



引用格式: 廖明明, 赵波. 基于面部表情和双流网络的驾驶员疲劳检测[J]. 科学技术与工程, 2022, 22(2): 614-619.

Liao Mingming, Zhao Bo. Driver fatigue detection based on facial expression and double-stream network[J]. Science Technology and Engineering, 2022, 22(2): 614-619.

基于面部表情和双流网络的驾驶员疲劳检测

廖明明, 赵波*

(上海工程技术大学机械与汽车工程学院, 上海 201620)

摘 要 提出了一种基于面部表情的驾驶员疲劳检测方法, 该方法结合了传统特征提取和双流卷积神经网络。首先, 对采集的驾驶员图片进行预处理, 使用 Dlib 进行人脸检测以及人脸特征点定位。然后, 根据人脸特征点获取人脸表情感兴趣区域以及嘴部区域, 并分别提取 Gabor 特征和局部二值模式(local binary patterns, LBP) 特征。最后, 利用设计的疲劳表情识别网络对获取的两种传统特征进行信息融合以及疲劳表情识别。结果表明, 提出的方法具有较高准确率, 能够适应驾驶室内不同光照条件的场景, 具有较强的鲁棒性。

关键词 表情识别; 卷积神经网络; 疲劳检测; 特征融合

中图分类号 TP391.41;

文献标志码 A

Driver Fatigue Detection Based on Facial Expression and Double-stream Network

LIAO Ming-ming, ZHAO Bo*

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] A method for driver fatigue detection was proposed based on facial expressions, which combines traditional feature extraction and double-stream convolutional neural networks. Firstly, the collected driver pictures were preprocessed, then face and facial feature points were detected through Dlib toolbox. Secondly, the facial expression interest area and the mouth area were obtained according to the facial feature points, and Gabor features and local binary patterns (LBP) features were extracted respectively. Finally, the designed fatigue expression recognition network was used to perform information fusion and fatigue expression recognition on the two traditional features obtained above. The results show that the proposed method achieved high accuracy, and can adapt to scenes with different lighting conditions in the cab, and had strong robustness.

[Keywords] expression recognition; convolutional neural network; fatigue detection; feature fusion

疲劳驾驶是威胁道路安全和造成人员伤亡的主要因素之一^[1]。中国交通部的一项调查显示, 每年有 90 多万人死于疲劳驾驶, 占交通事故死亡人数的 6% 以上^[2]。因此, 疲劳检测算法的研究对减少交通事故引起的人员伤亡和财产损失具有重要意义。

疲劳检测的方法可分为三种: 基于生理信息, 基于车辆行为, 基于视觉特征。基于视觉特征的方法大多是通过分析眼部和嘴部的特征或者整体行为来识别驾驶员的疲劳状态。Zhao 等^[3] 出了一种基于卷积神经网络(convolutional neural network, CNN) 的疲劳检测算法。该算法的第一个网络将人眼和非人眼进行分类, 第二个网络检测人脸特征点的位置, 根据获得的特征点位置计算眼睛睁开程

度, 并采用 PERCLOS 准则来判断疲劳状态。You 等^[4] 利用改进的深度神经网络进行人脸检测, 使用 Dlib 工具获取人脸的特征点, 通过眼睛长宽比(eye aspect ratio, EAR) 对眼睛状态进行识别, 进而得到驾驶员的疲劳状态。王霞等^[5] 通过对嘴部的内轮廓进行 Gabor 特征提取, 再通过嘴部的开合频率进行疲劳判别。Xie 等^[6] 利用最先进的卷积神经网络结构作为迁移学习过程中的特征提取器, 利用长短期记忆网络(long short-term memory, LSTM) 学习打哈欠动作时间维度的特征, 该模型的优点是能够将打哈欠动作与说话区分开。Liu 等^[7] 提出了结合面部特征和模糊推理理论的方法。使用多块局部二进制模式(MB-LBP) 和 Adaboost 算法对驾驶员的眼睛和嘴部区域进行定位。并通过模糊推理系统对驾

