

基于Dlib和变种Transformer的哈欠检测方法

廖冬杰^{1,2}

(1.华东交通大学,南昌 330013;2.江西省先进控制与优化重点实验室,南昌 330013)

【摘要】为及时检测驾驶员的疲劳驾驶情况,提出了一种基于Dlib和变种Transformer的哈欠检测方法。首先,基于Dlib的人脸关键点模型构建驾驶员眼部和嘴部的哈欠特征矩阵,然后在视频检测领域提出一种变种Transformer模型对哈欠特征矩阵进行提取和分类,最后基于YawDD数据集进行验证,结果表明,所提出算法的哈欠检测准确率达96.8%,高于已有算法,适用于驾驶员疲劳驾驶时哈欠行为的检测。

关键词:哈欠检测 Dlib 哈欠特征矩阵 变种Transformer YawDD

中图分类号:TP391.4 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20220453

Yawn Detection Method Based on Dlib and Variant Transformer

Liao Dongjie^{1,2}

(1. East China Jiaotong University, Nanchang 330013; 2. Key Laboratory of Advanced Control and Optimization of Jiangxi Province, Nanchang 330013)

【Abstract】In order to detect the driver's fatigue driving situation in time, this paper proposed a yawn detection method based on Dlib and variant Transformer. First, the yawn feature matrix of the driver's eyes and mouth was constructed based on the face key point model of Dlib. Then a variant Transformer model was proposed in the field of video detection, to extract the yawn feature matrix and classify the results. Finally, it was verified based on the YawDD dataset. The results show that the yawn detection accuracy of the proposed algorithm is 96.8%, which is higher than the existing algorithms, and is suitable for the detection of yawning behavior when the driver is fatigued.

Key words: Yawn detection, Dlib, Yawning features matrix, Variant Transformer, YawDD

【引用格式】廖冬杰. 基于Dlib和变种Transformer的哈欠检测方法[J]. 汽车技术, 2023(3): 42-48.

LIAO D J. Yawn Detection Method Based on Dlib and Variant Transformer[J]. Automobile Technology, 2023(3): 42-48.

1 前言

疲劳驾驶是造成交通事故的主要原因之一^[1],而打哈欠是驾驶员疲劳初期的主要表现形式之一^[2]。因此,研究驾驶员疲劳初期的哈欠特征,从而进行准确检测,对保障道路交通安全具有重要意义。

国内外研究人员针对哈欠检测已进行了大量的研究。马素刚等^[3]充分关注人脸的各种特征,以驾驶员面部图片作为输入,通过卷积核进行特征提取,并利用Softmax分类器判断驾驶员是否打哈欠,但只关注了单帧图像的分类,没有充分利用多帧图像在时间维度上的关联性,可能把讲话、唱歌、大笑等张嘴行为误检测为打哈欠。Mateusz Knapik等^[4]对热成像视频进行人脸区域检测,提出眼角检测算法,实现了人脸对齐,通过检测嘴部区域的快速温度变化判断驾驶员的打哈欠行为,但未考虑大幅度张嘴呼吸以及咳嗽等特殊情

况。瑞鹏等人^[5]提出了一种基于多任务卷积神经网络(Multi-Task Convolutional Neural Network, MTCNN)的加速优化算法,对图像中驾驶员是否存在张嘴行为进行分类,以嘴部持续张开时间作为评判驾驶员是否打哈欠的标准,但此类方法分类标准无法准确度量,无法确定嘴部持续张开时间阈值,只能根据经验给出,无法达到最优效果。王超等^[6]提出只关注嘴部图像,利用卷积神经网络(Convolutional Neural Network, CNN)和长短期记忆(Long Short-Term Memory, LSTM)网络对视频进行空间和时间的特征提取,从而实现哈欠检测,但是这类方法哈欠检测特征单一,未考虑眼部等特征与哈欠行为的关联,缺乏全面性。

本文基于Dlib和变种Transformer^[7]模型,针对哈欠检测特征单一的问题,同时关注左右眼部、嘴部的特征变化,构建包含直接哈欠特征和隐含哈欠特征的哈欠特征矩阵,有效避免驾驶员嘴部呼吸、咳嗽等嘴部特殊情

况的误检;针对无法度量阈值的问题,采用深度学习的方法,利用变种Transformer模型进行隐含哈欠特征提取,并实现哈欠分类,提高算法的准确性;针对帧与帧之间缺乏联系的问题,引进多头注意力机制和序列编码,以降低检测中大笑、说话等情况的误检可能性。

