Title
SVD-Based Broad Learning System with Optimized Network Structure for EEG Emotion Recognition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Zhang, Qihuai
Zhao, Xiaojie

Publication Date
2026

SVD-Based Broad Learning System with Optimized Network Structure for EEG Emotion Recognition

Qihuai Zhang[1] & Xiaojie Zhao (zhaox86@163.com)[1]
[1]School of Artificial Intelligence, Beijing Normal University

Abstract

In recent years, EEG-based emotion recognition has emerged as a research hotspot in affective computing, yet mainstream deep learning approaches demand considerable time and substantial computational resources. To address these challenges, this paper proposes the HTK-BLS model, which accelerates both training and testing while maintaining competitive accuracy and reducing resource consumption. First, high-order singular value decomposition (SVD) is employed to enhance the network's capacity to capture tensor-level information from EEG frequency domain features. To further improve BLS's feature extraction capabilities, a novel stacked architecture employing approximate kernel functions is introduced with an optimized objective function to eliminate redundant nodes. In addition, weight-matrix computations are refined through truncated SVD, improving the robustness of the network. Experimental results on the DEAP dataset demonstrate that HTK-BLS achieves competitive recognition accuracy while enabling efficient training without backpropagation.

Keywords: Broad learning system; Electroencephalography; Emotion recognition; Network structure; Computational efficiency

Introduction

Emotion recognition is the process of identifying and classifying an individual's emotional state based on physiological or behavioral signals (Pan & Bai, 2024). Electroencephalography (EEG) has demonstrated broad application potential in neural signal analysis, and its value is increasingly evident in emotion recognition research (Spyrou et al., 2016; Liu et al., 2021). Deep learning has become the most successful approach in EEG-based emotion recognition (Wei et al., 2020); however, its frequently complex network architectures require extensive iterative training, resulting in long training times and high resource consumption (Cheng et al., 2022). In scenarios requiring rapid processing and timely feedback, such as health monitoring and human–computer interaction, systems should support efficient incremental learning with minimal delay (Gao et al., 2021; Kavoosi et al., 2022).

The Broad Learning System (BLS) features a simple architecture that effectively addresses the challenges faced by deep learning (Chen & Liu, 2018). The basic architecture of the BLS is shown in Figure 1. Recent studies have shown

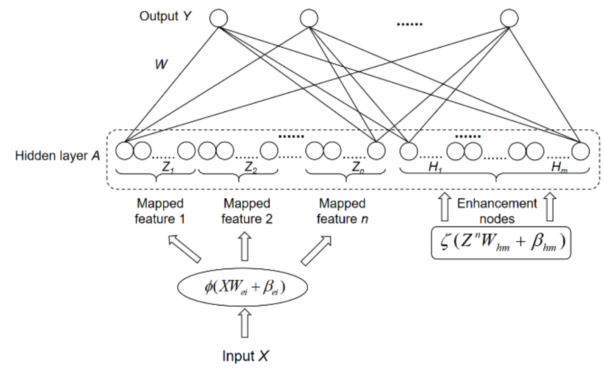

Figure 1: Basic Broad Learning System network structure.

that BLS has demonstrated strong potential in rapid EEG-based emotion recognition tasks. For example, Issa et al. (2021) proposed extracting grayscale image features from single-channel EEG data via continuous wavelet transform and then applying BLS for emotion classification; Zhong et al. (2024) introduced a BLS enhancement inspired by gated recurrent units, that employs EEG differential entropy features for emotion recognition.

However, BLS still faces several challenges (Li et al., 2022). First, like many machine learning algorithms, BLS was originally designed for vectorized data and relies on randomly generated, fixed hidden layer parameters (Li, Wang, & Liu, 2023). Nonetheless, EEG frequency domain features typically include information across multiple frequency bands for each channel and are commonly represented as tensors, making it difficult for BLS to capture the inherent relationships among EEG channels and frequency bands (Zhang, Fan, & Xu, 2015). This issue has been observed both in other application domains and with alternative algorithms. Pan et al. (2024) employed uncorrelated multilinear principal component analysis to transform tensor-form EEG time and frequency features into vectors for classification by an extreme learning machine. Zhang et al. (2019) developed an enhanced ELM that applies Tucker decomposition to tensor-level inputs to refine input representations, while incorporating inner-product optimizations into its network structure.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

