

面向运动想象解码的双尺度时空特征融合网络

陈志城

五邑大学电子与信息工程学院, 广东 江门

收稿日期: 2026年1月21日; 录用日期: 2026年2月20日; 发布日期: 2026年2月26日

摘要

卷积神经网络(Convolutional neural networks, CNNs)已在运动想象(Motor Imagery, MI)脑机接口(Brain-Computer Interface, BCI)领域取得了广泛应用和良好效果。然而, 基于单一尺度卷积结构的模型难以从脑电信号中充分挖掘多样化信息; 而现有的多尺度CNN虽然能够提取不同尺度的特征, 但通常仅通过简单拼接的方式进行融合, 难以实现多尺度信息之间的深度协同建模。为了解决上述问题, 相关研究提出了双尺度时空特征融合网络(Dual-Scale Spatiotemporal Feature Fusion Network, DSSFFN)。该方法通过双尺度卷积分支以提取脑电信号不同尺度的特征, 并通过Transformer模块对来自不同尺度的特征进行融合, 从而提升特征的判别性。本文采用BCI竞赛IV的2a数据集进行实验, 同时与多个前沿的运动想象算法进行比较。结果显示, DSSFFN在数据集上平均准确率为72.65%, 高于所有的对比模型, 展现出了DSSFFN模型在运动想象脑电识别任务上的优异性能。此外, 研究还通过消融实验分析了双尺度时空卷积分支与Transformer模块对整体性能的贡献, 验证了各个关键模块的有效性。同时比较了不同融合模型对模型影响, 进一步验证了Transformer模块融合双尺度的有效性。

关键词

脑电信号, 运动想象, 卷积神经网络, 脑机接口

Dual-Scale Spatiotemporal Feature Fusion Network for Motor Imagery Decoding

Zhicheng Chen

School of Electronics and Information Engineering, Wuyi University, Jiangmen Guangdong

Received: January 21, 2026; accepted: February 20, 2026; published: February 26, 2026

Abstract

Convolutional neural networks (CNNs) have been widely applied in motor imagery (MI) brain-computer interface (BCI) systems and have achieved promising performance. However, models based

on single-scale convolutional structures are often insufficient to fully exploit the diverse and rich information embedded in EEG signals. Although existing multi-scale CNNs can extract features at different scales, they usually fuse multi-scale features through simple concatenation, which makes it difficult to achieve deep and effective collaboration among features from different scales. To address these issues, related research proposes a Dual-Scale Spatiotemporal Feature Fusion Network (DSSFFN). This method employs dual-scale convolutional branches to extract EEG features at different scales, and further integrates the features from different branches using a Transformer module, thereby enhancing the discriminative capability of the feature representations. In this study, experiments are conducted on the BCI Competition IV-2a dataset, and the proposed method is compared with several state-of-the-art MI recognition algorithms. The results show that DSSFFN achieves an average classification accuracy of 72.65% on this dataset, outperforming all competing methods, thus demonstrating its superior performance in motor imagery EEG classification tasks. Furthermore, ablation studies are performed to analyze the contributions of the dual-scale spatiotemporal convolution branches and the Transformer fusion module, which verifies the effectiveness of each key component in the proposed network. In addition, the impacts of different fusion models on the proposed model's performance were compared, which further verifies the effectiveness of the Transformer module in fusing dual-scale features.

Keywords

Electroencephalogram, Motor Imagery, Convolutional Neural Network, Brain-Computer Interface

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑机接口(BCI)系统可通过采集大脑的神经响应信号,识别不同模式的脑部活动,并生成各类控制信号或指令,从而实现大脑与外部设备之间的直接通信与操控[1][2]。当前记录脑神经活动的技术众多,其中脑电图(Electroencephalogram, EEG)技术凭借无创性、低成本、便携性及高时间分辨率等核心优势,成为BCI领域的主流数据采集方式。运动想象是BCI经典范式之一,本质是运动行为的意念模拟过程,不涉及肢体实际动作的神经活动模式。运动想象BCI(MI-BCI)已在医疗辅助领域实现规模化应用,例如作为残障人士的交互工具,助力其完成对周边环境的指令操控。同时,该技术在虚拟现实、游戏娱乐及机械臂控制等非医疗场景中也展现出巨大应用潜力。然而,EEG信号本身存在低信噪比、非平稳性等固有缺陷,再叠加个体生理差异显著等现实问题,给运动想象EEG的解码工作带来了诸多技术挑战。

