

基于脑电图的情绪识别，通过通道式注意和自我注意来实现

陶伟、李昌、宋仁成、程娟、刘宇、万锋和陈勋

摘要--基于脑电图（EEG）的情感识别是脑机接口领域的一项重要任务。最近，许多基于深度学习的情感识别方法被证明优于传统方法。然而，提取脑电图情感识别的鉴别性特征仍然具有挑战性，而且大多数方法忽略了通道和时间方面的有用信息。本文提出了一种基于注意力的卷积递归神经网络（ACRNN），以从EEG信号中提取更多的鉴别性特征，提高情绪识别的准确性。首先，所提出的ACRNN采用了一种通道导向的注意力机制来自适应地分配不同通道的权重，并采用了一个CNN来提取编码的EEG信号的空间信息。然后，为了探索脑电信号的时间信息，扩展的自我注意被整合到一个RNN中，根据脑电信号的内在相似性重新编码重要性。我们在DEAP和DREAMER数据库上进行了广泛的实验。实验结果表明，提出的ACRNN优于最先进的方法。

索引词-脑电图(EEG)，情绪识别，通道式注意，自我注意。

1 简介

• 通讯作者:Chang Li.

动作分析在日常生活中非常重要，尤其是在人机交互领域[1]-[5]。情感分析可以帮助提高人机交流的质量，改善计算机的智能。此外，情绪分析在医疗保健中发挥着重要的作用，可以了解病人的行为和认知功能[6]，[7]，生理信号通常被用来测量情绪状态，包括皮肤电反应、肌电图、心率、呼吸率和脑电图（EEG）[8]。

在过去的十年中，人们对情绪和脑电信号之间的关系进行了广泛的研究[9]-[11]。脑电图

这项工作得到了国家重点研发计划（2019YFA0706203）、国家自然科学基金（61922075、41901350和61701160）的资助。安徽省自然科学基金（2008085QF285和1808085QF186），安徽省重点研发计划（1804h08020244）和中央高校基础研究基金（JZ2019HGBZ0151和JZ2020HGPA0111）。

- W.Tao在合肥工业大学生物医学工程系，合肥230009，中国，同时也在澳门大学电子和计算机工程系，澳门，中国。电子邮件：taovi1996@mail.hfut.edu.cn。
- C.Li, R. Song, J. Cheng and Y. Liu是合肥工业大学生物医学工程系的学生，合肥230009，中国。电子邮件：{changli, rcsong, chengjuan, yuliu}@hfut.edu.cn。
- F.Wan在中国澳门大学电气和计算机工程系工作。电子邮件：fwan@um.edu.mo。
- X.Chen在中国科技大学第一附属医院神经外科，生命科学与医学部，同时也在在中国科技大学电子工程与信息科学系，安徽合肥，230001，中国。电子邮件：xunchen@ustc.edu.cn。

脑电图信号很容易获得，它测量的是大脑神经元中的离子流产生的电压波动[12]。脑电图是无创的、快速的、廉价的，因此使它成为大脑对情绪刺激反应的首选方法。此外，脑电图信号被广泛用于情绪分析，因为脑电图可以从频段、电极位置和时间信息中探索出有关情绪的各种信息[13]-[15]。

一般来说，大多数EEG情绪识别方法首先从EEG信号中设计特征，并采用分类器对情绪特征进行分类。例如，Li等人从伽马频段提取特征，并采用线性支持向量机（SVM）对提取的特征进行分类[13]。Patil等人采用高阶交叉点作为特征，比其他统计特征更好地对情绪进行分类[16]。Shi等人首次提出了来自五个频段的微分熵（DE）特征，并评估了DE特征在表示EEG信号方面的优越性[17]。此外，Duan等人从多通道EEG数据中提取DE特征，并结合SVM和k-Nearest Neighbor（KNN）对提取的特征进行分类[18]。

最近，深度学习已经被证明能够超越在许多领域进行传统的机器学习，例如计算机视觉[19]、自然语言处理[20]和生物医学信号处理[21]-[23]。此外，许多基于深度学习的方法已经被广泛用于基于EEG的情绪识别。一方面，深度学习方法可以被认为是特征提取后的分类器。例如，Yang等人将多个频段的DE作为EEG特征，并采用连续卷积神经网络作为分类器[24]。Song等人根据电极位置关系设计DE特征，并采用图卷积神经网络作为分类器[25]。另一方面，许多深度学习方法都是数据驱动的，以端到端的方式运作，这不需要手工制作的特征

从EEG信号中。例如, Alhagry 等人提出了一个端到端的深度学习神经网络来识别原始EEG信号的情绪, 该网络使用LSTM-RNN从EEG信号中学习特征, 并使用密集层进行分类[15]。Yang 等人提出了一个平行卷积递归神经网络用于EEG情绪识别, 并取得了良好的性能[7]。然而, 为EEG情感识别提取更多的鉴别性特征仍然是一个挑战。因此, 设计一个有效的深度学习框架, 能够直接从原始EEG信号中提取特征并进行分类是非常重要的。

受级联卷积递归网络 (CRNN) 的启发, 该网络结合CNN和RNN从EEG信号中提取空间和时间特征[26], 我们使用CNN来提取EEG信号的空间信息。然后, 我们采用两个长短时记忆 (LSTM) 层来提取时间信息, 这比标准的RNN更善于存储和访问信息[27]。与传统的CRNN不同, 我们采用了一个框架, 利用两种注意机制, 即通道式注意机制[28]和扩展的自我注意机制[29], 来提取更具辨别力的时空信息。一般来说, CNN被用来提取EEG信号的空间特征[7], 然而, 这忽略了不同通道之间特征的重要性。为了从空间信息中提取更多的判别特征, 一些方法采用通道选择来选择更多相关通道[30]。与传统方法需要首先人为地选择相关通道不同[31], 在本研究中, 我们首先采用了一种自适应通道机制, 将通道转换为概率分布作为权重, 并根据转换后的权重对EEG信号进行重新编码。然后, 采用CNN来提取重新编码的信号的鉴别性空间特征。此外, 采用RNN来探索EEG信号的时间信息, 然而, 这也忽略了不同EEG样本的重要性。请注意, 扩展的自我注意可以应用于LSTM来利用长距离的依赖关系[32]。我们将扩展的自我注意机制整合到RNN中, 以探索不同EEG样本的重要性, 因为这种机制可以根据EEG信号的相似性来更新权重。因此, 通过在我们的框架中整合两种注意力机制, 可以获得脑电信号更多的时间和空间特征的判别。

在本文中, 我们提出了基于注意力的对话。递归神经网络 (ACRNN) 处理基于脑电的情感识别问题。原始脑电信号可以通过不同通道之间的内在关系和时间片之间的时间依赖性来包含空间信息, 因此, 所提出的ACRNN可以在卷积层中学习多通道脑电的空间特征, 并使用LSTM网络探索不同时间片的时间特征。此外, 通道明智的注意和扩展的自我注意机制可以分别提取更具辨识度的空间和时间特征。我们在两个公开可用的数据库, 即DEAP[2]和DREAMER[3]上对所提出的模型进行了评估, 所提出的方法在两个数据库中表现出与识别精度相关的优异性能。我们的主要

贡献总结如下。

1) 我们开发了一个数据驱动的ACRNN框架, 用于基于EEG的情绪识别。该框架将信道注意力机制整合到CNN中以探索空间信息, 通过信道注意力考虑不同信道的重要性, 通过CNN考虑多信道EEG信号的空间信息。此外, ACRNN在RNN中集成了扩展的自我注意机制来探索脑电信号的时间信息, 可以考虑LSTM的不同时间信息和扩展的自我注意的每个脑电样本的内在相似性。

2) 我们对DEAP和实验结果表明, 在DEAP数据库的情感价值和唤醒分类任务中, 平均情感识别准确率分别为92.74%和93.14%。此外, 所提出的方法在DREAMER数据库的价值、唤醒和支配性分类任务中的平均准确率分别为97.79%、97.98%和97.67%。

本文的其余部分组织如下。第二节介绍了相关工作, 第三节介绍了拟议的方法。第四节讨论了为证明所提出的ACRNN的有效性而进行的大量实验。最后, 第五节给出了一个讨论, 第六节是本文的结论。

2 相关的工作

在这里, 我们介绍了传统EEG情感识别框架的一般流程。然后, 我们介绍通道式注意和自我注意的机制。

2.1 脑电图情感识别的一般流程

最近, 从EEG信号中进行的情感识别已经重新得到了极大的关注。脑电图情感识别框架的一般流程总结如下 (图1)。

(i) 测试协议。首先, 使用的刺激物类型, 试验记录持续时间、受试者人数、他们的性别和要识别的情绪。然后, 受试者被暴露在刺激下, 如音乐或电影[2], [3]。

