

基于机器学习与双向时间序列数据的 车辆行驶行为预测模型

陈佳钦, 袁 健

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2023年3月21日; 录用日期: 2023年5月24日; 发布日期: 2023年5月31日

摘 要

自驾车辆跟驰时需密切关注前车的行驶状态, 以此采取加减速、制动等驾驶决策。现有的行为预测模型未考虑与前导车的耦合关系, 且持续感知具有一定局限。因此本文提出基于机器学习与双向时间序列数据的车辆行驶行为预测模型。该模型以轻量方式感知前车, 且引入优化的注意力机制, 解决交通流密集时跟车频繁出现遮挡的误检问题。提取出二维行为向量以双向序列形式输入双层预测模型中, 更精确预测跟驰行为数据变化趋势。实验结果表明, 提出的模型感知精度达84.56%、召回率达80.37%, 相比基准模型平均每帧耗时降低4.95 ms, 并且对前导车的行为预测准确率较高, 能为自驾车辆跟驰提供有效的行驶决策。

关键词

自动驾驶车辆, YOLOv3, 轻量级模型, 行驶行为预测

Vehicle Driving Behavior Prediction Model Based on Machine Learning and Bi-Directional Time Series Data

Jiaqin Chen, Jian Yuan

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 21st, 2023; accepted: May 24th, 2023; published: May 31st, 2023

Abstract

Self-driving vehicles need to pay close attention to the driving status of the preceding vehicle when following, so as to take driving decisions such as acceleration, deceleration and braking. The existing behavior prediction model does not consider the coupling relationship with the leading vehicle, and the continuous perception has certain limitations. Therefore, this paper proposes a prediction model of vehicle driving behavior based on machine learning and bi-directional time series data. The model senses the leading vehicle in a lightweight manner and introduces an optimized attention mechanism to solve the false detection problem of frequent occlusion of following vehicles when the traffic flow is dense. The two-dimensional behavior vectors are extracted and input into the two-layer prediction model in the form of two-way sequences to predict the change trend of following behavior data more accurately. The experimental results show that the proposed model has 84.56% perception accuracy and 80.37% recall rate, and the average time per frame is reduced by 4.95 ms compared with the benchmark model, and the prediction accuracy of the leading vehicle behavior is higher, which can provide effective driving decision for following vehicles.

Keywords

Autonomous Vehicles, YOLOv3, Lightweight Models, Driving Behavior Prediction

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

汽车保有量的增速惠及普罗大众, 给大家带来出行便利的同时, 道路安全始终面临严峻的挑战。由于人为驾驶过程会受主观思维影响, 故制定的驾驶决策不一定能做到安全最优解。而自动驾驶技术在执行制动、加减速、换道、停车等操作时, 可通过智能技术规避人为不可控因素, 做到更精准的感知、决策与执行, 因此自动驾驶技术是目前的研究热点。

跟驰任务作为最基本的行驶行为, 模拟了驾驶者无法改变车道的情况下跟随前车的纵向运动, 是微观交通建模的基本算法[1]。通过检测定位引导车, 实时获取到跟驰车辆间的行驶行为相关变量, 具象化非自由状态下行驶的车辆间的相互影响和作用, 对得到的信息进一步分析理解, 辅助车辆制定后续行车决策。预测跟驰工况中行驶行为的可降低道路交通事故的发生率, 为智能交通发展提供了无限可能。

2. 相关研究

车辆行为预测对车辆感知、辨析车辆行驶行为的能力提出较高的要求。随着自动行驶的发展, 车载传感器等设备的性能大幅度提升, 主要利用视觉传感器、激光雷达等设备实现对前导车辆的识别, 获取行驶过程中的大量行为信息与轨迹数据。其中, 激光雷达能够探测 3D 环境信息[2], 但在多车辆遮挡覆盖时可能会错漏关键特征信息, 且成本原因也成为掣肘激光雷达规模化上车应用的关键。而视觉技术模拟人类视觉系统, 所获取的数据信息占总信息的 80%左右, 更为丰富真实, 因此本文选用视觉技术作为自动驾驶车辆跟驰下的感知方式。

由于搭载在自动驾驶车辆上的视觉设备成本较低、操作相对简易、便于产品化推广等优势, 在环境感知领域得到了广泛的研究。Alvarez 等[3]利用单目彩色摄像机提出利用阴影进行道路检测的方法。Wu 等[4]提出的方法准确拟合车道边界, 并可用于获得有关车道线方向的稳健信息。Jen 等[5]提出了一种基于鲁棒视觉的尾灯信号自主检测方法有效检测刹车灯和转弯信号。以上方法对于环境复杂的非结构化道路检测精度不理想, 在交通流密集道路环境下车辆易频繁覆盖, 无法支撑自动驾驶车辆对于前车准确检测。其次, 自驾车辆需要持续性进行车辆检测, 如何在有限的行车硬件下进行大量数据处理是不容忽视的一点, 检测模型在保证准确性的基础上同样需要考虑实时性。因此本文提出一种 YOLO-SA 检测网络满足检测精度的同时减少行车计算成本。

为保证自动驾驶车辆在跟驰工况下对前车进行持续性的行为预测, 在感知行为信息后需要对该信息进行理解分析。随着硬件计算能力的提升, 近年来研究者们利用深度神经网络对预测算法进行了一番探索。Ding 等[6]提出基于时延反向传播神经网络实时预测变道轨迹。Zhou 等[7]提出了基于递归神经网络的微观汽车跟驰模型, 能够准确捕捉和预测交通流动态。Altche 等[8]使用广泛应用于行人轨迹预测的长短时记忆网络来预测高速公路上的汽车轨迹。以上方法能合理显示与预测复杂的交通现象, 但是参数过多导致行车计算耗时长, 无法满足自动驾驶在实际应用中对实时性的要求。其次, 用卷积网络提取数据可能会出现数据损失的情况从而影响预测精度。最后, 跟驰下的行为预测不能忽略车辆间的耦合关系, 需要基于本车视角实时探测前车的与本车间的相对信息。因此本文以本车作为数据采集载体, 通过基于神经网络的方法处理行为特征量, 并在行车过程中持续预测前车状态以提前应对行车风险。

