

Feature Selection for Autism Spectrum Disorder via a Multi-Pack Cooperative Grey Wolf Optimization Framework

Alicia Fei *

Mira Costa High School, Manhattan Beach, United States

* Corresponding Author Email: aliciafei2011@gmail.com

Abstract. Autism Spectrum Disorder (ASD) is a prevalent neurodevelopmental condition in childhood and adolescence. However, its heterogeneous genetic mechanisms make early diagnosis particularly challenging. Identifying robust molecular biomarkers from high-dimensional gene expression data remains a critical bottleneck. In this study, we proposed the Multi-pack cooperative grey wolf optimization (MPC-GWO) algorithm to select robust gene biomarkers for ASD using two public blood-based transcriptomic datasets (GSE25507 for training, GSE18123 for independent testing). MPC-GWO was applied to GSE25507 to identify the minimal gene subset that reliably discriminated ASD from controls. The selected biomarker genes were then evaluated on GSE18123 to test cross-dataset generalizability. After statistical preselection and MPC-GWO refinement, we identified a 24-gene signature that achieved superior classification performance on the discovery cohort (accuracy=0.808, AUC=0.823), outperforming both the all-features baseline (accuracy=0.713) and filter-based t-test selection (accuracy=0.678). Compared with LASSO (accuracy=0.732, AUC=0.756) and Random Forest (accuracy=0.678, AUC=0.737), MPC-GWO demonstrated superior classification performance. The algorithm reduced the feature set by >88% and converged within 50 iterations. On the independent cohort, the signature achieved accuracy=0.674 and AUC=0.733. Functional enrichment analysis revealed that the selected genes are strongly associated with nervous system development, Ig-like C2-type, neurodevelopmental disorders, and extracellular space, most of which have been repeatedly implicated in ASD. These results demonstrate that MPC-GWO is effective for ASD biomarker discovery. The identified gene signature showed improved classification accuracy, and its enriched biological functions provided insights into ASD-related molecular mechanisms, suggesting potential value for future ASD-related genomic research and non-invasive diagnostic exploration.

Keywords: Autism Spectrum Disorder (ASD), Gene Expression Feature Selection, Multi-Pack Cooperative Grey Wolf Optimization (MPC-GWO), Metaheuristic Optimization Algorithm, Biomarker Discovery.

1. Introduction

Feature selection remains a central challenge in modern machine learning, particularly for high-dimensional settings such as transcriptomic profiles with over 20,000 genes, where exhaustive screening for disease-related signatures like autism is impractical [1,2]. Classical approaches including filter, wrapper, and embedded methods [3-5] often struggle with strong feature dependencies, non-linear interactions, or limited computational budgets [6]. Nature-inspired metaheuristics like Genetic Algorithms, PSO, and Grey Wolf Optimization (GWO) offer scalable alternatives by balancing exploration and exploitation [7]; however, standard single-population GWO can suffer from reduced diversity and premature convergence.

To address discrete feature selection, recent GWO variants incorporate binary encodings, such as improved binary GWO approaches and adaptive mechanism-based GWO (AMGWO) that uses self-adaptive mutation and fitness-distance balancing to mitigate premature convergence [8]. Despite these advances, gaps remain in GWO variants: premature convergence persists in high-dimensional settings [9], single-pack search limits solution space coverage [10], fixed $\alpha/\beta/\delta$ leadership reduces adaptability [11], and binarization distorts continuous update dynamics [12]. While IGWO and AMGWO introduce useful improvements, they do not fundamentally redesign the search architecture

around multi-population collaboration, motivating a fully cooperative multi-pack framework that explicitly promotes diversity and structured inter-pack communication.

This paper introduces Multi-pack Cooperative Grey Wolf Optimization (MPC-GWO), a binary feature-selection method using parallel multi-pack searches, periodic elite solution exchange, and dynamic leadership updates. These mechanisms enable compact gene set selection from blood samples while maintaining classification performance through enhanced global search and controlled information sharing.

2. Materials and Methods

2.1. Datasets

Two public microarray datasets were used to evaluate MPC-GWO: GSE25507 [13] (82 ASD, 64 control) and GSE18123 [14] (170 ASD, 115 control), both profiled on the Affymetrix U133 Plus 2.0 Array with 54,675 gene expression features per sample. The high-dimensional nature (54,675 features vs. limited samples) motivates effective feature selection to improve classification performance and interpretability.

2.2. Problem Definition

Gene expression datasets typically contain tens of thousands of features per sample, but only a small subset is biologically relevant for discriminating ASD from controls [15,16]. Including all features leads to high computational costs and degraded predictive performance due to redundancy, noise, and overfitting [17]. Therefore, this study aims to identify a compact, informative gene subset that achieves high classification accuracy while substantially reducing dimensionality, balancing predictive performance against model efficiency and interpretability.

Formally, let

$$D = \{(x_i, y_i)\}_{i=1}^n \quad (1)$$

where $x_i \in \mathbb{R}^d$ denotes the gene expression profile of the i^{th} sample with d gene features, and $y_i \in Y$ denotes the corresponding class label. Let the full feature set be defined as

$$F = \{f_1, f_2, \dots, f_d\}. \quad (2)$$

The goal is to determine an optimal subset ($S \subseteq F$) that provides an effective trade-off between classification accuracy and subset compactness.

This trade-off is formulated as a single-objective optimization problem:

$$J(S) = \alpha \cdot (1 - \text{Acc}(S)) + (1 - \alpha) \cdot \frac{|S|}{|F|}, \quad (3)$$

where $\text{Acc}(S)$ denotes the cross-validation classification accuracy obtained using the selected feature subset S , $|S|$ is the number of selected features, $|F|$ is the total number of available features, and $\alpha \in [0, 1]$ controls the relative importance of predictive accuracy versus subset size.

In summary, this work addresses the above optimization problem using a MPC-GWO framework, with the aim of identifying compact, discriminative gene subsets that support accurate classification of ASD while addressing the challenges posed by high-dimensional gene expression data.

2.3. Overview of the MPC-GWO Framework

Three-stage framework progressively reduces dimensionality while preserving discriminative information and ensuring robustness.

In **Stage 1** (statistical feature pre-selection), a t-test identifies genes with significant differential expression between classes, removing non-informative features and reducing the search space from

tens of thousands to 200 candidates per dataset, substantially improving computational efficiency for subsequent optimization.

In **Stage 2** (optimization-based feature selection using MPC-GWO), the pre-selected genes are refined using a multi-population extension of standard GWO that overcomes single-population limitations such as limited diversity and premature convergence. Multiple grey wolf packs evolve in parallel, exploring distinct regions of the solution space, with periodic elite migration to encourage information exchange. Each wolf encodes a binary feature selection vector, evaluated by a composite objective balancing classification accuracy (via KNN with 10-fold CV) and feature compactness. Experimental results show that MPC-GWO effectively identifies compact and highly discriminative gene subsets.

In **Stage 3** (cross-dataset validation and robustness assessment), feature subsets optimized on the discovery dataset (GSE25507) are directly evaluated on an independent dataset (GSE18123) without retraining. As MPC-GWO is inherently stochastic, independent test set validation is critical to ensure that performance reflects genuine predictive generalization rather than dataset-specific overfitting or favorable random initialization. Multiple random runs are also conducted to evaluate robustness and consistency.

2.4. Summary of Methodology

The proposed framework consists of three sequential stages: (1) Statistical feature preselection using an independent t-test to eliminate non-informative genes and reduce dimensionality; (2) Metaheuristic feature optimization via MPC-GWO, which leverages multi-population co-evolution and periodic elite migration to enhance diversity and avoid premature convergence; (3) Cross-dataset validation to assess robustness and generalization. Metaheuristic algorithms like MPC-GWO are inherently stochastic due to random initialization, probabilistic updates, and leader selection [18]. Consequently, repeated runs do not guarantee identical feature subsets, though predictive performance remains comparable—a fundamental property that enhances exploration and mitigates local optima convergence [19-21].

Algorithm: Multi-Pack Cooperative Grey Wolf Optimization (MPC-GWO)

Input: Pre-selected feature matrix $X \in \mathbb{R}^{n \times d}$, class labels y , number of grey wolf populations G , wolves per population N , maximum iterations T , migration interval M , trade-off parameter α .

Output: Optimal feature subset S^*

Begin

1. