❖特约专栏❖

基于多形态红外特征与深度学习的实时驾驶员疲劳检测

耿 磊[1,2],梁晓昱[1,2],肖志涛[1,2]*,李月龙[1,3]

(1. 天津市光电检测技术与系统重点实验室,天津 300387;2. 天津工业大学 电子与信息工程学院,天津 300387;3. 天津工业大学 计算机科学与软件学院,天津 300387)

摘 要:疲劳驾驶是导致车祸的重要诱因,严重危害道路交通安全,而车辆行驶过程中的光照条件变化、驾驶员姿态调整和眼镜遮挡等因素将对疲劳检测任务产生不利的影响。针对以上问题,提出了基于深度学习的驾驶员疲劳检测算法。首先,使用 850 nm 红外光源补光,在复杂光照和遮挡形态下采集驾驶员的面部图像;其次,利用红外图像中的多种特征,通过级联 CNN 确定人脸边框和特征点位置,提取眼睛区域并识别眼睛的睁闭状态;最后,将眼睛状态识别结果和连续图像中的特征点坐标差值输出至 LSTM 网络,检测驾驶员疲劳状态。实验结果表明:该疲劳检测算法的准确率可达 94.48%,平均检测时间为 65.64 ms。

关键词:疲劳检测; 红外图像; 深度学习; CNN; LSTM

中图分类号:TP391 **文献标志码**:A **DOI**:10.3788/IRLA201847.0203009

Real-time driver fatigue detection based on morphology infrared features and deep learning

Geng Lei[1,2], Liang Xiaoyu[1,2], Xiao Zhitao[1,2]*, Li Yuelong[1,3]

(1. Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tianjin 300387, China;
2. School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin 300387, China;
3. School of Computer Science and Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China)

Abstract: Fatigue driving is the main cause or reason for traffic accidents, which has a huge influence on social safety. Considering the fact that light change and glasses could significantly increase the difficulty to monitor human eyes, fatigue detection was still an unsolved problem. A new driver fatigue method based on morphology infrared features and deep learning were proposed. Using 850 nm infrared light source, the facial image was obtained. Human faces and landmarks which indicated the area of eyes were located by Convolution Neural Network (CNN) with morphology features in infrared image. In the next step, a filter module which measured head displacement was added, aiming at reducing the impact of posture change. In the following, the collected facial states were transformed into sequential data. Finally, the sequential data was passed to the Long Short Term Memory (LSTM) network to detect fatigue state by analyzing the sequential correlations. Experimental results show that the accuracy of the fatigue detection algorithm can reach 94.48% with an average detection time of 65.64 ms.

Key words: fatigue detection; infrared image; deep learning; CNN; LSTM

收稿日期:2017-08-10; **修订日期**:2017-10-28

基金项目:国家自然科学基金(61601325,61771340)

作者简介:耿磊(1982-),男,副教授,硕士生导师,主要从事测试测量技术及仪器方面的研究。Email:genglei@tjpu.edu.cn

通讯作者:肖志涛(1971-),男,教授,博士生导师,主要从事图像处理与模式识别方面的研究。Email:xiaozhitao@tjpu.edu.cn

0 引 言

2005~2015 年,我国境内可统计的车祸数量超过 300 万起,导致的直接经济损失达 127 亿人民币。与此同时, 接近 30%的驾驶员有过疲劳驾驶行为,因此, 驾驶员疲劳检测方法研究有着重要的社会价值和现实意义。

驾驶员疲劳检测的主要方法有三类: 通过驾驶员的生理参数[1-3]检测疲劳状态、通过驾驶行为[4-6]检测疲劳状态和通过视觉特征检测疲劳状态[7-8]。其中,通过生理参数[1-3]的驾驶员疲劳检测方案的检测精度较高, 但相关传感器与驾驶员的接触部位和接触方式有着较多限制, 将对驾驶员驾驶活动产生直接影响;通过驾驶行为[4-6]的驾驶员疲劳检测方案其系统集成度高、检测便捷,但不易直接反映疲劳状态,且驾驶习惯差异、车辆型号和道路状况等因素将影响检测准确率;通过视觉特征的驾驶员疲劳检测方案有着直接检测疲劳特征、对用户非侵入等优点,作为疲劳检测重要依据,眼睛状态可以反映驾驶员的疲劳状态, 该类别方法以眼睛状态识别结果为基础,利用 PERCLOS、眨眼频率及眨眼时间等参数[7]判断疲劳状态,主要的眼睛定位及识别方法有模板匹配[8]、Adaboost 分类器[9]以及颜色空间聚类[10]等,但这些方法对姿态和光照变化下的疲劳状态检测准确率较低,并且依赖人工选取的疲劳参数阈值判决疲劳状态。

高质量的人脸图像是基于视觉特征检测疲劳状态方法进行准确检测的必要条件,但是在驾驶过程中,因为车辆的位置、光照条件和人脸姿态都会发生变化, 还有可能伴有不同类型的眼镜遮挡,成像环境复杂,增加了疲劳检测任务的难度。针对恶劣天气和夜间光照不足等情况,需要额外的红外光源照射协助成像,同时,在墨镜遮挡下,红外光补偿后可采集到驾驶员的眼睛区域图像,因此围绕红外光源搭建图像采集设备。滤光片作为滤光和选择谱线的器件[11]广泛应用于各个测试系统和分析仪器,衰减其他波段光线在感光元件的接受比例,可降低其他无关波段的光对采集结果的干扰。使用如图 1(a)所示的采集系统: 以焦距为 16 mm 的工业相机为主,配合 850 nm 红外光源和滤光片,采集驾驶员在多种光照和多种眼镜遮挡下的多形态红外图像。深度学习在图像识别[12-14]、目标检测[15-18]、自然语言理解和图像理解[19]等任务领域有着重要的应用,其中长短时记忆网络(Long Short Term Memory, LSTM)可以充分利用数据中的前后联系,提高算法的准确率。

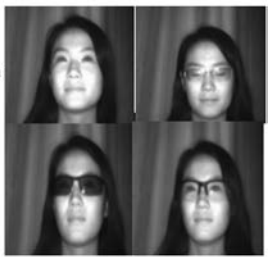
1 面部状态识别

为了在光照不足和多种眼镜遮挡条件下准确地检测驾驶员疲劳状态,文中利用如图 1(b)所示的多状态红外图像和疲劳发生的连续性, 通过多任务级联卷积神经网络[20](Cascaded Multi-Task Convolution Neural Networks, MTCNN) 定位驾驶员人脸图像,使用基于卷积神经网络(Convolution Neural Networks, CNN) 的面部状态识别算法提取多状态红外图像中包含的相应特征, 识别驾驶员面部状态并将结果输出为序列化数据, 在该状态序列数据上提出驾驶员疲劳长短时记忆网络 (Driver Fatigue-Detection with Long Short Term Memory, DF-LSTM), 完成驾驶员疲劳检测,图 2 为算法结构示意图。

其中, 人脸检测和特征点定位是完成疲劳检测的重要步骤,通过人脸检测和特征点定位,完成眼睛“睁闭”状态识别和头部俯仰位移差值计算,将得到的眼睛

(a) 红外系统

(a) Infrared system



(b) 多形态红外图像

(b) Morphology infrared image

图 1 红外采集系统及结果

Fig.1 Infrared acquisition system and result

“睁闭”状态和头部俯仰位移差值整合为序列化面部状态集合 S 来描述驾驶员面部状态。

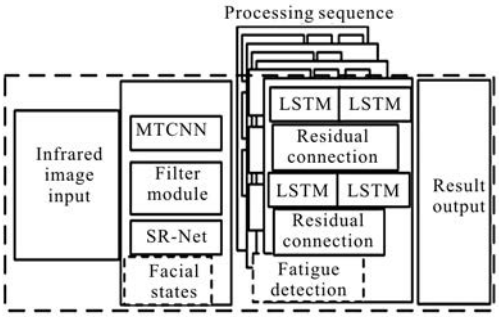


图 2 算法结构示意图

Fig.2 Algorithm structure diagram

为了获得该面部状态序列，通过 MTCNN 定位人脸边框和特征点位置。MTCNN 针对人脸检测这一专门的任务进行设计，将人脸检测和特征点定位两个问题统一到多任务框架下，并利用三个子网络和非极大值抑制在训练过程中完成难样本挖掘，提升了检测准确率和检测算法速度，图 3 为人脸及特征点检测结果。

图 3 人脸及特征点检测结果

Fig.3 Result of face and landmarks detection

眼睛中心位置即 MTCCN 中双眼瞳孔位置的特征点坐标，眼睛感兴趣区域(Region of Interest, ROI)定位可以依据得到的特征点坐标结合眼睛分布的对称性和实验测试结果确定位置约束条件来求解：(1) 根据“三庭五眼”[21-24]的面部分布规律，两瞳孔距离约为单只眼睛宽度的 2 倍，如图 4(a)所示；(2) 眼睛宽高比的选取为 2:1；(3) 待提取的眼睛区域顶点与瞳孔位置之间横纵坐标的距离差分别为一半的眼睛宽高值如图 4(b)所示。眼睛区域求解规则如公式(1)所示：

$$\begin{cases} d=x_B-x_A \\ h_e=d/2 \\ d_e=h_e/2 \\ x_E=x_A-h_e/2 \\ y_E=y_A-d_e/2 \\ x_F=x_B-h_e/2 \\ y_F=y_B-d_e/2 \end{cases} \tag{1}$$