针对视觉传感器无法准确描述行驶行为信息的问题, 本文提出基于机器视觉与时间序列数据的车辆行驶行为预测模型(Vehicle Driving Behaviour Prediction Model Based on Machine Vision and Time Series Data, 简称 VDBP)。该模型融合图像处理与时间序列数据思想, 以速度、相对距离等参数信息体现行驶车辆的行驶状态, 并通过深度学习算法对输出的数据信息源进行数学拟合, 根据变化趋势预测行驶车辆的未来行驶行为。VDBP 模型的创新工作如下: 1) 将视觉技术引入自动驾驶跟驰应用中, 一体化目标检测、数据处理、行驶行为规律探索三个维度, 以此实现自驾车辆行为预测闭环。2) 为解决持续检测造成的数据冗余与计算速度低下的问题, 本文提出轻量级 YOLO-SA 目标检测算法, 在特征提取网络中利用深度可分离卷积替换标准卷积, 并添加改进后的注意力机制, 规避降维的同时使网络赋予感兴趣特征更高的权重; 3) 针对现有预测算法处理自动驾驶车辆不断迭代行为参数时易出现数据丢失的情况, 本文将行为预测当作序列生成任务, 从正反两方向学习时间序列下过去与未来的双向行为参数, 获取更为精确的跟驰行为数据预测变化趋势。

3. VDBP 模型结构

车辆跟驰任务首先实现对单行车道中引导车辆的检测定位, 其次需要计算该车辆的位置和速度信息, 最后以自学习的方式对相关参数信息进行后续纵向行驶行为预测, 辅助车辆实时调整自身跟驰行驶状态。因此需要三个模块分别实现。VDBP 模型结合环境感知与数据预测, 将二维图像信息转换为时序数据信息, 为跟驰下前导车辆的行为作出预测辅助自驾车辆后续驾驶决策规划。该模型由三个采用层进式的模块组成: 前导车辆视觉检测模块(Leading Vehicle Vision Inspection Block, 简称 LVVIB)、参数计算模块(Behavior Information Processing Block, IPB)和行为预测模块(Behavioral Prediction Block, BPB), 总体结构如图 1 所示。车载摄像头采集该车跟驰工况下引导车图像, 在 LVVIB 模块中轻量化卷积网络, 并引入本文改进的注意力机制, 确定得到最终引导车目标矩形区域, IPB 模块计算检测到的引导车目标与本跟驰车的行为参数, 生成二维参数行为特征量, 最后 BPB 模块通过 PVBP (Pioneering vehicle behavior prediction algorithm, 简称 PVBP 算法)模型学习车辆在时序下两个参数的变化, 挖掘轨迹数据的深层信息。

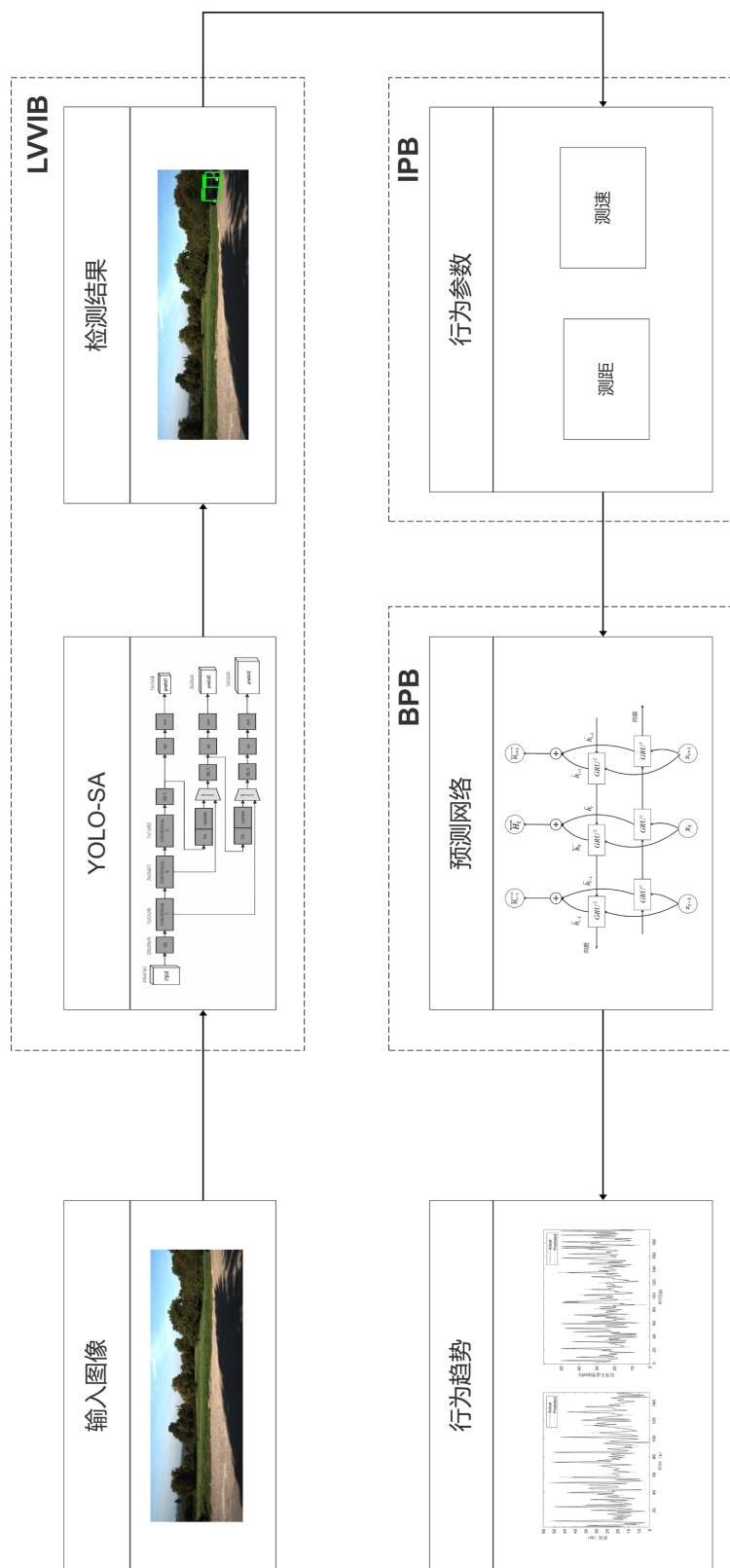


Figure 1. General structure of VDBP
 图 1. VDBP 总体结构

通过基于机器视觉下目标感知、数据处理、行为预测的自动驾驶跟驰应用闭环方案, 在感知层面上可以平衡跟驰工况中对前导车检测的精度与实时性, 相较基准检测模型更关注行车环境出现的多车辆遮挡重叠等非线性条件因素, 除此之外基于车辆本身特性探究单车道上跟驰车辆间的行驶行为, 辅助自动驾驶车辆提前做出加减速、制动等行车策略, 确保行驶安全性与舒适性, 提前应对碰撞风险。

