

基于2D-CNN和Cox-Stuart早停机制的癫痫预测模型

张喜珍^{1,2}, 张晓莉^{1,2}, 吕洋³, 陈扶明²

1. 甘肃中医药大学医学信息工程学院, 甘肃 兰州 730050; 2. 中国人民解放军联勤保障部队第940医院医疗保障中心, 甘肃 兰州 730050; 3. 中国人民解放军联勤保障部队第940医院眼科, 甘肃 兰州 730050

【摘要】针对如何有效预测癫痫患者是否将要发病这一问题,提出一种基于非独立患者的2维卷积神经网络(2D-CNN)和Cox-Stuart检验法的癫痫预测模型方法。首先对脑电数据做归一化处理,使用陷波滤波器和高通滤波器滤除脑电信号的噪声;将滤波后的信号输入到2D-CNN模型中进行特征提取和分类,使用Cox-Stuart方法检测是否需要早停,从而降低模型的计算复杂度和时间复杂度。此外,分别在发作前期为10、30、60 min的情况下对模型进行测试,结果显示,发作前期为10 min时,模型的效果最优。在测试集上的准确率为97.70%,灵敏度为97.36%,特异性为98.04%,具有良好的性能。

【关键词】癫痫;预测;Cox-Stuart检验法;2D-CNN;深度学习

【中图分类号】R318;TP912.35

【文献标志码】A

【文章编号】1005-202X(2025)01-0082-13

Epilepsy prediction model based on 2D-CNN and Cox-Stuart early stopping mechanism

ZHANG Xizhen^{1,2}, ZHANG Xiaoli^{1,2}, LÜ Yang³, CHEN Fuming²

1. School of Medical Information Engineering, Gansu University of Chinese Medicine, Lanzhou 730050, China; 2. Medical Security Center, The 940th Hospital of Joint Logistics Support Force of Chinese People's Liberation Army, Lanzhou 730050, China; 3. Department of Ophthalmology, The 940th Hospital of Joint Logistics Support Force of Chinese People's Liberation Army, Lanzhou 730050, China

Abstract: An epilepsy prediction model based on two-dimensional convolutional neural network and Cox-Stuart test for non-independent patients is proposed to address the problem of how to effectively predict whether epilepsy patients are going to have an attack or not. After EEG data normalization and EEG signal noise removal using a notch filter and a high-pass filter, the filtered signals are inputted into the two-dimensional convolutional neural network model for feature extraction and classification, and Cox-Stuart test is used to determine whether an early stopping is needed or not, thereby reducing the computational and time complexities of the model. The model is tested under the conditions with pre-seizure periods of 10, 30 and 60 min, respectively, and the results show that the model performs best when the pre-seizure period is 10 min. The model has an average accuracy, sensitivity and specificity of 97.70%, 97.36% and 98.04% on the test set, demonstrating its superior performance.

Keywords: epilepsy; prediction; Cox-Stuart test; two-dimensional convolutional neural network; deep learning

前言

癫痫是一种脑部慢性疾病,它的发作由大脑神经元突发性异常放电引起,发作病因复杂。《国际抗

癫痫联盟指南》总结以下几种病因:遗传性结构性、感染性、结构性、免疫性、代谢性和未知性病因^[1]。癫痫的临床表现有:意识丧失、肌肉抽搐、呼吸困难等。研究表明癫痫的反复发作经常伴有焦虑、抑郁等症状^[2]。目前,全世界约有5000万癫痫患者,且不限年龄。虽然已经提出多种癫痫治疗手段,但仍有约30%的患者会复发^[3]。癫痫的发病具有突发性和反复性,给患者及其家属带来身心上的巨大困扰^[4]。虽然可以根据癫痫日记预测癫痫发作,但准确率不到50%^[5]。因此,有效的癫痫预测手段具有重大的意义。

【收稿日期】2024-09-11

【基金项目】国家自然科学基金(61901515,82000926);甘肃省自然科学基金(22JR5RA002)

【作者简介】张喜珍,研究方向:生物医学信号检测与处理,E-mail:2252403239@qq.com

【通信作者】陈扶明,高级工程师,研究方向:生物医学信号检测与处理,E-mail:cfm5762@126.com

近年来,随着神经网络的不断发展,研究人员开始使用深度学习方法尝试对癫痫脑电信号进行分类。与机器学习相比,深度学习可以处理大数据、提取数据的高维特征,以此提高模型的分类效果。Pan等^[6]将原始脑电(Electroencephalography, EEG)数据、经过快速傅里叶变换、短时傅里叶变换、离散小波变换(Discrete Wavelet Transform, DWT)处理的EEG数据作为卷积神经网络(Convolutional Neural Network, CNN)的全部输入,并使用特征融合机制整合学习到的特征,提出的方法可达到99%以上的精度,但是该实验并没有设置测试集数据,实验结果均来自交叉验证,无法证明模型泛化能力的好坏;Takahashi等^[7]在对原始数据进行高通、低通、陷波滤波之后,利用自动编码器(Autoencoder, AE)将发作间期中连续10 s以上的高AE误差定义为非癫痫发作但不正常的数数据,然后将经过AE得到的数据作为CNN的输入,该方法将误报率降至0.034/h,比原始CNN的虚警率低了1/5。预处理步骤并不是必要的,比如Golmohammadi等^[8]直接使用线性频率倒谱系数及一阶、二阶倒数进行特征提取,然后输入到长短时记忆网络(Long Short-Term Memory, LSTM)中去,虽然特异性高于90%,但灵敏度低于35%。Jana等^[9]使用非支配排序遗传算法(Non-dominated Sorting Genetic Algorithm II, NSGA-II)选取3个最优通道的数据,降低计算复杂度,直接使用1维卷积神经网络(One-Dimensional Convolutional Neural Network, 1D-CNN)对信号进行特征提取和分类,该模型的准确率、灵敏度和特异性高达96%以上,但该实验没有设置测试集数据,且由于计算时间复杂度高,仅选取5例患者的数据进行通道选择和分类。Li等^[10]证明具有Waxman相似图的CNN具有最高的准确率。脑电图1 s的epoch在98%以上的时间里被正确地分配到癫痫发作期、癫痫发作前期或癫痫发作间期类别,但实验数据较少。Toraman^[11]比较3种预训练的CNN模型:VGG16、残差网络(ResNet)和密集连接网络(DenseNet)。利用频谱图图像来区分癫痫发作前期和发作间期的状态。研究发现,ResNet模型的性能最佳,准确率为90.32%,灵敏度为91.05%,特异性为89.76%,该方法可解决实验数据较少的问题。考虑到脑电信号的序列性,Aslam等^[12]采用结合长短期记忆的卷积神经网络(CNN-LSTM)对脑电图信号进行分类并预测癫痫发作,准确率为94%,灵敏度为93.8%,特异性为91.2%。为了加强对长期时间依赖性的处理,Ma等^[13]引入基于多通道特征融合的卷积双向长短期记忆网络模型(MCFF-CNN-Bi-LSTM),该模型整合注意力机制和信道融合,能有效管理长

期时间信号,同时降低计算复杂度。在波士顿儿童医院收集的顽固性癫痫头皮脑电数据集(CHB-MIT)数据集上可实现94.83%的准确率和94.94%的灵敏度,但在波恩数据集上的准确率和灵敏度不到80%。同时,Palanichamy等^[14]使用带有循环神经网络的时间注意力卷积神经网络(TA-CNN-LSTM),在CHB-MIT数据集和波恩数据集的准确率、灵敏度均高于94%。单绍杰等^[15]从多个脑电通道中选择C3-P3通道,并对此通道信号用零相移滤波器进行滤波,由DWT得到细节系数并求取其能量值,然后使用LSTM得到分类结果,并对分类结果进行卡曼滤波,从而减少误报率,最终的预测结果达到98.95%的精确度和98.73%的敏感度,但仅设置了训练集和测试集数据。

现有的文献中,大部分都采用独立于患者的实验,虽然此方法得到的结果较好,但在面对从未见过的数据时,需要通过脑电记录得到患者的数据,再由医生标注患者的癫痫发作状态,并将标注好的数据放入模型训练,费时费力,且加重患者的痛苦和对身体的伤害;再者,大部分脑电数据处理方法和模型都较为复杂,导致较高的时间复杂度,影响临床的最佳救治时间。因此,将所有患者数据作为模型输入非常重要,针对从未见过的患者数据,将此数据输入到模型中可以直接预测患者是否将要发生癫痫,而不需要收集该患者癫痫发作前期和癫痫发作期的数据。因此,本文提出一种基于CNN和Cox-Stuart的模型,可以在只对数据进行预处理的情况下,直接使用数据的时域信号作为模型的输入,减少时间复杂度和计算复杂度。

1 CNN与Cox-Stuart方法

1.1 CHB-MIT头皮脑电数据集介绍

本文使用CHB-MIT数据集^[16]。该数据集共23名受试者,病例1和病例21来自同一患者,所以共24个文件夹。数据集由10-20国际标准导联系统采集,采样率为256 Hz。 .edf是一种专门用于记录脑电信号的文件格式。每名受试者有9~42个连续的 .edf文件,记录癫痫期和非癫痫期大脑的波形信号,癫痫发作的起止时间已由专家标注。

1.2 数据选取方法

1.2.1 通道选择 在对患者进行脑电数据的采集时,部分患者 .edf文件的脑电通道数和电极名称略有不同,且有重复的电极名称。为了确保预测结果的良好,最终决定使用多数受试者数据都具有的22个脑电通道:FP1-F7、F7-T7、T7-P7、P7-O1、FP1-F3、F3-C3、C3-P3、P3-O1、FP2-F4、F4-C4、C4-P4、P4-O2、FP2-F8、

F8-T8、T8-P8-0、P8-O2、FZ-CZ、CZ-PZ、P7-T7、T7-FT9、FT9-FT10、FT10-T8、T8-P8-0和T8-P8-1是重复通道，删除T8-P8-1。由于第24个文件夹中没有标注每个.edf文件的起止时间，无法划分数据的发作前期和发作间期，不予考虑。因此，共筛选出23个文件夹的数据，分别是：chb01、chb02、chb03、chb04、chb05、

chb06、chb07、chb08、chb09、chb10、chb11、chb12、chb13、chb14、chb15、chb16、chb17、chb18、chb19、chb20、chb21、chb22、chb23。表1截取部分.edf文件的脑电通道数和电极名称。chb是所有.edf文件的前缀名，chb01指第1例患者，chb01_01指第1例患者的第1个.edf文件，以此类推。

表1 部分患者的通道数和电极名称

Table 1 Number of channels and electrode names for some patients

文件名	通道数	电极名称
chb01_01	23	['FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', 'FP2-F8', 'F8-T8', 'T8-P8-0', 'P8-O2', 'FZ-CZ', 'CZ-PZ', 'P7-T7', 'T7-FT9', 'FT9-FT10', 'FT10-T8', 'T8-P8-1']
chb01_02	23	['FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', 'FP2-F8', 'F8-T8', 'T8-P8-0', 'P8-O2', 'FZ-CZ', 'CZ-PZ', 'P7-T7', 'T7-FT9', 'FT9-FT10', 'FT10-T8', 'T8-P8-1']
chb12_24	28	['FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', '--0', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', '--1', 'FZ-CZ', 'CZ-PZ', '--2', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', '--3', 'FP2-F8', 'F8-T8', 'T8-P8-0', 'P8-O2', '--4', 'P7-T7', 'T7-FT9', 'FT9-FT10', 'FT10-T8', 'T8-P8-1']
chb12_32	29	['FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', '--0', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', '--1', 'FZ-CZ', 'CZ-PZ', '--2', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', '--3', 'FP2-F8', 'F8-T8', 'T8-P8-0', 'P8-O2', '--4', 'P7-T7', 'T7-FT9', 'FT9-FT10', 'FT10-T8', 'T8-P8-1', 'LOC-ROC']

1.2.2 癫痫发作状态划分方法 由于公开数据集CHB-MIT仅标注癫痫发作的开始和结束时间，因此需要人工定义癫痫发作间期和发作前期数据段。本文根据癫痫发作的开始和结束时间，将发作前期定义为癫痫发作前15 min到癫痫发作前5 min之内，发作后期定义为癫痫发作结束后30 min之内，发作间期定义为此次癫痫发作后2 h到下次癫痫发作前2 h之内。由于发作前期和发作间期的脑电数据分布不均衡，因此对某名患者数据量较少的一类数据进行10 s窗口重叠。图1是癫痫发作状态示意图。



图1 癫痫状态划分

Figure 1 Classification of different states in epilepsy

1.2.3 数据集划分方式 目前，共有两种癫痫脑电数据划分方式：患者独立的数据和非患者独立的数据。(1)患者独立的数据划分：患者独立是指将每名患者的数据划分为训练集、验证集和测试集。然后将对应的数据集放入模型中进行训练、验证、测试。该方法和非患者独立的数据集相比，评估结果更好。但是，该方法需要提前录入患者的数据，只有该患者有足够的不同发作时期数据，才能训练好该患者的模

型，从而检测或预测患者的发作时期。(2)非患者独立的数据划分：非患者独立是指将所有患者的数据整合在一起，再将数据集划分为训练集、验证集、测试集。该方法只需要使用训练集提前训练好模型，便可对从未见过的受试者数据进行分类，更适合临床。但是，该数据划分方法对模型有较高要求，只有模型的泛化能力较强，才能得到较好的分类结果。

1.3 CNN

CNN是一种受生物拓扑学启发的机器学习方法，由于它们的数据点可以以类似于图像的矩阵形式组织，因此适合分析脑电信号^[17]。基本结构有：输入层、卷积层、激活层、池化层、全连接层、输出层，如图2所示。输入层负责接收原始数据；输出层根据网络的任务类型，产生最终的输出。

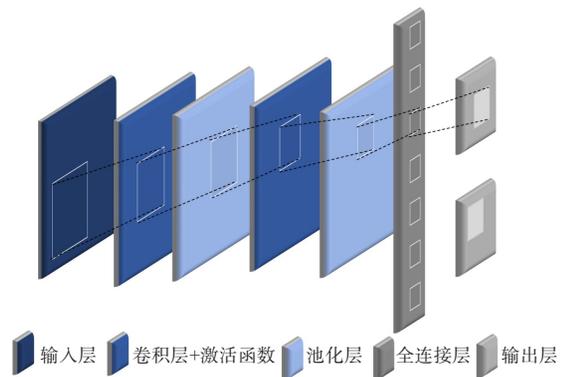


图2 CNN基本结构图

Figure 2 Basic structures of CNN

卷积层是 CNN 中最关键的步骤之一。卷积层通过卷积核对数据进行点乘运算, 由于卷积层的稀疏连接作用, 使卷积核的大小远小于输入数据的大小, 从而可以利用小规模的卷积核提取到大规模的数据特征, 减少参数运算量, 提高模型的处理效率。而卷积层之间的权值共享功能显著降低模型的存储需求, 也提高模型的运算效率。随着卷积层数的增加, 卷积核便能够提取到数据的高级特征^[18]。在进行卷积运算时, 除了利用卷积核提取到特征, 还允许为特征加偏置, 增强模型的泛化能力。有时, 为避免出现输出数据过小导致无法进行下一步卷积操作的情况

或避免信息丢失, 在进行卷积处理前, 会向输入数据的周围填充固定数据。但是, 不当的填充可能会导致数据量变多, 增加计算复杂度, 也可能降低模型的泛化能力。为了降低计算复杂度、减少特征大小, 允许卷积核在输入上每隔几步计算一次特征, 这意味着总体上所需的计算次数显著减少, 得到的特征图较小。但当步幅较大时, 会导致卷积核错过部分数据, 无法捕捉该信息特征, 造成特征提取不充分, 模型泛化能力减弱^[19]。本文不考虑卷积的偏置, 没有使用填充机制, 所有卷积操作的步幅均为 1。图 3 为卷积运算过程。

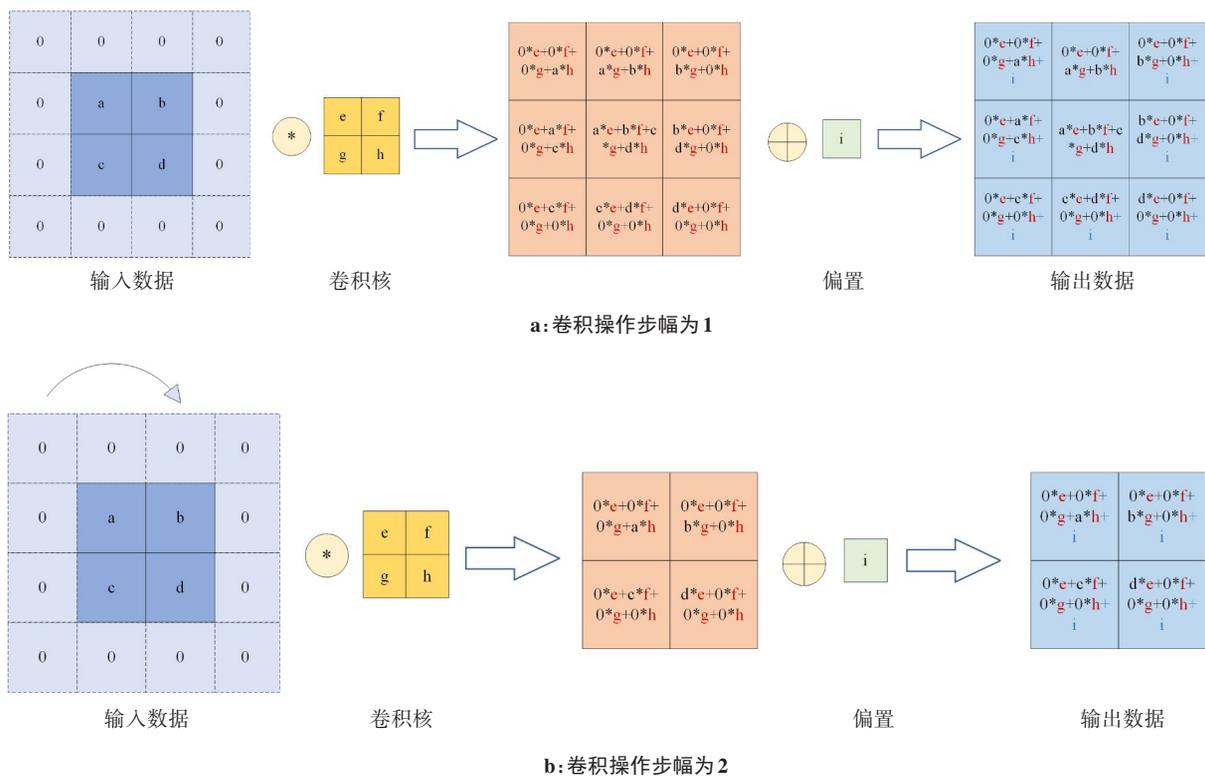


图 3 卷积运算过程

Figure 3 Convolutional operation process

激活层大多使用 Sigmoid、ReLU、LeakyReLU 等函数对数据进行非线性转换, 为神经网络注入非线性神经元, 从而提高神经网络的拟合能力^[20]。Sigmoid 函数容易出现梯度消失的问题, 而 ReLU 函数解决梯度消失问题且收敛速度较快, 但在 $x < 0$ 时可能会导致某些神经元在训练过程中永远不会被激活, 造成“死亡神经元”, LeakyReLU 在 $x < 0$ 时加入修正因子, 使得负数部分的输出不再为 0, 解决 ReLU 函数的这一缺点^[21]。式 (1)、式 (2) 为 ReLU 函数和 LeakyReLU 函数。图 4 为两者的函数图。

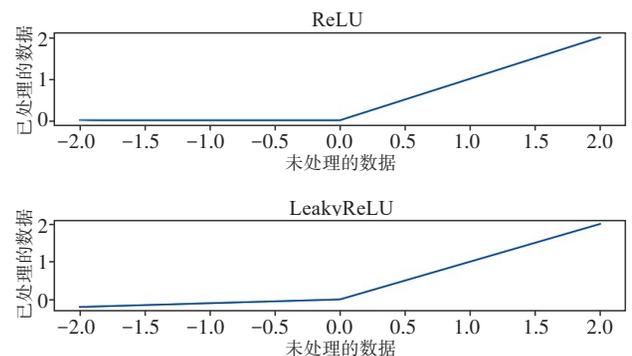


图 4 ReLU 和 LeakyReLU 函数图

Figure 4 ReLU and LeakyReLU function diagrams

$$h_{\text{ReLU}}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

$$h_{\text{LeakyReLU}}(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (2)$$

池化操作在于缩小高、长方向上的空间,降低维度,从而减少参数,降低计算复杂度,防止发生过拟合现象。此外,当输入数据发生微小变化时,池化操作不会影响输出数据的大小,因此池化对数据的微小变化具有鲁棒性^[19]。最常用的池化方法有最大池化、平均池化、自适应平均池化等。最大池化找出输入特征图选定区域中的最大值作为池化层的值;平均池化计算出输入特征图选定区域中所有值的平均值作为池化层的值;自适应平均池化不用考虑输入特征图的大小,根据所需要的输出形状大小自适应地对输入特征图进行平均池化。最大池化和平均池化的工作示意图如图5所示。全连接层通过将上一层的输出展平成一组一维向量,并与权重矩阵相乘得到全局特征,或用于最终分类。本文提出的模型使用批归一化层,批归一化层并不属于CNN的基本结构组成部分,但是它通过对每个小批次的输入数据进行归一化处理,使模型的学习速度加快、减少初始值的影响以及抑制过拟合^[19]。

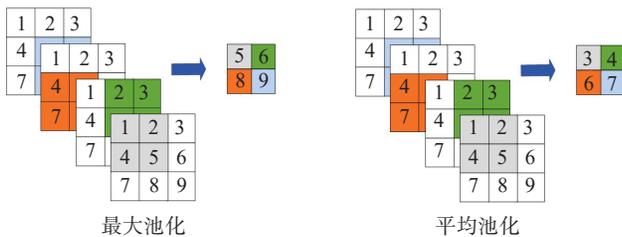


图5 最大池化和平均池化工作示意图

Figure 5 Schematic diagrams of maximum pooling and average pooling operations

1.4 Cox-Stuart 检验法

Cox-Stuart 检验法利用正负符号判断一组序列是否存在某种趋势,该方法适应于多种类型的数据,且不依赖数据的分布情况,仅根据符号检验判断是否有上升或下降的趋势^[22]。该检验原理可理解为将序列中的数据分成前后两部分,用后半部分序列值减去前半部分序列值,得到若干有正有负的差值。正差值表明后半部分的序列值大于前半部分的序列值(可能有上升趋势);负差值表明后半部分的序列值小于前半部分的序列值(可能有下降趋势)。根据假设检验,认为原假设是无变化趋势的,备择假设是有变化趋势的。

因此,需比较正差值和负差值的数量关系。当正差值较少(可能有下降趋势)且此时正差值出现的概率很低,说明序列有下降趋势,拒绝原假设;同理,当负差值较少(可能有上升趋势)且此时负差值出现的概率很低,说明序列有上升趋势,拒绝原假设。若正差值较多,但正差值的概率函数值并不低,不能判断未来的序列呈下降趋势,则无法推翻原假设;同理,若负差值较多,但负差值的概率函数值不低,不能判断未来的序列呈上升趋势,也无法推翻原假设;若正负差值数量同样多,说明该序列无趋势,接受原假设。一般认为概率函数阈值为0.05。图6为Cox-Stuart检验法的流程图。具体的计算方法如下所示。(1)假设: H_0 : 序列无趋势性; H_1 : 序列有趋势性;(2)输入序列 $\{x_1, x_2, \dots, x_n\}$ 并计算 c ,数组长度为 n , n 为偶数时, $c = n/2$, n 为奇数时, $c = [(n + 1)/2] - 1$;(3)计算 c 组数对值: $\{d_1 = x_c - x_0, d_2 = x_{c+1} - x_1, \dots, d_c = x_n - x_{c-1}\}$,统计 d_i 的正符号个数 num_+ 和负符号个数 num_- ,令 $k = \min(\text{num}_+, \text{num}_-)$;(4)使用二项分布的累积概率函数计算概率 $p_{\text{Cox-Stuart}}$, p 为出现一次正号或一次负号的概率值: $p_{\text{Cox-Stuart}} = p(x \leq k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}$;(5)若 $\text{num}_+ > \text{num}_-$ 且 $p \leq 0.05$,认为函数值有上升趋势,若 $\text{num}_+ < \text{num}_-$ 且 $p \leq 0.05$,认为函数值有下降趋势,否则认为没有明显趋势。

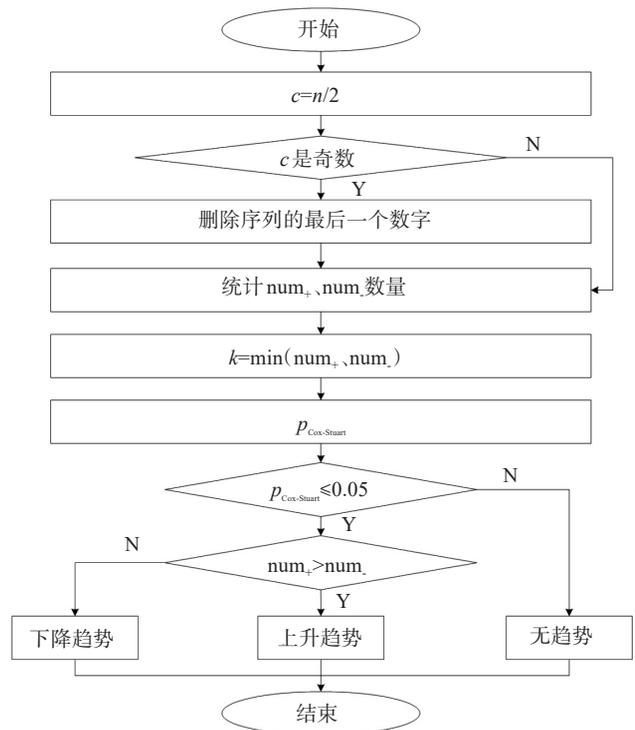


图6 Cox-Stuart 检验法流程图

Figure 6 Flowchart of Cox-Stuart test

2 CNN 癫痫预测模型

本实验使用非患者独立的数据划分方法, 总共 25 436 个数据, 随机选取 20% 的数据作为测试集, 将测试集数据和剩余数据分别保存至固定文件夹中, 每次运行代码时, 随机选取剩余数据的 20% 作为验证集数据, 其他作为训练集。训练集、验证集、测试集数据的划分比例为 6:2:2。

