



控制工程
Control Engineering of China
ISSN 1671-7848, CN 21-1476/TP

《控制工程》网络首发论文

题目：一种基于卷积神经网络的疲劳驾驶检测方法
作者：吴天昊，张震，王家曜
DOI：10.14107/j.cnki.kzgc.20230854
收稿日期：2023-09-21
网络首发日期：2023-12-21
引用格式：吴天昊，张震，王家曜. 一种基于卷积神经网络的疲劳驾驶检测方法[J/OL]. 控制工程. <https://doi.org/10.14107/j.cnki.kzgc.20230854>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

一种基于卷积神经网络的疲劳驾驶检测方法

吴天昊^{a,b}, 张震^{a,b}, 王家曜^{a,b}

(青岛大学 a. 自动化学院; b. 山东省工业控制重点实验室, 山东 青岛 266071)



摘要: 疲劳驾驶是导致交通事故的主要原因之一, 对驾驶员和道路安全构成严重威胁。为了解决这一问题, 本研究提出一种 Conv-Res 模块, 使用实例标准化代替批量标准化等操作对 XCM(eXplainable Convolutional Neural Network)模型进行改进, 实现了基于脑电信号的疲劳检测方法, 并使用两个权威数据集进行验证, 对疲劳状态分类准确率最高可达 99.6%。实验结果表明, 所提出的改进 XCM 网络分类器可在四种特征处理数据集以及 SPIS 数据集上达到主流时间序列分类器的水平, 并具有明显优势, 平均准确率可达到 97.1%。

关键词: 疲劳驾驶检测; 脑电信号; 卷积神经网络; 深度学习

中图分类号: TP399

文献标识码: A

A novel fatigue driving detection method based on Convolutional Neural Networks

WU Tian-hao^{a,b}, ZHANG Zhen^{a,b}, WANG Jia-yao^{a,b}

(a. College of Automation, Qingdao University; b. Shandong Provincial Key Laboratory of Industrial Control, Qingdao 266071, China)

Abstract: Fatigue driving is one of the major causes leading to traffic accidents, posing a significant threat to both drivers and road safety. In order to address this issue, this study proposes a Conv-Res module, which enhances the eXplainable Convolutional Neural Network (XCM) model by employing instance normalization instead of batch normalization and other operations. This improvement enables the development of a fatigue detection method based on electroencephalogram (EEG) signals. The proposed method was validated using two authoritative datasets, achieving a maximum accuracy of 99.6% in classifying fatigue states. Experimental results demonstrate that the improved XCM network classifier outperforms mainstream time series classifiers on four different feature-processed datasets as well as the SPIS dataset, with a notable advantage and an average accuracy of 97.1%.

Key words: Fatigue driving detection; EEG; Convolutional Neural Network; Deep learning

1 引言

疲劳驾驶是指驾驶者在长时间驾车或睡眠不足的情况下上路行驶的行为, 这一问题在道路安全和公共健康方面带来了严重的威胁。随着现代生活节奏的加快和长时间驾驶活动的增加, 疲劳驾驶的问题日益突出。疲劳驾驶表现为驾驶员的注意力下降、反应迟缓、睡意增加等症状, 严重时可能导致事故发生, 威胁生命安全。目前, 对于疲劳检测的

方法大致分为主观检测和客观检测^[1]。主观检测主要基于驾驶员的自我感知, 依赖于驾驶员自己的感觉和自我评价来判断疲劳水平。客观检测一方面基于生理和行为指标, 例如脸部表情分析、头部姿势、车辆行驶轨迹等, 但这些行为容易受到外界环境的影响^[2]; 另外一方面可以基于传感器与检测设备进行检测, 例如通过检测驾驶员的眼睛运动或是心电

收稿日期: 2023-09-21;

基金项目: 国家自然科学基金资助项目(61903209)

作者简介: 吴天昊(2000-), 男, 安徽黄山人, 硕士研究生, 主要研究方向为深度学习、脑机接口; 张震(1985-), 男, 山东淄博人, 副教授, 博士, 主要从事强化学习、多智能体系统等方面的教学与科研工作(本文通信作者, Email:527621327@qq.com); 王家曜(1999-), 男, 山东淄博人, 硕士研究生, 主要研究方向为深度学习、脑机接口。