在EEG信号解码研究早期,研究人员普遍采用神经生理学先验知识与传统机器学习分类算法相结合的方案,开展运动想象脑电信号(MI-EEG)解码工作。其中,共空间模式(common spatial pattern, CSP)[3]是应用广泛的运动想象特征提取方法,核心思路为通过构建空间滤波器提取最具判别力的空间特征,最大化不同类别信号间的差异。此后,基于CSP的衍生方法被提出(如FBCSP[4]、R-CSP[5]),进一步强化了特征的判别性能。最终,通过将支持向量机、线性判别分析等分类算法与上述特征提取方法结合,完成对目标特征的分类。传统机器学习方法高度依赖领域专家知识,这一局限性制约了其分类性能的提升。此外,机器学习在复杂数据和非线性建模方面的能力有限。

为了解决这些问题,近年来深度学习被广泛应用在脑机接口领域。深度学习是一种的端到端建模框架,能够直接从原始脑电信号中自动学习判别性特征,并将特征提取与分类过程统一。在众多深度学习

模型中, 卷积神经网络(Convolutional neural networks, CNNs)凭借其优越的特征表示能力, 能够同时建模脑电信号的时间与空间结构信息, 因此已成为基于运动想象脑机接口(MI-BCI)系统的核心技术之一。Schirrneister 等人[6]设计了两种具有代表性的卷积网络结构: 浅层结构 ShallowConvNet 与深层结构 DeepConvNet, 并在 MI 任务中验证了二者良好的解码性能。随后, Lawhern 等人[7]提出了 EEGNet, 该模型通过引入深度可分离卷积和逐点卷积, 在显著降低网络参数规模的同时, 依然保持了较高的分类精度, 从而提高了模型的效率与泛化能力。在此基础上, 研究者们不断尝试通过结构改进与特征建模机制增强模型性能。例如, EEG-TCNet [8]将 EEGNet 与时间卷积网络(Temporal Convolution Network, TCN)相结合, 利用 TCN 的大感受野特性加强对长时间序列信息的建模能力, 从而进一步提升了分类精度。Mane 等人[8]提出了基于滤波器组 CSP 的 FBCNet, 其中 CNN 从多频段信号中提取信息, 对 MI 任务进行分类。此外, 为了进一步提高 MI 解码的准确性, Altaheri 等人[9]提出了 ATCNet, 该模型融合了注意力机制、时间卷积模块以及滑动窗口策略, 在特征建模能力方面取得了显著提升。在此基础上, D-ATCNet [10]通过引入动态卷积和多层级注意力机制, 对网络结构进行了进一步优化, 从而实现了更高的分类性能。Zhao 等人[11]提出 CTNet, 融合卷积模块和 Transformer 模块以提取局部全局特征。

尽管基于单尺度卷积结构的卷积神经网络在运动想象脑机接口任务中已取得一定进展, 但由于脑电信号在时间演化、频谱分布以及空间拓扑结构等方面均呈现出显著的多尺度特性, 固定感受野的卷积算子往往难以同时兼顾不同尺度特征模式的有效建模, 从而在复杂脑电模式刻画能力方面仍存在不足。

针对上述问题, 近年来研究者逐渐将多尺度建模思想引入脑电解码网络结构设计中, 通过构建包含不同尺寸卷积核或不同时间感受范围的并行分支, 以实现多时间尺度、多频段特征信息的联合建模, 从而提升网络对复杂脑电信号结构的表征能力。例如, Dai 等人[12]提出了 HS-CNN 结构, 以缓解不同被试在特征尺度分布上的个体差异问题; Ko 等人[13]设计了 MSNN, 从多个时间与频率尺度提取特征表示, 并进一步刻画电极之间的空间相关性, 从而获得更具判别力的特征表达; Salami 等人[14]提出的 ITNet 结合 Inception 模块与扩张因果卷积结构, 能够同时捕获多通道脑电信号中的时序、频谱及空间信息; 此外, Santamaria-Vazquez 等人[15]提出的 EEG-Inception 通过引入不同尺度卷积核分支, 有效提升了脑电信号的解码性能。单尺度 CNN 在脑电信号处理中存在信息提取局限。传统多尺度 CNN 却忽视了不同尺度间信息的有机整合, 导致特征利用不充分。

