

DOI: 10.13382/j.jemi.B2306900

驾驶员视线区域自适应半监督学习标定方法^{*}

程 鸣 严运兵

(武汉科技大学汽车与交通工程学院 武汉 430072)

摘 要:目前监测驾驶员视线区域的算法通常采用深度学习端模型直接对图像特征分类,此方法依赖固定座舱视角下采集的驾驶员视线区域数据,但由于驾驶员外形特征差异、坐姿习惯差异和摄像头安装位置差异的影响,难以获取大量且全面的数据,导致分类精度降低的问题,如何仅采用小样本数据集提升视线区域识别精度成为难题。本文将基于半监督学习理论设计一种自适应的视线区域标定方法。首先采用 L2CS 模型回归小样本数据中驾驶员视线角度二维向量,再通过统计分析挖掘驾驶员视线角度和视线区域映射的泛化先验知识,利用该知识进行视线区域标定,剔除非待检区域无效视线落点,并以滑动窗口方式完成针对驾驶员个人的视线区域精细化分类。经试验证明,该方法解决了端模型数据跨域能力低下的问题,准确率和召回率分别提升 22.4% 和 10.3%,且标定结果具有自适应修复能力。

关键词: 视线区域;半监督学习;小样本数据;标定技术;自适应调节

中图分类号: TN911.8;TP183 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Adaptive semi supervised learning calibration method for driver's line of sight region

Cheng Ming Yan Yunbing

(School of Automotive and Traffic Engineering, Wuhan University of Science and Technology, Wuhan 430072, China)

Abstract: At present, algorithms for monitoring the driver's line of sight area usually use deep learning models to directly classify image features. This method relies on the driver's line of sight area data collected from a fixed cockpit perspective. However, due to differences in driver appearance, sitting habits, and camera installation positions, it is difficult to obtain a large amount of comprehensive data, resulting in a decrease in classification accuracy. How to improve the accuracy of line of sight recognition using only small sample datasets has become a challenge. This article will design an adaptive line of sight region calibration method based on semi supervised learning theory. Firstly, the L2CS model is used to regress the two-dimensional vector of driver's line of sight angle in small sample data. Then, statistical analysis is used to mine the generalization prior knowledge of driver's line of sight angle and line of sight area mapping. This knowledge is used for line of sight area calibration, removing invalid line of sight landing points in non-inspection areas, and completing fine classification of driver's personal line of sight area in a sliding window manner. Through experiments, it has been proven that this method solves the problem of low cross domain capability of end model data, improving accuracy and recall by 22.4% and 10.3% respectively, and the calibration results have adaptive adjustment ability.

Keywords: line of sight area; semi supervised learning; small sample data; calibration techniques; adaptive adjustment

0 引 言

驾驶员视线区域可以直观的表征其是否处于正常驾驶或分心驾驶中^[1],监控驾驶员视线区域有利于提升驾

乘安全^[2]。同时它也常常被用于智慧座舱中的人机交互场景,例如通过视线点亮屏幕,防止手动触碰屏幕造成的二次分心,又如使用抬头显功能时,根据视线向量定位抬头显投影到前车窗上信息的位置高度。驾驶员的视线监测和跟踪技术已受到工程界的关注,亟待转化成高价值

