

基于动态卷积和多级注意力的MI-EEG图像分类

摘要

脑机接口（BCI）是一种创新技术，它利用人工智能（AI）和可穿戴脑电图（EEG）传感器来解码脑信号，并提高生活质量。基于 EEG 的运动想象（MI）脑信号在许多 BCI 应用中被使用，包括智能医疗、智能家居和机器人控制。然而，解码脑信号的能力有限是阻碍 BCI 技术大规模扩展的主要因素。在这项研究中，我们首先介绍了一种用于解码基于 EEG 的 MI 信号注意力时序卷积网络（ATCNet），并在此模型的基础上改进。改进后的模型使用动态卷积（Dy-conv）和多级注意力来提高 MI 分类的性能，同时参数数量相对较少。改进后的模型有两个主要模块：1）动态和 2）时序卷积。Dy-conv 使用多级注意力来编码低级 MI-EEG 信息，时序卷积使用偏移窗口和自我注意力来从编码信号中提取高级时序信息。与现有方法相比，所提出的模型在 BCI 竞赛 IV-2a 数据集上，对于非特定受试者的准确率为 71.3%，对于特定受试者的准确率为 87.08%。

关键词：EEG；MI；动态卷积；多级注意力机制

1 引言

随着深度学习的发展，人们已经可以对许多信号进行识别并对其进行相关的利用。例如人类已经可以使用对音频信号进行识别，并在此基础上开发出了翻译和语音控制系统；对图像进行识别，从而可以进行人脸识别，物体分割；对自然语言进行识别，从而可以进行语言的情感分析、文本分类、机器翻译、问答系统等任务。从这些任务中可以看出人们可以利用现在的深度学习技术来对多种不同的信号进行识别，并在此基础上进行利用。前面的这些信号都是外部的信号，人们最终还是想对大脑的信号来进行识别，并对其运用到生活中来，而这种技术就是脑机接口。大脑是人类最重要的器官，理解大脑的结构与功能是目前最具挑战性的前沿科学问题。中国也非常重视对这方面的研究，在“十三五”规划的基础上又设定了相应的“十四五”脑电计划，力争在该领域中取得重大成果。脑机接口(BCI)是一项具有改变世界能力的技术，旨在建立一种直接连接人脑和计算机或其他外部设备的通信渠道。它允许人们通过思维、意图或感觉来控制外部设备，或者获取来自脑部的信息。脑机接口的发展对于医学、神经科学和人机交互领域具有重要意义。BCI系统已被应用于多个领域，如医疗、安全、娱乐、智能环境控制、情感教育等领域。BCI最广泛的应用之一是在医学领域治疗和诊断神经疾病，如癫痫、抑郁症、阿尔茨海默氏症、脑卒中等[1]。此外，它还被用于识别和分类情绪和睡眠阶段，以及为运动障碍者提供进行正常运动的机会[2]。在脑机接口领域里面最重要的是需要将信号识别出来才能够对这些信号进行利用。现在使用得最多的非侵入式脑

电信号是FMRI和EEG信号[3]。但是由于FMRI需要专门的测量人员和昂贵的机器，覆盖的脑电信号的范围也比较少，所以在现代的研究当中都是使用EEG信号来进行相关的研究。



图 1. EEG信号测量

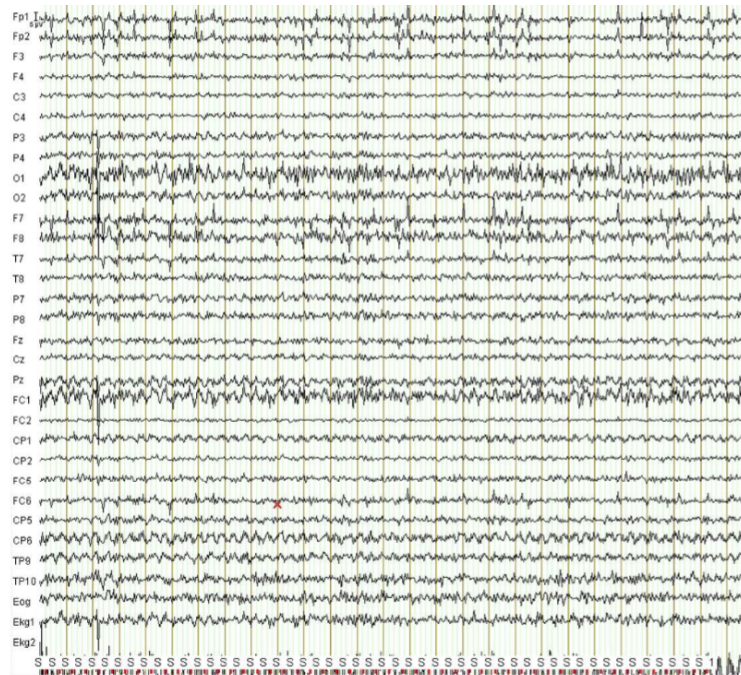


图 2. EEG图像

图1是一个人坐在电脑前面看着电脑上面的文字带着电极帽测量EEG信号。通过测量到的这些脑电信号，就可以在处理过后将其进行分类利用。现在的EEG信号的研究方向有MI、SSVEP、P300、ERS等。在这里面MI是最常见的一种应用。MI基于一个重要的神经科学观念，即想象运动与实际执行运动时在大脑中的活动模式相似。当我们想象进行运动时，大脑中的运动执行区域（MotorExecutionArea）会激活，尽管没有真实的肌肉动作发生。这个动作通常是一个物理运动，比如想象一只手或一条腿的运动[4]。这些想象运动会触发特定的神经活动，导致EEG信号中的特征变化。通过分析这些变化，可以识别出参与者的意图和所想象的运动类型。MI也在很多的领域中得到了应用，但是由于EEG信号只是一个波形的信号，所以想要通过这个来对其进行识别分类并加以应用还是有许多的困难需要去克服。MI信号分类

的准确率和图像分类的90%以上的准确率还有很大的距离，这也是EEG信号本身噪声多有很大关系，图2是一张EEG的原始图像，可以看出有很明显的噪声。限制脑机接口发展的就是对EEG信号的解码。所以本文就是聚焦于MI信号的分类准确率的提升。

