Robust Broad Learning System with Wave Loss for Classification under Data Uncertainty

Mushir Akhtar
Department of Mathematics
Indian Institute of Technology Indore
phd2101241004@iiti.ac.in

Abhishek Varshney
Department of Mathematics
Indian Institute of Technology Indore
phd2401141003@iiti.ac.in

A. Quadir
Department of Mathematics
Indian Institute of Technology Indore
mscphd2207141002@iiti.ac.in

A. Rahaman
Department of Mathematics
Indian Institute of Technology Indore
phd2401141001@iiti.ac.in

M. Tanveer*
Department of Mathematics
Indian Institute of Technology Indore
mtanveer@iiti.ac.in

Mohd. Arshad
Department of Mathematics
Indian Institute of Technology Indore
arshad@iiti.ac.in

***Abstract*—Broad Learning System (BLS) offers an efficient alternative to deep architectures by enabling fast learning through randomized feature mapping and closed-form solutions. However, its reliance on squared error loss makes it highly sensitive to noise, outliers, and corrupted labels, limiting its reliability in real-world scenarios. To address this limitation, we propose Wave-BLS, a robust broad learning framework that integrates the wave loss function, which is asymmetric, bounded, and smooth, enabling controlled penalization of large errors. The proposed formulation replaces the standard least-squares objective with a wave-loss-based optimization problem, solved efficiently using a Nesterov accelerated gradient (NAG)-based scheme without requiring matrix inversion, thereby improving scalability. Extensive experiments on 30 UCI benchmark datasets demonstrate that Wave-BLS consistently outperforms classical BLS and several robust variants. Statistical validation using Friedman and Nemenyi post-hoc tests confirms the significance of the observed improvements. Furthermore, robustness evaluations under controlled noise and outlier injection reveal that Wave-BLS exhibits substantially slower performance degradation compared to BLS, even in challenging contamination settings. These results establish Wave-BLS as a stable and robust alternative to existing broad learning models for learning under data uncertainty. Code and Supplementary are available at https://github.com/mtanveer1/Wave-BLS.**

***Index Terms*—Randomized Neural Network, Broad Learning System, Wave Loss Function, Robust Classification, Noise and Outlier Robustness**

I. Introduction

RANDOMIZED neural networks (RaNNs) have emerged as efficient alternatives to deep neural networks, offering reduced computational complexity, fewer parameters, and less sensitivity to hyperparameter tuning while maintaining competitive performance [1]. By fixing randomly generated hidden-layer parameters and learning only the output weights in closed form, RaNNs enable fast training with reduced computational overhead. Among them, the random vector functional link (RVFL) network [2] has gained prominence due to its direct input-output connections, which allow simultaneous modeling of linear and nonlinear relationships and yield strong generalization performance [3].

Despite these advantages, the shallow architecture of RVFL limits its representational capacity for complex data distributions [4]. To overcome this limitation, the Broad Learning System (BLS) was introduced as a flat yet expressive architecture that expands network width instead of depth [5]. BLS constructs multiple groups of randomized feature and enhancement nodes whose outputs are concatenated to form a rich representation, while retaining fast training through a closed-form least-squares solution. This design avoids back-propagation, supports incremental learning, and performs well in limited-data regimes.

Nevertheless, BLS suffers from two fundamental limitations: (i) its dependence on the squared-error loss makes it highly sensitive to noise and outliers, and (ii) the computation of output weights relies on matrix inversion, resulting in cubic-time complexity that limits scalability.

