

10-18-2022

Fatigue Detection Method Based on Facial Features and Head Posture

Rongxiu Lu

1.School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China;2.Key Laboratory of Advanced Control and Optimization of Jiangxi Province, Nanchang 330013, China;; ecjtu_rxlu@163.com

Bihao Zhang

1.School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China;2.Key Laboratory of Advanced Control and Optimization of Jiangxi Province, Nanchang 330013, China;

Zhenlong Mo

3.School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Fatigue Detection Method Based on Facial Features and Head Posture

Abstract

Abstract: Aiming at the of the single fatigue characteristics, low robustness and inability to customize fatigue thresholds for different drivers of fatigue detection methods, a method based on facial features and head posture is proposed. *In face detection and face key point positioning HOG feature operator and regression tree algorithm are used. In head posture estimation, head posture Euler angle is estimated by combining the face key points with the coordinate system transformation. In fatigue feature extraction, a deep residual neural network model is established to extract the eye fatigue features, which the eye, mouth aspect ratio and head posture Euler angle. The fatigue characteristics of eyes, mouth and head are used to establish the support vector machine models for different drivers to provide the early fatigue driving warning.* Experiments show that on YawDD and self-built fatigue simulation data sets, the method shows high accuracy and robustness, and can provide better fatigue warning when a certain fatigue feature detection is blocked.

Keywords

fatigue driving, face detection, head posture, deep learning, support vector machine

Recommended Citation

Rongxiu Lu, Bihao Zhang, Zhenlong Mo. Fatigue Detection Method Based on Facial Features and Head Posture[J]. Journal of System Simulation, 2022, 34(10): 2279-2292.

基于脸部特征和头部姿态的疲劳检测方法

陆荣秀^{1,2}, 张笔豪^{1,2}, 莫振龙³

- (1. 华东交通大学 电气与自动化工程学院, 江西 南昌 330013;
2. 江西省先进控制与优化重点实验室, 江西 南昌 330013;
3. 华东交通大学 交通运输与物流学院, 江西 南昌 330013)

摘要: 针对疲劳驾驶检测方法存在疲劳特征单一、鲁棒性低和不能因驾驶员不同定制疲劳阈值等问题, 提出了一种基于脸部特征和头部姿态的疲劳检测方法。利用HOG(histogram of oriented gradients)特征算子和回归树算法进行人脸检测和人脸关键点定位; 通过脸部关键点结合坐标系变换估计头部姿态欧拉角; 建立深度残差神经网络模型对眼部疲劳特征进行提取, 同时结合眼部、嘴部纵横比和头部姿态欧拉角进行疲劳特征提取; 利用眼部、嘴部和头部姿态疲劳特征建立针对不同驾驶员的支持向量机模型对疲劳驾驶进行预警。实验表明: 在YawDD和自建疲劳模拟数据集上, 该方法均表现出较高的准确率和鲁棒性, 在某一疲劳特征检测受阻时依然能进行较好的疲劳预警。

关键词: 疲劳驾驶; 人脸检测; 头部姿态; 深度学习; 支持向量机

中图分类号: TP391

文献标志码: A

文章编号: 1004-731X(2022)10-2279-14

DOI: 10.16182/j.issn1004731x.joss.21-0583

Fatigue Detection Method Based on Facial Features and Head Posture

Lu Rongxiu^{1,2}, Zhang Bihao^{1,2}, Mo Zhenlong³

- (1. School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China;
2. Key Laboratory of Advanced Control and Optimization of Jiangxi Province, Nanchang 330013, China;
3. School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China)

Abstract: Aiming at the of the single fatigue characteristics, low robustness and inability to customize fatigue thresholds for different drivers of fatigue detection methods, a method based on facial features and head posture is proposed. In face detection and face key point positioning HOG feature operator and regression tree algorithm are used. In head posture estimation, head posture Euler angle is estimated by combining the face key points with the coordinate system transformation. In fatigue feature extraction, a deep residual neural network model is established to extract the eye fatigue features, which the eye, mouth aspect ratio and head posture Euler angle. The fatigue characteristics of eyes, mouth and head are used to establish the support vector machine models for different drivers to provide the early fatigue driving warning. Experiments show that on YawDD and self-built fatigue simulation data sets, the method shows high accuracy and robustness, and can provide better fatigue warning when a certain fatigue feature detection is blocked.

Keywords: fatigue driving; face detection; head posture; deep learning; support vector machine

收稿日期: 2021-06-23

修回日期: 2021-08-09

基金项目: 国家自然科学基金(61733005, 61863014, 61963015)

第一作者: 陆荣秀(1976-), 女, 博士, 副教授, 研究方向为智能检测、机器视觉技术应用等。E-mail: ecjtu_rlxu@163.com

引言

随着社会经济的飞速发展,我国汽车保有量不断增加,安全隐患也日益剧增。据统计,疲劳驾驶是引发交通事故的主要因素,在疲劳驾驶导致交通事故前,驾驶员会有明显的先前疲劳特征,此时给予驾驶员疲劳预警可以极大程度避免交通事故。疲劳驾驶具有基数大、占比高和可预见的特点。研究如何对驾驶员的疲劳特征进行快速、准确、稳定地检测,对于交通安全驾驶问题的改善有着重大意义。

早期的疲劳检测是通过使用相关医疗器械对人体的生理指标进行检测进而判断是否疲劳。目前,在交通领域对驾驶员进行疲劳检测的方法主要有 4 类:①基于主观评定法,通过主观自评或主观他对驾驶员疲劳进行量化,该方法是离线分析方法,不能在发生事故前对驾驶员进行疲劳预警;②基于生理参数特征检测,通过对人体接入电极或仪器来检测心电、脑电等信号进行疲劳检测^[1-2],这种方法虽然能够实时准确地发出疲劳预警,但是由于采集特征信号会对驾驶员产生入侵性,影响驾驶员在驾驶过程中的舒适感;③基于车辆行为特征检测,通过车载计算机记录分析汽车的方向盘转角和角速度^[3],或通过传感器检测方向盘握力对驾驶员是否疲劳进行判定,这类方法会随着车辆使用时间的增加,硬件设备的老化而精度下降,且成本高不易于普及;④基于面部特征检测,是使用机器视觉技术获取驾驶员人脸图像特征,通过机器学习或深度学习算法对人脸状态进行分析并判断是否疲劳。如 Wang M 等^[4]使用 Adaboost 和轮廓圆结合的机器学习算法,根据疲劳判定准则 PERCLOS 阈值判断眨眼。Mandal B 等^[5]通过检测驾驶员头肩和眼睛睁闭判定公交车驾驶员是否疲劳。Zhao Xuepeng 等^[6]通过建立级联卷积神经网络判断驾驶员是否疲劳驾驶。Omidyeganeh M 等^[7]使用反投影理论检测驾驶员嘴部的变化,建立了一个哈欠检测系统。这类方法具有精度高、成本低、易普及、无干扰等优点,但存在疲劳特征参数单一、网络层数较浅和不能根

据驾驶员的不同定制疲劳阈值等问题。

随着计算机运算能力的不断加强和数据量的成倍增长,深度学习在数据挖掘、分类预测、模式识别等方面表现出巨大的优势,其强大的特征学习功能,对驾驶员有部分遮挡、光影变化等干扰都具有较强的适应性和鲁棒性。本文针对疲劳特征参数单一、疲劳特征提取不够深和不能定制疲劳阈值等问题,采用多种疲劳特征辨识的方法,使用深度残差神经网络来深层次提取特征,并使用 SVM 模型来实现定制疲劳阈值和提高系统鲁棒性和检测精度。

1 基于脸部特征和头部姿态的疲劳检测原理

本文提出的基于脸部特征和头部姿态疲劳检测原理如图 1 所示,主要包括:①人脸检测环节,通过方向梯度直方图(histogram of oriented gradients, HOG)算法和 ERT(ensemble of regression trees)算法对实时图像进行人脸检测并定位人脸关键点;②特征提取环节,根据关键点信息提取眼部、嘴部疲劳特征,并通过图像坐标系与世界坐标系的转换求解出驾驶员头部欧拉角;③疲劳检测环节,在提取的眼部 ROI(region of interest)图像基础上,采用深度残差神经网络进行特征提取并对其进行眨眼的判断,同时结合经过驾驶员个性化预训练的支持向量机模型输出的眼部、嘴部和头部姿态分类结果完成对驾驶员疲劳驾驶的预警。

2 人脸定位与头部姿态估计

对驾驶员进行人脸检测和关键点定位,并计算头部姿态欧拉角,可为后续眼部、嘴部和头部多疲劳特征提取做准备。

2.1 基于 HOG 的人脸检测算法

方向梯度直方图 HOG 算法是计算机视觉领域较为优秀的图像检测算法,使用 HOG 算法进行人脸检测的具体过程如图 2 所示。

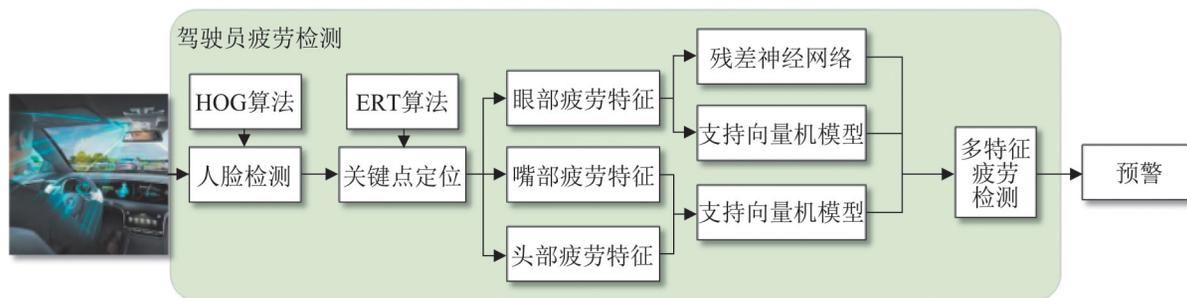


图 1 基于脸部特征和头部姿态疲劳检测方法原理图

Fig. 1 Schematic diagram of fatigue detection method based on facial features and head posture

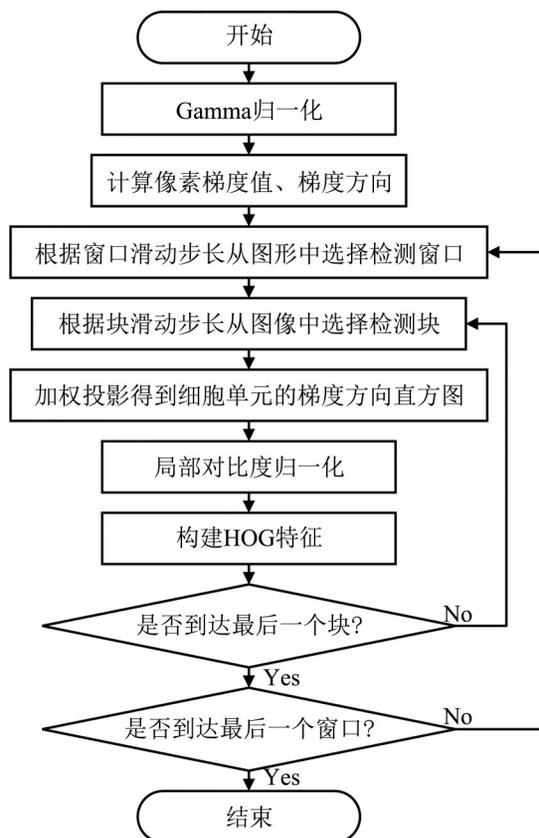


图 2 HOG 算法流程图

Fig. 2 HOG algorithm flowchart

具体步骤如下:

step 1: 为提高算法对外界环境的抗干扰性, 对摄像头输入的每一帧图像进行灰度图转化, 再利用 Gamma 归一化算法减少和抑制光照、颜色、阴影对图像的干扰, Gamma 归一化公式为

$$I(x, y) = I(x, y)^{\text{Gamma}} \tag{1}$$

式中: (x, y) 为图像中的像素点。

step 2: 通过导数计算图像中每个像素点 (x, y) 的梯度值 $G(x, y)$ 和梯度方向 $\alpha(x, y)$, 要计算像素点 (x, y) 的横、纵坐标方向的梯度:

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \tag{2}$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \tag{3}$$

式中: $G_x(x, y)$ 和 $G_y(x, y)$ 为像素点 (x, y) 水平方向和垂直方向上的梯度值; $H(x, y)$ 为像素值, 图像中每个像素点的梯度值和梯度方向为

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \tag{4}$$

$$\alpha(x, y) = \arctan \left[\frac{G_y(x, y)}{G_x(x, y)} \right] \tag{5}$$

step 3: 将输入的图像划分为 N 个 cells 单元, 每个 cells 中有 $n \times n$ 个像素, 再将 cells 单元的梯度方向从 $0 \sim 360^\circ$ 划分为 M 个区间, 对每个 cells 单元根据式(4)、式(5)进行加权投影输出 N 个 cells 单元的梯度方向直方图。

step 4: 进行局部对比度归一化后, 对每个 cells 单元进行滑动扫描提取 HOG 特征, 并送入分类器进行分类, 在输出结果中采用非极大值抑制 (non-maximum suppression, NMS) 的方法剔除多余的人脸候选框。

鉴于疲劳检测的视频图像中, 驾驶员的人脸占比较大, 采用基于 HOG 特征的人脸检测算法可以有效地识别有部分遮挡的人脸并能精确地标出轮廓候选框, 为下一步的人脸关键点定位奠定基础。

2.2 基于ERT的人脸关键点定位算法

人脸关键点定位就是使用基于优化或回归的方法将人脸的特征点进行准确定位。鉴于基于回归树集合(ensemble of regression trees, ERT)^[8]的人脸对齐算法,相比于传统的算法具有可以在执行形状不变特征选择的同时最大范围地减少训练时相同的损失函数,并能解决标签缺失或不确定的问题,本文采用基于回归树集合算法建立的预训练人脸关键点检测模型进行关键点定位。该模型由来自AFW^[9]、HELEN^[10]、LFPW^[11]和IBUG^[12]数据集组成的300W^[13]人脸基准数据集训练而成。其中基于ERT算法建立预训练模型原理如下:

首先,建立一个级联回归器:

$$\hat{\mathbf{S}}^{(t+1)} = \hat{\mathbf{S}}^{(t)} + r_t(L, \hat{\mathbf{S}}^{(t)}) \quad (6)$$

式中: $\hat{\mathbf{S}}^{(t)}$ 为人脸形状索引; $r_t(L, \hat{\mathbf{S}}^{(t)})$ 为回归器,每个回归器由 n 棵树组成,每一棵树都是由当前形状和标定形状坐标差与随机选取的像素对进行训练所得。若 N_i 为第 i 张人脸图像, \mathbf{S}_i 为图像的形状向量,训练数据集 $(N_1, \mathbf{S}_1), (N_2, \mathbf{S}_2), \dots, (N_n, \mathbf{S}_n)$, 设输入参数 $\{(N_i, \hat{\mathbf{S}}_i^{(t)}, \Delta \mathbf{S}_i^{(t)})\}_{i=1}^n$, 对回归器进行初始化:

$$f_0(N, \hat{\mathbf{S}}^{(t)}) = \arg \min_{\gamma \in \mathbb{R}^{2P}} \sum_{i=1}^n \|\Delta \hat{\mathbf{S}}_i^{(t)} - \gamma\|^2 \quad (7)$$

$$\Delta \mathbf{S}_i^{(t)} = \mathbf{S}_i - \hat{\mathbf{S}}_i \quad (8)$$

对回归器中 r_t 根据式(9)进行迭代并更新

$f_k(N_i, \hat{\mathbf{S}}_i^{(t)})$:

$$r_{ik} = \Delta \mathbf{S}_i^{(t)} - f_{k-1}(N_i, \hat{\mathbf{S}}_i^{(t)}) \quad (9)$$

$$f_k(N_i, \hat{\mathbf{S}}_i^{(t)}) = f_{k-1}(N_i, \hat{\mathbf{S}}_i^{(t)}) + g_k(N, \hat{\mathbf{S}}_i^{(t)}) \quad (10)$$

式中: $g_k(N, \hat{\mathbf{S}}^{(t)})$ 为弱分类器的回归方程, k 为迭代次数,最终训练完成的树的回归器为

$$r_t(N, \hat{\mathbf{S}}^{(t)}) = f_k(N_i, \hat{\mathbf{S}}_i^{(t)}) \quad (11)$$

该算法通过优化损失函数和误差总和来学习Tree,最终通过预训练模型可获得人脸图像关键点坐标位置。

2.3 驾驶员头部姿态估计

在真实驾驶环境中,驾驶员的头部姿态估计

主要运用于判断驾驶员是否处于疲劳打瞌睡状态,以保证安全,减少事故。在利用预训练的人脸关键点检测模型获得人脸图像关键点位置的基础上对驾驶员头部姿态进行估计,需将以世界坐标系表示的3D人脸坐标映射到以图像坐标系表示的2D关键点坐标,转换原理如图3所示。

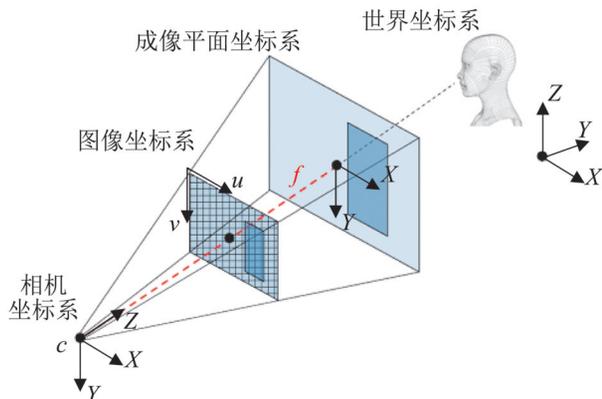


图3 坐标系统原理图

Fig. 3 Principle of coordinate system

在图3中,从左自右依次为相机坐标系、图像坐标系、成像平面坐标系和世界坐标系。由于系统是通过相机提取人脸图像,因此需要首先将世界坐标系映射到相机坐标系或者说是将坐标系与摄像机对齐^[14],其中世界坐标系中的 x 轴和 y 轴与图像的 x 轴和 y 轴对齐,正 z 轴与摄像机的光轴对齐,坐标系转换公式为

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (12)$$

式中: \mathbf{R} 为旋转矩阵; \mathbf{T} 为平移矩阵。由图3可知,从相机坐标系到图像坐标系之间的转换其实是一种透视变换,在透视变换的过程中产生了一个成像平面坐标系,其中成像平面在相机坐标系中的深度为焦距 f ,转换关系为

$$\lambda \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (13)$$

成像平面坐标系通过式(14)映射到图像坐标系中

$$\begin{bmatrix} X_p \\ Y_p \\ 1 \end{bmatrix} = \begin{bmatrix} m_x & s & x_0 \\ 0 & m_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \quad (14)$$

式中: m_x, m_y 为每个图像方向的像素比例因子; s 为传感器的倾斜因子; x_0, y_0 为光轴和像素平面的交叉点。

最终将式(12)~(14)相结合, 得到图像坐标系与世界坐标系的映射关系:

$$\begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \lambda \begin{bmatrix} fm_x & fs & x_0 \\ 0 & fm_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} X_p \\ Y_p \\ 1 \end{bmatrix} \quad (15)$$

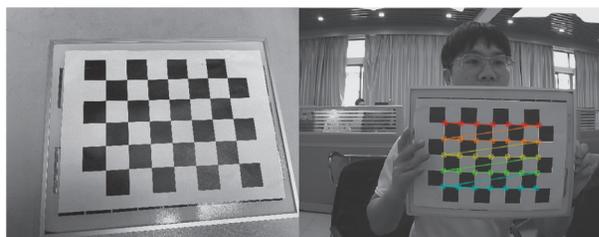
简化为

$$\begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \mathbf{X}_w = \lambda \mathbf{K}^{-1} \mathbf{X}_p \quad (16)$$

式中: 矩阵 \mathbf{K} 为摄像机参数矩阵, 通过相机标定的方法可以获取相机的内参矩阵。

本文通过使用张正友提出的相机标定方法进行标定^[15], 首先准备一张角点为 6×4 的棋盘(如图4), 对棋盘图片的角点和亚像素角点信息进行提取, 最后标定函数使用这些点生成矩阵, 并且生成一个畸变系数矢量, 用于消除相机的径向畸变, 最终输出的相机内参矩阵为

$$\mathbf{K} = \begin{bmatrix} 552.971 & 707 & 0 & 328.309 & 116 \\ 0 & 553.432 & 949 & 236.908 & 246 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$



(a) 国际象棋棋盘 (b) 相机标定实验图

图4 棋盘格与相机标定实验

Fig. 4 Checkerboard and camera calibration experiment

在 3DMM^[16] 获取到世界坐标系中心位置后, 得到像素坐标位置和相机的内参便可以利用旋转矩阵 \mathbf{R} 和平移矩阵得到驾驶员头部的欧拉角:

$$\mathbf{R} = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \cos \phi \cos \gamma \\ \cos \phi \sin \gamma + \sin \phi \sin \phi \cos \gamma \\ \sin \phi \sin \gamma - \cos \phi \sin \phi \cos \gamma \end{bmatrix} + \begin{bmatrix} -\cos \phi \cos \gamma \\ \cos \phi \sin \gamma + \sin \phi \sin \phi \sin \gamma \\ \sin \phi \cos \gamma + \cos \phi \sin \phi \sin \gamma \end{bmatrix} + \begin{bmatrix} \sin \phi \\ -\sin \phi \sin \phi \\ \cos \phi \sin \phi \end{bmatrix} \quad (17)$$

$$\begin{cases} \phi = \arctan(-r_{12}, r_{22}) \\ \phi = \arctan(r_{02}, \sqrt{r_{12}^2 + r_{22}^2}) \\ \gamma = \arctan(-r_{01}, r_{00}) \end{cases} \quad (18)$$

至此便完成了驾驶员人脸检测和关键点定位, 并实时地估算出驾驶员头部姿态的欧拉角, 接下来将对驾驶员脸部关键点和头部欧拉角进行疲劳特征的提取。

3 驾驶员疲劳特征提取

3.1 眼部疲劳特征提取

在疲劳检测中驾驶员眼睛在一定时间序列中的眨眼次数可以作为判断驾驶员疲劳的生理指标, 一般在正常驾驶中, 驾驶员两次眨眼的间隔时间约为 4 s, 眼睛闭合到睁开的时间约为 0.2 s, 在驾驶员产生困倦时, 眨眼速度会变慢约 0.6 s, 在相同时间序列中的眨眼频率也会增加, 所以可以根据单位时间序列中的眨眼次数和闭眼时长来判断驾驶员是否疲劳。PERCLOS 是国际上公认的疲劳状态判定准则, 其表示在单位时间内, 眼睛处于闭合状态的时间所占的比例。由于在视频中, 帧数序列在时间和内容上都具有连续性, 故可以利用眼睛闭合状态的帧数来近似求解 PERCLOS 值, 本文将其定义为眨眼频率:

$$f_{eye} = T_{eye} / M \quad (19)$$

式中: M 为单位时间序列视频总的帧数; T_{eye} 为闭眼的帧数。根据眼部几何特征和神经网络检测单

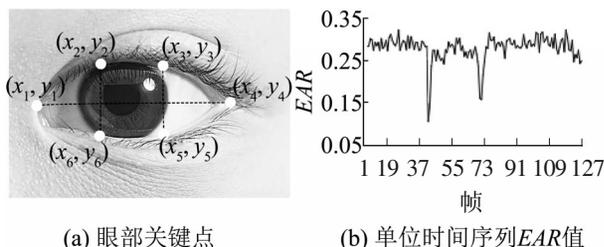
位时间序列的眨眼次数后，便可提取眼部疲劳特征。

3.1.1 基于眼睑纵横比EAR的疲劳特征提取

经过上文人脸检测和关键点定位后，提取驾驶员眼部 12 个关键点坐标，由于两只眼睛的眨眼反射一致，故取一只眼睛的 6 个关键点进行计算，其坐标由左至右顺时针分别为 $(x_1, y_1)(x_2, y_2) \dots (x_6, y_6)$ ，如图 5(a)所示。本文利用人眼的几何特征引入眼睑纵横比(eye aspect ratio, EAR)对驾驶员眼睛睁闭状态进行估计：

$$EAR = \frac{\|y_2 - y_6\| + \|y_3 - y_5\|}{2\|x_1 - x_4\|} \quad (20)$$

对于由红外摄像头输入的每一个视频帧，检测出驾驶员眼部的 12 个关键点，基于 EAR 计算眼睛高度和宽度欧式距离比，结果如图 5(b)所示，当驾驶员的眼睛睁开时，EAR 的值围绕 0.28 上下波动，基本处于恒定，当眼睛闭上时 EAR 的值会迅速下降并趋于 0。



(a) 眼部关键点 (b) 单位时间序列EAR值

图5 眼部关键点与EAR值

Fig. 5 Key points of the eye and EAR value

由图 5(b)，根据计算 EAR 值判断驾驶员的实时眨眼行为，结合眨眼频率 f_{eye} ，可以判定驾驶员眼部是否处于非正常驾驶状态。

由于驾驶员的眼部行为较为复杂，依据单一的 EAR 值作为眼部特征会导致检测系统精度和鲁棒性低。因此本文提出使用一种深度残差神经网络与 EAR 相补充的方法对眼部进行特征提取，以减少对驾驶员眨眼的漏检。

3.1.2 基于深度残差神经网络的眼部特征提取

传统的卷积神经网络由于卷积的层数不能深

度叠加^[17]，导致提取的特征层次不够深，网络精度不够高。本文采用一种深度残差神经网络(deep residual network, ResNet)进行眼部特征的提取，相比于传统神经网络因层数不断叠加而产生的梯度消失和网络退化的问题，残差神经网络在其基础上增加残差单元用以解决此问题，其网络结构如图 6 所示。

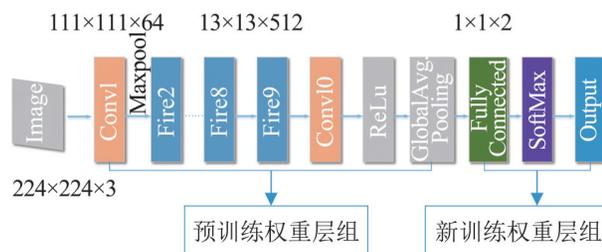


图6 残差神经网络结构图

Fig. 6 Residual neural network structure diagram

设卷积神经网络的损失函数为 $f(X, W)$ ，在进行训练时反向传播算法的公式为

$$\frac{\partial Loss}{\partial X} = \frac{\partial f(X, W)}{\partial X} \quad (21)$$

若往多层神经网络拓展将变为

$$Loss = f_n(X_n, W_n), L_n = f_{n-1}(X_{n-1}, W_{n-1}), \dots, L_2 = f_1(X_1, W_1) \quad (22)$$

计算可得第 i 层的梯度表达式为

$$\frac{\partial Loss}{\partial X_1} = \frac{\partial f_n(X_n, W_n)}{\partial X_n} \times \frac{\partial f_{n-1}(X_{n-1}, W_{n-1})}{\partial X_{n-1}} \times \dots \times \frac{\partial f(X_{i+1}, W_{i+1})}{\partial X_i} \quad (23)$$

从式(23)不难看出当神经网络在进行误差反向传播时，浅层网络的梯度随之减小最后导致网络的梯度消失。为了解决这一问题，残差神经网络在此基础上引入一个残差单元如图 7(b)所示。

相较于传统的神经网络输出为 $H(x)=F(x)$ ，图 7(b)中的残差神经网络通过加入一个同等映射 x 将输出变为 $H(x)=F(x)+x$ ，其中这一映射将 x 作为网络训练过程中的底线，若 x 经过卷积单元 $F(x)$ 训练后不满足网络精度要求，网络可以将这一单元块的权重置为 0，保证网络的梯度不会减小。基于此，式(23)将变为

$$\frac{\partial X_{i+1}}{\partial X_i} = \frac{\partial X_i + \partial f(X_n, W_n)}{\partial X_i} = 1 + \frac{\partial f(X_{i+1}, W_{i+1})}{\partial X_i} \quad (24)$$

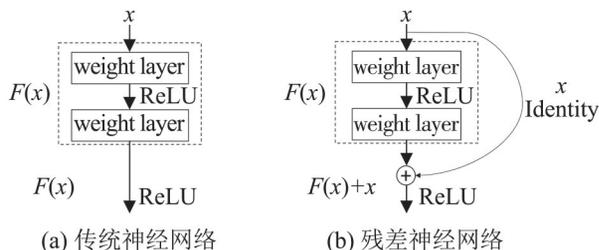


图 7 残差神经网络与传统神经网络对比
Fig. 7 Comparison of residual neural network and plain network

由式(24)可知, 随着网络深度的增加, 浅层网络梯度也不会消失。

深度学习网络的层数越深需要的数据量越大, 以往疲劳驾驶检测的研究中所使用的 CNN 大多使用人脸数据集而非真实驾驶环境中的人脸图像进行卷积神经网络的特征提取和训练, 虽然在数量上满足了深度学习网络的训练, 但是在质量、真实性上并不能满足疲劳驾驶检测的要求。针对目前公开的真实驾驶环境数据集少的问题, 本文提出一种基于 ResNet 的迁移学习模型^[18]。迁移学习模型就是将源领域和目标领域的数据从原始特征空间映射到新的特征空间中去, 在该空间中, 源领域的数据与目标领域的分布相同, 从而可以在新的空间中, 更好地利用源领域已有的有标记数据的样本进行分类训练, 最终对目标领域的数据进行分类测试。本文通过迁移学习将 ResNet-50 网络权重参数迁移到眼部残差神经网络中, 提高模型的学习效率, 该模型利用多层特征融合, 和层间判别特征, 并通过 softmax 回归将其融合进行分类。

从图 6 可知, 残差神经网络的输入大小为 224×224 , 输出层为 2 维向量, 分别表示驾驶员眼睛的睁眼和闭眼 2 个状态。首先, 在对眼睛区域进行 ROI 提取分割后, 将 ROI 区域尺度变换为 $224 \times 224 \times 3$ 的 RGB 颜色空间, 使图像经过一个

stride 为 2, 像素填充 pad 为 3, kernel 为 7×7 的卷积层进行特征提取并将图像的大小减小为原来的一倍; 其次, 经过激活函数 ReLU 后用 stride 为 2, 3×3 的最大池化层将眼部图像进行降采样; 然后, 驾驶员的眼部 ROI 图像依次通过 52 次 kernel 为 3×3 的卷积和最大池化操作进行深度特征提取和压缩, 其中多个小的 kernel 堆叠使用相比于使用一个大的 kernel 能减小大量计算参数, 并且使用多个 kernel 可以带来更深层次的网络、学习到更复杂的非线性关系^[19]。实验证明 3 个 3×3 的小 kernel 堆叠的效果与使用一个 7×7 的 kernel 相同, 但是在一个含有 n 个 kernel 的卷积层, 前者的参数个数为 $3 \times (3 \times 3 \times n)$, 而后者参数为 $7 \times 7 \times n$, 由此残差神经网络中卷积层的 kernel 都为 3×3 ; 最后将经过深层次提取的眼部特征输入到平均池化层和全连接层, 全连接层将特征数据映射到一个二维向量中, 再将该向量输入到 softmax 函数中进行回归, 最终输出睁眼和闭眼的概率值用于判断驾驶员的眨眼情况, 与 EAR 算法相补充, 在某一方法出现漏检时也能较好地检测到驾驶员的眨眼, 以提高系统的鲁棒性。

3.2 嘴部疲劳特征提取

驾驶员处于疲劳状态时, 大脑通常会释放打哈欠等特征性行为信号, 以提醒人体状态疲劳需要休息。因此, 在驾驶车辆过程中, 驾驶员的嘴部有三种状态: 松弛闭合、说话和哈欠疲劳。根据 2.2 节人脸关键点定位可知, 驾驶员嘴部外轮廓定位的坐标由左至右顺时针分别为 $(x_1, y_1), (x_2, y_2), \dots, (x_{12}, y_{12})$ 共 12 个坐标点如图 8(a) 所示, 本文利用嘴部的几何特征引入嘴部纵横比 (mouth aspect ratio, MAR) 的概念对驾驶员嘴部状态进行估计:

$$MAR = \frac{\|y_3 - y_{11}\| + \|y_5 - y_9\|}{2\|x_1 - x_7\|} \quad (25)$$

对于由红外摄像头输入实时图像帧, 在一定时间序列上计算 MAR 的实时值如图 8(b) 所示。

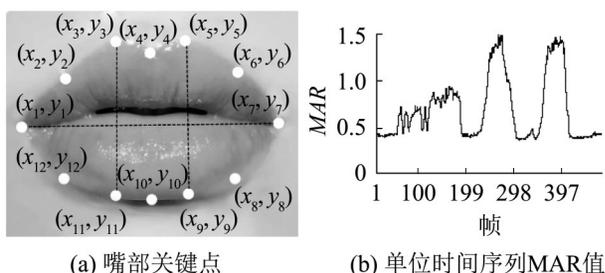


图8 嘴部关键点与MAR值
Fig. 8 Key points of mouth and MAR value

由图8可知,当驾驶员处于说话时MAR值大约稳定在[0.6, 1.0],且持续时间一般较长,当驾驶员嘴部出现哈欠状态时, MAR会迅速上升到1.2~1.5,也就是说,驾驶员嘴部的3种状态(即哈欠、说话、闭嘴松弛)特征较为明显,只要对嘴部的定位精准,不需要通过深度学习进行特征提取,即可快速计算MAR判断驾驶员的嘴部状态,以节省车载计算机CPU的算力。

由此可根据实时计算的MAR判断驾驶员的说话、哈欠行为,结合前文对驾驶员眨眼的判断,在驾驶员出现眼部遮挡时(佩戴墨镜),提供辅助的检测手段用于疲劳预警。

3.3 头部姿态疲劳特征提取

驾驶员处于严重疲劳状态时,身体和头部会出现前倾、左右摇晃等危险姿势,使用低头频率 f_{head} 和头部异常角度可以判断驾驶员的疲劳状态:

$$f_{head} = T_{head} / M \quad (26)$$

式中: M 为单位时间序列视频总的帧数; T_{head} 为低头的帧数。将驾驶员头部姿态的欧拉角分为俯仰角、偏航角、翻滚角,即驾驶员头部分别绕 x 、 y 、 z 轴旋转的欧拉角,如图9所示。根据研究^[20],一般人体头部绕中轴的 x 轴旋转的最大角度范围是 $[-60.4^\circ, 69.6^\circ]$,绕 y 轴旋转的最大角度范围是 $[-90^\circ, 75^\circ]$,绕 z 轴旋转的最大角度范围是 $[-40.9^\circ, 36.3^\circ]$ 。

由此可以根据驾驶员实时的俯仰角、偏航角、翻滚角来计算异常角度,本文对于异常角度的判

断参考PERCLOS中P80标准,即当驾驶员头部角度超过正常范围的80%时判断为异常角度。驾驶员瞌睡行为导致的头部姿态异常主要体现在俯仰角和翻滚角,经实验后取驾驶员头部俯仰角平均变化大于 25° ,翻滚角平均变化大于 15.4° 判定为驾驶员的头部异常角度。

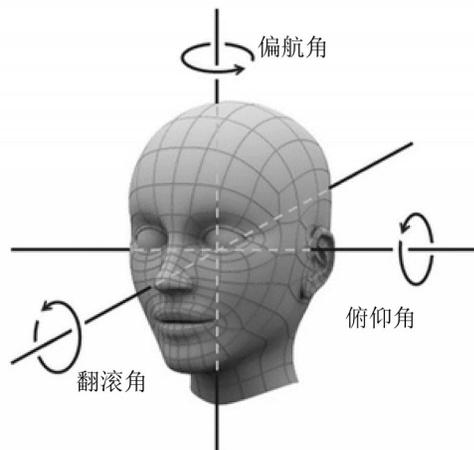


图9 驾驶员头部转角
Fig. 9 Driver's head angle

4 基于SVM的疲劳辨识模型

通常一辆汽车的常驻驾驶员有1~3名,由于每位驾驶员的眼睛大小、嘴部大小和头部姿态角度并不相同,若对不同的驾驶人员采用单一且统一的阈值判定驾驶员眨眼、说话、哈欠和头部低头动作不仅精度和鲁棒性低,还不能满足不同驾驶员对车载检测系统的个性化需求。因此本文提出使用基于支持向量机(SVM)的机器学习算法建立疲劳辨识模型,以提高系统的检测精度和鲁棒性,并且满足不同驾驶员对疲劳检测系统的个性化需求。