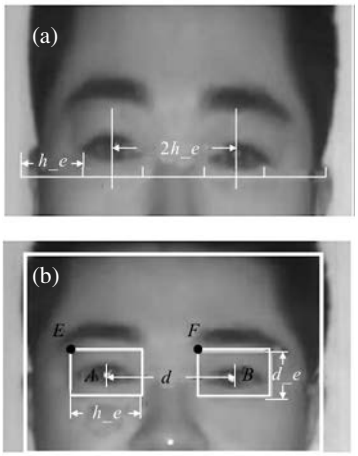


图 4 眼睛感兴趣区域定位

Fig.4 ROI locating of eyes

公式 (1) 中，眼部特征点 A、B 的坐标分别为 MTCNN 输出的特征点坐标(x_A, y_A)、(x_A, y_B)，确定 x_B 与 x_A 的差为两点之间的距离 d；根据距离 d 和位置约束条件求解眼睛区域宽度 h_e 和高度 d_e，相应地由 h_e 和 d_e 求解两只眼睛区域左上方顶点 E、F，坐标分别为(x_E, y_E)、(x_F, y_F)。眼睛区域定位如图 4(b)所示。

将提取的眼睛图像归一化尺寸并送入到 SR-Net[25]中进行眼睛状态识别，网络结构如图 5 所示，眼睛的“睁闭”状态输出为 1/-1 的序列 E_S，t 时刻右眼和左眼的识别结果表示为 e_{rt} 和 e_{lt}。

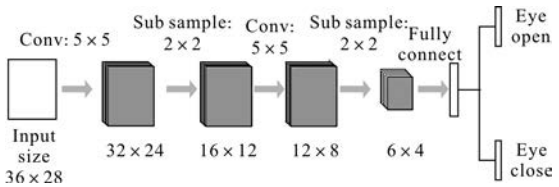


图 5 SR-Net 结构[25]

Fig.5 Construction of SR-Net[25]

在疲劳发生期间，在俯仰位移上，驾驶员的头部

出现缓慢移动，某些被测个体上出现了明显的点头动作。利用特征点坐标计算图像序列中驾驶员头部俯仰的位移差值，图 6 为头部俯仰差值的计算过程。公式(2)为头部俯仰差值计算公式，式中人脸特征点在铅直方向上的两眼、中心点(鼻尖)以及嘴巴坐标分别表示为 y_{leye}、y_{reye}、y_{center}、y_{lmouth}、y_{rmouth}，则顶部坐标为两眼坐标的均值 y_t，底部坐标为嘴巴两点坐标的均值 y_b，顶部至中心点纵轴距离为二者之差的绝对值 D_{ne}，中心点到底部纵轴距离可表示为 D_{nm}，图像的头部俯仰表示为 D_h，上一帧图像的头部俯仰表示为 D'_h，Δd_t 表示 t 时刻头部俯仰差值。

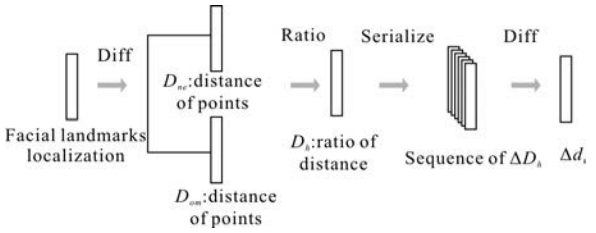


图 6 头部俯仰差值计算

Fig.6 Difference calculation about head elevation

最终将多个时刻的差值结果序列化后表示为 ΔD_S。

$$\begin{cases} y_t = \mathrm{mean}(y_{leye} - y_{reye}) \\ y_b = \mathrm{mean}(y_{lmouth} - y_{rmouth}) \\ D_{ne} = \left| \left| y_{center} - y_t \right| \right| \\ D_{nm} = \left| \left| y_{center} - y_b \right| \right| \\ D_h = D_{ne} / D_{nm} \\ \Delta d_t = D_h - D_h' \end{cases} \tag{2}$$

