



高阶Takagi-Sugeno-Kang模糊知识蒸馏分类器及其在脑电信号分类中的应用

蒋云良, 印泽宗, 张雄涛, 申情, 李华

引用本文:

蒋云良, 印泽宗, 张雄涛, 等. 高阶Takagi-Sugeno-Kang模糊知识蒸馏分类器及其在脑电信号分类中的应用[J]. [智能系统学报](#), 2024, 19(6): 1419-1427.

JIANG Yunliang, YIN Zezong, ZHANG Xiongtao, et al. TSK fuzzy distillation classifier with negative Euclidean probability and High-order fuzzy dark knowledge transfer and its application on EEG signals classification[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(6): 1419-1427.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202307029>

您可能感兴趣的其他文章

[任务间共享和特有结构分解的多任务TSK模糊系统建模](#)

Multi-task TSK fuzzy system modeling based on inter-task common and special structure decomposition
智能系统学报. 2021, 16(4): 622-629 <https://dx.doi.org/10.11992/tis.202007009>

[深度自编码与自更新稀疏组合的异常事件检测算法](#)

Abnormal event detection method based on deep auto-encoder and self-updating sparse combination
智能系统学报. 2020, 15(6): 1197-1203 <https://dx.doi.org/10.11992/tis.202007003>

[一种基于模糊划分和模糊加权的集成深度信念网络](#)

Ensemble deep belief network based on fuzzy partitioning and fuzzy weighting
智能系统学报. 2019, 14(5): 905-914 <https://dx.doi.org/10.11992/tis.201809018>

[基于改进卷积神经网络的多标记分类算法](#)

A multi-label classification algorithm based on an improved convolutional neural network
智能系统学报. 2019, 14(3): 566-574 <https://dx.doi.org/10.11992/tis.201804056>

[基于模糊超网络的知识获取方法研究](#)

Fuzzy hypernetwork-based knowledge acquisition method
智能系统学报. 2019, 14(3): 479-490 <https://dx.doi.org/10.11992/tis.201804055>

[多层递阶融合模糊特征映射的模糊C均值聚类算法](#)

Fuzzy C-means clustering algorithm for multilayered hierarchical fusion fuzzy feature mapping
智能系统学报. 2018, 13(4): 594-601 <https://dx.doi.org/10.11992/tis.201703047>

DOI: 10.11992/tis.202307029

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20240910.1054.002>

高阶 Takagi-Sugeno-Kang 模糊知识蒸馏分类器及其在脑电信号分类中的应用

蒋云良^{1,2,3}, 印泽宗^{1,2}, 张雄涛^{1,2}, 申情^{1,2}, 李华^{2,4}

(1. 湖州师范学院 信息工程学院, 浙江 湖州 313000; 2. 浙江省现代农业资源智慧管理与应用研究重点实验室, 浙江 湖州 313000; 3. 浙江师范大学 计算机科学与技术学院, 浙江 金华 321004; 4. 浙江师范大学 数理医学院, 浙江 金华 321004)

摘要: 在脑电信号 (electro encephalo gram, EEG) 的分类检测任务中, 低阶 TSK (Takagi-Sugeno-Kang) 模糊分类器的学习速度较快, 但性能表现不理想, 高阶 TSK 模糊分类器虽然具有较强的性能优势, 但极其复杂的模糊规则后件严重影响模型的运行速度。为此, 提出一种基于负欧氏概率和高阶模糊隐藏知识迁移的新型 TSK 模糊蒸馏分类器 (solved TSK-least learning machine-knowledge distillation classifier, STSK-LLM-KD)。首先, 利用所提出的基于知识蒸馏的最小学习机 (LLM-KD) 对教师模型的后件参数进行快速求解并得到相应的负欧氏概率用于生成软标签; 然后, 通过计算软标签之间的 Kullback-Leibler 散度提取教师模型的高阶模糊隐藏知识并迁移至低阶学生模型中, 使模型性能优于高阶 TSK 模糊分类器的同时保持更快的训练速度。在运动想象脑电数据集和新德里 Hauz Khas 癫痫脑电数据集上的实验结果充分验证了 STSK-LLM-KD 的优势, STSK-LLM-KD 相较于其他模糊分类器表现更加优异, 与深度知识蒸馏模型相比, STSK-LLM-KD 能够更好地提升学生模型的性能。

关键词: TSK 模糊分类器; 知识蒸馏; 高阶模糊隐藏知识; 脑电信号; 最小学习机; 癫痫; 运动想象; 模糊系统

中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2024)06-1419-09

中文引用格式: 蒋云良, 印泽宗, 张雄涛, 等. 高阶 Takagi-Sugeno-Kang 模糊知识蒸馏分类器及其在脑电信号分类中的应用 [J]. 智能系统学报, 2024, 19(6): 1419-1427.

英文引用格式: JIANG Yunliang, YIN Zezong, ZHANG Xiongtao, et al. TSK fuzzy distillation classifier with negative Euclidean probability and High-order fuzzy dark knowledge transfer and its application on EEG signals classification[J]. CAAI transactions on intelligent systems, 2024, 19(6): 1419-1427.

TSK fuzzy distillation classifier with negative Euclidean probability and High-order fuzzy dark knowledge transfer and its application on EEG signals classification

JIANG Yunliang^{1,2,3}, YIN Zezong^{1,2}, ZHANG Xiongtao^{1,2}, SHEN Qing^{1,2}, LI Hua^{2,4}

(1. School of Information Engineering, Huzhou University, Huzhou 313000, China; 2. Zhejiang Province Key Laboratory of Smart Management and Application of Modern Agricultural Resources, Huzhou 313000, China; 3. School of Computer Science and Technology, Zhejiang Normal University, Jinhua 321004, China; 4. College of Mathematical Medicine, Zhejiang Normal University, Jinhua 321004, China)

Abstract: In the classification and detection task of electroencephalogram (EEG) signals, the low-order Takagi-Sugeno-Kang (TSK) fuzzy classifier runs faster but performs poorly, while the high-order TSK fuzzy classifier demonstrates strong prediction performance. However, the extremely complex fuzzy rules in the consequent part notably affect the running speed of the model. Therefore, this study proposes a novel TSK fuzzy distillation classifier, STSK-LLM-KD, based on negative Euclidean probability and high-order fuzzy dark knowledge transfer. First, the least learning machine based on knowledge distillation (LLM-KD) is used to quickly solve the consequent parameters of the teacher model and obtain corresponding negative Euclidean probabilities to generate soft labels. Then, the high-order fuzzy dark knowledge of the teacher model is extracted by calculating the Kullback-Leibler divergence between soft labels and transferred to the low-order student model. This approach enhances the performance of the model beyond that of the high-order TSK fuzzy classifier while maintaining a faster training speed. Experimental results on the motor imagery EEG dataset and Hauz Khas epilepsy EEG dataset in New Delhi fully verify the advantages of the proposed STSK-LLM-KD. Compared to other fuzzy classifiers, STSK-LLM-KD performs better; compared to deep knowledge distillation models, STSK-LLM-KD more effectively improves the performance of the student model.