首先对原始数据进行预处理, 包括数据归一化和滤波; 将预处理后的训练集数据和验证集数据放入设计好的 CNN 模型中进行训练并验证; 对验证集的损失函数做 Cox-Stuart 检验, 判断是否有上升趋势并早停; 保存训练结束后的模型参数, 并使用测试集数据测试模型性能, 得到测试集的灵敏度、特异性、准确率; 实验结束。图 7 为实验流程图。

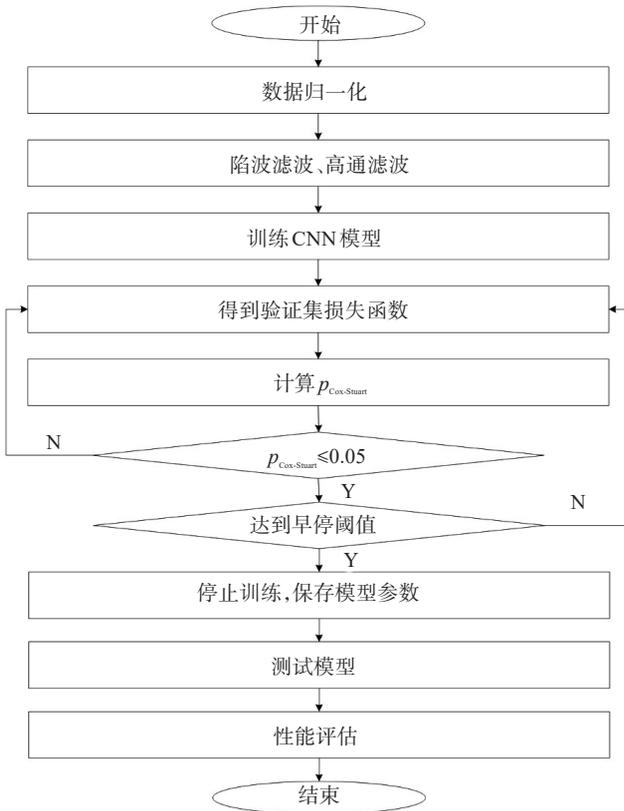


图7 实验流程图

Figure 7 Experimental flowchart

2.1 预处理

处于不同发作状态的脑电信号在幅度上可能有着较大或较小的差异, 不同脑电通道的信号幅度也有差异, 为了更好地区分癫痫发作间期和癫痫发作前期, 对原始数据进行归一化处理是非常有必要的。本文使用 Scikit-learn 库中的 MinMaxScaler 标准化方

法。MinMaxScaler 可以将数据线性转换到指定的范围, 公式如下:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \times (\max - \min) + \min \quad (3)$$

其中, x 为原始数据, x_{min} 和 x_{max} 分别是序列的最小值和最大值, \max 和 \min 分别是目标缩放范围的上限和下限, 在本文中默认指定范围为 $[0, 1]$ 。 x_{scaled} 是标准化后数据。

环境中的电磁干扰和设备本身存在的热噪声会对本身就比较微弱的脑电信号造成极大的噪声干扰。针对工频干扰噪声, 一般采用 50 Hz 或 60 Hz 陷波滤波器处理^[23]。本文对每个 .edf 文件的数据用 mne 库自带的陷波滤波函数和高通滤波函数分别滤除 60 Hz 的工频噪声和小于 1 Hz 的噪声^[24]。图 8~图 10 分别是 chb01_01 文件的原始功率谱密度图、经陷波滤波后的功率谱密度图、经高通滤波后的功率谱密度图。

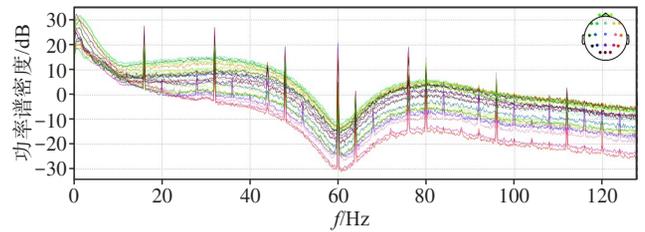


图8 chb01_01原始功率谱密度图

Figure 8 Original power spectral density map of chb01_01

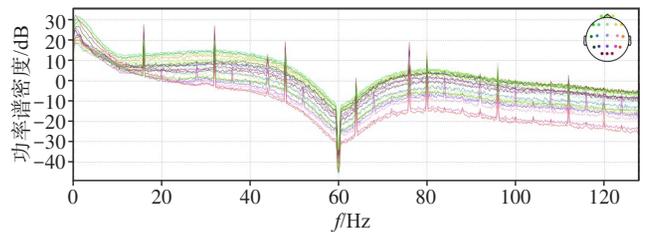


图9 chb01_01陷波滤波后的功率谱密度图

Figure 9 Power spectral density map of chb01_01 after notch filtering

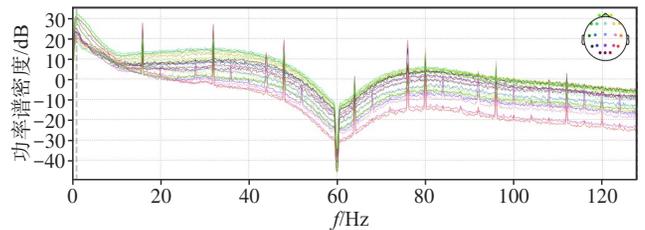


图10 chb01_01高通滤波后的功率谱密度图

Figure 10 Power spectral density map of chb01_01 after high-pass filtering

2.2 CNN Cox-Stuart 模型

对数据进行预处理后,直接放入设计好的网络中进行特征提取和分类。本文使用3个2维卷积层、2个批归一化层、2个最大池化层、2个激活函数设计一个时间复杂度低、计算复杂度低、泛化能力强的2维卷积神经网络(Two-Dimensional Convolutional Neural Network, 2D-CNN)模型,模型结构如图11所示。表2为模型结构的具体参数。模型的激活函数为LeakyReLU,损失函数为CrossEntropyLoss,优化器为Adam,学习率为0.0001,批大小为128。

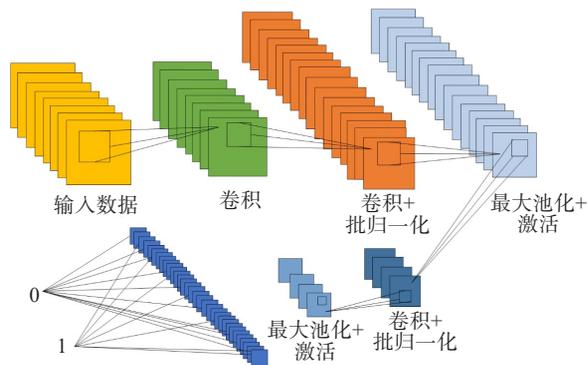


图11 2D-CNN 模型结构

Figure 11 2D-CNN model structure

表2 模型结构参数

Table 2 Model structural parameters

层	输入通道	输出通道	卷积核	输出维度	步长
输入	22	22	-	(128, 22, 10, 256)	-
2D 卷积	22	22	1×1	(128, 22, 10, 256)	1
2D 卷积	22	44	3×3	(128, 44, 8, 254)	1
批归一化	44	44	-	(128, 44, 8, 254)	-
最大池化	44	44	2×4	(128, 44, 4, 63)	1
激活函数	-	-	-	(128, 44, 4, 63)	-
2D 卷积	44	5	3×3	(128, 5, 2, 61)	1
批归一化	5	5	-	(128, 5, 2, 61)	-
最大池化	5	5	2×4	(128, 5, 1, 15)	1
激活函数	-	-	-	(128, 5, 1, 15)	-
全连接	-	75	-	(128, 2)	-
输出	75	2	-	(128, 2)	-

2.3 基于Cox-Stuart方法的模型早停机制

在训练模型时,由于每次的训练集和验证集不同,参数的初始化数据也有差异,因此很难确定训练次数。Cox-Stuart检验法用来判断序列的趋势是下降还是上升,且并不要求清楚数据的分布情况,假定

条件少,可以通过Cox-Stuart检验法判断验证集损失函数是否有上升趋势从而使训练早停。将前20次训练后得到的20个验证集损失函数作为Cox-Stuart检验法的输入序列,计算该序列的 $p_{Cox-Stuart}$ 值。当 $p_{Cox-Stuart}$ 值 ≤ 0.05 ,且由序列值得到的正差值数量大于负差值数量时,说明损失函数有上升趋势,此时将学习率降低到原来的0.1倍,当上升趋势的次数达到设定阈值时,不再训练模型,开始进入测试阶段;当 $p_{Cox-Stuart}$ 值 ≤ 0.05 ,且由序列值得到的正差值数量小于负差值数量,说明损失函数有下降趋势,继续训练,若不满足以上两种条件,继续训练模型。经过多次实验,确定阈值为20。