收稿日期: 2021-04-20; 修订日期: 2021-08-12

第一作者: 廖明明(1996—) 男, 汉族, 安徽安庆人, 硕士研究生。研究方向: 图像处理。E-mail: 318489188@qq.com。

* 通信作者: 赵波(1963—) 女, 汉族, 上海人, 硕士, 副教授。研究方向: 智能汽车。E-mail: 1131689745@qq.com。

投稿网址: www.stae.com.cn

驾驶员的疲劳状态进行评估。Ji 等^[8]使用多任务级联卷积神经网络进行疲劳检测,而且为了进一步分析驾驶员疲劳状态,将眼睛和嘴部的两个特征结合起来,构成疲劳判别模型。戎辉等^[9]提出使用迁移学习的方法将驾驶员分心行为分为7类,进而更加深入地分析驾驶员状态。

如上所述,目前大多数研究^[3-8]主要通过提取人脸的局部特征或者整体行为特征^[9]来进行疲劳识别。然而,这样对面部整体的疲劳表情特征并没有充分地挖掘,会导致部分面部特征丢失,从而难以将面部的整体疲劳状态表现出来。因此,现以驾驶员面部表情为研究对象,提出一种结合传统特征和卷积神经网络的疲劳状态识别网络。该网络结合整体面部特征和局部特征,有效提高疲劳状态识别的准确率。通过对驾驶员面部特征进行融合的疲劳识别算法的研究,旨在推动驾驶员疲劳检测的智能化发展。

1 图像预处理

疲劳检测算法流程如图1所示。算法的第一步是人脸图像的预处理,预处理是人脸表情识别的重要步骤,理想的预处理应该去除纯表情图像中不相关的信息(如光照、背景),得到预处理后的图像具有均匀的大小和归一化的亮度。利用Dlib进行人脸检测并提取表情感兴趣区域。对采集的图片进

行直方图均衡化处理,减小真实驾驶场景中光照条件变化的干扰。

1.1 人脸检测

Dlib 是一个开源工具包,包含机器学习算法和工具,被广泛应用于工业和学术界。Dlib 适用于人脸检测,因为不仅可以描述各种人脸的轮廓和边缘特征,而且由于在局部区域上执行操作,使得主体的运动被忽略。在 Dlib 工具包中,原始图片如图2(a)所示,人脸的姿态由68个关键点表示,如图2(b)所示。

使用 Dlib 进行人脸检测,主要有三个原因:

①Dlib具有检测速度快、使用方便的优点;②不需要测试者的先验人脸检测;③Dlib相对于其他的人脸检测方法能够快速准确地分割出用于表情识别的人脸区域。表情感兴趣区域分割效果如图2(c)所示。

1.2 直方图均衡化

在真实的驾驶环境中,驾驶室内的光照条件时常会发生变化,驾驶员脸部的各个部分也会呈现出不同的亮度,这将不可避免地对识别结果造成很大干扰。直方图均衡化是图像处理中一种简单而有效的算法,可以使不同图像的灰度值分布更加均匀,减小不同光照条件带来的干扰。如图3所示,直方图均衡化优化了更亮的面部部分,人脸的重要特征被更好地呈现。

