



文章编号:1001-9081(2019)07-2098-05

DOI:10.11772/j.issn.1001-9081.2018122441

基于人眼信息特征的人体疲劳检测

罗 元¹, 云明静^{1*}, 王 艺¹, 赵立明²

(1. 重庆邮电大学 光电工程学院, 重庆 400065; 2. 重庆邮电大学 先进制造工程学院, 重庆 400065)

(* 通信作者电子邮箱 975727433@qq.com)

摘要:人眼状态是反映疲劳程度的重要指标,头部姿势变化、光线等因素对人眼定位造成很大影响,从而影响人眼状态识别以及疲劳检测的准确性,为此提出了一种利用级联卷积神经网络通过检测人眼6个特征点来识别人眼状态进而识别人体疲劳的方法。首先,一级网络采用灰度积分投影结合区域—卷积神经网实现人眼的检测与定位;然后,二级网络将人眼图片进行分割后采用并联子卷积系统进行人眼特征点回归;最后,利用人眼特征点计算人眼开闭度识别当前人眼状态,并根据单位时间闭眼百分比(PERCLOS)准则判断人体疲劳状态。实验结果表明,利用所提方法实现了在归一化误差为0.05时,人眼6特征点的平均检测准确率为95.8%,并根据模拟视频帧的PERCLOS值识别疲劳状态验证了该方法的有效性。

关键词:灰度积分投影;卷积神经网络;人眼定位;人眼状态识别;疲劳检测

中图分类号: TP391.413 **文献标志码:**A

Human fatigue detection based on eye information characteristics

LUO Yuan¹, YUN Mingjing^{1*}, WANG Yi¹, ZHAO Liming²

(1. College of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. College of Advanced Manufacturing Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The eye state is an important indicator reflecting the degree of fatigue. Changes in head posture and light have a great influence on human eye positioning, which affects the accuracy of eye state recognition and fatigue detection. A cascade Convolutional Neural Network (CNN) was proposed, by which the human eye state could be identified by detecting six feature points of human eye to identify human body fatigue. Firstly, grayscale integral projection and regional-convolution neural network were used as the first-level network to realize the positioning and detection of human eyes. Then, the secondary network was adopted to divide the human eye image and parallel sub-convolution system was used to perform human eye feature point regression. Finally, human eye feature points were used to calculate the human eye opening and closing degree to identify the current eye state, and the human body fatigue state was judged according to the PERcentage of eyelid CLOSure over the pupil time (PERCLOS) criterion. The experimental results show that the average detection accuracy of six eye feature points reaches 95.8% when the normalization error is 0.05, thus the effectiveness of the proposed method is verified by identifying the fatigue state based on the PERCLOS value of analog video.

Key words: grayscale integral projection; Convolutional Neural Network (CNN); eye positioning; eye state identification; fatigue detection

0 引言

疲劳是指在一定环境条件下因机体长时间从事体力或脑力劳动而引起的劳动效率趋向下降的状态。疲劳不仅是多种慢性疾病的起源,更重要的是疲劳在某些领域会对社会安全造成重大危害,尤其在高空建筑作业、车辆驾驶、大型复杂工业等高风险作业中,每年因人员疲劳操作造成事故数量巨大。目前,疲劳检测主要分为基于生理特征信号检测、基于视觉特征检测^[1],对于驾驶疲劳检测还包括驾驶行为的检测方式。基于生理特征信号的检测方式具有较高的准确率,但是该方法信号采集设备复杂,相关设备的小型化与实用化也仍有不足,因此目前的主要研究还是在实验室进行^[2];而基于