A range of BLS variants have been proposed to improve robustness and mitigate sensitivity to noise. For example, Jin and Chen [6] introduced a regularized BLS (R-BLS) by incorporating ℓ_1, ℓ_2, and elastic-net penalties into the least-squares formulation, which helps control model complexity but remains sensitive to severe noise and outliers due to its reliance on the squared-error loss. To address uncertainty, Feng and Chen [7] proposed a neuro-fuzzy BLS (NF-BLS) that integrates fuzzy IF-THEN rules via k-means clustering, improving interpretability and robustness at the cost of additional computational overhead. More recently, the SC-KSBLS framework [8] replaces the squared-error loss with a kernel risk-sensitive mean p-power loss and incorporates adaptive label correction. While effective under label noise, its collaborative structure and kernel-based optimization substantially increase computational cost. Additional extensions have addressed specific challenges such as imbalanced noisy data via graph-fuzzy embeddings [9], privacy-preserving multiparty learning [10], and fractional-order BLS formulations [11]. Recently, copula-aligned weight initialization (CAWI) has been

*Corresponding author.

proposed to incorporate dependency-aware copula modeling into weight initialization, aligning randomized weights with the intrinsic structure of the input data and thereby improving representation quality and learning performance [12].

Overall, existing BLS variants enhance robustness through regularization, fuzzy inference, kernelization, or collaborative correction mechanisms, often at the expense of increased architectural complexity or computational overhead. These limitations motivate the need for a robust yet computationally efficient BLS framework that improves noise tolerance while preserving the simplicity of the original architecture.

The choice of loss function plays a central role in robustness against noise and outliers [13], yet most BLS-based methods continue to rely on the conventional least-squares loss. While robustness can be improved through auxiliary structures such as clustering or fuzzy membership evaluation, these additions inevitably increase computational burden. In contrast, directly adopting a robust loss function within the BLS optimization framework offers a principled and efficient alternative.

Motivated by this perspective, we consider the wave loss function [14], which has demonstrated strong robustness properties due to its boundedness, smoothness, and controllable asymmetry. We investigate whether the wave loss can be effectively integrated into the BLS framework to address its sensitivity to noise and outliers. Building on this idea, we propose Wave-BLS, i.e., a broad learning system with wave loss function, which replaces the traditional least-squares loss while retaining the flat architecture of BLS. In addition, we reformulate the parameter optimization process to eliminate the matrix inversion step required in conventional BLS, thereby significantly enhancing scalability. The resulting optimization problem is efficiently solved using a Nesterov accelerated gradient (NAG)-based algorithm, which ensures stable updates and fast convergence. In the following sections, we demonstrate that the proposed Wave-BLS model achieves improved generalization performance compared with existing baseline models.

The key contributions of the work are as follows:

1) We propose Wave-BLS, a robust broad learning system that integrates the wave loss function to effectively mitigate the impact of noise and outliers.
2) We formulate the optimization framework in Wave-BLS to eliminate the matrix inversion step required in traditional BLS, significantly enhancing its scalability.
3) We employ a NAG-based optimization algorithm to solve the proposed Wave-BLS model, ensuring efficient parameter updates and faster convergence.
4) We extensively evaluate the proposed Wave-BLS on 30 UCI benchmark datasets using accuracy and rank metrics, and further establish its statistical significance through Friedman and Nemenyi post-hoc tests.
5) We further assess the robustness of Wave-BLS under controlled data corruption by injecting varying levels of noise and outliers. The experimental results confirm that Wave-BLS maintains superior performance compared to the standard BLS.

The remainder of this paper is organized as follows. Section II reviews the architecture and formulation of BLS. Section III presents the proposed Wave-BLS model. Experimental evaluations are reported in Section IV. Finally, Section V concludes the paper.

II. Related Work

This section introduces the notations and presents an overview of the architecture and formulation of BLS.

A. Notations

Let $X \in \mathbb{R}^{m\times d}$ represent the input data matrix, where m is the number of samples and d is the number of features. The target matrix is denoted by $Y \in \mathbb{R}^{m\times n_{\text{out}}}$, where n_{out} is the number of classes. The outputs from the feature layer are represented by $Z \in \mathbb{R}^{m\times pq}$, where p and q denote the number of nodes per window and the number of feature node windows, respectively. Similarly, the outputs from the enhancement layer are represented by $H \in \mathbb{R}^{m\times rs}$, where r and s denote the number of nodes per window and the number of enhancement node windows, respectively. The combined representation matrix, obtained by concatenating the feature and enhancement layer outputs, is denoted as $A \in \mathbb{R}^{m\times (pq+rs)}$. Finally, the unknown weight matrix is denoted by $W \in \mathbb{R}^{(pq+rs)\times n_{\text{out}}}$.

B. Broad Learning System (BLS): Architecture and Formulation [5]

BLS is a flat network architecture composed of three main components: a feature layer, an enhancement layer, and an output layer. Unlike deep architectures, BLS expands the network width by incrementally adding feature and enhancement nodes, enabling efficient learning via closed-form solutions. Fig. 1 illustrates the overall architecture of the BLS framework.

Feature Layer: Given an input data matrix $X \in \mathbb{R}^{m\times d}$, the feature layer generates q groups (windows) of feature nodes, each containing p nodes. The output of the i^{th} feature window is computed as

$$Z_{f_i} = \phi(XW_i + B_i), \tag{1}$$

where $W_i \in \mathbb{R}^{d\times p}$ and $B_i \in \mathbb{R}^{m\times p}$ are randomly initialized weights and biases, respectively, and $\phi(\cdot)$ denotes a nonlinear activation function. Concatenating all feature windows yields

$$Z = [Z_{f_1}, Z_{f_2}, \ldots, Z_{f_q}] \in \mathbb{R}^{m\times pq}. \tag{2}$$

Enhancement Layer: The enhancement layer further enriches the representation by applying random nonlinear transformations to the feature output Z. Specifically, s enhancement windows are generated, each with r nodes:

$$H_{e_j} = \psi(ZW_j + B_j), \tag{3}$$

where $W_j \in \mathbb{R}^{pq\times r}$, $B_j \in \mathbb{R}^{m\times r}$, and $\psi(\cdot)$ is a nonlinear activation function. The overall enhancement output is

$$H = [H_{e_1}, H_{e_2}, \ldots, H_{e_s}] \in \mathbb{R}^{m\times rs}. \tag{4}$$

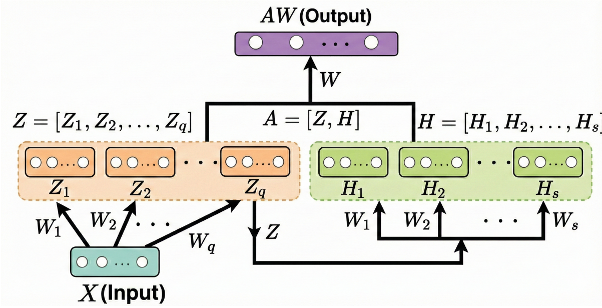

Fig. 1: Architecture of the standard BLS.

Output Layer: The final representation matrix is obtained by concatenating the feature and enhancement outputs:

$$A = [Z, H] \in \mathbb{R}^{m\times(pq+rs)}. \tag{5}$$

BLS determines the output weight matrix $W \in \mathbb{R}^{(pq+rs)\times n_{\text{out}}}$ by solving a regularized least squares problem:

$$\arg\min_{W} \ \frac{1}{2}\|W\|_F^2 + \frac{C}{2}\|AW - Y\|_F^2, \tag{6}$$

where Y is the target matrix and $C > 0$ is the regularization parameter. The closed-form solution is given by

$$W = \begin{cases} (A^\top A + \frac{1}{C}I)^{-1}A^\top Y, & (pq+rs) \leq m, \\ A^\top(AA^\top + \frac{1}{C}I)^{-1}Y, & m < (pq+rs), \end{cases} \tag{7}$$

with I denoting the identity matrix of appropriate dimension.

III. Proposed Work

In this section, we introduce Wave-BLS, i.e., the broad learning system based on the wave loss function. We first discuss the wave loss function and provide its key characteristics. Subsequently, we detail the formulation and optimization of the proposed Wave-BLS model.

A. Wave Loss Function

The wave loss function [14] is a recently introduced asymmetric, bounded, and smooth loss designed to provide controlled error penalization. It is defined as

$$L_{\text{wave}}(u) = \frac{u^2 e^{au}}{1 + \lambda u^2 e^{au}}, \tag{8}$$

where $u \in \mathbb{R}$ denotes the prediction error, $a \in \mathbb{R}$ is a shape parameter governing the asymmetry and growth behavior of the loss, and $\lambda \in \mathbb{R}^+$ is a bounding parameter that limits the maximum penalty induced by large errors. Fig. 2 illustrates the behavior of the wave loss function under different values of a and λ.