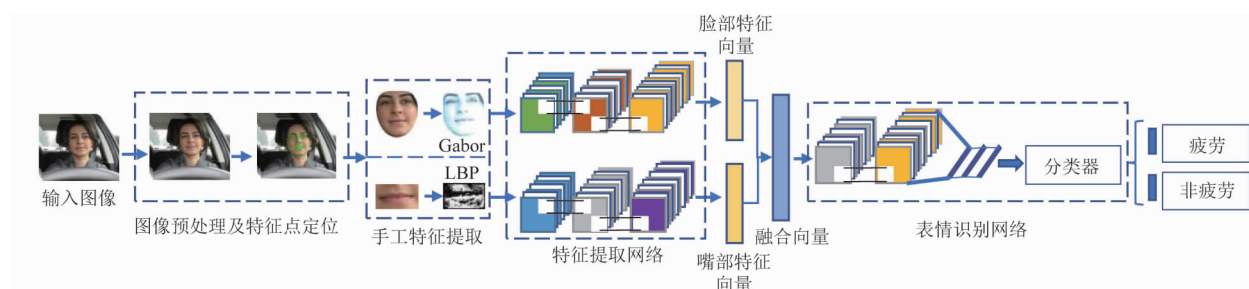


图1 疲劳表情检测流程

Fig. 1 Fatigued expression detection flow chart

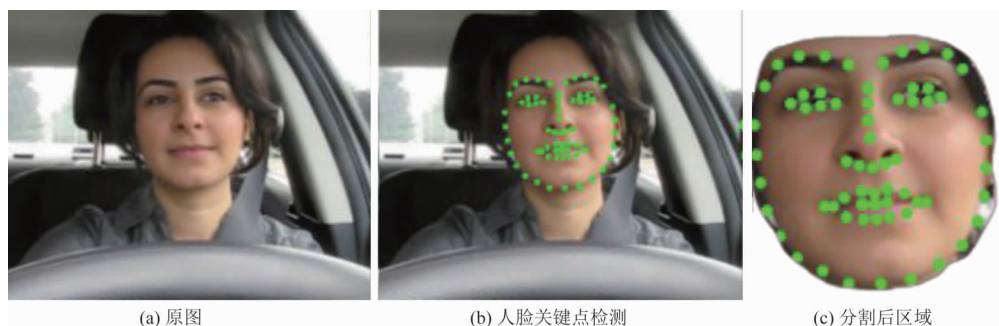


图2 Dlib人脸检测及分割效果

Fig. 2 Results of Dlib face detection and segmentation

投稿网址: www.stae.com.cn



图3 直方图均衡处理

Fig. 3 Histogram equalization processing

1.3 手工特征提取

1.3.1 Gabor 特征

Gabor 滤波器^[10]被广泛用于局部特征提取、纹理分析和边缘检测。使用 Gabor 滤波器对图像进行处理,图像的边缘和纹理变化能得到最高的响应,并能同时对空间和频率信息进行定位。Gabor 滤波器可以捕捉显著的视觉特性,如空间定位、方向选择性和空间频率特性。考虑到这些优点,选择 Gabor 特征来表示人脸表情。Gabor 滤波器的定义为

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} \times [e^{ik_{u,v}z} - e^{-\sigma^2/2}] \quad (1)$$

式(1)中: $e^{ik_{u,v}z}$ 为震荡波函数,其实部和虚部分别是余弦函数和正弦函数; $\sigma = 2\pi$ 为常量; i 为虚数。且 $z = (x, y)$ 给定了空间域中的位置,频域向量 $k_{u,v}$ 定义为

$$k_{u,v} = \frac{k_{\max}}{f} e^{i\varphi_u} \quad (2)$$

式(2)中: $\varphi_u = u\pi/u_{\max}$, $\varphi_u \in [0, \pi]$; f 为 Gabor 表征各种频率; v 为 Gabor 滤波器的比例,该比例主要确定 Gabor 滤波器在频域中的中心; u 为 Gabor 滤波器的方向。

为了降低参数维度,使用了包含 4 个尺度和 6 个方向的 Gabor 滤波器组,如图 4 所示。Gabor 滤波器的参数设为: 4 个尺度 $v = \{0, 1, 2, 3\}$, 6 个方向 $u = \{0, 1, \dots, 5\}$, $\sigma = 2\pi$, $k_{\max} = \pi/2$, 且 $f = \sqrt{2}$ 。

1.3.2 LBP 特征

局部二值模式(LBP)最初是由 Ojala 等^[11]在引入用于纹理的描述,后来应用于人脸分析。LBP 算子通过用中心值对每个像素的 3×3 邻域取阈值来标记图像的像素,并将结果视为二进制数,即相邻像素值等于或高于中心像素值,则用 1 表示,否则用 0 表示。然后生成的直方图可以用作纹理描述,图 5 为原始 LBP 算子的图解。LBP 是一种能够描述灰度图像局部纹理特征的算子,它具有旋转不变性和灰度不变性的特性。一般情况下不直接使用 LBP 作为分类识别的特征向量,而是使用 LBP 特征直方图作为分类识别的特征向量。