针对这些局限性, 本文设计了一种双尺度时空特征融合网络(Dual-Scale Spatiotemporal Feature Fusion Network, DSSFFN)。该网络可同步提取脑电信号在不同尺度下的时间特征与空间特征, 并通过有效的融合机制整合多尺度信息。主要贡献包括: 我们提出了一种新的基于 Transformer 的双尺度时空特征融合神经网络来识别运动想象脑 EEG 信号, 该网络捕获不同尺度的时间和空间信息, 从而充分挖掘运动想象脑电信号的丰富表征; 为进一步挖掘多尺度融合特征的潜在价值, 引入 Transformer 实现双尺度信息的深度耦合, 进而生成判别力更强的特征。

2. 实验设置

2.1. 数据集描述

本研究采用的数据集为 BCI 竞赛 IV 的 2a 数据集。该数据集包含 9 名受试者, 所有受试者均需完成左手、右手、足部、舌头四类运动想象任务。数据集中每位受试者 j 进行了两个独立测试场次, 两场测试在不同日期, 22 个 Ag/AgCl 电极在 250 Hz 采样。每个场次包含 288 次实验试次, 四类任务各分配 72 次试次, 试次分布均衡。原始数据集共设 25 个数据通道, 其中 3 个为眼电(EOG)通道, 22 个为脑电(EEG)通道, 本研究选用了这 22 个脑电通道的有效数据进行后续分析。在实验中, 我们将 2 个场次合并共 576 个样本, 取 2~5 秒的数据, 并采用带通滤波器将脑电数据过滤到 8~30 Hz, 随后进行归一化。

2.2. 实验环境设置

本研究中，所提模型的训练阶段采用交叉熵损失函数，并使用 Adam 优化器，学习率为 0.0001，同时将批量大小固定为 16。对比模型的训练参数遵循其原始文献中的设定。考虑到 EEG-Inception 与 EEGNet 模型的设计最初是匹配 128 Hz 采样率的数据，而本研究采用的脑电数据采样率为 250 Hz，因此参考研究 [8]，我们将这两个模型的时间卷积核长度与时间池化层尺寸均放大 2 倍，以实现近似匹配 250 Hz 采样率。本研究所有模型采用五折交叉验证方法进行性能评估，训练总轮数设为 200。实验基于 PyTorch 深度学习框架，在单张 GeForce RTX 4060 GPU 上执行。

2.3. 对比模型

Shallow ConvNet (ShallowNet): ShallowNet 是专门设计用于对振荡信号进行分类的，它能够对 MI-EEG 进行分类，仅需要两个卷积层分别执行时间卷积和空间滤波，以及一个平均池化层。

EEGNet: EEGNet 是一种轻量级网络模型。它沿时间维度进行时间卷积，沿空间维度进行深度卷积，最后通过点向卷积混合特征映射。

EEGInception: EEG-Inception 使用 Inception 模块提取多尺度特征，在保持模型轻量级的架构同时实现卓越的性能。

ITNet: ITNet 融合 Inception 模块和因果卷积兼顾多尺度时空特征提取与低复杂度，同时提升了分类性能。

CTNet: CTNet 是一种卷积与 Transformer 融合的混合网络模型。它通过卷积模块提取脑电信号的局部与空间特征，再利用 Transformer 捕捉高层特征的全局依赖关系。

