathcal{F}_{GT}$ 中前 k 个特征的相似度。图 5(a) 展示了结果。对于前 3 个特征，SOM 和 RF+LIME 的平均相似度分别为 69% 和 40%。对于前 5 个特征，SOM 和 RF+LIME 的平均相似度分别为 79% 和 48%。由于内部数据集包含更复杂的微事件，使用前 3 个特征时，在内部数据集上的性能略有下降，这可以接受。直观而言，模型能够捕捉到更多人工标注者所感知的多样性；因此，从 $k = 3$ 转向 $k = 5$ 时，相似度有所提升。然而，由于 RF+LIME 独立处理行程中的每个时间实例，其性能会下降。它捕捉的是当前时间窗口内导致驾驶行为变化的主要特征，而不是考察过去时间窗口的影响。

表 I：人工标注与模型生成输出之间的相似度度量

实例编号	人工标注 $\mathcal{F}_{GT}$	模型生成输出 $\mathcal{F}_{GEN}$	相似度 $\mathcal{N}$(%)	ATE
1	恶劣天气条件（强降雨、雾等）、急转向， 拥堵、超车、突然停车	拥堵、前车制动、 蛇形行驶、急停、严重颠簸	40%	1.96
2	侧滑、突然停车、交通信号灯：红灯	交通信号灯：红灯、拥堵、急停	66.67%	2.5
3	穿越行人、车辆间的高速变化、蛇形行驶	严重颠簸、穿越行人、蛇形行驶	66.67%	1.35

为便于了解，我们在表 I 中列出了三个测试实例样本的解释性特征（$\mathcal{F}_{GEN}$）与人工标注特征（$\mathcal{F}_{GT}$）；这些实例的相似度（Dice 系数）相对较低。有趣的是，当两者不匹配时，我们发现模型生成的特征与人工标注的特征在概念上大多相关。此外，模型生成的不匹配特征具有较高的正平均ATE值，这表明模型认为这些特征比一般人的感知更具因果性。值得注意的是，ATE 值 ≥ 1 表示特征与相应结果（驾驶行为变化）之间存在较强的因果关系。例如，在测试实例 #2 中，不匹配的特征分别是侧滑（人工生成）和拥堵（模型生成）；其中，拥堵的因果性相对更强，影响了驾驶行为的变化。通过手动分析该实例并采访相应的驾驶员，我们发现他确实在拥堵路段发生了轻微侧滑。事实上，这名驾驶员在拥堵路段驾驶手动挡汽车时感到不太自在。

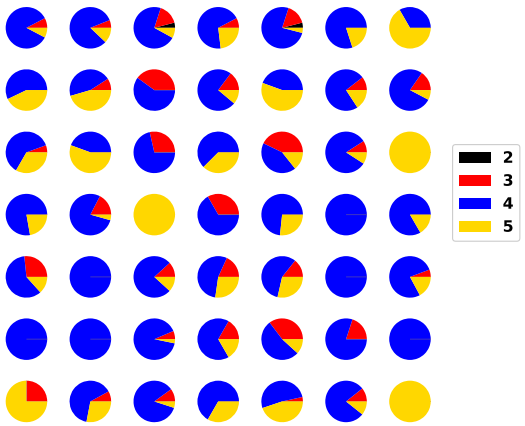


图 6：为 7×7 网络生成的 SOM 映射（缩小）

F. 消融研究

接下来，我们分析 §IV-A 中所述、与驾驶操作和道路空间事件相对应的不同特征类别对 DriCon 整体性能的重要性。为研究驾驶操作和空间特征的影响，我们实现了 SOM，每次排除上述某一类特征，并评估 $\mathcal{N}$，以考察各类特征的重要性。除 DriCon 外的两个变体按以下方式构建。(a) DriCon-man.：此处，我们排除驾驶操作 $\mathcal{F}_M$，并保留 $\mathcal{F}_S$ 。 (b) DriCon-spat.：接下来，我们排除空间特征 $\mathcal{F}_S$，并仅保留 $\mathcal{F}_M$。我们在排名前 3 和排名前 5 的生成特征两种设置下评估这两个变体，并与包含全部特征的 DriCon 一同进行比较，如图 5(b) 所示。排除驾驶操作和空间特征后，在排名前 5 的特征设置下，性能分别降至 45% 和 31%。这一显著下降表明空间特征至关重要，因为这些特征变化频繁，会导致驾驶行为波动。