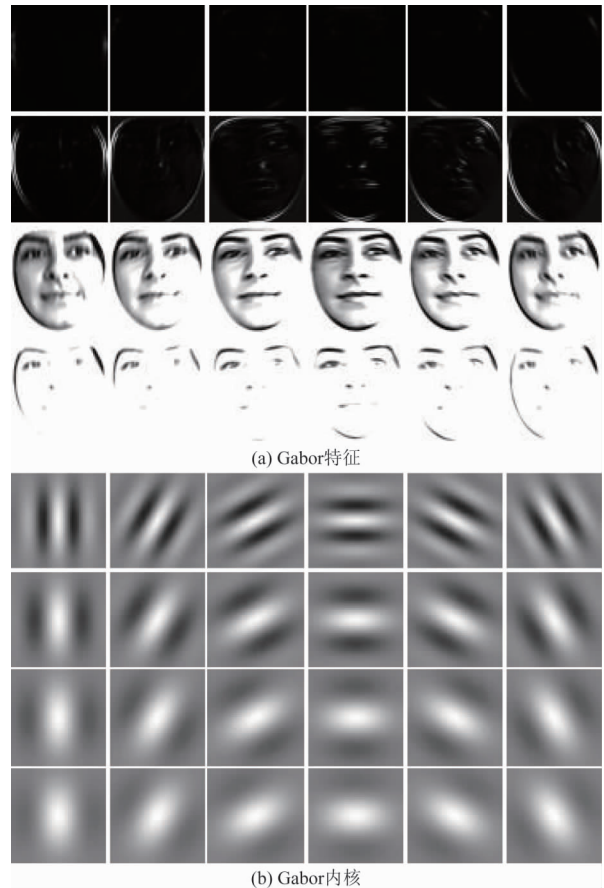


图4 Gabor 滤波器处理结果

Fig. 4 Results of Gabor filters processing

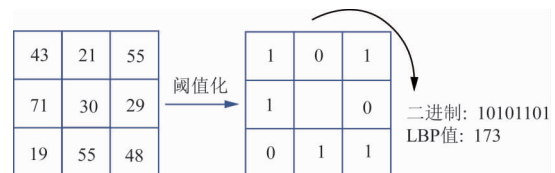


图5 原始 LBP 值定义

Fig. 5 Origin definition of LBP value

2 疲劳表情识别网络

为提高疲劳表情识别的准确率,设计的疲劳识别网络融合了驾驶员面部的全局和局部的特征。全局特征倾向于表现表情的完整性,而局部特征侧重于局部区域的细节描述,可以直接表现面部的活跃表情区域。这两个独立的分支从两个不同的方面表示表情,具有很好的互补性。

提出的疲劳表情识别网络包括特征提取网络 FE-Net 以及表情识别网络 ER-Net,具体结构如图 6(a) 所示。FE-Net 具有双分支结构,第一个分支以表情感兴趣区域的 Gabor 特征作为输入,提取面部全局特征;另一个分支以嘴部 LBP 特征作为输入,提取局部纹理特征。将 FE-Net 提取获得的融合

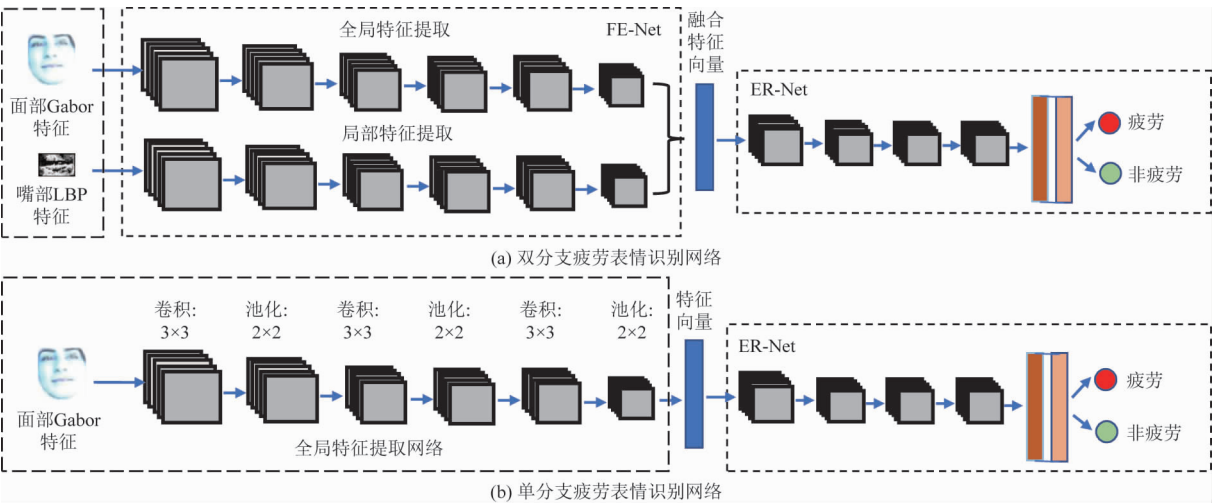


图6 疲劳表情识别网络

Fig. 6 Fatigued expression recognition network

特征向量作为 ER-Net 的输入 ,最终得到驾驶员的疲劳状态。

为验证在网络中加入嘴部特征对表情识别精度的提高 ,将 FE-Net 嘴部特征提取分支去掉 ,构造了一个基于全局面部特征的表情识别网络 ,用于与原网络进行对比 ,结构如图 6(b) 所示。