在驾车期间，驾驶员头部会发生水平转动或水平平面上的快速位移，这些姿态变化不利于面部状态的识别。针对文中研究的疲劳检测任务，剧烈头部动作主要有水平转动或水平平面上的快速位移，滤除含有剧烈动作的数据，只处理稳定位移状态下的待测图像。根据特征点坐标结果，头部水平转动滤除可表示为：

$$\begin{cases} d = \left| \left| x_{lt} - x_{rt} \right| \right| \\ d_l = \left| \left| x_{center} - x_{lt} \right| \right| \\ d_r = \left| \left| x_{center} - x_{rt} \right| \right| \\ r_d = d_l - d_r \end{cases} \tag{3}$$

$$\begin{cases} F = 1, r_d \in (T_{turn1} - T_{turn2}) \\ F_{turn} = 0, \text{others} \end{cases} \tag{4}$$

公式(3)中，人脸特征点在水平方向上的两眼、鼻尖坐标分别表示为 x_{lt}、x_{rt}、x_{center}，顶部水平距离为两眼横坐标之差的绝对值 d，顶部左边界至中心点水平距离表示成 d_l，中心点到顶部右边界距离可表示为 d_r，驾驶员的头部水平旋转表示为 r_d。

在公式(4)中，r_d 为头部水平旋转，T_{turn1} 和 T_{turn2} 为初始化时设定的转动阈值，F_{turn} 则表示头部水平转动的判定，当 F_{turn} 为 1 时则驾驶员头部无剧烈的水平转动；当 F_{turn} 为 0 时则判定为发生剧烈水平转动，将不对这一时刻的数据进行序列化和疲劳检测。

小范围的水平平移对疲劳检测的影响有限，为了规避驾驶员超出图像有效检测范围等情况引发的检测错误，将相邻两帧人脸中心点位置移动在相应方向上超过一倍的人脸长度或人脸宽度的水平移动判定为快速位移，滤除过程可表示为：

$$\begin{cases} \Delta Y = \left| \left| y_{center} - y'_{center} \right| \right| \\ \Delta X = \left| \left| x_{center} - x'_{center} \right| \right| \\ y_{mt} = \left| \left| y_{face\ top} - y_{face\ Bt} \right| \right| \\ x_{mt} = \left| \left| x_{face\ L} - x_{face\ R} \right| \right| \end{cases} \tag{5}$$

$$\begin{cases} F_{drift} = 1, (\Delta X < x_{mt}, \Delta Y < y_{mt}) \\ F_{drift} = 0, \text{others} \end{cases} \tag{6}$$

公式(5)中，前后两帧中人脸中心点，即鼻尖坐标分别表示为($x'_{acenter}$, $y'_{acenter}$)，($x_{acenter}$, $y_{acenter}$)表示当前帧的人脸鼻尖坐标，则前后两帧头部水平移动是 ΔX，铅直移动是 ΔY；上式中($x_{face\ L}$, $x_{face\ R}$)表示当前帧的人脸边框的左侧边界和右侧边界，($y_{face\ top}$, $y_{face\ Bt}$)表示当前帧的人脸边框上方边界和下方边界，x_{mt} 表示水平移动阈值，y_{mt} 表示铅直方向移动阈值。在公式(6)中，F_{drift} 为快速位移判定值，当水平和铅直方向上的位移都小于初始化的阈值 x_{mt} 和 y_{mt} 时，F_{drift} 的值为 1，继续处理；否则 F_{drift} 的值为 0，判定发生快速位移，滤除这一时刻的数据。

当 F_{turn} 和 F_{drift} 都判定为 1 时，将继续处理该时刻的面部数据。以时间为变量将眼睛状态和头部位移状态这两个疲劳特征进行统一，根据头部俯仰差值 Δd_t 和眼睛状态识别结果 e_{rt}、e_{lt} 建立量化的驾驶员面部状态集合。根据对应的时序顺序，将眼睛状态序列

E_S 和头部纵轴俯仰差值 ΔD_S 合并，则驾驶员面部状态序列可表示为如公式(7)的集合 S：

$$S=\{(e_{r0},e_{l0},\Delta d_0),\cdots,(e_{rt},e_{lt},\Delta d_t)\cdots\},(e_{rt}\in E_S,e_{lt}\in E_S,\Delta d_t\in\Delta D_S,\ t=0,1,2,\cdots) \tag{7}$$

构成完整的面部识别模块 (Facial Status of Drivers cascaded Convolution Neural Networks, FSD-CNN)的输出。

2 驾驶员疲劳检测

LSTM 网络[26]通过设计输出门、输入门和遗忘门来构建的循环神经网络能有效提高长序列处理能力，图 7 为 LSTM 结构示意图。发生疲劳是一个过程性事件，为了利用疲劳发生的前后联系，提出了包含 LSTM 单元、双向[27]和残差结构[28]的 DF-LSTM 网络，使用该网络处理驾驶员面部状态数据集合 S。通过关联前后时序两个方向的内容，获得驾驶员面部状态在清醒和疲劳状态下的特征，完成疲劳检测。