的产品^[3]。

目前,已有的视线区域检测方法从实现手段上可以分为接触式和非接触式^[4]。在接触式系统中,由于需要在待测者身上安装传感设备,不利于驾驶员的实际使用,所以在研究驾驶员视线时都是采用的非接触式方式。依靠视觉系统监测视线主要是以数字视频图像为分析介质,运用图像处理和模式识别等方法,提取图像中脸部和眼部的特征信息从而完成视线估计,基于视觉的非接触式视线估计方法已成为主流。文献[5]基于卷积神经网络 CNN 从眼睛图像学习用于凝视估计的回归网络,此方法可以为不同头部姿态的用户提供准确的注视估计,因为它明确地将头部姿态信息包括在所提出的注视框架中。Bernard 等^[6]基于胶囊网络在分类任务中的良好表现,重建简化的眼睛图像,然后从中估计注视区域。Shah 等^[7]首先采用 Inception-V3 优化 YOLOV4 的人脸检测模型,再基于多任务学习的多头回归思想同时对头部姿态和视线区域进行分类,以减弱头部运动所导致的视线区域分类精度降低的问题。文献[8]基于人脸面部关键点位置和 3D 面部模型参数的隐式相关性,引入了级联回归框架来实现视线区域分类,在保持精度的要求下减少了冗余模型压缩了算力消耗。Vora 等^[9]将人脸上半部分的图片作为深度学习模型输入,实验证明此方法有利于降低人脸外形差异带来的干扰,并公开了实车中采集的视线区域数据集 Lisat_gaze。也有一部分学者将端到端模型拆解为两步法,首先回归出视线角度,再获取视线角度和视线区域的映射关系。目前,这种映射关系的获取方法包括引导式标定、监督学习、无监督学习。Jiannan 等^[10]基于 PCCR (pupil center-corneal reflection) 瞳孔角膜反射技术的视线估计方法^[11-14],首先提出了一种瞳孔中心、眼角矢量与屏幕视线关注点的二维几何模型估计方法,对系统和人眼视线区域进行了标定以提高准确性。然而这类标定方法要求头部位置相对固定,不符合当前任务的需求。文献[15]提出的基于单相机双光源系统的二维映射标定方法中允许头部在左右 5 cm,上下 8 cm,转动 10° 的范围内运动,但是引导式标定方法的操作繁琐不便于驾驶员使用。单兴华^[16]设计了一种改进的随机森林模型,首先通过高斯混合聚类模型对驾驶员特征进行预处理,再根据该特征完成视线区域分类。闫秋女^[17]基于 3D 人眼模型的方法由 2D 关键点估计驾驶员视线方向,再结合驾驶员头部姿态及人眼凝视方向特征,利用改进的随机森林算法对注视区域进行估计。Wang 等^[18]设计了一种无监督学习聚类方法,此方法可以逐步提取驾驶员视线落点,由此建立非线性视线区域映射关系,然而该方法对非待检区域的离群点敏感,离群点会改变聚类中心导致精度下降。

随着驾驶员视线区域监测算法的用户人群和搭载车

型的激增,数据跨域问题愈发显著。在驾驶员视线监测领域,数据跨域指的源域和目标域的语义内容差别巨大,模型的泛化能力急剧下降,其中语义内容的差别可描述为如下 3 类典型的实际应用场景:驾驶员的外形特征差异、驾驶习惯差异和车内摄像头的位置差异。驾驶员的外形差异又分为身高差异和容貌差异,驾驶习惯差异指不同驾驶员会根据自身习惯调节座椅的位置。车内摄像头的位置差异在不同车型中表现的最为明显,内饰结构的差异导致车内摄像头的安装位置不一致。另外,摄像头的硬件指标也千差万别^[19],尤其是红外光和可见光的成像效果将完全不同,这些典型的场景都引起了数据跨域问题。结合此客观存在的问题,分析研究现状如下:1) 视线区域端模型方法,过度依赖固定座舱视角下采集的驾驶员视线区域数据,而现实应用中难于寻找大量且全面的志愿者在指定车辆中采集数据,导致分类精度提升受限。2) 两步法较端模型方法而言,已对数据跨域问题进行了解耦,即该方法允许在野外数据集下回归视线角度,大幅度降低了数据收集的难度,削弱了由于驾驶员容貌差异造成的过拟合风险。同时该方法将复杂的眼部特征信息降维成视线角度二维向量,可解释性及可操作性提升。然而,在视线角度和视线区域的映射关系重构阶段,仍然需要使用固定视角的车辆环境中采集的数据。如何仅采用小样本数据集,进一步减弱由于驾驶员身高差异、坐姿习惯差异、摄像头安装位置差异对检测精度造成的影响,成为该技术路线的难点。

考虑到半监督学习方法优势,它可以从少量的已标注样本数据中获取先验知识,指导未标注数据完成分类,这与本研究任务中的客观条件最为契合。本文将以两步法技术路线,基于半监督学习理论,设计一种操作性强、鲁棒可靠且精度较高的自适应标定方法,提升小样本数据集中,驾驶员视线区域分类精度。