(ii) 脑电图的记录。电极的数量和测试持续时间被记录下来, 然后通过电极记录脑电信号。然后受试者在每次试验后通过标记脑电图记录来评估他们的情绪状态[2], [3]。

(iii) 预处理。为了避免脑电图信号中的伪影。如眨眼, 脑电图信号应使用去除伪影的方法进行预处理, 如盲源分离和独立成分分析[33]。

(iv) 特征提取。为了提取相关的情感特征从脑电图信号中获得的信息, 探索有关脑电图信号的信息, 例如, 在时间、频率和空间域的脑电图特征[9]。

(v) 各种分类器可用于对前者进行分类。牵引的特征, 例如贝叶斯、支持向量机、决策树和深度学习分类器[34]。根据分类器是否在用户依赖的数据上进行训练, EEG情感识别也可以分为用户依赖和用户独立的任务。

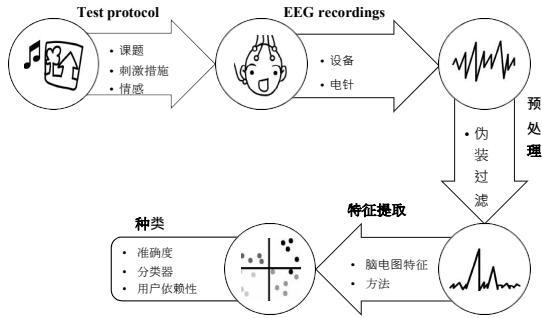


图1：EEG情感识别的一般流程。

2.2 渠道方面的关注

注意力在人类的感知中起着重要的作用[35], [36]。例如，人类可以利用一连串的部分瞥见，有选择地集中在突出的部分，以更好地捕捉视觉结构[37]。受人类注意机制的启发，人们为各种视觉任务提出了空间注意机制，例如，语义注意[38]、多层注意[39]和通道明智注意[28]。通道明智的注意表现出超强的或性能，因为它可以改变不同通道的权重来探索特征图的信息；因此，它可以提取更多关于通道的重要信息。因此，通道明智的注意机制被用来利用特征通道之间的相互依赖关系。例如，Hu 等人引入了一个紧凑的模块来利用特征图的通道间关系[40]，Chen 等人将空间和通道明智的注意力结合起来用于图像字幕[28]。

一般来说，渠道方面的关注可以挤压全球的空间信息，并产生信道的统计信息[28]。此外，它可与CNN一起训练，因此，它可以被整合到CNN架构中[41]。考虑到多通道EEG信号通过通道包含空间信息，通道明智的关注可以被整合到CNN中，以探索EEG信号的通道之间的重要性，并且可以通过CNN提取更多的鉴别性的空间信息。

2.3 自我关注

自我关注是一种内部关注机制，它将单个序列的不同位置联系起来，根据重要性评分对序列数据进行编码[20]。此外，自我注意机制之所以受欢迎，是因为它可以改善长距离的依赖性建模[42]。一个注意力函数可以被描述为将一个查询和一组键值对映射到一个输出，其中查询、键、值和输出都是向量。输出被计算为数值的加权和，其中分配给每个数值的权重是由查询与相应的键的兼容性函数计算的。自我关注已经被证明在简单语言的问题回答和语言建模任务中表现良好。例如，Vaswani 等人提出了一个基于注意力的机器翻译架构[32]，Shen 等人提出了方向性的自我注意力网络，以关注输入序列中elements之间的注意力[29]。在EEG识别任务中。

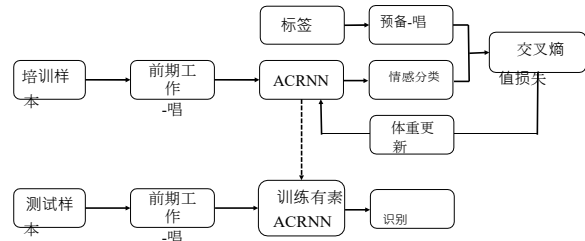


图2：基于注意力的卷积递归神经网络对基于EEG的情绪识别的概述。

为了增加训练样本的数量，一个脑电图试验常常被分割成几个输入样本。然而，许多方法忽略了不同脑电图样本的重要性。受自我注意的启发，我们采用这种技术来进一步探索EEG样本之间的时间依赖性。

3 建议的方法

在本节中，我们首先介绍了所提出的EEG情绪识别框架，然后介绍了我们的原始EEG信号预处理技术。最后，我们详细描述了拟议的ACRNN的构造。

3.1 拟议的ACRNN的框架

一般来说，大多数基于EEG的情绪识别研究都集中在首先提取相关的特征，然后用提取的特征对受试者的电子情绪状态进行分类[8], [14], [25]。在实践中，原始EEG信号包含丰富的空间和时间信息，这些信息可以被提取出来以识别受试者的情绪状态。所提出的ACRNN是一种数据驱动的方法，它同时将通道明智[28]和扩展的自我注意[29]机制整合到CNN-RNN。此外，ACRNN可以提取空间和时间信息作为情绪特征，并使用softmax函数对提取的特征进行分类。因此，这种端到端的技术提高了基于EEG的情绪识别的准确性(图2)。首先，我们将EEG样本分为训练和测试样本。然后，通过去除基线信号，分别对训练和测试样本进行预处理。此外，使用切片窗口技术对标签进行预处理。接下来，我们使用训练样本来训练拟议的ACRNN模型，计算交叉熵损失并使用Adam optimizer[43]更新网络参数。最后，用训练好的模型来识别测试样本的情绪状态，分类准确率被认为是最终的识别结果。

3.2 拟议的ACRNN的预处理

在拟议的ACRNN中，预处理包括去除基线信号和滑动窗口。一般来说，记录的EEG信号包含基线和试验信号[2], [3]。Yang 等人提出，去除基线的预处理可以提高DEAP数据库的EEG情绪识别率

[7]。这里，让 $\mathbf{X}_R = [\mathbf{X}_B, \mathbf{X}_T]$ $R \in N$ 是记录的EEG信号，采样频率为 H Hz，持续时间为 T_1 ，其中 M 是EEG电极节点的数量， N 是

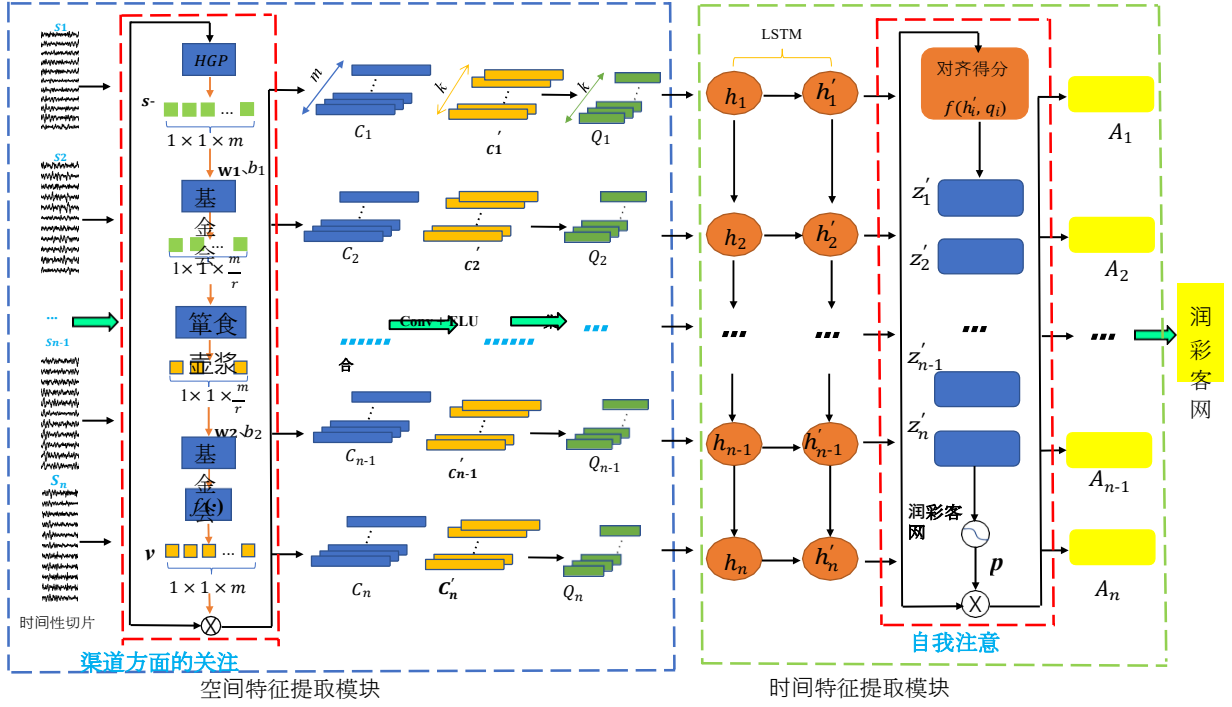


图3：基于注意力的卷积递归神经网络的结构图。

采样点的数量。此外， $\mathbf{X}^{RM \times B \times L}$ 表示持续时间为 T 的基线信号 2 ， L 表示采样点的数量， X_i ($i=1, 2, \dots, T_2$) $^{RM \times H}$ 表示第 i 次基线信号。因此，每秒钟基线信号的平均值可以表述如下。