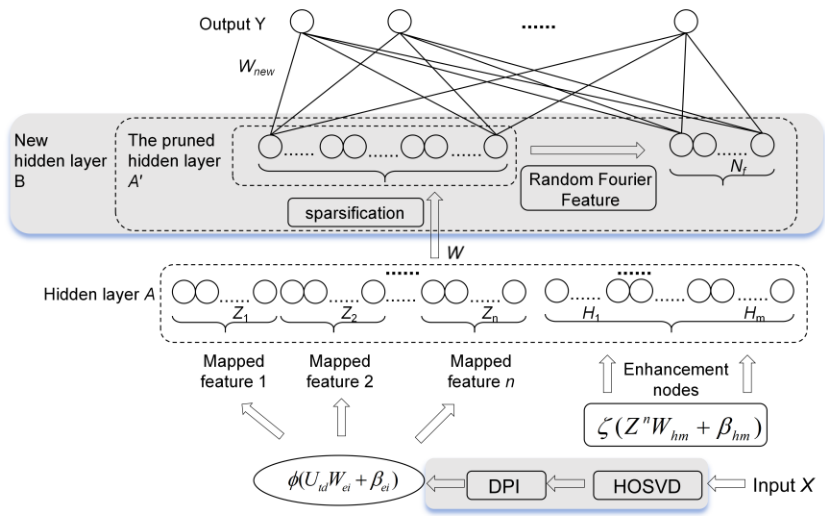

Figure 2: HTK-BLS (High-order Tensor Kernel Broad Learning System) network structure.

Secondly, the hidden layer of BLS still inadequately extracts features. One reason is that although a single hidden layer can theoretically approximate any function, its finite number of nodes limits nonlinear representation in practical applications (Liu et al., 2019). Another reason is that the hidden-layer parameters are generated by random mappings, so not all nodes are effective, leading to redundancy (Chu et al., 2020). Some researchers have addressed these limitations by stacking the original hidden layer or integrating kernel functions into BLS (Ding et al., 2021), and some by optimizing the objective function with Lasso or elastic-net to prune redundant nodes (Chu et al., 2024; Zhao et al., 2023). Moreover, random mapping can leave the weight matrix connecting the hidden and output layers ill-conditioned, undermining robustness and overall performance (Wang et al., 2015).

To address these issues based on prior work, this paper proposes a new framework for EEG emotion recognition, called HTK-BLS. First, a high-order singular value (HOSVD) integrated with the discriminant power index (DPI) generates tensor decomposition nodes, modeling the inherent relationships among EEG channels and frequency bands. Second, to optimize the network structure, the $L_{2,1}$ norm is employed to eliminate redundant node groups, and a smoothing regularizer is incorporated to enhance robustness and reduce overfitting. In addition, we use random Fourier features to approximate the kernel function and further map the pruned hidden layer, forming kernel nodes that enhance BLS's feature extraction capability. Moreover, truncated SVD optimizes the pseudo-inverse computation of the BLS weight matrix, addressing numerical instability and improving accuracy, thus yielding a fast BLS-based EEG emotion recognition model.

(1) An efficient emotion recognition model requiring no backpropagation is proposed.
(2) HOSVD is employed to construct tensor decomposition nodes, thereby addressing BLS's limitation in capturing tensor-level information.
(3) Both redundant nodes removal and approximate kernel mapping are employed to develop a new, more compact, and more efficient BLS architecture.

The rest of this paper is organized as follows. Section II presents the HTK-BLS framework and its algorithmic details. Section III describes the experimental datasets and reports the results. Finally, Section IV offers concluding remarks.

The Proposed Method

HTK-BLS Framework

The HTK-BLS framework (Figure 2) first generates tensor-decomposition nodes through HOSVD and DPI integration. These nodes are then mapped into the BLS's hidden layer, where redundant nodes are pruned by optimizing the objective function, while hierarchical stacking of hidden layers is implemented via an approximate kernel mapping. Finally, truncated SVD optimizes the BLS weight matrix computation.

