视觉分散自适应补偿的 SSVEP 多域协同解码研究

贾书婷，温昕，郝雁嵘，曹锐
（太原理工大学软件学院，山西 晋中 030600）

摘　要： 基于稳态视觉诱发电位（SSVEP）的脑机接口（BCI）由于个体差异和非目标刺激干扰而面临分类性能瓶颈，且现有方法尚未深入探究视觉分散干扰与个体差异之间的量化关系。为此，提出一种视觉分散自适应补偿的 SSVEP 多域协同解码算法，其核心包含视觉分散自适应标签平滑技术与多域联合解码模型两部分。首先，基于视觉拥挤效应理论，构建信号幅值与标签噪声的自适应量化模型，通过动态调节标签平滑强度实现个体视觉分散程度的量化表征，在缓解模型过拟合的同时，削弱非目标刺激的干扰影响，抑制个体差异的不利影响；然后，提出一种多域联合解码模型，先通过特征提取框架实现时频空域深度协同，再引入双向长短时记忆（BiLSTM）网络对时序全局依赖关系进行建模，形成兼具局部感受野与长程上下文感知能力的复合特征表达。在 3 个公开 SSVEP 数据集上以 0.5 s 和 1.0 s 两个不同长度的时间窗分别进行验证，结果表明，在所有实验设置下，该算法的平均识别准确率和平均信息传输率相较对比方法都更具优势。消融实验表明，视觉分散自适应补偿机制在各实验设置下均能发挥作用，尤其是在短时间窗口下，准确率最高能提升 18 百分点，表明该方法为神经解码中的个体适应性优化与时频特征融合提供了新的思路。

关键词： 脑机接口；稳态视觉诱发电位；视觉分散；标签平滑技术；多域协同
源代码链接： https://github.com/shutingjia/Multi-Domain-Collaborative-Decoding-of-SSVEP
中图分类号： TP391　　**文献标志码：** A　　**DOI：** 10.19678/j.issn.1000-3428.0070595

Multi-Domain Collaborative Decoding of SSVEP with Adaptive Visual Distraction Compensation

JIA Shuting, WEN Xin, HAO Yanrong, CAO Rui
(School of Software, Taiyuan University of Technology, Jinzhong 030600, Shanxi, China)

【Abstract】 Brain-Computer Interface (BCI) systems based on Steady-State Visual Evoked Potential (SSVEP) show classification performance limitations due to individual differences and interference from non-target stimuli. Moreover, existing methods have not thoroughly explored the quantitative relationship between visual distraction interference and individual differences. To address these issues, we propose an SSVEP multi-domain collaborative decoding algorithm with adaptive compensation for visual distraction. The algorithm includes an adaptive label smoothing technique for visual distraction and a multi-domain joint decoding model. First, based on the theory of the visual crowding effect, an adaptive quantitative model linking signal amplitude and label noise is constructed. By dynamically adjusting the intensity of label smoothing, the model achieves a quantitative representation of individual visual distraction levels, thereby mitigating model overfitting while reducing interference from non-target stimuli and suppressing the adverse effects of individual differences. Subsequently, a multi-domain joint decoding model is proposed. This model first achieves deep collaboration across time-frequency-spatial domains through a feature extraction framework and then introduces a Bidirectional Long Short-Term Memory (BiLSTM) network to model temporal global dependencies, forming a composite feature representation that combines local receptive fields with long-range contextual awareness. Validation on three publicly available SSVEP datasets using time windows of 0.5 s and 1.0 s demonstrates that, under all experimental settings, the proposed algorithm exhibits superior average recognition accuracy and average information transfer rate compared to other state-of-the-art methods. Ablation experiments reveal that the adaptive compensation mechanism for visual distraction is effective across all experimental settings, with accuracy improvements of up to 18 percentage points in short time windows. The findings indicate that the proposed approach provides new insights for optimizing individual adaptability and fusing time-frequency features in neural decoding.

基金项目： 国家自然科学基金（62206196）；山西省自然科学基金（202103021223035，202303021221001）；山西省研究生科研创新基金（2023KY291）。

作者简介： 贾书婷，女，硕士研究生，主研方向为稳态视觉诱发电位、脑机接口；温昕，副教授、博士；郝雁嵘，讲师、博士；曹锐（通信作者），副教授、博士。

收稿日期： 2024-11-08　　**修回日期：** 2025-02-27　　**E-mail：** caorui@tyut.edu.cn

【Key words】Brain-Computer Interface (BCI); Steady-State Visual Evoked Potential (SSVEP); visual distraction; label smoothing technique; multi-domain collaboration

0 引言

脑机接口(BCI)是一种新型的人机交互技术,在完成人脑通信外部电子设备的过程中,不需要借助人脑常规信息输出通道[1]。脑电图(EEG)由于其无创性、高分辨率而被广泛用作 BCI 成像[2]。与运动想象和 P300 相比,基于稳态视觉诱发电位(SSVEP)的 BCI 具有高信噪比、高信息传输率的特点[3]。因此,基于 SSVEP 的 BCI 在过去十年中快速发展[4],广泛应用于医疗[5]、军事[6]、娱乐[7]等领域。

然而,SSVEP 信号的高度个体差异性长期制约着解码模型的性能。由于不同个体在脑电信号特征和反应模式上的差异,导致同一解码系统在不同个体之间的性能表现存在显著波动。实验研究表明,不同被试在相同刺激条件下产生的信号幅值差异可达 3～5 倍[8],基频谐波分布模式页也存在明显差异。在 40 Hz 高频刺激下,约 32%的被试的二次谐波能量超过基频成分[9],而在低频段则呈现相反趋势[10]。最新研究[11]发现,个体差异存在的一大主要原因是大脑处理视觉信息的方式不同。简言之,当被试视觉中心周围出现干扰时,大脑顶枕叶皮层的视觉处理区抑制干扰的能力因人而异。具体来说,当被试专注目标刺激时,大脑枕区越活跃,产生的 SSVEP 信号质量越高[12],但如果周围出现高频干扰,大脑中传递视觉信号的外侧膝状体便会抑制目标信号,这种现象被称为“视觉拥挤”[13],它会直接改变脑电波的频谱特征,实验[14]明确显示它与 FFT 频谱变化有关。在 SSVEP 信号采集过程中,也存在“视觉拥挤”现象,被试注视目标时易受周围非目标刺激的干扰,导致视觉分散,有研究表明,周围刺激干扰可使分类准确率下降 22%[15]。上述发现表明,周围非目标刺激对 SSVEP 分类性能产生巨大干扰,控制这种干扰带来的影响可有效缓解个体差异问题,这为抑制个体差异研究提供了新视角。

对于基于 SSVEP 的 BCI 来说,分类精度直接影响系统的信息传输速率与用户体验。现有 SSVEP 分类器主要沿袭时空特征驱动与频域特征驱动两条技术路线。然而,SSVEP 信号本身具有多模态耦合特性,在时间维度上呈现节律振荡特征,在空间维度上反映头皮电位分布规律,在频域维度上则与刺激频率及其谐波成分紧密关联。这种跨域耦合特性要求分类器具备多尺度特征解耦与融合能力,而当前方法往往因建模视角单一导致关键信息流失。

针对上述问题,本文提出一种解码算法——视觉分散自适应补偿的 SSVEP 多域协同解码算法,该算法主要由视觉分散自适应标签平滑技术和多域联合解码模型两部分组成。视觉分散自适应标签平滑技术基于视觉拥挤效应理论,建立信号幅值与标签噪声的自适应量化模型,更好地捕捉和调节个体差异,提高方法的灵活性。此外,考虑到现有深度学习模型较容易受限于局部感受野和特征域单一的问题,本文还提出一种多域联合解码模型。本文主要贡献如下:

1)提出一种基于视觉拥挤效应的自适应标签平滑方法,建立信号幅值与标签噪声的量化模型,缓解视觉分散带来的不利影响。

2)针对 SSVEP 信号的多维特性,提出一种多域联合解码模型,用于 SSVEP 的分类任务。

3)在 3 个公开数据集上进行实验,验证所提方法的性能。

1 相关工作

1.1 SSVEP 信号中的个体差异研究

在针对 SSVEP 信号个体差异的研究中,DNN 模型[16]用微调的思想,使用单个被试的少量数据,对全局性的初始模型进行针对单个被试的适应性训练,针对个体差异进行单独训练。TEGAN 模型[17]使用生成对抗网络(GAN)延长 SSVEP 信号的数据长度,降低了个体差异噪声,具有更多判别性特征。文献[18]也利用 GAN 实现个体生物特征解耦,通过神经风格迁移技术构建主体不变特征空间。