The key properties of the wave loss function, which are directly leveraged in the proposed Wave-BLS framework, are summarized below.

- **Asymmetry:** The wave loss is asymmetric, enabling unequal penalization of overestimation and underestimation errors through the shape parameter a. Positive values of a emphasize overpredictions, while negative values accentuate underpredictions; the loss becomes symmetric when $a = 0$.
- **Boundedness:** The wave loss is upper-bounded by $1/\lambda$, ensuring that large errors exert limited influence on the optimization process. This bounded nature prevents excessive domination by outliers and enhances robustness in noisy learning environments.
- **Smoothness:** The wave loss is continuously differentiable, facilitating efficient and stable optimization. This smoothness enables the use of efficient gradient-based solvers.

These properties enable the wave loss function to address the fundamental issues of traditional loss functions.

B. Formulation & Optimization: Broad Learning System with Wave Loss Function (Wave-BLS)

In this subsection, we incorporate the wave loss function into the BLS by replacing the conventional squared error loss, resulting in a robust learning framework termed broad learning system with wave loss function, i.e., Wave-BLS. The proposed formulation improves robustness against noise and outliers while enabling scalable optimization by avoiding explicit matrix inversion during parameter estimation.

The optimization problem of Wave-BLS is formulated as

$$J(W) = \frac{1}{2}\|W\|_F^2 + C\, L_{\text{wave}}(AW - Y), \tag{9}$$

where $C > 0$ is the regularization parameter controlling the trade-off between model complexity and empirical risk, and $L_{\text{wave}}(AW - Y)$ represents the aggregated wave loss over all training samples.

The wave loss for the i^{th} sample and jth class is defined as

$$L_{\text{wave}}(u_{ij}) = \frac{u_{ij}^2 e^{au_{ij}}}{1 + \lambda u_{ij}^2 e^{au_{ij}}}, \tag{10}$$

where $u_{ij} = (AW)_{ij} - Y_{ij}$ denotes the prediction error for sample i and class j.

Accordingly, the total wave loss over the training set is computed as

$$L_{\text{wave}}(AW - Y) = \sum_{i=1}^{m}\sum_{j=1}^{n_{\text{out}}} L_{\text{wave}}(u_{ij}). \tag{11}$$

Hence, the objective function in (9) can be equivalently rewritten as

$$J(W) = \frac{1}{2}\|W\|_F^2 + C\sum_{i=1}^{m}\sum_{j=1}^{n_{\text{out}}} L_{\text{wave}}(u_{ij}). \tag{12}$$

The objective function in (12) consists of two components:

1) **Regularization term** $\left(\frac{1}{2}\|W\|_F^2\right)$, which controls model complexity and promotes generalization in accordance with the principle of structural risk minimization.
2) **Empirical loss term** $\left(\sum_{i=1}^{m}\sum_{j=1}^{n_{\text{out}}} L_{\text{wave}}(u_{ij})\right)$, which aggregates the wave loss over all samples and classes, thereby enhancing robustness through its asymmetric, bounded, and smooth characteristics.

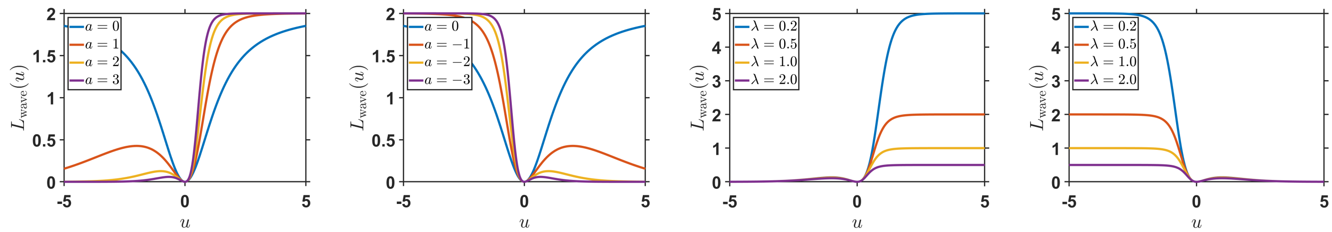

Fig. 2: Illustration of wave loss function, showing effect of asymmetry parameter a (positive, negative, and $a = 0$) and the bounding parameter λ.

To efficiently optimize the objective function in (12), we exploit its smoothness and use a gradient-based optimization strategy. This approach eliminates the need for explicit matrix inversion, which is a computational bottleneck in conventional BLS, thereby significantly improving scalability.

Specifically, we adopt the Nesterov accelerated gradient (NAG) algorithm due to its fast convergence and stable update behavior. The look-ahead mechanism of NAG enables accurate gradient estimates, while its momentum-based formulation effectively suppresses oscillations in ill-conditioned optimization landscapes.

Further, to ensure stable and efficient convergence, the learning rate is dynamically adjusted using an exponential decay schedule. At iteration t, the learning rate is updated as $\mu_t = \mu_0 \, e^{-\eta t}$, where μ_0 is the initial learning rate and η controls the decay rate. This adaptive strategy balances convergence speed and stability, leading to reliable optimization of the proposed Wave-BLS model.

To optimize the objective function in (12), we compute its gradient with respect to the output weight matrix W. The resulting gradient consists of contributions from the regularization term and the wave loss term and is used within the NAG-based optimization scheme.

For a given class $j \in \{1, 2, \ldots, n_{\text{out}}\}$, the gradient of the wave loss term with respect to $W(:, j)$ is given by

$$\nabla_{W(:,j)} L_{\text{wave}} = \sum_{i=1}^{m} \frac{u_{ij}\left(2 + a u_{ij}\right) e^{a u_{ij}}}{\left(1 + \lambda u_{ij}^2 e^{a u_{ij}}\right)^2} A(i,:)^\top. \tag{13}$$

This gradient is subsequently used to update the model parameters within the NAG optimization framework.

The overall gradient of the objective function $J(W)$ is given by

$$\nabla J(W) = W + C \, \nabla_W L_{\text{wave}}, \tag{14}$$

where the first term arises from the regularization component $\frac{1}{2}\|W\|_F^2$, and the second term corresponds to the contribution of the wave loss.

At the t^{th} iteration, the gradient is computed using the entire training dataset, ensuring stable and deterministic updates. Algorithm 1 provides the complete procedure for computing the output weight matrix W in the proposed Wave-BLS model using the NAG-based optimization scheme.

The computational complexity of the proposed Wave-BLS model is detailed in Section S.I of the supplementary file.

Algorithm 1 NAG-based optimization for Wave-BLS

Input: $A \in \mathbb{R}^{m \times (pq+rs)}$: representation matrix, $Y \in \mathbb{R}^{m \times n_{\text{out}}}$: target matrix, C: regularization parameter, $I_{\max}$: maximum number of iterations, $\mu^{(0)}$: initial learning rate, γ: momentum coefficient, η: learning rate decay factor, δ: convergence tolerance.

Output: $W \in \mathbb{R}^{(pq+rs) \times n_{\text{out}}}$: output weight matrix.

```
Initialize t ← 0, W^(0), v^(0) ← 0, μ^(0).
repeat
    Compute: W^look-ahead = W^(t) + γ v^(t).
    Compute stochastic gradient g^(t) = ∇J(W^look-ahead).
    Velocity update:
        v^(t+1) = γ v^(t) − μ^(t) g^(t).
    Model parameters update:
        W^(t+1) = W^(t) + v^(t+1).
    Learning rate update:
        μ^(t+1) = μ^(t) e^(−ηt).
    t ← t + 1.
until t ≥ I_max or ||W^(t) − W^(t−1)||_F ≤ δ
return W.
```

IV. Experiments and Discussion

This section empirically evaluates the proposed Wave-BLS against baseline models on 30 UCI benchmark datasets under both clean and noisy conditions. The proposed Wave-BLS model is evaluated against several baseline methods, including RVFL [15], ELM, i.e., RVFL without direct link(RVFLwoDL) [16], BLS [5], Wave-RVFL [17], NF-BLS [7], F-BLS [18], IF-BLS [18], and KRP-BLS [8]. The detailed experimental setup is provided in Section S.II of the supplementary material.