2 相关工作

从低信噪比和非平稳的EEG数据中提取出有用的信息，并使用其来准确地预测人类的意图是一件比较困难的事，因为收集到的EEG信号中有非常多的伪影。下面的图2就是一个原始的EEG信号图像，从这里可以看到在60HZ及其倍数的地方有明显的噪声，这个就是电源带来的噪声。这会很明显的影响我们对EEG图像的分类。在最初的阶段,MI-EEG图像都是通过人工来进行分类的，在此之后基于传统的机器学习(ML)和深度学习(DL)的方法被用来对MI-EEG图像进行特征的提取和分类。通过使用机器学习和深度学习的方法可以使分类的准确率得到很大的提高，许多研究者也在不断的进行相关的研究，希望可以将分类的准确率提升[5]。

2.1 使用机器学习对MI分类

在近年来，运动想象（MI）任务的脑电图（EEG）分类领域，多种机器学习（ML）技术得到了广泛研究和应用。这些技术不仅包括基于深度学习的方法，也涵盖了传统的机器学习方法。传统机器学习方法在这一领域的应用主要集中在特征提取和分类算法上[6]。例如，公共空间模式（CSP）及其变种，如滤波器组公共空间模式（FBCSP），被广泛用于从EEG信号中提取区分不同MI任务的特征[7]。这些特征随后可以被多种传统ML分类器利用，包括但不限于支持向量机（SVM）、决策树、随机森林等。这些方法依赖于人工设计的特征和经典统计学习算法，以实现有效分类。

2.2 使用深度学习对MI分类

机器学习对MI的分类虽然有一定的效果，但是深度学习对于MI的分类效果更好。在过去几年里，多种深度学习（DL）技术已经被开发出来，以优化运动想象（MI）任务在脑电图（EEG）中的分类。这些技术包括但不限于卷积神经网络（CNN），这类网络在多项研究[8]、[9]、[10]、[11]、[12]、[13]中被广泛应用于MI分类任务。CNN以其多样化的架构获得了广泛关注，包括轻量级[12]和深层[19]结构、以初始化为基础的CNN[10]、[11]、残差网络[20]、三维CNN[20]、多尺度CNN[13]、多层CNN[18]、多分支CNN[9]、[20]以及集成注意力机制的CNN[8]、[9]、[10]、[11]、[13]。此外，循环神经网络（RNN）[14]、[15]和深度信念网络（DBN）[16]也在这个领域中展现了其潜力。例如，Xu等[16]开发了一个基于受限玻尔兹曼机（RBM）的DBN特征提取模型，结合支持向量机（SVM）进行分类。Hassanpour等[17]则设计了一个使用频率特征进行MI任务分类的堆叠自动编码器（SAE）。另一方面，RNN被用来从MI-EEG信号中提取时间序列信息。例如，Kumar等[14]结合了FBCSP特征和SVM分类器的长短期记忆网络（LSTM）。Luo和Chao[15]则采用了FBCSP特征作为门控循环单元（GRU）模型的输入，展示了GRU在性能上优于LSTM的结果。在此之后，还有很多人使用了利用基于注意力的DL模型对MI-EEG信号[8]，[9]，[10]，[11]，[13]进行分类。例如，Zhanget[8]利

用LSTM和图神经表示的自我注意来解码MI任务Aminetal.[10], [11]将注意层与CNN和LSTM结合。总体来看，CNN在MI任务分类中相较于其他DL模型（如RNN、SAE和DBN）表现更佳。因此，结合CNN与其他深度学习模型，如LSTM[8]和SAE[18]，已经成为许多研究者的重点，并且取得了鼓舞人心的成果。

3 本文方法

3.1 本文方法概述

由于MI图像是根据时间来进行记录的图像，所以本文也是使用一种CNN的变体，被称为TCN的时间卷积网络来进行MI图像特征的分类。最终本文根据TCN模型提出了一个高性能的ATCNET模型，该模型结合使用了该模型利用了TCN、SciML、注意机制和基于卷积的滑动窗口的强大功能。利用卷积进行滑动窗口，通过并行化处理，减少计算量，增加MI数据，有效提高精度。这里还使用了自我注意，这样有助于DL模型关注脑电信号数据中最有效的信息，而多个头有助于关注多个位置，从而产生多个注意力表征。最终将有用的信息提取出来对图像进行分类。

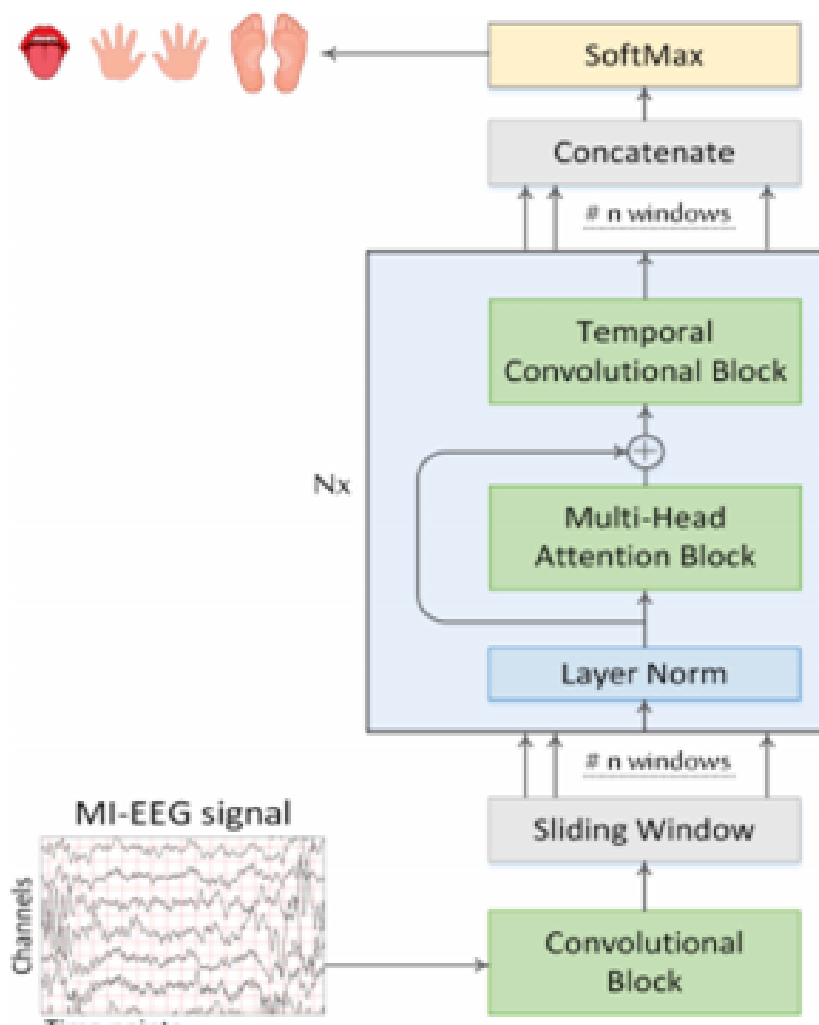


图 3. ATCNET模型

本文的ATCNet模型从结构上是由三个主要的卷积块组成：卷积（CV）模块、注意力（AT）模块，以及时间卷积（TC）模块，结构如图3所示。首先，CV模块通过三层卷积处理，对MI-EEG信号进行时空低阶信息的捕获，包括时间维度卷积、通道维度深度卷积和空间维度卷积。这一过程的输出是一个经过高阶表征的时间序列数据。接着，AT模块采用多头自注意力机制（MSA）对时间序列中的关键信息进行强调和提取。最终，TC模块利用时间卷积网络（TCN）来提取时间序列的高级时间特征。这些特征随后通过全连接（FC）层送入SoftMax分类器进行分类。为了进一步提升模型的性能，从CV模块得到的时间序列被划分为若干个小窗口，每个窗口单独经过AT和TC模块的处理。这些窗口的输出会被整合在一起，形成SoftMax分类器的输入。这种方法有助于有效地增强数据集，并提升模型的分类准确性[21]。

3.2 卷积模块

CV块与EEGNET架构类似，但是不是使用可分离卷积，而是使用了二维卷积来代替，这样可以体现出更好的性能。如图4所示，CV块是由三个卷积层组成的，其中第一层是时间卷积，是由一个尺寸为 $(1, KC)$ 的滤波器组成。这一层是使用四分之一的采样率来提取出4HZ以上的时间信息。第二层是一个深度卷积，是一个大小为 $(C, 1)$ 的滤波器，其中其中C为脑电图通道数。使用深度卷积，每个滤波器从单一的时间特征图中提取空间特征。这里最后输出的是一个 $F1 \times D$ 的特征图，是确定卷积视觉（CV）块输出维度的关键因素。模型中的深度卷积之后接入了一个尺寸为 $(1, 8)$ 的平均池化层，旨在从时间数据中提取出8倍的特征，进而将信号采样率降至约32Hz。此外，模型包括第三个卷积层，该层由尺寸为 $f2$ 的滤波器 $(1, KC2)$ 组成。我们将 $KC2$ 设定为16，以便在500毫秒的时间窗内解码运动想象（MI）活动，这适用于采样率为32Hz的数据。最后，采用尺寸为 $(1, P2)$ 的第二个平均池化层进一步降低采样率至约 $32/P2$ Hz。 $P2$ 的设置用于调节CV块产生的时间序列长度。最终CV块输出的是一个由维数为 $R^{T_C \times d}$ 的时间序列， $d = F2 = F1 \times D$ ，根据经验将 d 设置为32，并且时间序列的长度是由 $T_C = \frac{T}{8P_2}$ 来确定的。其中这里的T为原始EEG的时间点。

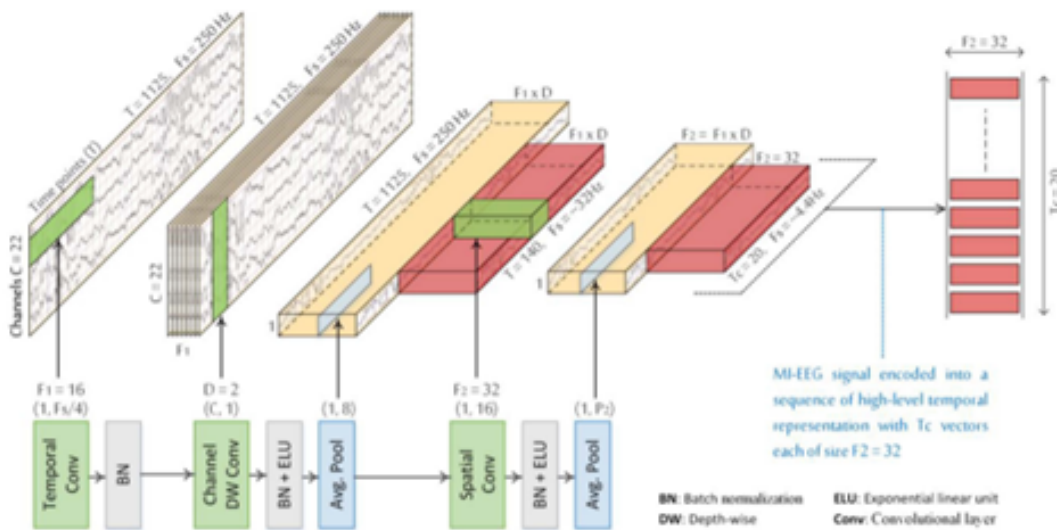


图 4. ATCNET模型

3.3 基于卷积的滑动窗口

MI数据是时间段的数据，直接对时间段的数据进行分析会有很大的困难，准确率不高，所以这里使用了滑动窗口来扩充数据，提高译码精度。由于单独使用滑动窗口会增加大量的计算，所以这里就使用了滑动窗口在卷积块里面进行集成。在这里将一个长度为 T_w 的滑动窗口划分为多个小的时间段数据。这里的窗口长度 T_w 是由下面的这两个式子决定的。

$$T_w = T_c - n + 1, T_c > n \geq 1 \quad (1)$$

$$T_w = \frac{T}{8P_2} - n + 1 \quad (2)$$

如果使用CV块来对大小 $P_1=8$ 和 $P_2=7$ 进行两次池化，就会产生一个由 $T_c=20$ 的向量组成的时间序列。每个向量都是表示原始的MI-EEG信号中的56个时间点。因此这里就相当于进行了56次时间步的滑动。