1 半监督学习驾驶员视线滤波估计框架

该框架的主要目的是,仅采用小样本数据集通过标定技术,解决由于驾驶员外形特征差异、驾驶习惯差异和车内摄像头的位置差异对视线区域分类造成的影响。框架包括监督学习、先验知识提取、实时标定和实时监测 4 个部分,流程框图如图 1 所示。

监督学习:主要通过人脸检测模型和视线角度模型对人脸图片进行降维操作,将复杂的容貌特征简化为俯仰角和偏航角的二维向量特征。人脸检测模型采用 Mediapipe Facemesh 接口,视线角度模型采用 L2CS-Net (L2 loss cross-entropy loss softmax net)^[20]。

先验知识提取:线下离线操作,通过人工标注小样本数据集,建立固定车辆视角中不同人群的视线角度和视

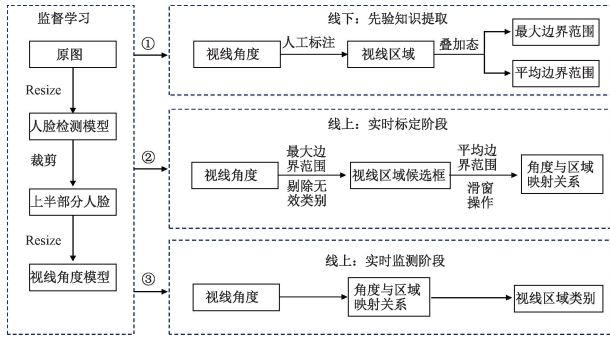


图 1 半监督学习驾驶员视线滤波框架

Fig. 1 Supervised learning driver's gaze filtering framework

线区域的初步映射关系。将这些人群的每一类视线区域中的视线角度统计分析计算,以得到每类视线区域所包络的最大视线落点边界和平均视线落点跨度。该先验知识将用于实时标定。

实时标定阶段:驾驶员自行调整至舒适的座椅位置后进入实时标定阶段。当前驾驶员的每一类视线区域的视线落点已达到阈值后。首先通过最大视线落点边界与当前驾驶员的视线落点叠加,以剔除非目标区域的无效点。再使用平均视线落点跨度在每类视线区域中执行滑动窗口操作,获取密度最大区域,即完成了当前驾驶员视线角度与视线区域的精细化映射。

实时监测阶段:针对驾驶员个人的标定结束后,进入实时监测阶段。算法将根据标定的结果,直接由模型输出的视线角度转化为视线区域。

2 视线角度模型

L2CS-Net 网络最大的贡献是仅依靠单目摄像头在野外数据训练,获得了自由移动人体的视线方向,并以视线角度作为的输出结果。它的主干部分采用 ResNet 网络,分别对视线偏航角和俯仰角设计了独立的损失函数,每种损失函数由交叉熵损失和均方差损失融合组成,模型结构如图 2 所示。

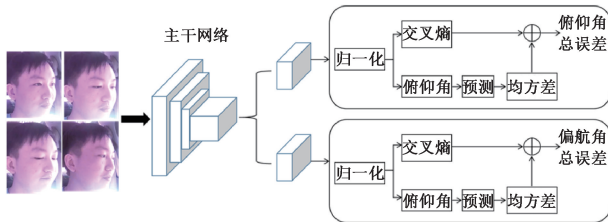


图 2 L2CS-Net 模型结构

Fig. 2 L2CS Net model structure

其中交叉熵和均方差损失函数如式(1)所示:

$$\begin{cases} H(y, p) = - \sum_i y_i \log p_i \\ MSE(y, p) = \frac{1}{N} \sum_0^N (y - p)^2 \end{cases} \quad (1)$$

式中: y 和 p 分别代表真实值和预测值,融合后的分类损失函数如式(2)所示:

$$CLS(y, p) = H(y, p) + \beta \cdot MSE(y, p) \quad (2)$$

式中: β 为权重参数。

L2CS 使用两个无约束条件下的数据集评估模型 MPIIGaze^[21] 和 Gaze360^[22] 分别取得了 3.92° 和 10.41° 的成绩。图 3 为使用 L2CS-Net 在 YawDD 数据中的实验表现。