Broad Learning System

BLS was proposed by Chen and Liu in 2018. Assuming the training dataset is $\{X \in R^{P\times M}\}$ and $\{Y \in R^{P\times N}\}$, where P denotes the number of training samples, and M and N denote the dimensions of the input and output data, respectively. Initially, BLS maps the input matrix X to n sets of feature nodes. The ith group of feature nodes is defined as follows:

$$Z_i = \phi(XW_{ei} + \beta_{ei}) \tag{1}$$

where W_{ei} is a randomly generated weight matrix and β_{ei} is a randomly generated bias vector ($i = 1,\ldots,n$), and the function ϕ denotes a mapping operation. All feature nodes are collected as $Z^n=[Z_1,\ldots,Z_n]$. Assume there are m groups of enhancement nodes defined as follows:

$$H_j = \xi\left(Z^n W_{hj} + \beta_{hj}\right) \quad (2)$$

where W_{hj} is a randomly generated weight matrix and β_{hj} is a randomly generated bias vector (j=1,...,m). ξ represents a nonlinear activation function. All enhancement nodes are represented as $H^m=[H_1,...,H_m]$.

The hidden layer A of the network is formed by concatenating the feature nodes and the enhancement nodes as $A=[Z^n|H^m]$. The output of the BLS is given by $Y=AW$, where W is the output weight matrix, defined as follows:

$$\arg\min_{w} \| Y - AW \|_2^2 + \lambda \| W \|_2^2 . \quad (3)$$

Using the pseudo-inverse, W can be obtained from (4), which represents the computation to a classical convex optimization problem:

$$W = (\lambda I + AA^T)^{-1} A^T Y \quad (4)$$

where λ represents the regularization parameter and I is the unit matrix.

Tensor Decomposition Nodes Creation

For a given matrix, SVD factorizes it into the product of two orthogonal matrices, and a diagonal matrix.

HOSVD was applied to the tensor, involving mode unfolding and SVD for each dimension (channel and frequency-band), by projecting the EEG spectral features onto a reduced set of principal components, we achieve a fused representation that preserves the intrinsic relationships both between channels and across frequency bands. For a three-dimensional sample tensor X of size (samples × channels × frequency bands), we first perform mode unfolding and then apply a standard SVD to each of the resulting unfolded matrices:

$$X_{(1)} = U^{(1)}\Sigma^{(1)}V^{(1)T},\ X_{(2)} = U^{(2)}\Sigma^{(2)}V^{(2)T},\ X_{(3)} = U^{(3)}\Sigma^{(3)}V^{(3)T} \quad (5)$$

where each $U^{(i)}$ is an orthogonal matrix. Since this study does not perform fusion along the sample mode, we set $U^{(1)}$ to the identity matrix. The core tensor of the EEG features S is defined as follows:

$$\mathcal{S} = X \times_1 \left(U^{(1)}\right)^T \times_2 \left(U^{(2)}\right)^T \times_3 \left(U^{(3)}\right)^T . \quad (6)$$

Each element of $\mathcal{S}$ represents an EEG feature that integrates relationships among channels and frequency bands. Although the core tensor of EEG features obtained via HOSVD captures multilinear interactions, it may still contain redundant components and, being an unsupervised method, does not account for class labels. To address this, we employ the DPI based on the Fisher score that selects features maximizing between-class separation and minimizing within-class variance. For binary classification, the DPI for feature i is defined as follows:

$$DPI_i = (\mu_{0,i} - \mu_{1,i})^2 / \sigma_{0,i}^2 + \sigma_{1,i}^2 + \epsilon \quad (7)$$

where the labels 0 and 1 denote the two classes, μ denotes the mean, σ denotes the variance, and ϵ is a very small constant. A higher DPI value indicates greater discriminative power of the corresponding feature. Features with low DPI values are removed, and the remaining features constitute the tensor-decomposition nodes U_{td}. Subsequently, U_{td} is used as input and mapped to produce the feature nodes and enhancement nodes, thereby obtaining the hidden layer A and the initial weight matrix W.

BLS with Optimized Objective Function

The $L_{2,1}$ norm encourages row-wise sparsity, pruning redundant node groups and producing a compact, discriminative hidden layer. To mitigate overfitting caused by random weight variability, we add a smoothness regularizer, compute the correlations among node groups to preserve or prune similar nodes in groups, smooth the output weights, thereby enhancing generalization. The optimized BLS objective function is reformulated as follows:

$$\arg\min_{w} \| Y - AW\|_2^2 + \lambda(\| DW\|_2^2 + \| W\|_{2,1}) \quad (8)$$

where D is a first-order difference matrix. We apply the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) to solve (9), first decomposing the objective function into its smooth and non-smooth components, as follows:

$$\arg\min_{w} \{ \underbrace{\| Y - AW\|_2^2 + \lambda\| DW\|_2^2}_{f(W)} + \underbrace{\lambda \| W\|_{2,1}}_{g(W)} \} \quad (9)$$