$$\bar{\mathbf{X}}\mathbf{B} = \frac{\sum_{i=1}^L \mathbf{X}_i}{L}, \quad (1)$$

其中 $\bar{\mathbf{X}}\mathbf{B}^{RM \times B \times H}$ 表示每秒钟的基线信号的平均值。让 $\mathbf{X}^{RM \times T \times J}$ 代表持续时间为 T 的试验脑电信号 3 ，其中 J 是采样点的数量。为了去除试验脑电信号的基线，用一秒钟的非重叠滑动窗口 τ 将 \mathbf{X} 分割成若干片 \mathbf{X}_j ($j=1, 2, \dots, T_3$) $^{RM \times H}$ ，每秒钟去除基线的信号可以表述为：

$$\mathbf{X}_j^I = \mathbf{X}_j - \bar{\mathbf{X}}\mathbf{B}_0. \quad (2)$$

最后，这些基线信号去除时间片 \mathbf{X}_j^I 被串联成一个新的矩阵 $\mathbf{X}\tau^I \in^{RM \times J}$ 。一般来说，为了增加训练数据的数量，一个脑电图试验 $\mathbf{X}\tau^I \in^{RM \times J}$ 通常被分割成几个时间片 $\mathbf{S} = \{\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n\}$ 的滑动窗口。这里， \mathbf{S}_i ($i=1, 2, \dots, n$) $^{RM \times J}$ 代表第 i 个EEG样本， T 表示每个滑动窗口的采样点数量。一般来说，人类的情绪状态持续时间为1秒到12秒，以前的研究表明，3秒的滑动窗口可以达到良好的分类精度[44]，因此，我们采用3秒的滑动窗口， $J/T/H=3$ 。

3.3 拟议的ACRNN的构建

拟议的ACRNN包括信道明智的注意机制、CNN、RNN和扩展的自我注意机制。拟议的ACRNN-N的结构如图3所示。图的左边显示了

空间特征提取模块。首先，为了探索多通道脑电信号中不同通道之间的重要性，我们将注意力机制以通道的方式运用到脑电信号中（图3）。在实际的脑电信号采集中，多通道设备中的不同脑电通道往往包含多余的或不太重要的信息。为了提高情绪识别的准确性，一些方法采用通道选择的方法来选择更多相关通道[30]。与传统的需要人为选择相关通道的方法不同[31]，我们采用了自适应通道选择机制，它可以考虑所有通道的信息，并根据重要性给不同的通道分配权重。在我们的框架中， $\mathbf{S} = \mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n$ 代表预处理后的EEG样本， $\mathbf{S}_i = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m]$ ($i=1, 2, \dots, m$) 是第 i 个EEG样本，其中 \mathbf{s}_j ($j=1, 2, \dots, m$) 代表EEG样本 \mathbf{S} 的第 j 个通道， m 是每个样本的总通道数。在这个模型中，我们首先对脑电图样本的每个通道进行均值汇集，以获得如下的通道明智的统计数据。

$$\mathbf{s}^- = \mathbf{s}_1^- \mathbf{s}_2^- \dots \mathbf{s}_m^-, \quad (3)$$

其中 \mathbf{s}_j^- ($j=1, 2, \dots, m$) 是第 j 个通道的平均值。为了降低模型的复杂性和提高泛化能力，通道明智的注意力机制在非线性周围采用了两个全连接（FC）层， 1 一个具有参数 \mathbf{W}_1 和偏置项 b_1 的降维层，以降维比 r 和tanh函数作为激活函数，以及一个具有参数 \mathbf{W}_2 和偏置项 b_2 的增维层 2_0 。因此，通道式注意力的门控机制表示如下。

$$\mathbf{v} = \text{softmax}(\mathbf{W}_2 - \tanh(\mathbf{W}_1 \mathbf{s}^- + b_1 + b_2)). \quad (4)$$

其中softmax函数将通道的重要性转化为概率分布 $\mathbf{v} = [v_1, v_2, \dots, v_m]$ ，表示不同通道的重要性。最后，我们认为概率是对每个通道中的EEG样本 $\mathbf{S} = [s_1, s_2, \dots, s_m]$ 的信息进行重新编码的重量。因此，通过通道明智的注意力提取的第 j 个 ($j = 1, 2, \dots, m$) 周到的通道feature可以表示如下。

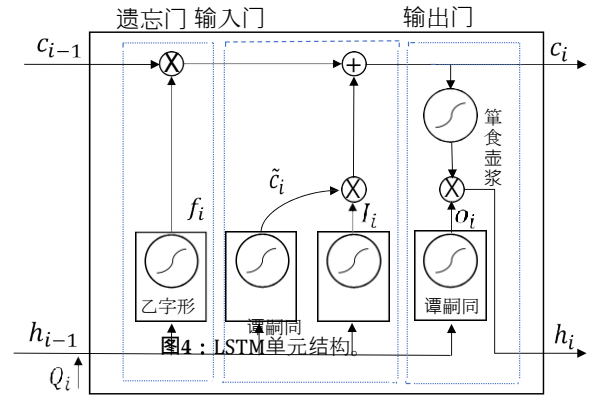
$$C_j = v_j \cdot S_j, \quad (5)$$

因此， $\mathbf{C} = C_1, C_2, \dots, C_m$ 代表提取的信道特征，第 i 个提取的特征 $C_i = [c_1, c_2, \dots, c_m]$ 可以通过 $\mathbf{S} = [s_1, s_2, \dots, s_m]$ 的每个信道和 $\mathbf{v} = [v_1, v_2, \dots, v_m]$ 的每个元素之间进行信道的叠加得到。

然后，我们使用CNN来进一步提取空间信息。脑电信号的核算，其中卷积核的数量为 k ，核的高度与电极的数量相同。这里，核的宽度也是为了探索脑电信号的时间信息。此外，我们使用指数线性单元 (ELU) 函数作为卷积运算的激活函数，这比常用的整流线性单元 (ReLU) 函数更好[45]。因此，第 i 个特征 C^{it} ($i=1, 2, \dots, n$) 可以在卷积和激活操作后从第 i 个通道的殷实特征 \mathbf{C} 得到。

之后，我们采用池化层来减少参数的数量，进一步提取特征。这里，池化后的第 i 个编码表示为 $Q_i = \text{MaxPool}(C^{it})$, $i = 1 \dots n$ 。

结构图的右侧显示的是温度。孔隙特征提取模块 (图3)，它包括一个两层的LSTM和扩展的自我注意机制。LSTM网络可以学习序列的上下文信息，因为它是基于递归结构的[27]。LSTM网络已经成功地用于EEG情绪识别，因为它可以从EEG数据中学习基于时间依赖性的特征[15]。如图4所示，一个LSTM单元接收三个输入，即当前时间 i 的输入 Q_i ，先前时间 $i-1$ 的输出 c_{i-1} ，以及 $i-1$ 代表先前时间 $i-1$ 的隐藏状态的 h 。然后，LSTM单元输出两个输出，即当前时间 i 的输出 c_i 和代表从LSTM提取的第 i 个时间特征的隐藏状态 h_i 。LSTM单元包含三个门，即输入门、遗忘门和输出门，通过sigmoid和tanh激活函数控制数据流。虽然这里的训练样本和测试样本是不同的，而且在时间上不是连续的，但在空间特征提取模块之后，编码的样本包含空间信息。同时，输入



网络采用两个叠加层来记忆和编码所有扫描的空间和时间区域[47]，因此，我们将LSTM层的数量设置为两个。因此，LSTM网络的第 i 个输出是第二个递归层的隐藏状态 h_i ， $h_i = \text{lstm}(Q_i)$, $i = 1 \dots n$ 。