3.1. 前导车辆视觉检测模块 LVVIB

引导车检测是自动驾驶车辆跟驰模型研究的关键技术, 其检测结果的时空轨迹数据作为跟驰模型的输入变量。由于道路环境的复杂性和传感器的限制, 现有算法仍有改进优化之处。由于自动驾驶车辆在中交通流密度相对较大的行驶环境中需要持续性动态定位引导车辆, 初步识别和定位待测目标车辆就需具有较高准确性与实时检测能力。因此本文在 YOLOv3 [9]的基础上提出一种改进的轻量级检测网络 YOLO-SA 进行实时目标检测, 利用深度可分离卷积大幅度降低计算量, 并引入优化后的通道注意力机制一自适应调整特征权值提升通道与通道间特征相关性, 保证检测的鲁棒性。

YOLO-SA 网络在 YOLOv3 模型基础上进行优化, 具体工作在于: 1) 通过深度可分离卷积轻量化特征提取网络结构, 大幅度减少参数; 2) 引入改进的注意力机制 ECSKNet, 降低交通背景对车辆检测的干扰, 优化模型检测精度。

3.1.1. 轻量化特征提取网络结构

YOLOv3 采用多尺度预测结构, 在某些卷积层之后进行下采样, 然后以上采样和张量拼接的方式产生三种不同尺寸大小的特征图, 最终融合多尺度特征的空间信息和语义信息。YOLOv3 特征提取网络 DarkNet-53 结构复杂, 以堆叠残差单元的方式提取特征, 导致有大量权重参数, 提取速度效果不佳。为了满足自动驾驶车辆在交通流密集的行车条件下持续性检测前导车辆, 并在有限行车计算资源的条件下提高模型处理速度, 本文常规卷积替代为深度可分离卷积大大减少参数数量。

$D_k \times D_k \times 1$ 深度卷积和标准点卷积组成普通卷积, 如图 2 所示。

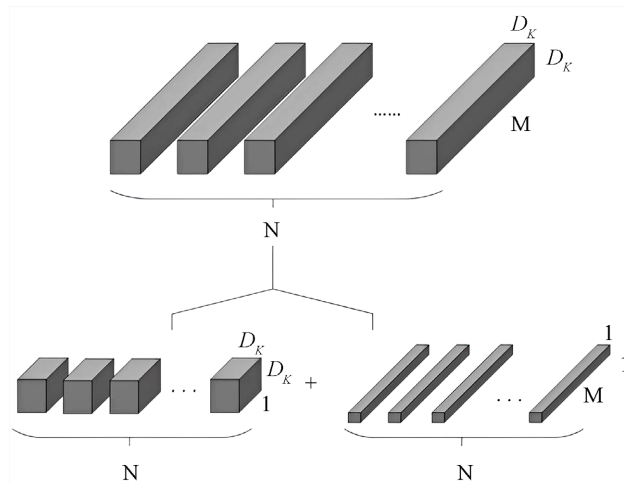


Figure 2. Depthwise separable convolution structure
图 2. 深度可分离卷积分解图示

利用深度可分离卷积与使用普通卷积的参数数量比值如式(1)所示:

$$\frac{D_k \cdot D_k \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_k \cdot D_k \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_k^2} \tag{1}$$

其中 D_F 输入特征矩阵的宽高, D_K 指卷积核尺寸, M 、 N 分别为输入输出特征矩阵的通道数。若采用 2×2 的卷积核即 $D_K = 2$, 通过深度可分离卷积可将参数量缩减为原来的四分之一, 而输出图像尺寸未发生变化, 效果十分显著。

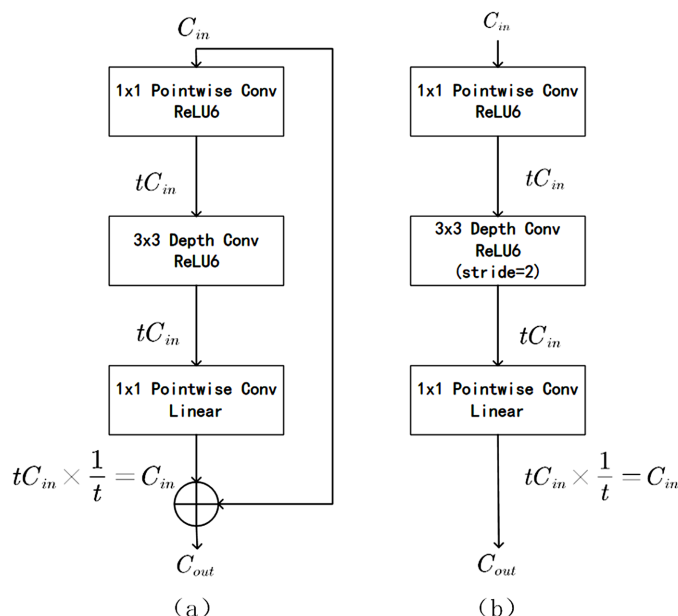


Figure 3. The basic residual unit ECSK-Residual

图 3. 基本残差单元 ECSK-Residual

除此之外, 在残差单元最后添加线性激活函数 ReLU6, 减轻输出通道数较少时的信息耗损。最终组成提取网络的基本残差单元 ECSK-Residual 设计结构如图 3 所示, t 表示升维倍数。

其中 a 为步长为 1 时, 若输入矩阵通道数等于输出矩阵通道数, 则做 shortcut 分支直接将输入与输出卷积做加法融合, 最后添加改进后的注意力机制 ECSKNet (于 3.1.2 详述)。b 为步长为 2 时, 首先输入特征通过 1×1 点卷积与 ReLU6 激活函数对特征进行升维, 其次利用 3×3 深度卷积层提取特征, 然后通过点卷积进行降维, 最后还是经过一个注意力机制单元加强特征提取效果, 低维信息特征通过该方法可有效表征。

3.1.2. 添加改进的通道注意力机制 ECSKNet

轻量化改造通过减少参数量使得计算速度有明显提升, 但参数数量的缩减也意味着特征表征能力会相对下降。因此在残差单元最后一个点卷积后添加改进的注意力机制模块 ECSKNet 构成 ECSK-Residual。通过在输出不同尺度特征图的三个分支后各添加该基本单元, 赋予通道特征权重系数体现各个通道特征图与重要特征的相关性, 突出重要特征。