信号。这种方法更加客观和准确,不受驾驶员主观意识的影响,因此利用生理信号进行疲劳检测是十分可靠的。通过收集可穿戴设备采集驾驶员相应部位的生理信号,例如脑电信号(Electroencephalogram, EEG^[3])、心电信号(Electrocardiogram, EMG^[4])、眼电信号(Electroculogram, EOG^[5])等,基于深度学习算法,设计模型判断驾驶员是否处于疲劳状态。脑电信号反应大脑的活动,异常信号在疲劳发生的前20秒就会被记录在脑电信号中,故此本文基于脑电信号,利用深度学习方法进行疲劳驾驶检测。

研究表明,人们在进入疲劳状态后, α 和 θ 等低频段脑电波信号活动会增加^[6],文献^[7]提出了一种基于EEG的心理状态分类模型进行疲劳驾驶检测。Wang等人^[8]运用全卷积网络(Fully Convolutional Neural Network, FCN)和循环神经网络(Recurrent Neural Network, RNN)对单变量时间序列进行分类。文献^[9]设计了一种基于整流线性单元层的改进双向长短期记忆深度神经网络进行疲劳驾驶检测。Chinara^[10]等人使用时-频转换法,对脑电信号的时域特征进行分析、训练。文献^[11]提出了一种基于慢速眼动的驾驶员睡意检测方法。Du等人^[12]使用了一种结合脑电信号和眼电信号的多模态融合方法进行疲劳驾驶检测。

卷积神经网络^[13]被广泛应用于机器视觉领域,并在图像分类、人脸识别等任务上取得了广泛的应用。Bai等人于2018年提出了TCN(Temporal Convolutional Network, 时间卷积网络^[14]),将CNN应用于时间序列领域。XCM网络(eXplainable Convolutional Neural Network^[15])提出了一种紧凑式的卷积神经网络,在UEA数据集上取得了较好的时间序列分类效果。但卷积神经网络在多维时间序列分类,特别是基于脑电信号的疲劳检测研究仍然较少。本文针对疲劳驾驶检测任务提出了一种改进的基于XCM网络的方法。该方法通过在XCM网络中引入两种卷积模块,对模型进行优化,旨在有效检测疲劳状态。我们使用了两个权威的脑电数据集进行性能验证,以验证该方法的有效性。

2 改进后的 XCM 网络模型

2.1 XCM 网络

可解释卷积神经网络(XCM)是 Fauvel K, Lin 等人提出的一种紧凑式卷积神经网络,它直接从输入数据中提取与观测变量和时间相关的信息,并且在公开时间序列数据集 Time Series Classification and Clustering - The UEA & UCR Time Series

Classification Repository(简称为 UEA)^[16]上展现了较为先进的 MTS(Multivariate Time Series, 多维时间序列)分类性能。XCM 提出了一种思想:采用不同种类的卷积核对时间序列特征分别进行提取:对于时间序列的观测变量特征,XCM 网络采用 2D 卷积核进行特征提取;对于时间序列的时间特征,采用 1D 卷积核进行特征提取,最后再将提取的特征图进行并行合并。首先,XCM 利用 2D 卷积滤波器(如图 1 中绿色部分所示)提取观测变量的相关信息。该部分包含一个 2D 卷积块,通过这个块将输入转换成特征图,并使用 1×1 的卷积滤波器来减少参数数量。具体来说,这个卷积块由一个 2D 卷积层,随后的批量归一化层^[17]和 ReLU 激活层^[18]组成。本文将 2D 卷积滤波器的卷积核大小设置为(窗口大小 $\times 1$),其中窗口大小是一个超参数,用于指定期望从多变量时间序列(MTS)子序列中提取特征的大小,以及每个观测变量的均值。因此,这些 2D 卷积滤波器(图 1 中的参数:F,默认值 128)允许提取每个观察变量的特征,使用滑动窗口(Strides=1)提取特征。窗口大小作为超参数,用于调节卷积核的尺寸。

与此同时,XCM 利用 2D 卷积滤波器(如图 1 中蓝色部分所示)提取相对于时间的信息。这一部分与右侧部分相似,唯一的区别在于 2D 卷积滤波器被 1D 卷积滤波器所替代。其次,我们将 1D 卷积滤波器的卷积核大小设为(窗口大小 $\times D$),其中窗口大小与 2D 卷积滤波器的超参数相同, D 为输入数据的观测变量个数。1D 卷积滤波器仅在时间轴上滑动(Strides=1),以捕捉不同时间序列之间的相互作用。在接下来的步骤中,将左右两个部分的输出特征图串联起来,形成一个新的特征图,然后应用相同的 1D 卷积块,沿时间轴滑动,以捕捉提取的特征之间的交互信息。最后,XCM 网络在输出特征图上添加一个 1D 全局平均池化层,并使用 Softmax 层进行分类。