A. Performance Evaluation

Table I summarizes the average accuracy and average rank of the proposed Wave-BLS and competing methods over 30 UCI benchmark datasets. The detailed results are provided in Table S.I of the supplement file. Overall, Wave-BLS achieves the best performance in both criteria, indicating not only

TABLE I: Comparison of average classification performance of the proposed Wave-BLS and baseline models 30 UCI benchmark datasets.

Model →	**RVFL [15]**	**RVFLwoDL [16]**	**BLS [5]**	**Wave-RVFL [17]**	**NF-BLS [7]**	**F-BLS [18]**	**IF-BLS [18]**	**KRP-BLS [8]**	**Wave-BLS†**
Average Accuracy	79.8876	79.4376	82.0751	81.5155	80.0399	82.0376	<u>83.9102</u>	83.002	**86.7448**
Average Rank	6.5333	7.1833	4.8	4.0667	6.0833	5.2667	<u>3.4833</u>	4.8333	**2.75**

† indicates the proposed model; **bold** denotes the best performance, and <u>underline</u> denotes the second-best.

higher accuracy but also more consistent dominance across datasets. Among the classical randomized baselines, RVFL and RVFLwoDL obtain average accuracies of 79.8876% and 79.4376%, respectively. Moving to the broad-learning family, the standard BLS improves the average accuracy to 82.0751%. Several robust BLS variants further enhance performance: NF-BLS achieves 80.0399%, F-BLS reaches 82.0376%, IF-BLS attains the second-best average accuracy of 83.9102%, and KRP-BLS achieves 83.002%. In contrast, the proposed Wave-BLS yields the highest average accuracy of 86.7448%, providing a clear improvement over the strongest baselines. Specifically, Wave-BLS improves upon the best-performing baseline IF-BLS by 2.8346% (86.7448% vs. 83.9102%) and over KRP-BLS by 3.7428% (86.7448% vs. 83.002%). Moreover, comparing wave-loss-based models highlights the benefit of embedding wave loss into the broader BLS architecture: Wave-BLS outperforms Wave-RVFL by a large margin of 5.2293% (86.7448% vs. 81.5155%), indicating that the proposed formulation leverages both the robustness of wave loss and the stronger representation capacity of BLS. While average accuracy provides a useful aggregate summary, it can hide dataset-wise variability: a method may perform extremely well on a subset of datasets and poorly on others while still achieving a competitive mean. Average rank complements this by measuring relative performance per dataset and then aggregating, thereby better reflecting how consistently a method performs across heterogeneous benchmarks. Lower average rank indicates that a model attains top performance more frequently across datasets, rather than relying on a few large gains. The rank results in Table I further confirm the advantage of Wave-BLS. RVFL and RVFLwoDL obtain average ranks of 6.5333 and 7.1833, respectively, whereas the standard BLS improves to 4.8. Among robust BLS variants, IF-BLS achieves the second-best average rank of 3.4833, while KRP-BLS yields 4.833. Wave-BLS achieves the best average rank of 2.75, demonstrating the most consistent superiority across datasets. Quantitatively, Wave-BLS improves over IF-BLS by 0.7333 (3.4833 → 2.75), and over KRP-BLS by 2.0833 (4.8333 → 2.75). A similar trend is observed when comparing the wave-loss counterparts: Wave-BLS significantly improves over Wave-RVFL in rank (4.0667 → 2.75), confirming that the proposed integration of wave loss within the BLS framework not only boosts mean accuracy but also yields more consistent dataset-wise wins. Taken together, the lowest average rank and highest average accuracy indicate that Wave-BLS delivers both stronger predictive performance and better cross-dataset stability than existing baselines. This advantage is primarily driven by (i) replacing the squared-error criterion with the robust wave loss and (ii) coupling it with the richer feature-enhancement representation of BLS, yielding substantial gains over both the strongest BLS competitors (IF-BLS, KRP-BLS) and the wave-loss-based RVFL baseline (Wave-RVFL).