ECSKNet 在通道注意力 SKNet [10] 的基础上进行优化, 减少全连接操作涉及降维对通道信息造成的损耗, 如图 4 所示。ECSKNet 在 Fuse 阶段融合不同尺度与路径的语义信息, 构成特征图 U 后对每个通道进行全局平均池化操作, 压缩成一维目标车辆特征 S 统计各通道特征信息, 如式(2)、(3)所示:

$$U = U_1 + U_2 \quad (2)$$

$$S_c = F_{gp}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (3)$$

其中, S_c 表示 S 的第 c 个通道, U_c 表示 U 的第 c 个通道, F_{gp} 为全局平均池化。之后通过大小为 k 的一维卷积处理输出维度为 $C \times 1 \times 1$ 的特征 Y 。 k 随着通道维数 C 动态调整, 代表该通道交互信息的覆盖范围, 捕获每个通道与其相邻 k 个通道实现局部跨通道交互。映射关系 φ 如式(4)所示, 其中, $|x|_{\text{odd}}$ 代表参数取最近的奇数。

$$k = \varphi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (4)$$

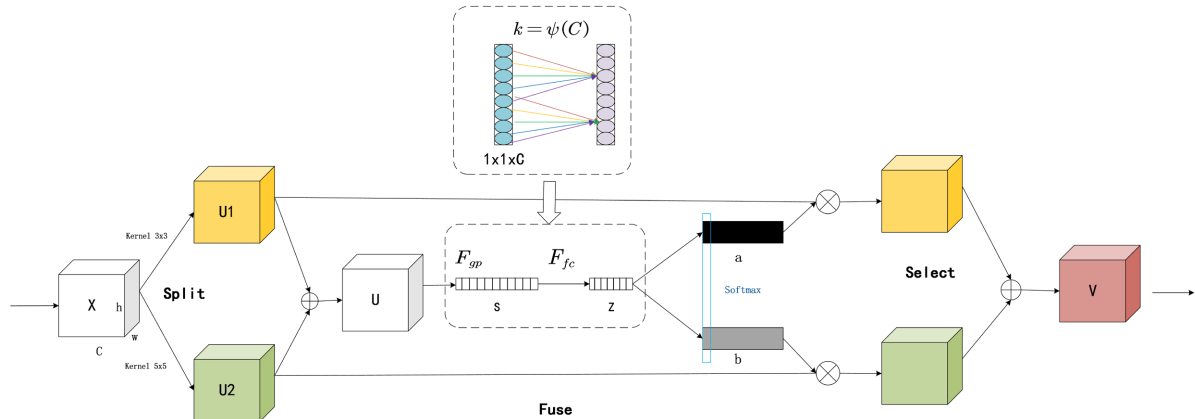


Figure 4. Advanced attention mechanism structure
图 4. 改进的注意力机制模块结构

ECSKNet 注意力模块无需降维即可以极轻量化的方式获取跨通道交互信息, 显式地建模通道之间的关系以此来强化特征学习, 使网络在降低计算复杂度和优化模型性能之间达到平衡。最终优化后的目标检测网络 YOLO-SA 结构如图 5 所示。

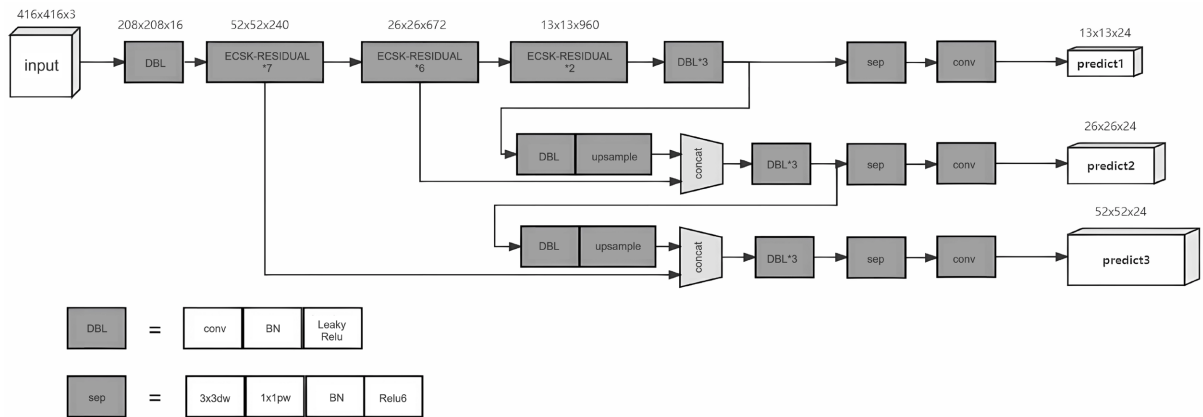


Figure 5. YOLO-SA network structure
图 5. YOLO-SA 网络结构

3.2. 参数计算模块 IPB

为满足自动驾驶车辆持续关注前方引导车行为状态, 需要基于目标定位区域有效提取行为信息。而对于交通流较密集的跟驰场景中, 更需要实时利用当前行为信息作为预测下一时刻行为信息的数据来源。因此本文利用 LVVIB 模块生成的目标区域, 计算出前车距离与两车相对速度并共同组成二维行为特征需

求 $P(Z,T)$, 用于表示前车的相对信息。

3.2.1. 前车距自车距离 Z

单目测距可以理解为将现实三维世界投射到图像二维平面中, 可采用小孔成像模型进行描述此类投影, 如图 6 所示。

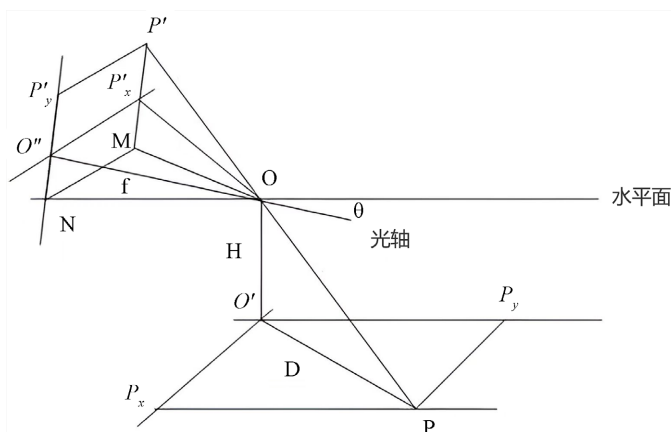


Figure 6. Keyhole imaging model

图 6. 小孔成像模型

其中 f 为单目相机焦距, (x,y) 为目标车辆在二维图像中的像素坐标, H 为相机高度, h 为目标像素点到水平线的垂直距离, Z 为至目标的水平距离, θ 为相机的俯仰角。根据图 6 与相似三角形原理, 可得出前车距自车的距离 Z , 如式(5)~(7)所示:

$$\frac{H}{\sqrt{H^2 + D^2}} = \frac{h}{\sqrt{x^2 + y^2 + f^2}} \quad (5)$$

$$f \cdot x \cdot (y + f \cdot \tan \theta) = x \cdot \sqrt{f^2 + h \cdot (\tan \theta)^2} \quad (6)$$

$$Z = \frac{H \cdot \sqrt{x^2 + (x^2 + y^2) \cdot (\tan \theta)^2 + f^2} - 2f \cdot y \cdot \tan \theta}{f \cdot \tan \theta + y} \quad (7)$$