2 基于Dlib和变种Transformer的哈欠检测方法原理

Dlib是一个包含众多机器学习算法、图像处理和数值计算等丰富功能的开源工具箱,广泛应用于人脸识别

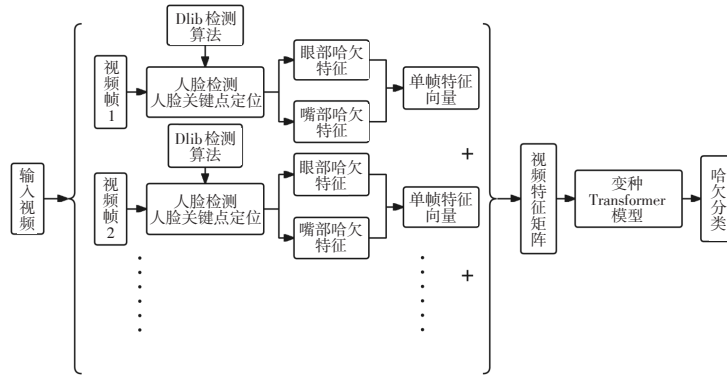


图1 哈欠检测原理

b. 哈欠特征矩阵构建。根据定位关键点坐标提取眼部、嘴部直接哈欠特征,得到每帧的哈欠特征向量,在时间维度进行堆积,构建包含直接哈欠特征和隐含哈欠特征的视频哈欠特征矩阵。

c. 哈欠检测。在得到视频哈欠特征矩阵的基础上,利用变种Transformer模型进行特征提取和视频分类,完成对视频中是否有打哈欠行为的检测。

3 基于Dlib的哈欠特征矩阵构建

3.1 Dlib人脸关键点定位

Dlib人脸关键点定位使用了Vahid Kazemi等人提出的ERT级联回归算法^[8],该算法构建了级联的梯度提升决策树(Gradient Boosting Decision Tree,GBDT),使得预测的人脸形状即关键点不断回归到真实位置^[9],其基本思想是:每一个GBDT的每个叶子节点均储存着残差回归量,当输入经过某一叶子节点时,将当前输入和叶子节点储存的残差回归量相加,起到回归作用,所有残差回归量相加后,即可获得预测的人脸形状。与传统算法相比,ERT算法具有较出色的准确率和检测速度,同时还能够处理数据缺失的问题^[8]。ERT算法核心表达式为:

$$S = (\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_p^T)^T \in \mathbf{R}^{2p} \quad (1)$$

$$\hat{S}^{(u+1)} = \hat{S}^{(u)} + \gamma_i(I, \hat{S}^{(u)}) \quad (2)$$

别、分类等领域。Transformer模型是一种自然语言处理模型,由编码器(Encoder)和解码器(Decoder)组成^[7],可以完成特征提取,同时捕获远距离的依赖关系,实现分类和生成等功能。本文旨在实现驾驶员哈欠检测,故提出基于Dlib和变种Transformer的哈欠检测方法,其原理如图1所示。

本文提出的哈欠检测方法主要包含3个部分:

a. 人脸检测和关键点定位。通过基于集成回归树(Ensemble of Regression Trees,ERT)的Dlib算法进行人脸检测以及68点关键点定位。

式中, S 为由 p 个关键点坐标组成的人脸真实形状; $\mathbf{x}_i^T \in \mathbf{R}^2 (i=1,2,\dots,p)$ 为面部图像 I 的第 i 个关键点的坐标; $\hat{S}^{(u)}$ 为第 u 次的人脸估计形状; $\gamma_i(I, \hat{S}^{(u)})$ 为第 u 次计算得到的残差回归量。

考虑到Dlib在人脸检测和人脸关键点检测上的出色能力,本文利用Dlib提供的人脸关键点模型^[10]对车内驾驶员的68个人脸关键点进行检测,结果如图2所示,试验证明,Dlib人脸关键点检测算法在车内环境中有着优秀的关键点定位能力。因此,利用Dlib人脸关键点模型分析驾驶员疲劳时的眼部和嘴部哈欠特征是可行的。