TABLE II: Robustness comparison of BLS and Wave-BLS under outlier and noise contamination.

Dataset	Level	Outliers		Level	Noise	
		BLS	Wave-BLS†		BLS	Wave-BLS†
blood	Clean	79.6272	**82.9184**	Clean	79.6272	**82.9184**
	5%	75.8421	**80.4067**	5%	77.1034	**81.7749**
	10%	72.5186	**78.3925**	10%	74.8652	**79.9368**
	15%	69.2847	**75.8106**	15%	71.4928	**77.1053**
	20%	66.0319	**72.4678**	20%	68.1074	**74.6389**
	Avg.	72.6609	**78.3992**	Avg.	74.3398	**79.2749**
horse_colic	Clean	**86.1422**	85.7919	Clean	**86.1422**	85.7919
	5%	82.6374	**84.9563**	5%	83.9046	**85.1187**
	10%	78.9421	**82.4018**	10%	80.2873	**82.9365**
	15%	75.6189	**79.6674**	15%	76.9548	**80.3412**
	20%	72.3046	**76.1897**	20%	73.6285	**77.5403**
	Avg.	79.1290	**81.8014**	Avg.	80.1835	**82.3457**
Overall Average		75.8949	**80.1003**		77.2617	**80.8103**
Average Rank		2.0	**1.0**		1.75	**1.25**

Bold values denote the best performance for each row.

To assess whether the observed differences in average ranks among the competing models are statistically significant, we employ the nonparametric Friedman test followed by Nemenyi post-hoc test. The detailed results are provided in Section S.III of the supplementary. The sensitivity analysis of the wave loss hyperparameters for the proposed Wave-BLS model is reported in Section S.IV of the supplementary material.

B. Robustness Analysis under Noise and Outliers

To evaluate the robustness of the proposed Wave-BLS, we conduct controlled experiments under outlier and noise contamination on two representative datasets. Specifically, artificial perturbations are introduced by contaminating 5%, 10%, 15%, and 20% of the training samples with outliers and label noise, respectively. The blood dataset is selected as a case where Wave-BLS consistently outperforms the standard BLS, while the horse_colic dataset represents a complementary scenario where BLS exhibits slightly better performance on the clean data. This selection enables a balanced and unbiased assessment of robustness behavior. As reported in Table II, Wave-BLS demonstrates clear robustness advantages on the blood dataset across all contamination levels. Under both outlier and noise perturbations, Wave-BLS maintains higher classification accuracy than BLS, with performance degradation occurring at a noticeably slower rate as contamination increases. This behavior highlights the effectiveness of the wave loss in bounding the influence of corrupted samples and mitigating the impact of extreme deviations. For the horse_colic dataset, although BLS achieves marginally higher accuracy on the clean data, Wave-BLS consistently surpasses BLS once outliers or noise are introduced. As contamination levels increase, Wave-BLS preserves stronger performance stability, leading to higher average accuracy under both outlier and noise settings. This indicates that even in cases where the baseline model performs well on clean data, the proposed wave loss enhances resilience under adverse conditions. Overall, Wave-BLS attains superior average accuracy and lower average rank across both datasets and contamination types. These results confirm that the proposed Wave-BLS framework offers improved robustness to data corruption, making it more suitable for real-world learning scenarios where noise and outliers are unavoidable.