3. 模型架构

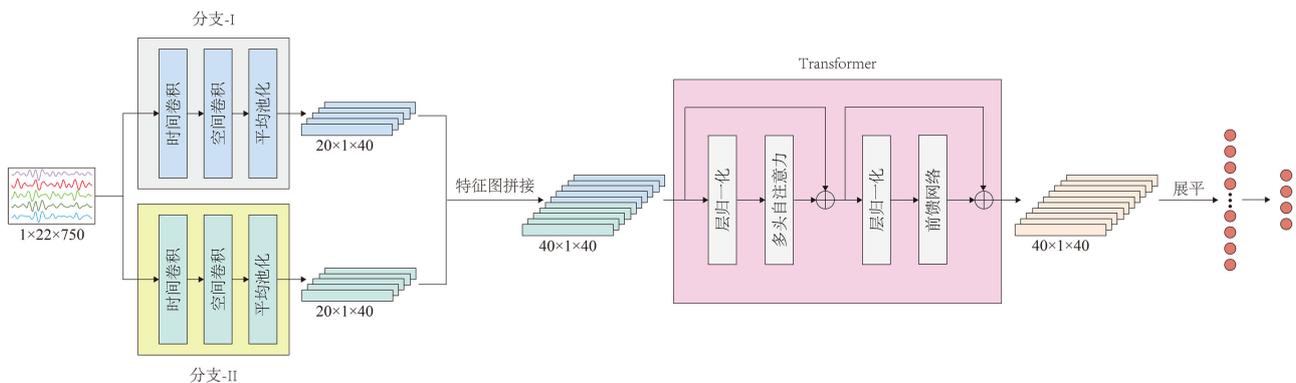


Figure 1. The architecture of DSSFFN

图 1. DSSFFN 的架构

大尺寸时间卷积核能够从脑电信号中捕获更宽泛的频率成分，而小尺寸核则更擅长提取高频特征信息。为充分发挥不同尺度下时间卷积与空间卷积的互补优势，本文设计了一种双尺度时空卷积神经网络模块，该模块可高效捕获脑电信号的多尺度时间特征与空间分布信息。在卷积核尺寸的设计上，本研究参考前人工作 [16]，将双尺度卷积核的尺寸设为 25 与 51，设计依据实验所用数据集 250 Hz 的采样率。其中 25 作为小尺度卷积核，恰好对应 0.1 秒的时间窗口，这一尺度与前人对 EEG 信号局部特征提取的核尺寸设计逻辑一致；51 作为大尺度卷积核，适配该采样率下的长时程特征提取需求，同时奇数核设计实现了时序特征的中心对齐。此外，为解决双尺度特征的有效融合问题，本文提出一种基于 Transformer

的特征融合模块，以进一步融合双尺度时间信息并增强特征的判别能力。所提框架如图 1 所示，包含双尺度时空卷积神经网络模块、Transformer 模块和分类模块。

分支一聚焦于大尺度时间与空间信息的提取，因此采用大尺度时间卷积与空间卷积的组合结构。采用 F 个尺寸为 (1×51) 的大尺度卷积核 (F 为 20)，对原始脑电信号进行滤波处理。紧接着，空间卷积层使用 F 个尺寸为 $(C \times 1)$ 的核， C 为脑电信号电极通道数，提取特征图的全局空间信息，该设计不仅能大幅降低卷积运算的可训练参数规模，还能提升模型泛化能力。最后，使用尺寸为 (1×75) 的平均池化层压缩时间维度特征。分支一的处理流程可概述如下：

$$x_{b1} = \text{Avg}(\text{SConv}(\text{TConv}(x))) \quad (1)$$

其中输入特征 $x \in \mathbb{R}^{1 \times 22 \times 750}$ ，输出特征 $x_{b1} \in \mathbb{R}^{20 \times 1 \times 40}$ ， TConv 为时间卷积， SConv 为空间卷积， Avg 为平均池化层。

分支二聚焦于小尺度时间与空间信息的提取，因此采用小尺度时间卷积与空间卷积的组合结构。时间卷积层采用 F 个尺寸为 (1×25) 的小尺度卷积核 (F 为 20)，提取脑电信号中的高频特征成分；随后，空间卷积层采用 F 个尺寸为 $(C \times 1)$ 的核，从所有特征图中学习全局空间关联信息；最后，引入一个尺寸为 (1×75) 的平均池化层，在完成特征维度压缩的同时，最大限度保留学习特征中的时序信息。分支二的特征学习过程可形式化描述如：

$$x_{b2} = \text{Avg}(\text{SConv}(\text{TConv}(x))) \quad (2)$$

其中输入特征 $x \in \mathbb{R}^{1 \times 22 \times 750}$ ，输出特征 $x_{b2} \in \mathbb{R}^{20 \times 1 \times 40}$ ， TConv 为时间卷积， SConv 为空间卷积， Avg 为平均池化层。