为了提取更多的鉴别性时间信息，我们采用扩展的自我注意机制[29]，通过探索每个样本的内在重要性，为每个EEG信号样本分配权重。自我注意的结构如图5所示。与传统的自我注意机制不同，传统的自我注意机制是用来给每个再当前编码的片段分配重要性，并将这些信息汇总形成最终的表示[48]，扩展注意是加性注意在多维特征层面的自然延伸。它可以通过计算每个样本内部不同点的相似性来更好地描述具体的含义，得到的 z^{it} 可以看作是第 i 个样本 h_i 特征明智的分数向量。此外，扩展的自我注意机制在激活函数的内侧和外侧增加了两个偏置项，第 i 个特征明智的分数向量 z^{it} 可以表示为：

$$z^{it} = f(h_i, q_i) = W^T \sigma(W_1 h_i + W_2 q_i + b_1) + b, \quad (6)$$

其中， $f(h_i, q_i)$ 表示第 i 个编码的EEG样本的内在相似性， q_i 是基于特征向量 h_i 通过线性转换生成的对齐模式向量，其中维度与特征向量相同。这里，激活函数 $\sigma(\cdot)$ 是一个指数线性单元 (ELU)， W 和 b 分别是 σ 函数的权重和偏置项， W_1, W_2 是权重参数， b_1 是偏置项。那么， $\mathbf{p} = p_1, p_2, \dots, p_n$ 表示所有样本的概率，第 i 个EEG样本的概率可表示为：

$$p_i = \frac{\exp z_i^{it} \cdot h_i^t}{\sum_{i=1}^n \exp z_i^{it} \cdot h_i^t} \quad (7)$$

最后， $\mathbf{A} = A_1, A_2, \dots, A_n$ 表示由扩展的自我注意机制提取的特征，以及由扩展的自我注意机制提取的第 i 个注意特征。

机制可以计算如下。

$$A_i = p_i \cdot h_i^t, \quad (8)$$

门权重的组织是为了学习空间信息，而遗忘权重组织学习更多的时间性内容信息，输入门和遗忘门的竞争。

相互之间分别向单元输入新的信息或保持当前的时间信息[46]。因此，LSTM网络可以提取时空特征。

在本研究中，每层的LSTM单元数与EEG样本数相同，每个时间步长的输出可视为从每个样本中提取的时间信息。一般来说，LSTM

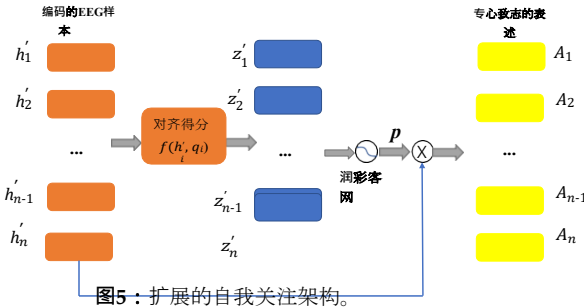


图5：扩展的自我关注架构。

在拟议的ACRNN的最后部分，我们采用softmax层作为分类器。提取出的空间情感特征为 $\mathbf{A} = \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ ，而softmax分类器接收这些提取的特征作为输入来识别情绪，如下所示。

$$\mathbf{P} = \text{softmax}(\mathbf{W}\mathbf{A} + \mathbf{b}), \quad (9)$$

其中 $\mathbf{P} = \mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n$, \mathbf{P}_i ($i = 1, 2, \dots, n$) 表示第 i 个 EEG 样本的预测概率， \mathbf{W} 和 \mathbf{b} 分别是 softmax 函数的权重和偏置项。然后，对所有标记的样本的交叉熵误差进行评估。

$$= \mathbf{L} - \sum_{i=1}^n \hat{Y}_i \log(\mathbf{P}_i). \quad (10)$$

其中， \hat{Y}_i 是第 i 个 EEG 样本的标签，交叉熵误差越小，说明情绪识别的准确性越高。

综上所述，我们设计了一个从原始 EEG 信号中提取特征并进行情感分类的框架。我们采用信道关注机制来自适应分配不同信道的权重，这可以提取信道之间的内在信息，然后采用 CNN 来提取编码后的 EEG 信号的空间信息。此外，我们采用两层 LSTM 来探索不同 EEG 样本的时间信息，并结合扩展的自我注意机制，根据每个样本的重要性为 EEG 样本分配权重。最后，可以得到用于 EEG 情绪识别的时空注意特征。

4 实验

在这里，我们介绍了两个广泛使用的数据库。然后，我们介绍了六种深度学习方法和两种传统方法进行比较。然后，我们在实验中演示了模型的实现。最后，我们介绍并比较了所提出的方法和所比较的方法得到的实验结果。

4.1 数据材料

为了验证所提出的 ACRNN 的性能，我们在两个广泛使用的数据库上进行了实验，即使用生理信号的情绪分析数据库 (DEAP) [2] 和通过 EEG 和 ECG 信号的情绪识别数据库 (DREAMER) [3]。DEAP

数据库包括 32 名参与者在观看 40 段音乐视频时记录的脑电图和周边生理信号。该数据库包含 32 个通道的脑电信号和 8 个通道的外周生理信号，其中脑电信号用于情绪识别，而外周生理信号则被弃用。在这个实验中，脑电信号以 512 赫兹采样，然后降频到 128 赫兹。此外，使用盲源分离技术去除脑电图 (EOG) 伪影。经过预处理的 EEG

每个试验的数据包含 60 秒的试验数据和 3 秒的基线数据。情感音乐视频包括 40 个一分钟的片段，参与者被要求记录他们对每个视频的唤醒、价值、喜欢和支配的水平，从 1 到 9。在我们的实验中，我们选择了价值和唤醒作为情感评价标准，根据唤醒和价值的额定水平将试验分为两类的阈值被设定为 5。每个主题文件包含两个数组，这些文件的数据格式详见表 1。

表1
DEAP数据库

阵列名称	阵列形状	阵列中心点
数据	40 × 40 × 8064	视频/试验 × 频道 × 数据
标签	40 × 4	视频/试用 × 标签

第二个数据库是 DREAMER (表 2)，它是一个多模态数据库，记录了通过视听刺激激发情感时的 EEG 和 ECG 信号。记录了 23 名参与者 (14 名男性和 9 名女性) 的信号，并要求参与者记录每次刺激后的唤醒程度、价值和主导地位。脑电图信号以 128 赫兹的采样率记录，使用情感 EPOC 系统 [49]。每个电影片段都是 65 至 393 秒，这足以引起单一情绪。此外，记录的 EEG 信号包含基线信号，通常在每个电影片段之前持续 4 秒。此外，大多数眼部伪影 (眨眼、眼球运动、心脏干扰等) 都已用线性相位 FIR 滤波器去除。此外，为了避免多种情绪对数据的污染，在每个片段的最后 180 秒内拍摄的记录被用于进一步分析。评级值的阈值被置于中间，其中小于或等于 3 的值代表低价值、唤醒和支配，大于 3 的值代表高价值、唤醒和支配。

4.2 模型实施

经过预处理阶段，我们在 DEAP 数据库中为每个受试者共获得了 800 个 EEG 样本，其中每个样本为 \mathbf{S}_i ($i=1, 2, \dots, 800$) $\in \mathbb{R}^{32 \times 384}$ 。对于 DREAMER，我们为每个受试者总共获得了 1250 个 EEG 样本。其中每个样本是 \mathbf{S}_i ($i=1, 2, \dots, 1250$) $\in \mathbb{R}^{14 \times 384}$ 。我们对每个子项目的不同试验中的所有样本进行了洗牌。然后，我们用 10 倍交叉验证法来评估所提出的方法和基线方法的性能。10 倍验证过程的平均性能被作为最终的实验结果。该模型采用了 TensorFlow 框架，并在 NVIDIA TITAN Xp pascal GPU 上进行训练。此外，Adam

采用优化器来最小化交叉熵损失函数，并对网络参数进行了优化。学习率为 10^{-4} ，辍学正则化被设定为0.5。为了在训练过程中获得更好的性能，采用了批量归一化。卷积核的大小为 a b ，DEAP的高度被设定为 $a = 32$ ，而 a DREAMER=14，宽度为 $b=40$ 。核的数量为 $k = 40$ ，池子的大小为1 75，跨度为10。此外，我们将LSTM中隐藏状态的维度设置为64。

4.3 结果和分析

为了验证所提方法的有效性，我们在两个数据库上进行了广泛的实验。为了评估注意力机制的性能，我们设计了三个模型来证明通道式注意力和扩展的自我注意力机制的影响，包括CNN+RNN（CNN-RNN），通道式注意力机制+CNN+RNN（A-CNN-RNN）和CNN+RNN+扩展的自我注意力机制（CNN-RNN-A）。这些模型的细节显示在表3中。CNN-RNN模型由一个CNN和LSTM网络组成，旨在验证基线框架的有效性，它可以使用级联框架从原始EEG信号中提取情绪特征。A-CNN-RNN模型由通道明智的注意机制、CNN和LSTM网络组成，旨在验证基线框架的通道明智注意机制的有效性。CNN-RNN-A模型由一个CNN、LSTM网络和前述的自我注意组成，旨在验证基线框架的扩展自我注意机制的有效性。此外，我们将所提出的方法与最近的三种深度学习方法进行了比较：连续卷积神经网络（Conti-CNN）[24]，图卷积神经网络（GCNN）[25]，以及卷积递归注意力模型（CRAM）[48]。Conti-CNN可以结合多个波段的特征来提高识别精度[24]，GCNN可以采用不同的熵（DE）特征作为输入，并使用谱图过滤来提取特征和识别情绪[25]，而CRAM可以利用CNN来编码EEG信号的高层表示，并利用递归注意机制来探索时间动态[48]。此外，我们采用了两种传统的基于特征的分类器进行比较，包括支持向量机（SVM）和决策树（DT）[24]。所有的方法都经过与ACRNN相同的预处理，即去除基线信号和滑动窗口。