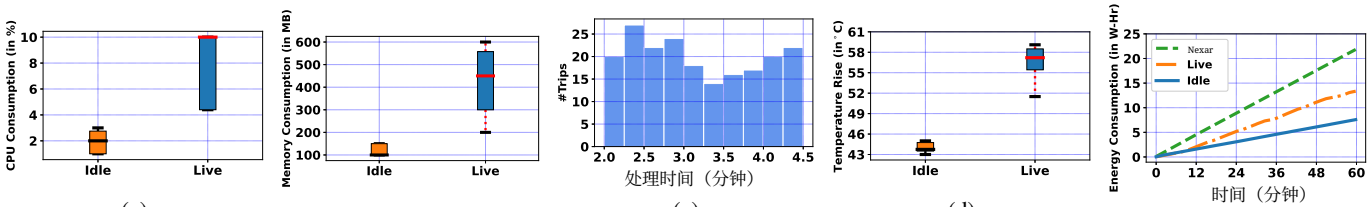
G. 模型洞察

为了解不同特征之间对应于驾驶操作和各种道路空间微事件的时空依赖关系是如何得出的，我们使用了分布在一个 7×7 二维阵列上的49个神经元（这是原始用于开发该模型的SOM网络的一个较小变体，因为原始模型包含147个神经元，难以可视化），并用200次行程进行拟合。该实例生成了如图6所示的映射，其中所有给定的行程都被分配给各个神经元。分数 $\mathcal{R}_U^P$ 仅用于直观展示各行程在映射中的位置。每次行程都通过特征变化捕捉驾驶行为的变化。与单色神经元相比，多色神经元更为重要，因为其中最常观察到分数波动。在一次独立行程中，行程各实例对应的特征值会保持相近，直到驾驶行为发生变化，因此这些实例会被分配到同一个神经元（单色）。然而，驾驶行为的差异会导致特征值不同于先前的实例，因此该实例会被分配到映射中的另一个神经元。如图6所示，多色神经元映射了驾驶行为发生突然变化的行程实例。

H. 剖析 DriCon

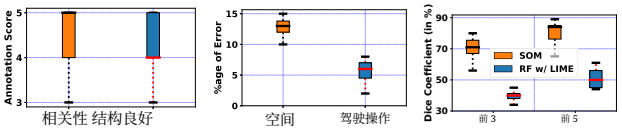
接下来，我们对 DriCon 的资源消耗情况进行基准测试，并分析该模型的重要性和敏感性。

1) 边缘设备资源消耗：我们针对两种情况进行基准测试，测量中央处理器和内存使用率、处理时间、温升及能耗：(a) 设备处于空闲状态；(b) DriCon 正在运行。从图 7(a) 可以看出，空闲模式下平均使用 2% 的中央处理器（使用“top”命令测得）。相比之下，运行 DriCon 时最多占用 10% 的处理器资源，这是可以接受的。不过，如图 7(b) 所示，内存使用率略高（$\approx$ 500MB），主要是由视频处理开销造成的。接下来，我们按行程展示从数据采集到输出生成所需的处理时间。



(a) (b) (c) (d) (e)

图 7：边缘设备上的资源消耗 (a) CPU 使用率 (b) 内存使用率 (c) 处理时间直方图（相对于行程数）(d) 温升 (e) 能耗



(a) (b) (c)