3 结果

3.1 评价指标

在对模型进行评估时,混淆矩阵是必不可少的,通过混淆矩阵可以计算准确率、特异性、灵敏度等评价指标的值。混淆矩阵由真阳、真阴、假阳、假阴组成。TP代表真阳性,指模型预测结果为发作前期且真实标签也为发作前期;TN代表真阴性,指模型预测结果为发作间期且真实标签也为发作间期;FP代表假阳性,指模型预测结果为发作前期但真实标签为发作间期;FN代表假阴性,指模型预测结果为发作间期但真实标签为发作前期^[25]。根据混淆矩阵的定义可知,TP+FN为阳性的标签总数,也就是所有发作前期的标签数量;TN+FP为阴性的标签总数,也就是所有发作间期的标签数量。

本文用到的评价指标有准确率、灵敏度、特异性,公式分别如式(4)~式(6)所示。其中,准确率指所有预测标签中,预测正确的概率;灵敏度指在所有真实的发作前期标签中,模型预测结果也为发作前期的概率,即阳性的准确率;特异性指在所有真实的发作间期标签中,模型预测结果也为发作间期的概率,即阴性的准确率^[26]。

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

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

$$\text{specificity} = \frac{TN}{TN + FP} \quad (6)$$

3.2 性能评估

在不使用Cox-Stuart检验法的情况下,将训练的epoch设为150,共进行4次实验。训练集和验证集的结果如图12~图15所示。由图12可知,训练次数约在80次以后,验证集损失函数有上升趋势。此时,验证集的准确率、特异性均可达到95%以上,灵敏度可达到90%以上。由图13可知,epoch在40次以后,验证集损失函

数有上升趋势。此时,验证集的准确率、灵敏度、特异性均可达到95%以上。由图14可知,epoch在40次以后,验证集损失函数有上升趋势。此时,验证集的准确率、灵敏度、特异性均趋于稳定且大于95%。由图15可知,epoch在65次左右,验证集损失函数有上升趋势。此时,验证集的准确率和灵敏度可到95%以上,特异性

趋近于95%。由图12~图15可知,损失函数下降到一定程度时有上升趋势,每次实验时,下降到最低点时的epoch差距较大。4次实验的验证集准确率、灵敏度、特异性均可大于90%。使用测试集对CNN模型进行性能评估,得到的4个混淆矩阵如图16所示。表3为4次测试集结果及其平均值。

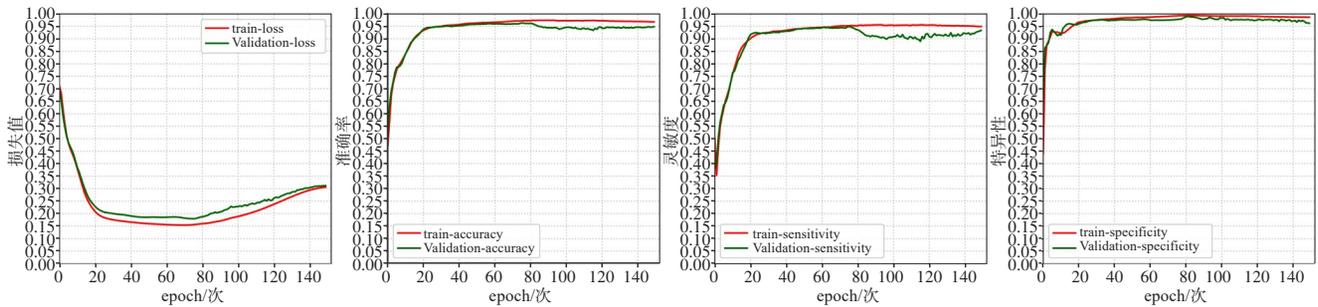


图12 第1次实验的训练集和验证集结果图

Figure 12 Results on the training and validation sets for the 1st experiment

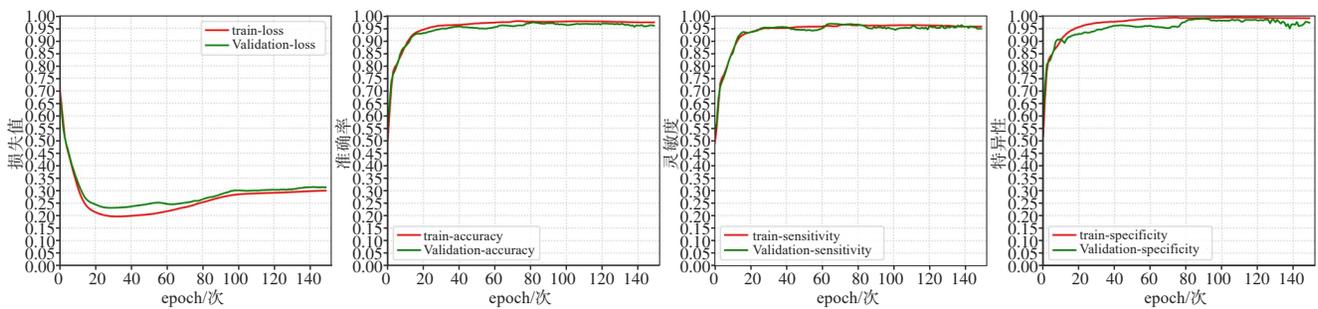


图13 第2次实验的训练集和验证集结果图

Figure 13 Results on the training and validation sets for the 2nd experiment

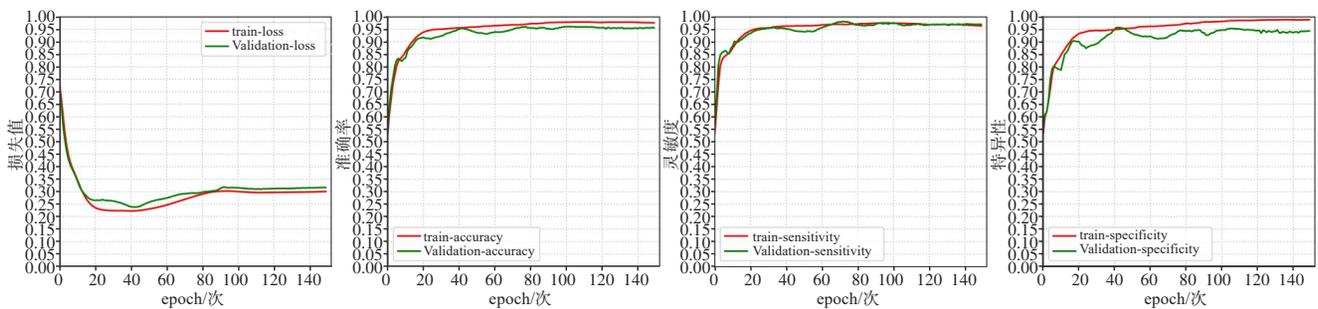


图14 第3次实验的训练集和验证集结果图

Figure 14 Results on the training and validation sets for the 3rd experiment

使用Cox-Stuart检验法判断验证损失函数是否有上升趋势并早停,其他不变。使用训练集和验证集对模型进行训练及验证的结果如图17~图20所示。由图17可以看出,epoch在40次左右时,验证集损失函数有上升趋势,此时验证集的准确率、灵敏度、特

异性均可达到95%以上。由图18可以看出,epoch在60次左右时,验证集损失函数有上升趋势。此时,验证集的准确率、特异性大于95%,灵敏度大于90%。由图19可知,epoch在约55次以后,验证集损失函数出现上升趋势。此时,验证集的准确率、灵敏度、特