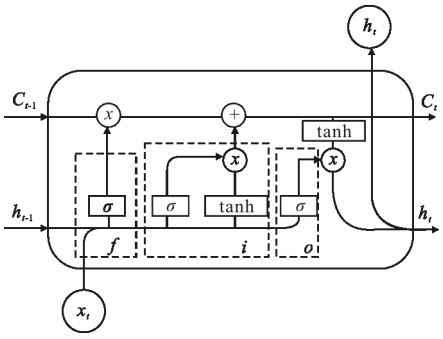


图 7 LSTM 结构示意图

Fig.7 LSTM structure

构成网络的 LSTM 单元中包含遗忘门、输入门和输出门结构，通过比较当前数据流与内部状态来控制更新的序列长度，进而保障梯度的合理变化，可较为有效地规避梯度消失和梯度爆炸。图 7 中 σ 表示 sigmoid 函数，用来完成所连接量中可以继续传递部分的控制；x_t 为该时刻的输入，h_t 为隐层输出，C_t 为这一时刻的控制流数据。

在该时刻，各参数间的关系可表示为：

$$f=\sigma(W_f[h_{t-1},x_t]+b_f) \tag{8}$$

式中：W 为对应的权重矩阵；b 表示对应的偏置量。公式(8)为遗忘门的处理过程，通过 sigmoid 函数对上一时刻隐层状态 h_{t-1} 和本时刻输入 x_t 进行部分舍弃，在不同时刻的 LSTM 中 f 的结果会发生相应的变化。

$$i=(\sigma(W_i[h_{t-1},x_t]+b_i)*(\tanh(W_c[h_{t-1},x_t]+b_c))) \tag{9}$$

式中：i 表示该时刻 LSTM 中与输入门相关的数学表达，sigmoid 函数所在的这一部分是输入门的核心。输入门主要有两个操作，利用 sigmoid 函数特性更新数据，并由 tanh 函数所在的模块计算出新的候选向量，二者逐点相乘之后得到输出值 i。

$$C_t=f*C_{t-1}+i \tag{10}$$

通过公式(10)更新控制流信息 C_t。将上一时刻的控制流信息 C_{t-1} 与遗忘门处理过的输入信息 f 相乘，叠加上公式(9)中输入门得到新数据，通过包含上一隐层和本时刻输入，完成该时刻的控制流信息更新。

$$o_t=\sigma(W_o[h_{t-1},x_t]+b_o) \tag{11}$$

公式(11)和公式(12)确定 LSTM 单元的输出值 h_t。公式(11)中的 O_t 为输出门结构的更新结果。与遗忘门和输入门类似，同样利用一个包含 sigmoid 函数的结构来对上一时刻隐层状态 h_{t-1} 和本时刻输入 x_t 进行部分舍弃，将结果选择性的输出到单元内部。

$$h_t=o_t*\tanh(C_t) \tag{12}$$

在公式(12)中，C_t 为更新过的控制流信息，并通过 tanh 进行作用，使其落在(-1,1)的区间内，O_t 为公式(10)得到的输出门结果，逐点相乘运算后得到的 h_t 则为本时刻 LSTM 结构输出的最终节点值，分别传输到下一时刻 LSTM 单元和本时刻更深的网络结构中。

图 8 是驾驶员疲劳长短时记忆网络。对某一确定时间节点 t 而言，x_t 为整个网络的输入，网络可以看成由多个 h_i 隐层组成的残差网络，而在 h_i 内部则包含一对连接的 LSTM 结构。中间状态输出和最终状态输出前采取了批归一化处理(Batch Normalization, BN)。双向功能由内部成对存在的 LSTM 单元承担。整个过程如公式(13)~(16)所示：

$$h_{lf}=g(U_{hl}x_t+W_{hl}h'_{lf}+b_{hl}) \tag{13}$$

$$h_{rb}=g(U_{hb}x_t+W_{hb}h'_{rb}+b_{hr}) \tag{14}$$

$$\begin{cases} h_i=g(V_1\,\mathrm{concat}(h_{lf},b_{rb})+b_i) & i=1 \\ h_i=g(V_i h_{i-1}+b_i)+h_{i-2} & i=2,\ \cdots,L-1 \end{cases} \tag{15}$$