图2 车内环境人脸检测关键点

3.2 哈欠特征构建

眼睛和嘴是驾驶员在打哈欠时表现最为突出的部分。本文利用Dlib人脸关键点模型得到人脸关键点,提取眼部的12个关键点坐标信息、嘴部的6个关键点坐标

信息,如图3、图4所示,以眼部和嘴部张度作为直接哈欠特征,得到视频帧哈欠特征向量,再按照时间维度堆积,构建包含嘴部持续张开时间、眼睛闭合时间比例(Percentage Eyelid Closure over the Pupil over Time, PERCLOS)等隐含哈欠特征的视频哈欠特征矩阵,从而进行哈欠行为检测。

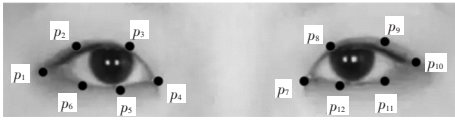


图3 眼部关键点

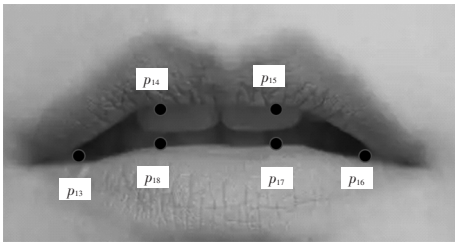


图4 嘴部关键点

3.2.1 眼部哈欠特征

根据Soukupová等人^[11]提出的观点,驾驶员眼部开合程度的改变可以通过眼部纵横比(Eye Aspect Ratio, EAR)的变化来表征。用 p_i 表示图3、图4中编号为 i 的点,左、右眼部纵横比 R_{cal} 和 R_{car} 的计算公式为:

$$R_{cal} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (3)$$

$$R_{car} = \frac{\|p_8 - p_{12}\| + \|p_9 - p_{11}\|}{2\|p_7 - p_{10}\|} \quad (4)$$

John Sofia Jennifer提出,为了定位角度更为宽广的朝向角,不必同时考虑2只眼睛,仅使用单眼检测即可达到出色效果^[12],但本文考虑到驾驶员可能存在揉眼、扭头等行为造成实际仅能检测到1只眼睛的情况,将左、右眼部纵横比均作为哈欠特征。

3.2.2 嘴部哈欠特征

鉴于人在打哈欠时,嘴部的张合程度会发生明显改变,因此类似于Soukupová提出的眼部纵横比,本文引入一种评估嘴部开合程度的新指标——嘴部纵横比(Mouth Aspect Ratio, MAR)。考虑到不同驾驶员嘴部的厚度差别,根据嘴唇的相关关键点进行特征提取:

$$R_{MA} = \frac{\|p_{14} - p_{18}\| + \|p_{15} - p_{17}\|}{2\|p_{13} - p_{16}\|} \quad (5)$$

式中, R_{MA} 为嘴部纵横比。

至此,已经根据人脸关键点模型提取出眼部和嘴部的哈欠特征,将这些特征组合在一起,形成视频帧哈欠特征向量。

3.2.3 特征矩阵构建

基于驾驶员驾驶状态的视频采集图像,提取每个视频帧 R_{cal} 、 R_{car} 、 R_{MA} 特征值,获得单帧的哈欠特征向量 V_{ec} :

$$V_{ec} = [R_{cali} \ R_{cari} \ R_{MAi}], i \in [1, n] \quad (6)$$

式中, R_{cali} 、 R_{cari} 、 R_{MAi} 分别为第 i 个视频帧的左、右眼部纵横比和嘴部纵横比; n 为视频所含视频帧数量。

若将特征向量按照时间维度进行堆积,得到视频特征矩阵 H_Q :

$$H_Q = \begin{bmatrix} R_{cal1} & R_{car1} & R_{MA1} \\ R_{cal2} & R_{car2} & R_{MA2} \\ R_{cal3} & R_{car3} & R_{MA3} \\ \vdots & \vdots & \vdots \\ R_{caln} & R_{carn} & R_{MA n} \end{bmatrix} \quad (7)$$

由于存在时间信息, H_Q 除了包含视频每帧的左、右眼部纵横比和嘴部纵横比等直接哈欠特征以外,还包含嘴部持续张开时间、短时间内是否存在嘴部张度大幅度改变、单位时间眨眼次数、PERCLOS等隐含哈欠特征,为后续分辨驾驶员是否存在哈欠行为奠定了基础。

4 变种Transformer模型

4.1 Transformer模型

李小平等^[13]通过LSTM对得到的驾驶员疲劳特征矩阵进行分类,从而判断驾驶员是否处于疲劳状态。卢喜东等人^[14]通过深度森林方法对方向梯度直方图(Histogram of Oriented Gradient, HOG)特征矩阵进行恶意代码分类。本文选择变种Transformer模型提取 H_Q 的隐含哈欠特征,从而完成对视频中驾驶员是否打哈欠的分类,因为Transformer通过注意力机制可以捕获远距离依赖关系,并可实现循环神经网络(Recurrent Neural Network, RNN)不能实现的并行训练功能,结构如图5所示。

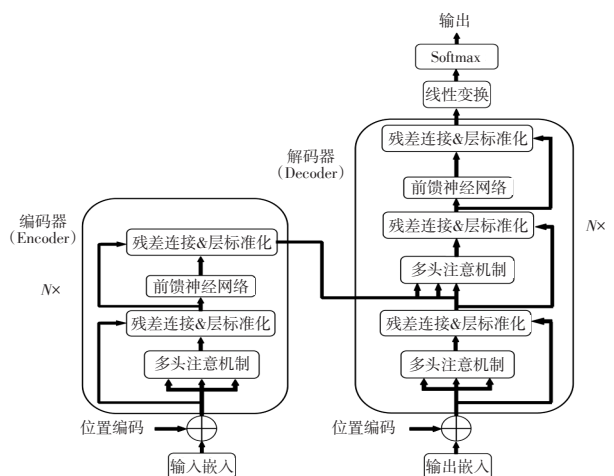


图5 Transformer模型结构

图5中,编码器部分的输入为词向量矩阵,考虑到单词间的顺序,为每个词向量添加了位置编码,计算公式分别为:

$$P_{(s,2i)} = \sin\left(s/10\,000^{2i/d_{\text{model}}}\right) \quad (8)$$

$$P_{(s,2i+1)} = \cos\left(s/10\,000^{2i/d_{\text{model}}}\right) \quad (9)$$

式中, $P_{(s,2i)}$ 、 $P_{(s,2i+1)}$ 分别为句子中第 s 个单词偶数维度/奇数维度的位置编码; d_{model} 为位置编码的维度。

将经过位置编码的词向量矩阵作为自注意力(Self-Attention)层的输入,得到查询矩阵、键矩阵和值矩阵,如图6所示。

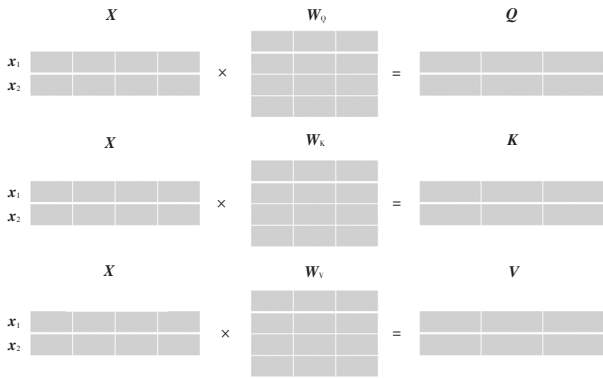


图6 矩阵实现自注意力机制示意

图6中, X 为位置编码后的词向量矩阵; x_1 、 x_2 分别为经过位置编码第1个和第2个单词的词向量; W_q 、 W_k 、 W_v 为训练好权重的线性变换矩阵; Q 、 K 、 V 分别为 X 的查询矩阵、键矩阵和值矩阵。

Q 与 K^T 相乘,得到词向量与其他词向量之间的关联性,为了维护梯度的稳定性,除以 $\sqrt{d_k}$,通过Softmax函数得到该词向量与其他词向量关联性的权重,再与词向量包含的信息 V 相乘,可得:

$$A(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (10)$$

式中, $A(Q, K, V)$ 为自注意力层输出; d_k 为键矩阵 K 的维度。

为了学习到更多独立信息,充分关注不同子空间,引进多头注意力机制^[7],设 h_i 为第 i 个子空间的自注意力层输出,将 m 个子空间的 h_i 进行横向拼接得到多头拼接矩阵,通过线性变换矩阵 W 将多头拼接矩阵变换成与 X 相同形状:

$$h_i = A(XW_q^i, XW_k^i, XW_v^i) \quad (11)$$

$$M = \text{Concat}(h_1, \dots, h_m)W^o \quad (12)$$

式中, XW_q^i 、 XW_k^i 、 XW_v^i 分别为第 i 个子空间的查询矩阵、键矩阵和值矩阵; M 为多头注意力值;Concat为横向

拼接函数。

考虑到网络中可能出现的退化问题,将 X 与 M 进行连接,再进行层标准化,得到前馈神经网络输入 O_f :

