

基于 RCNN-LSTM 的脑电情感识别研究

柳长源¹ 李文强¹ 毕晓君²

摘要 情感作为人脑的高级功能,对人们的个性特征和心理健康有很大的影响,利用网上公开的脑电情感数据库(DEAP (Database for emotion analysis using physiological signals) 数据库),根据心理效价和激励唤醒度等级进行情感划分,对压力和平静等 5 种情感进行研究分析.针对脑电信号时空特征结合的特点,把深度学习中的卷积神经网络(Convolutional neural network, CNN)和长短期记忆网络(Long short term memory, LSTM)两者作为基本前提,并在此基础上设计了一个 RCNN-LSTM 的脑电情感信号分类模型.利用循环卷积神经网络(Recurrent convolutional neural network, RCNN)自动提取脑电信号中的抽象特征,省去了人工选择与降维的过程,然后结合 LSTM 网络对脑电情感信号进行分类识别.实验结果表明,利用该方法对 5 种情感类别的平均分类识别率达到了 96.63%,证明了该方法的有效性.

关键词 脑电信号,情感识别,循环卷积神经网络,长短期记忆神经网络

引用格式 柳长源,李文强,毕晓君.基于 RCNN-LSTM 的脑电情感识别研究.自动化学报,2022,48(3):917-925

DOI 10.16383/j.aas.c190357

Research on EEG Emotion Recognition Based on RCNN-LSTM

LIU Chang-Yuan¹ LI Wen-Qiang¹ BI Xiao-Jun²

Abstract As an advanced function of the human brain, emotion has a great influence on people's personality characteristics and mental health. Using the online public EEG (electroencephalogram) database (DEAP (database for emotion analysis using physiological signals) database), emotions are divided according to psychological valence and arousal level. Five emotions, including stress and calm, are studied and analyzed. According to the combination of temporal and spatial features of EEG signals, CNN (convolutional neural network) and LSTM (long short term memory) in deep learning are taken as the basic prerequisites. On the basis of this, a new and more powerful EEG signal classification model is constructed. Use the recurrent convolutional neural network (RCNN) to automatically extract the abstract features of the EEG signal, eliminate the process of manual selection and dimensional reduction, and then send it to the LSTM network to classify and identify EEG signals. The experimental results show that the average accuracy of this method is 96.63%, which demonstrates the effectiveness of this method.

Key words Electroencephalogram (EEG), emotion recognition, recurrent convolutional neural network (RCNN), long short term memory (LSTM)

Citation Liu Chang-Yuan, Li Wen-Qiang, Bi Xiao-Jun. Research on EEG emotion recognition based on RCNN-LSTM. *Acta Automatica Sinica*, 2022, 48(3): 917-925

情感在人们的理性行为和日常生活中起着重要作用,与我们息息相关.通过情感的识别,我们可以去看一些患有自闭症和抑郁症的儿童,为他们提供心理咨询与辅导,这不仅能够更好地促进人们之间的交流,在人机交互的应用中也具有重要意义,这是一个涉及现代医学心理学领域、计算机领域和人

工智能等众多领域内的重要研究课题^[1].

目前,对于情感分析的分类方法已有一定的规模,一般我们在研究过程中经常使用的方式就是在研究生理特点的基础之上对人的各种情感差异进行对比分析,例如通过人的面部表情、行为动作和语音语调等外部生理特征来进行情感分析,但是这种研究方法易导致情感识别的精度有所下降^[2].因为人的面部表情、声音等又容易受到自己的主观控制,与这些外在表现出来的生理信号特征相比,脑电信号存在于人们的中枢神经系统中,能够体现出不同时刻的差异,因此,与情感的关联性远远超过其他信号.基于脑电信号的情感识别有着非常广阔的应用前景和研究价值,越来越多的科研人员投入到了这项工作中.Krisnandhka 等^[3]提取了脑电信号的小波能量,并计算出其相对利用率作为最后的脑电

收稿日期 2019-05-12 录用日期 2020-03-25
Manuscript received May 12, 2019; accepted March 25, 2020
国家自然科学基金(5177090001)资助
Supported by National Natural Science Foundation of China (5177090001)
本文责任编辑 张军平
Recommended by Associate Editor ZHANG Jun-Ping
1. 哈尔滨理工大学测控技术与通信工程学院 哈尔滨 150080 2. 中央民族大学信息工程学院 北京 100081
1. College of Measurement and Control Technology and Communication Engineering, Harbin University of Science and Technology, Harbin 150080 2. School of Information Engineering, Minzu University of China, Beijing 100081

特征进行情感识别, 达到了 76% 的情感识别率. Duan 等^[4]对常见的正向情感和负向情感做了二分类, 提出采用脑电信号的差分熵作为脑电特征, 准确率达到 83.28%. Murugappan 等^[5]利用三种不同的“db8”, “sym8”和“coif5”小波基函数来提取脑电信号的统计特征, 对常见的快乐、惊讶、恐惧、厌恶和中性情绪进行分类识别, 最大平均准确率达到 83.04%. 黄柠檬^[6]基于共同空间模式(Common spatial pattern, CSP)算法对脑电信号中的多种特征进行选择与融合, 分类识别率为 80.5%. 杨默涵等^[7]利用总体经验模态分解(Ensemble empirical mode decomposition, EEMD)和希尔伯特变换提取了边际谱和瞬时能量作为脑电特征, 利用线性判别分类器(Linear discriminant analysis, LDA)对脑电信号进行分类, 平均识别率为 82.74%. Zhang 等^[8]提出了利用变分模态分解(Variational mode decomposition, VMD)和自回归(Auto regressive, AR)相结合的方法进行脑电特征提取, 并采用随机森林分类器处理三分类任务, 取得了不错的效果. 另外, 也有研究者通过非线性动力学特征对脑电情感信号进行了分析, 孙颖等^[9]利用功率谱熵和非线性全局特征相结合构造出脑电情感特征, 采用支持向量机对脑电情感进行四分类, 平均识别率达到了 86.42%. 但是, 上述文献中的这些脑电特征都是依据经验进行设计手动提取的特征, 所得到的脑电特征的质量得不到保证且容易出现特征丢失. 利用传统的分类方法支持向量机, 随机森林和 LDA 等进行情感识别时, 这些算法的分类结果很大程度上依赖于特征质量, 且泛化性不强. 所以, 构造出一个能够自动提取情感特征和进行分类的模型是提高脑电情感分类识别率的关键.

深度学习(Deep learning, DL)作为机器学习领域的一个重要分支, 其目标是创造一个模仿人脑学习的网络. 与传统机器学习方法相比, 其省去了手动提取特征的环节, 通过半监督或者无监督的方式来满足特征自动提取的需要, 为数据的高级别抽象表示提供了可能^[10]. 如今, 深度学习在语音识别、图像识别等领域中都得到一定程度的运用, 并且为这些领域的发展做出了一定的贡献. 由于深度学习具有自动学习提取样本抽象特征的能力, 避免了特征选择与降维的过程, 在情感识别方面, Jirayuchareonsak 等^[11]应用堆叠自编码器(Stack autoencoder, SAE)方法建立情感识别模型, 除此之外, 并与一些我们传统使用的分类计算方式支持向量机(Support vector machine, SVM)、朴素贝叶斯(Naive Bayesian, NB)等进行对比分析, 以此验证了深度网络的优越性. 杨豪等^[12]采用深度信念网络(Deep

belief network, DBN)来识别三种脑电情感状态, 平均准确率达到 89.12%, 但是 DBN 训练所需的时间较长, 且训练过程中易产生冗余特性. Cheng 等^[13]利用深度卷积神经网络对脑电情感进行分类识别, 但是随着情感分类类别数的增加准确率出现下降, 此时体现出单一卷积神经网络(Convolutional neural network, CNN)模型的局限性. 此后, Wang 等^[14]引入了 3D-CNN 对脑电信号进行情感识别, 唤醒度和效价的分类准确率分别为 73.3% 和 72.1%, 但是识别脑电信号时所需的时间较长, 因为在识别前, 需要将一维的脑电信号转成三维数据形式. 阚威等^[15]利用 LSTM 网络对脑电情感信号进行二分类, 愉悦度分类的准确率为 73.5%, 唤醒度的分类准确率为 73.87%. 把脑电时序信号的各个时间点直接输入长短期记忆网络(Long short term memory, LSTM)中, 并没有达到理想的分类结果, 主要原因可能是将脑电信号的时序信息直接输入 LSTM 网络并不能获得与情感相关的空间频率信息特征. 在脑电情感信号处理时, 前后的关联和表征能力仍有不足之处. 所以, 有必要在 LSTM 的前层嫁接一个合适的用于提取脑电情感特征的神经网络结构.

因此, 本文采用深度学习中的卷积神经网络 CNN 和长短期记忆网络 LSTM 为基础, 设计了 RCNN-LSTM 混合深度学习模型来提高脑电情感信号的识别率, 首先, 改进传统的 CNN 模型, 利用循环卷积神经网络(Recurrent convolutional neural network, RCNN)来提取脑电信号特征, 以充分挖掘脑电信号的内在情感信息, 再结合 LSTM 对时序特征信息的融合与建模, 使其综合考虑了脑电信号的时间和空间特性, 更全面地保留了脑电信号的抽象和深层特征, 通过对最后实验结果的评估, 证明了 RCNN-LSTM 模型的有效性.

1 卷积神经网络与循环神经网络

1.1 卷积神经网络

卷积神经网络主要含有两大重要特性: 1) 局部连接特性; 2) 权值共享特性. 所谓的局部连接特性的优势在于可以在一定程度上降低拟合过程中所遇到的风险; 权值共享性特征的优势在于可以降低训练参数的数量, 使网络更易于优化^[16].

卷积层通过前向操作可使不同的卷积核与输入数据来进行运算, 输出的就是输入数据所产生的不同特征. 卷积计算公式为

$$Y_n = \sum_{i=1}^M [(W_n^i * x_i) + b_n] \quad (1)$$

其中, x_i 表示输入数据, W_n^i 表示第 i 通道上的第 n 个卷积核, b_n 是偏置值, $*$ 表示卷积运算, Y_n 是计算得到的第 n 个特征图.

池化层也叫降采样层. 为了把权值控制在一定范围之内, 降低网络拟合中产生的风险, 卷积神经网络将池化层引用其中, 以此来对特征图来进行降采样处理, 保障数据的精确度和可靠性.

1.2 循环神经网络

循环神经网络 (Recurrent neural network, RNN) 在神经网络的领域里已有一定历史^[17], 与卷积神经网络不同, RNN 具有环状结构, 可以使信息得以维持. RNN 每次处理一个时序输入, 更新一种向量状态, 该向量含有序列中所有过去的元素. 图 1 展示了一个输入为 X_t , 输出为 O_t 的神经网络.

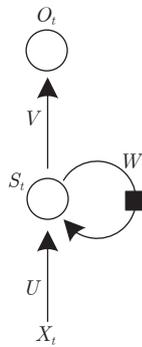


图 1 一个含有内部环的 RNN
Fig.1 An RNN with an inner ring

S_t 是网络中的一个向量状态, 可以视为系统的一种记忆模式. 它含有输入序列中所有之前的元素中的信息. 另一方面, 图中的循环使得信息可以由网络中的每一步移动到下一步. 状态向量 S_t 由当前输入和之前的状态向量通过矩阵 U 和 W 计算.

$$S_t = f(U \times X_t + W \times S_{t-1}) \quad (2)$$

式中, f 是一个非线性函数, 与 \tanh 和 ReLU 类似. 可以看到, 函数中的两项要先求和, 然后由 f 处理. 最后, O_t 是网络输出, 由矩阵 V 计算得来.

$$O_t = V \times S_t \quad (3)$$

RNN 在时间结构上是存在循环的一个状态, 这就意味着可以保证数据信息的长时间保存, 可以使神经网络对时间序列的前后关联问题进行处理. 但是梯度的消失-爆炸问题也会影响 RNN, RNN 网络的时序不能太长就是出于这个原因. 如果梯度在数层之后开始消失或爆炸, 网络就不能学习数据之间比较高的时间距离关系.

2 RCNN-LSTM-EEG 结构

2.1 循环卷积神经网络

对于 RCNN 循环卷积神经网络, 其结合了循环神经网络的特性^[18], 对传统的 CNN 中层与层之间自上而下的连接方式进行了改进, 通过在同一层上递归连接多个迭代的卷积子得到了循环卷积层 (Recurrent convolutional layer, RCL), 使 CNN 每一层前馈网络上进行了反复的循环计算, 而相邻的 RCL 层之间只有前馈连接. RCNN 结构主要包含一个标准的卷积层、数层循环卷积层以及最终输出层. 图 2 是整个 RCNN 构成的结构分布图.

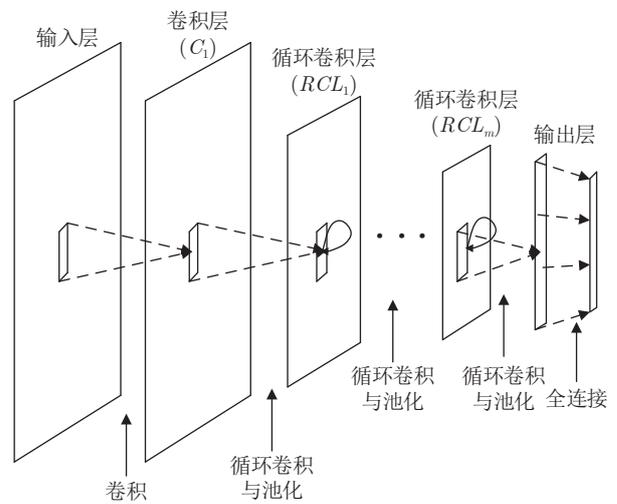


图 2 RCNN 结构分布图
Fig.2 RCNN structure distribution

第 1 层 (C_1) 是标准的前馈卷积层, 没有递归连接. 第 1 层 C_1 之后跟随着多个循环卷积层 ($RCL_1 \sim RCL_m$), 并且每个循环卷积层都是由众多的卷积子组建而成, 每两个循环卷积层中间有一个 maxpooling 层来进行降维, 并且在 $T + 1$ 个时刻 RCL 层中的多个卷积子都可以对权值进行共享. 假如把一个 RCL 层进行展开, 可以把其作为有 $T + 1$ 深度的前馈子网, 图 3 就是 RCL 层的具体展开图. 但是我们需要明确的是一个 RCL 层和具有 $T+1$ 深度的 CNN 是有一定差别的, 主要体现在: RCL 层主要是通过卷积来进行具体计算的, 它的输入是卷积运算在不同时间点上的数值, 而 $T + 1$ 深度的 CNN 的输入的时间点是一致的. 因此, 在参数量固定的状态下, 可以把 RCL 取任意时间点进行展开, 从而得到相应的神经网络.

确定分类效果的一个非常重要的因素就是模型所采用的深度, 层次越深, 越可以让网络学习到更

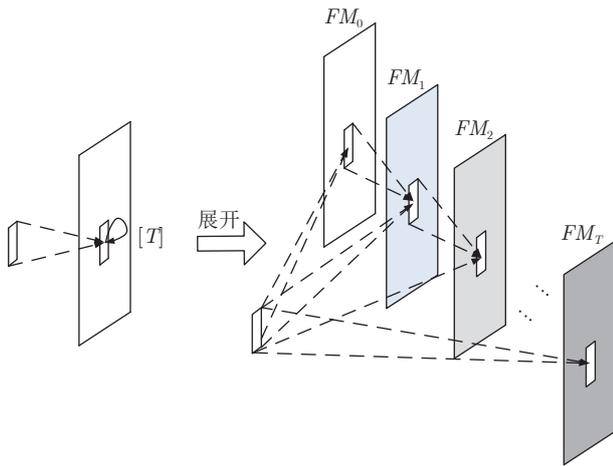


图 3 RCL 展开图

Fig.3 RCL expansion diagram

加复杂的特征. 相反, 网络深度越浅, 那么越有利于梯度回传. RCNN 能够更加高效合理地权衡这两点, 它可以看作是一个具有弹性路径的 CNN. 不仅可以对网络深度进行一定程度的加深, 而且还能够有效地对它进行训练. 也就是说 RCL 的这种迭代的计算, 实现了路径长度可以进行延展, 从 1 层可以延展到 $T + 1$ 层. 因此, 若在一个完整的 RCNN 网络中 RCL 层为 4 个, 那么就可以计算出总路径可以从 6 转变为 $4(T + 1) + 2$ (包括第 1 层的卷积层和最后的输出层). 正是因为 RCL 此特性, 从而使网络深度得到了更好的延展, 并且能够大大地提高权值共享效果, 使得参数量大幅度减少. 本文在设计卷积核时避免使用较大尺寸的核, 主要是由于脑电情感信号的时间序列不同于图像中的特征点很密集. 也就是说在一个脑电信号的区域内会存在很多无关的特征, 卷积核尺寸越大, 其感受野就越大, 就会提取到很多无用的脑电特征. 同时卷积核尺寸增大, 也会增加计算量的开销, 不利于模型深度的增加, 计算性能会大大降低. 一般情况下在更深的模型中用较小的卷积核能够实现正则化的效果. 在多次实验后得到了各层具体参数如表 1 所示.

2.2 长短期记忆神经网络

作为时间递归神经网络的一种形式, LSTM 网络不但能够体现出脑电信号信息的前后关联性, 并且更能够处理长时间依赖的问题. 在许多时间序列的分类问题上, 如语音分类、自然语言处理 (Natural language processing, NLP) 等, LSTM 网络都获得了很大的进步. 由于脑电情感信号的时间特征和频谱特性与上述的时序信息类似, 考虑将 LSTM 神

表 1 RCNN 网络模型结构

Table 1 RCNN network model structure

| 网络层 | 参数设置 |
|-------|-----------------------|
| 输入层 | 输入结构: (数据长度, 1) |
| 标准卷积层 | 核数量: 90; 大小: 3; 步长: 1 |
| 最大池化层 | 核大小: 4 步长: 2 |
| RCL层 | 核数量: 90; 大小: 3; 步长: 1 |
| 最大池化层 | 核大小: 4 步长: 2 |
| RCL层 | 核数量: 90; 大小: 3; 步长: 1 |
| 最大池化层 | 核大小: 4 步长: 2 |
| RCL层 | 核数量: 90; 大小: 3; 步长: 1 |
| 最大池化层 | 核大小: 4 步长: 2 |
| RCL层 | 核数量: 90; 大小: 3; 步长: 1 |
| 最大池化层 | 核大小: 4 步长: 2 |

经网络应用于脑电情感信号分类. LSTM 具有 RNN 的优势, 能够比较精准地对序列数据进行准确的分析^[19]. 为了保证整个循环网络能够更加稳定地进行, LSTM 模型增添了其他机制, 以此来降低梯度消失和发生爆炸的可能性. 记忆单元 c 、输入门 i 、遗忘门 f 及输出门 o 这些机制相互组合运用, 可以大幅度地提高神经网络的数据处理能力. LSTM 模型具体运算为

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (4)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (5)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (6)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (7)$$

$$h_t = o_t \times \tanh(c_t) \quad (8)$$

在上述函数式中, i_t 表示输入门, f_t 表示遗忘门, c_t 表示记忆单元, o_t 表示输出门向量值, W 表示权值响应数据, b 是偏置数值, \tanh 是双曲正切函数. 其中输入门主要的作用就是对记忆单元 c 的强度进行控制; 输入门、遗忘门以及输出门三者之间各自有着其相应的用途, 当有新的信息输入到 LSTM 网络当中时, 可以根据输入门的规则来进行选择性的添加, 并通过遗忘门对无用的信息进行选择性过滤遗忘. 它们之间的计算方式存在相似之处, 但是又有着不同的参数, LSTM 结构如图 4 所示.

在实验过程中, 主要利用的是单层 LSTM 结构, 选择 \tanh 函数作为模型的激活函数, 设置隐层的节点数为 30 个, 输出层采用使用 Softmax 函数来对情感进行分类识别. LSTM 在训练和识别的过程中通过输入当前信息和前一时间隐层的状态以此来改变隐层的状态, 并且对这一步骤进行不断地循环, 一直到处理结束.

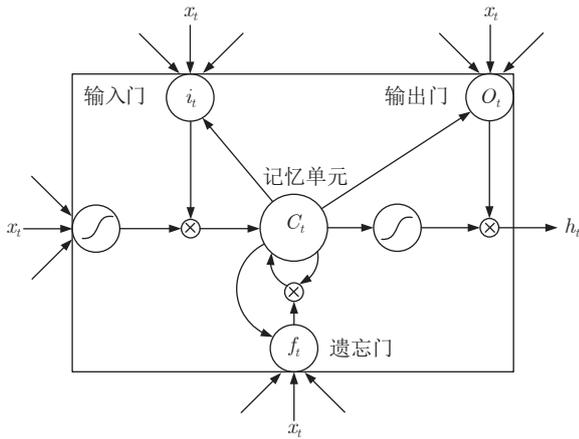


图 4 LSTM 结构示意图

Fig.4 LSTM structure diagram

2.3 RCNN-LSTM 结构

脑电信号在经过 RCNN 网络之后, 我们可以得出时序信号的具体特征及其所承载信息的特点. 因为在这个过程中我们所得到的特征也是一个相应的时间序列, 所以, 需要借助 LSTM 建立模型来对这些时序信息进行详细全面的分析, 图 5 为 RCNN-LSTM 模型整体结构示意图.

在图 5 所示结构中, EEG_{input} 为送入网络训练或测试的单条脑电信号长度, 样本长度的设置非常重要, 它和网络结构、分类数目、训练速度都紧密相关. 考虑到实验者观看视频时需要一定的反应时间, 并且在最后阶段难免会出现倦怠感, 因此把实验中前 23 s 和后 20 s 的数据去除, 只保留中间 20 s 的数据, 共 2560 个时间节点. ee_{len} 为对长度 EEG_{input} 的脑电信号进行再切分的单段长度, 也是 RCNN 网络进行特征提取的脑电信号长度. 在各个切分的脑电信号段数互不重叠的情况下, 可切分的段数 $step_n$ 可由式 (9) 计算得到

$$step_n = \frac{EEG_{input}}{ee_{len}} \quad (9)$$

$step_n$ 就是脑电信号切分的段数, 也是获取到脑电特征的总步数, 是输入到 LSTM 的总步数. 当依次将长度为 ee_{len} 的脑电信号送入 RCNN 后, 会分别提取到各段脑电信号的特征. 将得到的多时段脑电特征送入 LSTM 后将获得最后一步的隐层输出 h_{step_n} . 将隐层数据与全连接层相连, 图 5 中所示的 $N_{classes}$ 为分类类别, 将全连接层的输出再通过 Softmax 计算可获得预测输出的情感类别. 本文将 ee_{len} 分别定为 64, 128, 160 和 256, 将所有样本数据分为 10 折, 在其中的 9 折数据内部进行了交叉

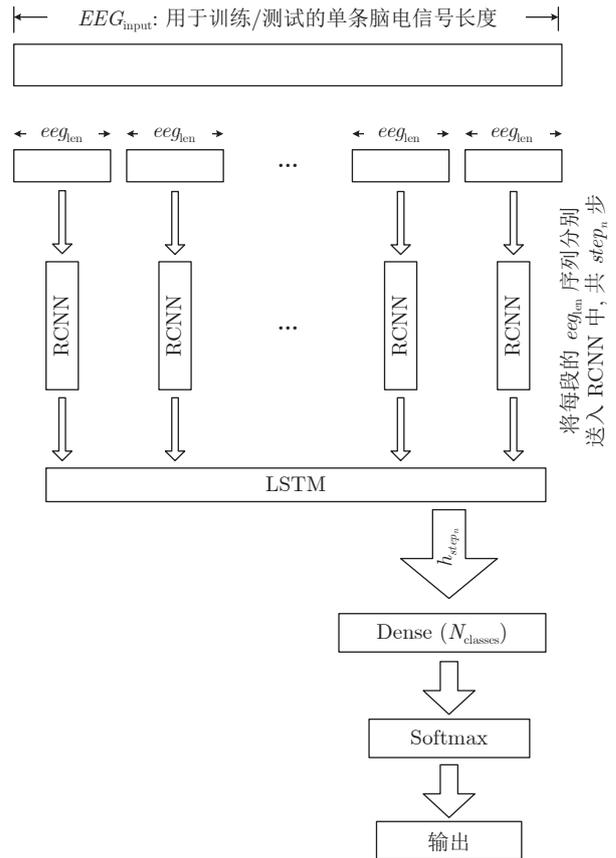


图 5 RCNN-LSTM 结构示意图

Fig.5 RCNN-LSTM structure diagram

验证实验, 尝试不同的 ee_{len} 和 $step_n$, 以此来寻找最优的 ee_{len} 和 $step_n$ 组合, 将剩下的 1 折数据作为最终的测试数据. 通过比较 9 折实验后脑电情感信号的平均识别率可以看到, 当 ee_{len} 为 160, $step_n$ 为 16 时, 能够获得较好的分类效果.

3 实验结果分析

3.1 实验数据

DEAP (Database for emotion analysis using physiological signals)^[20] 数据库中主要采集了 32 名身体健康参与人员的脑电数据, 在这些样本人员之中, 男女性别之间的比例是 1:1, 参与实验人员的平均年龄为 26 岁, 并且在实验之前把整个实验的细节和过程详细地告知了受试人员. 这些受试人员的身体和心理都是呈现良好的健康状态, 在收集样本数据的过程中主要采用 32 导 AgCl 电极, 此外, 还有 8 个电极采集外周生理信号. 每个实验时长为 63 s, 前 3 s 为实验准备阶段, 随后的 60 s 是被试者观看的 MV 时间. 每位受试者每次实验信号采集过程如表 2 所示.

表 2 DEAP 数据收集统计表
Table 2 DEAP data collection statistics

| 顺序 | 内容 |
|----|---|
| 1 | 把电极放在适当的实验位置, 通过铃声来通知开始实验 |
| 2 | 接着调整基线约 2 min, 直到符合要求 (主要目的就是为了使受试人员能够保持放松) |
| 3 | 显示 2 s 的实验编号来提示具体的进程 |
| 4 | 收集 3 s 的基线数据信息 |
| 5 | 播放 1 min 的音乐, 并且详细地记录这个时间段的数据信息 |
| 6 | 对收集到的数据信息进行统计整理, 并且对各个相应的指标进行全面评估 |

本文在实验过程中所采用的 DEAP 数据库中的数据是经过预处理之后得到的, 经过预处理后的数据采样频率为 128 Hz, 并且去除了眼电等噪声, 从而使得数据的准确性和精确度都能够保持在可靠范围之内. 根据相关的研究成果表明, 大脑的额区和中央区域主要参与人们的情感处理^[21]. 因此, 本文主要提取集中在大脑额区和中央区的脑电通道数据来进行研究, 分别是 FP1、FP2、F3、Fz、F4、F7、F8、C3、C4^[22]. DEAP 数据库主要是运用效价和唤醒度所构成的二维情感模型^[23], 该模型可以把人类的情感划分为两种对应极端, 一端为正极, 另一端为负极. 当情感处于正极意味着是正性情感, 能够为人带来良好的比较愉悦的内心感受; 而情感处于负极意味着是负性情感, 即给人带来不好的情感感受, 例如: 焦虑、暴躁等.

3.2 结果分析

将本文提出的 RCNN-LSTM 模型应用到 DEAP 数据集上, 对常见的 5 类情感: 轻松、沮丧、快乐、压力和平静进行研究分析, 轻松、沮丧、快乐、压力和平静的样本分别为 1 250、1 178、1 153、1 132 和 1 228 个. 此外, 又比较了采用传统的识别算法和浅层神经网络 (SVM、BP、CNN、LSTM) 对 5 种脑电情感信号的分类进行实验, 取 eeg_{len} 为 160, $step_n$ 为 16, 不同算法均对最后的 1 折数据进行 10 次测试实验, 并利用整体情感的平均分类识别率和单个情感分类识别率既召回率 (Recall) 对模型进行分类结果进行评估, 结果如表 3 所示.

分类问题中, 最常见的评价指标是准确率 (Accuracy, ACC), 它能够直接反映出分正确的比例, 同时计算非常简单. 但是实际的分类问题中, 可能各个类别的数据量有一定差异, 导致有时整体的准确率 (ACC) 较高, 但是部分类别的分类效果较差, 这时单纯的准确率 (ACC) 不能完全作为模型的评价标准. 为此, 计算模型分类准确率的方差和 kappa 系数来衡量分类模型的好坏, 结果如表 4 所

表 3 不同算法下的召回率和分类识别率 (%)
Table 3 Classification recognition rate under different algorithms (%)

| | 轻松 | 沮丧 | 快乐 | 压力 | 平静 | 总体识别率 |
|-------------|------|------|------|------|------|-------|
| RCNN + LSTM | 100 | 96.6 | 95.7 | 96.5 | 94.3 | 96.63 |
| CNN + LSTM | 79.2 | 77.1 | 86.1 | 92.0 | 85.4 | 83.83 |
| CNN | 72.8 | 76.3 | 82.6 | 84.9 | 81.0 | 79.46 |
| LSTM | 73.6 | 73.7 | 77.4 | 81.4 | 86.2 | 78.45 |
| SVM | 68.0 | 75.4 | 83.4 | 75.2 | 77.2 | 75.92 |
| BP | 76.0 | 72.8 | 75.6 | 84.1 | 74.7 | 76.59 |

表 4 不同算法下的 kappa 值和方差
Table 4 Kappa values and variances under different algorithms

| | kappa 系数 | 方差 ($\times 10^{-2}$) |
|-------------|----------|-------------------------|
| RCNN + LSTM | 0.928 | 0.04 |
| CNN + LSTM | 0.763 | 0.06 |
| CNN | 0.70 | 0.12 |
| LSTM | 0.718 | 0.16 |
| SVM | 0.658 | 0.27 |
| BP | 0.675 | 0.16 |

示, kappa 用于一致性检验, 具体计算为

$$k = \frac{p_0 - p_e}{1 - p_e} \quad (10)$$

式中, p_0 是每一类正确分类的样本数量之和除以总样本数, 也即总体分类识别率; p_e 是预期与真实结果一致的概率.

在表 3 中, SVM 与 BP 神经网络两种方法是在 MATLAB2015 平台中调用相应的分类函数分析得出的结果, SVM 是一种有监督的学习模型, 可以对脑电情感信号进行分类识别, 使用的核函数为高斯径向基核函数, 其形式为

$$K(x, y) = \exp \left\{ -\frac{\|x - y\|^2}{\sigma^2} \right\} \quad (11)$$

分类决策函数为

$$f(x) = \text{sgn} \left(\sum_{i=1}^m \alpha_i y_i K(x_i \times x) + b^* \right) \quad (12)$$

SVM 的分类性能主要依赖于惩罚因子 c 和径向基函数 (Radial basis function, RBF) 核函数的参数 g , 在许多情况下, 找到合适的参数一般需要人为进行大量的尝试, 这样计算量较大且不易获得最佳参数. 为了解决参数选择的问题, 本文使用遗传算法 (Genetic algorithm, GA) 对两个参数进行优化. 得到的最优的惩罚因子 $c = 1.3055$, 核参数 $g = 0.82359$. GA-SVM 的适应度曲线如图 6 所示.

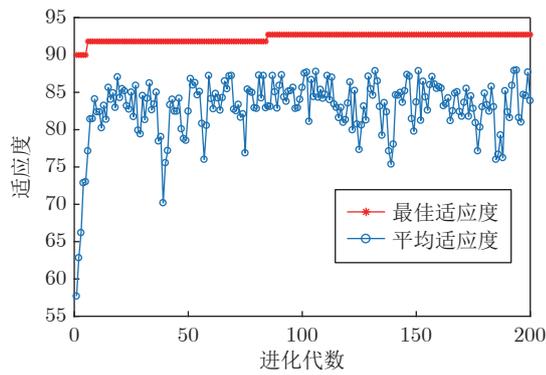


图6 GA-SVM 的适应度曲线

Fig.6 GA-SVM fitness curve

BP 神经网络的参数设置目前并没有明确的规则来确定, 根据以往的经验 and 多次实验, 本文将 BP 神经网络隐藏层设置为两层, 每层的节点数分别为 30、10, 用 $-1 \sim 1$ 的随机数来初始化 BP 神经网络中的权重, 用反向传输算法进行 BP 神经网络训练, 采用带有动量项的梯度下降法进行参数优化. CNN、LSTM 与本文的 RCNN-LSTM 模型是在开源框架 Tensorflow 中进行分类训练和测试. 表 1 中的 RCNN 循环卷积神经网络里有 1 个标准的卷积层和 4 个循环卷积层 RCL, 因此, 表 3 中的 CNN 模型设为 5 层标准卷积层, 每层的卷积核的数量和核大小与 RCNN 中的参数相同, 表 3 中的 LSTM 模型设为单层, 参数与 RCNN-LSTM 模型中的 LSTM 相同, 表 3 中 CNN-LSTM 网络由上述的 CNN 与 LSTM 进行组合, 分别与本文的 RCNN-LSTM 模型进行对比实验.

由表 3 可以看出, 传统的神经网络和机器学习算法对脑电情感分类准确率较为一般, 且单个情感类别的召回率波动较大, 本文构建的 RCNN-LSTM 模型整体平均识别率达到 96.63%, 并且对每种情感的召回率都比较高且比较稳定, 最低的分类精度为 94.3%, 最高的分类准确率达到了 100%, 与传统方法相比, 具有明显的优势. 由表 4 可以看出, 本文提出的 RCNN-LSTM 模型在进行多类的脑电信号分类时的 kappa 系数高于其他 5 种算法, 并且方差最小, 鲁棒性较优. 由于脑电信号具有复杂的时空特性, 传统的 CNN 网络通过不断训练能够学习脑电信号的空间频率信息, 但却忽略了时间域的信息. LSTM 方法则正好相反, LSTM 作为 RNN 的一种特殊形式, 可处理各部分输入的数据并进行统一分类, 但是, 直接将脑电信号以各个采样点输入 LSTM 并未取得良好的效果. 因此, 本文构建了 RCNN-LSTM 模型来提高脑电情感识别率. 在送入 LSTM 前, 利用 RCNN 循环卷积神经网络用于

对 EEG 信号的频域进行分析, 相较于传统的 CNN, RCNN 循环卷积神经网络不但能够学习脑电信号的空间频率信息, 通过循环卷积层联系网络隐层中当前时刻和前几个时刻, 还能获得脑电信号序列的时序关系. 再通过对脑电信号的单步切分 eeg_{len} , 依次将每一步的 eeg_{len} 送入 RCNN 中, 以此来获得更丰富的脑电情感特征, 最后采用 LSTM 对时序特征信息进行融合建模与分类, 更全面地保留了脑电信号的内在情感特征. 因此, 将 RCNN-LSTM 模型用于多类脑电情感信号分类相较于传统的神经网络和机器学习算法具有更好的识别效果.

将本文的研究方法与已有研究成果进行对比, 比较结果如表 5 所示, 文献 [24] 提取了脑电信号的功率谱密度为特征, 利用 SVM 进行二分类, 识别率达到了 70.1%, 文献 [25] 将近似熵、样本熵、排列熵进行组合作为特征, 利用 K-最近邻 (K-nearest neighbor, KNN) 进行分类, 分类精度达到了 77.8%, 文献 [26] 中利用了卷积神经网络和循环神经网络的方式对 DEAP 脑电数据进行了情感分类, 整体识别率为 73.09%. 文献 [27] 将堆栈自编码器与循环神经网络相融合, 得到的情感识别率为 79.26%. 文献 [28] 采用新型动态图卷积神经网络 (Dynamical graph convolutional neural networks, DGCNN) 对脑电情感信号进行分类识别, 平均识别率达到了 90.4%. 文献 [29] 采用 3D-CNN 模型对脑电信号进行情感识别, 在唤醒度和代价上的准确率分别为 88.49% 和 87.44%, 整体平均识别率为 87.965%.

表 5 已有研究成果与本文的对比

Table 5 The existing research results are compared with our method

| 方法 | 分类类别 | 识别率 (%) |
|--------------|------|---------|
| SVM | 2 | 70.1 |
| KNN | 2 | 77.8 |
| CNN + RNN | 2 | 73.09 |
| SAE + RNN | 4 | 79.26 |
| DGCNN | 3 | 90.4 |
| 3D-CNN | 2 | 87.965 |
| 本文 RCNN-LSTM | 5 | 96.63 |

本文通过构建新的 RCNN-LSTM 模型对脑电情感信号进行 5 类识别研究, 整体的平均分类识别率达到了 96.63%, 比上述已有的相似研究成果都有了一定程度的提高, 验证了本文方法的可行性.

4 结束语

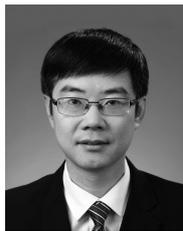
本文在研究的过程中主要使用了 DEAP 数据

库中的脑电数据,对沮丧、轻松、快乐、压力和平静5种情感脑电信号进行研究分析.针对目前基于脑电信号情感分类准确率较低且分类类别少的问题,提出了一种将循环卷积神经网络(RCNN)与长短期记忆神经网络(LSTM)模型相融合的脑电情感识别方法.通过递归连接的RCNN循环卷积神经网络,可以在每个RCL层参数量固定的条件下,增加网络的深度,使模型能够自动学习到高度复杂的脑电情感特征,避免了手动特征选择与降维等步骤,并对原始的脑电信号序列进行单步切分,通过RCNN特征提取后进行多时段特征融合,并借助LSTM对时间序列分类识别的优势,使其对脑电情感信号的时、空、频域特征能有很好的泛化能力和强鲁棒性,平均整体情感分类识别率达到了96.63%,结果表明本文所提的RCNN-LSTM模型可以帮助人们有效地识别脑电情感信号,为能够得到性能完善可靠的情感识别脑-机接口提供了一定的参考.

References

- Geng Xue-Qing, She Qing-Shan, Han Xiao, Meng Ming. Classification of motor imagery EEG based on Gaussian process optimized with artificial bee colony. *Chinese Journal of Sensors and Actuators*, 2017, **30**(03): 378-384
(耿雪青, 余青山, 韩笑, 孟明. 基于人工蜂群优化高斯过程的运动想象脑电信号分类. 传感技术学报, 2017, **30**(03): 378-384)
- Wang Wei-Rong, Zhang Xue-Ying, Sun Ying, Chang Jiang. Emotion optimization identification emulation for EEG signal. *Computer Simulation*, 2018, **35**(6): 426-431
(王薇蓉, 张雪英, 孙颖, 畅江. 关于脑电信号的情感识别优化仿真. 计算机仿真, 2018, **35**(6): 426-431)
- Krisnandhika B, Faqih A, Pumamasari P D, Kusumoputro B. Emotion recognition system based on EEG signals using relative wavelet energy features and a modified radial basis function neural networks. In: Proceedings of the 2017 International Conference on Consumer Electronics and Devices (ICCED). London, UK: IEEE, 2017. 50-54
- Duan R N, Zhu J Y, Lu B L. Differential entropy feature for EEG-based emotion classification. In: Proceedings of the 6th International IEEE EMBS Conference on Neural Engineering (NER). San Diego, California, USA: IEEE, 2013. 81-84
- Murugappan M, Nagarajan R, Yaacob S. Combining spatial filtering and wavelet transform for classifying human emotions using EEG signals. *Journal of Medical and Biological Engineering*, 2011, **31**(1): 45-51
- Huang Ning-Meng. Emotion Recognition Based on EEG [Master thesis]. South China University of Technology, China, 2016.
(黄柠檬. 基于EEG的情绪识别 [硕士学位论文]. 华南理工大学, 中国, 2016.)
- Yang Mo-Han, Chen Wan-Zhong, Li Ming-Yang. Multiple feature extraction based on ensemble empirical mode decomposition for motor imagery EEG recognition tasks. *Acta Automatica Sinica*, 2017, **43**(5): 743-752
(杨默涵, 陈万忠, 李明阳. 基于总体经验模态分解的多类特征的运动想象脑电识别方法研究. 自动化学报, 2017, **43**(5): 743-752)
- Zhang T, Chen W Z, Li M Y. AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier. *Biomedical Signal Processing and Control*, 2017, **100**(31): 550-559
- Sun Ying, Ma Jiang-He, Zhang Xue-Ying. EEG emotion recognition based on nonlinear global features and spectral feature. *Computer Engineering and Applications*, 2018, **54**(17): 116-121
(孙颖, 马江河, 张雪英. 结合非线性全局特征和谱特征的脑电情感识别. 计算机工程与应用, 2018, **54**(17): 116-121)
- Wu Zhi-Yong, Ding Xiang-Qian, Xu Xiao-Wei, Ju Chuan-Xiang. A method for ECG classification using deep learning and fuzzy C-means. *Acta Automatica Sinica*, 2018, **44**(10): 1913-1920
(吴志勇, 丁香乾, 许晓伟, 鞠传香. 基于深度学习和模糊C均值的心电信号分类方法. 自动化学报, 2018, **44**(10): 1913-1920)
- Jirayucharoensak S, Setha P N, Pasin I. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation [Online], available: <https://doi.org/10.1155/2014/627892>, March 25 2020
- Yang Hao, Zhang Jun-Ran, Jiang Xiao-Mei, Liu Fei. Research of electroencephalography representational emotion recognition based on deep belief networks. *Journal of Biomedical Engineering*, 2018, **35**(02): 182-190
(杨豪, 张俊然, 蒋小梅, 刘飞. 基于深度信念网络脑电信号表征情绪状态的识别研究. 生物医学工程杂志, 2018, **35**(02): 182-190)
- Cheng C L, Wei X W, Zhou J. Emotion recognition algorithm based on convolution neural network. In: Proceedings of the 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Nanjing, China: IEEE, 2017. 1-5
- Wang Y, Huang Z Y, Mccane B, Neo P. EmotioNet: A 3-D convolutional neural network for EEG-based emotion recognition. In: Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN). Rio de Janeiro, Brazil: IEEE, 2018. 1-7
- Kan Wei, Li Yun. Emotion recognition from EEG signals by using LSTM recurrent neural networks. *Journal of Nanjing University (Natural Science)*, 2019, **55**(1): 110-116
(阚威, 李云. 基于LSTM的脑电情绪识别模型. 南京大学学报(自然科学), 2019, **55**(1): 110-116)
- Li Yong, Lin Xiao-Zhu, Jiang Meng-Ying. Facial expression recognition with cross-connect LeNet-5 network. *Acta Automatica Sinica*, 2018, **44**(1): 176-182
(李勇, 林小竹, 蒋梦莹. 基于跨连接LeNet-5网络的面部表情识别. 自动化学报, 2018, **44**(1): 176-182)
- Dutta K K. Multi-class time series classification of EEG signals with recurrent neural networks. In: Proceedings of the 9th International Conference on Cloud Computing, Data Science and Engineering (Confluence). Noida, India: IEEE, 2019. 337-341
- Liang M, Hu X L. Recurrent convolutional neural network for object recognition. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 3367-3375
- Tang Peng-Jie, Wang Han-Li, Xu Kai-Sheng. Multi-objective layer-wise optimization and multi-level probability fusion for image description generation using LSTM. *Acta Automatica Sinica*, 2018, **44**(7): 1237-1249
(汤鹏杰, 王瀚漓, 许恺晟. LSTM逐层多目标优化及多层概率融合的图像描述. 自动化学报, 2018, **44**(7): 1237-1249)