FE-Net 每个分支网络均由三个卷积层和三个池化层组成 ,且每个网络的通道数依次为 32、64、128 个 ,卷积核大小为 3×3 ,步长为 1×1 ,用零填充。使用的激活函数是训练速度较快的 ReLU 激活函数。池化层使用最大池化 ,大小为 2×2、步长为 2×2 ,这极大减少了模型的参数。

SE-Net 包含两层卷积两层池化 ,通道数分别为 128 和 256 ,卷积核大小 3×3 ,步长为 1×1。使用两个全连接层 ,分别有 1 024 和 512 个单元 ,随机舍弃的概率为 0.5 ,以防止过拟合。最后 ,使用 softmax 分类器进行二分类任务 ,即疲劳和非疲劳两类。损失函数为交叉熵损失函数 ,即

$$\text{Loss} = - \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (3)$$

式(3)中: y_i 为样本 i 的标签 ,正例为 1 ,负类为 0; \hat{y}_i 为样本预测为正的概率。

3 实验

3.1 实验平台与数据集

实验硬件配置为: core i7-9800X CPU (3.80 GHz) ,32 GB 内存 ,RTX2080 显卡。使用的操作系统为 Ubuntu18.04。

由于目前没有公开疲劳表情数据集 ,构建了一个疲劳表情数据集。该数据集来源于自行采集的样本和用于疲劳检测的公开数据集。自行采集数据是实验者在驾驶室不同光照条件下采集的疲劳表情图片。选用的公开数据集是 YawDD^[12] 打哈欠数据集和 CEW(closed eyes in the wild) 闭眼数据集 ,在其中选取合适的疲劳表情图片 ,选择的部分不同光照条件的样本如图 7 所示。最终获得疲劳表情样本 2 610 张以及非疲劳表情样本 2 646 张。为减小数据样本较少对模型训练的影响 ,提取出数据集中驾驶员脸部 Gabor 特征和嘴部 LBP 特征 ,并将特征图片归一化到 216×216 大小用作模型的训练和测试。



图7 不同光照条件下的样本

Fig. 7 Samples under different light conditions

投稿网址: www.stae.com.cn

3.2 疲劳表情识别

3.2.1 评估指标

在分类问题上,通常用精确度(Precision)和召回率(Recall)来评估网络模型的好坏,精确度表示在预测结果为正样本中,正确预测为正样本的概率。召回率表示在原始样本的正样本中,最后被正确预测为正样本的概率。模型的精确度和召回率的计算方法为

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

式中: TP(true positive) 为样本的真实类别为正,预测结果也为正; FP(false positive) 为样本的真实类别为负,预测结果却为正; FN(false negative) 为样本的真实类别为正,预测得到为负。

3.2.2 疲劳表情识别网络

在训练过程中,设置学习率为 0.000 1,训练样本批次的大小为 32,网络使用随机梯度下降优化算法进行训练。模型迭代的周期为 100 epochs(训练集中全部样本训练一次为一个 epoch)。为了提高评估模型的性能,采用交叉验证法来划分数据集,训练集和测试集的比例为 8:2。如图 8 所示,随着迭代次数的增加网络识别的准确率不断提高,在 30 epochs 后,准确率达到约 89.1%。网络在训练最初的损失约为 0.20,经过 50 epochs 的迭代之后,损失值基本稳定在 0.01。