上述方法多聚焦于数据驱动层面的优化,但它们将个体差异视为黑箱噪声进行处理,忽视了其背后的神经生物学机制。因此,本文在建模时考虑神经生物学机制,考虑外侧膝状体对视觉信号的抑制规律,从干扰刺激对目标识别的生物性影响入手。基于这种生物学影响,本文将视觉拥挤效应引入 SSVEP 范式解释。非目标刺激的高频闪烁会诱发神经抑制,导致被试对目标刺激的响应强度减弱,可以理解为非目标刺激与目标刺激进行视觉注意力资源争夺,非目标刺激的高频闪烁导致目标刺激的视觉注意力被分散。针对这种视觉分散,本文构建了

信号幅值与标签噪声的自适应量化模型，通过动态量化 SSVEP 幅值来表征不同被试受到的非目标刺激的视觉分散情况。该方法在特征层面区别个体抑制能力差异，自适应调整权重；在标签层面引入正则化，缓解模型过拟合问题。

1.2 SSVEP 信号分类器研究

在有关分类器的架构中，主流深度学习方法常聚焦 SSVEP 信号的时空特征提取[19]。SSVEPNet[20]模型融合卷积神经网络（CNN）时空特征学习与长短时记忆（LSTM）网络依赖关系建模，在 12-JFPM 数据集上取得了 98%的准确率。文献[21]提出的时空卷积注意力网络（ST-CAN）通过联合优化空域滤波器与时域动态特征，在 40 个目标系统中实现了 92.3%的平均识别准确率。DDGCNN[22]将图卷积方法应用到 SSVEP 领域，主要用于提取脑电空间拓扑结构特征。文献[23]也引入动态图神经网络（DGNN），有效捕捉脑电信号的非平稳特性，将 16 Hz 高频 SSVEP 信号解码效率提升 27%。上述网络主要提取时空域特征，对频域信息利用不足。

鉴于 SSVEP 由不同频率刺激产生[24]，具有明显频谱特性，还有许多研究侧重频域特征提取。为了利用谐波中的频率信息，有研究提出一种基于滤波器组的 CNN 架构 FBtCNN[25]，将时域信号通过滤波器组，整合来自不同频带的特征信息，从而提升分类准确率。AttentCNN 模型[26]也加入了这种滤波器组技术以学习频率信息，进一步提高了短时间窗 SSVEP 的分类性能。SEC-Net 方法[27]通过自适应频段选择机制，成功地将信息传输速率提升至 480 bit/min。文献[28]设计了一个复杂频谱卷积神经网络（C-CNN），在 7 分类 SSVEP 数据集上准确率达到 92.5%。SSVEPformer 模型[29]借鉴了 C-CNN 的思想，以 FFT 变换后的信号频谱作为网络输入，引入 Transformer 的基础模块对 SSVEP 特征进行编码，在 12-JFPM 和 Benchmark 数据集上的准确率分别达到 84.16%和 80.40%。文献[30]开发的 DeepFreqNet 通过多尺度频域金字塔结构，融合 FFT 与残差网络特征，将 12 分类数据集的信息传输率提升至 498 bit/min。上述网络主要学习频带信息，但存在频率分辨率受限、提取特征偏局部化等问题。

现有分类器方法中仍存在两方面局限：其一，时空特征网络对频域利用不足；其二，频域方法缺乏对神经信号动态时序特性的精细建模。本文提出一种多域联合解码分类器。将原始信号和模板信号作为输入，通过一个多阶段特征提取模块，分别在时空和频域上进行深度特征挖掘。首先，模型利用 CNN 提取信号的时空特征，充分挖掘局部特征和空间信息；接着，通过 FFT 将时空特征转换到频域，从频域角度进一步提取信号的频谱特征。为了克服 CNN 在捕捉长距离依赖关系时的局限性，引入 BiLSTM，利用其对序列数据的建模能力，有效补充卷积层的局部特征提取，从而实现全局依赖的捕捉，提升模型的时序感知能力，形成一种全新的时频空特征提取框架。

2 解码算法

2.1 视觉分散自适应标签平滑技术

标签平滑技术是一种有效的正则化方法，本文针对 SSVEP 信号提出一种自适应的基于视觉分散的标签平滑技术作为损失函数。通过引入个性因子 α，针对不同被试评估其个体差异，更好地适应不同被试的独有生物特征，从而提升信号分类准确率。

标签平滑技术通过对目标标签进行改进，来缓解模型对训练数据过拟合的问题。在 SSVEP 信号收集的过程中，所有刺激同时闪烁，势必会有非目标刺激对被试造成干扰，影响采集到的信号。如图 1 所示（彩色效果见《计算机工程》官网 HTML 版，下同），以 12-JFPM 数据集范式为例，红框标记的方格代表目标刺激 5，考虑视域因素，周围一圈的方格 1、2、3、4、6、7、8、9 的刺激也会对采集信号产生负面影响，将单标签扩展为多标签，可以缓解这种不利影响。

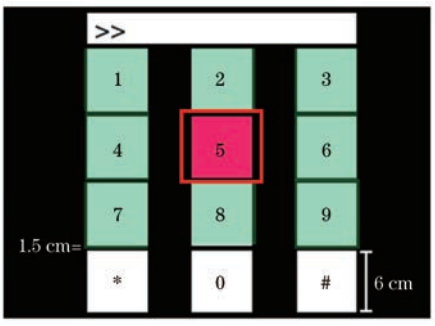

图 1 12-JFPM 数据集范式

Fig. 1 12-JFPM dataset paradigm

假设 N_f 个目标以 $n\times m$ 的形式排列成矩阵，那么第 t_g 个刺激目标的坐标位置 (x,y) 可以表示为：

$$(x,y)=\left(\frac{t_g}{m},t_g\%m\right) \tag{1}$$

假设周围单个非目标刺激对被试的影响相同，则所有非目标刺激的注意力分值计算为：

$$\beta_{t_g}=1.0/N_{\text{around}} \tag{2}$$

$$N_{\text{around}}=\sum_{i\neq j\neq 0,i=-1}^{1}\sum_{j=-1}^{1}((0\leqslant(x+i)\leqslant n-1)\&\&(0\leqslant(y+j)\leqslant m-1)) \tag{3}$$

式中：β 为注意力得分；N_{around} 表示周围影响信号采集结果的非目标刺激数量。可以得到基于注意力机制的多标签矩阵，如式(4)所示：

$$\boldsymbol{A}_{\text{ALS}}=\begin{pmatrix}1 & \beta_0 & \cdots & \beta_0\\ \beta_1 & 1 & \cdots & \beta_1\\ \vdots & \vdots & & \vdots\\ \beta_{t_g} & \beta_{t_g} & \cdots & 1\end{pmatrix} \tag{4}$$

式中：$\boldsymbol{A}_{\text{ALS}}$ 表示基于注意力机制计算得出的软标签矩阵，为了避免过度正则化，本文采用软硬标签相结合的方式训练模型。考虑到被试个体差异，对自身注意力的控制程度不同，导致在采集数据时受到非目标刺激的影响各不相同，所以软硬标签不能简单求和。引入个性因子 α，来描述被试受到周围刺激影响的程度大小，如式(5)所示：

$$\alpha=\text{Min-Max}\left(\sum_{t_g=1}^{N_f}\left(\sum|A_{t_g}-A_{\text{around}}|/N_{\text{around}}\right)\right) \tag{5}$$

式中：A 表示振幅；Min-Max(x)表示归一化，如式(6)所示。本文借助目标间的振幅差，来表示被试受影响的程度，令其自适应地选择个性因子。最终的损失函数如式(7)所示。

$$\text{Min-Max}(x)=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{6}$$

$$L=\alpha L_{\text{hard}}+(1-\alpha)L_{\text{soft}} \tag{7}$$

2.2 网络模型

本文提出一种多域联合解码模型来提高 SSVEP 的分类性能，模型主要包括特征提取模块和决策模块，完整模型如图 2 所示。

模型以原始信号 $S_{\text{sig}}\in\mathbb{R}^{N_s\times N_c\times 1}$ 和模板信号 $S_{\text{temp}}\in\mathbb{R}^{N_s\times N_c\times N_f}$ 作为输入。其中，N_s 表示采样时间点数，N_c 表示通道数，N_f 表示刺激类别数，N_{train} 表示每个刺激的训练试次数，$S_j^{(n)}$ 表示第 j 个刺激的第 n 次实验试次。模板信号表示为式(9)：

$$\overline{S}_j=\frac{1}{N_{\text{train}}}\sum_{n=1}^{N_{\text{train}}}S_j^{(n)} \tag{8}$$

$$S_{\text{temp}}=[\overline{S}_1,\overline{S}_2,\cdots,\overline{S}_j]\in\mathbb{R}^N \tag{9}$$

特征提取模块由 2 个卷积和 1 个双向长短时记忆网络组成。首先，令输入进来的信号经过空间卷积，利用大小为 8 的卷积核对信号的信道维度进行卷积，提取空间特征；随后，输出信号经过时间卷积，使用大小为 10 的卷积核来对时间维度进行卷积并提取时间特征。为了缓解训练过程中的过拟合等问题，在卷积层之后加入 BatchNorm、PReLU、Dropout 等操作。由于卷积感受野的局限性，CNN 对整体信号特征的感知具有局限性。为解决这一问题，又加入 BiLSTM 来提取相邻时间点之间的关系特征。同时，考虑到将 SSVEP 识别为时间序列会对固定频率刺激及其谐波做出响应，因此频域特征至关重要。在进入 BiLSTM 前，构建 2 个输入：一个是上层卷积输出的时空特征；另一个是通过 FFT 将时空特征转换为频域特征的输入。时空支线通过 BiLSTM 捕捉时空特征之间的相互依存关系，频域支线通过 BiLSTM 网络来提取融合频域特征。2 条支线输出结果通过 Add 操作融合后再经过 Flatten，最后送入决策模块。特征提取模块可以看作一个非线性函数 $F(\cdot)$，则 2 个分支的输出结果分别表示为 $X_{\text{sig}}=F(S_{\text{sig}})\in\mathbb{R}^{N_s\times N_c\times 1}$ 和 $X_{\text{temp}}=F(S_{\text{temp}})\in\mathbb{R}^{N_s\times N_c\times N_f}$。

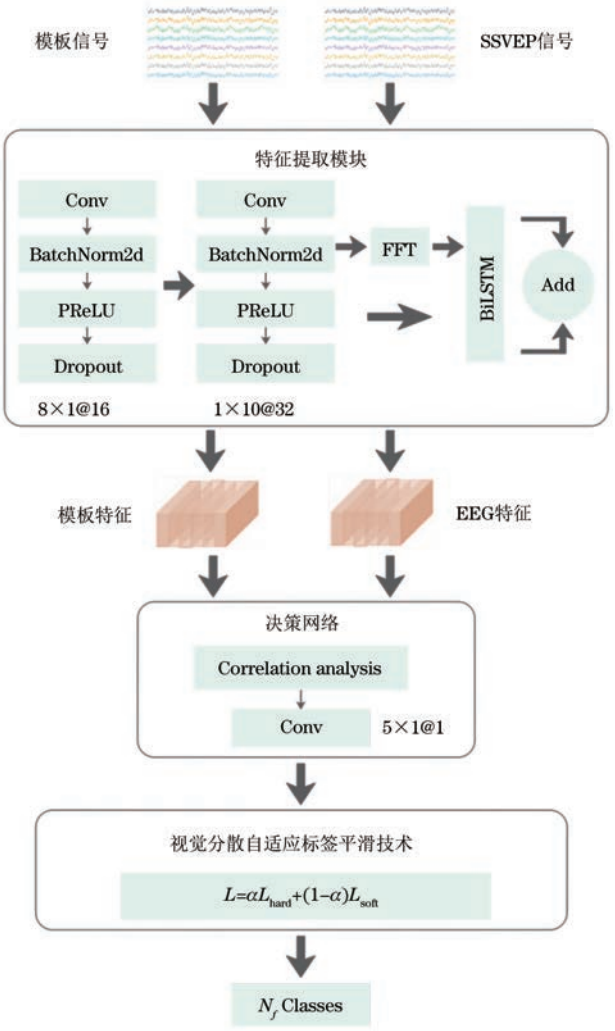

图 2 本文模型结构

Fig. 2 The structure of the model in this paper

决策模块由一个相关性分析和一个卷积组成。

在相关性分析层中，计算脑电特征信号 X_{sig} 和第 j 个模板信号的第 i 个通道 $X_{\text{temp}}^{(\cdot,i,j)}$ 的平均值作为相关性系数：

$$s_{\text{sim},j}=\frac{1}{N_c}\sum_{i=1}^{N_c}\text{Dist}(X_{\text{sig}}^{(\cdot,i,1)},X_{\text{temp}}^{(\cdot,i,j)}) \tag{10}$$

$$\text{Dist}(x,y)=\frac{x^T y}{\|x\|_2\|y\|_2} \tag{11}$$

式中：$i\in\{1,2,\cdots,N_c\}$ 代表通道索引；$j\in\{1,2,\cdots,N_f\}$ 代表刺激目标索引；$\|\cdot\|_2$ 表示 L2 范数。

相关性分析层的输出为：

$$s_{\text{sim}}=[s_{\text{sim},1},s_{\text{sim},2},\cdots,s_{\text{sim},N_f}]\in\mathbb{R}^{N_f} \tag{12}$$

然后，加入一个卷积核大小为 5 的卷积操作，实验表明，在输出前加入一个卷积层，可以使模型的训练过程更加稳定。

模型中的具体参数设置如表 1 所示，k 表示神经网络中卷积核的大小。

表 1 模型参数设置

Table 1 Model parameters setting

Module	Layer type	Kernel size	#Filters	Output
特征提取模块	Input	—	—	$(1@N_c\times N_s)$
	Conv2d	$(N_c,1)$	$2\times N_c$	$(2\times N_c@1\times N_s)$
	BatchNorm2d	—	—	$(2\times N_c@1\times N_s)$
	PReLU	—	—	$(2\times N_c@1\times N_s)$
	Dropout	—	—	$(2\times N_c@1\times N_s)$
	Conv2d	$(1,10)$	$4\times N_c$	$(4\times N_c@1\times(N_s-k)/2+1)$
	BatchNorm2d	—	—	$(4\times N_c@1\times(N_s-k)/2+1)$
	PReLU	—	—	$(4\times N_c@1\times(N_s-k)/2+1)$
	Dropout	—	—	$(4\times N_c@1\times(N_s-k)/2+1)$
	BiLSTM	—	—	$(4\times N_c\times((N_s-k)/2+1)\times 2)$
	FFT	—	—	$(4\times N_c\times((N_s-k)/2+1)\times 2)$
	BiLSTM	—	—	$(4\times N_c\times((N_s-k)/2+1)\times 2)$
	Flatten	—	—	$(4\times N_c\times((N_s-k)/2+1)\times 2)$
决策网络	Correlation analysis	—	—	N_f
	Conv2d	$(5,1)$	1	N_f

3 数据集与实验设置

在 12-JFPM[31]、Benchmark[32] 和 Wearable[33] 这 3 个公开数据集上分别进行实验。

对于 12-JFPM 数据集，12 个目标分别排布在 4×3 的视觉刺激矩阵中，频率为 9.25～14.75 Hz，间隔为 0.5 Hz。相同行的刺激目标使用相同相位，不同行的相位为 0.0～1.5，间隔为 0.5。实验使用 Ag/AgCl 电极，共 8 个通道（PO7、PO3、POZ、PO4、PO8、O1、Oz 和 O2），从 10 名被试中采取数据，数据包含 15 个区块，每块包含分别对应 12 个不同刺激的 12 个试次。每次实验开始后，目标刺激变为红色提示，被试需要在这 1 s 内将视线移至目标上，之后所有刺激进行持续 4 s 的闪烁。所有数据被下采样到 256 Hz，然后用无限脉冲响应（IIR）滤波器进行 6～80 Hz 的滤波。

对于 Benchmark 数据集，数据采集自 35 名被试，40 个目标分别排布在 5×8 的视觉刺激矩阵中，使用 JFPM 方法对其进行编码，频率为 8.0～15.8 Hz，间隔为 0.2 Hz，相邻目标间的相位差为 0.5。对于单个被试来说，数据包含 6 个区块，每块包含分别对应 40 个不同刺激的 40 个试次。实验的提示时间为 0.5 s，刺激闪烁时间为 5 s，紧接着是 0.5 s 的被试休息时间。实验共收集 64 个电极的信号。所有数据被下采样到 250 Hz，然后进行 6～80 Hz 的高低通滤波和 50 Hz 的凹陷滤波，实验中仅适用视觉枕区的 9 个电极（Pz、PO5、PO3、POz、PO4、PO6、O1、Oz 和 O2）。

对于 Wearable 数据集，12 个目标分别排布在 3×4 的视觉刺激矩阵中，频率为 9.25～14.75 Hz，间隔为 0.5 Hz。相同行的刺激目标使用相同相位，不同行的相位为 0.0～1.5，间隔为 0.5。实验采用可穿戴式的干电极和传统湿电极 2 种方式采集数据，共 8 个通道（POz、PO3、PO4、PO5、PO6、Oz、O1 和 O2），数据从 102 名被试中采集。数据包含多个区块，每个区块包含分别对应 12 个不同刺激的 10 个试次。每次实验开始后，有 0.5 s 的提示期，目标刺激会以特殊方式提示被试将视线移至目标上，随后所有刺激进行持续 2 s 的闪烁，还有 0.14 s 的视觉延迟期以及 0.2 s 的视觉后效期，每次实验总时长为 2.84 s。

所有数据原始采样率为 1 000 Hz，被下采样到 250 Hz。本文仅使用湿电极采集的数据。

原始 EEG 信号按 0.5 s 与 1.0 s 时间窗截取，训练时采用步长 50 ms 的滑动窗口增强以扩充样本多样性，测试阶段禁用滑动以避免数据泄漏。数据划分以被试为单位，按照不同比例进行划分，以验证训练数据量较少时的模型性能。实验使用 5 折交叉验证，每折交替作为测试集，其余用于训练。

在优化策略上，使用 AdamW 优化器（$\beta_1=0.9$，$\beta_2=0.999$），结合权重衰减系数 0.01 与梯度裁剪（阈值 5.0）抑制过拟合。在不同数据集上，学习率设置存在差异：12-JFPM 与 Wearable 数据集上初始学习率为 0.003，Benchmark 数据集因类别数多、信号变异大，初始学习率降为 0.001。训练批次大小固定为 30，最大迭代次数为 500，并设置早停机制。

模型通过正则化缓解过拟合问题：全连接层丢弃率设为 0.5，卷积层丢弃率为 0.2；标签平滑技术改进为动态机制，个性因子 α 根据个体视觉分散水平在 0～1 间自适应调整。实验基于 PyTorch 1.13.1 实现。

4 实验结果与分析

4.1 数据分析

本文使用单因素重复测量方差分析（ANOVA）统计方法来评估分类方法在不同条件下的显著性差异，使用事后配对 t 检验来评估相同条件下方法对之间的差异。其中，$\alpha=0.05$。下文通过星号来表示显著性水平，“*”表示结果在统计学上达到显著性水平（$p<0.05$），“**”表示更强的显著性（$p<0.01$），“***”表示高度显著性（$p<0.001$）。

4.2 消融实验

消融实验证明了模型特征提取模块中的时空侧分支、频域侧分支、视觉分散自适应标签平滑损失函数（损失）、模板信号输入这 4 个主要模块的有效性。(a)、(b)、(c)、(d) 4 组分别对应剔除其中一个主要模块，(e)组是本文提出的完整模型。

在表 2 中（最优结果加粗表示），对于 12-JFPM 数据集，使用 ANOVA 统计方法计算的不同模块影响显著（0.5 s 时间窗下 $F(3,28)=539.818$，$p<0.001$；1.0 s 时间窗下 $F(3,28)=475.813$，$p<0.001$）。在同一时间窗口下，缺少时间侧分支的模型性能均明显降低（Case(e) vs. Case(a)：all $p<0.001$），证明时间侧分支不可忽视的作用。在同一时间窗口下，完整模型的准确率也均高于缺少频率侧、使用普通损失和缺少模板的模型，尤其是在 0.5 s 的短时间窗口下，能发挥更大影响（Case(e) vs. Case(b)：$p_1<0.001$，$p_2<0.01$；Case(e) vs. Case(c)：$p_1<0.001$，$p_2<0.01$；Case(e) vs. Case(d)：$p_1<0.001$，$p_2<0.05$）。本文提出的完整模型不仅能提升分类准确率，其标准差也远低于其余 3 组实验结果，这表明该模型大幅提升了网络稳定性。

在表 2 中，对于 Benchmark 数据集，使用 ANOVA 统计方法计算的不同模块影响显著（0.5 s 时间窗下 $F(3,28)=974.266$，$p<0.001$；1.0 s 时间窗下 $F(3,28)=749.569$，$p<0.001$）。任何时间窗口下，完整模型与其余 4 组模型相比结果都极其显著（Case(e) vs. Case(a)：all $p<0.001$；Case(e) vs. Case(b)：all $p<0.001$；Case(e) vs. Case(c)：all $p<0.001$；Case(e) vs. Case(d)：$p_1<0.001$，$p_2<0.05$）。在这 4 个模块中，缺少时间侧对模型结果影响最大，其次是自适应正则化方法，影响最小的是模板。

针对 Wearable 数据集，表 2 的 ANOVA 结果表明，不同模块对分类性能均有显著影响。在 0.5 s 时间窗下，组间差异的显著性检验结果为 $F(3,28)=892.417$，$p<0.001$；在 1.0 s 时间窗下，$F(3,28)=684.329$，$p<0.001$。这些结果进一步验证了模型架构的全局有效性。完整模型 Case(e)与所有消融模型相比均表现出显著优势（Case(e) vs. Case(a)：all $p<0.001$；Case(e) vs. Case(b)：all $p<0.001$；Case(e) vs. Case(c)：all $p<0.001$；Case(e) vs. Case(d)：$p_1<0.001$，$p_2<0.05$）。在该数据集上，尽管被试数量大幅增加，多域协同解码框架仍能维持强统计显著性，表明其对个体差异的鲁棒性较强。

表 2 消融实验结果

Table 2 Ablation experiment results %

分组	时空	频域	损失	模板	12-JFPM		Benchmark		Wearable	
					0.5 s	1.0 s	0.5 s	1.0 s	0.5 s	1.0 s
(a)		√	√	√	67.67±24.63***	83.00±17.92***	26.93±14.29***	45.87±22.18***	70.50±19.25***	86.41±11.63***
(b)	√		√	√	92.67±7.57***	98.20±2.13**	63.93±23.41***	88.73±21.63***	78.33±17.26***	92.33±9.26***
(c)	√	√		√	91.33±9.09***	98.00±2.21**	56.19±22.24***	78.67±27.52***	82.45±15.23***	94.72±8.53***
(d)	√	√	√		95.00±5.43***	99.00±1.53*	67.05±23.03***	90.33±18.50*	86.53±13.24***	97.26±2.31*
(e)	√	√	√	√	**97.33±2.49**	**99.33±1.32**	**74.19±16.73**	**90.95±17.15**	**90.12±10.42**	**97.56±2.11**

根据消融实验结果，模型的各个组件对分类性能有着显著影响。去除时空侧分支会导致性能明显下降，尤其是在短时间窗下，表明时空建模对于任务至关重要。而去除频域侧分支的影响相对较小，尤其在较长时间窗下，说明频域信息对该任务的贡献较为有限。去除自适应标签平滑损失函数会显著降低分类准确率，特别是在短时间窗下，说明该损失函数对于降低被试个体差异和防止过拟合具有重要作用。模板信号输入的缺失则会导致较大性能波动，特别是在0.5 s时间窗下，显示出模板信号在提供先验知识和引导分类方面的关键作用。最终，完整模型综合了所有模块的优势，在各时间窗下均展现了最优的性能，证明了这些设计在提升模型准确性方面的协同作用。