where λ denotes the regularization parameter. For the differentiable component of the objective function, its gradient is computed as follows:

$$\nabla f(W) = 2A^T (AW - Y) + 2\lambda D^T (DW). \quad (10)$$

The non-smooth term is the $L_{2,1}$-norm, whose proximal operator can be computed by processing the matrix row-wise as follows:

$$[prox_{\lambda g/L}(W)]_i = \max\left(1 - \lambda / L\| W_i\|_2, 0\right) w_i \quad (11)$$

where $L=\lambda_{max}(A^TA)$ denotes the Lipschitz constant, i=1……M, w_i denotes the i-th row of W. To solve the optimized objective function using FISTA, we begin by randomly initializing the network weight matrix W and then perform iterative updates. During the iterative process, we first execute the algorithm's acceleration step as follows:

$$Z^{(k)} = W^{(k)} + t_{k-1} - 1(W^{(k)} - W^{(k-1)}) / t_k , \quad (12)$$

$$t_{k+1} = 1 + \sqrt{1 + 4t_k^2} / 2, \quad t_0 = 1. \quad (13)$$

Next, we perform the gradient-descent update followed by the proximal mapping using the group-sparse proximal operator as follows:

$$\tilde{W} = Z^{(k)} - \nabla f(Z^{(k)}) / L, \quad (14)$$

$$W^{(k+1)} = prox_{\lambda g/L}(\tilde{W}). \quad (15)$$

Formula (12) - (15) are iteratively applied until the change in the weight matrix between successive updates falls below the predefined threshold, or the maximum number of iterations is reached, yielding the optimized BLS weight matrix W. After obtaining the updated weight matrix W, the weights corresponding to each node group are ranked, and those with the smallest values are deemed redundant. By removing these redundant nodes from A, the pruned hidden layer A' is obtained.

Stacked Structure Utilizing Approximate Kernel Functions and Truncated SVD

According to Bochner's theorem, if the kernel function $k(x-y)$ is continuous and positive definite, the Fourier transform $p(\omega)$ constitutes a valid probability density function. A sample ω can then be drawn from $p(\omega)$ at random. The kernel function $k(x, y)$ can be approximated as follows:

$$k(x,y) = \mathbb{E}_\omega[\phi(\omega,x)^T \phi(\omega,y)]^T, \quad (16)$$

$$\phi(\omega,x) = (\cos(\omega^T x), \sin(\omega^T x))^T. \quad (17)$$

By means of (16) and (17), the kernel function can be approximated by drawing L independent and identically distributed samples from the distribution $p(\omega)$ and computing their empirical mean. The kernel nodes N_f in BLS are defined as follows:

$$N_f = \begin{pmatrix} f_n(x_1) \\ \vdots \\ f_n(x_N) \end{pmatrix}, \quad (18)$$

$$z(x) = 1/\sqrt{L}\left(\cos(\omega_1^T x),\ldots,\cos(\omega_L^T x),\sin(\omega_1^T x),\ldots,\sin(\omega_L^T x)\right) \quad (19)$$

where the weights ω are obtained from the given probability density $p(\omega)$. By substituting the new hidden layer A' into (19) for x, the kernel nodes N_f are derived. The original BLS hidden layer A with redundant nodes removed, is concatenated with the kernel node group to form the final BLS hidden layer B, B=[A'|N_f]. The stacked BLS architecture designed in this study is illustrated in Figure 3.

For the final new hidden-layer B, we first perform an SVD as follows:

$$B = U\Sigma V^T, \quad \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r). \quad (20)$$

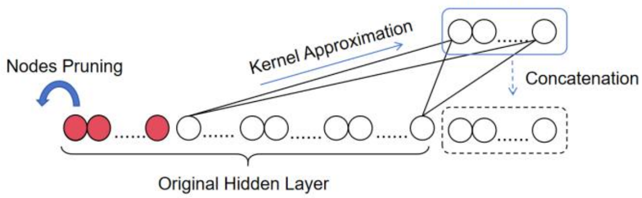

Figure 3: Architecture of the stacked Broad Learning System.