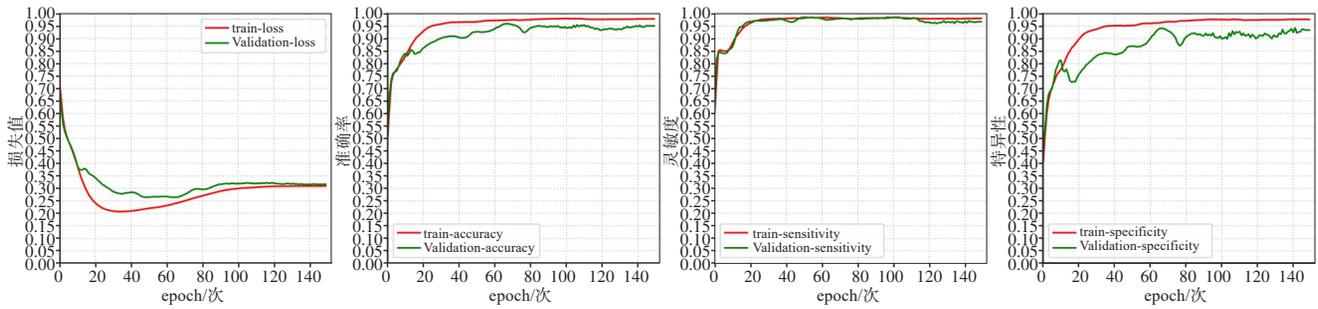


图15 第4次实验的训练集和验证集结果图

Figure 15 Results on the training and validation sets for the 4th experiment

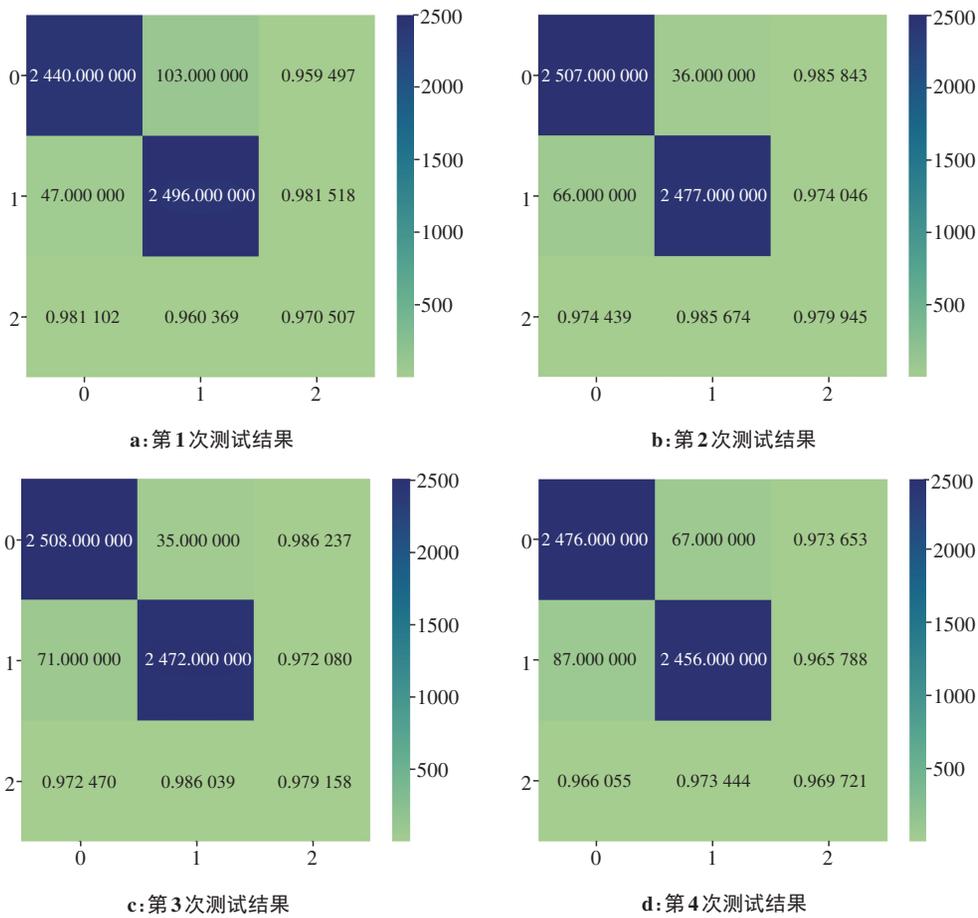


图16 4次实验的混淆矩阵图

Figure 16 Confusion matrices for 4 experiments

表3 4次测试集结果及其平均值(%)

Table 3 Results on 4 test sets and their averages (%)

实验次数	准确率	灵敏度	特异性
1	97.05	95.95	98.15
2	97.99	98.58	97.40
3	97.92	98.62	97.21
4	96.97	97.37	96.58
平均值	97.48	97.63	97.34

异性均大于95%。由图20可知, epoch在约55以后, 验证集损失函数有上升趋势。此时, 准确率, 灵敏度, 特异性均可到达90%及以上。通过对比分析图17~图20发现, epoch均在150次以内, 训练集和验证集的准确率、灵敏度、特异性均在90%以上。使用测试集对Cox-Stuart和CNN模型进行性能评估, 4次测试集分类结果的混淆矩阵分别如图21所示。表4为4次实验结果及其平均值。

通过实验可知, 加入Cox-Stuart早停机制后的

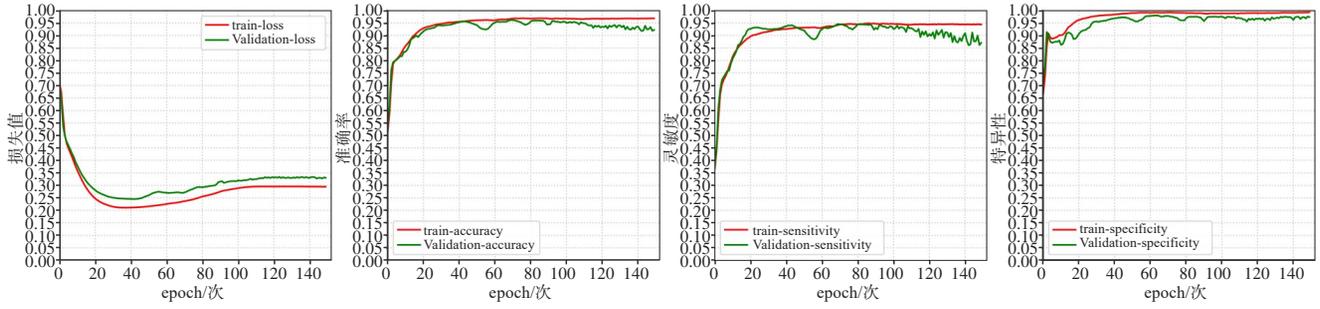


图 17 加入 Cox-Stuart 检验法后第 1 次实验的训练集和验证集结果图

Figure 17 Results on the training and validation sets for the 1st experiment after combining with Cox-Stuart test

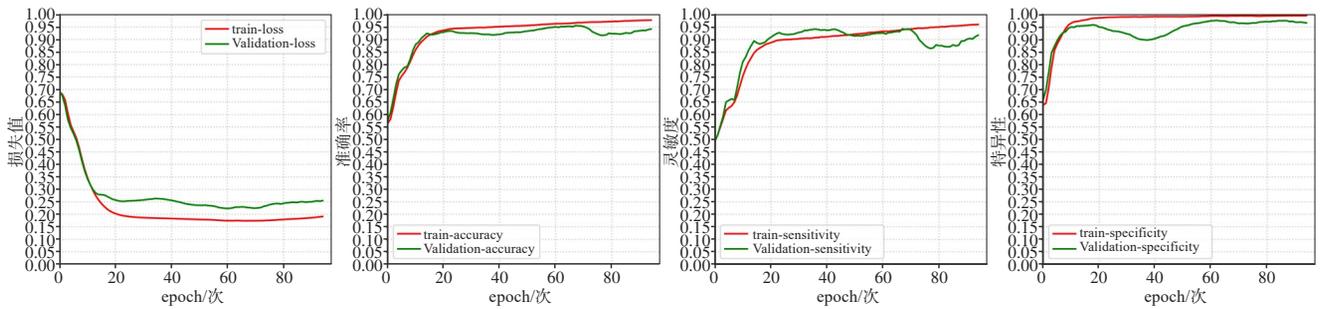


图 18 加入 Cox-Stuart 检验法后的第 2 次实验的训练集和验证集结果图

Figure 18 Results on the training and validation sets for the 2nd experiment after combining with Cox-Stuart test

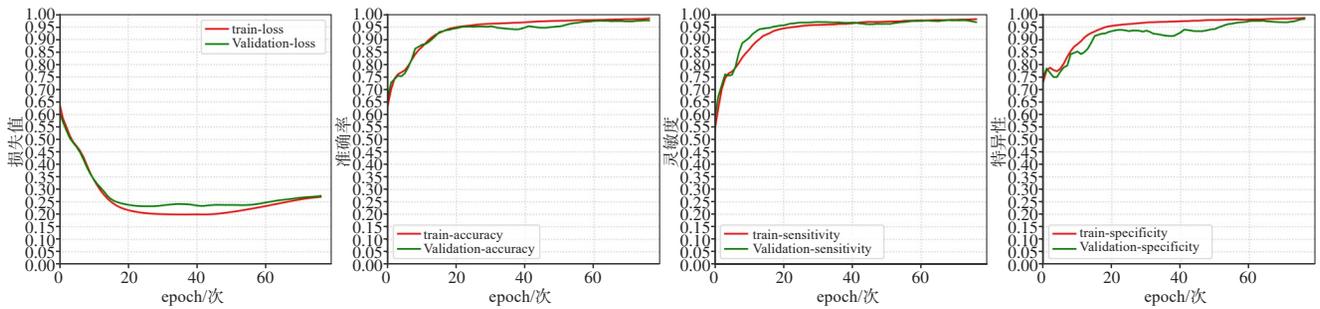


图 19 加入 Cox-Stuart 检验法后的第 3 次实验的训练集和验证集结果图

Figure 19 Results on the training and validation sets for the 3rd experiment after combining with Cox-Stuart test