4.3 不同时间窗下的平均准确率

分别在0.5 s和1.0 s时间窗下，将多域联合解码模型与其他SSVEP分类算法进行比较，3个数据集上的分类结果如表3所示。

对于12-JFPM数据集，整体来看，多域联合解码模型在任何时间窗下都具有更高的分类准确率。在1.0 s时间窗下，与其余8种方法相比，差异均更显著（SSVEPNet：$p<0.05$；其他：$p<0.001$）。在0.5 s时间窗下，各个方法的分类准确率均有所下降，但多域联合解码模型仍然最优，且与对比方法均有显著差异（all $p<0.001$）。

对于Benchmark数据集，多域联合解码模型在1.0 s时间窗下的平均分类准确率是90%，SSVEPNet模型的平均分类准确率也是90%，根据事后配对t检验显示，分类结果无明显差异（$p>0.05$）。在0.5 s时间窗下，本文模型结果高于SSVEPNet模型，且具有显著优势（$p<0.05$）。与另外5种方法相比，在任何时间窗下，本文模型性能都具有极显著的优势（all $p<0.001$）。

对于Wearable数据集，在所有条件设置下，本文模型都更具优势，且差异显著（SSVEPNet：$p<0.05$；其他：$p<0.001$）。在0.5 s短时间窗条件下，本文模型的预测准确率达90.12%，较次优模型SSVEPNet提升2.5百分点，且标准差降低3.14%。在长时间窗场景下，本文模型达到97.56%的准确率，较最优传统方法DG-Conformer提升3.44百分点。

表3 不同时间窗下的平均准确率比较

Table 3 Comparison of average accuracy under different time windows %

模型	12-JFPM		Benchmark		Wearable	
	0.5 s	1.0 s	0.5 s	1.0 s	0.5 s	1.0 s
DG-Conformer[34]	89.33±14.32***	97.67±25.33***	57.76±16.59***	80.63±21.36***	83.56±18.29***	94.12±12.74***
C-CNN[28]	85.33±13.27***	90.67±13.15***	35.48±19.75***	56.38±25.03***	72.85±21.93***	85.42±20.16***
SSVEPformer[29]	87.67±16.13***	94.00±11.04***	46.33±26.19***	76.86±27.62***	80.36±19.88***	91.75±15.36***
FBtCNN[25]	81.33±16.41***	92.67±9.40***	37.10±24.52***	53.57±30.72***	67.92±24.62***	82.56±23.63***
SSVEPNet[20]	92.33±7.61***	98.20±2.13*	64.48±22.26*	90.33±17.15	87.62±13.56***	95.88±8.23*
AttentCNN[26]	88.35±15.86***	96.67±4.71***	55.76±24.39***	76.33±24.99***	79.25±18.12***	89.03±16.10***
TTCCA[35]	86.11±9.21***	95.00±5.24***	48.93±24.46***	82.68±19.68***	76.42±20.93***	88.36±17.54***
DDGCNN[22]	64.33±24.13***	75.33±22.47***	18.62±12.03***	32.00±19.10***	54.12±26.78***	67.32±24.57***
Ours	**97.33±2.49**	**99.33±1.32**	**74.19±16.73**	**90.95±17.15**	**90.12±10.42**	**97.56±2.11**

表3比较了0.5 s和1.0 s时间窗下多个模型的分类准确率。结果表明，1.0 s时间窗下整体准确率较高，模型能够更好地捕捉信号的长期依赖性。尽管如此，本文提出的模型在2个时间窗下均表现优异，分类准确率领先其他方法，尤其在1.0 s时间窗下进一步拉开了差距，展现了其在时序信号分类中的优势。因此，无论是在短时间窗还是长时间窗下，本文模型都具有较强的适应性和优越性，表明自适应标签平滑技术通过动态调节噪声抑制强度，有效缓解了可穿戴设备信号的高变异特性。

4.4 不同时间窗下的单被试准确率

模型在12-JFPM数据集上对单个个体的分类效果如表4、表5所示。从平均分类准确率来看，模型整体优于对比方法，尤其是在0.5 s时间窗口下，与对比方法差距更大，优势更明显。对于10个被试来说，被试S1、S2、S3、S9、S10的性能明显弱于其余被试，可能是由于这些被试在采集数据过程中眼睑震颤过于剧烈或注意力不集中导致。本文模型对于这些较弱被试数据的分类明显优于其余模型，尤其是对于被试S1，在0.5 s时间窗下，本文模型分类准确率达到96.67%，比其余8种方法的平均准确率高33百分点。

在0.5 s时间窗下，各模型的准确率呈现出一定的波动，且表现差异较大。本文模型在多个被试中的表现显著优于其他方法。虽然其他模型（如SSVEPNet和AttentCNN）也在部分被试中表现出色，但本文模型在大多数被试中表现出了更为稳定

和优越的分类能力，尤其是在被试 S1 和 S2 等较低准确率的情况下，表现仍然突出。在 1.0 s 时间窗下，整体准确率普遍有所提升，特别是在 S5、S6、S7、S8 等被试中，所有模型的表现均达到了较高水平。然而，即便在这种较长时间窗下，本文模型依然具有优势，在除 S9 外的所有被试中达到最高准确率。总体来看，尽管不同模型在不同被试中表现各异，但本文模型在 2 个时间窗下均展现出了较为稳定和较高的准确率，尤其在较长时间窗下，分类准确率更高。这表明，在处理时序信号任务时，较长时间窗能够更好地体现模型能力，尤其是在捕捉信号中的长期依赖关系方面。

表 4　0.5 s 时间窗下单被试准确率

Table 4　Individual subject accuracy in a 0.5 s time window　%

被试	DG-Conformer	C-CNN	SSVEPformer	FBtCNN	SSVEPNet	AttentCNN	TTCCA	DDGCNN	Ours
S1	85.33	73.33	80.00	66.67	90.00	86.67	60.26	53.33	**96.67**
S2	85.33	70.00	73.33	50.00	85.33	60.00	64.32	26.67	**93.33**
S3	87.33	90.00	93.33	73.33	90.00	**100.00**	93.33	73.33	**100.00**
S4	96.67	96.67	96.67	93.33	96.67	**100.00**	96.67	70.00	**100.00**
S5	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	80.00	**100.00**
S6	**100.00**	**100.00**	**100.00**	96.67	96.67	**100.00**	**100.00**	83.33	93.33
S7	96.67	96.67	**100.00**	96.67	**100.00**	**100.00**	**100.00**	96.67	**100.00**
S8	86.67	86.67	**96.67**	93.33	**96.67**	93.30	**96.67**	90.00	**96.67**
S9	86.67	80.00	90.00	80.00	93.33	86.67	89.43	43.33	**96.67**
S10	68.63	60.00	46.67	63.33	73.33	56.67	60.32	26.67	**96.67**

表 5　1.0 s 时间窗下单被试准确率

Table 5　Individual subject accuracy in a 1.0 s time window　%

被试	DG-Conformer	C-CNN	SSVEPformer	FBtCNN	SSVEPNet	AttentCNN	TTCCA	DDGCNN	Ours
S1	98.03	90.00	86.67	76.67	**100.00**	86.67	87.58	76.67	**100.00**
S2	56.67	56.67	63.33	83.33	**96.67**	90.00	73.42	30.00	**96.67**
S3	95.33	76.67	96.67	76.67	**100.00**	93.33	95.67	83.33	**100.00**
S4	**100.00**	**100.00**	**100.00**	96.67	**100.00**	**100.00**	**100.00**	86.67	**100.00**
S5	96.67	96.67	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	73.33	**100.00**
S6	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	96.67	**100.00**
S7	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
S8	**100.00**	93.33	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	96.67	**100.00**
S9	**100.00**	96.67	**100.00**	**100.00**	93.33	96.67	**100.00**	70.00	96.67
S10	96.67	96.67	93.33	93.33	**100.00**	**100.00**	93.33	40.00	**100.00**