Here, r denotes the rank of matrix B. Among these singular values, only the leading n are retained and each augmented with a regularization term λ, while all remaining singular values are set to zero:

$$\gamma_i = \begin{cases} 1/\sigma_i + \lambda, & 1 \le i \le n, \\ 0, & n < i \le r. \end{cases} \quad (21)$$

From (20) and (21), the pseudo-inverse of the truncated singular value matrix can be expressed as follows:

$$\Sigma^+_{trunc} = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n, 0, \ldots, 0). \quad (22)$$

The final output weight matrix of the network, W_{new}, is given as follows:

$$W_{new} = V\Sigma^+_{trunc} U^T Y. \quad (23)$$

By computing the HTK-BLS weight matrix via truncated SVD, the influence of small singular values can be effectively eliminated, thereby enhancing the model's performance.

Experimental Materials and Results

Dataset and Data Preprocessing

In this paper, we selected the DEAP dataset (Koelstra et al., 2012). This dataset includes 32 participants. For the four dimensions—valence, arousal, dominance, and liking, each dimension is divided into high and low categories. 32 EEG channels were used for analysis. Each trial consists of a 63-second signal, with the first 3 seconds designated as the baseline. The remaining EEG signals were downsampled to 128 Hz and processed using a band-pass filter ranging from 4 to 45 Hz.

We used a 256-sample sliding window with a 32-sample step to segment the continuous EEG and applied a Fourier transform to each segment. For each channel, we then calculated the total amplitude within each of the five frequency bands as frequency-domain features (Jiang et al., 2022). These bands were defined as follows. θ (4–8 Hz), α (8–13 Hz), low β (13–18 Hz), high β (18–30 Hz), and γ (30–45 Hz).

In the experiments, the samples were divided into two classes using a threshold value of 5. Samples with scores lower than 5 were labeled as low, whereas samples with scores greater than or equal to 5 were labeled as high.

Parameter Setup

Following prior studies (Jia et al., 2020), we set the regularization parameter λ to 2^{-16}. During weight matrix computation, iterations are terminated when either the Frobenius norm of the difference between consecutive weight matrices drops below 10^{-2} or twenty iterations are completed. The 4% of nodes with the smallest absolute weights are then pruned as redundant. Finally, model performance is evaluated via ten-fold cross-validation.

All experiments ran in JupyterLab 3.2.1 on an Intel Core i9-12900 CPU (8 GB RAM) and an NVIDIA GeForce RTX 3070 Ti GPU (8 GB RAM).

Table 1: Classification accuracy results of HTK-BLS under different numbers of nodes.

Feature nodes and groups	Enhance-ment nodes	Kernel nodes	Train time(s)	Test time(s)	Acc (%)
10×10	1000	500	2.67	0.046	87.84
15×15	1500	750	8.86	0.080	92.77
20×20	***2000***	***1000***	***12.28***	***0.101***	***93.98***
25×25	2500	1250	30.11	0.192	94.07
30×30	3000	1500	78.12	0.303	94.10

We first evaluated the model's parameter settings. Drawing on existing literature (Chu et al., 2024), multiple models were constructed by varying the number of feature nodes. For each model, a discrete grid search was conducted to explore both the number of tensor decomposition nodes and the number of truncated singular values, thereby identifying the optimal parameter combination, and grid search used only the training folds. Results for each model under its optimal configuration were shown in Table I as the average performance across all four DEAP dimensions.

Table I indicates that when the number of feature nodes reaches twenty, test accuracy plateaus while both training and testing times increase markedly. Therefore, we selected the third configuration for subsequent experiments.

When selecting the number of tensor decomposition nodes, choosing too few may lead to the loss of useful information, while choosing too many may not only introduce features with little impact on the classification task or even cause adverse effects, but also increase the computational burden. Regarding the choice of truncated singular values, truncating too many values can result in the loss of valuable information, whereas truncating too few may render the weight matrix ill-conditioned or numerically unstable, thereby reducing classification accuracy. Finally, Figures 4 and 5 present the classification accuracy curves when one parameter varies while the other is fixed at its optimal value.

As shown in Figures 4 and 5, in the selected model, a discrete grid search identified the optimal performance when the number of tensor decomposition nodes was set to 120 and the number of truncated singular values to 200.

Experimental Results

To evaluate the model's stability among different participants, we conducted comparative experiments between BLS and HTK-BLS on 32 participants from the DEAP dataset, using the average accuracy across its four dimensions; the results are presented in Figure 6.

Figure 6 shows that HTK-BLS outperforms the original BLS across all subjects, with lower accuracy for subjects 2 and 22 and higher accuracy for subjects 13, 16, 23, and 27.