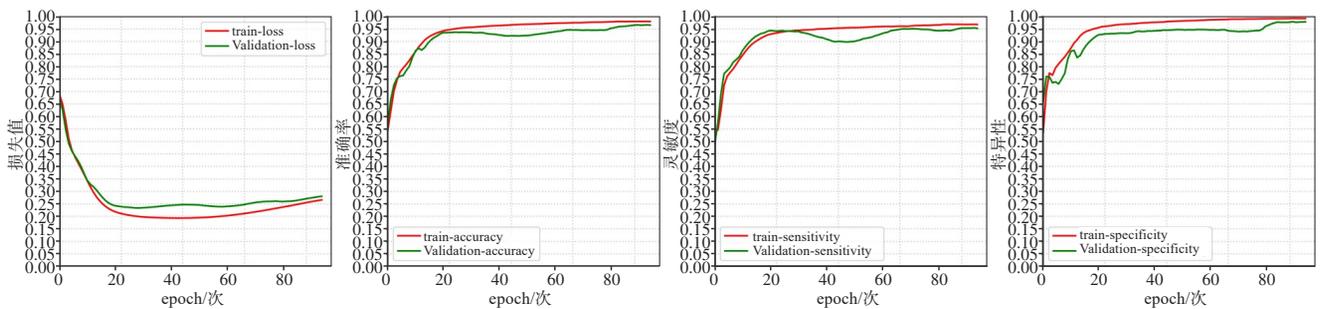


图 20 加入 Cox-Stuart 检验法后的第 4 次实验的训练集和验证集结果图

Figure 20 Results on the training and validation sets for the 4th experiment after combining with Cox-Stuart test

实验结果和不加 Cox-Stuart 早停机制的实验结果相差很小,说明加入 Cox-Stuart 早停机制是可行的。Cox-Stuart 早停机制在验证集损失函数有上升趋势

时应及时停止训练,可大大降低模型的计算复杂度和时间复杂度。表 5 为两者的实验结果对比。

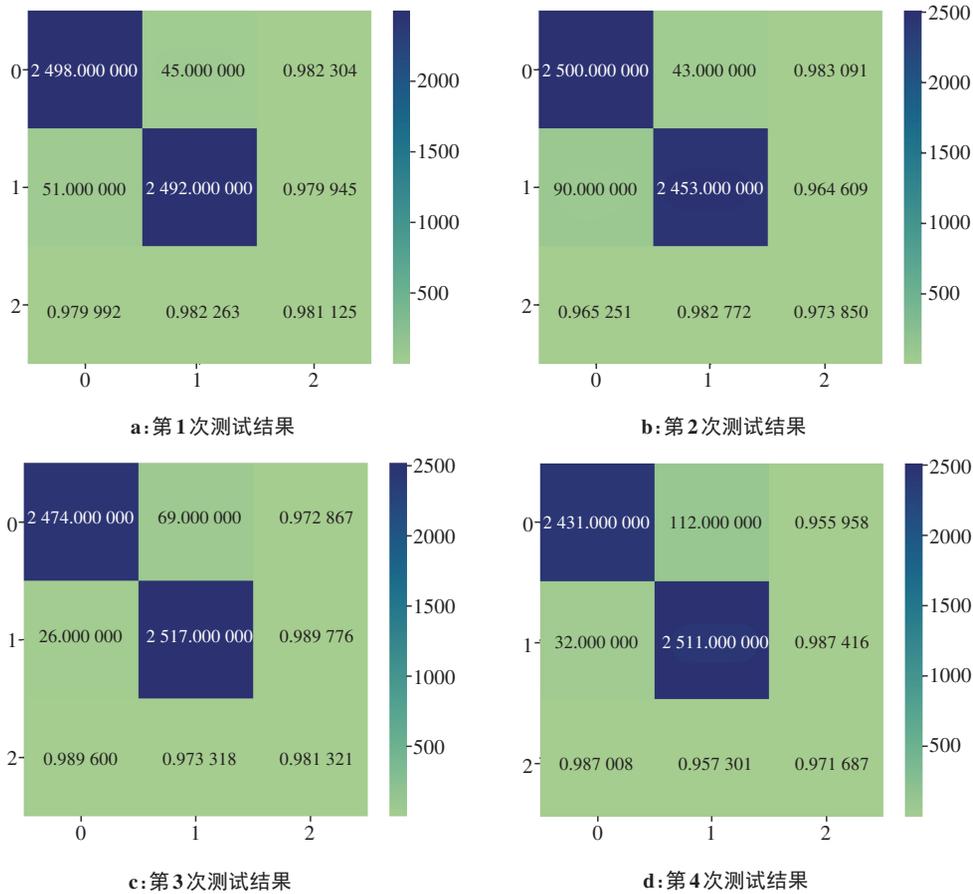


图21 加入Cox-Stuart检验法后4次实验的混淆矩阵图

Figure 21 Confusion matrixes for 4 experiments after combining with Cox-Stuart test

表4 4次测试集结果及其平均值

Table 4 Results on 4 test sets and their averages

实验次数	训练次数/次	准确率/%	灵敏度/%	特异性/%
1	134	98.11	98.23	97.99
2	148	97.39	98.31	96.46
3	142	98.13	97.29	98.98
4	102	97.17	95.60	98.74
平均值	132	97.70	97.36	98.04

表5 不加Cox-Stuart早停机制和加入Cox-Stuart早停机制的测试集实验结果对比

Table 5 Comparison of experimental results on the test set with and without Cox-Stuart early stopping mechanism

方法	训练次数/次	准确率/%	灵敏度/%	特异性/%
2D-CNN	150	97.48	97.63	97.34
2D-CNN和Cox-Stuart	132	97.70	97.36	98.04

3.3 对比实验

表6对比了基于整体癫痫患者的研究结果和本

文的研究结果。表7对比了基于独立癫痫患者的研究结果和本文的研究结果。为了突出模型性能,本文还做了发作前期为30和60 min的实验,均添加在表6和表7中^[27-32]。由表6可知,当发作前期为10 min时,本文的模型效果最好,均优于其他文献的实验结果。当发作前期为30 min时,本文的特异性低于文献[30]的特异性,除此之外,本文的模型效果最优。当发作前期为60 min时,本文的模型在整体上次于文献[30],优于其他文献的实验结果。文献[30]将脑电数据转换为图片形式,进行3维卷积处理,其计算复杂度和时间复杂度高于本文。由表7可知,当发作前期为10 min时,本文的模型效果最好,均优于其他文献的实验结果。当发作前期为30和60 min时,本文的模型无明显优势,但本文时间复杂度和计算复杂度均小于其他模型。此外,本文使用基于整体患者的数据划分方式,相较于表7使用独立于患者的数据划分方式,本文的实验难度更大。通过分析表6和表7可知,发作前期为10 min时,本文提出的模型效果最好,均优于表6和表7中的其他文献,且本文模型的时间复杂度和计算复杂度较低。

表6 基于非患者独立数据划分方式下的不同模型的实验结果对比

Table 6 Comparison of experimental results of different models based on non-independent data partitioning method

方法	数据集	预处理	特征提取+模型	结果	发作前期时间/min
文献[27]	CHB-MIT	陷波滤波,带通滤波	CDAN	准确率:70.90%	60
	Siena			准确率:60.27%	
文献[28]	峨山医疗中心儿童医院	陷波滤波,高通滤波,低通滤波	AC-GRU	准确率:82.86% 灵敏度:80.00% 特异性:85.50%	10
文献[29]	CHB-MIT	小波变换	1D-CNN	准确率:86.50% 灵敏度:84.40%	10
文献[30]	新疆医科大学	3D图像	3D-CNN	准确率:92.37% 灵敏度:88.90% 特异性:93.78%	60
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:97.70% 灵敏度:97.36% 特异性:98.04%	10
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:93.79% 灵敏度:94.35% 特异性:93.22%	30
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:88.84% 灵敏度:91.17% 特异性:86.51%	60

表7 基于患者独立数据划分方式下的不同模型实验结果对比

Table 7 Comparison of experimental results with different models based on patient independent data partitioning method

方法	数据集	预处理	特征提取+模型	结果	发作前期时间
文献[31]	CHB-MIT	-	自相关+Lightweight-2D-CNN	准确率:89.98% 灵敏度:92.90% 特异性:87.04%	15 min
文献[9]	CHB-MIT	-	NSGA-II+CNN	准确率:96.51% 灵敏度:96.55% 特异性:96.47%	10 min
文献[32]	CHB-MIT	-	振幅谱+CNN-SVM	准确率:86.25%	60 min
文献[12]	CHB-MIT	高通滤波,短时傅里叶变换	CNN-LSTM	准确率:94.00% 灵敏度:93.80% 特异性:91.20%	19.5 min
文献[14]	CHB-MIT	-	CGAN-TACNN-LSTM	准确率:94.60% 灵敏度:94.50%	-
	Bonn			准确率:94.80% 灵敏度:94.90%	
文献[13]	CHB-MIT	-	MCFF-CNN-Bi-LSTM	准确率:94.83% 灵敏度:94.94%	小于1s
	Bonn			准确率:77.62% 灵敏度:77.62%	
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:97.70% 灵敏度:97.36% 特异性:98.04%	10 min
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:93.79% 灵敏度:94.35% 特异性:93.22%	30 min
本文方法	CHB-MIT	陷波滤波,高通滤波	2D-CNN和Cox-Stuart	准确率:88.84% 灵敏度:91.17% 特异性:86.51%	60 min