3.4 注意力模块

注意力机制由可以用三个组成部分来模拟:值(感官输入)、键(非意志线索)和查询(意志线索)。查询和键的交互产生了注意力集中，从而导致了值的选择，如图5所示。这里使用MSA层作为注意力层，MSA由图6中的几个称为头的自注意层（即缩放的点积注意）组成。这样就可以通过查询和键之间的交互产生注意分数，引导选择偏向于值。

$$q_t^h = W_h^Q \text{LN}(z_{i,t}^w) \in \mathbb{R}^{d_h}, \quad W_h^Q \in \mathbb{R}^{d \times d_h} \quad (3)$$

$$k_t^h = W_h^K \text{LN}(z_{i,t}^w) \in \mathbb{R}^{d_h}, \quad W_h^K \in \mathbb{R}^{d \times d_h} \quad (4)$$

$$v_t^h = W_h^V \text{LN}(z_{i,t}^w) \in \mathbb{R}^{d_h}, \quad W_h^V \in \mathbb{R}^{d \times d_h} \quad (5)$$

其中这里的LN代表层归一化， $t=1, \dots, T_w$ 是窗口 w 中的时间矢量的索引，并且 T_w 是该窗口的长度。 H 是多个注意力头上的索引， H 是头的总数。根据经验将注意力头部的减小设置为 $dH=d/2H$ 。给定一个输入向量 $q_t \in \mathbb{R}^{d_h}$ 和一组键值对 $(k_t(1), v_t(1)), \dots, (k_t(T), v_t(T))$ ，其中 $k_t(i) \in \mathbb{R}^{d_k}$ 和 $v_t(i) \in \mathbb{R}^{d_v}$ ，注意力机制在计算时，下文简称 C_t 就是关注 $V_t(i)$ 的权重和。

给定一个查询 $q_t \in \mathbb{R}^{d_q}$ 和 T 个键值对 $(k_t(1), v_t(1)), \dots, (k_t(T), v_t(T))$ ，其中 $k_t(i) \in \mathbb{R}^{d_k}$ 和 $v_t(i) \in \mathbb{R}^{d_v}$ 。注意力池化 f 生成上下文向量 c_t 被定义为 $v_t(i)$ 的加权和。查询向量 q_t^h 和键向量 k_t^v 的注意力权重（标量） $\alpha_{t,u}$ 是通过对应的对齐 $e_{t,u}$ 得到的。

$$\alpha_t^h = \text{softmax}(e_{tt'}) = \frac{\exp(e_{tt'}^h)}{\sum_{k=1}^{T_u} \exp(e_{tk}^h)} \in \mathbb{R}. \quad (6)$$

加性注意力和乘性注意力是经常使用的，在本文中我们使用了加性注意力，并且在这里进行了缩放点积的处理，这样可以确保无论矢量的长度如何，点击的方差都保持不变。

$$e_{tt'}^h = a(q_t^h, k_{t'}^h) \in \mathbb{R}. \quad (7)$$

$$a = \frac{(q_t^h)^\top k_v^h}{\sqrt{d_H}} \in \mathbb{R}. \quad (8)$$

对于每个头，在缩放点积注意力机制中，我们计算一个上下文向量，其中有 n 个查询 m 个键值对。查询和键的维度 d_h ，而值的维度 v （在本文中 d_h 的设定值是8）。这些上下文向量用来捕捉序列内部元素之间的依赖关系。在本文中，我们选择自注意力机制，它能同时对所有输入进行并行计算。

$$C^h = \text{softmax} \left(\frac{(Q^h (K^h)^\top)}{\sqrt{d_H}} \right) V^h \in \mathbb{R}^{(n-T_w+1) \times (v=d_h H)} \quad (9)$$

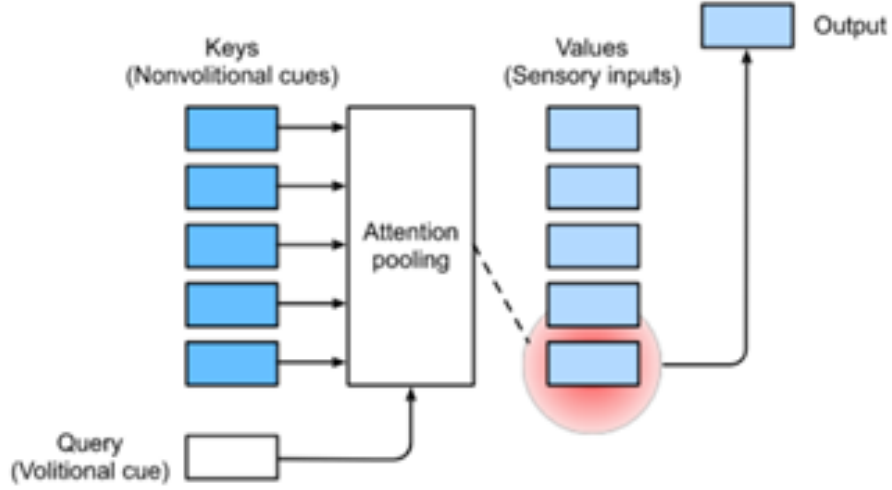


图 5. 通过注意力查询使值的选择产生偏差

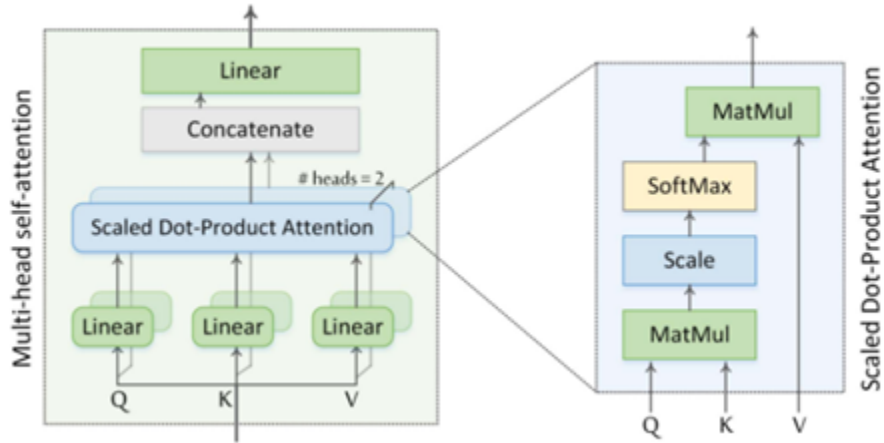


图 6. 多头自我注意力

3.5 时间卷积块

本文中的时间卷积块与TCN具有相同的结构，TCN由一堆剩余块组成。残差块由两个扩展的因果卷积层组成，每个卷积层之后是批量归一化和ELU激活，如图7所示。由于其学习身