对于传统的分类器，我们使用DE特征作为in...把[17], [18]。DE特征具有区分EEG模式的低频和低频能量的平衡能力，通常在EEG情绪识别中作为频域特征使用[14], [47], [50]。根据文献[18]和[51]，带通滤波器被应用于EEG信号以获得子带信号，这些信号近似于高斯分布。因此，定义了五个子波段。1) delta（1-3赫兹）；2) theta（4-7赫兹）。3) α （8-13赫兹）；4) β （14-30赫兹）；和5) γ （31-50赫兹）。请注意，我们从侧面提取了DE特征。

表2
梦想家数据库

视听刺激	
视频的数量	18
视频	内容音频-视频
视频	持续时间65-393秒
(M = 199秒)。	
实验信息	
参与者的数量	23
男性人数	14
女性的数量	9
参与者的年龄	22-33
评级表	唤醒、价值、支配力
评级值	1-5
记录的信号	14通道128赫兹脑电图

表3
基线模型和基于注意力的EEG情绪识别模型

模型 \ 组成部分	渠道方面的关注	有线电视新闻网	LSTM网络	自我注意
CNN-RNN	×	✓	✓	×
A-CNN-RNN	✓	✓	✓	×
CNN-RNN-A	×	✓	✓	✓
ACRNN	✓	✓	✓	✓

四个子频段的信号，因为较高的频段（大约30-100赫兹）更适合于脑电的情绪识别[13]。最后的特征向量是所有通道的特征的串联。对于DEAP，最终的特征向量为 432=128维，每个受试者产生800个样本，其中每个样本 $\mathbf{X}^{R32 \times 128}$ ($i=1, 2, ..., 800$)。对于DREAMER，最终的特征向量是414=56维，每个主题产生1250个样本，其中每个样本 $\mathbf{X}^{R14 \times 56}$ ($i=1, 2, ..., 1250$)。

在我们的工作中，我们用提出的ACRNN在同一受试者身上进行了实验，并比较了依赖受试者的EEG情绪识别方法。我们将样本数据分为训练集和测试集，然后使用10倍交叉验证法[52]。通常情况下，10倍交叉验证将数据分为10个相等的数据子集，其中一个子集作为测试集，其他9个子集构成训练集。这个过程被重复了10次。对于DEAP数据库，训练样本的数量为720个，每个主体的剩余80个样本被用作测试样本。对于DREAMER数据库，训练和测试样本的数量分别为1125和125。

为了进一步分析渠道性关注的贡献，我们进行了计算渠道性关注的实验。

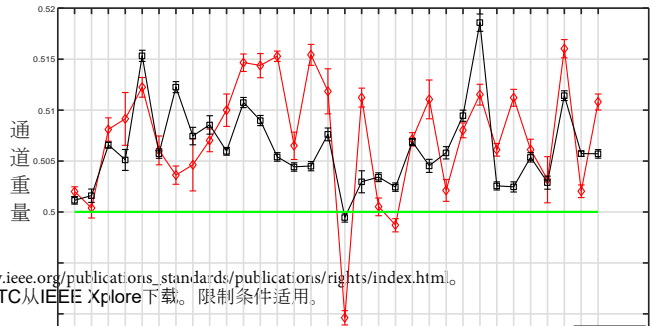




图6：DEAP数据库的平均通道权重。

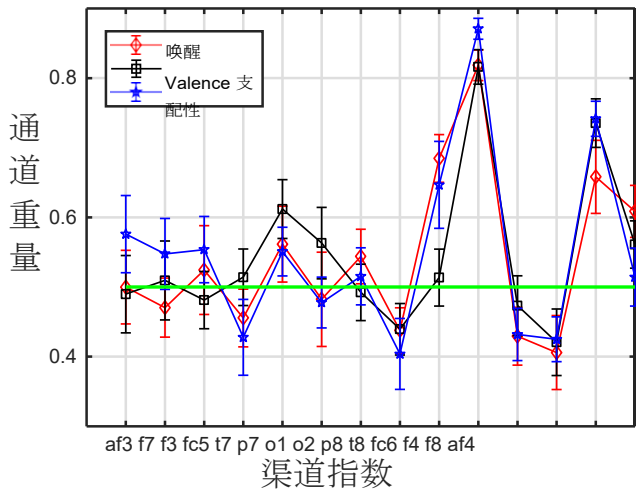


图7：DREAMER数据库的平均通道权重。

表4

在DEAP和DREAMER中，更大的重量通道和大脑区域之间的关系

数据库	频道	大脑区域
DEAP	F3	正面
	FC5	正面
	P3	顶叶
	P7	顶叶
	F8	正面
	Cz	中央
	C4	中央
DREAMER	T8	时间性
	F8	正面

通过通道的关注，脑电信号的权重。图6和图7分别显示了DEAP和DREAMER的平均通道权重。在DEAP数据集中有32个通道，在DREAMER数据集中有14个通道。如图所示，DEAP和DREAMER中的EEG信号的通道权重在通道导向的注意机制中都是不同的。图6显示，在DEAP数据库中，FC5、P3、C4、P8的通道权重在两个维度上明显大于其他通道。图7显示，在DREMAER数据库中，T8和F8的通道权重在三维空间上明显大于其他通道。在两个数据库中，电极是按照国际10-20系统放置的，较大权重的通道和脑区之间的关系如表4所示。该结果也与一些研究一致，这些研究表明与情绪有关的脑电信号主要分布在额叶、颞叶和顶叶[30], [31], [53]。更大的通道权重也表明给定的通道与情绪更相关，在EEG信号中更重要。

表5显示了在DEAP数据库中提出的和比较的方法在情感和唤醒方面的平均识别准确率。可以看出，与基线框架CNN-RNN相比，A-CNN-RNN的平均识别准确率在两个维度上分别提高了约30%和25%，因为通道导向的注意力集中在不同通道的空间特征上。此外，提议的ACRNN提高了平均的

与A-CNN-RNN相比，识别准确率提高了0.6%和0.5%，因为ACRNN结合了通道明智和扩展的自我注意机制来提取EEG信号的时空注意信息。此外，我们发现CNN-RNN-A与CNN-RNN相比，在两个维度上的识别精度提高了约27%，因为CNN-RNN-A的扩展自我注意机制根据每个样本的重要性来提取注意信息。此外，与CNN-RNN-A相比，提议的ACRNN在两个维度上的识别精度提高了约3%，因为提议的模型同时利用了两种注意力机制。与三种深度学习方法（Conti-CNN、CRAM和GCNN）相比，实验结果表明，所提出的ACRNN将平均识别准确率提高了10%，8%。

和5%，分别。与传统方法相比，ACRNN取得了比传统方法（如DT和SVM）更好的识别性能。

表6显示了在DREAMER数据库上得到的比较方法的平均识别精度。如图所示，在DREAMER数据库中，情感识别的准确性有明显的提高。例如，与Conti-CNN、GCNN和CRAM相比，所提出的ACRNN的平均识别精度分别提高了15%、9%和5%。因此，所提出的ACRNN可以在所有比较的方法中达到最好的识别精度，实验结果证明了将两种注意力机制整合到CNN-RNN中的有效性。

表5

不同方法在deap数据库的情绪和唤醒分类任务中的平均准确率（%）（平均值±std.）

	缬氨酸	唤醒
邓小平	75.95 ± 4.76	78.18 ± 5.45
证券公司	89.33 ± 7.41	89.99 ± 6.74
康蒂-康尼	82.77 ± 4.47	81.55 ± 6.55
CRAM	87.09 ± 7.49	84.46 ± 9.27
GCNN	88.24 ± 3.18	87.72 ± 3.32
CNN-RNN	62.75 ± 7.53	67.12 ± 9.13
A-CNN-RNN	91.48 ± 5.02	91.59 ± 5.42
CNN-RNN-A	89.15 ± 6.66	89.96 ± 5.93
ACRNN	93.72 ± 3.21	93.38 ± 3.73