3.2.2. 相对速度 V

设摄像机以一定帧速采集行车视频图像, 采集周期为 T 。从第 p 帧到第 q 帧历经一段行驶时间, 该行驶间隔 Δt 如式(8)所示:

$$\Delta t = (n_p - n_q) \times T \quad (8)$$

N 帧内引导车在世界坐标系中映射的坐标分别为 $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, 利用帧差法可计算得到两车相对行驶距离 ΔS 如式(9)所示, 式中 S_0 为两车初始位置距离。

$$\Delta S = S_0 + \frac{1}{N-1} \times \sum_{i=2}^N \sqrt{(X_i - X_{i-1})^2 + (Y_i - Y_{i-1})^2} \quad (9)$$

结合式(8)、(9)计算可得跟驰自动车与引导车之间的相对运动速度 V , 如式(10)所示:

$$V = \frac{1}{2} \times \left(\frac{\Delta S_1}{\Delta t_1} + \frac{\Delta S_2}{\Delta t_2} \right) \quad (10)$$

3.3. 行为预测模块 BPB

为确保行车安全与舒适, 实时预测前导车行为可辅助自动驾驶车辆第一时间采取相应驾驶决策, 避免突然刹车与碰撞情况的发生, 减轻对交通流的负面影响。因此本模块提出一种适用于交通流密集情况下的跟驰前导车行为预测算法 PVBP 从正反两个方向提取并学习 IBP 模块所输出时间序列下引导车的行为参数, 拟合预测值与真实值, 获取更为精确的跟驰行为数据预测变化趋势。

GRU [11]改善了 RNN 梯度消失导致的短时记忆问题, 且 GRU 拥有更少参数不易产生过拟合, 保持预测效果的同时收敛性与训练效率更优, 在短时交通流预测方面广泛适用。因此, PVBP 算法以 GRU 神经网络作为基础预测网络, 并在此基础上进行优化, 如图 7 所示。PVBP 算法核心思想为: 搭建两层双向预测模型, 且选取鲁棒性更高的激活函数, 将未来与过去的背景信息进行有效融合, 而非依赖正向时间序列或逆向时间序列导致数据缺失。通过该双层双向的 PVBP 算法训练顺序与逆序两类二维行为参数序列 $P(Z, T)$ 可进一步分析时序下行为参数间的相关性, 减少循环神经网络伴随的遗忘性与梯度消失等问题, 具体工作流程如下:

Step1: 构建基础 GRU 神经网络。鉴于传统 GRU 模型采用的隐状态激活函数存在软饱和性会导致梯度消失, 故采用 ReLU 线性非饱和激活函数将其替代, 极大提高梯度下降的收敛速度。在计算完更新门 z_t 和重置门 r_t 后, 利用 r_t 重置前的隐藏层数据 h_{t-1} 输出得到候选隐藏层 \tilde{h}_t , 引入 ReLU 激活函数后数据基础传输过程如式(11)~(14)所示:

$$z_t = \delta(W_z \cdot [h_{t-1}, x_t]) \quad (11)$$

$$r_t = \delta(W_h \cdot [h_{t-1}, x_t]) \quad (12)$$

$$\tilde{h}_t = \text{ReLU}(W_h \cdot x_t + U_h(h_{t-1} \odot r_t)) \quad (13)$$

$$h_t = (1 - z_t) \odot h_{t-1} + \tilde{h}_t \odot z_t \quad (14)$$

式中, W_h 、 U_h 为 GRU 权值矩阵, \odot 为矩阵的哈达马积, x_t 为当前时刻输入, h_{t-1} 为上一时刻输出, z_t 、 r_t 分别为更新门和重置门, $[\]$ 为矩阵相连。

Step2: 考虑到预测跟驰车辆行驶行为的相关参数与时间具有一定关联性, 即这一时刻的速度与相对距离不仅与当前行驶状态有关, 还与上一时刻相应参数有紧密关系。为防止预测过分依赖单向数据流导致的数据遗忘问题, 更好提取过去与未来双重背景信息提高预测精度, 本文利用 Step1 中构建的基础神经网络对正序二维行为参数序列 $P(Z, T)$ 进行正反两个方向训练。

正向输出 \bar{h}_t 表示按时序从前往后预测未来数据, 反向输出 \tilde{h}_t 表示按时序从后往前预测过去数据, 在每一个时刻点 t 内对二者进行融合得到结果 \bar{H}_t , 如式(15)~(17)所示, 其中 c 为比例系数。

$$\bar{h}_t = \text{GRU}^1 = \text{GRU}(\bar{h}_{t-1}, x_t) \quad (15)$$

$$\tilde{h}_t = \text{GRU}^2 = \text{GRU}(\tilde{h}_{t-1}, x_t) \quad (16)$$

$$\bar{H}_t = [\bar{h}_t \oplus \tilde{h}_t] = [(1-c)\bar{h}_t + c\tilde{h}_t, c\bar{h}_t + (1-c)\tilde{h}_t] \quad (17)$$

Step3: 搭建第二层双向预测网络, 同时重复 Step2 的方法对逆序序列 $\hat{P}(\hat{Z}, \hat{T})$ 进行正反方向的训练, 得出 $\tilde{\hat{H}}_t$ 。