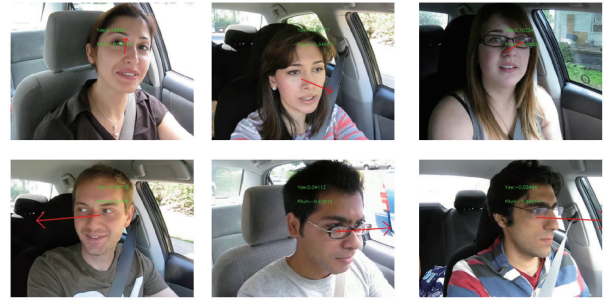


图 3 L2CS-Net 回归驾驶员视线角度

Fig. 3 L2CS-Net obtain the driver's line of sight angle

3 先验知识提取

将驾驶员视线偏航角和俯仰角投射到二维坐标系中形成视线落点离散信号图,并人工标注每个视线落点的视线区域。令 C 代表样本数据集中所有参与者视线落点的集合,如式(3)所示。

$$\begin{cases} C = \{C_1, C_2, \dots, C_{n-1}, C_n\} \\ C_n = \{p_1^n, p_2^n, \dots, p_{k-1}^n, p_k^n\} \\ p_k^n = (yaw_k^n, pitch_k^n) \end{cases} \quad (3)$$

其中, n 为待检测的类别数量, p_k^n 为第 n 类的第 k 个点, $yaw_k^n, pitch_k^n$ 分别代表 n 类的第 k 个点的偏航角和俯仰角。

计算样本数据每一类别的的最大视线落点边界和平均视线落点跨度是本阶段的核心工作。最大边界和平均跨度分别用 B_n 和 S_n 表达,如式(4)和(5)所示。

$$B_n = (YAW_{\min}^n, YAW_{\max}^n, PITCH_{\min}^n, PITCH_{\max}^n) \quad (4)$$

式中: $YAW_{\min}^n, YAW_{\max}^n, PITCH_{\min}^n, PITCH_{\max}^n$ 分别代表第 n 类的最小横坐标,最大横坐标,最小纵坐标和最大纵坐标。

$$\begin{cases} S_n = (W_n^s, H_n^s) \\ W_n^s = \sum_{i=1}^{R_n} ({}_nYaw_{\max}^i - {}_nYaw_{\min}^i) / R_n \\ H_n^s = \sum_{i=1}^{R_n} ({}_nPitch_{\max}^i - {}_nPitch_{\min}^i) / R_n \end{cases} \quad (5)$$

式中: W_n^s, H_n^s 分别表示第 n 类平均边界的长度和宽度, R_n 表示样本数据中第 n 类自愿者的人数, ${}_nYaw_{\min}^i, {}_nYaw_{\max}^i, {}_nPitch_{\min}^i, {}_nPitch_{\max}^i$ 表示第 n 类中第 i 个自愿者俯仰角和偏航角的最小值和最大值。

本文认为,最大视线落点边界的意义在于,它表示了样本数据集中,不同身高、不同坐姿的人群在凝视某一待检区域的最大视线允许范围;而视线落点平均跨度则反映了每类设定的待检区域的长宽距离。由此先验知识,即可指导未标注中视线角度和视线区域的映射关系重构。因为该先验知识是由视线角度二维向量和与之对应的视线区域统计得到,而这种映射关系的可解释性强,它主要受到驾驶员身高差异及坐姿习惯的影响,所以在制作该小样本数据集时,仅需邀请身高存在差异且身高已涵盖正常范围的志愿者,即可得到有效的先验知识,而不再需要邀请大量存在容貌差异的志愿者,这使得数据收集的难度大幅下降。