在分支一和分支二后，我们得到 2 个脑电信号特征图。我们先将 2 个特征图沿着滤波器维度进行拼接，得到 $x_c \in \mathbb{R}^{40 \times 1 \times 40}$ 。随后，我们送入 Transformer 模块取融合双分支特征，以进一步融合双尺度时间信息并增强特征的判别能力。Transformer 模块的核心在于多头自注意力(Multi-Head Self-Attention, MHSA)机制。具体而言，假设输入特征序列 $x_l \in \mathbb{R}^{T \times F}$ ， F 为滤波器维度， T 为时间维度。首先通过三组线性映射分别生成查询(Q)、键(K)和值(V)向量。随后，通过缩放点积注意力机制计算注意力权重：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

其中 d_k 表示键向量的维度。多头自注意力则将滤波器维度分成 h 份，分别送入到自注意力模块，然后将结果串联得到输出，这个过程可以表述为：

$$\text{MHSA}(Q, K, V) = [\text{head}_0; \dots; \text{head}_{h-1}] \quad (4)$$

$$\text{head}_l(Q, K, V) = \text{Attention}(Q_l, K_l, V_l) \quad (5)$$

其中， Q_l ， K_l ， $V_l \in \mathbb{R}^{T \times F/h}$ 分别表示对第 l 个注意力头中，对特征 x_l 完成分块处理后，经线性变换后得到的查询矩阵、键矩阵与值矩阵。最后 Transformer 模块输出融合后的特征 $x_a \in \mathbb{R}^{40 \times 1 \times 40}$ 。

分类器模块包含一个全连接层和一个 softmax 层。将融合后的特征输入到分类器得到最后的分类结果。

4. 实验结果

图 2 展示了各模型在 BCI2A 数据集的分类结果。其中所提模型 DSSFFN 实现最高分类准确率为 72.65%，超过了所有对比模型。相较于第二好的 EEGNet，准确率提高了 1.24%。相较于 ShallowNet、

EEGInception、ITNet、CTNet，准确率分别提高了 6.78%、8.82%、11.06%、6.87%。表 1 展现了各模型所有被试的分类准确率。其中，DSSFFN 在 5 个被试中分类准确率超过了所有模型。EEGNet 在 4 个受试者中取得最好结果。

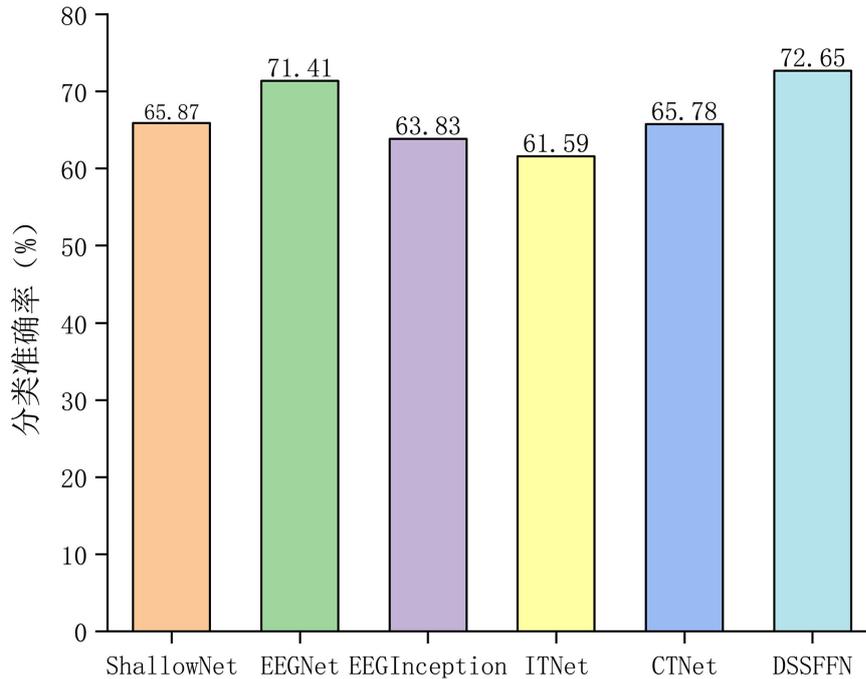


Figure 2. Classification accuracy of different models on the dataset
图 2. 各模型在数据集上的分类准确率

Table 1. The classification accuracy of each subject for the DSSFFN and the compared models
表 1. DSSFFN 与对比模型在各受试者上的分类准确率