份函数的能力。在残差块中，因为输入和输出维度相同，所以使用单位映射。感受野(RFS)由两个参数控制：剩余块的数量L和犬舍大小KT，如中所定义：

$$RFS = 1 + 2(K_T - 1)(2^L - 1). \quad (10)$$

本文的TC块由具有残差块的TCN和用于所有卷积层的大小为KT=4的32个滤波器组成。有了这个TCN，RFS是19，即TCN可以在一个序列中处理多达19个元素，如图7所示。对于TC块中时间序列小于等于19的数据，在处理时可以保证其在处理时时间信息不被丢失。这样最后输出的是一个大小为FT的向量。

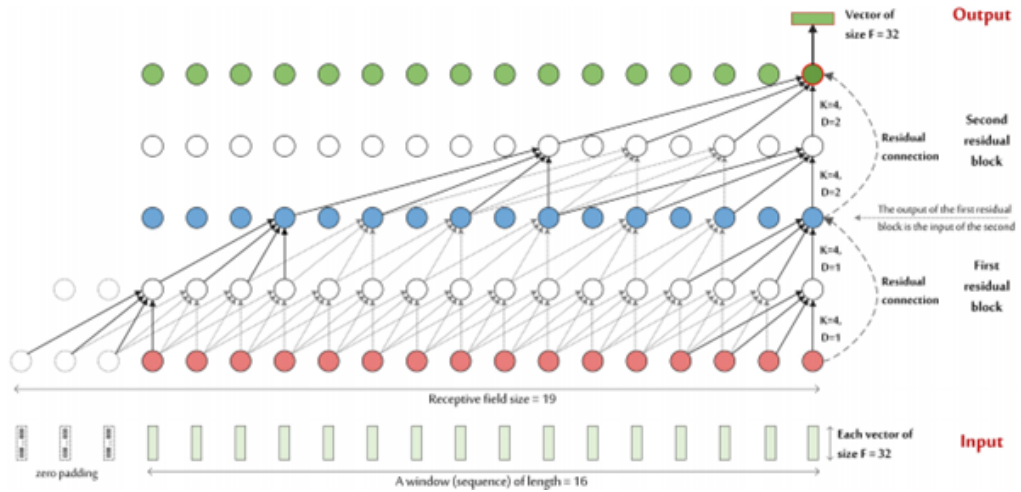


图 7. TCN结构

4 复现细节

4.1 与已有开源代码对比

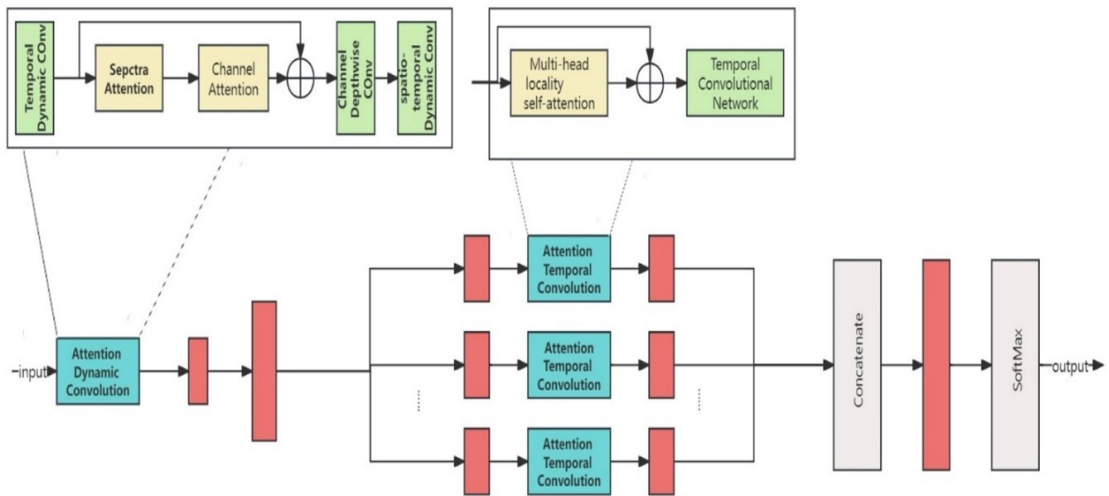


图 8. 改进模型架构

这个项目有公开在github上面的源代码，作者就是使用上面的ATCNET网路架构来对MI-EEG数据进行分类，通过三个主要的层来进行特征的提取，并且将提取到的特征进行训练、评估和可视化结果。在原来的代码里面是在attention-models模块里面实现了对注意力机制的实现，该函数支持多头注意力、多头局部注意力、多头自注意力等方法的实现。在preprocess模块里面实现了对MI-EEG数据的预处理的步骤，将数据进行导入和标准化操作，然后还实现了LOSO交叉验证方法的实现。在models模块里面实现了对ATCNET模型的定义，并且对其参数进行了相应的定义操作。最后在main模块的代码里面定义了相关的优化函数、损失函数、回调函数等。通过对模型的训练、验证、测试得到相关的曲线，再对每个受试者绘制其相关的混淆矩阵图。在原有的代码的基础上，我对attention模块的代码进行了改进，将原来的多头自注意力机制改为了多级注意力机制，这样就可以对多个不同的方面进行关注，也就可以观测到数据里面的更复杂和微妙的特征。然后还在models模块里面将原来的卷积层改为了动态卷积，这样可以使模型更具输入的数据动态调整其卷积滤波器，使得每个信号都具有更独特的特征。