支持向量机算法是主要用于解决模式识别领域的分类问题的一种高精度且高效的有监督机器学习算法。本文通过将图像数据转换为关键点坐标,再将其转换为EAR、MAR值和头部姿态欧拉角俯仰角、偏航角、翻滚角,然后利用SVM算法针对不同驾驶员建立不同的疲劳辨识模型,原理如图10所示。

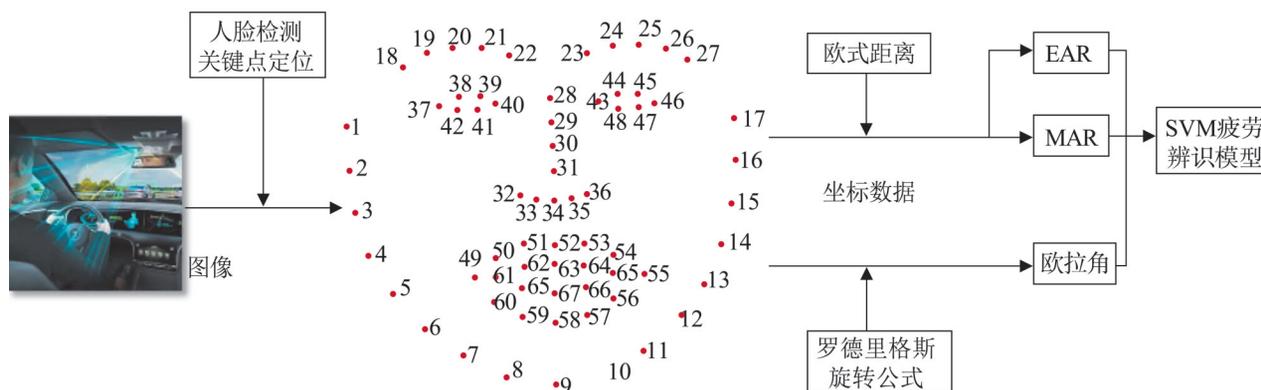


图 10 SVM 疲劳辨识模型原理
Fig. 10 SVM fatigue identification model principle

使用 SVM 机器学习算法建立疲劳辨识模型的具体步骤为:

(1) 人脸信息采集。在驾驶员第一次上车时, 要求在采集系统的提示下配合系统进行眨眼、张嘴、低头等动作进行一次短暂的人脸信息采集, 如图 11 所示。目的是采集不同驾驶员在不同状态的眼睑长宽比、嘴部长宽比和头部姿态的欧拉角。



图 11 采集系统实时图
Fig. 11 Real-time diagram of acquisition system

(2) 特征提取。对驾驶员眼部、嘴部、头部数据进行特征提取并组成 n 维向量。本文通过提取 5 名实验样本在实际驾驶环境中睁眼、闭眼的 EAR, 嘴部闭嘴松弛、正常交谈、哈欠的 MAR 和头部正常驾驶时和疲劳异常动作时的俯仰角、偏航角、翻滚角欧拉角特征数据进行训练。

(3) 训练 SVM 模型。采用编码方式将眼部的

睁闭分别标签为 0、1, 以此类推, 将带标签的数据送入 SVM 分类器进行训练。

5 实验结果与分析

为了验证本文方法的有效性, 实验分为 3 部分: ①眼部深度残差神经网络模型的训练与验证; ②针对不同驾驶员训练的 SVM 疲劳辨识模型的训练与方法对比; ③利用自建疲劳数据集对多特征疲劳辨识模型的验证。

5.1 基于深度残差神经网络模型的眼部特征有效性验证

驾驶员眼部特征提取使用的深度残差神经网络使用 CEW^[21]数据集、YawDD^[22]数据集进行训练。其中 CEW 数据集包含 2 423 张人脸图像; YawDD 数据集是由渥太华大学在真实驾驶环境下利用自然光采集的不同年龄不同种族的视频数据集, 记录了驾驶员正常、说话、哈欠等驾驶状态, 是为数不多的面向疲劳驾驶检测建立的公开数据集。

对 YawDD 视频数据集进行关键帧截取, 最后得到 4 538 张真实驾驶环境的图像, 部分如图 12(a) 所示, 并根据驾驶员状态进行人工标签。对 CEW 数据集和标签好的 YawDD 数据集进行眼部 ROI 裁剪并将维度 resize 到 $224 \times 224 \times 3$, 将颜色三通道降为单通道后部分结果如图 12(b) 所示。



图 12 YawDD 数据集预览与眼部 ROI
Fig. 12 YawDD dataset preview and eye ROI

对深度残差神经网络模型进行训练的具体步骤如下：首先将图像维度 **resize** 到 224×224 ，进行数据增强后，在每个像素中减去在 ImageNet 数据集上计算的平均 RGB 值，做归一化处理，目的是使得迁移学习的精度和效率更加有保障；其次基于 ResNet-50 的网络架构构建深度卷积神经网络 (DCNN)，用 2 维全连接和 softmax 层代替原先的网络层级，并将 ResNet-50 卷积层的参数传递给 DCNN 模型的卷积层；最后通过 Transfer learning 的训练方式进行训练。

Transfer learning 训练是通过冻结传输卷积层参数，这些层可以看作是网络的特征提取层，训练输出为 2 维的全连接和 softmax 层。训练的具体参数调节如下：① 采用动态学习率。动态学习率的采用是使网络更好的逼近最优值，初始的学习率定为 0.001 随后每 7 个 epoch 衰减为原来的 1/10。② 优化器选择 Adam(adam optimizer)来进行网络权重的更新。相比于 SGD 优化器，Adam 有更快的收敛速度和更少内存需求。③ Batch 为 32，epoch 为 100。训练过程中训练集和验证集准确率的变化如图 13 所示，训练结束后保存验证集表现最好的网络模型权重。

由图 13 可知，训练集在 Transfer learning 训练中经过 100 个 epoch 迭代后训练集准确率高达 98.9%，验证集为 97.5%，模型在训练集和验证集上均表现出较高的精度，可与 SVM 模型相结合用作驾驶员眼部疲劳特征辨识模型。

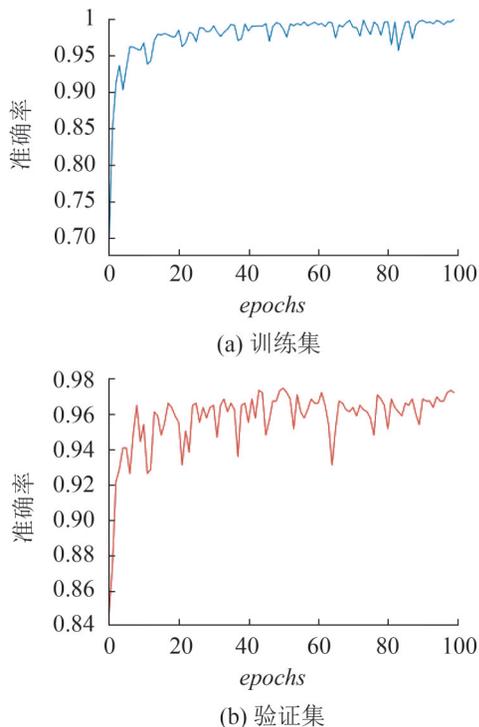


图 13 训练过程结果图
Fig. 13 Result of the training process

5.2 基于 SVM 的疲劳辨识模型

当更换驾驶员时，系统将要求驾驶员配合系统提示，以采集其眼部、嘴部和头部特征数据，然后对不同驾驶员采用基于 SVM 的疲劳辨识模型分别进行训练，以达到系统对不同驾驶员定制疲劳阈值的效果，其中实验采集的部分特征数据如表 1 所示。