4.5　不同时间窗下的平均信息传输率

分类信息传输率比较如表 6 所示。根据 3 个数据集在 0.5 s 和 1.0 s 时间窗下的 ITR 比较结果可知，本文方法在所有时间窗下均表现优异，领先于其他 8 种方法。例如，在 12-JFPM 数据集的 0.5 s 时间窗下，多域联合模型的 ITR 达到 425 bit/min，较第二名 SSVEPNet 提升近 9%，这一差距在 Benchmark 这种刺激类别多、噪声大的数据集中进一步扩大，多域联合模型达到 310 bit/min，SSVEPNet 是 280 bit/min。这种优势可能是因为多域模型对时空频特征的联合建模能力，通过动态调节软硬标签权重机制增强了高噪声环境下的信号鲁棒性。

时间窗长度对 ITR 也有一定影响。以 SSVEPNet 为例，在 Benchmark 数据集上，时间窗从 0.5 s 延长至 1.0 s 时，ITR 从 280 bit/min 变成 420 bit/min，增加了 50%。这是因为长时间窗口对信号周期性的充分学习，弥补了噪声干扰带来的负面影响。然而，在 12-JFPM 这种高质量数据集中，时间窗增加带来的信息传输率增加相对有限。例如，多域模型从 425 bit/min 上升到 455 bit/min，仅提高 7%。

数据集对信息传输率的影响也十分明显。在 Wearable 这类多被试数据中，多域模型在 2 个时间窗下的 ITR 分别达到 410 bit/min 和 450 bit/min，表明本文方法对噪声的抑制能力较强。对于 Benchmark 数据集，FBtCNN 的 ITR 普遍低于 260 bit/min，而多域模型仍能维持 430 bit/min 的

高性能，说明其具有泛化能力。此外，DDGCNN 在所有场景中表现最差，可能与其稀疏连接设计和梯度消失问题相关，这为未来模型轻量化设计提供了改进方向。

表 6 不同时间窗下的信息传输率比较

Table 6 Comparison of information transmission rates under different time windows 单位：bit/min

模型	12-JFPM		Benchmark		Wearable	
	0.5 s	1.0 s	0.5 s	1.0 s	0.5 s	1.0 s
DG-Conformer	365.52±12.15	420.41±11.72	220.44±21.22	340.78±20.14	340.71±15.74	415.86±25.03
C-CNN	340.52±15.23	385.42±13.45	150.23±25.63	230.56±24.68	280.46±19.63	370.47±13.78
SSVEPformer	355.85±13.42	410.12±10.25	190.43±23.78	320.78±22.57	324.65±17.45	405.23±15.25
FBtCNN	315.63±16.73	400.53±11.24	160.41±28.22	250.79±27.74	260.75±23.12	350.47±19.72
SSVEPNet	390.41±8.13	435.63±11.83	280.42±16.14	420.52±15.75	375.89±12.16	440.23±7.35
AttentCNN	360.57±12.34	415.26±13.25	210.42±21.59	310.57±20.54	315.25±17.49	390.58±14.59
TTCCA	350.74±14.65	405.85±17.54	200.47±24.38	335.23±17.12	300.15±19.49	380.51±13.52
DDGCNN	245.12±28.16	320.75±24.67	80.63±27.43	140.52±21.23	200.06±25.03	270.14±20.12
Ours	**425.74±6.26**	**455.27±1.56**	**310.77±18.64**	**430.83±15.73**	**410.69±10.13**	**450.65±6.75**

4.6 自适应机制对准确率的影响

本文提出的自适应标签平滑技术对 12-JFPM 数据集中每个被试的影响如图 3 所示。其中，每条线段代表不同被试，圆形标记指 α 取固定值 0.5，即软硬标签对所有被试均产生相同的影响，方形标记指自适应地产生 α，图 3 中给出每个被试分别对应的 α 值以及该被试的分类准确率。如果方形标记分布在圆形标记上方，表示训练该被试时硬标签占比大于软标签，即该被试受目标刺激周围的非目标刺激影响小；如果方形标记分布在圆形标记下方，表示训练该被试时，软标签占比大于硬标签，即该被试受目标刺激周围的非目标刺激影响大。图中短线的方向表示准确率的增减，斜率表示增减的幅度大小，其均呈现从左到右的趋势，这表明加入代表生物特性的因子后，所有被试的分类准确率均有不同程度的提升，尤其是对在固定比例情况下准确率较低的被试而言，准确率提升尤为明显。

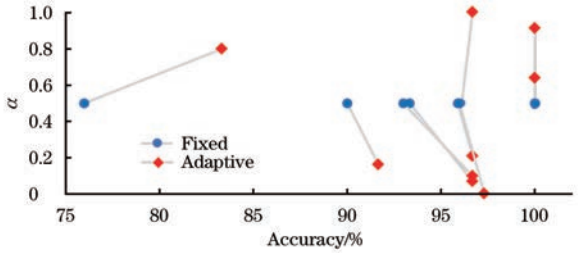

图 3 单个被试的 α 和准确率的关系

Fig.3 Relationship between α and accuracy for a single subject

为了探索自适应方法对每个被试的影响，在 12-JFPM 的 10 个被试上进行 4 组实验，前 3 组分别将 α 固定为 0（全使用软标签）、0.5（一半软标签和一半硬标签）和 1（全使用硬标签），第 4 组则根据被试特性自动生成。实验结果如图 4 所示，图中条柱依次表示这 4 组条件下的分类结果。

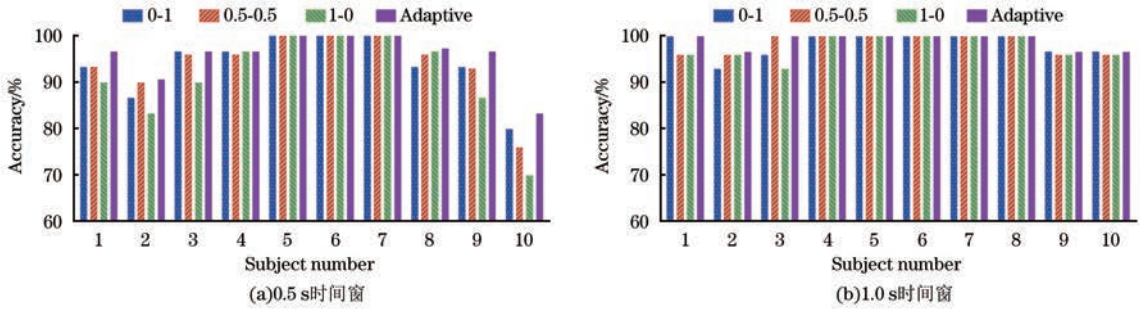

图 4 α 值对不同被试的影响

Fig.4 The effect of α values on different subjects

从图 4 中可见，对于任何时间窗下的任何被试，自适应组的分类精度都不低于其余 3 组，这表明该方法具有普遍有效性。对于被试 5、6、7 来说，在不同时间窗口下都能达到 100%的分类精度，证明本身信号质量较高。随着时间窗长度的增加，被试 4 和 8 的 4 组实验也能达到 100%的分类准确率，其余各被试的 4 组实验之间的精度差也在缩小。这表明数据量较少时特征不明显，生物特征干扰较大，所提自适应方法的效果也越明显。在较长的 1.0 s 时间窗口下，被试 1 和 3 的前 3 组实验中均有某一组

特别突出，达到 100%，远高于其余 2 组，与最后的自适应结果一致，表明自适应方法选到了最适合该被试的 α 取值。除此之外的其余被试的 4 组实验结果基本相等，没有明显差异(all：$p>0.05$)，侧面印证了本文所提自适应的、基于注意力的标签平滑技术在数据量较少的情况下更具优势。

4.7 训练量对准确率的影响

为了评估训练量对模型性能的影响，对 12-JFPM 数据集的前 5 个被试分别进行两种数据削减的实验：第 1 种是减少输入的数据长度，原始输入对应的是 4 s 的数据长度，每减少 1/4 代表减少 1 s 数据长度，即“−25%”表示 3 s 数据长度，“−50%”表示 2 s 数据长度，“−75%”表示 1 s 数据长度；第 2 种是减少训练试次，输入的原始数据的训练集和测试集的划分比例为 8∶2，每降低 1/4 表示训练集减少两份，即“−25%”表示比例为 6∶4，“−50%”表示比例为 4∶6，“−75%”表示比例为 2∶8。这两类实验在 0.5 s(方形标记)和 1.0 s(菱形标记)时间窗下均进行了测试，实验结果如图 5 所示。在图 5 中，缩短数据长度带来的变化并不明显，图中代表减少数据长度实验的实线较为平缓，几乎持平。在每个时间窗口下，将数据长度缩小一定比例，分类精度虽有微弱变化，但无明显差异(all：$p>0.05$)。缩短训练试次对模型有明显影响，图中代表减少训练试次的虚线十分陡峭，减少的比例越多，精度下降越明显。在同一时间窗口下，减少 25%和减少 50%的训练试次都会使分类精度降低，且差异显著(all：$p<0.05$)。特别地，将训练试次缩小 75%后，分类精度大幅降低，且具有显著差异(all：$p<0.001$)。