图 8：(a) DriCon 的显著性 (b) DriCon 的敏感性分析 (c) 在 BDD 数据集上的性能

对于大多数行程，DriCon 仅需 $\approx$ 3 分钟即可生成输出，这进一步验证了其响应时间较短（见图 7(c)）。为深入分析，我们还记录了运行 DriCon 时的温升（通过“vcgencmdmeasure temp”命令测得），并使用 Monsoon 高压功率监测仪 [35] 测量总能耗。从图 7(d) 和 (e) 可以看出，温度最高升至 59°C，平均能耗为 13 瓦时，这对于任何实时系统而言都属于正常水平。为了评估 DriCon 的性能，我们还测量了 Nexar 行车记录仪的能耗：其平均能耗为 22 瓦时，但在不具备任何上下文的情况下，只能捕捉极少数驾驶操作（例如急刹车）。这进一步说明，DriCon 不会耗尽边缘设备的资源，并且能够精准检测微事件。

2) 生成的解释的重要性：接下来，我们检查生成的解释有多重要。如 §VI-B 所述，我们绘制了两项指标——“相关性”和“结构良好”——的标注评分分布（由招募的标注人员给出）。“相关性”表示生成的解释在解释意外事件时的适用性。相比之下，“结构良好”表示模型生成的句子在多大程度上符合人类认知、具有良好的可解释性。图 8(a) 显示，“相关性”和“结构良好”的中位数分别为 5 和 4，这进一步证明了 DriCon 的可信度。我们还计算了人工标注句子与模型生成句子之间的相似度，并使用 TF-IDF 向量化器得出相似度的最小值、最大值和平均值分别为 51.33%、85.5% 和 70.57%。因此，DriCon 自动生成情境解释的能力与人类认知水平相近，达到人类与模型之间几乎难以区分的程度；这进一步表明，它能够为利益相关者的决策流程提供反馈。

3) DriCon 的敏感性：最后，我们检查 DriCon 未能捕捉到的微观事件。除了模型的效率外，我们还必须考察其不足，以分析这些不足对整体性能的影响程度。尤其是在利益相关者提升或惩罚驾驶员档案评分时，这一点非常重要。如图 8(b) 所示，未能捕捉空间特征和驾驶操作的情况都较少。尽管如 §VI-F 所述，这可能会导致模型性能下降，但由于驾驶操作 ($\mathcal{F}_M$) 依赖于空间特征 ($\mathcal{F}_S$)，它们对模型性能的贡献并不显著。但对于 $\mathcal{F}_S$，误差百分比仍为 $\leq$ 13%，因此该系统在生成易出错的情境解释时不那么敏感。

I. 离线性能

最后，我们报告了系统在 BDD 数据集上的准确率。该数据集包含 17 小时的驾驶数据，覆盖 $1.5k$ 次行程，使用了 $\mathcal{N}$。如图 8(c) 所示，DriCon 在预录制数据上的表现相当出色，前 3 和前 5 个特征的准确率分别为 $\mathcal{N} = \{71\%、84\%\}$。我们发现，在使用公开数据集进行离线分析时，SOM 能够更好地识别微事件。然而，与离线分析相比，在真实驾驶环境中实时运行系统至关重要，因此可以接受这点轻微的准确率下降。

VII. 结论

本文开发了一种名为 DriCon 的智能边缘设备系统，利用多模态检测导致驾驶行为意外波动的微事件。DriCon 生成的人类可解释的解释表明，该系统在识别此类情境方面具有相关性和可信度。此外，系统以无监督方式检查各种特征间的时空依赖关系，以捕捉多样化的驾驶场景。另外，在实时测试平台上进行的资源友好型部署进一步验证了 DriCon。尽管本研究捕捉了各特征贡献被分别考量的情境，但并未明确捕捉特征间依赖关系。例如，一名驾驶员为避免与穿越行人相撞，在转弯时突然变道，导致后方车辆的驾驶员猛踩制动器。在这里，第一名驾驶员的动作是由穿越行人引起的，而这又影响了第二名驾驶员的动作。要分析车辆之间此类复杂的群体交互，需要一种更复杂的