4 在线实时标定

实时标定的过程如图 4 所示,首先根据 B_n 的约束区

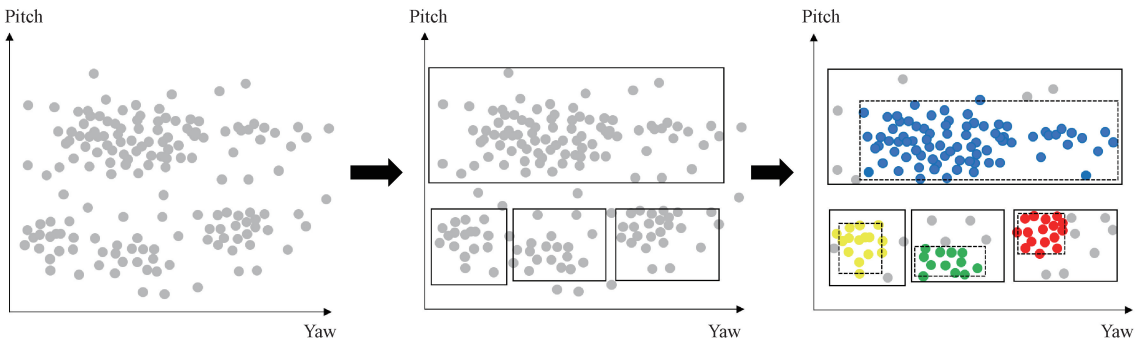


图 4 视线落点滤波估计视线区域

Fig. 4 Estimating gaze area using line of sight point filtering

5 监测损失函数及自修复性

客观现象告诉我们,驾驶员的视线点数量会随着驾驶时间的增长而变多,同时每个类别集中化的程度也会更高。只对每个类别设计一个滑窗操作的启始量 P_n ,当达到以后仅遍历一次就输出视线类别的角度范围,从整个驾驶环节来看,结果可能会缺乏稳定性。例如驾驶员

域快速的筛除掉非待检目标类的视线落点,然后将 S_n 作为滑动窗口计算 B_n 中密度最高的局部区域。滑动窗口的移动方向同 CNN 卷积运算,从左上角开始以设定的步长向右向下移动,边界采用自适应补 0 填充。

滑动窗口操作计算视线落点密度最大区域的索引,令滑窗步长为 ST ,由式(4)容易计算 B_n 的长度 W_n^B 和宽度 W_n^s 。将每一类索引总数记为 Q_n ,每个类别完成一次迭代要滑动的次数如式(6)所示。

$$Q_n = [(W_n^B - W_n^s) / ST]_{\text{down}} \times [(H_n^B - H_n^s) / ST]_{\text{down}} \quad (6)$$

其中, down 表示向下取整。 W_n^B 和 H_n^B 分别代表滑动窗口横向移动次数和纵向移动次数,如式(7)所示。

$$\begin{cases} W_n^B = [(W_n^B - W_n^s) / ST]_{\text{down}} \\ H_n^B = [(H_n^B - H_n^s) / ST]_{\text{down}} \end{cases} \quad (7)$$

窗口每滑动一个步长所包含的视线落点数量记为 N_n^T ,完成一次滑窗迭代后保存本次迭代每一步所包含的视线点数量,用集合 N_n 表示如式(8)所示。

$$N_n = \{N_n^1, N_n^2, N_n^3, N_n^4, \dots, N_n^T\} \quad (8)$$

式中: n 为视线类别,其中 $T = Q_n$ 。

记集合 N_n 中的最大值为 $N_n^{\max}$,其对应的索引为 $Q_n^{\max}$,所以滑动窗口移动至最大索引处即完成了视线角度和视线区域映射关系重构。

调整了座椅的前后位置,或者驾驶员在座椅上垫了一个更舒适的海绵垫等,都会让视线标定的结果出现偏差。

这就引出本文中标定算法的自修复性概念,已标定好的视线角度范围如果发生了超出设定阈值的变化,需要自我修复标定的结果。从前面的章节可知, B_n 和 S_n 是通过小样本数据得到,使用二范数计算滑窗迭代的损失函数,如式(9)所示。

$$Loss = \sum_{i=1}^e \sqrt{({}_nR_{pitch}^{e+1} - {}_nR_{pitch}^e)^2 + ({}_nR_{yaw}^{e+1} - {}_nR_{yaw}^e)^2} / e$$

(9)