表 1 采集数据表
Table 1 Collected data table

测试员编号	EAR		MAR		
	睁眼	闭眼	嘴闭合	说话	哈欠
1	0.274 1	0.085 9	0.372 8	0.640 0	1.521 6
2	0.262 0	0.107 8	0.380 2	0.750 4	1.428 1
3	0.250 3	0.139 4	0.155 2	0.500 3	1.063 1
4	0.291 9	0.159 3	0.295 5	0.867 1	1.160 4
5	0.236 6	0.145 8	0.427 7	0.794 6	1.238 9

由表 1 数据可知，驾驶员在初次上车根据要求进行动作时，采集的眼部、嘴部和头部姿态等特征数据具有趋势性，根据实际状态将数据进行编码标签，然后采用 SVM 疲劳辨识模型进行训练。其中采集驾驶员睁闭眼、张闭嘴、说话、头

部进行正下方低头、左右下方低头和正常姿态的疲劳特征数据各 500 个, 根据采集的特征数据, 每一个驾驶员对应的 SVM 模型训练的样本数量分别为: 眼部 1 000 个、嘴部 1 500 个、头部姿态 2 000 个, 训练模型时参数设置为: 惩罚参数 $C=1.0$; 分类器的核函数选择使用 RBF 径向基核函数; $probability=False$; 停止训练的误差值: 1×10^{-3} 。为验证使用 SVM 算法改进的疲劳检测系统有更高的精确度和鲁棒性, 实验使用自行采集的疲劳驾驶模拟视频作为测试样本。自建疲劳模拟视频数据集由招募的 4 男 1 女在实际驾驶环境中自然光下录制的模拟疲劳视频组成, 其中记录了驾驶员说话、哈欠和低头瞌睡的状态, 截取数据集部分如图 14 所示。



图 14 自建数据集部分预览
Fig. 14 Partial preview of self-built data set

招募 5 名观察员对视频中驾驶员眨眼、说话、哈欠等数量进行统计, 取统计后的均值作为视频中实际眨眼、说话、哈欠、低头瞌睡的数量。实验取单一阈值、视频连续帧的平均阈值同基于 SVM 算法改进的疲劳检测方法进行对比, 其中单一阈值法中眼部取 0.25, 检测到 EAR 值连续 2 帧小于阈值判断眨眼; 哈欠取 0.9 作为阈值, 检测到 MAR 值连续 5 帧大于 0.9 判断哈欠, 平均阈值法取驾驶员睁眼、张嘴时的 EAR 、 MAR 均值作为阈值。对比结果如表 2、3。

表 2 眼部检测算法对比结果

Table 2 Comparison results of eye detection algorithms

视频编号	实际闭眼数	检测闭眼数/准确率%		
		单一阈值	平均阈值	SVM 模型
1	90	78/86.7	83/92.2	89/98.9
2	157	148/94.3	142/90.4	154/98.1
3	131	97/74.0	120/91.6	129/98.5
4	87	60/69.1	77/88.5	85/97.7
5	106	74/69.8	98/92.5	103/97.2

表 3 哈欠检测算法对比结果

Table 3 Comparison results of yawn detection algorithms

视频编号	实际哈欠数	检测哈欠数/准确率%		
		单一阈值	平均阈值	SVM 模型
1	3	2/66.7	3/100	3/100
2	0	0	0	0
3	2	2/100	1/50	2/100
4	5	6/83.3	4/80	6/83.3
5	4	2/50	4/100	4/100

由表 2、3 可见采用单一阈值的疲劳检测系统对某一些(如视频编号 2)驾驶员眼部特征的检测精度很高(如 94.3%), 但对于不同的驾驶员由于个体的差异导致检测精度不稳定, 范围在 69.1%~94.3% 之间波动, 系统鲁棒性低、可移植性差; 采用平均阈值的方法虽然整体检测的准确度较稳定, 对不同驾驶员眼部的检测精度都稳定在大约 90%, 移植性较单一阈值有所提高, 但系统的检测精度仍不能满足要求; 采用本文提出的基于 SVM 算法改进的疲劳检测系统对不同驾驶员眼部特征的检测精度都在 97.2% 以上, 对哈欠的检测精度也较高, 系统鲁棒性高可移植性好, 与采用单一阈值和平均阈值的疲劳检测方法相比有更好的检测性能。

5.3 多特征疲劳辨识模型实验结果

为了验证多特征疲劳辨识模型的有效性, 选取在自然光条件下的真实驾驶环境作为实验条件, 从驾驶员进入车辆开始通过车载摄像头进行人脸检测并输出人脸候选框, 实验结果见图 15(a), 在候选框中对驾驶员人脸进行关键点检测实验结果见图 15(b)。

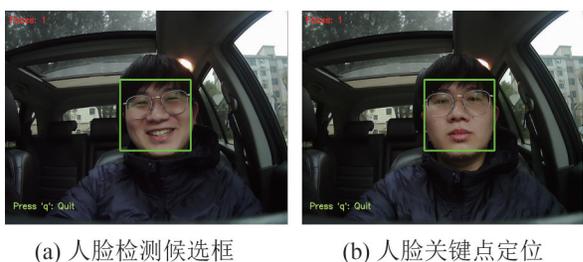


图 15 人脸识别与关键点定位实验结果图
Fig. 15 Experimental results of face recognition and key point positioning

根据关键点提取驾驶员眼部、嘴部、头部特征，这些特征检测之间独立工作互不影响、相互补充，通过分别预警驾驶员眼部、嘴部和头部姿态的疲劳特征，使得模型在某一特征检测受阻时，也能较好的对驾驶员的疲劳进行预警，检测实验结果如图 16 所示。



图 16 多特征疲劳检测实验结果
Fig. 16 Multi-feature fatigue test results

在驾驶员多特征的疲劳预警上，由文献[23]可知，驾驶员眼部PERCLOS与疲劳状态的关系为

$$PERCLOS \begin{cases} < 0.1, \text{ 正常驾驶状态} \\ [0.1, 0.3), \text{ 轻度疲劳状态} \\ [0.3, 0.5], \text{ 中度疲劳状态} \\ > 0.5, \text{ 重度疲劳状态} \end{cases} \quad (27)$$

(1) 在眼部疲劳预警上，本文经实验将驾驶员眼部疲劳阈值设置为 0.4，时间为 30 s，即在单位时间内检测驾驶员的眨眼频率 f_{eye} 和闭眼时长，若 $f_{eye} > 0.4$ ，或闭眼时长大于 3 s 则对驾驶员眼部进行疲劳预警。

(2) 在嘴部疲劳预警上，根据实时计算的 MAR 值输入到 SVM 模型，根据模型输出判定驾驶员是否打哈欠从而对其进行嘴部疲劳预警。

(3) 在头部姿态的疲劳预警上，将实时的欧拉角输入到模型，根据 SVM 模型输出判定驾驶员是否进行低头或者根据欧拉角判断头部姿态是否超过正常范围的 80% 从而进行头部疲劳预警，即 $f_{head} > 0$ 或者 pitch 和 roll 变化大于 25° 和 15.4° 时给予驾驶员疲劳预警。其中若检测到驾驶员眼部、嘴部和头部姿态特征处于疲劳状态时都会对驾驶员分别进行一次预警。