Keywords: TSK fuzzy classifier; knowledge distillation; high-order fuzzy dark knowledge; electro encephalo gram (EEG); least learning machine; epilepsy; motor imagery; fuzzy system

收稿日期: 2023-07-20. 网络出版日期: 2024-09-10.

基金项目: 国家自然科学基金项目 (U22A20102, 62376094); 浙江省“尖兵”“领雁”研发攻关计划项目 (2023C01150).

通信作者: 张雄涛. E-mail: 1047897965@qq.com.

脑电信号 (Electro encephalo gram, EEG) 是一种通过在头皮上放置多个电极, 记录大脑表面不同位置脑部神经元离子电流产生的电压波动, 从

而反映出不同脑区域的活动情况的图形。EEG被广泛应用于研究大脑活动和脑部相关疾病辅助诊断,如癫痫和睡眠障碍等。同时,还用于脑机接口技术(brain-computer interface, BCI)中,例如识别运动想象、意图和情感等^[1],从而实现大脑与计算机或其他外部设备的交互。EEG信号在记录过程中会受到多种因素干扰,例如头发、肌肉和眼球运动等,因此在处理脑电数据时往往首先需要对EEG信号使用快速傅里叶变换^[2]和小波包分解^[3]等方法进行预处理和特征提取,然后再采用经典的或最新的机器学习算法进行分类或预测。TSK(takagi-sugeno-kang)模糊分类器是最著名的模糊分类器之一^[4-5],作为基于模糊规则的可解释机器学习算法已被成功运用于EEG信号的分类检测中^[6-8],由包含模糊规则的前件和后件组成。前件将输入空间划分为多个模糊区域,其模糊规则通常由一组条件语句进行描述,这些条件语句可以很容易地被人类所理解和解释,后件将输出结果去模糊化为精确的数值^[9]。在各种TSK模糊分类器中,低阶TSK模糊分类器^[10-11]由于其简单的结构和较快的训练速度受到最广泛的关注。然而,低阶TSK模糊分类器的性能表现较差,分类结果难以令人满意。

近年来,越来越多的研究将低阶TSK模糊分类器和具有强大特征提取能力的深度学习模型相结合。Tian等^[12]通过卷积神经网络(convolutional neural network, CNN)提取EEG信号的深层特征,提出了一种融合浅层特征和深层特征的多视角一阶TSK模糊分类器;蒋云良等^[13]将深层特征和类别信息相融合,提出了一种基于增强深层特征的一阶TSK模糊分类器。深层特征的引入提升了低阶TSK模糊分类器的分类性能,但深度学习模型内部各非线性层的参数难以被解释,从而严重影响了TSK模糊分类器的可解释性。

高阶TSK模糊分类器^[14]的后件使用高阶多项式来提升模型的拟合能力,凭借较少的模糊规则获得了更好的分类性能。然而,高阶TSK模糊分类器后件中的高阶多项式的参数过于复杂,导致其运行速度较慢,可解释性低于低阶TSK模糊分类器。

知识蒸馏^[15](knowledge distillation, KD)是目前深度学习领域中最火热的研究方向之一,通过将教师模型(复杂的大模型)的隐藏知识提取并迁移给学生模型(简单的小模型),帮助提升学生模型的性能表现^[16-23]。目前已有研究将知识蒸馏和TSK模糊分类器进行结合^[24],Gu等^[25]将隐藏

知识从CNN转移到TSK模糊分类器中,解释TSK模糊分类器是如何做出决策的;Erdem等^[26]通过CNN来蒸馏区间2型模糊分类器,提升模型在大数据集上的分类性能。结合知识蒸馏,本研究构建了一种具有强大分类性能的新型TSK模糊蒸馏分类器STSK-LLM-KD,用于EEG脑电信号的分类检测,通过从作为教师模型的高阶TSK模糊分类器中提取高阶模糊隐藏知识,并迁移给作为学生模型的低阶TSK模糊分类器,得到了一种性能表现更加强大的模糊蒸馏分类器,以此解决低阶TSK模糊分类器在EEG信号分类检测过程中性能表现较差的问题。在STSK-LLM-KD中,基于最小学习机^[27]的知识蒸馏算法(least learning machine-knowledge distillation, LLM-KD)被用于快速求解教师模型的后件,通过计算教师模型的直接输出与每个类别标签之间的负欧氏概率,得到教师模型的概率输出。此外,STSK-LLM-KD将概率输出通过带有温度参数的softmax函数,得到教师模型和学生模型的软标签,计算两者软标签之间的KL散度(kullback-leibler divergence),从而提取教师模型中的高阶模糊隐藏知识并迁移给学生模型。最后,通过梯度下降算法对包含高阶模糊隐藏知识的STSK-LLM-KD进行求解。

本研究的主要贡献如下。

1) 提出了一种结合知识蒸馏的新型TSK模糊蒸馏分类器STSK-LLM-KD,通过LLM-KD提取教师模型的高阶模糊隐藏知识并迁移给学生模型的后件。与常用于求解后件的梯度下降算法相比,LLM-KD中的最小学习机计算速度更快,更适用于求解复杂的教师模型。

2) 在后件的去模糊化过程中,最小学习机的输出是一个精确值,LLM-KD通过计算输出值和各类别标签的负欧氏距离,将单个的精确值转化为一组负欧氏概率,包含更多类别信息,更有利于高阶模糊隐藏知识的提取和迁移。

3) 在运动想象脑电数据集和新德里HauzKhas癫痫脑电数据集上的实验结果证明了STSK-LLM-KD的优势。STSK-LLM-KD在所有数据集上都取得了最好的性能表现,具有比高阶TSK模糊分类器更加强大的泛化能力和更快的学习速度。