然而,XCM 网络在 UEA 数据集中的某些大型多维时间序列分类数据集上表现不佳,例如在 Ethanol Concentration 和 Hand Movement Direction 数据集上仅有 33.6%、44.6%的准确率。因此,本文针对 XCM 网络进行了优化,增加了 Conv2D-Res 模块、Conv1D 模块和全连接层模块(如图 1 中虚线框所示)。虽然增加了参数量(从 302,215 增至 514,454,提升幅度为 70.2%),但本文进一步提升了模型对时间和空间特征的提取能力。具体的模块设计将在第 2.2 节中详细介绍。另外,本文使用了 Instance Normalization(实例标准化)模块代替了原文中使用的批量标准化模块,具体的模块设计将在 2.3

节进行介绍。该设计可以增强模型提取特征的能力。此外，添加了全连接层模块，并在 Softmax 层进行分类前使用了 L2 正则化以及 Dropout^[19]防止模型过拟合。经过改进，最终使其能够在大型和小型的疲劳检测分类数据集上都能取得较好的效果。改进后的 XCM 模型如图 1 所示。

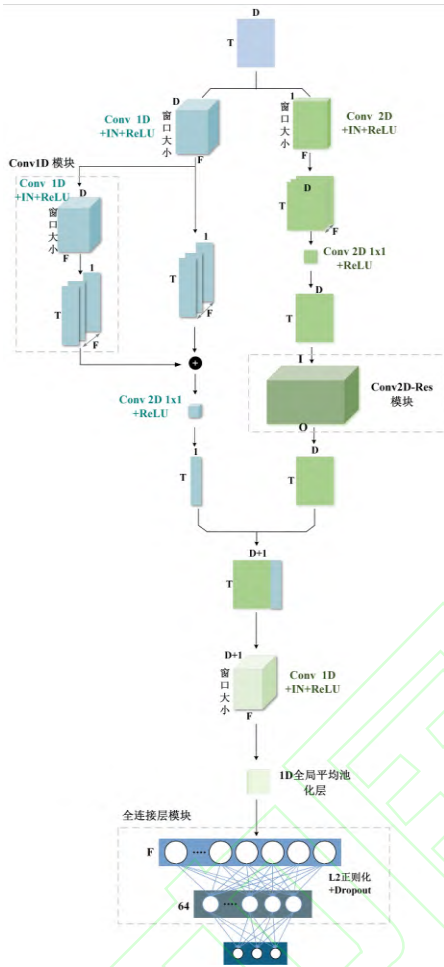


图 1 改进后的 XCM 模型
Fig. 1 Improved XCM model

图 1 中，本文所使用的脑电数据为三维张量，格式为(数据量，通道数，时间步长)。其中 D 为观测变量，即通道数； T 为时间步长。经过左边和右边模块中的多层卷积滤波器处理后，分别生成不同的特征图，最后将二者并行叠加生成最终形状为 $(T, D+1)$ 的特征图，随后经过了一层 1×1 卷积滤波器和 1D 全局平均池化层后进入全连接层，最后经过 softmax 函数输出分类。

2.2 Conv2D-Res 模块

本文主要设计了带有残差连接的卷积模块 Conv2D-Res 提高模型的性能。受到 Inception 网络^[20]的启发，我们采用了紧凑式的卷积模块设计，大大减少了网络参数的计算量，同时使网络的性能提升。

我们首先使用了 1×1 卷积核，从而减少计算复杂性，同时保留关键特征，控制网络的参数量和计算量。此外， 1×1 卷积可以改变特征映射，以获得更丰富的特征表示，有助于网络学习更复杂的特征。同时其在不同位置上共享参数，降低过拟合风险。

随后，我们采用了两个优化的卷积模块，它们的卷积核大小分别为 $(1 \times \text{window_size})$ 和 $(\text{window_size} \times 1)$ ， window_size 表示窗口大小，用以控制卷积核的尺寸，进而分析不同时间子序列的特征信息。其中， T 表示数据集的时间步长， D 表示观测变量的个数。这种设置能够构建多层次的特征表示，有助于模型更好地理解序列信息的细节和层次结构。

另外，在这些模块中引入了残差连接，旨在加速模型的收敛，并在一定程度上避免过拟合的发生。同时，我们在模型的左侧通道增加了 Conv1D 模块，类似于 Conv2D-Res 的操作，该模块能够增强模型对时间特征的提取能力。最后我们将经过处理的特征图与 Conv2D-Res 模块处理后的特征图进行 element-wise 相加(即两个同维度的向量/矩阵中的每个元素相加)后输出。我们将在第 4.5 节的消融实验中详细介绍 Conv2D-Res 和 Conv1D 模块的有效性证明。Conv2D-Res 模块的具体设计如图 2 所示。

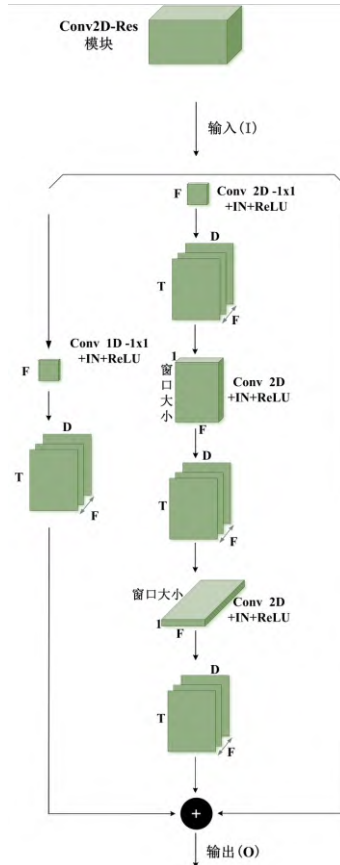


图 2 Conv2D-Res 模块
Fig.2 Conv2D-Res Module

2.3 实例标准化模块

实例标准化(Instance Normalization, IN^[21])模块是一种对于单个样本进行标准化的模块。XCM 原模型使用了批量归一化(Batch Normalization, BN)结构,其结构被广泛使用于深度学习中,有助于防止模型过拟合,并能结合样本之间的信息,但当训练数据集非常小或者小批量数据的大小较小时,BN的效果通常不如在大数据集上明显,并可能导致模型的过拟合。在多维时间序列分类问题中,IN有助于模型学习样本之间和通道之间的关系,特别是当每个通道代表不同的特征或传感器时,IN可以将每个通道的数据归一化,消除通道之间的偏差,从而有助于模型更好地理解每个通道的贡献。假定输入为 $[x_1, x_2, x_3, \dots, x_n]$,对于每个独立的特征 x_i , ($i \in n$) 计算其均值 μ 和方差 σ ,最后对每个特征进行标准化。对于多维时间序列,具体如公式(1)(2)所示:

$$\mu_{i,c} = \frac{1}{T} \sum_{t=1}^T x_{i,c,t} \quad (1)$$

$$\sigma_{i,c} = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_{i,c,t} - \mu_{i,c})^2} \quad (2)$$

其中 c 表示通道数, T 为时间步长。

区别于BN对于批量大小比较敏感,IN不仅可以加速模型收敛,并且可以保持每个样本之间的独立和训练过程的稳定,同时避免梯度爆炸或者消失。

2.4 优化器和损失函数

本文模型选用Adam(Adaptive Moment Estimation^[22])优化器,它是一种基于梯度的自适应优化算法,可以根据每个参数的历史梯度信息和平均值自适应地调整学习率。其基本思想是利用梯度的一阶矩和二阶矩来调整每个参数的学习率(learning rate, lr),从而使模型训练更加高效和准确。本文选择了多分类交叉熵损失函数(Categorical Cross Entropy Loss)对训练时的loss值进行计算。由于本文使用的脑电信号疲劳检测方法属于三分类问题,采用Softmax函数输出,交叉熵主要描述的是实际输出概率与期望输出概率之间的距离,交叉熵越小,两个概率分布越接近。损失函数计算公式如公式(3)所示:

$$L = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (3)$$

$M=3$ 为标签类别数, y_{ic} 为符号函数,样本 i 所对应的类别为 $c(c=1,2,\dots,M)$; p_{ic} 为观测样本 i 属于类别 c 的预测概率。

3 数据集

本文使用的数据集是来自上海交通大学吕宝粮教授团队收集的警觉度估计前额眼电和脑电公开数据集SEED-VIG^[23],以及SPIS静息状态数据集

^[24]。其中SEED-VIG数据集筛选收集了21个受试者的脑电数据。数据中包含以2Hz带宽计算的50Hz的脑电特征,含17个脑电通道,25个频带,885个样本以及4种脑电信号处理后的特征,分别为:功率谱密度-滑动平均特征、功率谱密度-线性动力系统特征、微分熵-滑动平均特征和微分熵-线性动力系统特征,作为四个脑电特征子数据集使用。64脑电电极分布位置图如图3所示。

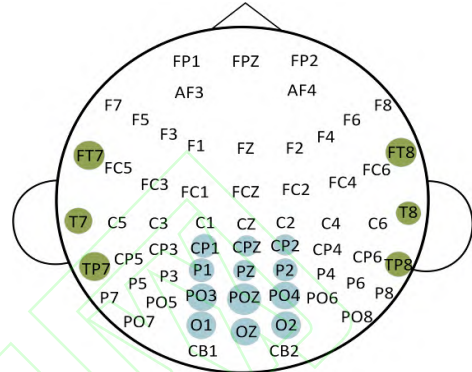


图3 64导联脑电电极图
Fig. 3 64-lead electroencephalogram

64电极脑电分布图是一种用于记录和分析脑电信号的电生理学工具。图中包括64个电极位置,这些电极分布在头皮上,按照国际10-20系统的命名方式进行标记,以捕获大脑活动。其中PZ(Parietal Midline),OZ(Occipital Midline)和POZ等,是其中一些电极名称。例如POZ(Parieto-Occipital Midline)电极位于头皮上的顶部和后部交界处,正好位于PZ和OZ之间。它用于监测顶部和后部的中线大脑活动,同时涵盖了与警觉度、视觉处理和感知相关的多个脑区。在疲劳检测中,POZ电极有助于捕捉整体的大脑状态,包括警觉度下降和认知功能受损等方面的变化。特征提取的17个通道如图5所示(除CPZ为参考电极),将每名被试者在实验过程中采集的 $perclos$ 值作为警觉性标签($lable$),值介于(0, 1)之间(后文中用 p 表示),意为眨眼($blink$)、闭眼($CLOS$)的时间在眨眼、注视($fixation$)、扫视($saccade$)、闭眼时间之和($interval$)中的占比,值越小表示警惕度越高,即被试驾驶时更清醒。Trejo和Chang等人认为有必要将 p 值划分为三个不同的疲劳等级,分别为清醒、疲劳、嗜睡,具体的计算公式如公式(4)(5)所示:

$$PERCLOS = \frac{blink + CLOS}{interval} \quad (4)$$

$$interval = blink + fixation + saccade + CLOS \quad (5)$$

考虑到 p 值大小与被测疲劳程度呈正相关,对于疲劳驾驶检测任务,Trejo^[25]等人认为应当分别以

0.35, 0.7 为阈值, 将其划分为三种不同的疲劳状态, [0-0.35]为清醒状态, [0.35-0.7]为疲劳状态, [0.7-1]为嗜睡状态, 并将其转化为新的标签, 分别以 1, 2, 3 表示。具体的疲劳状态特征见表 1, SEED-VIG 数据集中各疲劳状态的标签分布情况见表 2。

表 1 疲劳特征评测表

Tab.1 Fatigue characteristics evaluation table	
疲劳状态	状态特征
清醒	眼球状态活跃, 眨眼时间短, 注意力集中, 反应灵敏, 可以轻松应对突发事件。
疲劳	眼球活动频率下降, 眨眼时间较长, 出现打哈欠注意力不集中等现象, 反应缓慢, 对突发事件来不及做出反应。
嗜睡	出现睡意, 视觉模糊, 头晕等症状, 眼皮长时间闭合, 思维迟钝。

表 2 不同标签在 SEED-VIG 数据集中的数量

Tab. 2 The number of different tags in the SEED-VIG dataset				
	微分熵-线性动力系统	微分熵-滑动平均特征	功率谱密度-线性动力系统	功率谱密度-滑动平均特征
清醒	5991(1441)	5885(1517)	5937(1465)	5927(1475)
疲劳	7130(1774)	7179(1725)	7103(1801)	7189(1715)
嗜睡	3234(815)	3291(758)	3315(734)	3239(810)

SPIS 数据集采集了 10 名受试者在睁眼和闭眼状态下的脑电信号, 每种状态记录 150 秒。类似于 SEED-VIG 数据集中的处理, 根据睁眼和闭眼状态, 将 SPIS 数据集中的标签分成为 2 类, 以此判断被试者的疲劳状态, 具体的疲劳状态特征分布见表 3。

表 3 不同标签在 SPIS 数据集中的数量

Tab.3 The number of different tags in the SPIS dataset	
	SPIS 静息状态
睁眼状态	4756(1211)

表 4 各模型在微分熵-线性动力系统特征数据集上的准确率

Tab.4 Accuracy of each model on the differential entropy-linear dynamic system characteristic dataset								
	FCN	ResNet	CNN	MLP	Twiesn	Inception	T-LeNet	XCM
准确率	94.3	95.3	93.2	89.1	77.4	95.8	76.6	93.4

表 5 各模型在微分熵-滑动平均特征数据集上的准确率

Tab.5 Accuracy of each model on differential entropy-moving average feature data set								
	FCN	ResNet	CNN	MLP	Twiesn	Inception	T-LeNet	XCM
准确率	97.6	98.1	86.9	89.1	75.3	97.8	76.3	95.6

表 6 各模型在功率谱密度-线性动力系统特征数据集上的准确率

Tab.6 Accuracy of each model on power spectral density-linear dynamic system feature dataset								
	FCN	ResNet	CNN	MLP	Twiesn	Inception	T-LeNet	XCM
准确率	67.8	69.1	63.1	68.7	71.3	61.5	41.3	53.3

闭眼状态

4844(1189)

4 实验

4.1 实验数据和平台

为了验证本文提出的模型性能, 同时对上述两个疲劳驾驶检测数据集上进行实验。采用 Scikit-learn 库中的分层交叉验证方法进行 K 折交叉验证 (本文实验取 $K=5$), 即 5 折交叉验证。为了检测该网络的泛化性能以及鲁棒性, 将交叉验证后得到的最优超参数模型应用到整个数据集上进行训练, 获得平均准确率。

本文方法和实验的详细信息如下所示:

Windows11 系统, CPU:12th Gen Intel(R) Core(TM) i7-12700KF 3.6GHz, GPU:NVIDIA GeForce RTX 3080。编程语言为 python3.8, 深度学习框架为 Tensorflow2.7.0。

4.2 基于 SEED-VIG 数据集的对比实验

首先在 SEED-VIG 数据集上进行了对比实验, 进行对比试验的网络有: 全卷积网络(FCN)、ResNet、传统 CNN 网络, 多层感知机 MLP、Time Le-Net(T-LeNet)、时间扭曲不变回声状态网络(Time warping invariant echo state network, Twiesn^[26])、Inception 网络、XCM 网络。初始学习率 $lr=0.001$, 批量大小设置为 64, 利用 Keras 中的 ReduceLROnPlateau 模块进行学习率控制: 设置为若 5 个 epoch 后验证集 loss 不变, 则以衰减因子 $factor=0.5$ 衰减更新学习率, 如公式 6 所示:

$$lr^* = lr \times factor \quad (6)$$

公式 6 中, lr^* 为更新后的学习率, lr 为更新前的学习率。对于 XCM 网络和本文网络, 窗口大小均设置为 0.8, 全部模型在 epoch=200 的参数条件下进行交叉验证。各个模型在对应的四个特征取得的准确率(单位%)见表 4—表 7。

本文模型

96.5

表 7 各模型在功率谱密度-滑动平均特征数据集上的准确率
Tab.7 Accuracy of each model on power spectral density-moving average feature dataset

	FCN	ResNet	CNN	MLP	Twiesn	Inception	T-LeNet	XCM	本文模型
准确率	53.1	58.3	72.4	58.8	69.6	68.2	50.4	60.3	96.2

表 8 各模型在 SPIS 数据集上的准确率
Tab.8 Accuracy of each model on SPIS dataset

	FCN	ResNet	CNN	MLP	Twiesn	Inception	T-LeNet	XCM	本文模型
准确率	63.4	96.3	53.3	23.9	96.7	93.4	62.4	94.3	99.6

表 9 消融实验
Tab.9 Ablation experiment

	Module1	Module2	Module3	Module4	本文模型
微分熵-线性动力系统	95.8 (± 0.1)	96.7 (± 0.1)	95.9 (± 0.1)	96.7 (± 0.1)	96.5 (± 0.1)
微分熵-滑动平均特征	96.8 (± 0.1)	96.3 (± 0.1)	97.2 (± 0.1)	97.6 (± 0.1)	97.9 (± 0.1)
功率谱密度-线性动力系统	95.8 (± 0.1)	95.9 (± 0.1)	43.5 (± 0.1)	96.0 (± 0.1)	96.6 (± 0.1)
功率谱密度-滑动平均特征	95.3 (± 0.1)	95.0 (± 0.1)	50.9 (± 0.1)	96.3 (± 0.1)	96.2 (± 0.1)
SPIS	98.8 (± 0.1)	95.3 (± 0.1)	99.7 (± 0.1)	99.7 (± 0.1)	99.6 (± 0.1)
平均准确率	96.5 (± 0.1)	95.8 (± 0.1)	77.4 (± 0.1)	97.3 (± 0.1)	97.4 (± 0.1)

4.3 基于 SPIS 数据集的对比实验

其次在 SPIS 数据集上进行对比实验, 仍然选择了与 SEED-VIG 实验中相同的对比实验数据集, 设置 epoch=200 进行对比试验, 得到的平均准确率见表 8。

4.4 结果分析

根据表 4 到 7 的数据, 我们可以得出以下结论: 在微分熵特征的分类表现上, 除了 T-LeNet 和 Twiesn 之外, 各个模型均实现了 90% 以上的准确率。具体而言, 在微分熵-滑动平均特征方面, ResNet 取得了最高的准确率, 本文提出的模型仅略低于它 0.2% 的性能, 达到了 97.9% 的准确率。对于微分熵-线性动力平均特征, 本文的模型实现了 96.5% 的准确率。然而, 在功率谱特征上, 基准模型的表现都不尽如人意, 只有经过改进的 XCM 模型能够达到 96% 以上的准确率。综合表 8 的数据, 可以看到在 SPIS 数据集上, ResNet、T-LeNet 和 Twiesn 表现出色, 但本文提出的模型取得了最高的准确率, 达到了 99.6%。总体来看, 本文模型取得了与这几个基准模型相当的准确率, 而在功率谱密度-滑动平均特征数据以及功率谱密度-线性动力系统数据集上, 只有本文模型达到了 96% 以上的准确率, 并且本文模型在四个数据集上的平均表现最优。经过改进, 本文模型在功率谱密度的准确率相较于原始 XCM 模型, 提高了近 40%。

4.5 消融实验

为了评估本文模型中各模块的性能增益, 本节

进行了完全的消融实验。实验分别对四个特征数据集以及 SPIS 数据集进行准确率测试, 并统计平均准确率。其中 1, 2, 3, 4 号模型分别为去掉 Conv2D-Res 模块、去掉 Conv1D 模块、去掉实例标准化模块、去掉全连接层, Dropout 以及 L₂ 正则化后的本文模型。实验结果所示准确率(± 0.1 表示准确率误差在正负 0 到 0.1 之间, 单位%)如表 9 所示。

由 1, 2 号模型可知, Conv2D-Res 模块与 Conv1D 模块对于模型识别准确率都具有提升作用。而从微分熵-线性动力系统特征和功率谱密度-滑动平均特征的准确率来看, Conv2D-Res 模块对于这两个特征的提取能力要优于 Conv1D 模块。由 2, 3, 4 号模型可知, 对于 SPIS 数据集来说, Conv1D 模块对此特征的影响更大, 而对实例标准化模块 Dropout、L₂ 正则化操作不敏感, 但模型准确率并没有因为该操作明显衰减, 仍能达到 99.6% 的准确率。

由 3 号模型可知, 实例标准化模块对于功率谱密度特征的提取十分重要。功率谱密度特征通常由时间序列数据产生, 如脑电数据集, 每个时间戳或实例代表一个瞬时状态。不同的实例可能具有不同的统计特性, 因为它们反映了时间序列中的不同情况。传统的批量标准化倾向于对整个批次的数据进行标准化, 会导致实例间的特性差异被平滑化, 可能不适用于功率谱密度特征。实例标准化通过对每个实例进行独立标准化, 有助于保持实例间的差异性, 从而更好地捕获时间序列中的特征。实验结果表明, 本文引入模型和改进的部分在 SEED-VIG 数据集以及 SPIS 数据集的分类准确率上取得了较为

明显的提升。

5 结 论

本文采用改进后的 XCM 网络, 在疲劳驾驶检测的两个脑电数据集上进行实验, 取得了较为理想的结果。实验表明, 本模型已经取得了与当前时序分类模型相媲美的性能与结果, 同时进行了完全的消融实验, 证明添加的 Conv-Res 模块与 IN 模块等对于分类效果有较大的提升, 但在模型计算效率以及分类准确率方面仍有提升空间。在未来的研究中会更加注重不同生理特征对于疲劳状态的影响, 采用不同的分类器进行疲劳检测的研究, 同时探究类脑神经网络的研究, 例如脉冲神经网络等, 提高网络计算效率, 疲劳驾驶检测效率以及准确率。

参考文献(References)