$$y_t=g(V_y h_i+b_y)+h_{i-1} \quad i=L \tag{16}$$

式中：x_t 为当前时刻输入；h_i 为重复的隐层结构；y_t 为

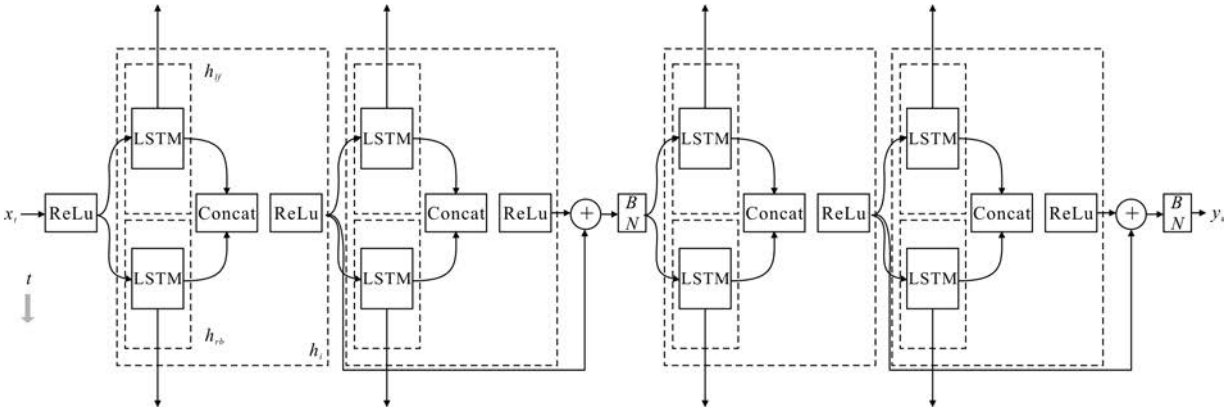


图 8 驾驶员疲劳长短时记忆网络

Fig.8 Driver fatigue long short term memory net

输出；h'_{lf} 表示当前时刻前向连接 LSTM 单元；h_{lf} 为相邻时刻的前向连接 LSTM 单元；h'_{rb} 表示向后连接的 LSTM 单元；h_{rb} 为上一相邻时刻的后向连接 LSTM 单元；UVW 为对应的权重矩阵；g 为 ReLu 激活函数。

3 实验结果与分析

实验平台处理器为 x86 架构的 Intel Core i7-6 700 K（主频为 4.0 GHz），配合 Pascal 架构的 GTX1080Ti（包含 3584 CUDA cores，Compute Capability6.1），搭载 Ubuntu16.04 操作系统，内存为 16 GB DDR4。

在实验室模拟驾驶和实际行驶场景下，采集 35 名被测人员不同光照条件下的疲劳视频，如图 9 所示。疲劳驾驶是违法行为，因此，采集被测对象在行车过程中副驾驶位的疲劳数据，建立多形态红外视频的数据集合 TJPU-FDIR，表 1 为数据样该统计，其中 myopia mirror 为近视眼睛、sunglasses 为墨镜、polarizer 为偏光镜、no glasses 为没有眼镜遮挡。算法的准确性由检测准确率 Accuracy 进行评估，如公式(17)所示：

图 9 不同光照条件下的数据样本

Fig.9 Data samples under different lighting conditions

$$\text{Accuracy}=\frac{G}{D_n}\times 100\% \tag{17}$$

式中：D_n 为测试样本中待检测帧数之和；G 表示测试样本中正确检测出疲劳状态的帧数；通过算法运行时间统计 T(Time)评估方法的实时性。

表 1 驾驶员疲劳检测数据样本

Tab.1 Data samples of driver fatigue detection

No.	Model	Speed/fps	Size/in	Number/frame
1	Myopia mirror	27	640×480	18 200
2	Sunglasses	27	640×480	17 540
3	Polarizer	27	640×480	14 590
4	No glasses	27	640×480	16 780

结合 CNN 检测结果和位置约束条件完成眼睛区域(ROI location)定位，结果见表 2，根据公式(17)计算在多形态红外图像中算法的平均定位准确率可达 98.16%，在数据集中不同光照和遮挡情况下有较高的定位能力。

但是由于不同眼镜光学性能的差异、实际行驶过程中的车辆震颤以及不同时刻驾驶员相对于红外光源位移产生的反光光斑，眼睛区域定位的准确率

表 2 眼睛 ROI 定位结果

Tab.2 Result of eyes ROI location

Model	Data set	D_n/frame	Accuracy
Myopia mirror	TJPU-FDIR	6 240	98.43%
Sunglasses	TJPU-FDIR	6 536	97.11%
Polarizer	TJPU-FDIR	5 808	97.38%
No glasses	TJPU-FDIR	6 968	99.71%