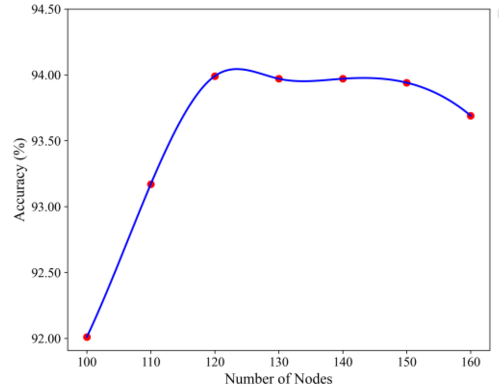

Figure 4: Relationship between number of tensor decomposition nodes and accuracy.

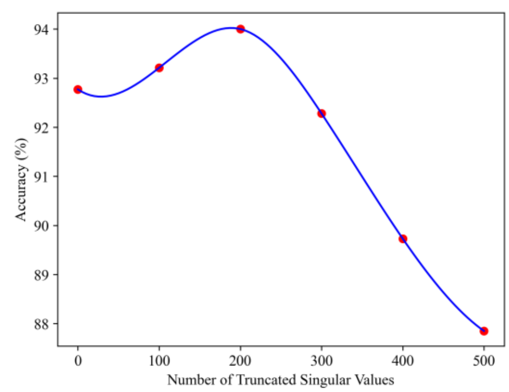

Figure 5: Relationship between the number of Truncated Singular Values and accuracy.

This pattern, consistent with previous work (Han, Zhang, & Yin, 2024), underscores significant inter-subject variability and confirms the model's effectiveness in capturing subject-specific information and the reliability of the results.

To validate the effectiveness of HTK-BLS, we conducted ablation experiments on all four dimensions of the DEAP dataset. In HK-BLS, truncated SVD was omitted; in TK-BLS, tensor-decomposition nodes were replaced by flattening the input signal; and in HT-BLS, kernel nodes were excluded. The average experimental results across the four DEAP dimensions are presented in Table II.

Table II shows that all three ablation variants outperformed the baseline, with HTK-BLS achieving the highest accuracy. The models with the optimized objective function significantly reduced the number of network nodes, and the node numbers reported in the table are averaged across multiple experimental runs. Although these enhanced models required slightly more training time and testing time than the baseline, the increases remained acceptable, and they achieved clear improvements in classification performance, striking a balanced trade-off between speed and accuracy.

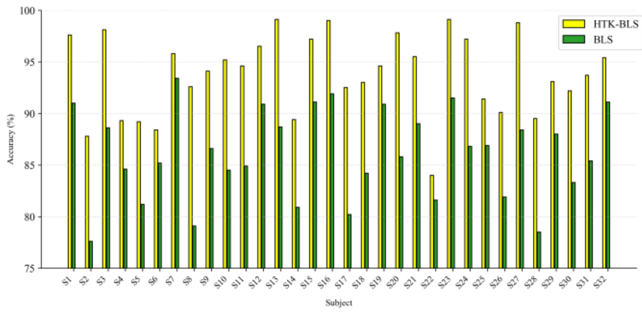

Figure 6: Average accuracy of 32 participants on the DEAP dataset.

Table 2: Ablation experimental results of HTK-BLS on the DEAP dataset.

MODEL	BLS	HT-BLS	HK-BLS	TK-BLS	HTK-BLS
Accuracy(%)	85.1	87.6	92.7	88.2	***94.0***
number of nodes	3400	3400	3210.7	***3204.1***	3215.9
train time(s)	***7.21***	9.94	11.97	10.12	12.28
test time(s)	***0.038***	0.078	0.099	0.060	0.101

Table 3: Comparative experimental results of HTK-BLS on the DEAP dataset.

MODEL	Valence	Arousal
CNN	87.96%	88.12%
MLP	89.12%	89.97%
ACRNN (Tao et al., 2023)	93.72%	93.38%
MS-TimesNet (Han et al., 2024)	90.45%	91.31%
ST-C shapes (Chu et al., 2023)	93.25%	93.16%
MLP-V2 (Neverlien et al., 2023)	91.19%	91.78%
Ours	***93.74%***	***93.97%***

We also compared HTK-BLS's performance on the valence and arousal dimensions of the DEAP dataset with that of classical machine learning models and recent state-of-the-art methods. For classical models, we selected Multilayer Perceptron (MLP) and Convolutional Neural Network (CNN). The MLP consists of four fully connected layers and two Dropout layers, while the CNN includes two convolutional layers, two pooling layers, one fully connected layer, and one Dropout layer. Both models were trained using Adam optimizer. The results are presented in Table III.