表6

不同方法在梦者数据库的情感、唤醒和支配力分类任务中的平均准确率（%）（平均值±标准差）。

	缬氨酸	唤醒	优势
邓小平	68.81 ± 6.87	67.50 ± 7.28	67.43 ± 6.73
证券公司	76.71 ± 5.89	77.54 ± 5.62	75.76 ± 5.63
康蒂-康尼	81.72 ± 5.24	82.48 ± 5.11	82.58 ± 5.28
CRAM	92.27 ± 2.95	93.03 ± 1.87	93.34 ± 1.78
GCNN	88.87 ± 3.58	88.79 ± 3.86	88.54 ± 3.89
CNN-RNN	78.59 ± 13.87	77.66 ± 13.34	77.75 ± 14.22
A-CNN-RNN	97.47 ± 2.32	97.92 ± 1.60	98.15 ± 1.76
CNN-RNN-A	96.61 ± 3.42	97.36 ± 2.63	97.54 ± 2.16
ACRNN	97.93 ± 1.73	97.98 ± 1.92	98.23 ± 1.42

为了证明所提出的方法的性能和比较每个主题的方法，我们对每个主题进行了实验。图8、9、10、11和12显示了每个主体在每个维度上的平均准确度和标准偏差。可以看出，传统的

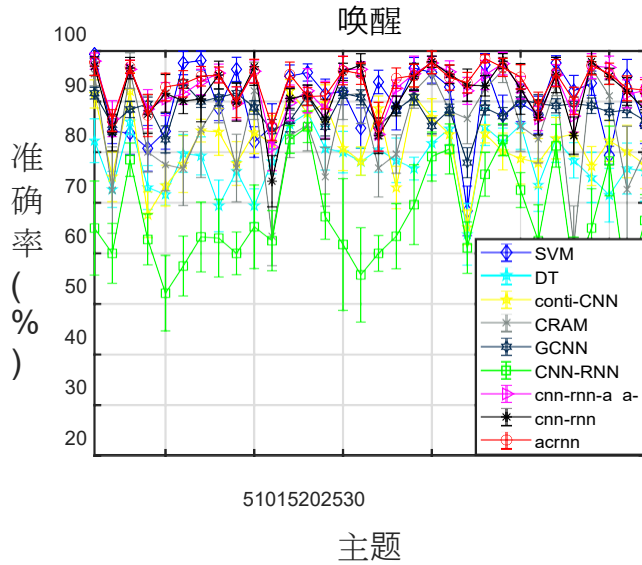


图8：不同方法在DEAP数据库的唤醒分类任务中对每个主题的平均准确率 (%)。

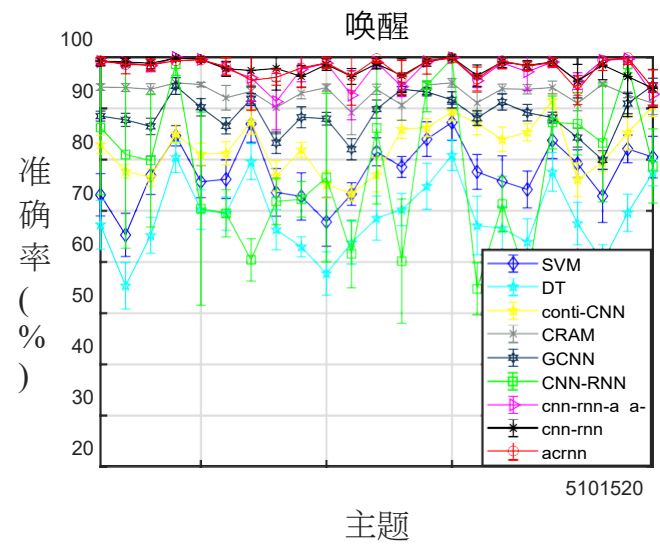


图10：不同方法在DREAMER数据库的唤醒分类任务中对每个主题的平均准确率 (%)。

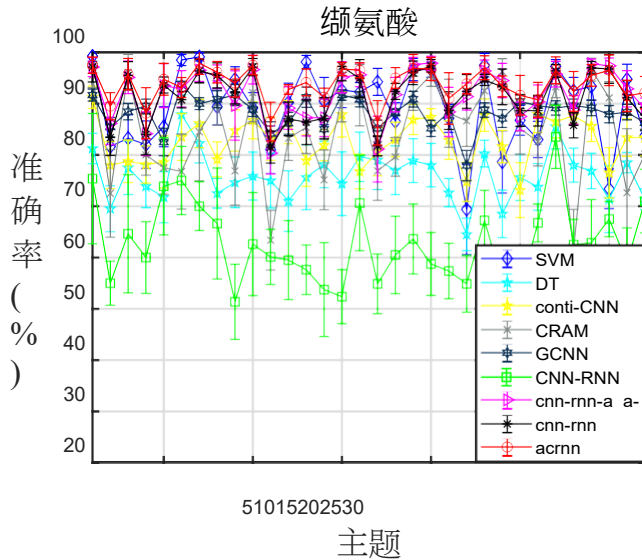


图9：不同的方法在DEAP数据库的情感分类任务中对每个主题的平均准确率 (%)。

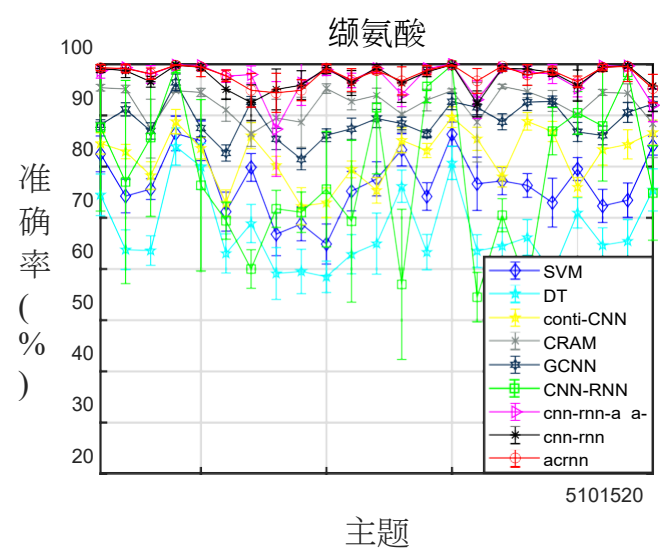
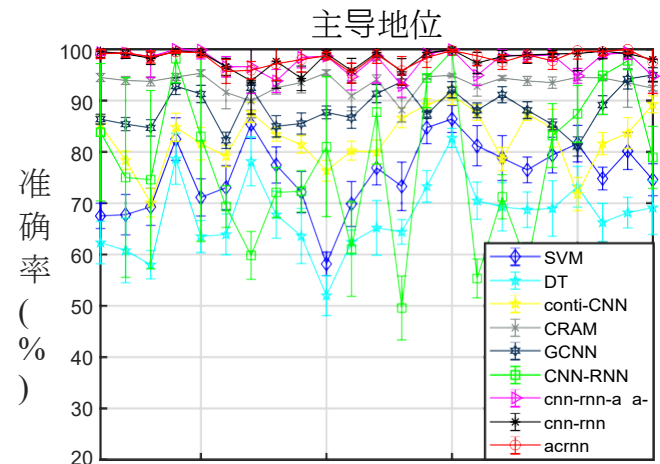


图11：不同方法在DREAMER数据库上的情感分类任务中对每个主题的平均准确率 (%)。

SVM和DT方法在一些受试者身上取得了良好的平均识别精度，然而，标准差非常大。此外，所比较的方法在一些受试者身上表现较差。然而，我们发现，三种基于注意力的方法在每个受试者身上都取得了较好的平均识别准确率，而且标准差也小于比较的方法。因此，实验结果表明，基于注意力的方法在每个受试者身上都能比比较的方法效果更好。此外，实验结果表明，所提出的ACRNN结合了通道明智的注意力模块和扩展的自我注意力，以利用更多的鉴别信息进行脑电情绪识别，并能在两个公共数据库上取得较好的识别精度。

5 讨论



基于脑电图的情绪识别被广泛用于帮助计算机更好地理解当

前的情绪状态。

主题

图12：不同方法在DREAMER数据库的支配力分类任务中对每个主题的平均准确率（%）。

用户[21]。传统的EEG识别方法首先从原始EEG信号中去掉手工制作的特征, 然后采用分类器来对这些特征进行分类。最近, 深度学习的方法采用了端到端的技术来识别原始EEG信号中的情绪。然而, 从原始信号中提取更具辨别力的特征用于EEG情感识别仍然是一个挑战。因此, 需要一个鉴别性的框架来识别原始EEG信号中的情绪状态。考虑到原始EEG信号通过不同通道之间的内在关系和时间片之间的时间差异包含了空间信息, 我们提出了ACRNN来提取空间和时间上的殷切信息, 并对受试者的情绪状态进行分类。