受到了一定影响。

在状态识别网络中提取多形态红外图像中的特征,识别眼睛的睁闭状态,计算头部位移,完成面部状态识别,将序列化后的面部状态数据集合 S 通过人工标注分为“清醒”和“疲劳”两类,每类下分别由三个分量描述在清醒和疲劳状态下的驾驶员左眼、右眼状态以及头部位移。为了减少过拟合现象,在网络训练过程中采取相应的 dropout 策略,随机屏蔽 50%的神经元连接;对 LSTM 网络,检测滑动步长和待检测序列的长度选择将影响检测准确率,而被测疲劳的对象一次完整眨眼动作的平均耗时在 250~300 ms,采集设备的速度为 27 fps,因此设定每次滑动的步长 4,可确保两个相邻检测窗口不错过一个完整的眨眼动作。检测当前时刻之前约 5 s 的历史窗口内驾驶员面部状态序列,故每组的训练序列长度选择为 128。训练与测试数据比例为 3:1,迭代次数 400 000,基础学习率为 0.05,每隔 2 000 次迭代对模型进行测试,观察测试准确率。

图 10 反映了 DF-LSTM 网络在训练集和测试集上的准确率与损失值的变化曲线。

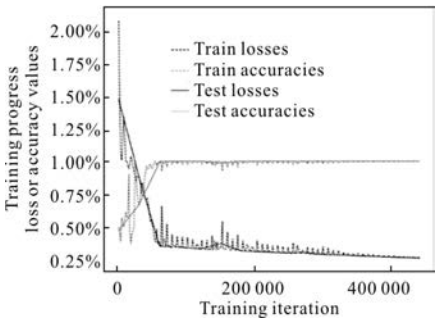


图 10 DF-LSTM 的准确率与损失值

Fig.10 Accuracy and loss of DF-LSTM

根据采集人员对被测人处于疲劳状态的时间记录作为真值。表 3 是各形态下疲劳检测算法的检测结果,文中算法的整体检测准确率较高,可检测出不同条件下的疲劳状态;在检测数据上的模型泛化能力良好,没有出现过拟合现象。但是结合表 2 定位结果可知,眼睛区域定位错误对降低了疲劳检测结果的准确性。

表 3 疲劳检测结果

Tab.3 Result of fatigue detection

Model	Data set	D_n/frame	Accuracy
Myopia mirror	TJPU-FDIR	6 240	95.57%
Sunglasses	TJPU-FDIR	6 536	92.16%
Polarizer	TJPU-FDIR	5 808	93.36%
No glasses	TJPU-FDIR	6 968	96.81%

表 4 为运算速度的统计数据,测试 6 个片段(clip)总计 28 782 帧的数据,测试结果显示,基于深度学习的疲劳检测算法的平均检测时间为 65.64 ms。

表 4 算法速度测试结果

Tab.4 Result of algorithm speed test

Clip number	D_n/frame	FSD-CNN /frame·ms^{-1}	DF-LSTM /frame·ms^{-1}	Total /frame·ms^{-1}
1	5 967	41.52	20.12	61.64
2	4 104	48.31	20.12	68.43
3	5 994	46.93	20.01	67.04
4	3 807	45.61	20.11	65.72
5	4 212	42.29	20.13	62.42
6	4 698	48.40	20.16	68.56
Mean	*	45.51	20.108	65.64

将疲劳检测结果的平均检测时间和平均检测准确率同 SR-Net+PERCLOS[25] 方法进行对比,结果如表 5 所示。通过利用疲劳过程中面部状态序列的前后关联性,使用 DF-LSTM 网络改进后的检测算法平均检测准确率提高了 1.38%,但平均检测时间增加了 7.03 ms。

表 5 疲劳检测结果对比

Tab.5 Comparison of fatigue detection result

Method	Speed/frame·ms^{-1}	Mean accuracy
CNN+PERCLOS[25]	58.61	93.10%
CNN+DF-LSTM	65.64	94.48%

4 结 论

文中通过级联 CNN 的方式定位眼睛区域并识别眼睛状态,将序列化的驾驶员面部状态数据送入 DF-LSTM 网络完成疲劳检测。实验结果表明,该疲劳检测方法能够有效利用多形态红外特征,疲劳检测准确率可达 94.48%,平均检测时间为 65.64 ms。

但算法对于车辆震颤、眼镜强烈反光状态下的检测鲁棒性不够好，之后的研究应该针对这些问题提出解决方案，使得疲劳检测算法的准确率和适用性进一步得到提升。

参考文献：

[1] Katsis C D, Ntouvas N E, Bafas C G, et al. Assessment of muscle fatigue during driving using surface EMG [C]// Proceedings of the IASTED International Conference on Biomedical Engineering, 2004: 259-262.

[2] Ohsuga M, Kamakura Y, Roongroj Nopsuwanchai, et al. Classification of blink waveforms toward the assessment of driver's arousal levels-an EOG approach and the correlation with physiological measures[C]//Engineering Psychology and Cognitive Ergonomics, 2007: 787-795.