In comparative experiments, HTK-BLS achieved classification accuracies of 93.74% for valence and 93.97% for arousal on the DEAP dataset. It outperformed traditional machine-learning algorithms while remaining competitive with recent deep-learning methods. Moreover, because HTK-BLS does not require backpropagation, its training time is substantially reduced, greatly simplifying model development and enabling reliable, rapid EEG emotion recognition.

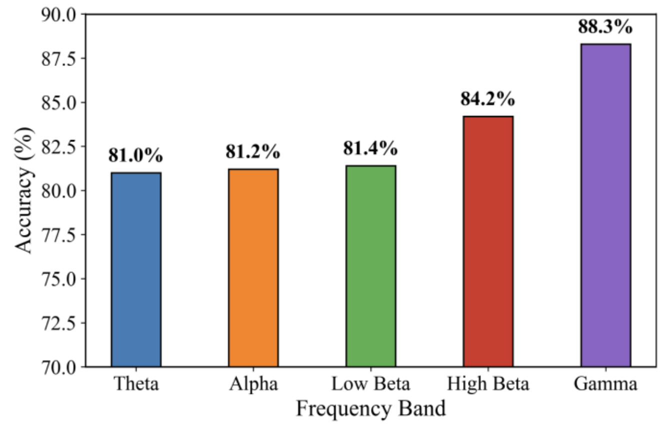

Figure 7: Classification results of single frequency bands on the DEAP dataset.

We further conducted single-band classification to evaluate the contribution of each EEG band. Figure 7 shows the average accuracy across the four DEAP dimensions. High beta and gamma outperformed the other bands, and gamma achieved the highest accuracy, indicating that high-frequency EEG bands contain more emotion-related information.

Conclusion

In this paper, we proposed HTK-BLS, a novel BLS-based model designed for rapid EEG emotion recognition without relying on backpropagation. To enhance BLS's ability to capture tensor-level EEG features, we constructed tensor decomposition nodes. The network structure was then optimized to remove redundant nodes, enhancing feature extraction. Moreover, a truncated SVD was employed to stabilize weight-matrix computation and enhance network robustness. Compared with current EEG emotion recognition methods, HTK-BLS offers a leaner structure and comparable accuracy without lengthy training. This efficiency supports fast EEG analysis and prompt system responses, especially in human–computer interaction, where swift emotional assessment is crucial. HTK-BLS therefore shows strong promise for broad deployment wherever quick, accurate EEG-based emotion recognition is needed. Our future work will further validate HTK-BLS under more diverse evaluation protocols and additional EEG emotion datasets.

References

Chen, C. L. P., & Liu, Z. (2018). Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Transactions on Neural Networks and Learning Systems, 29*(1), 10–24.

Cheng, W. X., Gao, R., Suganthan, P. N., & Yuen, K. F. (2022). EEG-based emotion recognition using random convolutional neural networks. *Engineering Applications of Artificial Intelligence, 116*, Article 105349.

Chu, F., Su, J., Liang, T., Chen, J., Wang, X., & Ma, X. (2020). Least angle regression adaptive incremental broad learning system. In *Proceedings of the 2020 Chinese Automation Congress (CAC)* (pp. 7153–7157). IEEE.

Chu, F., Liang, T., Chen, C. L. P., Wang, X., & Ma, X. (2024). Compact broad learning system based on fused lasso and smooth lasso. *IEEE Transactions on Cybernetics, 54*(1), 435–448.

Chu, W., Fu, B., Xia, Y., & Liu, Y. (2023). EEG-based emotion recognition using spatial–temporal connectivity. *IEEE Access, 11*, 92496–92504.

Ding, W., Tian, Y., Han, S., & Yuan, H. (2021). Greedy broad learning system. *IEEE Access, 9*, 79307–79315.

Gao, Z., Dang, W., Liu, M., Guo, W., Ma, K., & Chen, G. (2021). Classification of EEG signals on VEP-based BCI systems with broad learning. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51*(11), 7143–7151.

Han, L., Zhang, X., & Yin, J. (2024). EEG emotion recognition based on the TimesNet fusion model. *Applied Soft Computing, 159*, Article 111635.