4.2 创新点

在原有的模型架构里面是使用了普通的卷积来对原始的图像进行特征的提取，但是由于MI-EEG数据不是普通的图像或文字数据，所以普通的卷积处理效果会受到限制。这里使用动态卷积来提取特征。如图8所示，改进后的模型架构中的动态卷积具体过程为：先使用时间动态转换层来处理 EEG 数据的时间方面的特征；然后使用频谱注意力将注意力集中在脑电图信号中的特定频段，突出显示任务中最重要的频谱特征；接着使用通道注意力来专注于识别和强调脑电图数据中信息最丰富的通道；再在每个通道上独立运行的卷积，允许模型学习每个通道内更复杂的空间特征；最后使用时空动态卷积结合了空间和时间特征，动态调整卷积滤波器以更好地捕获与运动想象任务相关的脑电图信号特征。使用动态卷积处理有以下几个好处：

- 1.个性化:脑电图信号是高度个性化的。动态卷积使其处理适应每个信号的独特特征，增强信号分析的个性化。
- 2.复杂性管理:脑电图数据包含复杂和非线性的特征。动态卷积通过根据特定信号模式定制滤波器来有效管理这种复杂性，从而增强网络捕获和分析复杂细节的能力。
- 3.噪声处理:脑电图信号通常包含噪声。动态卷积的自适应特性有助于更好地辨别噪声，从而实现更清晰的信号分析。
- 4.计算效率:尽管动态卷积具有适应性，但可以设计得具有计算效率，使其适合实时应用。
- 5.改进的泛化能力:由于其动态特性，该模型可以在不同的主题和会话中更好地泛化从而使其在各种场景中更加稳健。[22]

除了上面的一个改进还将原来的多头自注意力机制变为多级注意力机制。如图8所示，这里注意力模块是由多头局部自注意力和时间卷积网络组成。这里的多头局部自注意力与原文中的MSA相比引入了引入了对角掩码和可学习温度。MLSA层由多个自我注意头组成，它们执行缩放的点积注意。每个注意头的三个主要部分是查询Q、键K和值v。查询与键交互，生成注意分数，突出值中的有价值元素。每个头对局部窗口计算其对应的注意力得分，得出每个头的上下文的向量。然后将所有的头部的上下文向量串联起来并将线性结果进行投影加到输

入序列 z 中，如下面的公式所示：

$$z^l = W_0 [c^1, \dots, c^H] + z^i \in \mathbb{R}^{T \times d}, \quad W_0 \in \mathbb{R}^{d_h \times d} \quad (11)$$

这里的时间卷积块与之前的相比有一个唯一的变化，即我们用Dy-conv替换了一维conv层以提高效果。TCN由一系列的残块组成，每个残块包含两个因果扩张的转换层，与批归一化（BN）和指数线性单元（ELU）相连。每个分量都是一个大小为 F 的向量，其中 F 等于ADC块中最后一个转换层的核数。TCN的输出序列中的最后一个分量是一个长度为 F 的向量，其中 F 等于TCN中内核的数量，我们的实验中设定为 $F=2=ft=32$ 。

通过这样的改进，理论上可以更好地关注时间片中的EEG信号在时间、空间和光谱维度上的注意力。这有助于深度学习模型自动关注多维度的有价值特征。

5 实验结果分析

5.1 数据集

我们在BCI-IV2a数据集上评估了D-ATCNet模型。BCI-IV2a是Graz大学于2008年推出的一套基于脑电图的MI数据集。BCI-IV2a是基于脑电图MI分类的基准数据集。使用这个数据集解码MI任务是具有挑战性的，因为在不受控制的环境中进行的MI试验的数量有限，并且有重要的人为干扰。BCI-IV2a数据集由属于4个MI任务的5184个MI试验组成。每次试验时长为8秒，中间的4秒进行心肌梗死活动。这组数据由9名受试者在两个时间段内使用22个脑电图传感器记录下来。该模型使用两个会话中的一个进行训练，使用另一个会话评估其性能。

5.2 评估方法和绩效指标

在评估所提出的模型时，使用了主体依赖和主体独立的方法。我们使用原始的竞赛训练和测试数据进行受试者依赖评估。具体来说，我们在第一阶段使用 9×288 次试验来训练模型，然后在第二阶段使用 9×288 次试验来评估模型。对于学科独立的评价，我们采用“留一名受试者(LOSO)”评价方法。Kappa评分和准确性度量被用来评估在中定义的提议和再现模型的性能。本研究中的所有训练和评估实验都在以下gpu上运行：GTX 4060 8gb，使用Python3.7的TensorFlow框架。最后改进后的模型优于改进前的模型，改进前的模型的k-score为0.805，准确度为85.38%；而改进后的模型的其k-score为0.828，准确度为87.08%。改进后的模型在两个指标上都取得了较为明显的进步。

5.3 多水平注意消融分析

我们进行了消融分析来评估多水平注意力的表现。在内核和输入特征级别上从改进后模型中删除注意层的效果如表II所示。在培训和测试程序之前，注意力层被移除。结果表明，使用Dy-conv的核水平注意提高了0.97%的总体精度，特征水平注意(包括时间、光谱和通道注意)提高了1.43%。结果表明，两个层次的注意，跨越核和输入特征，可以协作和单独提高D-ATCNet的性能。表1还显示了将注意力放在整个时间序列上(在SW之前)而不是放在每

个 SW 上的影响。结果表明，使用单独的时间注意对每个 SW 提高了 1.36% 的准确性。这是因为 SW 注意力(SWA)有助于在每个窗口中分别突出显示本地特性。

Removed attention	Accuracy %	k-score
none	87.08	0.828
kernel level attention (DC)	86.11	0.815
features level attention (TA, SA, CA)	85.65	0.809
multilevel attention (DC, TA, SA, CA)	85.46	0.806
shifted window attention (SWA)	85.72	0.809