在这一框架中, 一个基于通道的注意机制通过给不同的通道分配权重, 从EEG信号中提取通道间的差异, 而CNN的设计是通过将EEG信号的所有通道进行卷积运算, 提取作为空间信息的特征图。与一些基于大脑认知功能的方法不同, 他们需要关注EEG通道并设计相关特征[54]-[56]。例如, Li等人用锁相值(PLV)构建了情绪相关的脑网络, 并采用多种特征融合的方法, 将补偿性激活和连接信息结合起来进行情绪识别[55], Wang等人将脑定向连接(BDC)和脑区间不同频段的DE特征结合起来, 提取鉴别性信息以提高识别精度[56]。然而, 我们提出的方法是采用通道关注来分配不同通道的权重。为了进一步分析不同通道的重要性, 我们计算了平均通道权重。从图6和图7中, 我们可以发现DEAP和DREAMER中EEG信号的信道权重根据信道明智的关注机制而出现不同。在DEAP数据库中, FC5、P3、C4、P8的通道权重明显大于二维上的其他通道, 而在DREMAER数据库中, T8和F8的通道权重明显大于三维上的其他通道。该结果表明, 与情绪相关的脑电信号主要分布在额叶、颞叶和顶叶, 这与现有研究[30], [31], [53]一致。也可以看出, 计算权重较大的通道与情绪的关系更大, 因此在基于EEG的情绪识别中更为重要。为了证明通道明智的注意力机制的有效性, 我们将通道明智的注意力整合到基线CNN-RNN框架中, 实验结果表明, 在DEAP和DREAMER数据库中, A-CNN-RNN的通道明智的注意力比CNN-RNN模型的平均准确率提高了约30%, 因为通道明智的注意力机制可以将通道转化为概率分布作为权重, 并根据转化的权重对EEG信号重新编码。此外, 扩展的自我注意机制旨在探索不同EEG样本的重要性。为了证明扩展的自我注意机制的有效性, 我们将其整合到基线CNN-RNN框架中, 实验结果表明, 扩展的自我注意机制是一个很好的例子。

与CNN-RNN模型相比, 在这些数据库上, 注意力可以提高平均准确率27%和29%。实验结果还表明, 扩展的自我注意机制通过基于样本之间的相似性对概率进行打分, 关注更重要的EEG样本。

综上所述, 在这两个数据库中, 通道式注意和扩展的自我注意机制提高了平均准确率, 超过了25%。这表明这些注意机制可以提高脑电图的识别率, 并取得相当的识别效果。然而, 与扩展的自我注意相比, 在DEAP和DREAMER数据库上, 通道明智的注意分别提高了大约2%和1%的平均识别精度。这表明, 通道明智的注意力比扩展的自我注意力表现得略好。综上所述, 所提出的ACRNN是一个级联框架, 它整合了通道式注意力和扩展的自我注意力机制。它可以同时有效地提取时空注意的特征。此外, 我们所有的实验结果都是通过10倍交叉验证得到的, 高准确率和低标准差也证明了所提出的ACRNN可以实现卓越的识别精度。

6 结论

在本文中, 我们提出了一种端到端的深度学习的方法用于EEG情感识别。所提出的ACRNN将EEG信号的空间信息、时间信息和注意力信息都考虑在内。此外, 我们在CNN中集成了通道明智的注意力, 它可以提取空间上的注意力特征, 而通道明智的注意力可以提取通道间的注意力信息。我们还将扩展的自我注意集成到RNN中, 它可以根据每个样本的重要性来提取注意信息。最后, 大量的实验结果表明, 所提出的ACRNN在DEAP数据库的静脉和唤醒分类任务中分别达到了93.72%和93.38%的平均精度。此外, 在DREAM-ER数据库中, 所提出的ACRNN在情感、唤醒和支配性分类任务上的平均准确率分别为97.93%、97.78%和98.23%。与现有的方法相比, 很明显, 提出的ACRNN提高了DEAP和DREAMER数据库中EEG情绪识别的准确性。在未来的工作中, 我们将研究基于注意力机制的基于试验和主体间的EEG情绪识别。

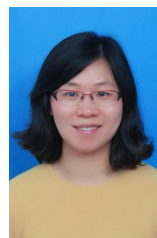
参考文献

- [1] R.W. Picard, *Affective Computing*, 1997.
- [2] S.Koelstra, C. Muhl, M. Soleymani, J. -S.Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp.18-31, 2011.
- [3] S.Katsigiannis and N. Ramzan, "Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp.98-107, 2017.
- [4] J.Cheng, M. Chen, C. Li, Y. Liu, R. Song, A. Liu, and X. Chen, "Emotion recognition from multi-channel eeg via deep forest," *IEEE Journal of Biomedical and Health Informatics*, 2020.

- [5] Y.Liu, Y. Ding, C. Li, J. Cheng, R. Song, F. Wan, and X. Chen, "Multi-channel eeg-based emotion recognition via a multi-level features guided capsule network," *Computers in Biology and Medicine*, vol. 123, p. 103927, 2020.
- [6] M.Ali, A. H. Mosa, F. Al Machot, and K. Kyamakya, "Eeg-based emotion recognition approach for e-healthcare applications," in *2016 eighth international conference on ubiquitous and future networks (ICUFN)*. IEEE, 2016年, 第946-950页。
- [7] Y.Yang, Q. Wu, M.Qiu, Y. Wang, and X. Chen, "Emotion recognition from multi-channel eeg through parallel convolutional recurrent neural network," in *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2018, pp.1-7。
- [8] S.M. Alarco和M. J. Fonseca, "Emotions recognition using eeg signals:A survey," *IEEE Transactions on Affective Computing*, vol. PP, no. 99, pp.1-1, 2017.
- [9] R.Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from eeg," *IEEE Transactions on Affective Computing*, vol. 5, no.3, pp. 327-339, 2014.
- [10] P. Ackermann, C. Kohlschein, J. Á. Bitsch, K. Wehrle, 和 S.Jeschke, "基于Eeg的自动情绪识别。特征提取、选择和分类方法," "在2016年IEEE第十八届电子健康网络、应用和服务国际会议 (Healthcom) 上。IEEE, 2016, pp.1-6.
- [11] N.Sulthan, N. Mohan, K. A. Khan, S. Sofiya, and M. S. PP, "Emotion recognition using brain signals," in *2018 International Conference on Intelligent Circuits and Systems (ICICS)*.IEEE, 2018, 第315-319页。
- [12] C.Li, W. Tao, J. Cheng, Y. Liu, and X. Chen, "Robust multichannel eeg compressed sensing in the presence of mixed noise," *IEEE Sensors Journal*, vol. 19, no. 22, pp.10 574-10 583, 2019.
- [13] M.Li和B.-L. Lu, "Emotion classification based on gamma-band eeg," in *2009 Annual International Conference of the IEEE Engineering in medicine and biology society*.IEEE, 2009, pp.1223-1226.
- [14] X.Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, "Emotion recognition from multi-channel eeg data through convolutional recurrent neural network," in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*.IEEE, 2016, 第352-359页。
- [15] S.Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on eeg using lstm recurrent neural network," *Emotion*, vol. 8, no. 10, pp.
- [16] A.Patil, C. Deshmukh, and A. Panat, "Feature extraction of eeg for emotion recognition using hjorth features and higher order crossings," in *2016 Conference on Advances in Signal Processing (CASP)*.IEEE, 2016年, 第429-434页。
- [17] L.-C.Shi, Y.-Y.Jiao, and B.-L. Lu, "Differential entropy feature for eeg-based vigilance estimation," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.IEEE, 2013, pp.6627-6630.
- [18] R.-N.Duan, J.-Y.Zhu, and B.-L. Lu, "Differential entropy feature for eeg-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*.IEEE, 2013, pp.81-84.
- [19] M.Jaderberg, K. Simonyan, A. Zisserman *et al.*, "Spatial transformer networks," in *Advances in neural information processing systems*, 2015, pp.2017-2025.
- [20] D.Hu, "An introductory survey on attention mechanisms in nlp problems," in *Proceedings of SAI Intelligent Systems Conference*.Springer, 2019, pp.432-448.
- [21] A.Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of neural engineering*, vol. 16, no.3, p. 031001, 2019.
- [22] J.Li, Z. Struzik, L. Zhang, and A. Cichocki, "Feature learning from incomplete eeg with denoising autoencoder," *Neurocomputing*, vol. 165, pp. 23-31, 2015.
- [23] S.K. Goh, H. A. Abbass, K. C. Tan, A. Al-Mamun, N. Thakor, A.Bezarianos, and J. Li, "Spatio-spectral representation learning for electroencephalographic gait-pattern classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 9, pp.1858-1867, 2018.
- [24] Y.Yang, Q. Wu, Y. Fu, and X. Chen, "Continuous convolutional neural network with 3d input for eeg-based emotion recognition," in *International Conference on Neural Information Processing*.Springer, 2018, pp.433-443.
- [25] T.Song, W. Zheng, S. Peng, and C. Zhen, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE*