1 模型介绍

1.1 STSK-LLM-KD模型结构

本研究提出了一种新型TSK模糊蒸馏分类器STSK-LLM-KD,其模型结构如图1所示。

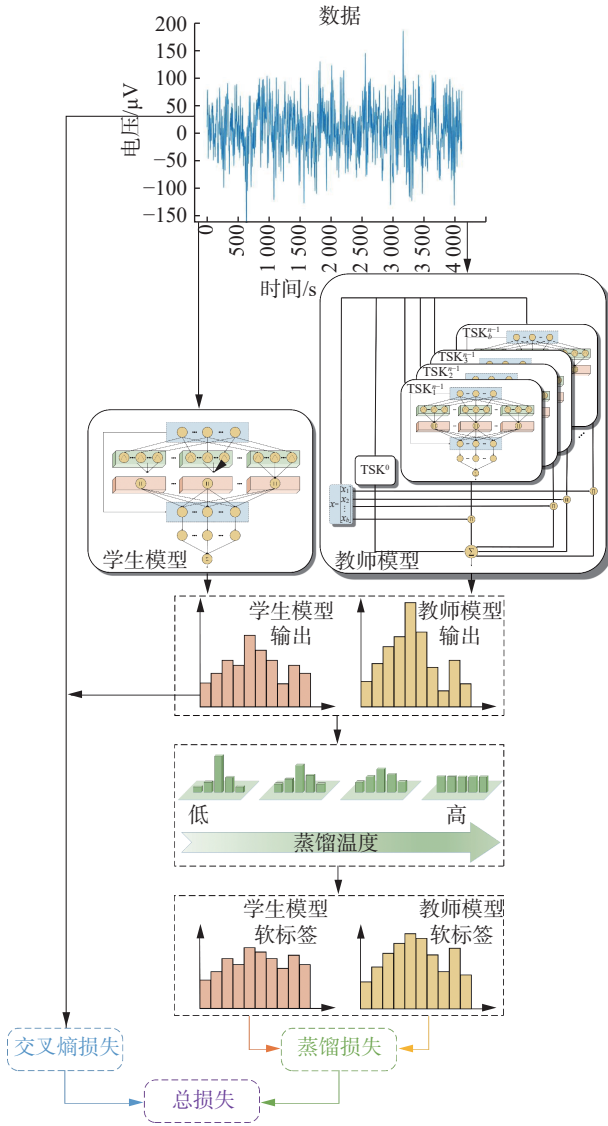


图 1 STSK-LLM-KD 模型结构

Fig. 1 Structure of STSK-LLM-KD

1.1.1 TSK 模糊分类器

TSK 模糊分类器可以通过模糊 IF-THEN 规则来进行描述, 该规则表示模糊分类器的输入-输出关系。对于 TSK 模糊分类器而言, 模糊规则可以表示为

$$\begin{aligned} \text{If } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \wedge \cdots \wedge x_b \text{ is } A_b^k, \\ \text{Then } y^k = f^k(\mathbf{x}) \end{aligned} \quad (1)$$

式中: $k = 1, 2, \dots, K$, K 为规则系统中模糊规则的总数; A_i^k 为 x_i 的第 k 条规则的模糊集; y^k 为第 k 条规则的输出; $f^k(\mathbf{x})$ 为模糊隶属程度和输入变量的多项式函数; \wedge 为模糊合取操作。

对于零阶 TSK 模糊分类器而言

$$f^k(\mathbf{x}) = P_0^k \quad (2)$$

式中 P_0^k 为第 k 条模糊规则隶属程度的系数。

对于一阶 TSK 模糊分类器而言

$$f^k(\mathbf{x}) = P_0^k + x_1^k P_1^k + x_2^k P_2^k + \cdots + x_b^k P_b^k \quad (3)$$

式中: P_0^k 为 $f^k(\mathbf{x})$ 的常数项; P_i^k 为多项式中第 k 条规则上 x_i 的系数, $i = 1, 2, \dots, b$ 。

高阶 TSK 模糊分类器将原始输入映射到更高维度的特征空间, 从而获得更加强大的拟合能力。对于高阶 TSK 模糊分类器而言

$$f^k(\mathbf{x}) = \sum_{\substack{(j_1 + j_2 + \cdots + j_i + \cdots + j_b) \leq n \\ j_1, j_2, \dots, j_i, \dots, j_b \geq 0}} (a_{j_1, j_2, \dots, j_i, \dots, j_b}^k) x_1^{j_1} \cdots x_i^{j_i} \cdots x_b^{j_b} \quad (4)$$

式中: $k = 1, 2, \dots, K$, K 为规则系统中模糊规则的总数; n 为高阶 TSK 模糊分类器的最高次多项式的阶次; j_i 为 x_i 的阶次; $a_{j_1, j_2, \dots, j_i, \dots, j_b}^k$ 为构成第 k 个规则的线性组合中包含 b 个自变量的最高次多项式的系数。

1.1.2 教师模型和学生模型的具体结构

STSK-LLM-KD 采用由多个零阶 TSK 模糊分类器堆叠而成的二阶 TSK 模糊分类器作为教师模型, 即其后件可表示为多个低阶 TSK 模糊分类器的后件的叠加^[28]:

$$\begin{aligned} y_n = y_0 + x_1(y_0^{(1)} + x_1 y_{n-2}^{(1,1)} + \cdots + x_b y_{n-2}^{(1,b)}) + \\ x_2(y_0^{(2)} + x_1 y_{n-2}^{(2,1)} + \cdots + x_b y_{n-2}^{(2,b)}) + \cdots + \\ x_b(y_0^{(b)} + x_1 y_{n-2}^{(b,1)} + \cdots + x_b y_{n-2}^{(b,b)}) \end{aligned} \quad (5)$$

式中: y_n 为 n 阶 TSK 模糊分类器的输出, y_0 为零阶 TSK 模糊分类器的输出。

式 (5) 通常使用加权平均来进行计算:

$$y = \sum_{k=1}^K \frac{\mu^k(\mathbf{x})}{\sum_{k'=1}^K \mu^{k'}(\mathbf{x})} y^k = \sum_{k=1}^K \tilde{\mu}^k(\mathbf{x}) y^k \quad (6)$$

式中: $\mu^k(\mathbf{x})$ 为第 k 条规则的模糊隶属程度, 规范化后得到 $\tilde{\mu}^k(\mathbf{x})$ 。

$$\mu^k(\mathbf{x}) = \prod_{i=1}^b \mu_{A_i^k}(x_i) \quad (7)$$

式中: $\mu_{A_i^k}(x_i)$ 是 x_i 在模糊集 A_i^k 上的隶属程度, 高斯函数被广泛地应用于计算模糊隶属程度:

$$\mu_{A_i^k}(x_i) = \exp\left(-\frac{(x_i - v_i^k)^2}{2\delta_i^k}\right) \quad (8)$$

式中: v_i^k 为中心参数, 是每个模糊规则的中心; v_i^k 从 $\{0, 0.25, 0.50, 0.75, 1.00\}$ 中随机选择, 可以被自然语言解释为 $\{\text{非常低、低、中、高、非常高}\}$, 从而确保 STSK-LLM-KD 的前件是可解释的; δ_i^k 为核宽, 通常被设定为一个正值。 v_i^k 和 δ_i^k 叫作前件参数, 定义了模糊集 A_i^k 的具体形式。

教师模型和学生模型的前件的计算公式为

$$\mathbf{x}_r = (1, \mathbf{x}^T)^T \in R^{K(b+1)} \quad (9)$$

$$\mathbf{x}_w = (1, \mathbf{x}_i \mathbf{x}_r^T)^T \in R^{K((b+1)b+1)} \quad (10)$$

$$\tilde{\mathbf{x}}_m^k = \tilde{\mu}^k(\mathbf{x}) \mathbf{x}_w \in R^{K((b+1)b+1)} \quad (11)$$

$$\mathbf{x}_m = [(\tilde{\mathbf{x}}_m^1)^T, (\tilde{\mathbf{x}}_m^2)^T, \dots, (\tilde{\mathbf{x}}_m^K)^T]^T \in R^{K((b+1)b+1)} \quad (12)$$

$$\tilde{\mathbf{x}}_h^k = \tilde{\mu}^k(\mathbf{x}) \mathbf{x}_r \in R^{K(b+1)} \quad (13)$$

$$\mathbf{x}_h = [(\tilde{\mathbf{x}}_h^1)^T, (\tilde{\mathbf{x}}_h^2)^T, \dots, (\tilde{\mathbf{x}}_h^K)^T]^T \in R^{K(b+1)} \quad (14)$$

本研究分别用 \mathcal{M} 和 \mathcal{H} 来表示教师模型和学生模型。STSK-LLM-KD 的教师模型是二阶 TSK 模糊分类器, 其性能表现优秀, 但是需要花费大量时间来训练, 因此本研究使用最小学习机来快速求解教师模型的后件。

$$\mathbf{y}^M = \mathbf{q}_m^T \mathbf{x}_m$$

$$\mathbf{q}_m = ((1/L)\mathbf{I} + \mathbf{X}_m^T \mathbf{X}_m)^{-1} \mathbf{X}_m^T \bar{\mathbf{Y}}$$

式中: \mathbf{y}^M 为教师模型的直接输出, \mathbf{q}_m 为教师模型的后件参数, L 为正则化参数, \mathbf{I} 为 $N \times N$ 的单位矩阵, N 是样本数, $\bar{\mathbf{Y}} = [\bar{Y}_1 \ \bar{Y}_2 \ \dots \ \bar{Y}_N]^T$ 为真实标签, $\mathbf{X}_m = [\mathbf{x}_m^1 \ \mathbf{x}_m^2 \ \dots \ \mathbf{x}_m^N]^T$ 。

STSK-LLM-KD 的学生模型是一阶 TSK 模糊分类器, 其运行速度较快, 但与高阶 TSK 模糊分类器相比, 性能表现较弱。本研究通过最小化交叉熵误差的梯度下降算法对学生模型的后件参数进行更新。

$$\mathbf{z}^H = \mathbf{Q}_h^T \mathbf{x}_h$$

$$H = - \sum_{i=1}^N \sum_{t=1}^C \bar{Y}_{i,t} \log(z_{i,t}^H) \quad (15)$$

$$\mathbf{Q}_h(d+1) = \mathbf{Q}_h(d) - \eta \frac{\partial H}{\partial \mathbf{Q}_h(d)} \quad (16)$$

式中: \mathbf{z}^H 为学生模型的概率输出; \mathbf{Q}_h 为学生模型的后件参数; H 为交叉熵损失; η 为给定的学习率。

1.1.3 基于最小学习机的知识蒸馏算法 (LLM-KD)

LLM-KD 通过计算教师模型的直接输出与每个类别标签之间的负欧氏距离^[29], 得到教师模型预测样本隶属于各个类别的概率, 称为负欧氏概率, 作为教师模型的概率输出:

$$z_i^M = -\sqrt{(\mathbf{y}^M - \hat{\mathbf{y}}_i)^2}$$

式中: $\hat{\mathbf{y}}_i$ 为第 i 类的标签; z_i^M 是教师模型的概率输出, $\mathbf{z}^M = [z_1^M \ z_2^M \ \dots \ z_i^M \ \dots \ z_C^M] \in \mathbf{R}^{1 \times C}$; C 是样本的类别数。

对于给定数据 \mathbf{x} , 软标签可以表示为 $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_i \ \dots \ u_C] \in \mathbf{R}^{1 \times C}$, 其中 u_i 是第 i 类别的软标签, \mathbf{u} 中的每个元素都可以通过带有温度参数 τ 的 softmax 函数计算得到:

$$u_i = \frac{\exp(z_i/\tau)}{\sum_{j=1}^C \exp(z_j/\tau)} \quad (17)$$

式中 z_i 为第 i 类别的概率输出。

STSK-LLM-KD 使用 KL 散度 (记为 K_L) 来迁移隐藏知识:

$$K_D = K_L(\mathbf{u}^M \parallel \mathbf{u}^H) = \sum_{i=1}^C u_i^M \log\left(\frac{u_i^M}{u_i^H}\right) \quad (18)$$

式中 K_D 为蒸馏损失。

结合式 (18), 可以得到 STSK-LLM-KD 的总损失函数 (记为 L_{oss}):

$$L_{\text{oss}} = \alpha K_L(\mathbf{u}^M \parallel \mathbf{u}^H) + (1 - \alpha) \cdot H$$

式中 α 为蒸馏损失所占的权重。

1.2 STSK-LLM-KD 学习算法

STSK-LLM-KD 学习算法包括教师模型和学生模型的构建过程以及知识蒸馏过程。

算法 1 构建教师模型和学生模型

输入 数据集 $\mathbf{X} = \{\mathbf{x}_i, \mathbf{x}_i \in \mathbf{R}^b, i = 1, 2, \dots, N\}$, 真实标签 $\bar{\mathbf{Y}} = \{\bar{Y}_i, \bar{Y}_i \in \mathbf{R}, i = 1, 2, \dots, N\}$, 模糊规则 K , 正则化参数 L , 最大迭代次数 θ , 阈值参数 ξ , 学习率 η 。