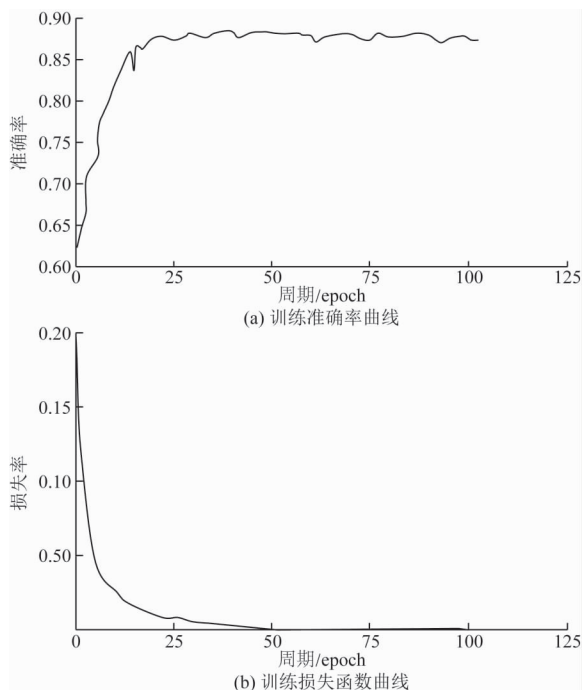


图 8 准确率与损失函数曲线

Fig. 8 Training accuracy and loss function curves

为验证双分支疲劳表情识别网络相比于单分支疲劳表情识别网络有性能的提升,绘制两种网络的 $P-R$ 曲线,如图 9 所示。由图 9 可知,加入局部特征之后的双分支网络相比于单分支网络的平均精确率提高约 3%,获得了更好的疲劳检测效果。

在数据集中提取出晴天、多云、阴天三种不同驾驶场景的样本。使用提出的双分支疲劳表情识别网络进行测试,结果如表 1 所示。提出的算法在光照条件较好的晴天以及光照条件一般的多云天均能达到 89% 以上的准确率,在光照条件较差的阴天仍能达到 85.5% 的准确率,可知提出的算法能够较好地适应不同的光照条件。

为进一步验证本文提出方法具有一定的优越性,选取近年来相关的疲劳检测方案进行比较。文献[13]采用 Adaboost 提取人脸有效的 LBP 特征,再采用支持向量机(support vector machines, SVM)进行二分类,从而判断测试者是否处于疲劳状态。文献[14]使用卷积神经网络提取图片空间域特征,再使用 LSTM 对其进行时域分析,判断驾驶员是否打哈欠。文献[15]通过提取驾驶员的眼部和嘴部特征得到驾驶员的眨眼、打哈欠的频率并建立数据库,再构建朴素贝叶斯分类器进行疲劳分析。文献[16]采用反向投影算法测量嘴部变化的频率和数量来判别测试者是否打哈欠。在数据集中选取 200 张不同驾驶环境的疲劳表情样本图片,使用本文提出方法和其他 4 种方法分别进行测试,疲劳检测性能对比结果如表 2 所示。根据表 2 可知,相比

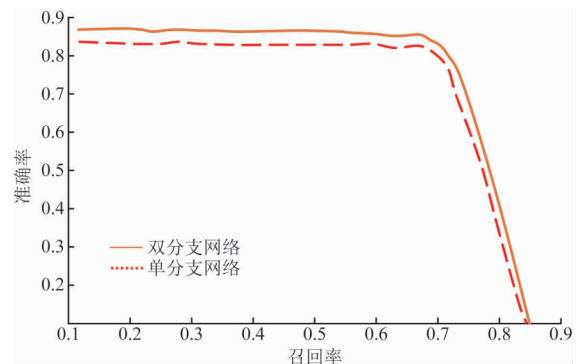


图 9 单分支和双分支网络 $P-R$ 曲线

Fig. 9 $P-R$ curves of single-branch and double-branch networks

表 1 不同光照条件下的疲劳检测结果

Table 1 Results of fatigue detection under different light conditions

光照条件	测试样本数	正确检测数	准确率/%
晴天	302	270	89.4
多云	350	312	89.1
阴天	215	184	85.5

表 2 本文方法与其他方法性能对比
Table 2 Comparison of performance between proposed method and other methods

方法	准确率 / %	精确度 / %	召回率 / %
LBP + SVM ^[13]	83.5	81.8	83.5
LSTM ^[14]	84.5	83.3	85.9
融合多特征 ^[15]	86.5	85.6	88.1
反向投影法 ^[16]	86.0	85.0	86.7
本文方法	88.0	87.1	88.8

