

具有多层次优化能力的 EEG 生成模型^①



张 达^{1,2,3,4}, 郭 特^{1,2,3,4}, 丁 瑞¹, 丁锦红⁵, 周文洁⁵, 李一凡^{1,2,3,4}, 张璐砚^{1,3},
张雨柔^{1,3}, 夏立坤^{1,2,3,4}

¹(首都师范大学 信息工程学院, 北京 100048)

²(首都师范大学 信息工程学院 电子系统可靠性与数理交叉学科国家国际科技合作示范型基地, 北京 100048)

³(首都师范大学 信息工程学院 神经计算与智能感知实验室, 北京 100048)

⁴(首都师范大学 信息工程学院 北京成像理论与技术高精尖创新中心, 北京 100048)

⁵(首都师范大学 心理学院, 北京 100048)

通信作者: 夏立坤, E-mail: xlk@cnu.edu.cn

摘 要: 基于生成对抗网络 (generative adversarial networks, GAN) 的脑电信号 (electroencephalogram, EEG) 生成技术存在生成样本特征单一、幅值差异过大以及拟合速度慢等问题, 其质量难以满足深度学习模型训练和优化的要求。因此, 本文通过对 WGAN-GP 的优化, 使其更适应脑电信号生成, 从而解决以上问题。具体而言: (1) 在 WGAN-GP 网络的框架的基础上, 通过将长短期记忆网络 (long short-term memory, LSTM) 代替卷积神经网络 (convolutional neural networks, CNN), 以保证时间相关特征的完整性, 从而解决脑电生成特征单一的问题; (2) 将标准化处理后的真实脑电信号输入至判别器, 以解决幅值差异过大问题; (3) 将脑电噪声部分作为先验知识输入至网络生成器, 以提高生成模型的拟合速度。本文分别通过 sliced Wasserstein distance (SWD)、mode score (MS) 以及 EEGNet 对生成模型做多层次定量评估。与目前已有生成网络 WGAN-GP 相比较, 基于本模型的生成数据更为接近真实数据。

关键词: 脑电图; 样本生成; 生成对抗网络; 长短期记忆网络; 先验知识

引用格式: 张达, 郭特, 丁瑞, 丁锦红, 周文洁, 李一凡, 张璐砚, 张雨柔, 夏立坤. 具有多层次优化能力的 EEG 生成模型. 计算机系统应用, 2022, 31(8): 369-379. <http://www.c-s-a.org.cn/1003-3254/8658.html>

EEG Generation Model with Hierarchical Optimization

ZHANG Da^{1,2,3,4}, GUO Te^{1,2,3,4}, DING Rui¹, DING Jin-Hong⁵, ZHOU Wen-Jie⁵, LI Yi-Fan^{1,2,3,4}, ZHANG Lu-Fan^{1,3},
ZHANG Yu-Rou^{1,3}, XIA Li-Kun^{1,2,3,4}

¹(College of Information Engineering, Capital Normal University, Beijing 100048, China)

²(International Science and Technology Cooperation Base of Electronic System Reliability and Mathematical Interdisciplinary, Information Engineering College, Capital Normal University, Beijing 100048, China)

³(Laboratory of Neural Computing and Intelligent Perception, Information Engineering College, Capital Normal University, Beijing 100048, China)

⁴(Beijing Advanced Innovation Center for Imaging Theory and Technology, Information Engineering College, Capital Normal University, Beijing 100048, China)

⁵(School of Psychology, Capital Normal University, Beijing 100048, China)

Abstract: Electroencephalogram (EEG) generation via generative adversarial networks (GANs) suffers from various issues including invariant features of samples generated, large amplitude differences, and slow fitting speeds. The quality of signals thus generated fails to meet the requirements of deep-learning model training and optimization. To address the issues above, this study optimizes the Wasserstein GAN gradient penalty (WGAN-GP) so that it can perform better in EEG generation. The details are as follows: (1) On the basis of the framework of the WGAN-GP network, the

① 基金项目: 北京自然科学基金面上项目 (4202011); 国家自然科学基金面上项目 (61572076); 首都师范大学交叉科学研究院引导研发课题 (JCKXYJY 2019018)

收稿时间: 2021-11-17; 修改时间: 2021-12-21; 采用时间: 2022-01-05; csa 在线出版时间: 2022-06-16

convolutional neural network (CNN) is replaced by the long short-term memory (LSTM) network to ensure the integrity of time-dependent features and thereby solve the issue of invariant features; (2) real EEGs are normalized and then applied to the discriminator to reduce the amplitude differences; (3) the noisy parts of EEGs are applied to the generator as prior knowledge to increase the fitting speed of the generation model. Sliced Wasserstein distance (SWD), mode score (MS), and EEGNet are applied to evaluate the proposed generation model quantitatively and hierarchically. Compared with the current generative network WGAN-GP, the proposed model provides data closer to their real counterparts.

Key words: electroencephalogram (EEG); signal generation; generate adversarial networks (GANs); long short-term memory (LSTM); prior knowledge

脑电图 (electroencephalogram, EEG) 作为一种包含大量人体信息的多通道生理信号, 已成为脑科学研究中的重要数据^[1]. 近年来, 深度学习的发展推动了计算机科学与脑科学的进一步融合, 其对 EEG 的特征提取方法成为 EEG 研究的重要手段^[2-4]. 这类技术通常需要训练大量的数据来训练和优化模型, 从而构建鲁棒模型. 但是, 在 EEG 采集过程中, 繁琐的操作极易导致被试人员受到主观意识的影响, 从而产生眼电、心电等掩盖真实 EEG 的波形的噪声. 因此, 获取满足深度学习模型训练所需的数据量极其困难. 针对上述问题, 研究人员通过设计针对于 EEG 的生成技术, 以实现样本数量的扩充^[5,6].

目前, 样本生成模型主要包括自回归模型 (autoregressive model, AR)^[6]、自动编码变分贝叶斯 (variational auto-encoding Bayes, VAE)^[7]、生成对抗网络 (generative adversarial network, GAN)^[8] 以及流模型 (flow-based model)^[9] 等. 相较而言, 前两者缺乏判别模型的指导, 因此在训练过程中极易出现生成样本单一的情况; 而流模型在训练过程中对计算资源的要求很高, 难以得到广泛应用. 综合以上分析, GAN 及其相关衍生技术成为 EEG 生成的最佳模型. 相对而言, 传统 GAN 及其衍生模型的构建主要通过全连接层 (fully connected layer, FC) 来实现, 而 FC 层无法捕捉时序信号的特征和相关属性之间的复杂关联^[10], 这就导致模型在处理 EEG 时容易丢失时间相关性信息.

为解决此问题, Abdelfattah 等人^[11] 通过使用循环神经网络 (recurrent neural network, RNN) 搭建循环生成对抗网络框架 (recurrent generative adversarial networks, RGAN), 从而实现多通道 EEG 的生成. 为验证其性能, 作者将生成 EEG 用于训练分类框架中并取得了较好的分类结果. 相较于其他神经网络, RNN 对

时序信号特性具有较强的学习能力, 能够保留时间相关信息; 但是, 由于其本身存在梯度爆炸等问题, 容易导致生成模型训练不稳定. Luo 等人^[12] 提出 conditional Wasserstein GAN (CWGAN) 框架, 通过在输入生成器的随机噪声中添加辅助条件, 以用于生成 EEG 的微分熵 (differential entropy, DE) 特征, 并将其应用到情感识别任务中. 其中, 辅助条件能够加速模型的收敛速度, 从而保证模型训练的稳定性. 然而, DE 只适用于情感分析任务, 具有一定的局限性. 另外, Panwar 等人^[13] 提出 class conditioned WGAN-GP (CC-WGAN-GP) 模型, 并将其用于多通道 EEG 信号的生成. 该模型在判别器中增加了分类功能, 即在生成 EEG 的同时能够对其分类, 之后采用分类准确率代替损失函数来判定该模型的训练情况. 相较于已有模型, CC-WGAN-GP 在模型稳定性以及样本多样性方面都有所提高, 但是, 额外的分类器增加了模型的复杂度, 从而造成模型拟合速度降低.

综上所述, 基于 GAN 及其衍生模型的 EEG 生成模型能够证明所生成样本在深度学习模型训练和优化研究中的可行性. 然而, 上述生成模型仍存在以下问题: (1) 模型通常添加批标准化层 (batch normalization, BN) 来防止模型过拟合, 但这会导致真实样本与生成样本之间的幅值差异过大. (2) 模型训练过程中易出现过拟合、模式崩溃及梯度消失等问题, 导致生成样本多样性不足; (3) 模型通过增加模型复杂度来提高模型训练的稳定性, 这往往导致模型的拟合速度降低.

本文拟从以下 2 个方面解决上述 3 个问题, 具体而言: (1) 在 WGAN-GP 网络框架下, 通过将长短期记忆网络 (long short-term memory, LSTM) 代替卷积神经网络 (convolutional neural networks, CNN), 从而解决脑电信号生成单一的问题; 这是由于 LSTM 能够保留

EEG 信号的时间相关性特征,从而提高生成样本中特征的多样性.在此模型的基础上,本文分别对网络框架的生成器和判别器的输入数据进行处理:(2)在 EEG 预处理阶段,我们将标准化处理后的真实脑电信号输入至判别器,即使其值域映射到 $[-1, 1] \mu\text{V}$,从而解决生成样本与真实样本幅值差异过大的问题.(3)考虑到 EEG 噪声中必然存在与真实 EEG 相似特征,我们将其代替传统的高斯噪声输入至生成器,即,将其作为先验知识来提高生成模型的拟合速度;

为综合评估生成模型的性能,本文采用多层次定量评估方法,从不同角度验证其性能,具体而言:首先,通过使用 sliced Wasserstein distance (SWD) 和 mode score (MS) 实现对生成模型的相似性和多样性的定量评估.其次,通过对比真实样本与生成样本的波形差异对其进行定性评估.最后,将生成样本与真实样本以不同比例混合并用于 EEGNet 模型的训练;通过对分类结果的定量分析,以验证生成模型的性能.

1 相关技术介绍

为更好地阐述所提出的 EEG 生成方法,本节主要阐述工作中所需的理论知识:首先,介绍 WGAN 以及 WGAN-GP 的相关知识;其次,介绍本文从损失函数、内部结构以及工作原理等角度对使用的网络框架—LSTM 进行详细描述.

1.1 Wasserstein GAN

Wasserstein GAN (WGAN)^[14] 通过将 Wasserstein 距离 (Wasserstein distance, WD)^[15] 代替标准 GAN 中的交叉熵来作为损失函数,以解决原始 GAN 中模式崩溃以及过拟合等问题.

WD 又称为推土机距离 (earth-mover's distance, EMD),定义如式 (1) 所示:

$$W(P_r, P_g) = \inf_{\gamma \sim \Pi(P_r, P_g)} E_{(x,y) \sim \gamma} [\|x_r - x_g\|] \quad (1)$$

其中, γ 为真实样本分布 P_r 与生成样本分布 P_g 的集合,由于该集合期望的下边界 (*inf*) 无法直接求出,故使用 Lipschitz 连续代替.

定义 1. Lipschitz 连续: 在连续函数 $f(x)$ 上额外施加一个限制,若存在一个常数 K 使得定义域内的任意两个元素 x_1 和 x_2 都满足式 (2),则称 K 为 $f(x)$ 的 Lipschitz 连续常数.

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2| \quad (2)$$

通过 Lipschitz 连续变换后,WD 的定义如式 (3):

$$W(P_r, P_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} E_{x \sim P_r}[f(x)] - E_{x \sim P_g}[f(x)] \quad (3)$$

$W(P_r, P_g)$ 即为函数 $f(x)$ 的 Lipschitz 常数 $\|f\|_L$ 在不超过 K 的条件下,对所有可能满足条件 f 的取到的上界 $E_{x \sim P_r}[f(x)] - E_{x \sim P_g}[f(x)]$. 因此,我们可以用一组参数来定义一系列可能的函数 f_ω , 此时求解式 (3) 可以近似变成求解如下形式:

$$K \cdot W(P_r, P_g) \approx \max_{\|f\|_L \leq K} E_{x \sim P_r}[f(x)] - E_{x \sim P_g}[f(x)] \quad (4)$$

综合以上分析, WGAN 的生成器和判别器的损失函数分别如式 (5) 和式 (6) 所示:

$$L_G = -E_{x \sim P_g}[f(x)] \quad (5)$$

$$L_D = E_{x \sim P_g}[f(x)] - E_{x \sim P_r}[f(x)] \quad (6)$$

WGAN 能够解决 GAN 所存在的梯度消失、生成样本单一等问题,但 Lipschitz 连续的引入也带来了梯度爆炸的问题.

1.2 Wasserstein GAN-Gradient Penalty

针对 WGAN 存在的问题, Gulrajani 等人^[16] 在 WGAN 的判别器的损失函数中加入 GP 项,如式 (7) 所示:

$$GP = \lambda E_{x \sim P_x} [\|\nabla_x D(x)\|_p - K]^2 \quad (7)$$

其中, λ 为超参数,表示在真实样本 x_r 和生成样本 x_g 的分布连线上的随机插值取样.

因此,其生成器的损失函数如式 (8) 所示:

$$L_D = E_{x \sim P_g}[D(x)] - E_{x \sim P_r}[D(x)] + \lambda E_{x \sim P_x} [\|\nabla_x D(x)\|_p - 1]^2 \quad (8)$$

其判别器的损失函数如式 (9) 所示:

$$L_G = -E_{x \sim P_g}[f(x)] \quad (9)$$

总体来说, WGAN-GP (Wasserstein GAN-gradient penalty) 保留了 WGAN 的优点,同时,在判别器 D 的损失函数上加入梯度惩罚项;这在提升训练速度的同时,也保证梯度更新在可控范围内,进而降低 WGAN 梯度爆炸的可能性.

1.3 长短期记忆网络

LSTM 网络具有前向传播链式结构.相较于其他处理时间序列的网络, LSTM 更适用于分析长期记忆的时序信号^[17,18],其结构如图 1 所示,其中, x 为输入到 LSTM 单元的训练数据, h 为 LSTM 单元的输出数据.

LSTM 单元包含遗忘门、输入门和输出门 3 个部分, 其结构如图 2 所示, 以控制 LSTM 网络的记忆与遗忘^[19].

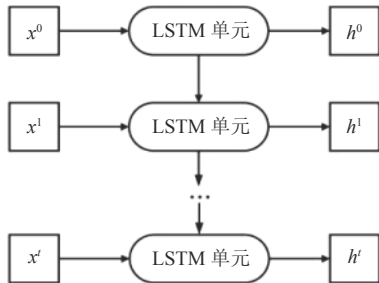


图 1 LSTM 网络结构

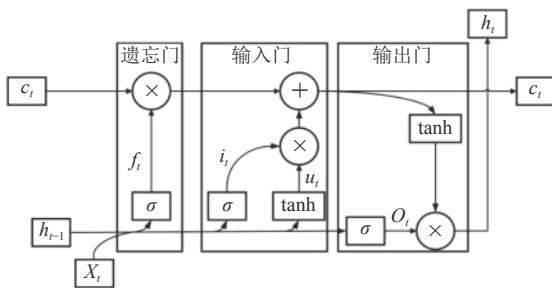


图 2 LSTM 单元结构

遗忘门通过接收前一个 LSTM 单元的输出, 决定是否丢弃前一时刻的信息 h_{t-1} ; 输入门决定当前 LSTM 单元需要保存的信息 i_t , 同时生成备选更新信息 u_t , 将 i_t 和 u_t 两部分相结合, 共同更新当前 LSTM 单元状态; 输出门决定是否输出当前 LSTM 单元的状态, 并获得当前 LSTM 的隐藏状态 h_t .

得益于特殊的单元结构, LSTM 网络能够缓解在模型训练过程中出现的梯度爆炸和梯度消失的问题, 并且能够提取更多的时间域特征.

2 框架设计

为解决生成 EEG 过程中存在的问题, 本文拟将 EEG 生成过程分为数据预处理阶段和数据生成两个阶段, 如图 3 所示. 在数据预处理阶段, 本文首先使用独立成分分析 (independent component analysis, ICA)^[20] 对真实 EEG 做滤波处理, 从而将 EEG 数据与 EEG 噪声分开. 生成器采用后者作为输入数据, 这是由于后者包括被试人员的眼动、呼吸以及肌电等信号, 而这些信号中依然存在重要的特征. 因此, 我们将 EEG 噪声作为先验知识输入到生成器中, 从而提高模型的拟合

速度; 其次, 我们对滤波后 EEG 信号做标准化处理, 使其值域映射到 $[-1, 1] \mu\text{V}$, 从而解决生成样本与真实样本幅值差异过大的问题.

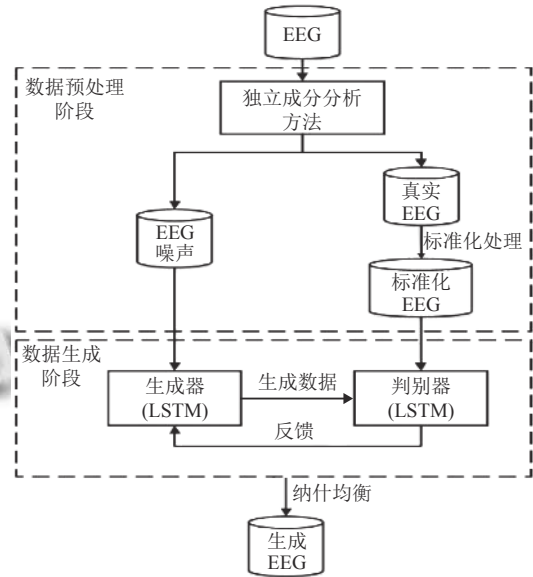


图 3 EEG 生成流程图

在数据生成阶段, 本文通过将 CNN 替换成 LSTM, 以构建新的生成器和判别器, 进而保证 EEG 中时间相关性信息的完整性, 这是由于 LSTM 网络具有前向传播链式结构, 包含多个 LSTM 单元. 相较于其他处理时间序列的网络, LSTM 更适用于需要长期记忆的时序信号. 最后, 当生成器和判别器的对抗训练达到纳什均衡状态时, 生成样本的质量达到最优.

2.1 EEG 噪声

生成器通常将具有高斯分布的随机噪声作为输入数据, 但是由于 EEG 信号的特殊性, 此类噪声会导致生成模型的拟合速度降低. 本文拟将 EEG 噪声来代替高斯噪声作为生成器输入数据, 这是由于 EEG 噪声中存在与 EEG 信号相似的特征信息. 因此, 当其代替随机噪声, 生成器能够更高效地拟合真实样本分布.

本文采用 ICA 对所采集的 EEG 信号进行过滤, 进而分离出 EEG 信号和干扰噪声. ICA 能够将相互独立的样本成分从混合样本中提取出来^[21], 其模型一般如式 (10) 所示:

$$X = A \times S \quad (10)$$

其中, X 是观测样本; A 为系数矩阵, 包含所需分析数据的隐含特征; S 代表独立源信号. 假设 X 服从独立程度

量原则^[22], 在已知 A 的情况下, 则有 $S=A^{-1} \times X=WX$, 其中 $W=A^{-1}$ 一般称为分离矩阵.

2.2 EEG 标准化

在搭建网络框架时, 本文在生成器中嵌入 BN 层^[23]. BN 层能够将上层神经网络的输出数据分布强行转换为标准的正态分布 (mean=0, std=1), 使得输出数据分布在非线性函数较为敏感的区域, 从而避免梯度消失并且加快模型的收敛速度. 然而, BN 层的转换操作能够将生成器的输出数据值域限制在 $[-1, 1]$, 这就导致生成样本的值域远远小于真实样本的值域范围. 因此, 本文通过使用最大最小值方法将 EEG 信号的值域范围映射到 $[-1, 1]$ μV , 一方面能够提高模型的收敛速度, 另一方面能够保证生成样本与真实样本的幅值在相同范围内. 最大最小值方法如式 (11) 所示:

$$y = \frac{(x - x_{\text{mean}})}{x_{\text{max}} - x_{\text{min}}} \quad (11)$$

其中, x_{mean} 为数据的均值, x_{min} 为数据中的最小值, x_{max} 为数据中的最大值, y 为标准化后的数据.

2.3 生成对抗网络设计

EEG 作为一种非线性、随机性强的时序信号, 其时间相关性比其他时序信号 (音频、语音等) 更为复杂^[24,25]. 另一方面, GAN 的衍生框架大多采用 FC 或者 CNN 来提取特征^[26,27], 但二者均难以捕捉到时间序列信号的时间特征和相关属性之间的复杂关联性, 致使生成样本丢失真实样本的时间相关性信息^[12], 从而导致生成样本特征单一. 本文将 WGAN-GP 作为基础网络框架, 使用 LSTM 代替 CNN, 并结合第 2.1 节和第 2.2 节的创新点, 以解决目前 EEG 生成模型存在的问题. 改进后的网络框架如图 4 所示.

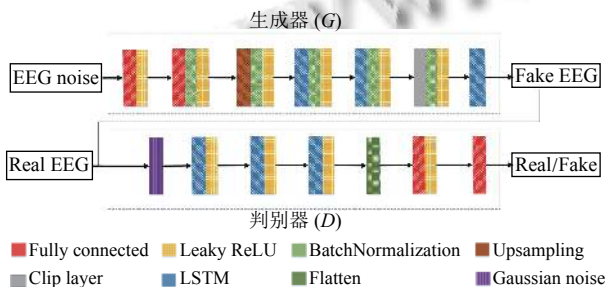


图 4 改进后的 WGAN-GP 网络框架

改进后的判别器体系结构如表 1 所示. 首先, 我们将与真实 EEG 维度 (231×1) 相同的 EEG 噪声信号作为生成器的输入数据, 并使用两层 fully connected 层将

其维度逐步扩充至 1×15360; 为与单元数为 128 的 LSTM 层对应, Reshape 层将上层网络的输出数据转换为维度为 1×120×128 的三维数据; 随后, 我们采用双次插值法对三维数据进行上采样操作, 这一操作使得数据提高到 1×240×128; 为满足 LSTM 的输入要求, Reshape 层将输出维度再次转换为 240×28, 然后将输出数据反馈到两层 LSTM 层中. 接下来, Clip 层将 LSTM 层的输出数据整合到真实 EEG 的维度; 最后, 使用一层 LSTM 对上层输出进行整合, 使其维度满足判别器的输入维度. 此外, 本文在生成器的每一层神经网络后加入 batch normalization (BN) 层并使用 Leaky ReLU 激活函数以防止模型在训练过程中出现梯度消失等问题.

表 1 生成器体系结构

Layer	Output Shape	Parameters size
Input layer	231×1	—
Fully connected	1×1024	1 024
Leaky ReLU	1×1024	—
Fully connected	1×15360	15 360
Batch Normalization	1×15360	—
Leaky ReLU	1×15360	—
Reshape	1×120×128	15 360
Upsampling	1×240×128	240
Batch Normalization	1×240×128	—
Leaky ReLU	1×240×128	—
Reshape	240×128	128
LSTM	240×128	128
Batch Normalization	240×128	—
Leaky ReLU	240×128	—
LSTM	240×64	64
Batch Normalization	240×64	—
Leaky ReLU	240×64	—
Reshape	1×240×64	15 360
Clip layer	1×231×64	231
Batch Normalization	1×231×64	—
Leaky ReLU	1×231×64	—
Reshape	231×64	29 568
LSTM	231×1	1

改进后的判别器体系结构如表 2 所示. 本文将维度为 231×1 真实 EEG 信号以及生成器的输出数据作为判别器的输入数据. 为提高判别器的稳定性, 本文通过添加高斯白噪声 (mean=0, std=0.05) 对输入数据进行腐蚀. Goodfellow 等人^[28,29] 认为, 模型不稳定的主要原因是真实样本分布和生成样本分布具有不相交的特性, 而在最优判别器的情况下, 这将导致梯度消失. 因此, 我们在判别器中加入高斯白噪声, 可以避免梯度消失, 从而提高生成模型的稳定性. 其次, 我们添加 3 个

具有 Leaky ReLU 的 LSTM, 以实现对 EEG 的时间特征提取操作; 其参数设置与生成器的 LSTM 层参数相对应, 从而保证判别器与生成器对数据的同步处理; 接下来, Flatten 层将 LSTM 层的输出数据的维度降维至 29 568, 最后, 经过具有单神经元的 Fully connected 层输出判别器的鉴别结果. 本文在判别器未添加 BN 层, 这是由于 BN 层会破坏 Lipschitz 连续的约束范围, 导致模型出现难以收敛的现象.

表 2 判别器体系结构

Layer	Output shape	Parameters size
Input layer	231×1	—
Gaussian noise	231×1	—
LSTM	231×64	64
Leaky ReLU	231×64	—
LSTM	231×128	128
Leaky ReLU	231×128	—
LSTM	231×128	128
Leaky ReLU	231×128	—
Flatten	29 568	—
Fully connected	1	1

3 实验结果分析

本文使用 Python 语言和 Keras (TensorFlow) 深度学习框架实现对模型的训练, 并使用 Bi2015a 数据集^[30]作为训练数据. Bi2015a 数据集包含了 50 名受试者在进行《大脑入侵者》游戏时的 EEG 记录. 该游戏在一个由 36 个符号 (1 个目标符号, 35 个非目标符号) 组成的网格上使用了 oddball 实验范例, 这些符号随机闪现以引起被试人员 P300 的脑电波形. 实验采用 32 个电极, 分别在 50 ms、80 ms 和 110 ms 三种条件下记录 EEG 数据. 经过初步筛选, 我们删除了 15 个存在损坏通道的 EEG 数据; 同时, 对比 32 个通道的 P300 特征, 发现 C3 通道的 P300 特征最明显. 因此, 本文使用 35 个被试的 C3 通道的数据进行实验. 在预处理阶段, 本文将未经处理的 EEG 信号进行降采样处理, 使其频率降至 256 Hz; 同时, 为保留 P300 特征波形, 我们按照事件标签, 截取事件发生前 100 ms 至事件发生后 500 ms, 共计 900 ms 的 EEG 数据, 并按照式 (12) 计算得到维度为 231×1 的训练数据.

$$l = f \times t \quad (12)$$

其中, l 为分割后的数据长度, f 为频率, t 为时间.

本文将每一类 EEG 数据分割并累加, 分别得到 3 411、2 862、2 436 个维度为 231×1 的数据. 因此, 本

文使用 8 709 个 EEG 样本作为训练数据.

3.1 模型评价指标

3.1.1 相似性评估方法

本文采用 sliced Wasserstein distance (SWD)^[31]对生成样本的相似性进行评估, 如式 (13) 所示:

$$SWD(g, x) = \int_{S^{d-1}} WD(R_\theta g, R_\theta x) d\theta \quad (13)$$

其中, S 表示将样本分布映射到一维空间的空间单元, 其满足 $\int_{S^{d-1}} d\theta = 1$, θ 表示映射角度, R_θ 表示在生成样本 g 和真实样本 x 上的一维线性投影运算.

SWD 通过计算两个样本分布的所有一维投影之间的 WD, 来拟合样本整体的分布. SWD 的值越低, 表示两种分布在外观和样本的变化上越相似.

3.1.2 多样性评估方法

Mode score (MS)^[32]通过计算样本标签分布的交叉熵来评估二者之间的差异, 如式 (14) 所示:

$$MS(P_g) = \exp(E_{x \sim P_g} [KL(p(y|x) \| p(y)) - KL(p(y) \| p(y^*))]) \quad (14)$$

其中, KL 表示交叉熵, $x \sim P_g$ 表示生成样本, $p(y)$ 表示生成样本标签 y 的分布, 满足 $p(y) = \int_x p(y|x) dp_g$, $p(y^*)$ 表示真实样本的标签分布, $p(y|x)$ 表示通过预测得到的 x 对标签 y 的分布. 若生成 EEG 分类特征越明显, 则 $p(y)$ 的熵值越大. 因此, MS 越高, 说明样本的多样性越好.

3.1.3 准确性评估方法

为评估生成样本的准确性, 本文将真实 EEG 与生成的 EEG 以不同比例混合, 然后通过使用 EEGNet^[33]对这些数据进行分类. EEGNet 将其作为激活函数以保证分类模型的稳定性^[34], 如式 (15) 所示:

$$f(x) = \begin{cases} \alpha(e^x - 1), & \text{if } (x < 0) \\ x, & \text{if } (x \geq 0) \end{cases} \quad (15)$$

其中, α 为可调整参数, 用于控制 ELU 负值的饱和位置.

本文使用 EEGNet 对真实 EEG 和生成的 EEG 的混合数据进行分类, 从准确率、召回率、精确率等方面对生成模型进行二次定量评估.

3.2 实验结果与分析

我们拟从多角度对本文所设计的生成模型进行验证, 具体包括 4 个实验: (1) 使用叠加噪声的正弦信号模拟 EEG, 验证 BN 层对生成样本幅值的影响; (2) 使用 CNN 和 LSTM 搭建生成模型并用于生成 EEG, 验

证不同神经网络对生成样本多样性和相似性的影响。

(3) 使用 EEG 噪声代替高斯噪声作为生成器的输入数据, 验证 EEG 噪声对生成器拟合速度的影响; (4) 将生成 EEG 与真实 EEG 以不同比例混合, 通过 EEGNet 的训练数据, 验证生成的 EEG 的准确性。

3.2.1 BN 层对模型的影响

在生成 EEG 时, 我们发现生成的 EEG 与真实 EEG 的值域具有明显差异. 由于其原因无法直接判定, 本文设计人工信号对生成模型进行训练, 以验证 BN 层对生成样本值域范围变化的影响. 我们以正弦函数 ($\text{range}=[-1.0, 1.0]$, $f=(10/\pi)$ Hz) 为基础函数, 并叠加一组值域为 $(-0.4, 0.4)$ 的高斯白噪声 ($\text{mean}=0$, $\text{std}=1$) 来增加输入数据的多样性以及提升生成模型的抗干扰能力. 最后我们对信号做标准化处理, 以得到值域为 $[-1.0, 1.0]$ 的输入信号. 合成信号如图 5 所示.

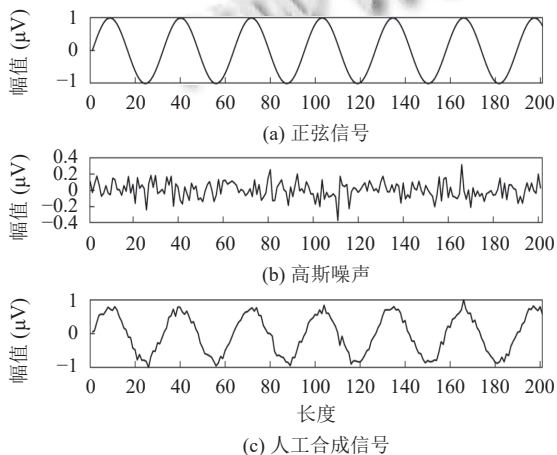


图5 人工合成正弦信号

我们将人工信号输入到由 CNN 和 LSTM 搭建的生成框架中, 从而对比在有/无 BN 层情况下生成样本的值域变化, 实验结果如表 3 所示。

表3 人工信号生成结果

网络框架	BN层	幅值范围 (μV)
CNN	√	$(-0.50, 0.50)$
	—	$(-1.00, 1.00)$
LSTM	√	$(-0.02, 0.04)$
	—	$(-1.00, 1.00)$

从表 3 可以看出, 在加入 BN 层的框架中, 生成样本的幅值都比真实样本的幅值小 2–50 倍; 而在去掉 BN 层的框架中, 与真实样本的值域范围相近。

我们发现, 在保留 BN 层的网络框架中, 生成样本存在非周期性的随机性波形, 如图 6(a)、图 6(b) 所示;

图 6(c)、图 6(d) 展示了在去除 BN 层后所生成的波形, 可以看出, 它们具有正弦信号的形状, 但丢失了随机噪声特征, 这说明模型并未生成较为真实的信号. 实验结果表明, BN 层对数据进行批标准化处理后将导致生成数据幅值变小. 但是, 在去掉 BN 层后, 模型出现过拟合的问题. 因此, 为解决生成样本与真实样本幅值差异的问题, 本文先对滤波后的真实 EEG 信号进行标准化处理, 将 EEG 信号的值域范围映射到 $[-1, 1]$ μV , 再输入到判别器中。

3.2.2 不同网络框架对生成模型的影响

我们通过使用 CNN 和 LSTM 搭建 4 种不同的生成框架 WGAN-GP ($G_D D_D$ 、 $G_D D_L$ 、 $G_L D_D$ 、 $G_L D_L$) 以验证神经网络对生成 EEG 质量的影响. 其中, G_D 表示由 CNN 搭建的生成器, G_L 表示由 LSTM 搭建的生成器, D_D 表示由 CNN 搭建的判别器, D_L 表示由 LSTM 搭建的判别器; 同时, 本文通过使用 *SWD* 和 *MS* 分别对不同框架生成的 EEG 的相似性和多样性进行评估, 评估结果如表 4 所示。