4 结论

本文提出的2D-CNN模型针对非独立患者,只需对数据做归一化和滤波处理便可输入到模型中对数据做特征提取和分类;其次,使用Cox-Stuart检验法设置早停机制,有效减少时间复杂度和计算复杂度。和其他文献相比,本文提出的模型结构简单,计算复杂度和时间复杂度较低,泛化能力强,模型效果较好。当发作前期为10 min时,在测试集上的准确率为97.70%,灵敏度为97.36%,特异性为98.04%。但本文还存在以下不足:(1)实验数据过于单调,本文使用的是顽固性癫痫患者的数据,虽然在此数据集上的效果较好,但尚不清楚在其他类型的癫痫患者脑电数据上的效果;(2)实验结果的灵敏度、特异性、准确率有待提高。

【参考文献】

- [1] Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: position paper of the ILAE commission for classification and terminology[J]. *Epilepsia*, 2017, 58(4): 512-521.
- [2] 周东, 吴欣桐, 李劲梅. 癫痫共病的诊断和治疗[J]. *西部医学*, 2019, 31(6): 821-827.
Zhou D, Wu XT, Li JM. Diagnosis and treatment of comorbidities in patients with epilepsy[J]. *Medical Journal of West China*, 2019, 31(6): 821-827.
- [3] Liu GD, Xiao RL, Xu LY, et al. Minireview of epilepsy detection techniques based on electroencephalogram signals[J]. *Front Syst Neurosci*, 2021, 15: 685387.
- [4] 帅国彬, 胡伟波. 基于新型能量算子和多尺度熵的癫痫自动检测[J]. *计算机应用与软件*, 2023, 40(3): 137-141.
Shuai GB, Hu WB. Automatic detection of epilepsy based on new energy operator and multi-scale entropy[J]. *Computer Applications and Software*, 2023, 40(3): 137-141.
- [5] Bao XQ, Xu YJ, Kamavuako EN. The effect of signal duration on the classification of heart sounds: a deep learning approach[J]. *Sensors (Basel)*, 2022, 22(6): 2261.
- [6] Pan YY, Zhou XY, Dong FY, et al. Epileptic seizure detection with hybrid time-frequency EEG input: a deep learning approach[J]. *Comput Math Methods Med*, 2022, 2022: 8724536.
- [7] Takahashi H, Emami A, Shinozaki T, et al. Convolutional neural network with autoencoder-assisted multiclass labelling for seizure detection based on scalp electroencephalography[J]. *Comput Biol Med*, 2020, 125: 104016.
- [8] Golmohammadi M, Ziyabari S, Shah V, et al. Gated recurrent networks for seizure detection[C]//2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). Piscataway, NJ, USA: IEEE, 2017: 1-5.
- [9] Jana R, Mukherjee I. Efficient seizure prediction and EEG channel selection based on multi-objective optimization[J]. *IEEE Access*, 2023, 11: 54112-54121.
- [10] Li Z, Fields M, Panov F, et al. Deep learning of simultaneous intracranial and scalp EEG for prediction, detection, and lateralization of mesial temporal lobe seizures[J]. *Front Neurol*, 2021, 12: 705119.
- [11] Toraman S. Preictal and interictal recognition for epileptic seizure prediction using pre-trained 2D-CNN models[J]. *Traitement Du Signal*, 2020, 37(6): 1045-1054.
- [12] Aslam MH, Usman SM, Khalid S, et al. Classification of EEG signals for prediction of epileptic seizures[J]. *Appl Sci*, 2022, 12(14): 7251.
- [13] Ma YH, Huang ZT, Su JY, et al. A multi-channel feature fusion CNN-Bi-LSTM epilepsy EEG classification and prediction model based on attention mechanism[J]. *IEEE Access*, 2023, 11: 62855-62864.
- [14] Palanichamy I, Sundaram V. Improving deep learning for seizure detection using GAN with cramer distance and a temporal-spatial-frequency loss function[J]. *Int J Recent Innovation Trends Comput Commun*, 2023, 11(6s): 424-432.
- [15] 单绍杰, 李汉军, 王璐璐, 等. 基于LSTM模型的单导联脑电癫痫发作预测[J]. *计算机应用研究*, 2018, 35(11): 3251-3254.
Shan SJ, Li HJ, Wang LL, et al. Epileptic seizure prediction from single channel scalp EEG based on LSTM model[J]. *Application Research of Computers*, 2018, 35(11): 3251-3254.
- [16] Shoeb AH. Application of machine learning to epileptic seizure onset detection and treatment[D]. Cambridge: Massachusetts Institute of Technology, 2009.
- [17] Lai CQ, Ibrahim H, Suandi SA, et al. Convolutional neural network for closed-set identification from resting state electroencephalography[J]. *Mathematics*, 2022, 10(19): 3442.
- [18] 伊恩·古德费洛, 约书亚·本吉奥, 亚伦·库维尔. 深度学习[M]. 赵中剑, 黎戎君, 符天凡, 等译. 北京: 人民邮电出版社, 2017: 203-210.
Goodfellow I, Bengio Y, Courville A. Deep learning[M]. Translated by Zhan SJ, Li YJ, Fu TF, et al. Beijing: Posts & Telecom Press, 2017: 203-210.
- [19] 斋藤康毅. 深度学习入门: 基于Python的理论与实现[M]. 陆宇杰译. 北京: 人民邮电出版社, 2018: 184-216.
Saito Y. Introduction to deep learning: theory and implementation based on Python[M]. Translated by Lu YJ. Beijing: Posts & Telecom Press, 2018: 184-216.
- [20] Zhou WC, Zheng W, Feng YB, et al. LMA-EEGNet: a lightweight multi-attention network for neonatal seizure detection using EEG signals[J]. *Electronics*, 2024, 13(12): 2354.
- [21] Lai DK, Cheng ES, So BP, et al. Transformer models and convolutional networks with different activation functions for swallow classification using depth video data[J]. *Mathematics*, 2023, 11(14): 3081.
- [22] 胡成雨, 张焕明, 朱家明. Cox-Stuart及Wilcoxon符号秩检验在产品质检中的应用[J]. *齐齐哈尔大学学报(自然科学版)*, 2020, 36(2): 80-85.
Hu CY, Zhang HM, Zhu JM. Application of Cox-Stuart and Wilcoxon symbol rank test in product quality inspection[J]. *Journal of Qiqihar University (Natural Science Edition)*, 2020, 36(2): 80-85.
- [23] Raghu S, Sriraam N, Vasudeva Rao S, et al. Automated detection of epileptic seizures using successive decomposition index and support vector machine classifier in long-term EEG[J]. *Neural Comput Appl*, 2020, 32(13): 8965-8984.
- [24] Liu YH, Chen L, Li XW, et al. Epilepsy detection with artificial neural network based on as-fabricated neuromorphic chip platform[J]. *AIP Adv*, 2022, 12(3): 035106.
- [25] Luque A, Carrasco A, Martín A, et al. The impact of class imbalance in classification performance metrics based on the binary confusion matrix[J]. *Pattern Recognit*, 2019, 91: 216-231.
- [26] Zeng GP. On the confusion matrix in credit scoring and its analytical properties[J]. *Commun Stat Theory Methods*, 2020, 49(9): 2080-2093.
- [27] Jemal I, Abou-Abbas L, Henni K, et al. Domain adaptation for EEG-based, cross-subject epileptic seizure prediction[J]. *Front Neuroinform*, 2024, 18: 1303380.
- [28] Choi W, Kim MJ, Yum MS, et al. Deep Convolutional gated recurrent unit combined with attention mechanism to classify pre-ictal from interictal EEG with minimized number of channels[J]. *J Pers Med*, 2022, 12(5): 763.
- [29] Halawa RI, Youssef SM, Elagamy MN. An efficient hybrid model for patient-independent seizure prediction using deep learning[J]. *Appl Sci*, 2022, 12(11): 5516.
- [30] Wei XY, Zhou L, Chen ZY, et al. Automatic seizure detection using three-dimensional CNN based on multi-channel EEG[J]. *BMC Med Inform Decis Mak*, 2018, 18(Suppl 5): 111.
- [31] Zhang SS, Chen D, Ranjan R, et al. A lightweight solution to epileptic seizure prediction based on EEG synchronization measurement[J]. *J Supercomput*, 2021, 77(4): 3914-3932.
- [32] Hu WB, Cao JW, Lai XP, et al. Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks[J]. *J Ambient Intell Humaniz Comput*, 2023, 14(11): 15485-15495.

(编辑:陈丽霞)