系统，可能需要采用能够关联车辆间交互的不同模态。不过，DriCon 提供了一种简单的独立解决方案，可单独部署在配备仪表盘安装的边缘设备或行车记录仪的车辆上。

参考文献

[1] “道路交通伤害”，世界卫生组织，https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries，2022年，（在线访问日期：2023年1月16日）。[2] “东京农工大学工学研究院（TUAT）智能移动研究中心——研究。”https://web.tuat.ac.jp/~smrc/research.html，2017年，（在线访问日期：2023年1月16日）。[3] N. H. T. S. A.（NHTSA），https://www.nhtsa.gov/，（在线访问日期：2023年1月16日）。[4] “自动驾驶车辆的激光雷达：一场技术军备竞赛”，https://www.automotiveworld.com/articles/lidars-for-self-driving-vehicles

2020年，（在线访问日期：2023年1月16日）。[5] Z. Li、C. Wu、S. Wagner、J. C. Sturm、N. Verma 和 K. Jamieson，“REITS：面向智能交通系统的反射表面”，载于第22届 ACM HotMobile，2021年，第78–84页。[6] R. Akikawa、A. Uchiyama、A. Hiromori、H. Yamaguchi、T. Higashino、M. Suzuki、Y. Hiehata 和 T. Kitahara，“基于智能手机的危险交通情境检测与分类”，载于 IEEE PerCom研讨会，2020年，第1–6页。[7] D. A. Ridel、N. Deo、D. Wolf 和 M. Trivedi，“利用车载视觉理解行人与车辆交互：长短期记忆网络模型与实证分析”，于2019年 IEEE 智能车辆研讨会（IV），第913–918页。[8] F. Yu、W. Xian、Y. Chen、F. Liu、M. Liao、V. Madhavan 和 T. Darrell，“BDD100K：一个多样化的驾驶视频数据库及可扩展的标注工具”，arXiv 预印本 arXiv:1805.04687，第2卷，5期，第6页，2018年。[9]“车辆检测与距离估计”，https://towardsdatascience.com/vehicle-detection-and-distance-estimation-a-technological-arms-race-7acde48236e1，2017，（在线访问日期：2023年1月16日）。[10] D. Das、S. Pargal、S. Chakraborty 和 B. Mitra，“Dribe：用于驾驶行为标注的地域中立型道路移动遥测”，载于第23届 IEEE MDM，2022年，第159–168页。[11] D. Mohan、G. Tiwari 和 K. Bhalla，“印度道路安全：2019年现状报告。新德里：印度理工学院德里分校交通研究与伤害预防计划。”http://tripp.iitd.ac.in/assets/publication/Road Safety in India2018.pdf，2019年，（在线访问日期：2023年1月16日）。[12] K. Fu、Z. Chen 和 C.-T. Lu，“StreetNet：基于卷积神经网络的城市犯罪感知偏好学习”，载于第26届 ACM SIGSPATIAL 会议论文集，2018年，第269–278页。[13] K. Patroumpas、N. Pelekis 和 Y. Theodoridis，“飞机轨迹上的实时移动事件检测”，载于第26届 ACM SIGSPATIAL 会议论文集，2018年，第259–268页。[14] I. Janveja、A. Nambi、S. Bannur、S. Gupta 和 V. Padmanabhan，“INSIGHT：利用智能手机在低光照条件下监测驾驶员状态”，《ACM交互式、移动、可穿戴与泛在技术汇刊》，第4卷，第3期，第1–29页，2020年。[15] X. Fan、F. Wang、D. Song、Y. Lu 和 J. Liu，“GazMon：基于眼动注视的驾驶行为监测与预测”，《IEEE移动计算汇刊》，2019年。[16] M. Walch、M. Woide、K. M¨uhl、M. Baumann 和 M. Weber，“协作超车：借助驾驶员帮助克服自动驾驶车辆受阻的传感器探测范围”，载于第11届 ACM AutomotiveUI，2019年，第144–155页。