输出 教师模型直接输出和学生模型概率输出。

从 5 个固定的模糊划分 $\{0, 0.25, 0.50, 0.75, 1.00\}$ 中随机选择高斯函数的中心 v_i^k , 将核宽 δ_i^k 设置为正值, 通过式 (6)~(8) 计算归一化的模糊隶属程度;

通过式 (9)~(14) 计算教师模型和学生模型的前件参数矩阵;

通过 $\mathbf{q}_m = ((1/L)\mathbf{I} + \mathbf{X}_m^T \mathbf{X}_m)^{-1} \mathbf{X}_m^T \bar{\mathbf{Y}}$ 计算教师模型的后件参数 \mathbf{q}_m ;

学生模型的后件参数 \mathbf{Q}_h 可以通过梯度下降算法进行更新:

初始化后件参数 \mathbf{Q}_h 并设定 $d = 1$;

Repeat

通过式 (18)~(19) 计算 $\mathbf{Q}_h(d+1)$;

$d = d + 1$;

Until $H(d) - H(d-1) \leq \xi$ or $d \geq \theta$

计算教师模型直接输出 $\mathbf{y}^M = \mathbf{q}_m^T \mathbf{x}_m$ 和学生模型的概率输出 $\mathbf{z}^H = \mathbf{Q}_h^T \mathbf{x}_h$ 。

算法 2 STSK-LLM-KD

输入 数据集 $\mathbf{X} = \{\mathbf{x}_i, \mathbf{x}_i \in \mathbf{R}^b, i = 1, 2, \dots, N\}$, 真实标签 $\bar{\mathbf{Y}} = \{\bar{Y}_i, \bar{Y}_i \in \mathbf{R}, i = 1, 2, \dots, N\}$, 教师模型直接输出 $\mathbf{y}^M = \{y_i^M, y_i^M \in \mathbf{R}, i = 1, 2, \dots, N\}$, 学生模型的概率输出 $\mathbf{z}^H = \{z_i^H, z_i^H \in \mathbf{R}, i = 1, 2, \dots, N\}$, 最大迭代次数 θ , 阈值参数 ξ , 学习率 η , 蒸馏参数 τ, α 。

输出 STSK-LLM-KD 的输出。

通过式 (20) 计算教师模型的直接输出 \mathbf{y}^M 与各类别标签之间的负欧氏概率,得到教师模型的概率输出 z^M ;

利用 softmax 函数,通过式 (21) 计算得到教师模型和学生模型的软标签 \mathbf{u}^M 和 \mathbf{u}^H ;

STSK-LLM-KD 的后件参数 \mathbf{Q}_h 可以通过梯度下降算法进行更新:

初始化后件参数 \mathbf{Q}_h 并设定 $d = 1$;

Repeat

$$\mathbf{Q}_h(d+1) = \mathbf{Q}_h(d) - \eta \frac{\partial L_{\text{oss}}}{\partial \mathbf{Q}_h(d)};$$

$d = d + 1$;

Until $L_{\text{oss}}(d) - L_{\text{oss}}(d-1) \leq \xi$ or $d \geq \theta$

计算 STSK-LLM-KD 的输出。

2 实验研究

本研究选择 7 个模糊分类器和 1 个深度知识蒸馏模型与 STSK-LLM-KD 进行对比实验。实验中,描述了运动想象脑电数据集^[30]和新德里 HauzKhas 癫痫脑电数据集^[31]的详细信息;描述了实验的各项设置;详细报告了 STSK-LLM-KD 在各项数据集上的实验结果和分析;讨论了模糊知识蒸馏的参数敏感性。实验运行的硬件环境为: Intel(R) Core(TM) i5-1135G7 at 2.4 GHz and 32 GB RAM with Microsoft Windows 10, 编程环境为: Python 3.6.2 with Torch 1.10.0.

2.1 数据集

本研究所用的运动想象脑电数据集 (Dataset III, BCI competition II) 是 BCI Competition 中广泛使用的一个数据集。在该数据集中,一位正常受试者通过左手或右手的想象动作来控制一个反馈杆,其脑电信号在反馈过程中被记录和整理,总共包含 3 个 EEG 通道,采样频率为 128 Hz,数据集实验次数为 280 次。根据 Wu 等^[32]的研究,本研究采用巴特沃斯滤波器和共空间模式算法 (common spatial pattern, CSP) 对数据进行预处理。

新德里 HauzKhas 癫痫脑电数据集使用 10~20 电极放置系统收集 10 名癫痫患者的 EEG 信号用于判断癫痫是否发作。数据集可分为癫痫发作前 (Pre-ictal)、癫痫发作间歇 (Interictal) 和癫痫发作中 (Ictal), 每个类别包含 256 条数据,采样频率为 200 Hz。根据 Zhang 等^[3]的研究,本次研究使用小波包变换提取时频域特征,采用 Scikit-learn 中的 MinMaxScaler 模块对所有采用的数据集进行规范化处理。数据集见表 1。

表 1 数据集

Table 1 Introduction of dataset

数据集	特征数	样本数
BCI	3	280
Ictal vs Interictal (IvI)	14	512
Ictal vs Preictal (IvP)	14	512

2.2 实验设置和性能指标

STSK-LLM-KD 是从二阶 TSK 模糊分类器提取高阶模糊隐藏知识的新型 TSK 模糊蒸馏分类器。因此,本研究将 n 阶 TSK 模糊分类器 ($n = 0, 1, 2$) 和基于最小二乘支持向量机的模糊系统 LSSVFS^[14] 加入对比模型中。TSKn v1 (n 为阶次, $n = 0, 1, 2$) 使用 LLM 求解后件参数; TSKn v2 (n 为阶次, $n = 0, 1, 2$) 采用梯度下降法更新后件参数; LSSVFS² 使用含有二阶多项式的最小二乘支持向量机 (LSSVM) 来求解后件参数。与此同时,将深度知识蒸馏模型 CNN-TSK-KD 加入到对比模型中,其将 CNN 作为教师模型,一阶 TSK 模糊分类器作为学生模型,与 STSK-LLM-KD 在准确率和加权 F_1 分数方面进行对比。对比模型见表 2, CNN 的具体结构见表 3。