$$O_f = \text{LayerNorm}(X + M) \quad (13)$$

式中,LayerNorm为层标准化函数。

将 O_f 通过由线性层、线性整流函数(Rectified Linear Unit, ReLU)层和线性层构成的前馈神经网络进行残差连接和层标准化,结束模型的编码器部分,公式为:

$$O_s = \max(0, O_f W_1 + b_1) W_2 + b_2 \quad (14)$$

$$O_i = \text{LayerNorm}(O_f + O_s) \quad (15)$$

式中, O_s 为前馈神经网络的输出; O_i 为编码器部分的输出; W_1 、 b_1 分别为第1层线性层的参数矩阵和偏置; W_2 、 b_2 分别为第2层线性层的参数矩阵和偏置。

O_i 经过模型的解码器部分,连接一个线性层,通过Softmax进行分类,完成整个Transformer模型的构建。

4.2 Transformer模型的改进

由于Transformer的解码器部分主要完成生成类功能,而疲劳驾驶哈欠检测只需要实现哈欠特征的分类任务即可,故本文提出一种只使用编码器部分的变种Transformer模型。其基本思想是:在继承Transformer模型编码器部分的基础上,对输入嵌入、位置编码以及输出层进行了改进,整体结构如图7所示。

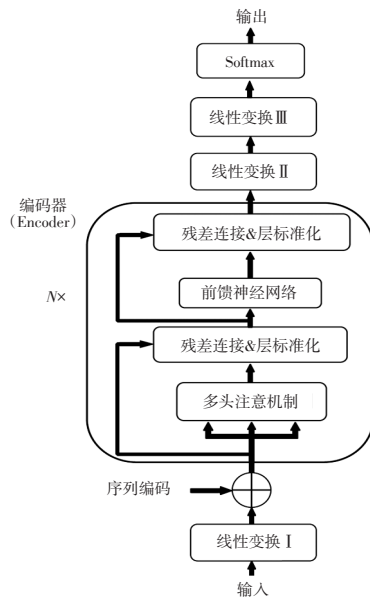


图7 变种Transformer模型结构

图7中,模型的输入不再是词向量矩阵,而是经过Dlib检测算法构建的如式(7)所示的疲劳特征矩阵 H_Q ,如图8所示,设 d_{model}^* 为序列编码的维度,利用 $3 \times d_{\text{model}}^*$ 维的矩阵 W ,得到线性变换I的输出 H_Q^* :

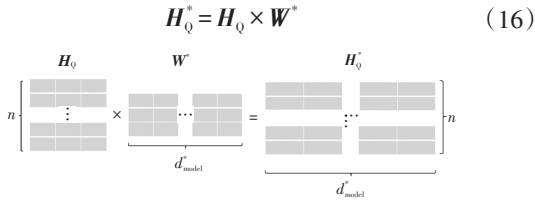
$$H_Q^* = H_Q \times W^* \quad (16)$$


 图8 H_Q 变换示意

为了充分考虑各视频帧之间的时间先后顺序,将 H_Q^* 进行序列编码:

$$S_{(q,2)} = \sin\left(q/10\,000^{2/d_{model}}\right) \quad (17)$$

$$S_{(q,2j+1)} = \cos\left(q/10\,000^{2/d_{model}}\right) \quad (18)$$

式中, $S_{(q,2)}$ 、 $S_{(q,2j+1)}$ 为视频中第 q 帧视频帧偶数维度、奇数维度的序列编码。

H_Q^{**} 与对应序列编码相加后,形成新哈欠特征矩阵 H_Q^{**} ,再通过训练好的矩阵 W_k^* 、 W_v^* 、 W_q^* 进行线性变换,得到新哈欠特征矩阵 H_Q^{**} 的查询矩阵 Q^* 、键矩阵 K^* 和值矩阵 V^* ,计算过程为:

$$Q^* = H_Q^{**} \times W_q^* \quad (19)$$

$$K^* = H_Q^{**} \times W_k^* \quad (20)$$

$$V^* = H_Q^{**} \times W_v^* \quad (21)$$

根据式(10),得到其他视频帧包含的该视频帧信息,通过式(11)和式(12),多个子空间充分挖掘学习视频帧相关信息,之后根据式(13)~式(15),得到变种Transformer模型编码器部分的输出 O_l^* ,通过图9所示的变换,利用 $d_{model} \times 2$ 维的线性变换矩阵 W^α 得到线性变换II的输出 O :