最后使用自建疲劳驾驶数据集进行疲劳检测模型的验证，其中，1、3、4、5 号为模拟疲劳视频，2 号为正常驾驶视频。疲劳辨识模型验证结果如表 4 所示。

表 4 疲劳驾驶检测结果
Table 4 Test results of fatigue driving

视频编号	眼部			嘴部				头部		预警次数
	实际 闭眼数	检测 闭眼数	准确率/%	实际 哈欠数	检测 哈欠数	实际说话 时间/s	检测说话 时间/s	实际 低头数	检测 低头数	
1	90	89	98.90	3	3	12	9	5	5	8
2	157	155	98.70	0	0	20	18	1	0	0
3	131	130	99.20	2	2	17	13	4	5	7
4	87	85	97.70	5	6	6	8	8	6	13
5	106	104	98.10	4	4	18	17	13	15	21

从表4中可以看出, 在眼部检测上SVM模型和神经网络模型相补充进一步提高了疲劳检测模型的精度, 对不同驾驶员眨眼检测的准确率都在97.7%以上。多特征疲劳辨识模型对模拟疲劳驾驶的1、3、4、5号视频都能进行一定次数的预警, 对正常非疲劳驾驶的2号视频没有出现误预警, 证明了疲劳辨识模型对驾驶员进行预警的有效性。

6 结论

本文针对现有疲劳检测特征单一、鲁棒性低等问题, 提出一种基于深度学习和机器学习相结合来实现疲劳检测的方法: ①利用HOG特征和ERT算法进行人脸检测和关键点定位。②根据关键点进行不同坐标系的转换求出旋转矩阵, 得出驾驶员头部姿态的欧拉角。③基于深度学习算法, 建立深度残差神经网络, 对驾驶员眼部特征进行特征提取并判断眨眼。④结合机器学习算法, 利用眼部、嘴部纵横比特征和头部姿态欧拉角建立针对不同驾驶员SVM模型。通过实验对比分析, 本文方法由于综合了眼部、嘴部和头部3方面、多特征的结合, 相较于其他的疲劳检测方法, 具有成本低、无接触、精度高、易普及和鲁棒性高等优点, 提高了实际应用时因某一特征检测受阻时的普适性。此外, 该方法以建立针对不同驾驶员进行训练的SVM疲劳检测模型, 进一步提高了在实际情况下针对不同驾驶员具体情况的适用性和鲁棒性, 对交通安全的改善具有现实意义。

参考文献:

- [1] Tuncer T, Dogan S, Ertam F, et al. A Dynamic Center and Multi Threshold Point Based Stable Feature Extraction Network for Driver Fatigue Detection Utilizing EEG Signals[J]. *Cognitive Neurodynamics*(S1871-4080), 2021, 15(7): 223-237.
- [2] 王博石, 吴修诚, 胡馨艺, 等. 基于单通道脑电信号的疲劳检测系统[J]. *计算机科学*, 2020, 47(5): 225-229.
Wang Boshi, Wu Xiucheng, Hu Xinyi, et al. Fatigue Detection System Based on Single Channel EEG Signal [J]. *Computer Science*, 2020, 47(5): 225-229.
- [3] Kyecheon, LEE, SUNG-AE, et al. Detecting Driver Fatigue by Steering Wheel Grip Force[J]. *International Journal of Contents*(S1738-6764), 2016, 12(1): 44-48.
- [4] Wang M, Guo L, Chen W. Blink Detection Using Adaboost and Contour Circle for Fatigue Recognition[J]. *Computers&Electrical Engineering*(S0045-7906), 2017, 58: 502-512.
- [5] Mandal B, Li L, Wang G S, et al. Towards Detection of Bus Driver Fatigue Based on Robust Visual Analysis of Eye State[J]. *IEEE Transactions on Intelligent Transportation Systems*(S1524-9050), 2017, 18(3): 1-13.
- [6] Zhao Xuepeng, Meng Chunming, Feng Mingkui, et al. Fatigue Detection Based on Cascade Convolutional Neural Network[J]. *Journal of Optoelectronics·Laser* (S1005-0086), 2017, 28(5): 497-502.
- [7] Omidyeganeh M, Shirmohammadi S, Abtahi S, et al. Yawning Detection Using Embedded Smart Cameras[J]. *IEEE Transactions on Instrumentation and Measurement* (S0018-9456), 2016, 65(3): 570-582.
- [8] Kazemi V, Sullivan J. One Millisecond Face Alignment with an Ensemble of Regression Trees[C]// *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus. New York: IEEE Press, 2014: 1867-1874.
- [9] Zhu X, Ramanan D. Face Detection, Pose Estimation, and Landmark Localization in the wild[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2012: 2879-2886.
- [10] Le V, Brandt J, Lin Z, et al. Interactive Facial Feature Localization[C]// *European Conference on Computer Vision*. Heidelberg, Berlin: Springer, 2012: 679-692.
- [11] Belhumeur P N, Jacobs D W, Kriegman D J, et al. Localizing Parts of Faces Using a Consensus of Exemplars [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(S0162-8828), 2013, 35(12): 2930-2940.
- [12] Sagonas C, Tzimiropoulos G, Zafeiriou S, et al. 300 Faces in-the-wild Challenge: The First Facial Landmark Localization Challenge [C]// *IEEE International Conference on Computer Vision Workshops*. Sydney: IEEE, 2013: 397-403.
- [13] Sagonas C, Antonakos E, Tzimiropoulos G, et al. 300Faces in-the-Wild Challenge: Database and Results [J]. *Image and vision computing*(S0262-8856), 2016, 47: 3-18.
- [14] Rafael C, Richard E. *Digital Image Processing*[M]. Beijing: PHI, 2017.
- [15] Zhang Z. A Flexible New Technique for Camera Calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(S0162-8828), 2000, 22(11): 1330-1334.

- [16] Huber P, Hu G, Tena R, et al. A Multiresolution 3D Morphable Face Model and Fitting Framework[C]//11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2016: 79-86.
- [17] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, NV: IEEE, 2016: 770-778.
- [18] Yosinski J, Clune J, Bengio Y, et al. How Transferable are Features in Deep Neural Networks?[C]// International Conference on Neural Information Processing Systems. Quebec, Canada: MIT Press, 2014: 3320-3328.
- [19] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition [C]// International Conference on Learning Representations. 2015.
- [20] Ferrario V F, Sforza C, Serrao G, et al. Active Range of Motion of the Head and Cervical Spine: A Three-Dimensional Investigation in Healthy Young Adults[J]. Journal of Orthopaedic Research (S1554-527X), 2002, 20 (1): 122-129.
- [21] Song F, Tan X, Chen S, et al. A Literature Survey on Robust and Efficient Eye Localization in Real-Life Scenarios [J]. Pattern Recognition(S0031-3203), 2013, 46 (12): 3157-3173.
- [22] Abtahi S, Omidyeganeh M, Shirmohammadi S, et al. YawDD: A Yawning Detection Dataset[C]// ACM Multimedia Systems (MMSys). Singapore: ACM, 2014: 24-28.
- [23] 汪磊, 孙瑞山. 基于面部特征识别的管制员疲劳监测方法研究[J]. 中国安全科学学报, 2012, 22(7): 66-71.
Wang Lei, Sun Ruishan. Study on Feature Recognition-based Fatigue Monitoring Method for Air Traffic Controller[J]. China Safety Science Journal, 2012, 22(7): 66-71.