

基于 Swin Transformer 的脑电通道注意力算法用于运动想象分类

摘要

本工作旨在复现 Wang 等人提出的基于 Swin Transformer 的脑电通道注意力算法，验证其在运动想象分类任务中的有效性。尽管未能完全重现原文的结果，但 Swin Transformer 在运动想象分类中的潜力仍然得到了验证。针对实验中遇到的问题，未来工作将聚焦于优化数据预处理流程、调整模型架构、以及引入数据增强技术等方面，以进一步提升分类效果和模型的鲁棒性。

关键词：脑电图；脑机接口；注意力机制；Swin Transformer

1 引言

脑机接口（Brain-Computer Interface, BCI）作为神经科学与人工智能领域的一项重要技术，已在医疗康复和辅助设备等领域得到了广泛应用。这项技术不仅为患有严重残疾的患者提供了有效的帮助 [1]，还在早期癫痫发作的识别中发挥了关键作用 [2] [3]。除此之外，BCI 的应用范围正不断拓展，包括先进假肢的控制 [4]、游戏与虚拟现实中的交互体验 [5]，以及推动科学研究的进一步发展。运动想象（Motor Imagery, MI）是 BCI 中常见的研究范式，通过分析脑电信号（Electroencephalogram, EEG）来识别用户的运动意图，从而为神经康复提供技术支持。然而，EEG 信号的高维度、低信噪比和复杂的时间-空间特性给传统分类方法带来了巨大的挑战。

Wang 等人 [6] 提出了一种结合 EEG 通道注意力机制（EEG Channel Attention, ECA）与 Swin Transformer 的新算法结构。该方法通过 ECA 模块捕捉 EEG 信号中重要通道的特征，并利用 Swin Transformer 的移动窗口机制对时序特征进行有效建模。实验表明，该模型在脑卒中患者的 MI 任务分类中表现出显著的准确性和稳定性，优于传统方法。

本工作旨在复现该论文提出的模型与实验，以验证其方法的有效性与稳定性，同时分析结果与原论文的差异，并探讨优化和扩展的可能性。

2 相关工作

近年来，随着 EEG 信号在脑机接口和医学辅助诊断等领域的广泛应用，如何高效地对 EEG 信号进行分类和解码成为了研究的热点。相关研究从早期的传统方法逐步发展到深度学习及注意力机制的应用，尤其是 Transformer 结构的引入，为 EEG 分类任务提供了全新的解决方案。

2.1 EEG 信号分类的传统方法

传统的分类方法主要针对 EEG 信号进行时频域分析，结合主成分分析 (PCA) 等方法提取特征，并通过支持向量机 (SVM) 等分类器完成运动想象任务的分类。例如，文献 [7] 利用功率谱密度和共振频率特征，结合 SVM 完成 EEG 分类任务。然而，这类方法过于依赖领域知识，难以全面捕捉 EEG 信号的复杂特性。

2.2 深度学习在 EEG 信号分类中的应用

深度学习方法的引入为 EEG 信号分类带来了显著的突破。文献 [8] 采用卷积神经网络 (CNN) 提取 EEG 信号的空间特征，并通过时间序列模型捕获信号的动态变化，显著提升分类性能。此外，文献 [9] 提出了一种混合 CNN-RNN 架构，结合了空间特征提取与时间依赖建模，但在长时间依赖和全局关系建模方面仍存在一定不足。

2.3 注意力机制在 EEG 信号分类中的应用

注意力机制被广泛应用于 EEG 信号分类任务，旨在突出关键信号区域。文献 [10] 结合了时间卷积网络 (Temporal Convolutional Network, TCN) 和注意力机制，以高效地捕获脑电信号中的时间模式和关键特征，展示了从原始 EEG 信号中提取 MI 特征的强大能力。

2.4 Transformer 在 EEG 信号分类中的应用

Transformer 架构凭借其出色的全局特征建模能力，近年来被广泛应用于 EEG 信号分类。例如，文献 [11] 提出了一种创新性架构，结合 CNN 和 Transformer 的优势，为 EEG 解码提供了高效解决方案，同时通过信号可视化技术加深了对脑电特征的理解。

3 本文方法

论文 [6] 设计了基于 EEG 通道注意力机制 (ECA) 和 Swin Transformer 的深度学习模型，其中 ECA 通过注意力机制增强重要特征的表达能力，而 Swin Transformer 利用滑动窗口卷积和分层特征提取机制处理 EEG 数据，并完成最终的分类任务。该架构的设计旨在有效捕获 EEG 信号的时空特征，从而增强模型的分类性能。以下是模型的具体架构与各模块的详细描述：