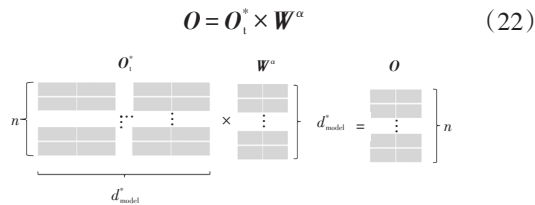
$$O = O_l^* \times W^\alpha \quad (22)$$


 图9 O_l^* 变换示意

如图10所示,利用 $(n \times 1)$ 维的矩阵 W^β 对 O 的转置矩阵 O^T 进行线性变换,得到线性变换III的输出 O_μ :

$$O_\mu = O^T \times W^\beta \quad (23)$$

最终分类结果为:

$$A_{ns} = \operatorname{argmax}\left(\operatorname{Softmax}\left(O_\mu^T\right)\right) \quad (24)$$

式中, A_{ns} 为视频分类结果; O_μ^T 为 O_μ 的转置矩阵。

本文提出的变种Transformer模型充分考虑视频帧之间的关联性,且引入序列编码使得视频帧时间维度得以关联,通过多头注意力机制和多层线性变换,可以实现对哈欠特征矩阵的分类。

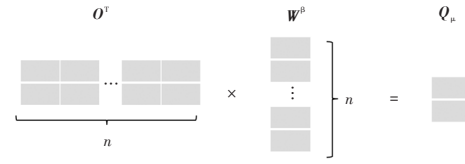
$$O^T \times W^\beta = O_\mu$$


 图10 O^T 变换示意

5 试验验证与分析

为了验证所提出方法的有效性,本文基于YawDD数据集^[15]进行训练和测试,与其他方法进行准确率的对比:

$$A_c = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (25)$$

式中, A_c 为准确率; T_p 、 F_p 分别为被模型预测为哈欠的哈欠视频、其他视频数量; F_n 、 T_n 分别为被模型预测为其他的哈欠视频、其他视频数量。

5.1 数据集构建

YawDD数据集是一个公开的视频数据集,视频分辨率均为 680×480 ,帧率为30帧/s,包含2个子数据集,分别由固定在后视镜下方的摄像头拍摄的视频和由固定在组合仪表上方的摄像头拍摄的视频组成^[15]。考虑到角度问题以及人脸关键点定位精度,本文主要以第2个子数据集为研究对象。该子数据集中的每个视频包含不同的驾驶员,且驾驶员在同一个视频里面包含讲话、大笑、唱歌、打哈欠等行为,不能直接用来进行驾驶员哈欠检测,因此本文采用人工重构数据集的方法对数据集进行处理,得到新的数据集DX-YawDD:即将每个视频中出现的打哈欠视频片段单独裁剪出来,作为哈欠数据集,将每个视频的其他视频片段按照每段视频3~10s进行随机切割,得到新的视频集,作为其他数据集。经过处理后,DX-YawDD数据集中共包含71个哈欠数据集视频,344个其他数据集视频。

5.2 试验环境

本文试验环境配置如下:服务器处理器为Intel Core i7-9750H CPU @ 2.60 GHz,安装内存为8 GB,在Windows10环境下,在Pytorch深度学习框架下搭建了基于Dlib和变种Transformer哈欠检测模型,使用数据集DX-YawDD进行模型训练和性能测试。其中,模型测试集共有125个视频,包括随机选取的22个打哈欠视频和103个其他视频,剩下的49个哈欠视频和241个其他视频组成模型训练集。

5.3 试验结果

基于训练数据集,选择变种Transformer模型的损失函数为交叉熵损失函数,采用随机梯度下降优化器,学

习率为0.001,动量为0.99,得到图11所示的损失函数曲线。

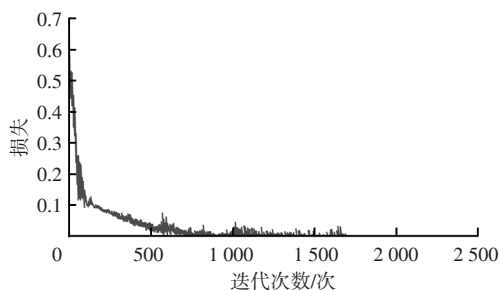


图11 训练损失函数曲线

由图11可知,当训练1700次时,损失函数收敛,此时模型训练的准确率为100%。模型在测试数据集上的哈欠检测结果如表1所示,准确率达到96.8%,召回率为95.5%,表明本方法哈欠漏检率较低。