视觉检测的方式在保持较高准确率的同时设备要求低,同时具有非侵入性特点^[3],是疲劳检测的主要研究方向之一。

基于眼睛状态的疲劳检测主要包括人眼定位以及人眼状态的识别^[4~5]。Deng等^[6]利用肤色模型结合人脸三庭五眼的布局定位人眼,并利用人眼的积分投影区域大小识别人眼状态;这种方法虽然算法简单,但是定位的准确率受环境光照影响较大,且由于人眼区域在图像中占的比例很小,利用积分投影的人眼状态识别准确率较低。李响等^[7]利用图像的矩特征通过计算人眼模板的Zernike矩特征向量与待识别的人脸区域作相似度计算,选取相似度最大的区域作为人眼区域;该方法虽然能够减小环境光照对于检测结果的影响,但计算量较大,并且结果受选取的人眼模板影响较大。在人眼状态

收稿日期:2018-12-11;修回日期:2019-02-25;录用日期:2019-03-21。

基金项目:重庆市科委基础与前沿研究计划项目(cstc2016jcyjA0537)。

作者简介:罗元(1972—),女,湖北宜昌人,教授,博士,主要研究方向:数字图像处理; 云明静(1994—),男,四川绵阳人,硕士研究生,主要研究方向:机器视觉; 王艺(1999—),女,重庆渝北人,主要研究方向:光信息处理; 赵立明(1981—),男,河北秦皇岛人,讲师,博士,主要研究方向:机器视觉。



识别中,传统人眼状态识别方法通过对检测的人眼区域进行人眼形状拟合^[8]或者对人眼区域采用积分投影根据投影区域的宽度^[9]识别人眼的张开程度。形状拟合的方法不仅计算复杂而且在头部姿态变化时容易失真,积分投影的方法要求检测的人眼区域完全匹配人眼的上下眼睑,实际情况下很难做到,容易造成投影区域宽度过大或过小影响检测结果。

针对这些传统算法在实际环境中易受头部姿势变化、光线等因素的干扰,以卷积神经网络(Convolutional Neural Network, CNN)为代表的深度学习在图像检测领域具有较高的准确性同时能具有较强的鲁棒性,在目标检测方面就有区域—卷积神经网络(Region-Convolutional Neural Network, R-CNN)^[10]、快速区域—卷积神经网(fast Region-Convolutional Neural Network, fast R-CNN)^[11]、多任务卷积神经网(Mutil-task-Convolutional Neural Network, Mutil-task-CNN)^[12]等,但是这些网络结构较为复杂,常用于图像中的多目标检测,本文针对人脸图像设计了一种级联神经网络结构来检测人眼以及人眼特征点,并提出两点改进:1)采用灰度积分投影进行人眼粗定位后将结果输入神经网络进行精定位,提高了检测速度;2)将人眼图像进行 2×2 分割后利用4个子卷积网络构成的卷积神经网进行特征点回归预测,同时考虑到眼睑与眼部其他特征的差异,在最后一层卷积层采用不共享权值,在一定程度上提高了人眼特征点的定位准确率。

1 算法流程

一般情况下,对于在自然环境中得到的人物图像,人眼在图像中所占比例较小,直接检测人眼往往比较困难且检测准确率不高。本文首先进行人脸检测,对得到的人脸图像利用改进的级联卷积神经网进行人眼的定位以及人眼特征点检测,通过6特征点计算人眼闭度从而识别人眼状态,最后根据计算模拟视频帧的单位时间闭眼百分比(PERcentage of eyelid CLOSure over the pupil time, PERCLOS)值识别疲劳状态验证本文方法,疲劳检测算法如图1所示。

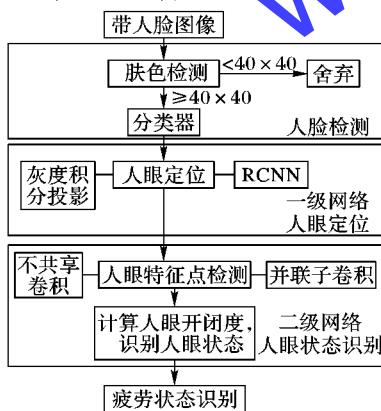


图1 疲劳检测流程

Fig. 1 Flow chart of fatigue detection

在人脸检测部分,本文采用了肤色模型^[13]与支持向量机(Support Vector Machine, SVM)分类器从全图中提取人脸图像。经过肤色检测后找到图像中可能存在肤色的连通区域的最小外接矩形,经过实验表明,小于 40×40 的人脸图像中人眼属于小目标,针对小目标人眼的状态识别非常困难且检测效率较低,因此舍弃小于 40×40 大小的矩形区域,将剩下的矩形区域内的图像从原图中分割作为候选区域输入SVM分类器进行分类识别。

2 数据集准备

训练数据 本文采用CEW人眼睁闭数据库来训练一级网络,CEW库是用于人眼睁闭检测的数据集,该数据集共包含2423张不同人眼开闭状态的带人脸图片,其中包括来自于网络的1192张闭眼图片以及来自于LFW人脸数据库的1231张睁眼图片。利用BBox-Label-Tool工具制作人眼位置框标签,每幅图片的标签信息包含实际矩形框的坐标信息、宽高信息以及框内目标的类别信息。训练二级网络从CEW库中裁剪人眼图片2133张,其中闭眼图片共1015张,睁眼图片共1118张,左眼图片1060张,右眼图片1073张,利用sloth工具为人眼图片添加特征点标签。

实验数据 由于目前缺少关于人体疲劳的公开数据集,本文根据自采集模拟人体清醒与疲劳两种状态的视频进行实验。首先定义当上下眼睑均未遮挡虹膜时的眼睛状态为完全张开状态,通过实验统计眼睛完全张开时开闭度在(0.35, 0.4]区间,并从清醒状态视频与疲劳状态视频中各选取200帧共400帧带人脸视频帧进行实验,选取的视频帧均包含不同人眼张开度,其中在清醒状态帧中选取人眼开闭度为(0.1, 0.2]共40帧,(0.2, 0.3]共100帧,(0.3, 0.4]共60帧,同时在疲劳帧中人眼完全闭合共50帧,开闭度为[0, 0.1]共150帧。

3 人眼检测与状态识别

3.1 人眼检测

由于人眼区域占人脸比例较小,若直接进行人眼定位需要对整幅人脸图像进行计算,造成大量计算资源的浪费,因此本文设计了一种结合灰度积分投影与卷积神经网的方法,先对得到的人脸区域图像进行人眼的粗定位再利用神经网络提取特征并进行人眼精定位。灰度积分投影法^[14]是一种简易的人眼粗定位方法,根据人脸先验知识,人眼及眉毛区域的灰度较人脸其他区域低,因此将人脸图像转换为灰度图像并将像素灰度值投影到水平或者垂直方向累加得到灰度曲线,对于待检测图像 $f(x, y)$,其水平灰度积分可表示为:

$$G(x) = \frac{1}{x_2 - x_1} \sum_{x_1}^{x_2} f(x, y) \quad (1)$$

其中 x_i 表示水平方向第*i*个像素的位置。

由图2水平灰度积分投影曲线可知,人脸灰度第一极大值与第二极大值点之间对应于人的额头位置与鼻中部位置,人眼位置就位于两个极大值点之间的极小值点附近,为了使截取的图片完整包含人眼,提取第一极大值与第二极大值点之间的区域即额头至鼻中部区域图片作为人眼的候选区域。

人眼定位网络结构为一种7层结构的网络结合2层RPN(Region Proposal Network),主要实现人眼分类与人眼框回归预测两种任务,网络结构如图3所示。

网络结构中第1层为输入层;第2层为灰度积分投影人眼粗定位;第3、4、5层为卷积层,第6层为全连接层,第7层为两个全连接层构成的输出层;RPN结构为两层全卷积层。各层的具体结构如下:

1) 第1层为输入层,由输入样本构成,实验采用的人脸图片大小均为 100×100 大小,即输入样本为 100×100 的矩阵,每一个矩阵中的点表示图像像素的大小。

2) 第2层为灰度积分投影人眼粗定位,将得到的人眼候选区域图像调整为 60×100 大小。



3) 第3、4、5层为卷积层;用来提取图像特征,在第3、4、5层分别采用了8、2、1种卷积核,卷积核大小均为 3×3 ,最终得到了16个特征图;在每个卷积层的后面采用了 2×2 的最大池化层,主要作用是用来简化卷积层输出的特征。

4) RPN为全卷积网络,其主要包含两个卷积层,采用 5×5 大小的卷积核,主要作用是提取人脸图像中人眼可能存在的区域,根据文献[15]建议采用了9种anchor,经过非极大值抑制(Non-Maximum Suppression, NMS)^[16]后保留60个人眼候选框,其中正负类比例为1:3。

5) 第6层为全连接层,其主要作用是将各个特征图映射成一个特征向量,便于后续的分类与人眼位置回归。

6) 第7层为输出层,对于人眼分类采用softmax分类最后输出分类概率,对于人眼框回归输出为人眼定位矩形框的长、宽以及中心点的坐标。

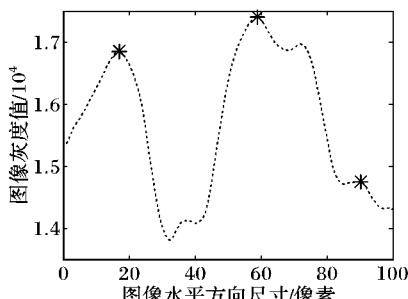


图2 人脸水平灰度积分投影

Fig. 2 Horizontal gray integral projection of human face

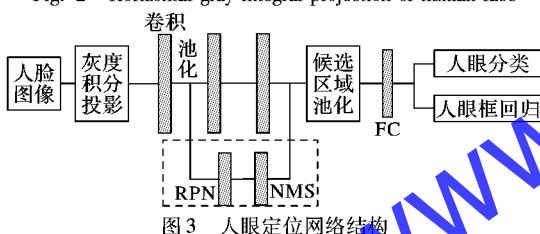


Fig. 3 Network structure of eye positioning

3.2 人眼状态识别

本文利用改进的卷积神经网络作为人眼检测后的2级网络进行人眼特征点回归,本文共检测了人眼的6个特征点,如图4所示。其中A、B点分别为上下眼睑的内交点与外交点,C、E点为AB连线三等分点靠近内交点处与上下眼睑的交点,D、F点为AB连线三等分点靠近外交点处与上下眼睑的交点。

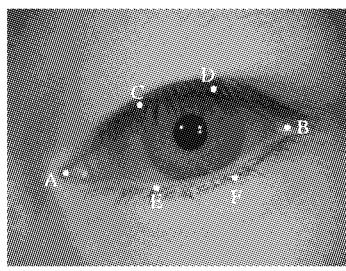


图4 人眼的6个特征点

Fig. 4 Six feature points of human eye

定义人眼的开闭度:

$$\frac{d_1 + d_2}{2d} \quad (2)$$

其中: d 为点A、B之间的直线距离, d_1 为点C、E之间的距离, d_2 为点D、F之间的距离。当比值 < 0.1 时,认为当前人眼状态

为闭眼状态,0.1为根据P80^[17]计算的经验值,人眼特征点检测网络结构如图5所示。

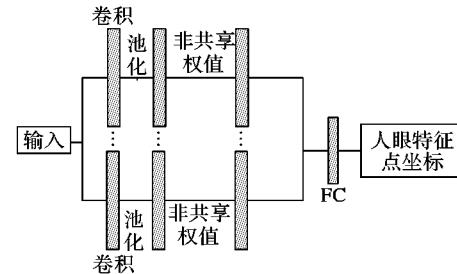


图5 特征点检测网络结构

Fig. 5 Network structure of feature points detection

网络具体结构如下:

1) 第1层为输入层,将得到的人眼图像调整为 48×48 大小,输入大小为 48×48 的矩阵,将输入进行 2×2 分割后分别连接4个并行子卷积结构。

2) 每个子卷积结构为3层网络;各个卷积层分别采用8、2、4种卷积核,在前两个卷积层后采用 2×2 的最大池化降维,最终得到了16个特征图,因为在最后一个卷积层采用了不共享卷积方式。共享卷积的主要优点是减少了参数个数,但是忽视了图像不同部分特征的差异性,对于人眼特征点的预测,眼睑的高层特征与眼睛其他部位差别比较大,且由于人眼图片较小,因此在最后一层卷积层采用非共享卷积能够更好地提取图像特征。

3) 第5层为全连接层,将子卷积的各特征图映射成一个特征向量。

4) 第6层输出6个人眼特征点的坐标共12个神经元。

3.3 卷积神经网络训练

卷积神经网的训练主要是通过梯度下降法与反向传播来实现的,主要原理输入数据,计算每一层的激活值,最后计算损失函数,根据损失函数大小是否满足设定的阈值判断是否结束迭代过程,若不满足条件则反向更新参数并继续进行迭代。

对于神经网中第k层第m个特征图,第j个神经元,其激活值表达式为:

$$a_j^{k,m} = g(x_j^{k,m}) \quad (3)$$

其中: $a_j^{k,m}$ 为该神经元激活值; $x_j^{k,m}$ 为该神经元输入值; $g(\cdot)$ 为激活函数。本文采用的激活函数为tanh函数,其表达式为 $(e^z - e^{-z}) / (e^z + e^{-z})$,其中e为自然常数,可以看出tanh函数的取值范围在 $[-1, 1]$,均值为0,在训练中较sigmod函数具有更快的收敛速度。

在一级网络的输出中主要存在人眼分类与人眼位置回归两种任务,对于分类任务,采用交叉熵损失函数,表达式为:

$$Loss = -[y \ln \hat{y} + (1-y) \ln(1-\hat{y})] \quad (4)$$

其中: y 为分类标签实际值; \hat{y} 为神经网络输出值。

在人眼位置回归任务中,假设人眼框标签实际值为 $X = (x, y, w, h)$,预测值为 $X' = (x', y', w', h')$,损失函数采用二次损失函数,表达式为:

$$L_{err} = \sqrt{(x - x')^2 + (y - y')^2 + (w - w')^2 + (h - h')^2} \quad (5)$$

其中:(x, y, w, h)分别表示实际回归框左上顶点的坐标与实际回归框的宽度与高度,而(x', y', w', h')则表示预测回归框的坐标与宽高。

在二级网络主要任务属于人眼特征点回归任务,因此也



采用一种二次损失函数,其表达式为:

$$poin_{err} = \sum_{i=1}^N \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} / l \quad (6)$$

其中: (x_i, y_i) 为第 i 个特征点实际标注点的坐标; (x'_i, y'_i) 为其预测点的坐标; i 的顺序为 $1 \sim N$, N 为特征点的个数; l 为人体图像的长度。

采用梯度下降法进行神经网络参数的更新,设标签实际值为 \mathbf{Y} , 神经网络输出值为 \mathbf{Y}' , 那么网络中权值以及偏置更新:

$$\mathbf{w} \leftarrow \mathbf{w} + \eta \frac{\partial(\mathbf{Y} - \mathbf{Y}')}{\partial \mathbf{w}} \quad (7)$$

$$\mathbf{b} \leftarrow \mathbf{b} + \eta \frac{\partial(\mathbf{Y} - \mathbf{Y}')}{\partial \mathbf{b}} \quad (8)$$

其中: \mathbf{w} 为权值矢量; \mathbf{b} 为偏置矢量; η 为学习率, 训练过程迭代 1000 次, 学习率采用以 0.1 为基数的 luong234 衰减, 即在总步数的三分之二后开始衰减, 在后面的步数中平均衰减 4 次, 每次衰减为上次的二分之一。

4 实验结果及分析

为了验证人眼定位算法的有效性,本文在人眼定位部分与 Haar + AdaBoost 算法、Gabor + SVM 等算法以及未作人眼粗定位的本文神经网络作了比较,同时分析了不同光照以及不同头部偏转姿态下的人眼定位,强光设定为平均灰度大于 180,弱光灰度设定为平均灰度小于 60。实验在 Matlab 2016b 环境下进行,计算机的 CPU 主频为 3.30 GHz, 实验结果分别如表 1~2 所示。