Step4: 将正序逆序二维行为参数序列的输出结果 \bar{H}_t 与 $\tilde{\hat{H}}_t$ 进行融合, 如式(18)所示:

$$Q_t = [\bar{H}_t, \tilde{\hat{H}}_t] \quad (18)$$

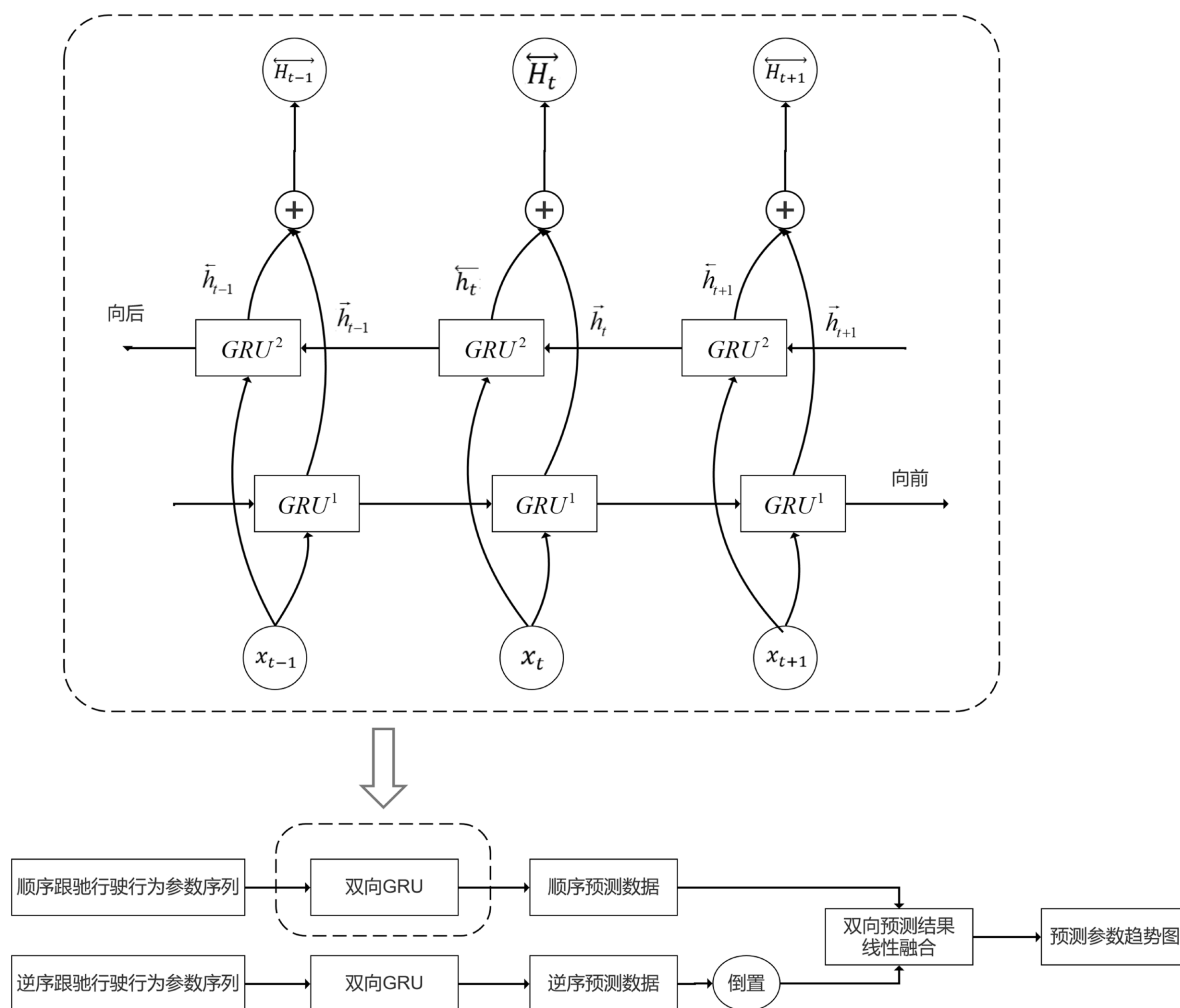


Figure 7. The process of pioneering vehicle behavior prediction algorithm
图 7. PVVP 算法工作流程

4. 实验结果与分析

4.1. 数据集及训练细节

本文模型采用 PyTorch 框架搭建, 操作系统为 Ubuntu18.04, GPU 为 NVIDIA RTX 2080Ti (显存 11 GB)。首先为验证 YOLO-SA 检测网络的有效性, 本文选择公开数据集 KITTI [12]-Object 进行模型训练。该数据集由 7481 个训练图像和 7518 个测试图像组成, 所有图像均为彩色并保存为 png 格式, 保留数据集中所标注的 Car, Van, Truck、Tram 四种车辆相关类别, 并且对原始分辨率为 1242×375 的图像做基本数据增强预处理。训练时采用 Adam 优化器, 设置学习率为 0.001, 且使用学习率衰减, 衰减系数为 0.1, 训练时批处理大小设置为 32, 总共迭代 1000 个 Epoch。

为验证跟驰工况下, 行为预测算法 PVBP 能持续处理前导车感知信息与预测其行驶行为, 本文选择 KITTI-Tracking 数据集进行模型训练。该数据集由 21 个训练序列和 29 个测试序列组成。由于序列的特殊性, 随机选择一个训练序列, 并将连续帧合成 10 fps 视频序列。对每一帧完成 YOLO-SA 检测后采用 SORT 跟踪算法[12]持续跟踪前车轨迹, 对当前视频帧进行目标检测获取检测框, 从而计算得出行为参数

数据集, 对其进行归一化, 即线性映射到区间[0,1], 并以 8:2 的比例划分训练数据集与测试数据集。将时长为 140 s 的数据输入模型进行预测评估, 预测往后 5 秒的行驶行为参数, 通过迭代训练调整参数以达到预测模型较高性能。训练时采用 Adam 优化器, 不断优化调试最终确定超参数, batch_size 设为 32, 隐含层神经元数为 768, 初始学习率设为 0.002。不断增加迭代次数, 学习率也会随着变化, 迭代 10 个 epoch 后, 最终学习率为 0.00005。

4.2. 评价指标

检测实验以精度(Precision)、召回率(Recall)作为算法性能定量评价指标。精度所有检测出的目标中检测正确的概率, 亦称查准率, 召回率是指所有的正样本中正确识别的概率, 亦称查准率, 公式如式(19)、(20)所示, 得出指标数值越高代表模型效果越好。

$$\text{Precision} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (19)$$

$$\text{Recall} = \frac{N_{TP}}{N_p} \quad (20)$$

其中 N_{TP} 为预测正确的车辆数量, $N_{TP} + N_{FP}$ 为所有的预测为正样本的数据, 即生成的预测框总数, N_p 为实际车辆总数。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (|y_i - \hat{y}_i|)^2} \quad (21)$$

预测实验中将采用均方根误差 RMSE (Root Mean Square Error)、平均绝对百分比误差 MAPE (Mean Absolute Percentage Error)作为性能指标评估算法。RMSE 值反映真实值与预测值的误差程度, 如式(21)所示, 该值越小, 偏差越小, 预测更为精准。