其中, $Loss$ 为损失函数, e 是滑窗操作的迭代次数, ${}_nR_{pitch}^e, {}_nR_{yaw}^e$ 表示第 e 次迭代第 n 类视线区域的中心点坐标。例如 $e = 10$, 即每 10 次迭代为一轮, 计算这一轮的 $Loss$, 当 $Loss$ 收敛至到一个较小值时, 才能说明此时的标定结果是相对稳定的。算法的自修复性也体现在其中, 假设当前 $Loss$ 已收敛到目标值, 视线标定阶段结束, 进入视线监测阶段。但此时, 滑动窗口算法依然在继续进行, 并在后台中计算 $Loss$ 。当发现 $Loss$ 达到抑制阈值时, 退出视线监测阶段, 重新进入在线标定阶段, 如图 5 所示。将激活条件和抑制条件的阈值设置成不相同, 可以起到防止系统抖动的作用。例如, 激活条件的 $Loss$ 阈值设计为 0.1, 抑制条件的 $Loss$ 阈值也设计为 0.1, 那么系统就可能在激活和抑制状态中频繁切换, 根据实际经验设定抑制条件的阈值高于激活条件 20%, 也就是将抑制条件的 $Loss$ 阈值设定为 0.12。

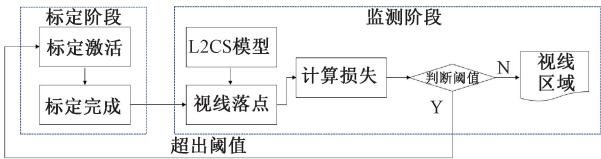


图 5 视线监测自修复性流程

Fig. 5 Gaze area monitoring self-healing process

6 实验与分析

6.1 环境及数据准备

本实验采用的控制器是 TITAN4C, 这一款以英伟达 XAVIER 和 NXP 的 MPC57XX 系列 MCU 为基础, 采用 GPU、CPU、MCU 异构方式的高算力、高性能、高可靠要去的控制器。其配备 2 个英伟达 Xavier 核心计算单元, 总算力达到 64TOPS, 操作系统采用 Ubuntu18. 04。摄像头水平角度和垂直角度分别为 $\pm 60^\circ$ 和 $\pm 40^\circ$, 帧率 30 fps, 固定在后视镜区域。

共邀请 12 位自愿者参与样本数据的录制, 参与者的身高范围在 155~190 cm 之间, 以 5 cm 为一个区间, 平均每个区间有 1~2 名参与者。要求参与者观察图 6 中所示的目标待检区域, 并有标注人员实时记录区域类别。每一类目标区域的单个视频长度约为 10~15 s, 按 0.2 s 的间隔对视频进行分帧操作, 总数据量如表 1 所示。为尽可能的体现此方法的便捷性, 仅收集足够且尽可能少的数据。

在小样本数据准备环节中, 根据实验发现, 当摄像头

在人脸正前方时, L2CS 模型可在头部左右偏转正负约 75° 范围内, 至少单眼可见的情况下正常输出视线角度。这点将决定本文算法的适用性, 当超出该头部姿态限制范围时, 本文算法不再适用。实际应用中, 可根据摄像头与驾驶员的相对位置关系, 剔除超出头部姿态限定范围的检测结果。



图 6 实验中待检视线区域

Fig. 6 Gaze area to be inspected in the experiment

表 1 小样本数据组成结构

Table 1 Small sample data composition structure

数据类型	用途	类别	各类别数据量/张
小样本数据	提取先验知识 B_n 和 S_n	前方	735
		中控屏	689
		仪表盘	658
		左方	791
		右方	802
		其他区域	1 358
测试集	测试验证 算法效果	前方	136
		中控屏	125
		仪表盘	125
		左方	142
		右方	139
		其他区域	625

6.2 对比实验

为与其他方法进行更公平的对比, 以对比各方法的便捷性、抗干扰能力以及准确性, 将设计如下试验条件: 1) 仅使用表 1 中的数据量完成实验, 不再额外扩增实车数据, 允许使用其他公开数据集做模型预训练; 2) 在视线区域监测阶段, 驾驶员可随意调整位置与坐姿, 使驾驶员与摄像头的相对位置关系发生改变。根据车辆座椅位置调整的最大行程, 前后调整幅度大于 3 cm, 上下调整幅度大于 2 cm; 3) 视线监测阶段和小样本数据标注阶段邀请不同的自愿者, 保证测试集与训练集的独立性。实验结果如表 2 所示, 其中 $Precision, Recall, F - Measure$ 的计算公式如式 (10) 所示。