V. Conclusions

In this work, we proposed Wave-BLS, a robust Broad Learning System that integrates the wave loss function into the BLS framework to address its inherent sensitivity to noise and outliers. By replacing the least-squares error formulation with a boundedand and asymmetric loss and optimizing the resulting objective via NAG-based scheme, the proposed approach eliminates the need for matrix inversion while preserving the flat and efficient architecture of BLS. Extensive experiments on several benchmark datasets show that Wave-BLS achieves consistently higher classification performance than baseline variants. The superiority of the proposed model is further supported by Friedman and Nemenyi statistical tests. These results indicate that replacing the squared-error criterion with a bounded loss can substantially improve the reliability of BLS-based models. Robustness studies conducted under controlled noise and outlier injection provide additional insight into the behavior of Wave-BLS in adverse learning conditions. While standard BLS may perform competitively on clean data, its performance degrades rapidly as contamination increases. Moreover, sensitivity analysis with respect to the wave loss hyperparameters reveals a wide region of stable performance, indicating that Wave-BLS does not depend on finely tuned parameter settings. Overall, the empirical evidence suggests that Wave-BLS offers a robust enhancement to the BLS framework without introducing architectural complexity.

References

[1] L. Zhang and P. N. Suganthan, "A survey of randomized algorithms for training neural networks," *Information Sciences*, vol. 364, pp. 146–155, 2016.

[2] A. K. Malik *et al.*, "Random vector functional link network: recent developments, applications, and future directions," *Applied Soft Computing*, vol. 143, p. 110377, 2023.

[3] M. Akhtar, A. Kumari, M. Sajid, A. Quadir, M. Arshad, P. Suganthan, and M. Tanveer, "Towards robust and inversion-free randomized neural networks: The XG-RVFL framework," *Pattern Recognition*, vol. 172, p. 112711, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320325013743

[4] C. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 1, pp. 10–24, 2017.

[5] ——, "Broad learning system: A new learning paradigm and system without going deep," in *2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE, 2017, pp. 1271–1276.

[6] J.-W. Jin and C. P. Chen, "Regularized robust broad learning system for uncertain data modeling," *Neurocomputing*, vol. 322, pp. 58–69, 2018.

[7] S. Feng and C. P. Chen, "Fuzzy broad learning system: A novel neuro-fuzzy model for regression and classification," *IEEE Transactions on cybernetics*, vol. 50, no. 2, pp. 414–424, 2018.

[8] W. Deng *et al.*, "Robust dual-model collaborative broad learning system for classification under label noise environments," *IEEE Internet of Things Journal*, vol. 12, no. 12, pp. 21 055–21 067, 2025.

[9] W. Chen *et al.*, "Adaptive broad network with graph-fuzzy embedding for imbalanced noise data," *IEEE Transactions on Fuzzy Systems*, vol. 33, no. 6, pp. 1949–1962, 2025.

[10] X.-K. Cao *et al.*, "Multiparty secure broad learning system for privacy preserving," *IEEE Transactions on Cybernetics*, vol. 53, no. 10, pp. 6636–6648, 2023.

[11] S. Wu *et al.*, "Fractional approximation of broad learning system," *IEEE Transactions on Cybernetics*, vol. 54, no. 2, pp. 811–824, 2024.

[12] M. Akhtar, M. Tanveer, and M. Arshad, "CAWI: Copula-aligned weight initialization for randomized neural networks," in *The 29th International Conference on Artificial Intelligence and Statistics*, 2026. [Online]. Available: https://openreview.net/forum?id=fwqydIjLn6

[13] Y. Tian *et al.*, "Recent advances on loss functions in deep learning for computer vision," *Neurocomputing*, vol. 497, pp. 129–158, 2022.

[14] M. Akhtar *et al.*, "Advancing supervised learning with the wave loss function: A robust and smooth approach," *Pattern Recognition*, p. 110637, 2024.

[15] Y.-H. Pao *et al.*, "Learning and generalization characteristics of the random vector functional-link net," *Neurocomputing*, vol. 6, no. 2, pp. 163–180, 1994.

[16] G. B. Huang *et al.*, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[17] M. Sajid *et al.*, "Wave-RVFL: A randomized neural network based on wave loss function," in *Neural Information Processing*. Springer Nature Singapore, 2025, pp. 242–257.

[18] ——, "Intuitionistic fuzzy broad learning system: Enhancing robustness against noise and outliers," *IEEE Transactions on Fuzzy Systems*, 2024.