从图 5 中观察到，无论在哪个时间窗口下，本文提出的模型更能接受数据长度的减少，减少数据长度的分类精度明显高于减少训练试次的分类精度，即同一降低比例下，减少训练试次对模型精度的影响高于减少数据长度。此外，在降低相同的比例时，较短的 0.5 s 时间窗下 2 种方法的精度差要高于 1.0 s 时间窗；在同一时间窗下，减少高数据比例时 2 种方法之间的精度差要高于减少低数据比例。

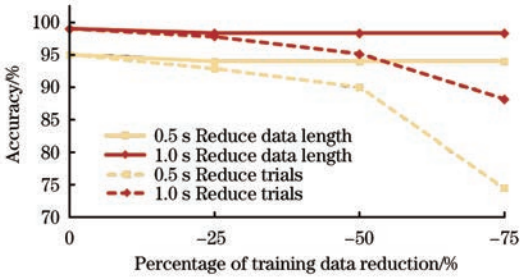

图 5 不同训练量下的平均分类准确率

Fig. 5 Average classification accuracy under different training amounts

考虑到分类模型高度依赖输入训练试次，在 12-JFPM 数据集上，按照不同比例划分输入数据，实验结果如表 7 和表 8 所示。随着输入训练试次的减少，每种方法分类准确率都有所下降，尤其是在时间窗较短的 0.5 s 窗口下，训练试次减少所带来的性能影响更大。但在同一划分比例下，本文模型分类准确率更高，表明该模型在数据量较少时同样具有优势。其中，本文方法和 SSVEPNet 模型受影响较轻，C-CNN 和 FBtCNN 受影响较重。在时间窗长度为 1.0 s、划分比例为 8∶2的情况下，本文模型与 SSVEPNet 模型相比，分类准确率都达到 99%以上，没有明显差异。在其余情况下，本文模型的分类准确率略高于 SSVEPNet 模型，结果具有显著性。与 DG-Conformer、C-CNN、FBtCNN、AttentCNN、SSVEPformer、TTCCA、DDGCNN 等 7 种方法相比，本文模型总是具有优势，且根据事后配对 t 检验显示，本文模型分类结果与这 8 种模型结果在 0.5 s 和 1.0 s 时间窗下均差异显著(all：$p<0.001$)。

表 7 不同训练量下的平均准确率比较(0.5 s)

Table 7 Comparison of average accuracy under different training amounts (0.5 s) %

比例	DG-Conformer	C-CNN	SSVEPformer	FBtCNN	SSVEPNet	AttentCNN	TTCCA	DDGCNN	Ours
2∶8	70.53	47.58	59.58	43.67	79.00	64.75	65.12	45.83	**94.75**
4∶6	79.63	69.89	72.78	66.65	90.33	78.56	75.63	53.33	**95.78**
6∶4	85.33	77.83	78.83	76.50	91.79	86.33	84.23	57.50	**96.33**
8∶2	89.33	85.33	87.67	81.33	92.33	88.33	86.11	64.33	**97.33**

表 8 不同训练量下的平均准确率比较(1.0 s)

Table 8 Comparison of average accuracy under different training amounts (1.0 s) %

比例	DG-Conformer	C-CNN	SSVEPformer	FBtCNN	SSVEPNet	AttentCNN	TTCCA	DDGCNN	Ours
2∶8	78.52	48.00	74.17	59.17	87.75	75.17	79.55	58.92	**89.58**
4∶6	87.56	71.78	85.22	80.33	95.33	88.56	89.65	70.00	**97.00**
6∶4	94.63	84.00	89.50	87.50	96.00	95.00	94.32	72.00	**98.83**
8∶2	97.67	90.67	94.00	92.67	**99.00**	96.67	95.00	75.33	**99.00**

4.8 层级特征可视化分析

为解析特征提取网络的层级学习机制，本文采用 T-SNE 降维技术对关键中间层特征进行可视化，结果如图 6 所示。以 12-JFPM 数据集中的被试 S1 为例，不同颜色分别代表 12 个刺激目标，4 张图分别代表对输入网络的原始 SSVEP 信号、时空滤波后、特征融合后及分类后的特征进行可视化的结果。使用滑动窗口来生成测试样本，有利于观测不同频率之间是否被区分。在图 6(a)中，各种颜色随机混合分布，表明原始信号的不同频率之间高度重叠，难以区分。经时空特征提取后，图 6(b)中相同颜色逐渐聚合成堆，各频率之间差距缩小。待多域特征混合后，图 6(c)的不同颜色堆逐渐线性分离，说明各频率之间差距扩大。最终图 6(d)的分类特征形成离散条带状分布，实现了高效的类间分离。

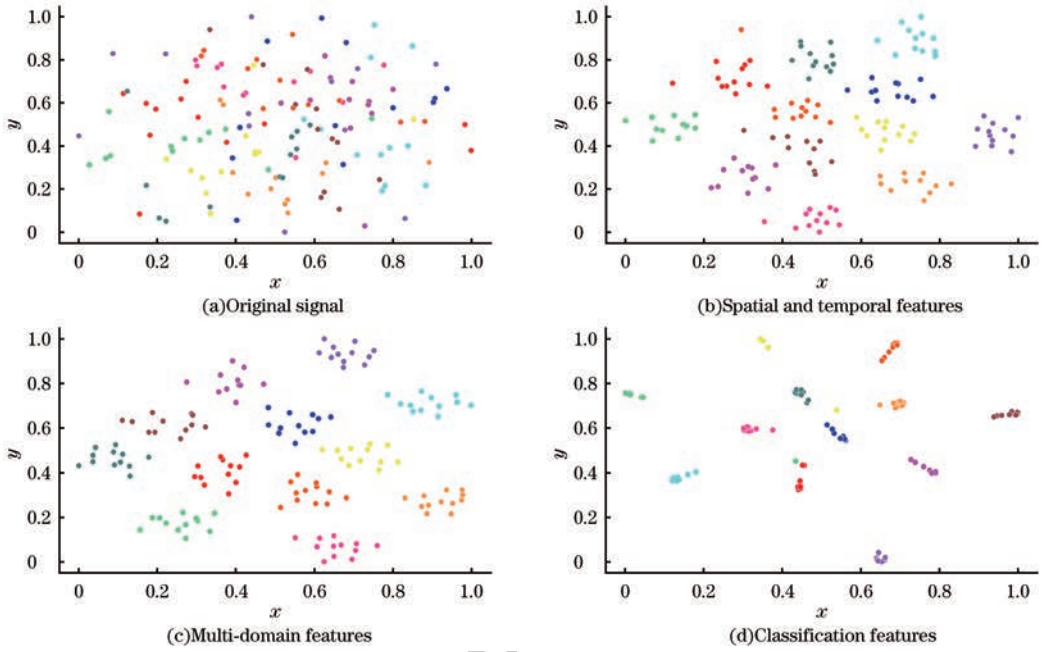

图 6 T-SNE 示意图

Fig. 6 Schematic diagram of T-SNE

5 结束语

本文提出了一种基于视觉分散自适应补偿的 SSVEP 多域协同解码算法，通过构建视觉分散量化模型与自适应标签平滑技术的协同机制，有效解决了高个体差异与非目标刺激干扰导致的解码性能瓶颈问题。本文的核心创新在于将视觉拥挤效应理论融入标签平滑过程，动态调节平滑强度，在抑制非目标神经干扰的同时缓解模型过拟合问题。同时，提出时频空三域特征协同解码框架，该算法在 3 个数据集的任何条件设置下均表现出显著优势。通过消融实验，验证了各个特征提取分支均能发挥重要作用，其中自适应补偿机制对视觉分散敏感被试的优化效果尤为显著，准确率最高可提升 18 百分点。该算法不仅为提升基于 SSVEP 的 BCI 解码性能提供了新的解决方案，也为解码过程中生物特征干扰的量化剔除开辟了新的研究思路。

本文提出的自适应的、基于注意力机制的标签平滑技术，仅针对被试内分类任务，后续考虑将其拓展到被试间的 SSVEP 分类任务。此外，使用生成网络生成脑电数据，实现数据增强，也是未来的研究方向。

参考文献

[1] ZHUANG M M, WU Q H, WAN F, et al. State-of-the-art non-invasive brain-computer interface for neural rehabilitation: a review[J]. Journal of Neurorestoratology, 2020, 8(1): 12-25.