表 2 算法性能对比

Table 2 Comparison of algorithm performance

算法	调整座椅前			调整座椅后		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
方法 1 ^[9]	0.82	0.75	0.78	0.66	0.58	0.62
方法 2 ^[16]	0.88	0.93	0.90	0.76	0.83	0.79
方法 3 ^[18]	0.68	0.96	0.80	0.55	0.92	0.71
引导式标定	0.97	0.99	0.98	0.76	0.87	0.81
本文方法	0.94	0.95	0.95	0.93	0.96	0.94

$$\left\{\begin{aligned} Precision &= \frac{TP}{(TP + FP)} \\ Recall &= \frac{TP}{(TP + FN)} \\ F-Measure &= 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)} \end{aligned}\right. \quad (10)$$

由表 2 可见,方法 1 采用端模型直接推理视线,当数据量不足时,模型效果将受到影响;方法 2 先回归视线角度,再使用监督学习模型获取视线角度和视线区域的映射关系,这种方式减弱了由于数据量不足引起的人脸特征提取不充分的问题。但是当座椅位置调整时,原有的映射关系发生改变,准确率大幅度下降。方法 3 同样也是先回归视线角度,与方法 2 的区别在于使用无监督学习完成的视线区域聚类。该方法难以区别非待检区域的视线落点,导致召回率很高但准确率下降。引导式标定的方法,在调整座椅位置前,监测性能最佳,但明显会受到驾驶员和摄像头相对位置关系的影响,同时此方法不具备便捷性。

本文中的方法,系统会根据条件重新进入标定阶段,算法具备很强的抗干扰能力,可以在无引导提示的条件下,实时标定驾驶员视线区域。准确率和召回率较引导式标定方法分别提升 22.4%和 10.3%。通过实验说明,在解决视线区域监测的数据跨域问题时,监督学习、无监督学习和引导式标定的方法从便捷性、稳定性的准确性的评价角度都不如半监督学习方法。

7 结 论

监测驾驶员视线可以帮助系统分析驾驶状态,也可以提高人机交互体验。但驾驶员相貌差异、身高差异、坐姿习惯差异和车内摄像头安装位置等差异都会导致数据难以全面收集,视线区域分类精度下降。本文提出了一种基于半监督学习的视线区域标定方法,简化了在固定座舱视角下数据采集的难度,仅根据小样本数据中统计分析得到的先验知识,完成针对驾驶员个人的视线角度和视线区域映射关系自适应标定。经实验验证,该方法较现有方法相比具有抗干扰能力强,便捷实用和准确率高的优势。

参考文献

[1] 杨易蓉. 基于头眼特征融合的驾驶员视线区域估计及驾驶场景关联方法研究[D]. 济南:山东大学,2022.
YANG Y R. Research on driver's line of sight area estimation and driving scene association method based on head eye feature fusion [D]. Jinan: Shandong University, 2022.

[2] PRATHA B, PREETAM S. Driver gaze tracking and eyes off the road detection [J]. Materials Today: Proceedings, 2023, 72(3): 1863-1868.

[3] 张捷,张晋崇. 车联网信息对驾驶员视觉注意的影响研究[J]. 现代电子技术,2021,44(11):135-140.
ZHANG J, ZHANG J CH. Research on the impact of connected vehicle information on driver visual attention [J]. Modern Electronic Technology, 2021, 44 (11): 135-140.

[4] 刘佳惠,迟健男,尹怡欣. 基于特征的视线跟踪方法研究综述[J]. 自动化学报,2021,47(2):252-277.
LIU J H, CHI J N, YIN Y X. A review of research on feature-based line of sight tracking methods [J]. Journal of Automation, 2021,47 (2): 252-277.

[5] SOCCINI A. Gaze estimation based on head movements in virtual reality applications using deep learning [C]. IEEE Virtual Reality, 2017: 413-414.

[6] BERNARD V, WANNOUS H, VANDEBORRE J. Eye-Gaze estimation using a deep capsule-based regression network[C]. 2021 International Conference on Content-Based Multimedia Indexing, 2021: 1-6.

[7] SHAH S M, SUN Z, ZAMAN K. A driver gaze estimation method based on deep learning[J]. Sensors, 2022, 22(10):3959.

[8] GOU C, ZHOU Y, XIAO Y, et al. Cascade learning for driver facial monitoring [J]. IEEE Transactions on intelligent vehicles, 2023, 8(1): 404-412.