MAPE 以百分比形式反映真实值与预测值的误差绝对值的平均值, 如式(22)所示, 可以用来对比不同算法预测相同数据时的预测差异性, 该值越小则说明预测值偏离真实值的程度越小, 预测效果越好。其中, N 代表预测实验样本量, y_i 代表每个行为参数真实值, \hat{y}_i 代表对应预测值。

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (22)$$

4.3. 前导车辆视觉检测实验结果与分析

图 8 可视化展示本文算法在 KITTI 数据集上的检测效果。图 8(a)选取出现大量重叠覆盖时的车辆行驶环境, 图 8(b)选取存在远距离小目标的车辆行驶环境, 且受光线影响多数车辆隐于阴影背景之下, 车辆外观线条并不清晰。而 YOLO-SA 较为准确地框定了车辆具体位置, 关注到频繁遮挡的车辆、感知范围内远距离的小目标车辆, 因此检测效果较为理想。这是因为 YOLO-SA 网络添加改进后的注意力机制, 使得在有限的计算资源中更加关注重要信息而非无关信息, 避免行车过程中帧与帧之间频繁切换车辆 ID、以及车辆覆盖等造成漏检问题, 在持续感知前车的过程中由不错的稳定性。

为验证 YOLO-SA 网络模型对于检测效能的提升, 本文以 YOLOv3 作为基准模型, 在此基础上加入 SKNet、本文改进后的 ECSKNet 以及深度可分离卷积, 通过比较精度、召回率、平均每帧耗时对该网络进行消融实验, 对比结果如表 1 所示。

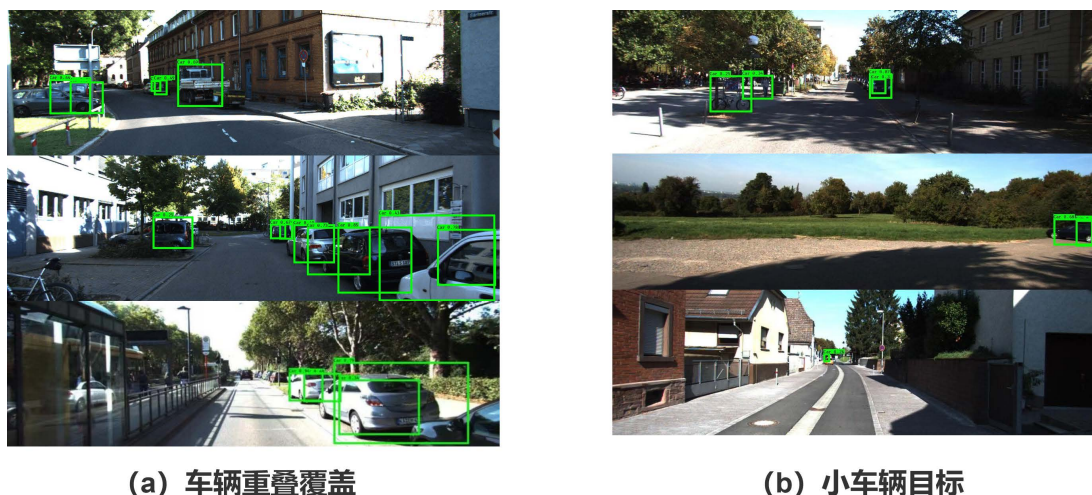


Figure 8. Detection results of vehicle object by YOLO-SA
图 8. YOLO-SA 目标车辆检测结果

Table 1. Comparison of algorithm performance
表 1. 算法性能对比

组件	是否使用				
YOLOv3	√	√	√	√	√
SKNet		√			
ECSKNet			√		√
深度可分离卷积				√	√
精度	72.31	78.29	80.85	69.04	84.56
召回率	70.10	76.64	77.09	70.82	80.37
平均每帧耗时(ms)	64.38	65.22	64.93	57.79	59.43

由表 1 可知, 在 YOLOv3 中引入改进后的注意力机制 ECSKNet, 相比基准模型 YOLOv3 与仅适用 SKNet 注意力机制, 在精测精度与召回率上有一定的优化, 分别提升 8.53% 与 6.99%、2.56% 与 0.45%。且通过一维卷积处理以代替全连接的操作在耗时上也有一定缩减。除此之外考虑到实际行车场景, 通过引入深度可分离卷积轻量化特征提取网络可以降低耗时与参数量, 相比基准 YOLOv3 平均每帧耗时缩减 6.59 ms。综合分析得出本文提出的 YOLO-SA 在检测准确性与实时性上均有不错的表现, 更适宜车辆实时持续检测。

4.4. 行为预测实验结果与分析

图 9 展示了本文提出的 PVBP 算法的预测走势, 可以看出预测值与真实值误差较小, 预测精度理想, 这是因为该算法在预测部分考虑将过去与未来双重信息进行融合, 数据复用提升预测精度, 并且通过引入 ReLU 激活函数来缓解原先模型易出现梯度消失的问题。

在预测部分选取采用 tanh 激活函数的传统 GRU 基准模型、采用 ReLU 的 GRU 模型与本文提出 PVBP 算法进行性能对比。表 2 列举了三种算法在 RMSE、MAPE 与训练时间上的不同表现。可以看出, 由于模型结构的增加, 训练时长高于传统 GRU 模型, 但在 RMSE 与 MAPE 指标上都有不错的表现, 皆小于其他两种算法。由定义可知, 本文提出的 PVBP 算法预测走势与真实值之间的误差最小。