由表 4 的客观指标可知, 对于 *SWD* 指标, 生成器或判别器任意一方使用 LSTM 代替 2DCNN 时, *SWD* 值都有所提高, 且 WGAN-GP ($G_D D_L$) 比 WGAN-GP ($G_L D_D$) 的 *SWD* 值高 0.2; $G_L D_L$ 的 *SWD* 值为 1.9, 明显优于其他 3 个框架的结果, 可以证明本文算法能够生成与真实 EEG 更为相似的样本. 此外, 从 *MS* 值可以看出, 基于 WGAN-GP ($G_D D_L$) 的 *MS* 值远高于其他 3 个框架, 但 WGAN-GP ($G_L D_L$) 的值最低, 这表明相较于 WGAN-GP ($G_L D_L$), WGAN-GP ($G_D D_D$) 能够生成更多样化的样本. 综上所述, 定量评估结果出现了相互矛盾的问题. 经过多次实验, 我们发现生成样本的多样性和相似性结果呈负相关, 即生成样本的相似性越高, 多样性就越差。

为进一步评估生成的 EEG 的质量, 本文从视觉角度对基于不同框架生成的 EEG 波形进行分析. 从图 7(a)–图 7(c) 可以看出, WGAN-GP ($G_D D_D$)、WGAN-GP ($G_D D_L$) 以及 WGAN-GP ($G_L D_D$) 生成的 EEG 的值域范围与真实 EEG 的值域范围明显不符. 从图 7(d) 可以看出, WGAN-GP ($G_L D_L$) 生成的 EEG 的值域范围更接近真实 EEG 的值域范围, 且生成 EEG 波形与真实 EEG 波形存在相似波段. 因此, 从视觉评估的角度分析, 基于 WGAN-GP ($G_L D_L$) 生成的 EEG 更接近于真实 EEG。

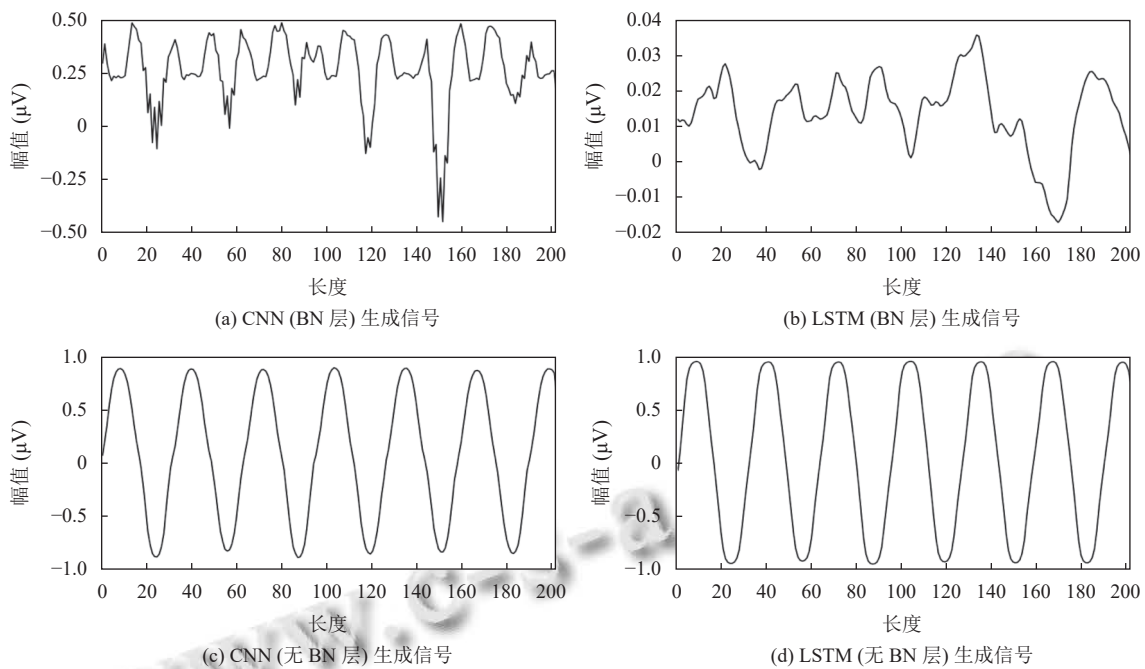


图6 生成的正弦信号

表4 不同网络框架生成 EEG 的定量评估结果

网络框架	框架选择		评估指标	
	生成器	判别器	<i>SWD</i>	<i>MS</i>
WGAN-GP ($G_D D_D$)	CNN	CNN	3.37	12.25
WGAN-GP ($G_D D_L$)	CNN	LSTM	2.83	570.44
WGAN-GP ($G_L D_D$)	LSTM	CNN	3.03	56.92
WGAN-GP ($G_L D_L$)	LSTM	LSTM	1.90	3.95

3.2.3 不同噪声对生成模型的影响

相较于 CNN 搭建的生成框架, LSTM 搭建的生成框架在处理速度上更为缓慢. 为进一步提升模型的拟合速度, 本文提出 WGAN-GP ($G_L D_L^{+N}$), 即将 EEG 噪声输入到生成器中以用于生成 EEG. 这是因为 EEG 噪声存在与真实 EEG 相似的信息, 本文将其作为先验知识来提升生成器的拟合速度. 我们记录了将不同噪声作为输入数据时, 模型的训练时间, 在结果如表 5 所示.

由表 5 可知, 相较于 WGAN-GP ($G_L D_L$), WGAN-GP ($G_L D_L^{+N}$) 的训练时间缩短了 1 h 42 min, 这表明 EEG 噪声的加入能够提高生成器的拟合速度.

为进一步评估 EEG 噪声对生成模型的影响, 本文使用 *SWD* 和 *MS* 对生成的样本进行定量评估, 结果如表 6 所示.

从表 6 可以看出, 相较于 WGAN-GP ($G_L D_L$), WGAN-GP ($G_L D_L^{+N}$) 的 *MS* 值要高出 15.56, 这证明本文的模型在应用 EEG 噪声训练过后可以获得多样性

更好的生成样本; 但是由于生成样本的多样性和相似性结果呈负相关, 这就导致 WGAN-GP ($G_L D_L^{+N}$) 的 *SWD* 值比 WGAN-GP ($G_L D_L$) 的 *SWD* 值高. 另一方面, 仅通过定量分析可能会存在片面性, 因此, 本文将视觉分析结果作为另一个重要指标, 其结果如图 8 所示. 由图可以看出, 基于 WGAN-GP ($G_L D_L^{+N}$) 生成的 EEG 在幅度和波形上更符合真实 EEG 的特征.

3.2.4 EEGNet 分类实验

本文将 CC-WGAN-GP 和 $G_L D_L^{+N}$ 生成的 EEG 以不同比例与真实 EEG 混合对 EEGNet 模型进行训练, 并将分类结果与基于真实样本的分类结果相对比, 验证生成样本对分类器性能的影响. 首先, 本文使用 8 709 个真实 EEG 训练 EEGNet; 然后, 分别使用两个不同框架的 50% 生成 EEG 和 50% 的真实 EEG 进行混合来训练 EEGNet; 最后, 在真实 EEG 数据集中加入相同数量的生成样本来训练 EEGNet. 实验结果如表 7 所示.

由表 7 结果分析可知, 在样本数量不变的情况下, 使用 50% 的 CC-WGAN-GP 框架生成数据代替真实数据训练模型, 其准确率降低了 10.04%, 精确率降低了 3.43%, 召回率降低了 9.47%, *F1-score* 降低了 10.88, *AUC* 降低了 7.86%. 实验结果表明, CC-WGAN-GP 生成的样本不能够代替真实样本来训练分类模型, 反而

会影响分类模型的性能. 使用 50% 的 $G_L D_L^{+N}$ 框架生成数据代替真实数据训练模型时, 其准确率、精确率、召回率相差 1% 左右, $F1$ -Score 相差 2.20%, AUC 相差 0.23%, 从而证明 $G_L D_L^{+N}$ 生成的 EEG 与真实 EEG 相似, 能够代替真实数据作为分类器的训练数据. 此外, 将相同数量的 $G_L D_L^{+N}$ 框架生成的数据加入到真

实数据中训练 EEGNet, 准确率提高了 8.5%, 精确率提高了 16.5%, 召回率提高了 9.25%, $F1$ -score 提高了 12.73%, ROC-AUC 提高了 4.37%. 由结果看出, 在增加生成数据的情况下, 模型在分类精度、稳定性以及泛化能力等方面都能够提升; 进而证明, 我们所设计的 EEG 生成模型能够提升分类模型的性能.

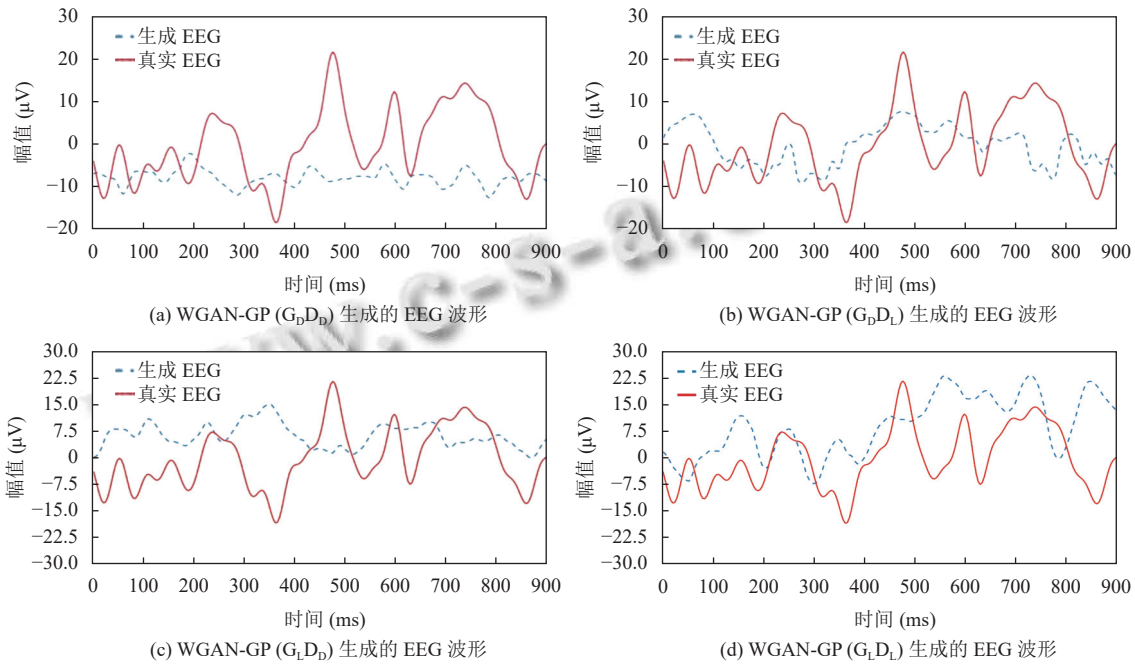


图 7 生成的 EEG 信号

表 5 不同框架的训练时间对比结果

框架	训练时间
WGAN-GP ($G_L D_L^{+N}$)	7 h 7 min
WGAN-GP ($G_L D_L$)	8 h 49 min

表 6 生成样本相似性以及多样性对比结果

框架	噪声	评估指标	
		SWD	MS
WGAN-GP ($G_L D_L^{+N}$)	EEG噪声	3.08	19.51
WGAN-GP ($G_L D_L$)	高斯噪声	1.90	3.95

表 7 基于不同混合比例的样本分类结果 (%)

数据	框架	准确率	精确率	召回率	$F1$ -score	AUC
真实数据 (8709)	—	56.79±1.45	56.02±3.60	55.28±34.79	49.86±21.12	78.66±1.13
真实样本+生成样本 (8709)	CC-WGAN-GP	46.75±0.99	52.59±7.12	45.81±38.01	38.98±16.50	70.80±0.89
真实样本+生成样本 (8709)	$G_L D_L^{+N}$	57.77±0.07	60.93±13.36	56.62±33.79	52.15±20.62	78.89±4.46
真实样本+生成样本 (17418)	$G_L D_L^{+N}$	65.29±7.68	72.52±18.82	64.53±27.97	62.59±14.20	83.03±6.31

4 结束语

本文提出一种新型 EEG 信号生成模型, 以解决目

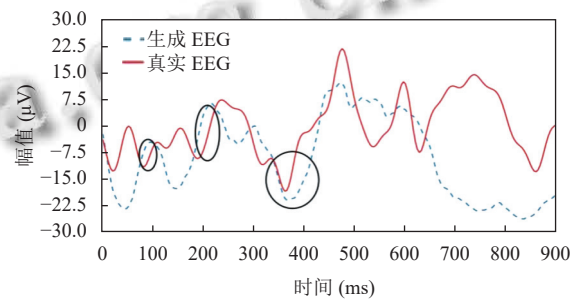


图 8 WGAN-GP ($G_L D_L^{+N}$) 生成的 EEG 波形

前存在的生成样本多样性不足、拟合速度慢以及幅值差异过大等问题. 为评估生成模型的性能, 本文从相似

性、多样性、准确性以及训练时间等方面对生成模型进行评估。与 WGAN-GP 相比较,本文提出的生成模型在相似性和多样性方面的性能均有所提高。此外,我们使用 EEGNet 对模型进行二次定量评估,分类结果表明此模型能够生成较为准确的样本,且能够提高分类模型的性能。本文所提出的 EEG 信号生成模型将为构建高性能的深度学习模型提供保证,同时能够极大降低人力资源消耗和成本。因此,本研究具有十分重要的实际意义和应用价值。