[3] Takahashi I, Yokoyama K. Development of a feedback stimulation for drowsy driver using heartbeat rhythms[C]// Engineering in Medicine and Biology Society, 2011: 4153-4158.

[4] Jin Xue. Research on the detection method of fatigue driving based on driving behavior[D]. Beijing: Beijing University of Technology, 2015. (in Chinese)

[5] Celenk M, Eren H, Poyraz M. Prediction of driver head movement via Bayesian learning and ARMA modeling[C]// Intelligent Vehicles Symposium, 2009: 542-547.

[6] Friedrichs F, Yang B. Drowsiness monitoring by steering and lane data based features under real driving conditions[C]// European Signal Processing Conference, 2010: 209-213.

[7] Tao H, Zhao Y. Real-time driver fatigue detection based on face alignment[C]//International Conference on Digital Image Processing, 2017: 1042003.

[8] Meng C N, Bai J J, Zhang T N, et al. Eye movement analysis for activity recognition based on one web camera[J]. ***Acta Physica Sinica***, 2013, 62(17): 116-121.

[9] Cheng Ruzhong, Zhao Yong, Dai Yong, et al. An on-board embedded driver fatigue warning system based on adaboost method [J]. ***Acta Scientiarum Naturalium Universitatis Pekinensis***, 2012, 48(5): 719-726. (in Chinese)

[10] Kuang Wenteng, Mao Kuancheng, Hong Jiacai, et al. Fatigue driving detection based on sclera Gaussian model[J]. ***Journal of Image and Graphics***, 2016, 21(11): 1515-1522.

[11] Wang Jizhou, Li Hong, Xiong Yuqing, et al. Design and preparation of one induced transmission filter with anlistatig and anti-infrared character[J]. ***Infrared and Laser Engineering***, 2015, 44(10): 3005-3009. (in Chinese)

[12] Boser B, Denker J S. Backpropagation applied to handwritten zip code recognition [J]. ***Neural Computation***, 1989, 1(4): 541-551.

[13] Farhan M, Yli-Harja O, Niemist A. A novel method for splitting clumps of convex objects incorporating image intensity and using rectangular window-based concavity point-pair search [J]. ***Pattern Recognition***, 2013, 46(3): 741-751.

[14] Alex Krizhevsky, Llya Sutskever, Geoffrey E, et al. ImageNet classification with deep convolutional neural networks[J]. ***Communications of the Acm***, 2012, 60(2): 2012.

[15] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Computer Vision and Pattern Recognition, 2014: 580-587.

[16] Girshick R. Fast R-CNN [C]//IEEE International Conference on Computer Vision, 2015: 1440-1448.

[17] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]// International Conference on Neural Information Processing Systems, 2015: 91-99.

[18] Redmon J, Divvala S, Girshick R, et al. You only look once: unified real-time object detection [C]//Computer Vision and Pattern Recognition, 2015, arxiv: 1506, 02640.

[19] Donahue J, Hendricks L A, Rohrbach M, et al. Long-term recurrent convolutional networks for visual recognition and description [J]. ***IEEE Transactions on Pattern Analysis & Machine Intelligence***, 2014, 39(4): 677-691.

[20] Zhang K, Zhang Z, Li Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. ***IEEE Signal Processing Letters***, 2016, 23 (10): 1499-1503.

[21] Li Rui. Facial expression recognition method based on feature fusion [D]. Hefei: Hefei University of Technology, 2013. (in Chinese)

[22] Hu Feng. Multi-source information fusion application to driving fatigue detection [D]. Hefei: Anhui University of Science and Technology, 2014. (in Chinese)

[23] Chen Mingchu. Research on driver fatigue detection based on eye state [D]. Chongqing: Chongqing University, 2012. (in Chinese)

[24] Ma Ying. Research on driver fatigue alarm system based on facial features [D]. Hefei: Anhui University of Technology, 2016. (in Chinese)

[25] Zhang F, Su J, Geng L, et al. Driver fatigue detection based on eye state recognition [C]//International Conference on

Machine Vision and Information Technology, 2017: 105–110.

[26] Sepp Hochreiter, Jürgen Schmidhuber. Long short –term memory[J]. ***Neural Computation***, 1997, 9(8): 1735–1780.

[27] Ma X, Hovy E. End –to –end sequence labeling via Bi –directional LSTM–CNNs–CRF [C]//Meeting of the Association for Computational Linguistics, 2016: 1064–1074.

[28] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing human –level performance on imageNet classification[C]//IEEE International Conference on Computer Vision, 2016: 1026–1034.