[9] VORA S, RANGESH A, TRIVEDI M. Driver gaze zone estimation using convolutional neural networks; A general framework and ablative analysis[J]. IEEE Transactions on Intelligent Vehicles,2018, 3(3): 254-265.

[10] JIANNAN C, ZUOYUN Y, GUOSHENG Z, et al. A novel multi-camera global calibration method for gaze tracking system [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69 (5): 2093-2104.

[11] CHENG S, PING Q, WANG J, et al. EasyGaze: Hybrid eye tracking approach for handheld mobile devices[J]. Virtual Reality and Intelligent Hardware, 2022,4(2): 173-188.

[12] SKODRAS E, KANAS V G, FAKOTAKIS N. On visual

- gaze tracking based on a single low cost camera [J]. Signal Processing Image Communication, 2015, 36: 29-42.
- [13] CHO D C, KIM W. Long-Range gaze tracking system for large movements [J]. IEEE Transactions on Biomedical Engineering, 2013, 60(12): 3432-3440.
- [14] NAQVI R A, ARSALAN M, BATCHULUUN G, et al. Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor [J]. Sensors, 2018, 18(2): 456-464.
- [15] 孟春宁, 白晋军, 张太宁, 等. 基于低成本眼动记录系统的视线估计研究 [J]. 光电子激光, 2013, 24(8): 1600-1605.
- MENG CH N, BAI J J, ZHANG T N, et al. Research on line of sight estimation based on low-cost eye tracking systems [J]. Optoelectron Laser, 2013, 24(8): 1600-1605.
- [16] 单兴华, 王增才, 范柏旺. 基于改进随机森林的驾驶员视线估计的方法 [J]. 传感器与微系统, 2021, 40(5): 33-37.
- DAN X H, WANG Z C, FAN B W. A method for driver line of sight estimation based on improved random forest [J]. Sensors and Microsystems, 2021, 40(5): 33-37.
- [17] 闫秋女, 张伟伟. 基于多模态特征融合的驾驶员注视区域估计 [J]. 计算机与数字工程, 2022, 50(10): 2217-2222.
- YAN Q N, ZHANG W W. Driver gaze area estimation based on multimodal feature fusion [J]. Computer and Digital Engineering, 2022, 50(10): 2217-2222.
- [18] WANG Y, DING X, YUAN G, et al. Dual-Cameras-Based driver's eye gaze tracking system with non-linear gaze point refinement [J]. Sensors, 2022, 22(6): 2326.
- [19] 张宝龙, 杜康熙, 刘小凡. 全天候车载驾驶员监控系统相机设计 [J]. 电子测量与仪器学报, 2021, 35(6): 109-116.
- ZHANG B L, DU K X, LIU X F. Camera design for all-weather in vehicle driver monitoring system [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(6): 109-116.
- [20] HU H, WU C, LIN K, et al. HG-Net: Hybrid coarse-fine-grained gaze estimation in unconstrained environments [C]. 9th International Conference on Virtual Reality (ICVR), 2023: 1-6.
- [21] ZHANG X, SUGANOY, FRITZ M. MPIIGaze: Real-World dataset and deep appearance-based gaze estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(1): 162-175.
- [22] KELLNHOFER P, RECASENS A, STENT S, et al. Gaze360: Physically unconstrained gaze estimation in the wild [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 6911-6920.

作者简介



程鸣, 武汉大学电子信息硕士, 武汉科技大学汽车交通学院博士研究生, 主要研究方向为智能交通、无人驾驶汽车、驾驶员人机交互、驾驶员疲劳及分心监测等。

E-mail: tc-chengm@dfmc.com.cn

Cheng Ming received a M. Sc. degree in Electronic Information from Wuhan University. He is now a Ph. D. candidate in the School of Automotive Transportation at Wuhan University of Science and Technology. His main research interests include intelligent transportation, autonomous vehicles, human-machine interaction for drivers, driver fatigue, and distraction monitoring.



严运兵 (通信作者), 博士, 教授, 主要研究方向为车辆动力学及其控制、电动汽车能源管理和智能汽车技术等。

E-mail: yanyunbing@wust.edu.cn

Yan Yunbing (Corresponding author) received a Ph. D. degree. He is now a professor. His main research interests include vehicle dynamics and control, electric vehicle energy management, and intelligent vehicle technology.