表 1 不同方法的人眼定位准确率和检测速率对比

Tab. 1 Comparison of human eye positioning accuracy and detection rate with different methods

检测方法	准确率/%			检测速率/ (帧·s ⁻¹)
	自然状态	强光	暗光	
本文方法(无人眼粗定位)	93.4	86.20	71.2	7
本文方法(有人眼粗定位)	96.2	88.70	75.7	11
Haar + AdaBoost	87.2	69.20	53.6	10
YOLO	79.2	70.60	62.3	14
Gabor + SVM	85.8	77.85	65.9	8
Spp-net	90.7	83.60	72.5	7

表 2 不同姿态下人眼定位结果

Tab. 2 Human eye positioning results with different postures

姿态	检测帧数	正确检测帧数	正确率/%
0°偏转	40	40	100.0
左偏 15°	40	35	87.5
左偏 35°	40	28	70.0
右偏 15°	40	36	90.0
右偏 35°	40	26	65.0

从表 1~2 可以看出,本文采用的方法与目前人眼定位中常用的 Haar 特征结合 AdaBoost 算法、Gabor + SVM 算法以及 Spp-net 目标检测算法相比具有最高的定位准确率和检测速率, 检测速率略低于 YOLO 目标检测算法但具有更高的定位准确率, 同时与未采用灰度积分投影进行人眼粗定位的神经网络以及相比, 在自然状态下检测准确率提升了 2.8 个百分点, 而检测速率由 7 帧/s(Frames Per Second, FPS) 提升到了 11 帧/s, 提升了 57% 左右, 说明采用灰度积分投影的粗定位在一定程度上提高了检测性能。同时在不同光照下的检测结

果表明本文对于光照环境变化的人眼定位具有一定鲁棒性, 对于人体头部姿态偏转角度在小幅度偏转时具有较好的检测效果, 但是当偏转角度过大时定位准确率会大幅下降。

从图 6 中可以看出, 本文方法能够实现较好的特征点检测定位, 在归一化误差允许范围为 0.05 时各个特征点的定位准确率均达到 92% ~ 98%, 平均定位准确率达 95.8%, 其中 A、B 点位于角点, 特征较为复杂, 定位准确率较高, 下眼睑两点因为闭眼时睫毛的遮挡导致特征模糊因此检测准确率较上眼睑两点低。同时本文与主动形状模型(Active Shape Model, ASM)、主动外观模型(Active Appearance Model, AAM)传统方法以及受限深度卷积神经网络(Tasks-Constrained Deep Convolutional Network, TCDCN)方法进行了比较, 本文方法、ASM、AAM、TCDCN 方法的特征点定位平均准确率分别为 95.8%、78.5%、72.4%、92.4%, 结果表明本文方法具有最高的平均定位准确率。

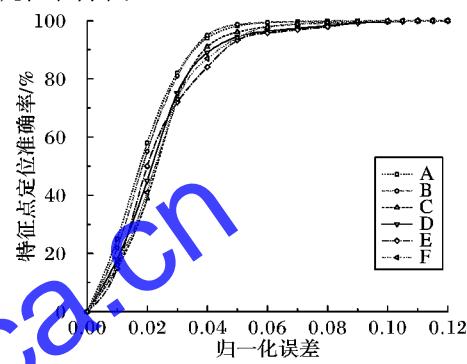


图 6 各特征点检测准确率

Fig. 6 Accuracy of each feature point

通过对选取的 400 帧带人脸视频帧检测人眼特征点后计算人眼开闭度识别人眼状态。表 3 中人眼状态平均检测准确率为 97.5%, 由表 3 可知, 在人眼状态检测中, 误检主要存在于阈值 0.1 附近, 人眼开闭度越接近阈值 0.1, 检测结果越容易出错, 而人在正常睁眼状态下开闭度在(0.2, 0.3]区间, 当人眼开闭度位于 0.1 附近时一般是人处于眨眼状态。