模型名称	S01	S02	S03	S04	S05	S06	S07	S08	S09
ShallowNet	75.88	53.30	87.32	54.52	42.53	44.47	79.50	77.43	77.95
EEGNet	77.96	55.02	92.53	60.42	56.94	48.10	84.90	83.33	83.50
EEGInception	71.88	47.38	84.54	58.34	42.71	50.35	67.70	77.95	73.61
ITNet	68.41	45.12	86.11	54.33	37.33	41.32	70.82	75.69	75.18
CTNet	74.14	52.26	84.90	57.46	42.36	47.22	75	80.73	77.96
DSSFFN	83.17	60.41	89.23	66.84	55.55	55.56	85.76	82.29	75.01

图 3 展现了 DSSFFN 模型对 4 类运动想象分类准确率的混淆矩阵。可以看出，DSSFFN 擅长分类左手运动想象和舌头运动想象任务，准确率分类达到 90.97%和 88.19%。相较于右手运动想象任务来说，其他 3 类的分类准确率远高于该类。

表 2 可以看出每个模块对 DSSFFN 模型的提升。具体而言，分支一和分支二的特征拼接能提升模型的性能，相对于单一分支都有不错的提升。Transformer 作为模型最重要的组成部分之一，使模型分类准确率提高了 4.44%，这充分的说明了 Transformer 模块能够有效的融合双尺度的特征，以进一步提高提取脑电信号特征表征的鲁棒性。



Figure 3. Confusion matrix of the DSSFFN
图 3. DSSFFN 的混淆矩阵图

Table 2. Accuracy contribution of each module
表 2. 各模块的准确率贡献

分支一	分支二	Transformer	准确率(%)
✓	✗	✓	70.70
✗	✓	✓	70.20
✓	✓	✗	68.21
✓	✓	✓	72.65

Table 3. Accuracy of different fusion modules
表 3. 不同融合模块的准确率

方法	准确率(%)
SE	68.66
Transformer	72.65

表 3 为不同融合模块下模型的分类型准确率对比结果。可以看出，采用 SE 模块[17]时，模型分类准确率达 68.66%，相较表 2 基础模型(仅使用分支一和分支二)的 68.21%提升 0.45%。而以 Transformer 作为融合模块时，模型准确率提升至 72.65%，较表 2 基础模型显著提高 4.44%。对比两种融合模块可见，Transformer 融合模块的效果远优于 SE 模块，分类准确率相对提升 3.99%，充分证明 Transformer 模块能更有效地实现双尺度特征的深度融合，大幅提升特征的判别能力。

总体来看，DSSFFN 在 BCI 竞赛 IV 的 2a 数据集取得了不错的性能，这得益于双尺度时空卷积模型对特征的有效提取和 Transformer 模块对双尺度的融合，以进一步提高脑电信号特征的可辨别性。

5. 结论

本文提出了一种双尺度时空特征融合网络(DSSFFN)，该模型采用双尺度来提取脑电信号的时空特征，以捕获不同尺度的时空信息。然后送入到 Transformer 模块融合双尺度信息，以进一步融合双尺度时间信息并增强特征的判别能力。本文在 BCI 竞赛 IV 的 2a 数据集对 DSSFFN 模型进行了测试。本研究中，将 DSSFFN 与 EEGNet、EEG-Inceptio 等主流脑电信号解码模型进行对比。结果表明，DSSFFN 表现

出更优的性能，其分类平均准确率显著优于对比算法，充分证明了双尺度特征提取与 Transformer 深度融合策略的有效性。本文提出的 DSSFFN 模型是一种具有很大潜力的运动想象解码模型，在未来可应用于运动想象脑机接口系统的解码任务之中，以提高系统的可靠性。