- [1] 张瑞, 朱天军, 邹志亮, 等. 驾驶员疲劳驾驶检测方法研究综述[J]. 计算机工程与应用, 2022, 58(21):53-66.
- [2] ZHANG R, ZHU T J, ZOU Z L, et al. Review of driver fatigue driving detection methods[J]. Computer Engineering and Applications, 2022, 58(21): 53-66.
- [3] 李照, 舒志兵. 基于 DSP 的驾驶员疲劳实时预警系统设计[J]. 控制工程, 2019,26(01):92-98.
- [4] LI Z, SHU Z B. Design of real-time warning system for driver fatigue based on DSP[J]. Control Engineering, 2019, 26(01): 92-98.
- [5] WANG H, DRAGOMIR A, ABBASI NI, et al. A Novel Real-time Driving Fatigue Detection System Based on Wireless Dry EEG[J]. Cognitive Neurodynamics, 2018, 12(4):365-376.
- [6] GROMER M, SALB D, WALZER T, et al. ECG sensor for detection of driver's drowsiness[J]. Procedia Computer Science, 2019, 159: 1938-1946.
- [7] 管凯捷, 姚康, 任谊文等. 基于头动与眼电信号的疲劳检测研究[J]. 计算机应用与软件, 2022, 39(02): 81-87.
- [8] GUAN K J, YAO K, REN Y W, et al. Research on fatigue detection based on head movement and eye elscrtical signal[J]. Computer Applications and Software, 2022, 39(02): 81-87.
- [9] MAHMOODI M, NAHVI A. Driver drowsiness detection based on classification of surface electromyography features in a driving simulator[J]. Proceedings of the Institution of Mechanical Engineers.Part H, Journal of Engineering in Medicine, 2019, 233(4): 395-406.
- [10] ZENG H, YANG C, DAI G, et al. EEG classification of driver mental states by deep learning[J]. Cognitive neurodynamics, 2018, 12: 597-606.
- [11] WANG Z, YAN W, OATES T. Time series classification from scratch with deep neural networks: A strong baseline[C]//2017 International joint conference on neural networks (IJCNN). IEEE, 2017: 1578-1585.
- [12] ANSARI S, NAGHDY F, DU H, et al. Driver mental fatigue detection based on head posture using new modified reLU-BiLSTM deep neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(8): 10957-10969.
- [13] CHINARA S. Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal[J]. Journal of neuroscience methods, 2021, 347: 108927.
- [14] 焦影影. 基于脑电和眼电信号的驾驶员睡意检测研究[D]. 上海: 上海交通大学, 2019: 101-103.
- [15] JIAO Y Y, Driver Sleepiness Detection Based on EEG and EOG Signals[D]. Shanghai: Shanghai Jiao Tong University, 2019:101-103.
- [16] DU L H, LIU W, ZHENG W L, et al. Detecting driving fatigue with multimodal deep learning[C]//2017 8th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2017: 74-77.
- [17] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series[J]. The handbook of brain theory and neural networks, 1995, 3361(10): 1995.
- [18] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv preprint arXiv:1803.01271, 2018.
- [19] FAUVEL K, LIN T, MASSON V, et al. Xcm: An explainable convolutional neural network for multivariate time series classification[J]. Mathematics, 2021, 9(23): 31-37.
- [20] BAGNALL A, DAU H A, LINES J, et al. The UEA multivariate time series classification archive, 2018[J]. arXiv preprint arXiv: 1811.00075, 2018.
- [21] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. International conference on machine learning. pmlr, 2015: 448-456.
- [22] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]//Proceedings of the 27th international conference on machine learning (ICML-10). 2010: 807-814.
- [23] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [24] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [25] ULYANOV D, VEDALDI A, LEMPITSKY V. Instance normalization: The missing ingredient for fast stylization[J]. arXiv preprint arXiv:1607.08022, 2016.
- [26] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [27] ZHENG W L, LU B L. A multimodel approach to estimating vigilance using EEG and forehead EOG[J]. Journal of Neural Engineering, 2017, 14(2): 026017.1-026017.14.
- [28] TORKAMANIAZAR M, KANIK SD, AYDIN S, et al. Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state EEG in a long sustained attention task[J]. IEEE J Biomed Health Inform. 2020, 24(9): 2550-2558.
- [29] TREJO L J, KNUTH K, PRADO R, et al. EEG-based estimation of mental fatigue: Convergent evidence for a three-state model[C]//Foundations of Augmented Cognition: Third International Conference, FAC 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007. Proceedings 3. Springer Berlin Heidelberg, 2007: 201-211.
- [30] TANISARO P, HEIDEMANN G. Time series classification using time warping invariant echo state networks[C]//2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2016: 831-836.