3.1 ECA

ECA 模块的目标是通过建模各通道之间的相关性，动态调整每个通道的权重。首先通过全局池化操作提取通道的全局特征，并使用一个简单的前馈网络生成每个通道的注意力权重。计算出的权重与原始 EEG 特征按通道逐一相乘，从而突出对分类任务更有贡献的通道特征。这种设计有效提升了信号特征的分辨能力，避免了不重要信号的干扰。

图 1 说明了 ECA 的结构和数据的流向。

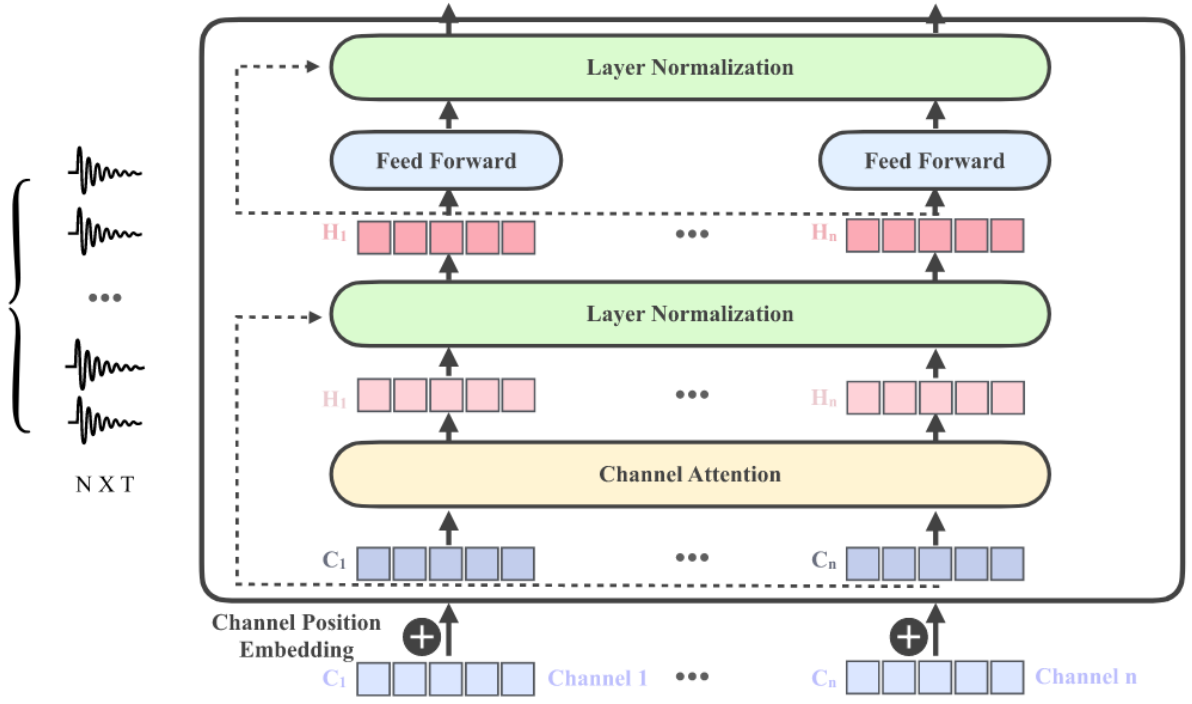


图 1. ECA 结构以及数据流示意图 [6]

3.2 Swin Transformer

Swin Transformer 是本文的核心模块，用于捕捉 EEG 数据的多尺度特征并建模其时空关系。Swin Transformer 的特点在于其分层窗口注意力机制和强大的全局建模能力。EEG 数据经过前一模块处理后，被划分为多个固定大小的窗口，每个窗口内的特征被单独输入到 Transformer 模块中计算自注意力。这样的窗口注意力机制能够显著降低计算复杂度，同时捕获局部特征。为了弥补固定窗口划分可能导致的信息丢失，Swin Transformer 采用窗口移动机制，使得窗口之间可以共享更多的全局信息。模型采用分层架构，每一层的输入特征维度逐渐减少，而特征表示的语义信息逐渐丰富。该设计可以有效建模 EEG 信号的多尺度特征。每个窗口内部采用位置嵌入技术，为特征向量加入位置信息；多头自注意力机制则进一步挖掘特征之间的复杂关系。