表 3 不同人眼开闭度下人眼状态检测

Tab. 3 Human eye state detection results with different human eye opening and closing degrees

人眼 开闭度	实际 帧数	检测 帧数	正确 检测	失 败 检测	准 确 率/%	状 态
0 (0,0.1]	50	50	49	1	98	闭眼状态
(0.1,0.2]	150	144	143	1	95.3	闭眼状态
(0.2,0.3]	40	36	35	1	92.5	闭眼状态
(0.3,0.4]	100	96	95	1	95.0	睁眼状态
(0.3,0.4]	60	60	59	1	98.3	睁眼状态

PERCLOS 是指一定时间内眼睛闭合时间占总时间的百分比, 是目前视觉疲劳检测方面最有效的指标。研究表明人平均每分钟眨眼 10 至 20 次, 即平均每 3 s 到 6 s 一次, 人的正常眨眼时间为 0.2 ~ 0.3 s, 此时 PERCLOS 值处于 3.3% 至 10% 之间, 而若眨眼时间为 0.5 ~ 3 s 则可视为疲劳状态, 此时 PERCLOS 值处于 16.7% 至 100% 之间。为了提高容错率并且更为准确地区分疲劳与清醒状态, 规定 PERCLOS ≥ 20% 时, 则判断当前人体处于疲劳状态。图 7 是分别模拟人体清醒与疲劳的两段视屏的人眼状态检测结果两段。其中图 7(a)是模拟人体清醒状态下的人眼状态检测结果, 检测到模拟视频眨眼过程中的闭眼时间达到 4 帧, 可以计算 PERCLOS



为8%;在疲劳状态下,结果如图7(b)所示,检测到模拟视频眨眼过程中的闭眼时间为22帧,可以计算PERCLOS为44%,均能够正确反映身体状态。测试结果证明了本文方法的有效性。

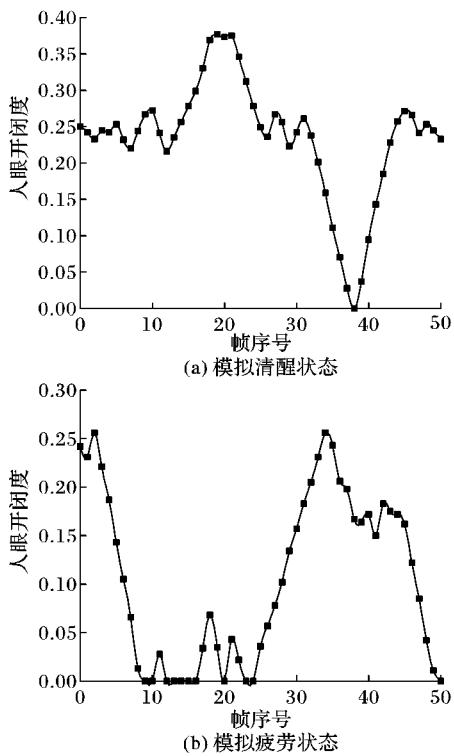


图7 模拟清醒与疲劳状态下人眼开闭度

Fig. 7 Human eye opening and closing degree under simulated fatigue and awake state

5 结语

本文针对视觉疲劳检测,提出了一种级联卷积神经网络。一级网络结合灰度积分投影与卷积神经网络定位人眼提高了检测速度与检测准确率,利用改进的卷积神经网络进行人眼特征点检测,通过人眼特征点计算人眼开闭度来识别人眼状态,并最终根据PERCLOS准则检测人体疲劳。结果表明本文方法具有较高的检测速度与检测准确率,能够满足疲劳检测的需要,但是本文方法针对头部姿势偏转过大时人眼定位效果还有待进一步提高,可以作为下一步的研究方向。

参考文献 (References)