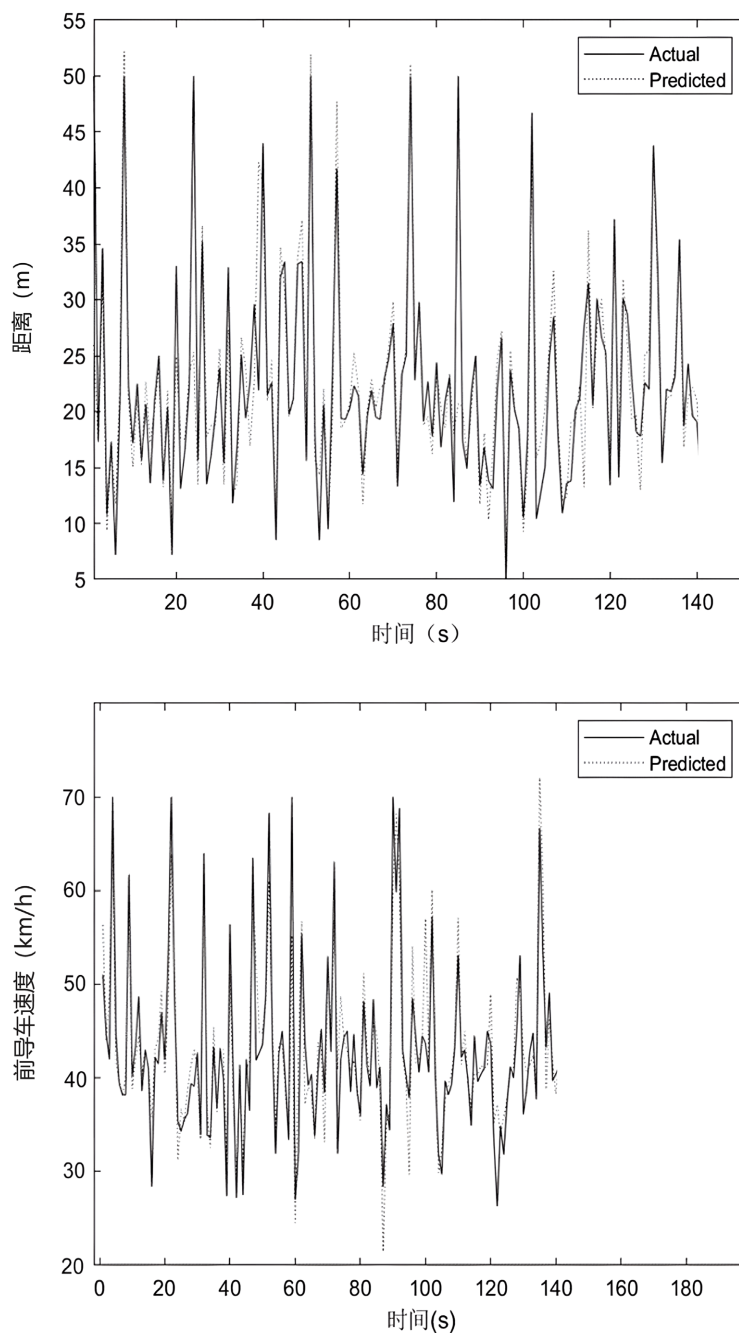


Figure 9. PVPB algorithm model prediction results grape
 图 9. PVPB 算法模型预测结果图

Table 2. Comparison of model prediction performance
 表 2. 预测性能对比

模型	RMSE	MAPE	训练时长(s)
GRU + tanh	2.26444	5.8910	157.22
GRU + ReLU	1.84814	5.7232	203.19
本文	1.11170	5.3184	194.52

5. 结论

本文提出的行驶行为预测模型 VDBP, 实现对自动行驶车辆在跟驰工况中的行驶行为预测, 辅助车辆实时调整自身驾驶策略。首先通过构建轻量级的提取网络模型, 并且引入优化后的注意力机制突出重要信息对网络的提取表达能力, 实时准确地对引导车辆进行定位与测距。其次采用帧差法计算行驶行为参数。最后通过 PVBP 算法预测模型学习车辆在时序下两个参数的变化, 挖掘轨迹数据的深层信息。实验结果表明, 本文提出的检测网络在精度与计算开销上优于基本检测网络, 且能更精确预测对引导车距离、速度, 有助于拓宽自动行驶运行策略, 具有一定实用价值。

由于实验条件限制, 今后将进一步结合安全预警机制, 通过预测出的行驶行为趋势为跟驰车辆行驶过程中出现的危险情况做提前预警, 实现该技术的应用落地。

参考文献

- [1] 魏福祿, 刘杨. 经典跟驰模型的发展综述[J]. 吉林建筑大学学报, 2018, 35(4): 19-23.
- [2] Alvarez, J. and Lopez, A.M. (2011) Road Detection Based on Illuminant Invariance. *IEEE Transactions on Intelligent Transportation Systems*, **12**, 184-193. <https://doi.org/10.1109/TITS.2010.2076349>
- [3] Wu, C.F., Lin, C.J. and Lee, C.Y. (2012) Applying a Functional Neurofuzzy Network to Real-Time Lane Detection and Front-Vehicle Distance Measurement. *IEEE Transactions on Systems Man and Cybernetics Part C*, **42**, 1-13. <https://doi.org/10.1109/TSMCC.2011.2166067>
- [4] Jen, C.L., Chen, Y.L. and Hsiao, H.Y. (2017) Robust Detection and Tracking of Vehicle Taillight Signals Using Frequency Domain Feature Based Adaboost Learning. *Proceedings of IEEE International Conference on Consumer Electronics*. Piscataway, 12-14 June 2017, 423-424. <https://doi.org/10.1109/ICCE-China.2017.7991176>
- [5] Ding, C., Wang, W., Xiao, W., et al. (2013) A Neural Network Model for Driver's Lane-Changing Trajectory Prediction in Urban Traffic Flow. *Mathematical Problems in Engineering*, **2013**, Article ID: 967358. <https://doi.org/10.1155/2013/967358>
- [6] Zhou, M., Qu, X. and Li, X. (2017) A Recurrent Neural Network Based Microscopic Car Following Model to Predict Traffic Oscillation. *Transportation Research Part C Emerging Technologies*, **84**, 245-264. <https://doi.org/10.1016/j.trc.2017.08.027>
- [7] Altche, F. and Fortelle, A. (2017) An LSTM Network for Highway Trajectory Prediction. *Proceedings of the 20th IEEE International Conference on Intelligent Transportation System*. Piscataway, 16-19 October 2017, 353-359. <https://doi.org/10.1109/ITSC.2017.8317913>
- [8] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, 18 April 2018, 89-95.
- [9] Li, X., Wang, W., Hu, X., et al. (2019) Selective Kernel Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, 15-18 March 2019, 510-519.
- [10] Cho, K., Merriënboer, B.V., Gulcehre, C., et al. (2014) Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of Conference on Empirical Methods in Natural Language Processing*. 25-29 October 2014, 1724-1734. <https://doi.org/10.3115/v1/D14-1179>
- [11] Geiger, A., Lenz, P. and Urtasun, R. (2012) Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, 16-21 June 2012, 3354-3361. <https://doi.org/10.1109/CVPR.2012.6248074>
- [12] Bewley, A., Ge, Z., Ott, L., et al. (2016) Simple Online and Realtime Tracking. *Proceedings of IEEE International Conference on Image Processing*. Piscataway, 25-28 September 2016, 3464-3468. <https://doi.org/10.1109/ICIP.2016.7533003>