表1 测试集哈欠检测结果

参数	22个哈欠视频	103个其他视频	全部测试集视频
预测哈欠视频/个	21	3	24
预测其他视频/个	1	100	101
预测准确率/%	95.5	97.1	96.8

将本文方法与其他哈欠检测算法基于DX-YawDD数据集进行对比验证,不同算法的哈欠检测准确率如表2所示。

表2 不同算法哈欠检测准确率

算法	哈欠检测准确率/%
基于双层专家的检测算法 ^[16]	94.0
基于图片输入的CNN算法 ^[3]	92.4
基于三维卷积和双向长短时记忆网络算法 ^[17]	93.6
基于卷积神经网络和优化监督下降方法(Supervised Descent Method,SDM)的算法 ^[18]	87.0
本文算法	96.8

由表2可知,本文算法的检测准确率为96.8%,高于其他检测算法。这是因为本文算法综合考虑了眼部、嘴部的哈欠特征,构建了包含有直接、隐含哈欠特征的视频特征矩阵,利用变种Transformer模型进行特征提取和分类,关注了各视频帧之间的联系,捕捉到远距离视频帧之间的依赖特性,提高算法的准确性。

6 结束语

针对现有驾驶员哈欠检测时,哈欠特征单一、分类阈值无法度量、视频图像中帧与帧之间缺乏联系等问题,本文提出了基于Dlib和变种Transformer算法,通过

公开数据集YawDD进行测试,本文算法的哈欠检测准确率高达96.8%,可以应用于驾驶员疲劳驾驶初期哈欠检测任务,对避免驾驶员疲劳驾驶,保障人员和道路安全具有积极作用。

同时,利用本文算法进行哈欠检测时,虽然多方面考虑了人脸特征点对于哈欠检测的作用,但还存在面部特征挖掘不充分等缺陷,下一步可以通过3D卷积等方法进行挖掘,避免显式的哈欠特征提取,进一步简化算法,提高准确率与鲁棒性。

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] LI L L, CHEN Y Z, LI Z L. Yawning Detection for Monitoring Driver Fatigue Based on Two Cameras[C]// International IEEE Conference on Intelligent Transportation Systems. St. Louis, MO, USA: IEEE Xplore, 2009: 1-6.

[3] 马素刚, 赵琛, 孙韩林, 等. 一种基于卷积神经网络的哈欠检测算法[J]. 计算机科学, 2018, 45(增刊1): 227-229+241.

[4] MA S G, ZHAO C, SUN H L, et al. A Yawn Detection Algorithm Based on Convolutional Neural Network[J]. Computer Science, 2018, 45(S1): 227-229+241.

[5] KNAPIK M, CYGANIEK B. Driver's Fatigue Recognition Based on Yawn Detection in Thermal Images[J]. Neurocomputing, 2019, 338: 274-292.

[6] 史瑞鹏, 钱屹, 蒋丹妮. 一种基于卷积神经网络的疲劳驾驶检测方法[J]. 计算机应用研究, 2020, 37(11): 3481-3486.

SHI R P, QIAN Y, JIANG D N. A Fatigue Driving Detection Method Based on Convolutional Neural Network[J]. Application Research of Computers, 2020, 37(11): 3481-3486.

[7] 王超, 徐楚昕, 王志锋. 面向空中交通管制员疲劳识别的哈欠检测[J/OL]. 安全与环境学报: 1-10[2022-05-14]. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=AQHJ20220309001&uniplatform=NZKPT&v=816eb9my9p0Bu6DEUKHOhjzqycQ_xssEGLzK8T51M05Lz-rc0oOeNNhpgPYZ0Yemr. DOI:10.13637/j.issn.1009-6094.2021.2353.

WANG C, XU C X, WANG Z F. Yawn Detection for Air Traffic Controller Fatigue Recognition[J/OL]. Journal of Safety and Environment: 1-10[2022-05-14]. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=AQHJ20220309001&uniplatform=NZKPT&v=816eb9my9p0Bu6DEUKHOhjzqycQ_xssEGLzK8T51M05Lzrc0oOeNNhpgPYZ0Yemr. DOI:10.13637/j.issn.1009-6094.2021.2353.