Swin Transformer 的架构如图 2 所示。通过各模块的协同作用，其能够有效提取 EEG 信号的多尺度空间和时间特征。

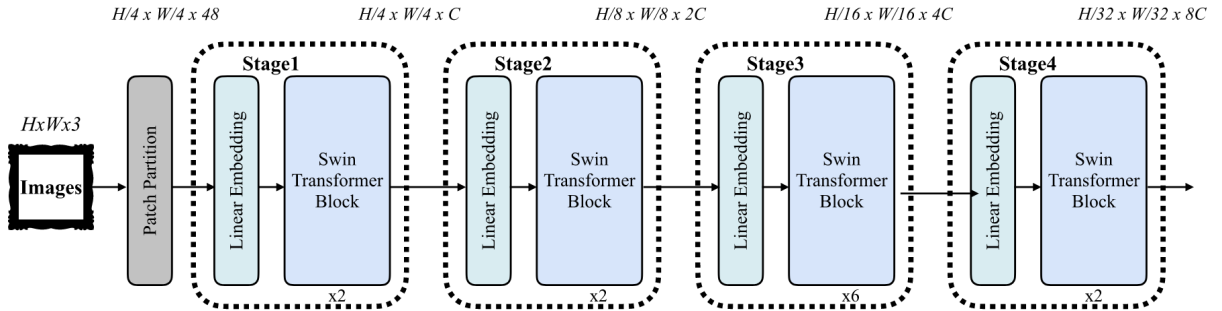


图 2. Swin Transformer 的架构示意图 [6]

4 复现细节

4.1 与已有开源代码对比

本文复现的论文并未提供开源代码。在我们的复现工作中，采用了文献 [12] 提出的 Swin Transformer 模型。模型的实现基于 PyTorch 框架 [13]，并对官方代码进行了适当调整，以满足特定的 EEG 运动想象识别任务需求。此外，在 EEG 数据预处理和模型训练中，使用了 TorchEEG 工具库 [14]。

4.2 数据集对比

4.2.1 论文所用数据集

本文复现的论文并未公开其使用的数据集。据论文 [6] 介绍，该数据集由上海华山医院提供，包含 7 名年龄在 25 至 75 岁之间的中风患者的 EEG 数据，主要用于研究运动想象和静息状态的脑活动模式。

每名患者完成 12 次会话 (session)，每次会话包含 3 次运行 (run)。每次运行随机呈现两种任务 (运动想象或静息状态) 的 30 个试次 (trial)。在每个试次的开始，屏幕显示一支白色箭头提示方向，持续 3 秒钟，随后出现红色圆圈 (代表运动想象) 或红色矩形 (代表静息状态)，持续 1.5 秒钟。

对于运动想象任务，患者需要尽快在脑海中想象手腕的运动，而无需实际动作。每试次结束后，患者休息约 1.5 秒。每个试次均标注任务标签，表明其对应运动想象或静息状态。

4.2.2 复现所用数据集

在我们的复现工作中，使用了文献 [15] 提供的脑卒中患者 EEG 运动想象数据集，包含 50 名年龄在 31 至 77 岁之间的中风患者的 EEG 数据。实验设计旨在探究脑卒中患者左右手运动想象的脑活动模式。

每名患者参与 40 个试次，每个试次持续 8 秒，包括指令阶段 (2 秒)、运动想象阶段 (4 秒) 和休息阶段 (2 秒)。在指令阶段，患者接收提示，要求想象用左手或右手抓住一个球形物体。在运动想象阶段，患者观看抓握动作的视频，并根据提示进行运动想象。休息阶段用于恢复状态。