- [1] YOU Z, GAO Y, ZHANG J, et al. A study on driver fatigue recognition based on SVM method [C]// Proceedings of the 2017 4th International Conference on Transportation Information and Safety. Piscataway, NJ: IEEE, 2017: 693–697.
- [2] CHAI R, NAIK G, NGUYEN T N, et al. Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system [J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(3): 715–724.
- [3] XU J, MIN J, HU J. Real-time eye tracking for the assessment of driver fatigue [J]. Healthcare Technology Letters, 2018, 5(2): 54–58.
- [4] 唐广发,张会林.人眼疲劳预测技术的研究[J].计算机工程与应用,2016,52(9):213–218.(TANG G F, ZHANG H L. Research on human eye fatigue prediction technology[J]. Computer Engineering and Applications, 2016, 52(9): 213 –218.)
- [5] ZENG S, LI J, JIANG L, et al. A driving assistant safety method based on human eye fatigue detection [C]// Proceedings of the 2017 Control and Decision Conference. Piscataway, NJ: IEEE, 2017: 6370–6377.
- [6] DENG Z, JING R, JIAO L, et al. Fatigue detection based on isophote curve [C]// Proceedings of the 2015 International Conference on Computer and Computational Sciences. Piscataway, NJ: IEEE, 2015: 146–150.
- [7] 李响,谭南林,李国正,等.基于Zernike矩的人眼定位与状态识别[J].电子测量与仪器学报,2015(3):390–398.(LI X, TAN N L, LI G Z, et al. Human eye localization and state recognition based on Zernike moment[J]. Journal of Electronic Measurement and Instrumentation, 2015(3): 390 –398)
- [8] ARAUJO G M, FML R, JUNIOR W S, et al. Weak classifier for density estimation in eye localization and tracking[J]. IEEE Transactions on Image Processing, 2017, 26(7): 3410–3424.
- [9] SONG M, TAO D, SUN Z, et al. Visual-context boosting for eye detection[J]. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), 2010, 40(6): 1460–1467.
- [10] LI J, WONG H C, LO S L, et al. Multiple object detection by a deformable part-based model and an R-CNN[J]. IEEE Signal Processing Letters, 2018, 25(2): 288–292.
- [11] LI J, LIANG X, SHEN S M, et al. Scale-aware fast R-CNN for pedestrian detection [J]. IEEE Transactions on Multimedia, 2018, 20(4): 985–996.
- [12] ABDULNABI A H, WANG G, LU J, et al. Multi-task CNN model for attribute prediction [J]. IEEE Transactions on Multimedia, 2015, 17(11): 1949–1959.
- [13] LUO Y, GUAN Y P. Adaptive skin detection using face location and facial structure estimation [J]. IET Computer Vision, 2017, 11(7): 550–559.
- [14] YANG W, ZHANG Z, ZHANG Y, et al. Real-time digital image stabilization based on regional field image gray projection[J]. Journal of Systems Engineering and Electronics, 2016, 27(1): 224–231.
- [15] KÖNIG D, ADAM M, JARVERS C, et al. Fully convolutional region proposal networks for multispectral person detection[C]// Proceedings of the 2017 Computer Vision and Pattern Recognition Workshops. Washington, DC: IEEE Computer Society, 2017: 243–250.
- [16] FU L, ZHANG J, HUANG K. Mirrored non-maximum suppression for accurate object part localization [C]// Proceedings of the 2015 3rd IAPR Asian Conference on Pattern Recognition. Piscataway, NJ: IEEE, 2015: 51–55.
- [17] 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, 2017, 18(3): 545–557.

This work is partially supported by the Basic and Frontier Research Project of Chongqing Municipal Science and Technology Commission (cstc2016jcyjA0537).

LUO Yuan, born in 1972, Ph. D., professor. Her research interests include digital image processing.

YUN Mingjing, born in 1994, M. S. candidate. His research interests include machine vision.

WANG Yi, born in 1999. Her research interests include optical information processing.

ZHAO Liming, born in 1981, Ph. D., lecturer. His research interests include machine vision.