以上通过局部特征以及局部特征融合的疲劳判别方法,提出的结合整体面部表情特征以及面部局部特征来进行疲劳判别的方法,具有相对较好的性能。

4 总结

提出疲劳检测方法结合了传统特征和神经网络进一步的特征提取。通过对传统特征提取获得面部形状外观和局部微妙的纹理细节,再使用设计的疲劳表情识别网络对传统特征进行深度融合以用于表情识别。提出的方法不仅对能对整体面部特征进行深入挖掘,而且能够考虑到包含丰富信息的面部局部特征。该算法能够适应不同光照条件的驾驶场景,并能保证较高的检测准确率。由于使用的数据集未包含戴眼镜的驾驶员图片,在未来工作中,需要进一步对有眼镜遮挡的面部疲劳表情识别进行研究。

参 考 文 献

[1] Ramzan M , Khan H U , Awan S M , et al. A survey on state-of-the-art drowsiness detection techniques [J]. IEEE Access , 2019 , 7: 61904-61919.

[2] Wei S , Zhang X , Peeta S , et al. A real-time fatigue driving recognition method incorporating contextual features and two fusion levels [J]. IEEE Transactions on Intelligent Transportation Systems , 2017 , 18(12) : 3408-3420.

[3] Zhao X , Meng C , Feng M , et al. Eye feature point detection based on single convolutional neural network [J]. IET Computer Vision , 2017 , 12(4) : 453-457.

[4] You F , Li X , Gong Y , et al. A real-time driving drowsiness detection algorithm with individual differences consideration [J]. IEEE Access , 2019 , 7: 179396-179408.

[5] 王霞,全美娇,王蒙军. 基于嘴部内轮廓特征的疲劳检测[J]. 科学技术与工程, 2016, 16(26) : 240-244.

Wang Xia , Tong Meijiao , Wang Mengjun. Fatigue detection based on the features of the inner contour of the mouth [J]. Science Technology and Engineering , 2016 , 16(26) : 240-244.

[6] Xie Y , Chen K , Murphey Y L. Real-time and robust driver yawning detection with deep neural networks [C]// IEEE Symposium Series on Computational Intelligence (SSCI) . Bengaluru: IEEE , 2018: 532-538.

[7] Liu Z , Peng Y , Hu W. Driver fatigue detection based on deeply-Learned facial expression representation [J]. Journal of Visual Communication and Image Representation , 2019 , 71: 102723.

[8] Ji Y , Wang S , Zhao Y , et al. Fatigue state detection based on multi-index fusion and state recognition network [J]. IEEE Access , 2019 , 7: 64136-64147.

[9] 戎辉,华一丁,张小俊,等. 基于迁移学习和 AlexNet 的驾驶员行为状态识别方法 [J]. 科学技术与工程, 2019 , 19(28) : 208-216.

Rong Hui , Hua Yiding , Zhang Xiaojun , et al. Driver behavior state recognition method based on transfer learning and AlexNet [J]. Science Technology and Engineering , 2019 , 19(28) : 208-216.

[10] Yang P , Shan S , Gao W , et al. Face recognition using ada-boosted gabor features [C]//Sixth IEEE International Conference on Automatic Face and Gesture Recognition. Seoul: IEEE , 2004: 356-361.

[11] Ojala T , Pietikäinen M , Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. Pattern Recognition , 1996 , 29(1) : 51-59.

[12] Abtahi S , Omidyeganeh M , Shirmohammadi S , et al. YawDD: a yawning detection dataset [C]//Proceedings of the 5th ACM Multimedia Systems Conference. New York: ACM , 2014: 24-28.

[13] Zhang Y , Hua C. Driver fatigue recognition based on facial expression analysis using local binary patterns [J]. Optik-International Journal for Light and Electron Optics , 2015: 4501-4505.

[14] Zhang W , Su J. Driver yawning detection based on long short term memory networks [C]// IEEE Symposium Series on Computational Intelligence (SSCI) . Hawaii: IEEE , 2017: 1-5.

[15] 黄占鳌,史晋芳. 多特征融合的驾驶员疲劳检测研究 [J]. 机械科学与技术, 2018 , 37(11) : 1750-1754.

Huang Zhanao , Shi Jinfang. Research on driver fatigue detection based on multi-feature fusion [J]. Mechanical Science and Technology , 2018 , 37(11) : 1750-1754.

[16] Omidyeganeh M , Shirmohammadi S , Abtahi S , et al. Yawning detection using embedded smart cameras [J]. IEEE Transactions on Instrumentation and Measurement , 2016 , 65(3) : 1-13.