- 20 Koelstra S, Muhl C, Soleymani M, Lee J S, Yazdani A, Ebrahimi T, et al. DEAP: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing*, 2012, **3**(1): 18–31
- 21 Chen Ming. EEG-based Emotion Recognition [Master thesis], Hangzhou Dianzi University, China, 2017.
(陈明. 基于脑电信号的情绪识别 [硕士学位论文], 杭州电子科技大学, 中国, 2017.)
- 22 Xie Jia-Li. Multi-Scale Entropy Algorithm and Its Application in Emotion Recognition [Master thesis], Yanshan University, China, 2016.
(谢佳利. 多尺度熵算法在情感脑电识别中的应用 [硕士学位论文], 燕山大学, 中国, 2016.)
- 23 Li Xia, Lu Guan-Ming, Yan Jing-Jie, Zhang Zheng-Yan. A survey of dimensional emotion prediction by multimodal cues. *Acta Automatica Sinica*, 2018, **44**(12): 2142–2159
(李霞, 卢官明, 闫静杰, 张正言. 多模态维度情感预测综述. 自动化学报, 2018, **44**(12): 2142–2159)
- 24 Bastos-Filho T F, Ferreira A, Atencio A C, Argunan S, Kumar D. Evaluation of feature extraction techniques in emotional state recognition. In: Proceedings of the 4th International Conference on Intelligent Human Computer Interaction (IHCI). Kharagpur, India: IEEE, 2012. 1–6
- 25 Su Jian-Xin. Emotion Recognition Research Based on EEG Signal [Master thesis]. Nanjing University of Posts and Telecommunications, China, 2015.
(苏建新. 基于脑电信号的情绪识别研究 [硕士学位论文], 南京邮电大学, 中国, 2015.)
- 26 Li X, Song D W, Zhang P, Yu G L, Hou Y X, Hu B. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. In: Proceedings of the 2016 IEEE International Conference on Bioinformatics and Bio-medicine (BIBM). Shenzhen, China: IEEE, 2016. 352–359
- 27 Li You-Jun, Huang Jia-Jin, Wang Hai-Yuan, Zhong Ning. Study of emotion recognition based on fusion multi-modal bio-signal with SAE and LSTM recurrent neural network. *Journal on Communications*, 2017, **38**(12): 109–120
(李幼军, 黄佳进, 王海渊, 钟宁. 基于 SAE 和 LSTM RNN 的多模态生理信号融合和情感识别研究. 通信学报, 2017, **38**(12): 109–120)
- 28 Song T F, Zheng W M, Song P, Cui Z. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*, 2020, **11**(3): 532–541
- 29 Shawky E, El-Khoribi R, Shoman M, Wahby Shalaby M. EEG-based emotion recognition using 3D convolutional neural networks. *International Journal of Advanced Computer Science and Applications*, 2018, **9**(8): 329–337



柳长源 哈尔滨理工大学测控技术与通信工程学院副教授. 主要研究方向为机器学习方法研究与改进, 智能优化算法及脑电信号智能诊断技术. 本文通信作者.

E-mail: liuchangyuan@hrbust.edu.cn

(LIU Chang-Yuan Associate professor at the College of Measurement and Control Technology and Communication Engineering, Harbin University of Science and Technology. His research interest covers research and improvement of machine learning methods, intelligent optimization algorithms, and EEG intelligent diagnosis technology. Corresponding author of this paper.)



李文强 哈尔滨理工大学测控技术与通信工程学院硕士研究生. 主要研究方向为深度学习与情感识别.

E-mail: lwqpost@163.com

(LI Wen-Qiang Master student at the College of Measurement and Control Technology and Communication Engineering, Harbin University of Science and Technology. His research interest covers deep learning and emotion recognition.)



毕晓君 中央民族大学信息工程学院教授, 中国人工智能学会自然计算专委会成员, 黑龙江省生物医学工程学会常务副理事长. 主要研究方向为信息智能处理技术, 深度学习及智能优化算法.

E-mail: bixiaojun@hrbeu.edu.cn

(BI Xiao-Jun Professor at the School of Information Engineering, Minzu University of China, member of the Natural Computing Committee of Chinese Artificial Intelligence Society, and Executive Vice Chairman of Heilongjiang Biomedical Engineering Society. Her research interest covers information intelligent processing technology, deep learning, and intelligent optimization algorithm.)