表 2 采用的模糊方法

Table 2 Adopted fuzzy methods

模糊分类器	前件参数	后件参数	阶次
TSK0 v1			0
TSK1 v1		最小学习机	1
TSK2 v1			2
TSK0 v2	等间距划分		0
TSK1 v2		梯度下降	1
TSK2 v2			2
LSSVFS ²		LSSVM	2

表 3 CNN 的具体结构

Table 3 Specific structure of CNN

网络层	特征图	核尺寸(步长)
输入层	—	—
卷积层	16	1×3(1)
池化层	16	1×2(1)
卷积层	8	1×3(1)
池化层	8	1×2(1)
卷积层	1	1×3(1)
全连接层	—	—
全连接层	—	—
输出层	—	—

所有模型均在所采用的数据集上进行五折交叉验证,所有可调参数均采用网格搜索策略进行选择。模糊规则 K 的搜索范围为 $\{1,2,\dots,20\}$;正则化参数 L 设置为100;蒸馏参数 τ 的搜索范围为 $\{1,5,10,20,100\}$;蒸馏参数 α 的搜索范围为 $\{0,0.25,0.50,0.75,1.00\}$;最大迭代次数 θ 设置为30;阈值参数 ξ 设置为 10^{-5} ;学习率 η 设置为0.01;其他参数设置为默认值^[14]。

为了评价所用模型的性能,实验中采用了3个常用的性能指标,即准确率、加权 F_1 分数和平均模糊规则数。准确率是最直观的性能指标,显示了正确预测的样本与总样本的比率;加权 F_1 分数是精确率和召回率的加权平均值,综合考量了

假阳性和假阴性。实验的最佳结果用粗体标记。

2.3 结果与分析

表4和表5给出了STSK-LLM-KD和所采用的7个模糊分类器的对比实验结果,得出以下结论。

1) 在准确率和加权 F_1 分数方面,STSK-LLM-KD在所有数据集中都取得了最好的性能表现。值得注意的是,与高阶TSK模糊分类器(TSK2 v1、TSK2 v2和LSSVFS²)相比,STSK-LLM-KD拥有更好的性能表现。本研究认为,知识蒸馏通过计算教师模型的负欧氏概率和软标签帮助STSK-LLM-KD学习到了来自教师模型的高阶模糊隐藏知识,从而有效地提升了模型的鲁棒性,进一步提升了模型的性能表现。

表4 STSK-LLM-KD与7个模糊分类器在各数据集的准确率和加权 F_1 分数(Acc/W-F)

Table 4 Accuracy and weighted F_1 score of STSK-LLM-KD and seven fuzzy classifiers on datasets

%

数据集	TSK0 v1	TSK1 v1	TSK2 v1	TSK0 v2	TSK1 v2	TSK2 v2	LSSVFS ²	STSK-LLM-KD
BCI	82.71/82.70	83.85/83.84	83.92/83.97	81.50/81.45	83.62/83.21	84.92/84.69	84.28/84.32	85.42/85.09
IvI	95.86/95.85	96.37/96.36	95.91/95.89	96.48/96.48	97.65/97.65	97.97/97.96	96.09/96.08	98.20/98.05
IvP	92.81/92.81	93.24/93.24	92.94/92.93	93.20/93.21	93.67/93.59	93.63/93.54	92.97/92.97	94.30/94.22

注:最优值加黑显示。

表5 STSK-LLM-KD在各数据集的平均规则数

Table 5 Average number of rules of STSK-LLM-KD on datasets

数据集	TSK0 v1	TSK1 v1	TSK0 v2	TSK1 v2	STSK-LLM-KD
BCI	18.6	9.6	16.2	12.0	9.4
IvI	14.6	8.8	15.6	9.4	7.2
IvP	17.6	14.2	16.2	13.6	13

注:最优值加黑显示。

2) 在平均模糊规则数上,与作为学生模型的TSK1 v2相比,STSK-LLM-KD凭借更少的模糊规则取得了更加优异的性能表现,这说明来自教师模型的高阶模糊隐藏知识提升了STSK-LLM-KD的泛化能力,从而通过更少的模糊规则获得了更强的性能表现。

表6给出了STSK-LLM-KD和CNN-TSK-KD与相应的学生模型在准确率和加权 F_1 分数上的对比实验结果,得出以下结论。

表6 STSK-LLM-KD与CNN-TSK-KD在各数据集上的准确率和加权 F_1 分数(Acc/W-F)

Table 6 Accuracy and weighted F_1 score of STSK-LLM-KD and CNN-TSK-KD on datasets

%

数据集	CNN-TSK-KD		STSK-LLM-KD	
	学生模型	蒸馏模型	学生模型	蒸馏模型
BCI	83.22/83.08	84.98/84.91	83.50/83.11	85.42/85.09
IvI	97.46/97.25	98.04/97.88	97.56/97.30	98.20/98.05
IvP	93.17/93.08	94.10/94.01	93.33/93.21	94.30/94.22
平均值	91.28/91.14	92.37/92.27	91.46/91.21	92.64/92.45

注:模型最优值加黑显示。

1) STSK-LLM-KD获得了最强的性能提升,平均准确率和平均加权 F_1 分数提升百分比分别为1.18%和1.24%,说明知识蒸馏可以通过从高阶TSK模糊分类器中提取高阶模糊隐藏知识来有效提升低阶TSK模糊分类器的性能。

2) 与CNN-TSK-KD相比,STSK-LLM-KD在

准确率和加权 F_1 分数上表现更强,这说明相比于CNN,高阶TSK模糊分类器所提取的高阶模糊隐藏知识更有利于提升低阶TSK模糊分类器的性能。

对各数据集的准确率实验结果进行弗里德曼排名(Friedman ranking)统计计算,进一步观察

STSK-LLM-KD 与其他模型的分分类准确率的差别。由图 2 可知, STSK-LLM-KD 取得最优等级, 说明 STSK-LLM-KD 与其他模型在统计学上存在显著差异。