- [7] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. New York, USA: ACM Press, 2017: 5998–6008.
- [8] KAZEMI V, SULLIVAN J. One Millisecond Face Alignment with an Ensemble of Regression Trees[C]// Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Press, 2014: 1867–1874.
- [9] 王迪. 基于人眼状态的疲劳检测算法研究与应用[D]. 成都: 电子科技大学, 2020.
WANG D. Research and Application of Fatigue Detection Algorithm Based on Human Eye State[D]. Chengdu: University of Electronic Science and Technology of China, 2020.
- [10] 刘春平, 高翔, 张建伟, 等. 基于人脸关键点的表情实时动态迁移[J]. 现代计算机(专业版), 2019(2): 41–45.
LIU C P, GAO X, ZHANG J W, et al. Real-Time Dynamic Migration of Expressions Based on Facial Key Points[J]. Modern Computer, 2019(2): 41–45.
- [11] SOUKUPOVA' T, ECH J C. Real-Time Eye Blink Detection Using Facial Landmarks[J]. Computer Vision Winter Workshop, 2016: 3–5.
- [12] JOHN S J, SHARMILA S T. Real Time Blink Recognition from Various Head Pose Using Single Eye[J]. Multimedia Tools and Applications, 2018, 77(23): 31331–31345.
- [13] 李小平, 白超. 基于深度学习的司机疲劳驾驶检测方法研究[J]. 铁道学报, 2021, 43(6): 78–87.
- LI X P, BAI C. Research on Driver Fatigue Driving Detection Method Based on Deep Learning[J]. Journal of the China Railway Society, 2021, 43(6): 78–87.
- [14] 卢喜东, 段哲民, 钱叶魁, 等. 一种基于深度森林的恶意代码分类方法[J]. 软件学报, 2020, 31(5): 1454–1464.
LU X D, DUAN Z M, QIAN Y K, et al. A Deep Forest-Based Malicious Code Classification Method[J]. Journal of Software, 2020, 31(5): 1454–1464.
- [15] ABTAHI S, OMIDYEGANEN M, SHIRMOHAMMADI S, et al. YawDD: A Yawning Detection Dataset [C]// Proceedings of the 5th ACE Multimedia Systems Conference. New York, USA: ACM, 2014: 24–28.
- [16] ANITHA C, VENKATESHA M K, ADIGA B S. A Two Fold Expert System for Yawning Detection[J]. Procedia Computer Science, 2016, 92: 63–71.
- [17] YANG H, LIU L, MIN W D, et al. Driver Yawning Detection Based on Subtle Facial Action Recognition[J]. IEEE Transactions on Multimedia, 2020, 23: 572–583.
- [18] LI X F, XIA J H, CAO L B, et al. Driver Fatigue Detection Based on Convolutional Neural Network and Face Alignment for Edge Computing Device[J]. Proceedings of the Institution of Mechanical Engineers, 2021, 235(10/11): 2699–2711.

(责任编辑 斛 畔)

修改稿收到日期为2022年6月30日。

《汽车工程师》征稿启事

《汽车工程师》于1974年创刊,是由中国第一汽车集团有限公司主管、中国第一汽车股份有限公司主办的、国内外公开发行的汽车应用技术类学术期刊,国家新闻出版广电总局第一批认定的学术期刊。

《汽车工程师》以报道汽车产品开发领域的成果为主,专注于技术开发与应用,提供汽车设计创新解决方案。《汽车工程师》将把握电动化、智能化、网联化、共享化的汽车技术主流发展趋势,努力在电池技术、电驱技术、电控技术、高压技术、补能技术、燃料电池技术、智慧控制、智能驾驶、智享座舱、智能悬架、线控转向、线控制动、NVH、功能安全、预期功能安全、信息安全、数据安全、被动安全、高效动力、高效传动、智能管理、低风阻、低滚阻、轻量化领域吸收优质稿源,为我国汽车工程技术创新能力提升贡献力量。

热忱欢迎汽车行业的专家学者不吝赐稿,反映国家重点扶持项目、自然科学基金项目和其他重点项目等研究成果的稿件将优先发表,我们期待与您共同践行“把论文写在祖国大地上”的指示精神,为强大中国汽车工业作贡献!

《汽车工程师》编辑部