[2] CHEN X, WANG Y, GAO S. EEG-based brain-computer interfaces: a review of recent progress in signal processing and applications[J]. IEEE Transactions on Cognitive and Developmental Systems, 2022, 14(3): 732-748.

[3] VEENA N, ANITHA N. A review of non-invasive BCI devices[J]. International Journal of Biomedical Engineering and Technology, 2020, 34(3): 205.

[4] NAKANISHI M, WANG Y, CHEN X. High-speed SSVEP-based BCI with filter bank canonical correlation analysis[J]. Journal of Neural Engineering, 2021, 18(4): 046028.

[5] 李明, 王伟, 张涛. 基于稳态视觉诱发电位的脑机接口系统设计与医疗应用[J]. 中国生物医学工程学报, 2021, 40(4): 432-439.
LI M, WANG W, ZHANG T. Design and medical applications of SSVEP-based brain-computer interface systems[J]. Chinese Journal of Biomedical Engineering,

2021, 40(4): 432-439. (in Chinese)
[6] 周晓燕, 刘洋, 高小榕. 可穿戴脑电设备在军事脑机接口中的最新进展[J]. 仪器仪表学报, 2023, 44(2): 156-165.
ZHOU X Y, LIU Y, GAO X R. Recent advances in wearable EEG devices for military brain-computer interfaces[J]. Chinese Journal of Scientific Instrument, 2023, 44(2): 156-165. (in Chinese)
[7] MÜLLER P G. Benchmarking brain-computer interfaces outside the laboratory: a case study with SSVEP-based gaming[J]. Frontiers in Human Neuroscience, 2022, 16: 812345.
[8] ZHANG Y. Enhancing SSVEP detection through cortical source reconstruction[J]. NeuroImage, 2018, 183: 897-908.
[9] WANG Y, JUNG T P. A hybrid EEG-fNIRS approach for robust SSVEP classification [J]. Journal of Neural Engineering, 2020, 17(3): 036028.
[10] CHEN X. Deep learning-based spatial filtering for high-speed SSVEP-BCI [J]. IEEE Transactions on Biomedical Engineering, 2021, 68(4): 1123-1132.
[11] ZHANG Q, WANG L, CHEN Z, et al. Cross-subject fluctuation analysis in SSVEP-based BCIs[J]. Journal of Neural Engineering, 2021, 18(4): 0460a3.
[12] CHEN X, LI H, WANG Y. Activation mechanisms of parieto-occipital cortex during SSVEP generation [J]. NeuroImage, 2021, 237: 118143.
[13] ROSENHOLTZ R. A unified model of crowding and visual search[J]. Vision Research, 2020, 173: 1-12.
[14] VAN DEN BERG R, ROERDINK J, CORNELISSEN F. FFT amplitude analysis of crowding effects in visual perception[J]. PLoS Computational Biology, 2020, 16(4): e1007842.
[15] NAKANISHI M, TANAKA T, WANG Y. Unsupervised frequency-recognition method of SSVEPs using a filter bank implementation of binary subband CCA [J]. IEEE Transactions on Biomedical Engineering, 2019, 66(3): 725-735.
[16] GUNEY O B, OBLOKULOV M. A deep neural network for SSVEP-based brain-computer interfaces [J]. IEEE Transactions on Biomedical Engineering, 2022, 69(2): 932-944.
[17] PAN Y D, LI N. Short-time SSVEP data extension by a novel generative adversarial networks based framework[EB/OL]. [2024-08-05]. https://www.researchgate.net/publication/367166060_Short-time_SSVEP_data_extension_by_a_novel_generative_adversarial_networks_based_framework.
[18] MING D, LI R, GAO X. SSVEP-GAN: subject-invariant feature learning via adversarial neural style transfer [J]. Medical Image Analysis, 2023, 88: 102856.
[19] PAN Y D, ZHANG Y. A survey of deep learning-based classification methods for steady-state visual evoked potentials[J]. Brain-Apparatus Communication: A Journal of Bacomics, 2023, 2(1): 2181102.
[20] PAN Y, CHEN J, ZHANG Y, et al. An efficient CNN-LSTM network with spectral normalization and label smoothing technologies for SSVEP frequency recognition[J]. Journal of Neural Engineering, 2022, 19(5): 102797.
[21] LIU Y, CHEN X, WANG Y. Deep frequency-domain residual network with multi-scale spectral pyramid for high-accuracy SSVEP decoding [J]. Journal of Neural Engineering, 2023, 20(3): 450-458.
[22] ZHANG S, AN D, LIU J, et al. Dynamic decomposition graph convolutional neural network for SSVEP-based brain-computer interface[J]. Neural Networks: the Official Journal of the International Neural Network Society, 2023, 172: 106075.
[23] 陈小刚, 李航, 王毅军. 基于时空卷积注意力网络的 SSVEP 解码方法研究[J]. 自动化学报, 2023, 49(5): 1021-1032.
CHEN X G, LI H, WANG Y J. Spatio-temporal convolutional attention network for SSVEP decoding [J]. Acta Automatica Sinica, 2023, 49(5): 1021-1032. (in Chinese)
[24] ZHANG Z, LI D D, ZHAO Y, et al. A flexible speller based on time-space frequency conversion SSVEP stimulation paradigm under dry electrode[J]. Frontiers in Computational Neuroscience, 2023, 17: 1101726.
[25] DING W, SHAN J, FANG B, et al. Filter bank convolutional neural network for short time-window steady-state visual evoked potential classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29: 2615-2624.
[26] CHEN J N, SUN F C. Attention-based multimodal CNN for classification of steady-state visual evoked potentials and its application to gripper control [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 21(2): 1124-1138.
[27] 明建国, 张伟, 刘洋. 混合神经网络在高速脑机接口中的应用[J]. 中国生物医学工程学报, 2024, 43(2): 156-165.
MING J G, ZHANG W, LIU Y. Application of hybrid neural networks in high-speed brain-computer interfaces[J]. Chinese Journal of Biomedical Engineering, 2024, 43(2): 156-165. (in Chinese)
[28] RAVI A, BENI N H, MANUEL J, et al. Comparing user-dependent and user-independent training of CNN for SSVEP BCI[J]. Journal of Neural Engineering, 2020, 17(2): 046080.
[29] CHEN J B, ZHANG Y S, PAN Y D, et al. A Transformer-based deep neural network model for SSVEP classification[J]. Neural Networks, 2023, 164: 521-534.
[30] NAKANISHI M, TANAKA T, WANG Y. DeepFreqNet: a multi-scale frequency-domain deep network for SSVEP classification [J]. IEEE Transactions on Biomedical Engineering, 2021, 68(12): 3565-3574.
[31] NAKANISHI M, WANG Y J. A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials[J]. PLOS ONE, 2015, 10(10): 18.
[32] WANG Y J, CHEN X G, GAO X R, et al. A benchmark dataset for SSVEP-based brain-computer interfaces[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017, 25(10): 1746-1752.
[33] ZHU F, JIANG L, DONG G, et al. An open dataset for wearable SSVEP-based brain-computer interfaces [J]. Sensors, 2021, 21(3): 1256.
[34] LIU J W, WANG R M, YANG Y K, et al. Convolutional Transformer-based cross subject model for SSVEP-based BCI classification[J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(11): 6581-6593.
[35] QIN K, XU R, LI S R, et al. A time-local weighted transformation recognition framework for steady state visual evoked potentials based brain-computer interfaces[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024, 32(2): 1596-1605.

文字编辑 吴云芳
栏目编辑 赖玉玲