表 1. 消融实验对比图

5.4 与最近的实验比较

在本小节中，我们将BCI-IV2a 数据集中的每个受试者的改进后模型的性能与改进前的ATCNET模型及EEG-TCNet、EEGNet模型进行了比较，这些模型的结果依赖于原始文章中定义的参数，而训练，测试都是依赖于本文中的方法进行。图9显示了基于10次随机运行的每个模型的最佳和平均性能。从表中结果可以看到，改进后的所有受试者平均性能优于其他的模型。即平均 90 次单独和随机运行(9 受试者×10 次运行)。这表明所提出的模型具有较强的学习能力，并能在多次运行后重现相似的稳定结果。该进后的模型也达到了最低的标准偏差的对象，表明它的表现是更一致的不同对象。

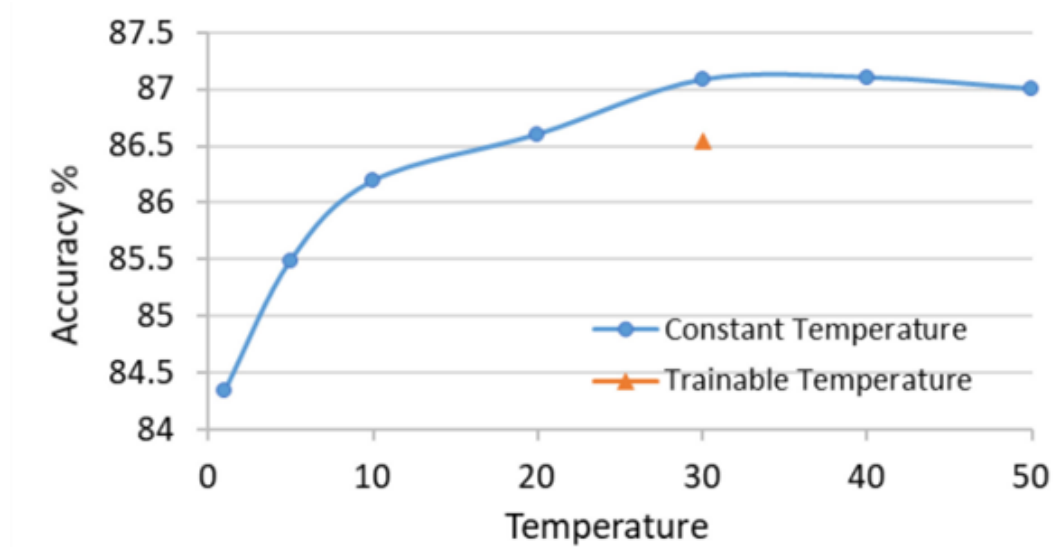


图 9. 模型平均性能

图10显示了改进后的模型和复制网络的混淆矩阵。与其他模型相比，改进后模型对所有MI活动的MI分类性能都有所提高。

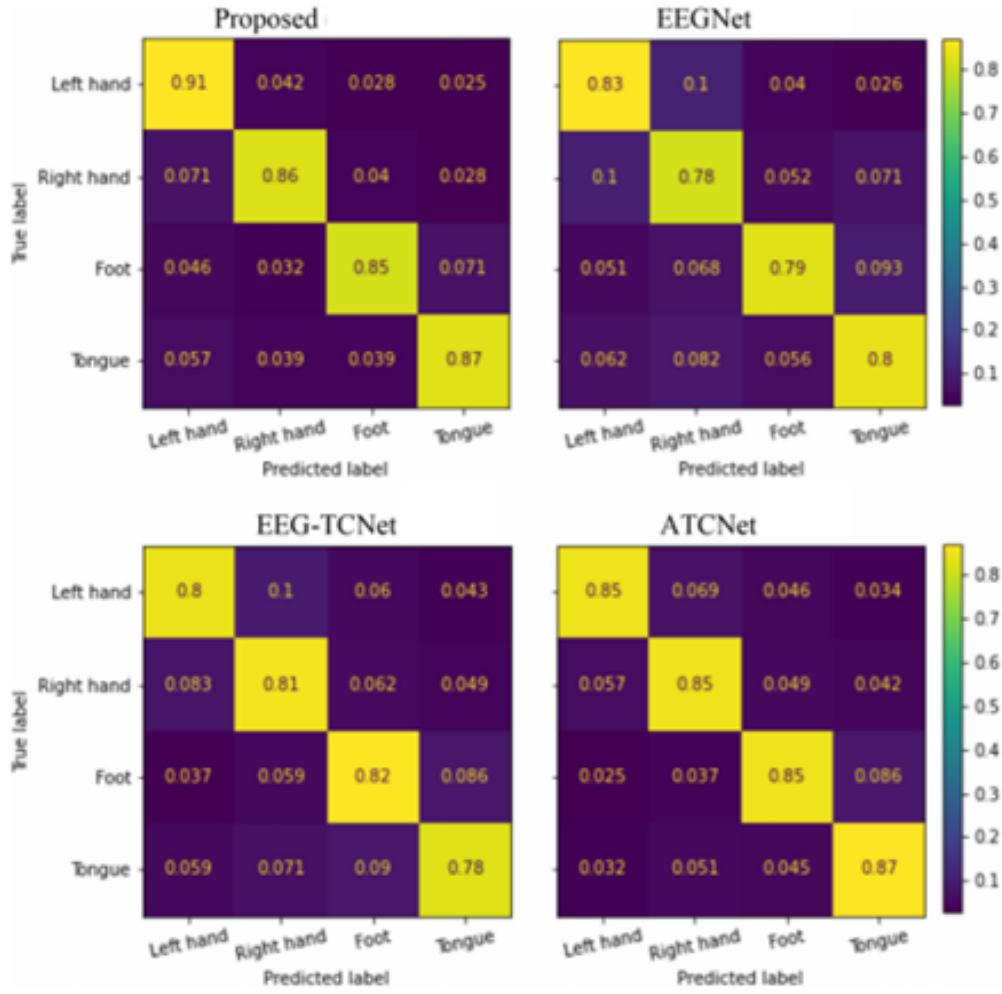


图 10. 改进模型与最近模型混淆矩阵对比

6 总结与展望

本文在原有的ATCNET模型上加以改进，提出了一种基于MI-EEG信号的解码分类器。该改进后的模型主要是由注意力ADC和ATC组成。ADC块使用具有多级注意力的三个conv层从MI-EEG数据中提取低级别时空特征。使用Dy-conv在内核级别实现多级注意力，使用MLSA和SE在特征级别实现多级注意。研究了EEG信号的时间、空间和频谱维度上的特征级别注意力。ATC块然后使用SW和自注意来提取高级时间信息。消融研究的结果证明了Dy-conv和多级关注对整体网络性能的重大贡献。所提出的模型在解码MI-EEG方面优于当前的DL架构使用BCI-IV2a基准数据集，独立受试者的准确率为71.3%，依赖受试者为87.08%。改进后的模型在对小而有挑战性的数据集进行训练时，表现出了从原始大脑信号中检测MI活动的强大能力，而无需去除伪影，也几乎无需预处理。改进后的模型提高了脑机接口-IV2a数据集中所有受试者和所有心肌梗死活动的MI-EEG分类性能，证明了其检测各种受试的能力。模型经过改进后，因为其结构简单，未来可以考虑将其布置在物联网设备上面。希望未来可以对模型继续改进，提高MI-EEG分类准确率，早日实现更高级的脑机接口的应用。

参考文献

- [1]H. Altaheri, M. Alsulaiman, and G. Muhammad, “Date fruit classification for robotic harvesting in a natural environment using deep learning,” *IEEE Access*, vol. 7, pp. 117115 – 117133, 2019.
- [2]G. A. Altuwaijri, G. Muhammad, H. Altaheri, and M. Alsulaiman, “A multi-branch convolutional neural network with squeeze-and-excitation attention blocks for EEG-based motor imagery signals classification,” *Diagnostics*, vol. 12, no. 4, pp. 1 – 16, 2022.
- [3]D. Li, J. Xu, J. Wang, X. Fang, and J. Ying, “A multi-scale fusion convolutional neural network based on attention mechanism for the visualization analysis of EEG signals decoding,” *IEEE Trans. Neural Syst.Rehabil. Eng.*, vol. 28, pp. 2615 – 2626, 2020.
- [4]T. Liu and D. Yang, “A densely connected multi-branch 3D convolutional neural network for motor imagery EEG decoding,” *Brain Sci.*,vol. 11, no. 2, pp. 1 – 24, 2021.
- [5]S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain, “Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion,” *Future Gener. Comput. Syst.*,vol. 101, pp. 542 – 554, Dec. 2019.
- [6]G. A. Altuwaijri and G. Muhammad, “Electroencephalogram-based motor imagery signals classification using a multi-branch convolutional neural network model with attention blocks,” *Bioengineering*, vol. 9,no. 7, p. 323, 2022.
- [7]Y. K. Musallam et al., “Electroencephalography-based motor imagery classification using temporal convolutional network fusion,” *Biomed.Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102826.
- [8]D. Zhang, K. Chen, D. Jian, and L. Yao, “Motor imagery classification viatemporal attention cues of graph embedded EEG signals,” *IEEE J. Biomed.Health Inform.*, vol. 24, no. 9, pp. 2570 – 2579, Sep. 2020.