Issa, S., Peng, Q., & You, X. (2021). Emotion classification using EEG brain signals and the broad learning system. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51*(12), 7382–7391.

Jia, X., Zhang, T., Chen, C. L. P., Liu, Z., Chen, L., Wen, G., & Hu, B. (2020). Multi-channel EEG based emotion recognition using temporal convolutional network and broad learning system. In *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 2452–2457). IEEE.

Jiang, H., Wu, D., Tang, X., Li, Z., & Wu, W. (2022). EEG emotion recognition using an attention mechanism based on an optimized hybrid model. *Computers, Materials & Continua, 73*(2), 2697–2712.

Kavoosi, A., Toth, R., Benjaber, M., Zamora, M. E., Valentín, A., Sharott, A., & Denison, T. (2022). Computationally efficient neural network classifiers for next generation closed loop neuromodulation therapy: A case study in epilepsy. In *Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 288–291). IEEE.

Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing, 3*(1), 18–31.

Li, G., Zhang, A., Zhang, Q., Wu, D., & Zhan, C. (2022). Pearson correlation coefficient-based performance enhancement of broad learning system for stock price prediction. *IEEE Transactions on Circuits and Systems II: Express Briefs, 69*(5), 2413–2417.

Li, Y., Wang, D., & Liu, F. (2023). The auto-correlation function aided sparse support matrix machine for EEG-based fatigue detection. *IEEE Transactions on Circuits and Systems II: Express Briefs, 70*(2), 836–840.

Liu, H., Zhang, Y., Li, Y., & Kong, X. (2021). Review on emotion recognition based on electroencephalography. *Frontiers in Computational Neuroscience, 15*, Article 758212.

Liu, Z., Chen, C. L. P., Zhang, T., & Zhou, J. (2019). Multi-kernel broad learning systems based on random features: A novel expansion for nonlinear feature nodes. In *Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (pp. 193–197). IEEE.

Neverlien, E. C. S., Lu, R., Kumar, M., & Molinas, M. (2023). Decoding emotions from EEG responses elicited by videos using machine learning techniques on two datasets. In *Proceedings of the 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 1–4). IEEE.

Pan, H., Song, W., Li, L., & Qin, X. (2024). The design and implementation of multi-character classification scheme based on EEG signals of visual imagery. *Cognitive Neurodynamics, 18*, 2299–2309.

Pan, J., & Bai, C. (2024). EEG-based emotion recognition via convolutional transformer with class confusion-aware attention. *Proceedings of the Annual Meeting of the Cognitive Science Society, 46*, 5287–5294.

Spyrou, L., Martín-López, D., Valentín, A., Alarcón, G., & Sanei, S. (2016). Detection of intracranial signatures of interictal epileptiform discharges from concurrent scalp EEG. *International Journal of Neural Systems, 26*(4), Article 1650016.

Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., & Chen, X. (2023). EEG-based emotion recognition via channel-wise attention and self attention. *IEEE Transactions on Affective Computing, 14*(1), 382–393.

Wang, J., Wang, X., Gu, S., Liao, W., & Fang, X. (2015). An improved extreme learning algorithm based on truncated singular value decomposition. In *Proceedings of the 2015 27th Chinese Control and Decision Conference (CCDC)* (pp. 1697–1701). IEEE.

Wei, C., Chen, L.-L., Song, Z.-Z., Lou, X.-G., & Li, D.-D. (2020). EEG-based emotion recognition using simple recurrent units network and ensemble learning. *Biomedical Signal Processing and Control, 58*, Article 101756.

Zhang, F., Fan, Y. L., & Xu, L. (2015). A report on uncorrelated multilinear principal component analysis plus extreme learning machine to deal with tensorial data. *Journal of Computational and Theoretical Nanoscience, 12*(7), 1258–1262.

Zhang, H., Ai, B., Xu, W., Xu, L., & Cui, S. (2019). Multi-antenna channel interpolation via Tucker-decomposed extreme learning machine. *IEEE Transactions on Vehicular Technology, 68*(7), 7160–7163.

Zhao, H., Zhang, P., Zhang, R., Yao, R., & Deng, W. (2023). A novel performance trend prediction approach using

ENBLS with GWO. *Measurement Science and Technology, 34*(2), Article 025018.

Zhong, X., Wu, F., Yin, Z., & Liu, G. (2024). An attention-enhanced retentive broad learning system for subject-generic emotion recognition from EEG signals. In *ICASSP 2024—2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2310–2314). IEEE.