目前,本文仅完成单通道 EEG 生成的工作,而在实际应用中,基于单通道的信息可能无法满足模型泛化能力的需求。未来,我们将继续优化生成模型,以用于多通道 EEG 信号生成,使其能够更加有效地应用于与 EEG 相关的工作中,例如,情感识别、脑控技术等。

参考文献

- 1 李颖洁,李玉玲,杨帮华.基于脑电信号深度学习的情绪识别研究现状.北京生物医学工程,2020,39(6):634-642. [doi: 10.3969/j.issn.1002-3208.2020.06.014.]
- 2 Bochet A, Sperdin HF, Rihs TA, *et al.* Early alterations of large-scale brain networks temporal dynamics in young children with autism. *Communications Biology*, 2021, 4(1): 968. [doi: 10.1038/s42003-021-02494-3]
- 3 Xu XY, Jia TY, Li Q, *et al.* EEG feature selection via global redundancy minimization for emotion recognition. *IEEE Transactions on Affective Computing*, 2021. [doi: 10.1109/TAFFC.2021.3068496]
- 4 Song TF, Liu SY, Zheng WM, *et al.* Variational instance-adaptive graph for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 2021. [doi: 10.1109/TAFFC.2021.3064940]
- 5 Jiao YY, Deng YN, Luo Y, *et al.* Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks. *Neurocomputing*, 2020, 408: 100-111. [doi: 10.1016/j.neucom.2019.05.108]
- 6 Roy S, Dora S, McCreddie K A, *et al.* MIEEG-GAN: Generating artificial motor imagery electroencephalography signals. *Proceedings of 2020 International Joint Conference on Neural Networks (IJCNN)*. Glasgow: IEEE, 2018. 1-8.
- 7 Kingma DP, Welling M. Auto-encoding variational Bayes. *Proceedings of 2nd International Conference on Learning Representations (ICLR)*. Banff: ICLR, 2014. 1-14.
- 8 Goodfellow IJ, Pouget AJ, Mirza M, *et al.* Generative adversarial nets. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2014. 2672-2680.
- 9 Kingma DP, Dhariwal P. Glow: Generative flow with invertible 1×1 convolutions. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Red Hook: MIT Press, 2018. 10236-10245.
- 10 Yu LT, Zhang WN, Wang J, *et al.* SeqGAN: Sequence generative adversarial nets with policy gradient. *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. San Francisco: AAAI Press, 2017. 2852-2858.
- 11 Abdelfattah SM, Abdelrahman GM, Wang M. Augmenting the size of EEG datasets using generative adversarial networks. *Proceedings of 2018 International Joint Conference on Neural Networks (IJCNN)*. Rio de Janeiro: IEEE, 2018. 1-6.
- 12 Luo Y, Lu BL. EEG Data augmentation for emotion recognition using a conditional wasserstein GAN. *Proceedings of 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu: IEEE, 2018. 2535-2538.
- 13 Panwar S, Rad P, Jung TP, *et al.* Modeling EEG data distribution with a wasserstein generative adversarial network to predict RSVP events. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020, 28(8): 1720-1730. [doi: 10.1109/TNSRE.2020.3006180]
- 14 Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. *arXiv*: 1701.07875v3, 2017.
- 15 Shi J, Zhang W, Wang YL. Shape analysis with hyperbolic wasserstein distance. *Proceedings of 2016 Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 5051-5061.
- 16 Gulrajani I, Ahmed F, Arjovsky M, *et al.* Improved training of wasserstein GANs. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook: MIT Press, 2017. 5767-5777.
- 17 Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 1994, 5(2): 157-166. [doi: 10.1109/72.279181]
- 18 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735-1780. [doi: 10.1162/neco.1997.9.8.1735]
- 19 Greff K, Srivastava RK, Koutník J, *et al.* LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2222-2232. [doi: 10.1109/TNNLS.2016.2582924]

- 20 Phadikar S, Sinha N, Ghosh R. Automatic EEG eye-blink artefact identification and removal technique using independent component analysis in combination with support vector machines and denoising autoencoder. *IET Signal Processing*, 2020, 14(6): 396–405. [doi: [10.1049/iet-spr.2020.0025](https://doi.org/10.1049/iet-spr.2020.0025)]
- 21 高军峰, 郑崇勋, 王沛. 基于独立成分分析和流形学习的眼电伪差去除. *西安交通大学学报*, 2010, 44(2): 113–118.
- 22 王爽. 基于独立成分分析的碳价格分析及预测 [硕士学位论文]. 西安: 陕西师范大学, 2019.
- 23 Luo P, Wang XJ, Shao WQ, *et al.* Towards understanding regularization in batch normalization. *Proceedings of 7th International Conference on Learning Representations*. New Orleans: OpenReview.net, 2019. 1–8.
- 24 Pascual S, Bonafonte A, Serr J. SEGAN: Speech enhancement generative adversarial network. *Proceedings of Interspeech 2017, 18th Annual Conference of the International Speech Communication Association*. Stockholm: ISCA, 2017. 3642–3646.
- 25 Donahue C, McAuley J, Puckette M. Adversarial audio synthesis. *Proceedings of 7th International Conference on Learning Representations*. New Orleans: OpenReview.net, 2019. 1–8.
- 26 Bousmalis K, Silberman N, Dohan D, *et al.* Unsupervised pixel-level domain adaptation with generative adversarial networks. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017. 95–104.
- 27 Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Proceedings of 26th Annual Conference on Neural Information Processing Systems*. Lake Tahoe: MIT Press, 2012. 1097–1105.
- 28 Sutskever I, Jozefowicz R, Gregor K, *et al.* Towards principled unsupervised learning. *arXiv: 1511.06440v1*, 2015.
- 29 Salimans T, Goodfellow IJ, Zaremba W, *et al.* Improved techniques for training GANs. *Proceedings of 30th International Conference on Neural Information Processing Systems*. Red Hook: Curran Associates Inc., 2016. 2234–2242.
- 30 Korczowski L, Cederhout M, Andreev A, *et al.* Brain Invaders calibration-less P300-based BCI with modulation of flash duration Dataset (bi2015a). GIPSA-lab. 2019.
- 31 Lee CY, Batra T, Baig MH, *et al.* Sliced wasserstein discrepancy for unsupervised domain adaptation. *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2020. 10277–10287.
- 32 Shmelkov K, Schmid C, Alahari K. How good is my GAN? *Proceedings of 15th European Conference on Computer Vision*. Munich: Springer, 2018. 218–234.
- 33 Lawhern VJ, Solon AJ, Waytowich NR, *et al.* EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 2018, 15(5): 056013. [doi: [10.1088/1741-2552/aace8c](https://doi.org/10.1088/1741-2552/aace8c)]
- 34 Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs). *Proceedings of 4th International Conference on Learning Representations (ICLR)*. San Juan: ICLR, 2016. 1–8.

(校对责编: 牛欣悦)