参考文献

- [1] Wei, F., Xu, X., Jia, T., Zhang, D. and Wu, X. (2023) A Multi-Source Transfer Joint Matching Method for Inter-Subject Motor Imagery Decoding. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **31**, 1258-1267. <https://doi.org/10.1109/tnsre.2023.3243257>
- [2] Tang, X., Yang, C., Sun, X., Zou, M. and Wang, H. (2023) Motor Imagery EEG Decoding Based on Multi-Scale Hybrid Networks and Feature Enhancement. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **31**, 1208-1218. <https://doi.org/10.1109/tnsre.2023.3242280>
- [3] Müller-Gerking, J., Pfurtscheller, G. and Flyvbjerg, H. (1999) Designing Optimal Spatial Filters for Single-Trial EEG Classification in a Movement Task. *Clinical Neurophysiology*, **110**, 787-798. [https://doi.org/10.1016/s1388-2457\(98\)00038-8](https://doi.org/10.1016/s1388-2457(98)00038-8)
- [4] Ang, K.K., Chin, Z.Y., Zhang, H., *et al.* (2008) Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. 2008 *IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Hong Kong, 1-8 June 2008, 2390-2397. <https://doi.org/10.1109/ijcnn.2008.4634130>
- [5] Schirrmester, R.T., Springenberg, J.T., Fiederer, L.D.J., Glasstetter, M., Eggenesperger, K., Tangermann, M., *et al.* (2017) Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization. *Human Brain Mapping*, **38**, 5391-5420. <https://doi.org/10.1002/hbm.23730>
- [6] Lawhern, V.J., Solon, A.J., Waytowich, N.R., Gordon, S.M., Hung, C.P. and Lance, B.J. (2018) EEGNet: A Compact Convolutional Neural Network for EEG-Based Brain-Computer Interfaces. *Journal of Neural Engineering*, **15**, Article 056013. <https://doi.org/10.1088/1741-2552/aace8c>
- [7] Ingolfsson, T.M., Hersche, M., Wang, X., Kobayashi, N., Cavigelli, L. and Benini, L. (2020) EEG-TCNet: An Accurate Temporal Convolutional Network for Embedded Motor-Imagery Brain-Machine Interfaces. 2020 *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Toronto, 11-14 October 2020, 2958-2965. <https://doi.org/10.1109/smc42975.2020.9283028>
- [8] Mane, R., Chew, E., Chua, K., *et al.* (2021) FBCNet: A Multi-View Convolutional Neural Network for Brain-Computer Interface.
- [9] Altaheri, H., Muhammad, G. and Alsulaiman, M. (2022) Physics-Informed Attention Temporal Convolutional Network for EEG-Based Motor Imagery Classification. *IEEE Transactions on Industrial Informatics*, **19**, 2249-2258. <https://doi.org/10.1109/tii.2022.3197419>
- [10] Altaheri, H., Muhammad, G. and Alsulaiman, M. (2023) Dynamic Convolution with Multilevel Attention for EEG-Based Motor Imagery Decoding. *IEEE Internet of Things Journal*, **10**, 18579-18588. <https://doi.org/10.1109/jiot.2023.3281911>
- [11] Zhao, W., Jiang, X., Zhang, B., Xiao, S. and Weng, S. (2024) CTNet: A Convolutional Transformer Network for EEG-Based Motor Imagery Classification. *Scientific Reports*, **14**, Article No. 20237. <https://doi.org/10.1038/s41598-024-71118-7>
- [12] Dai, G., Zhou, J., Huang, J. and Wang, N. (2020) HS-CNN: A CNN with Hybrid Convolution Scale for EEG Motor Imagery Classification. *Journal of Neural Engineering*, **17**, Article 016025. <https://doi.org/10.1088/1741-2552/ab405f>
- [13] Chakladar, D.D., Kumar, P., Roy, P.P., Dogra, D.P., Scheme, E. and Chang, V. (2021) A Multimodal-Siamese Neural Network (mSNN) for Person Verification Using Signatures and EEG. *Information Fusion*, **71**, 17-27. <https://doi.org/10.1016/j.inffus.2021.01.004>
- [14] Salami, A., Andreu-Perez, J. and Gillmeister, H. (2022) EEG-ITNet: An Explainable Inception Temporal Convolutional Network for Motor Imagery Classification. *IEEE Access*, **10**, 36672-36685. <https://doi.org/10.1109/access.2022.3161489>
- [15] Santamaría-Vázquez, E., Martínez-Cagigal, V., Vaquerizo-Villar, F. and Hornero, R. (2020) EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive Erp-Based Brain-Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **28**, 2773-2782. <https://doi.org/10.1109/tnsre.2020.3048106>
- [16] Zhang, C., Kim, Y. and Eskandarian, A. (2021) EEG-Inception: An Accurate and Robust End-to-End Neural Network for EEG-Based Motor Imagery Classification. *Journal of Neural Engineering*, **18**, Article 046014. <https://doi.org/10.1088/1741-2552/abed81>
- [17] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>