- [9]M.-T. Luong, H. Pham, and C. D. Manning, “Effective approaches to attention-based neural machine translation,” in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Lisbon, Portugal, Sep. 2015,pp. 1412 – 1421.
- [10]S. U. Amin, H. Altaheri, G. Muhammad, W. Abdul, and M. Alsulaiman, “Attention-inception and long short-term memory-based electroencephalography classification for motor imagery tasks in rehabilitation,” *IEEE Trans. Ind. Inform.*, vol. 18, no. 8, pp. 5412 – 5421, Aug. 2022.
- [11]S. U. Amin, H. Altaheri, G. Muhammad, M. Alsulaiman, and W. Abdul, “Attention based inception model for robust EEG motor imagery classification,” in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, 2021, pp. 1 – 6.
- [12]V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain – computer interfaces,” *J. Neural Eng.*, vol. 15, no. 5,2018, Art. no. 56013.
- [13]D. Li, J. Xu, J. Wang, X. Fang, and Y. Ji, “A multi-scale fusion convolutional neural network based on attention mechanism for the visualization analysis of EEG signals decoding,” *IEEE Trans. Neural Syst. Rehabil.Eng.*, vol. 28, no. 12, pp. 2615 – 2626, Dec. 2020.
- [14]S. Kumar, R. Sharma, and A. Sharma, “OPTICAL+: A frequency-based deep learning scheme for recognizing brain wave signals,” *Peer J. Comput.Sci.*, vol. 7, 2021, Art. no. e375.
- [15]T. Luo and F. Chao, “Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network,” *BMC Bioinf.*, vol. 19, no. 1, 2018, Art. no. 344.

- [16]J. Xu, H. Zheng, J. Wang, D. Li, and X. Fang, “Recognition of EEG signal motor imagery intention based on deep multi-view feature learning,” *Sensors*, vol. 20, no. 12, 2020, Art. no. 3496.
- [17]A. Hassanpour, M. Moradikia, H. Adeli, S. R. Khayami, and P. Sham sinejadbabaki, “A novel end-to-end deep learning scheme for classifying multi-class motor imagery electroencephalography signals,” *Expert Syst.*,vol. 36, no. 6, 2019, Art. no. e12494.
- [18]S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain, “Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion,” *Future Gener. Comput. Syst.*, vol. 101,pp. 542 – 554, 2019.
- [19]M.-A. Li, J.-F. Han, and L.-J. Duan, “A novel MI-EEG imaging with the location information of electrodes,” *IEEE Access*, vol. 8, pp. 3197 – 3211,2020.
- [20]T. Liu and D. Yang, “A densely connected multi-branch 3D convolutional neural network for motor imagery EEG decoding,” *Brain Sci.*, vol. 11, no. 2, 2021, Art. no. 197.
- [21]Hamdi Altaheri, Ghulam Muhammad, Mansour Alsulaiman. Physics-inform attention temporal convolutional network for EEG-based motor imagery classification[J], *IEEE Transactions on Industrial Informatics*, 2022, 19(2): 2249-2258.
- [22]J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proc.IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132 – 7141.