- [26] D.Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, and R. Boots, "Eeg-based intention recognition from spatio-temporal representations via cascade and parallel convolutional recurrent neural networks," *arXiv preprint arXiv: 1708.06578*, 2017.
- [27] S.Hochreiter和J. Schmidhuber, "长短期记忆". *神经计算*, 第9卷, 第8期, 第1735-1780页, 1997年。
- [28] L.Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "Sca-cnn: Spatial and channel-wise attention in convolutional net- works for image captioning," in *Proceedings of IEEE conference computer vision and pattern recognition*, 2017, pp.Chua, "Sca-cnn: Spatial and channel-wise attention in convolutional net- works for image captioning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp.5659-5667.
- [29] T.Shen, T. Zhou, G. Long, J. Jiang, S. Pan, and C. Zhang, "Dis- an 。 Directional self-attention network for rnn/cnn-free language understanding," in *Thirty-Second AAAI Conference on Artificial In- telligence*, 2018.
- [30] M. S. Özerdem和H. Polat, "基于EEG的情绪识别 feature in movie clips with channel selection," *Brain informatics*, vol. 4, no.4, p. 241, 2017.
- [31] L.Tong, J. Zhao, and W. Fu, "Emotion recognition and channel selection based on eeg signal," in *2018 11th International Conference on Intelligent Computation Technology and Automation (ICICTA)* .IEEE, 2018, pp.101-105.
- [32] A.Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł.Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp.5998- 6008.
- [33] X.Chen, X. Xu, A. Liu, S. Lee, X. Chena, X. Zhang, M. J. McKeown, and Z. J. Wang, "Removal of muscle artifacts from the eeg: a review and recommendations," *IEEE Sensors Journal*, 2019.
- [34] X.-W.Wang, D. Nie, and B.-L. Lu, "Emotional state classification from eeg data using machine learning approach," *Neurocomputing*, vol.129, pp.94-106, 2014.
- [35] L.Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no.11, pp.1254-1259, 1998.
- [36] R.A. Rensink, "The dynamic representation of scenes," *Visual cognition*, vol. 7, no. 1-3, pp.17-42, 2000.
- [37] H.Larochelle和G.E.Hinton, "用三阶波尔茨曼机学习结合眼窝瞥见的东西", 《神经信息处理系统进展》, 2010年, 第1243-1251页。
- [38] M.D. Zeiler和R. Fergus, "可视化理解卷积网络", 在*欧洲计算机视觉会议*上。Springer, 2014, pp.818-833.
- [39] P.H. Seo, Z. Lin, S. Cohen, X. Shen, and B. Han, "Hierarchical attention networks," *arXiv preprint arXiv:1606.02393*, Vol. 2, 2016.
- [40] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp.7132-7141.
- [41] S.Woo, J. Park, J.-Y.Lee, and I. So Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)* , 2018, pp.
- [42] M.Daniluk, T. Mischel, J. Welbl, and S. Riedel, "Frustratingly short attention spans in neural language modeling," *arXiv preprint arXiv: 1702.04521*, 2017.
- [43] D.P. Kingma和J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [44] Y.Li, J. Huang, H. Zhou, and N. Zhong, "Human emotion recog- nition with electroencephalographic multidimensional features by hybrid deep neural networks," *Applied Sciences*, vol. 7, no. 10, p. 1060, 2017.
- [45] D.-A.Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *arXiv preprint arXiv:1511.07289*, 2015.
- [46] S.Goldstein, Z. Hu, and M. Ding, "Decoding working mem- ory load from eeg with lstm networks," *arXiv preprint arX- iv:1910.05621*, 2019.
- [47] T.Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," *IEEE transac- tions on cybernetics*, Vol. 49, no.3, pp. 839-847, 2018.
- [48] D.Zhang, L. Yao, K. Chen, and J. Monaghan, "A convolutional recurrent attention model for subject-independent eeg signal anal- ysis," *IEEE Signal Processing Letters*, vol. 26, no.5, pp. 715-719, 2019.
- [49] N.A. Badcock, P. Mousikou, Y. Mahajan, P. De Lissa, J. Thie, and G. McArthur, "Validation of the emotiv eeg gaming system for measuring research quality auditory erps," *PeerJ*, vol. 1, p. e38, 2013.

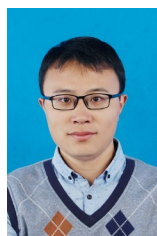
- [50] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, Vol. 7, no.3, pp.162-175, 2015.
- [51] W.-L. Zheng, Y.-Q.Zhang, J.-Y.Zhu, and B.-L. Lu, "Transfer components between subjects for eeg-based emotion recognition," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* .IEEE, 2015年, 第917-922页。
- [52] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215-223, 1979.
- [53] Z.-M.Wang, S.-Y.Hu, and H. Song, "Channel selection method for eeg emotion recognition using normalized mutual information," *IEEE Access*, vol. 7, pp. 143 303-143 311, 2019.
- [54] X.Liu, T. Li, C. Tang, T. Xu, P. Chen, A. Bezerianos, and H. Wang, "Emotion recognition and dynamic functional connectivity analysis- sis based on eeg," *IEEE Access*, vol. 7, pp. 143 293-143 302, 2019.
- [55] P.Li, H. Liu, Y. Si, C. Li, F. Li, X. Zhu, X. Huang, Y. Zeng, D.Yao, and Y. a. Zhang, "Eeg based emotion recognition by combining functional connectivity network and local activations," *IEEE Transactions on Biomedical Engineering*, pp.1-1, 2019.
- [56] H.Wang, X. wu, and L. Yao, "Identifying cortical brain directed connectivity networks from high-density eeg for emotion recogni- tion," *IEEE Transactions on Affective Computing*, vol. PP, pp.1-1, 07 2020.



程娟分别于2008年和2013年在中国科学技术大学电子科学与技术系获得学士和博士学位。

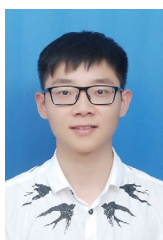
她目前是合肥工业大学生物医学工程系的副教授，中国合肥。她的研究兴趣包括生物医学信号/图像处理、非接触生理参数测量和机器测量。

学习。



刘学分别于2011年和2016年在中国科学技术大学自动化系获得学士和博士学位。

他目前是合肥工业大学生物医学工程系的助理教授。他的研究兴趣包括图像/信号处理、计算机视觉、信息融合和机器学习。



魏涛于2017年获得中国合肥工业大学测量、控制技术与仪器专业的学士学位，并于2020年获得中国合肥工业大学生物医学仪器专业的工程硕士学位。他目前正在澳门大学科技学院电子与计算机工程系攻读博士学位。他的研究兴趣包括基于脑电的情绪识别、压缩传感和机器

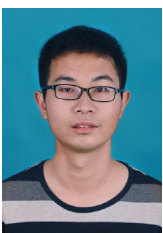
学习。



冯万在香港科技大学获得电气和电子工程博士学位。

他目前是科学与技术学院电子与计算机工程系的副教授，同时也是澳门大学合作创新学院认知与脑科学中心的初级教师。他的研究兴趣包括脑机接口。

人脸、生物医学信号处理、神经影像和神经反馈训练、深度学习和迁移学习、计算智能和指示性控制。



李昌于2012年在中国武汉理工大学获得信息与计算科学学士学位，并于2018年在武汉华中科技大学电子信息与通信学院获得电路与系统的博士学位。

他目前是合肥工业大学生物医学工程系的讲师，合肥。他的研究兴趣包括生物医学信号处理。

高光谱图像分析、计算机视觉、模式识别和机器学习。



陈勋于2009年在中国科学技术大学电子科学与技术系获得学士学位，并于2014年在不列颠哥伦比亚大学（UBC）电子与计算机工程系获得博士学位。

他在中国科技大学电子科学与技术系担任教授。他的研究兴趣包括统计信号处理和机器学习等广泛领域。

在生物医学方面的应用。他已经发表了100多篇有参考价值的科学论文。他担任《信号处理-图像通信》、《IEEE信号处理通讯》、《神经科学前沿》和《IEEE Access》的副主编。



宋任成于2005年在中国长春的吉林大学获得数学学士学位，并于2010年在中国杭州的浙江大学获得博士学位。

他目前是合肥工业大学生物医学工程系的一名副教授。他的研究兴趣包括生物医学信号处理、非接触式生理参数测量、机器学习和电磁成像。