每个试次的标签标注了任务类型，即左手或右手的运动想象。该数据集为脑卒中患者的运动想象研究提供了重要的数据支持，适用于该复现工作。

5 实验流程

本实验旨在利用脑电信号进行左右手运动想象的分类。实验流程主要分为数据预处理、模型训练和性能评估三个部分。如图 3 所示。具体如下：

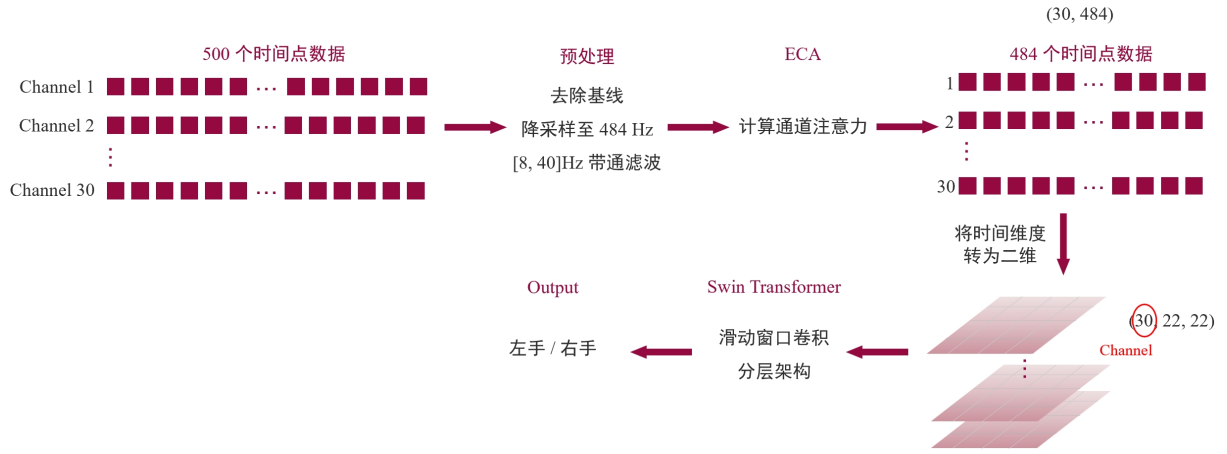


图 3. 实验流程图

5.1 数据预处理

对 EEG 数据进行预处理以提高信号质量并优化模型输入，包括基线校正以去除直流分量和减少基线漂移的影响，降采样至 484 Hz 以适应后续 Swin Transformer 的输入，同时尽量保留信息，以及带通滤波以提取 8 Hz 至 40 Hz 的频段信号，捕捉与运动想象任务相关的脑电活动特征。

5.2 模型训练

为了提升模型性能并避免过拟合，我们采用了提前停止和交叉验证策略。训练过程中通过监控训练损失，若在多轮内无显著改进，则提前停止训练以防止过拟合。同时，采用基于受试者的 5 折交叉验证 (KFoldPerSubject)，确保训练集与测试集数据相互独立，从而提高模型的泛化能力。

5.3 性能评估

实验采用分类准确率作为主要评估指标，用于衡量模型对左右手运动想象的分类能力。每折交叉验证的具体流程为：首先，将数据集按受试者划分，确保训练集与测试集完全独立；随后，在训练集上训练模型，最多进行 50 轮；然后，在测试集上评估模型性能，并记录每折实验的测试准确率；最后，计算所有折的平均测试准确率，作为模型的最终性能指标。

6 实验结果分析

表 1 展示了 50 名被试者的实验分类准确率，被试 45 的分类准确率达到 88.1%，是实验中的最高值。准确率超过 70% 的被试者主要集中在被试 37、39、40 和 45 上，其共同特点可能与被试个体差异、脑电信号质量或任务专注程度相关。另一方面，准确率较低的被试结果（如被试 6 和 26，准确率分别为 58.1% 和 50.6%）可能受限于数据量不足或信号特征不明显。模型的平均分类准确率为 61.9%。与原论文的结果相比，复现实验的分类性能不尽如人意。分析导致结果不理想的可能原因如下：

表 1. 各被试者的分类准确率，加粗为最高值，平均为 61.9%

Subject	1	2	3	4	5	6	7	8	9	10
Accuracy	0.606	0.638	0.631	0.638	0.538	0.581	0.512	0.631	0.500	0.638
Subject	11	12	13	14	15	16	17	18	19	20
Accuracy	0.550	0.600	0.675	0.612	0.600	0.531	0.650	0.581	0.600	0.569
Subject	21	22	23	24	25	26	27	28	29	30
Accuracy	0.525	0.622	0.650	0.656	0.538	0.506	0.619	0.641	0.606	0.709
Subject	31	32	33	34	35	36	37	38	39	40
Accuracy	0.625	0.638	0.631	0.531	0.587	0.625	0.725	0.594	0.700	0.737
Subject	41	42	43	44	45	46	47	48	49	50
Accuracy	0.600	0.581	0.513	0.722	0.881	0.712	0.544	0.663	0.694	0.713

首先，数据量的差异可能是关键影响因素之一。原论文中，每位被试者的数据量显著更多（12 个 session \times 3 个 run \times 30 个 trial = 1080 个试次），这为模型提供了丰富的特征表达，显著提升了模型对分类模式的学习能力。而本实验中，每位被试者仅有 40 个试次，数据量的不足可能导致模型难以全面捕获特征分布，从而降低分类性能。此外，较小的数据集易于引发模型的过拟合问题，尤其在复杂任务中泛化能力不足。