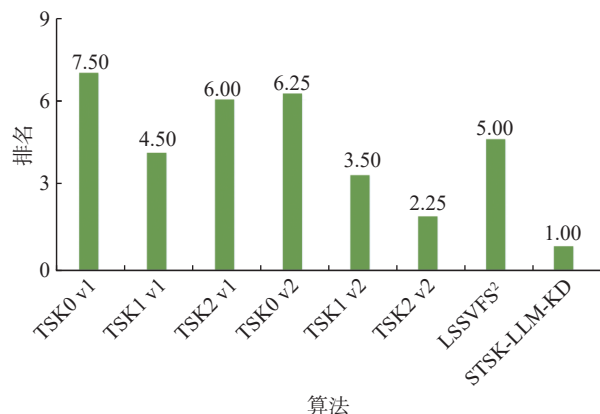


图 2 各模型排名值

Fig. 2 Ranking value of each model

2.4 STSK-LLM-KD 的参数敏感性

图 3 和图 4 给出了 STSK-LLM-KD 在选取不同蒸馏参数时对准确率产生的影响。由图 3 和图 4 可以清楚地看到, 随着参数的增大, 分类性能先提高后下降。蒸馏温度 τ 的选取范围为 $\{1, 5, 10, 20, 100\}$, 其对模型输出结果的柔软度有着重要影响。由图 3 可知, 在 $\{10, 20\}$ 范围内的蒸馏温度 τ 可能是一个更好的选择, 这说明较低的温度不能有效地蒸馏出类别之间的相似信息, 而较高的温度则会破坏模型对不同类别的预测, 所以中间的温度是最合适的。

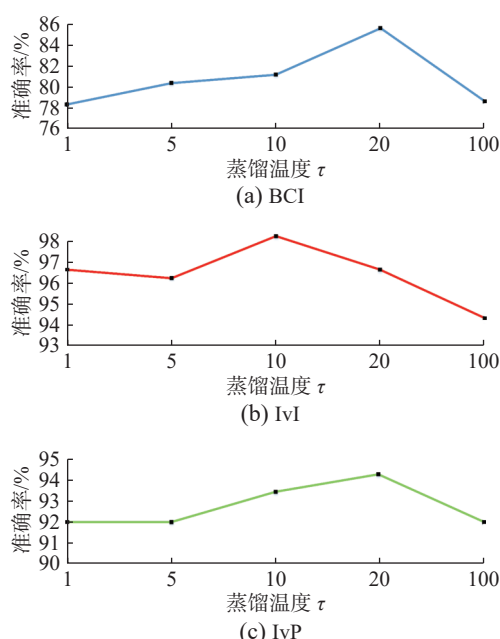


图 3 蒸馏温度 τ 在 BCI、IvI 和 IvP 中对准确率的影响

Fig. 3 Effectiveness of distillation temperature τ on Accuracy of BCI, IvI and IvP

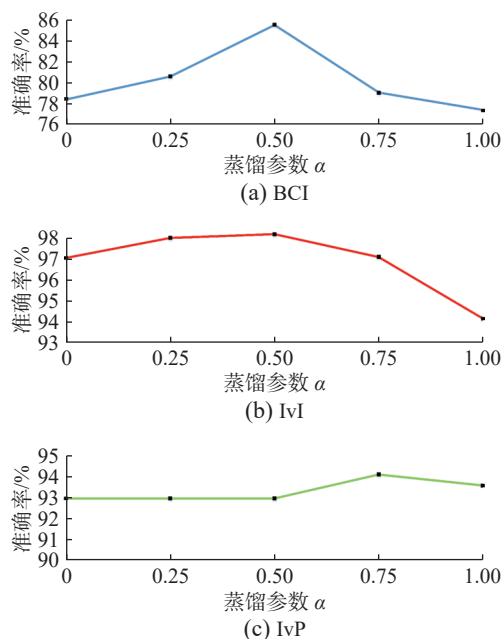


图 4 蒸馏参数 α 在 BCI、IvI 和 IvP 中对准确率的影响

Fig. 4 Effectiveness of distillation parameter α on Accuracy of BCI, IvI and IvP

蒸馏参数 α 的选取范围为 $\{0, 0.25, 0.50, 0.75, 1.00\}$, 表示模糊隐藏知识和包含在真实标签中的显性知识比例。由图 4 可知, α 在 $\{0.25, 0.75\}$ 内可能是一个更好的选择, 这说明适当的模糊隐藏知识可以提高分类性能。当模糊隐藏知识传递过少时, 教师模型不能有效地引导学生模型, 当模糊隐藏知识传递过多时, 学生模型会被教师模型所犯的错误所误导。

3 结束语

本研究主要关注如何将知识蒸馏和 TSK 模糊分类器相结合, 从而提升 TSK 模糊分类器在 EEG 脑电信号分类检测任务中的性能表现。本研究提出了一种结合知识蒸馏的新型 TSK 模糊蒸馏分类器 STSK-LLM-KD, 通过计算作为教师模型的高阶 TSK 模糊分类器的负欧氏概率得到相应的软标签, 从而实现高阶模糊隐藏知识的迁移。在运动想象脑电数据集和新德里 HauzKhas 癫痫脑电数据集上的实验结果证明了 STSK-LLM-KD 的优势。

此外, STSK-LLM-KD 还有一些地方值得进一步研究。首先, 将对癫痫检测和运动预测等实际应用进行更加深入的研究。其次, 如何使用更先进的知识蒸馏算法提升 TSK 模糊分类器的性能也将是今后研究的重点。

参考文献:

- [1] ZHANG Xiongtao, CHUNG F L, WANG Shitong. An in-

- interpretable fuzzy DBN-based classifier for indoor user movement prediction in ambient assisted living applications[J]. *IEEE transactions on industrial informatics*, 2020, 16(1): 42–53.
- [2] JIANG Yizhang, DENG Zhaohong, CHUNG F L, et al. Recognition of epileptic EEG signals using a novel multiview TSK fuzzy system[J]. *IEEE transactions on fuzzy systems*, 2017, 25(1): 3–20.
- [3] ZHANG Yong, LIU Bo, JI Xiaomin, et al. Classification of EEG signals based on autoregressive model and wavelet packet decomposition[J]. *Neural processing letters*, 2017, 45(2): 365–378.
- [4] WANG Jun, LIN Defu, DENG Zhaohong, et al. Multitask TSK fuzzy system modeling by jointly reducing rules and consequent parameters[J]. *IEEE transactions on systems, man, and cybernetics: systems*, 2021, 51(7): 4078–4090.
- [5] QIN Bin, CHUNG F L, WANG Shitong. Biologically plausible fuzzy-knowledge-out and its induced wide learning of interpretable TSK fuzzy classifiers[J]. *IEEE transactions on fuzzy systems*, 2020, 28(7): 1276–1290.
- [6] JIANG Yizhang, WU Dongrui, DENG Zhaohong, et al. Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system[J]. *IEEE transactions on neural systems and rehabilitation engineering*, 2017, 25(12): 2270–2284.
- [7] DENG Zhaohong, XU Peng, XIE Lixiao, et al. Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals[J]. *IEEE transactions on neural systems and rehabilitation engineering*, 2018, 26(8): 1481–1494.
- [8] JIANG Yizhang, ZHANG Yuanpeng, LIN Chuang, et al. EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system[J]. *IEEE transactions on intelligent transportation systems*, 2021, 22(3): 1752–1764.
- [9] ZHENG Yuanhang, XU Zeshui, WANG Xinxin. The fusion of deep learning and fuzzy systems: a state-of-the-art survey[J]. *IEEE transactions on fuzzy systems*, 2022, 30(8): 2783–2799.
- [10] XIE Runshan, WANG Shitong. A wide interpretable Gaussian Takagi-Sugeno-Kang fuzzy classifier and its incremental learning[J]. *Knowledge-based systems*, 2022, 241: 108203.
- [11] ZHOU Ta, ISHIBUCHI H, WANG Shitong. Stacked blockwise combination of interpretable TSK fuzzy classifiers by negative correlation learning[J]. *IEEE transactions on fuzzy systems*, 2018, 26(6): 3327–3341.
- [12] TIAN Xiaobin, DENG Zhaohong, YING Wenhao, et al. Deep multi-view feature learning for EEG-based epileptic seizure detection[J]. *IEEE transactions on neural systems and rehabilitation engineering*, 2019, 27(10): 1962–1972.
- [13] 蒋云良, 翁江玮, 申情, 等. 基于增强深度特征和TSK模糊分类器的癫痫脑电信号识别[J]. *控制与决策*, 2023, 38(1): 171–180.
- JIANG Yunliang, WENG Jiangwei, SHEN Qing, et al. TSK fuzzy classifier based on enhanced deep feature for epilepsy EEG signal recognition[J]. *Control and decision*, 2023, 38(1): 171–180.
- [14] QIN Bin, NOJIMA Y, ISHIBUCHI H, et al. Realizing deep high-order TSK fuzzy classifier by ensembling interpretable zero-order TSK fuzzy subclassifiers[J]. *IEEE transactions on fuzzy systems*, 2021, 29(11): 3441–3455.
- [15] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. (2015–03–09) [2021–01–01]. <http://arxiv.org/abs/1503.02531>.
- [16] ZHAO Wenda, TONG Tingting, WANG Haipeng, et al. Diversity consistency learning for remote-sensing object recognition with limited labels[J]. *IEEE transactions on geoscience and remote sensing*, 2022, 60: 5631510.
- [17] WANG Peng, WEN Jun, SI Chenyang, et al. Contrast-reconstruction representation learning for self-supervised skeleton-based action recognition[J]. *IEEE transactions on image processing*, 2022, 31: 6224–6238.
- [18] TU Zhigang, LIU Xiangjian, XIAO Xuan. A general dynamic knowledge distillation method for visual analytics[J]. *IEEE transactions on image processing*, 2022, 31: 6517–6531.
- [19] LIU Zhiwei, LYU Wentao, WANG Chengqun, et al. D-CenterNet: an anchor-free detector with knowledge distillation for industrial defect detection[J]. *IEEE transactions on instrumentation and measurement*, 2022, 71: 2518412.
- [20] SONG Jie, CHEN Ying, YE Jingwen, et al. Spot-adaptive knowledge distillation[J]. *IEEE transactions on image processing*, 2022, 31: 3359–3370.
- [21] BLAKENEY C, LI Xiaomin, YAN Yan, et al. Parallel blockwise knowledge distillation for deep neural network compression[J]. *IEEE transactions on parallel and distributed systems*, 2021, 32(7): 1765–1776.
- [22] ZHANG Libo, DU Dawei, LI Congcong, et al. Iterative knowledge distillation for automatic check-out[J]. *IEEE transactions on multimedia*, 2020, 23: 4158–4170.
- [23] ZHAO Qilu, DONG Junyu, YU Hui, et al. Distilling ordinal relation and dark knowledge for facial age estimation[J]. *IEEE transactions on neural networks and learning systems*, 2021, 32(7): 3108–3121.
- [24] JIANG Yunliang, WENG Jiangwei, ZHANG Xiongtao, et

- al. A CNN-based born-again TSK fuzzy classifier integrating soft label information and knowledge distillation[J]. *IEEE transactions on fuzzy systems*, 2023, 31(6): 1843–1854.
- [25] GU Xiangming, CHENG Xiang. Distilling a deep neural network into a takagi-sugeno-Kang fuzzy inference system[EB/OL]. (2020-10-10)[2021-01-01]. <http://arxiv.org/abs/2010.04974>.
- [26] ERDEM D, KUMBASAR T. Enhancing the learning of interval type-2 fuzzy classifiers with knowledge distillation[C]//2021 IEEE International Conference on Fuzzy Systems. Luxembourg: IEEE, 2021: 1–6.
- [27] ZHOU Ta, CHUNG F L, WANG Shitong. Deep TSK fuzzy classifier with stacked generalization and triplely concise interpretability guarantee for large data[J]. *IEEE transactions on fuzzy systems*, 2017, 25(5): 1207–1221.
- [28] DEMIRLI K, MUTHUKUMARAN P. Higher order fuzzy system identification using subtractive clustering[J]. *Journal of intelligent & fuzzy systems*, 2000, 9(3/4): 129–158.
- [29] BISIADA M. Empirical studies in translation and discourse (Volume 14)[M]. Berlin: Language Science Press, 2021: 165–198.
- [30] SCHLÖGL A, NEUPER C, PFURTSCHELLER G. Estimating the mutual information of an EEG-based brain-computer interface[J]. *Biomedizinische Technik Biomedical engineering*, 2002, 47(1/2): 3–8.
- [31] SAMEER M, GUPTA B. Detection of epileptical seizures based on alpha band statistical features[J]. *Wireless personal communications*, 2020, 115(2): 909–925.
- [32] WU Dongrui, JIANG Xue, PENG Ruimin. Transfer learning for motor imagery based brain-computer interfaces: a tutorial[J]. *Neural networks*, 2022, 153: 235–253.

作者简介:



蒋云良, 博士, 教授, 博士生导师, 主要研究方向为深度学习、智慧交通、智慧医疗和智能教育。发表学术论文 60 余篇。E-mail: jyl@zjhu.edu.cn。



印泽宗, 硕士研究生, 主要研究方向为模糊系统、深度学习。E-mail: 759960776@qq.com。



张雄涛, 博士, 副教授, 主要研究方向为深度学习、模糊系统、智慧交通和智慧医疗。E-mail: 1047897965@qq.com。