[17] H. T. Lam，“空间异常及其在高效实时驾驶行为监测中的应用简述”，载于第24届 ACM SIGSPATIAL 会议论文集，2016年，第1–9页。[18] S. Moosavi、B. Omidvar-Tehrani、R. B. Craig、A. Nandi 和 R. Ramnath，“根据驾驶员行为刻画驾驶情境”，载于第25届 ACM SIGSPATIAL 会议论文集，2017年，第1–4页。[19] Y. Shi、R. Biswas、M. Noori、M. Kilberry、J. Oram、J. Mays、S. Kharude、D. Rao 和 X. Chen，“利用地理空间数据和机器学习预测道路事故风险（演示论文）”，载于第29届 ACM SIGSPATIAL 会议论文集，2021年，第512–515页。[20] M. R. Samsami、M. Bahari、S. Salehkaleybar 和 A. Alahi，“用于自动驾驶的因果模仿模型”，arXiv预印本 arXiv:2112 03908，2021年。[21] V. Ramanishka、Y.-T. Chen、T. Misu 和 K. Saenko，“迈向驾驶场景理解：用于学习驾驶员行为和因果推理的数据集”，载于 IEEE计算机视觉与模式识别会议，2018年，第7699–7707页。[22] F. Codevilla、E. Santana、A. M. L´opez 和 A. Gaidon，“探索行为克隆在自动驾驶中的局限性”，载于 IEEE/CVF国际计算机视觉大会，2019年，第9329–9338页。[23] J. Redmon 和 A. Farhadi，“YOLOv3：渐进式改进”，arXiv，2018年。[24] J. Yu、Z. Chen、Y. Zhu、Y. Chen、L. Kong 和 M. Li，“利用智能手机检测并识别细粒度异常驾驶行为”，《IEEE移动计算汇刊》，第16卷，2212，2016 8，2198，第–页。[25] T.-Y. Lin、M. Maire、S. Belongie、J. Hays、P. Perona、D. Ramanan、P. Doll´ar 和 C. L. Zitnick，“Microsoft COCO：情境中的常见物体”，载于欧洲计算机视觉会议。施普林格，2014年，第740–755页。[26] D. Das、S. Pargal、S. Chakraborty 和 B. Mitra，“为什么猛踩刹车？通过自适应因果建模自动标注驾驶行为”，载于 IEEE PerCom研讨会，第587–592页。[27] T. Kohonen，“自组织映射”，《IEEE会刊》，第78卷，第9期，第1464–1480页，1990年。[28] C. Rudin，“对于高风险决策，应停止解释黑箱机器学习模型，转而使用可解释模型”，《自然·机器智能》，第1卷，第5期，第206–215页，2019年。[29] “邻域函数”，https://users.ics.aalto.fi/jhollmen/dippa/node21.html，（在线访问日期：2023年1月16日）。[30] “行人设施指南”，http://www.irc.nic.in/admnis/admin/showimg.aspx?ID=345，（在线访问日期：2023年1月16日）。[31] J. Pennington、R. Socher 和 C. D. Manning，“GloVe：用于词表示的全局向量”，载于自然语言处理经验方法会议，2014年，第1532–1543页。[32] A. Carass、S. Roy、A. Gherman、J. C. Reinhold、A. Jesson、T. Arbel、O. Maier、H. Handels、M. Ghafoorian、B. Platel 等，“采用改进的Sørensen-Dice分析评估白质病变分割结果”，《科学报告》，第10卷，第1期，第1–19页，2020年。[33] D. B. Rubin，“估计随机研究和非随机研究中治疗的因果效应。”《教育心理学杂志》，第66卷，第5期，第688页，1974年。[34] M. T. Ribeiro、S. Singh 和 C. Guestrin，“‘我为什么要相信你？’解释任意分类器的预测”，载于第22届 ACM知识发现与数据挖掘会议论文集，2016年，第1135–1144页。[35] “Monsoon 高压功率监测仪”，https://www.msoon.com/online-store/High-Voltage-Power-Monitor-p90002590，（在线访问日期：2023年1月16日）。