其次，任务目标的复杂性增加了实验难度。原论文中，分类任务为动作想象与静息状态的区分，这一任务相对简单，脑电特征更加显著。而本实验的分类任务为左手与右手想象的区分，这对模型的特征提取能力提出了更高的要求，可能进一步加剧了数据量不足对分类性能的负面影响。

最后，模型参数与测试策略的差异也可能对实验结果造成影响。由于原论文未详细阐述模型架构的具体细节，复现过程中无法完全还原其设置。此外，数据集划分方法的不明确（如训练集和测试集的具体比例与策略）可能进一步影响了模型的训练与评估过程。

7 总结与展望

本次论文复现以左右手运动想象分类为研究对象，复现了基于 EEG 通道注意力机制和 Swin Transformer 的深度学习模型，并对其在脑卒中患者 EEG 运动想象数据集 [15] 上的表现进行了评估。实验结果表明，模型在 50 名被试者上的平均准确率达到 61.9%。与原论文相比，复现结果在整体性能上存在差距。对比分析显示，数据量不足、任务目标复杂性以及模型参数未知是实验结果未达到期望的主要原因。

未来的研究方向主要包括以下几点：

- 增加数据量：通过扩充数据集规模或引入数据增强方法，为模型提供更丰富的训练样本，以提升泛化能力。
- 优化模型架构：探索更适合运动想象分类的模型结构，如结合更多时间序列特征或引入更精细的注意力机制。

- 改进任务设计：尝试简化分类任务（如二分类任务的类别划分）或采用多模态数据（如结合肌电信号），以进一步提升分类性能。

通过上述改进措施，相信模型在脑电信号分类任务中的表现将进一步提升，为脑机接口技术的发展提供有力支持。

参考文献

- [1] Shayan Jalilpour, Sepideh Hajipour Sardouie, and Amirmohammad Mijani. A novel hybrid bci speller based on rsvp and ssvep paradigm. *Computer methods and programs in biomedicine*, 187:105326, 2020.
- [2] Charmi Daftari, Jainish Shah, and Manan Shah. Detection of epileptic seizure disorder using eeg signals. In *Artificial intelligence-based brain-computer interface*, pages 163–188. Elsevier, 2022.
- [3] Vladimir Maksimenko, Annika Lüttjohann, Sabrina van Heukelum, Janita Kelderhuis, Vladimir Makarov, Alexander Hramov, Alexey Koronovskii, and Gilles van Luijtelaar. Brain-computer interface for the epileptic seizures prediction and prevention. In *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*, pages 1–5. IEEE, 2020.
- [4] Adrian Czech. Brain-computer interface use to control military weapons and tools. In *Control, Computer Engineering and Neuroscience: Proceedings of IC Brain Computer Interface 2021*, pages 196–204. Springer, 2021.
- [5] R Raja Subramanian, Konduru Yaswanth Varma, Kondapalli Balaji, Musalappagari Devendra Reddy, Adepudi Akash, and Kurapati Nithish Reddy. Multiplayer online car racing with bci in vr. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1835–1839. IEEE, 2021.
- [6] Han Wang, Lei Cao, Chenxi Huang, Jie Jia, Yilin Dong, Chunjiang Fan, and Victor Hugo C De Albuquerque. A novel algorithmic structure of eeg channel attention combined with swin transformer for motor patterns classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.
- [7] Fabien Lotte, Marco Congedo, Anatole Lécuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.
- [8] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420, 2017.

- [9] Pouya Bashivan, Irina Rish, Mohammed Yeasin, and Noel Codella. Learning representations from eeg with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448*, 2015.
- [10] Hamdi Altaheri, Ghulam Muhammad, and Mansour Alsulaiman. Physics-informed attention temporal convolutional network for eeg-based motor imagery classification. *IEEE transactions on industrial informatics*, 19(2):2249–2258, 2022.
- [11] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2023.
- [12] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022, 2021.
- [13] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [14] Zhi Zhang, Sheng-hua Zhong, and Yan Liu. Torcheegemo: A deep learning toolbox towards eeg-based emotion recognition. *Expert Systems with Applications*, 249:123550, 2024.
- [15] Haijie Liu, Penghu Wei, Haochong Wang, Xiaodong Lv, Wei Duan, Meijie Li, Yan Zhao, Qingmei Wang, Xinyuan Chen, Gaige Shi, et al. An eeg motor imagery dataset for brain computer interface in acute stroke patients. *Scientific Data*, 11(1):131, 2024.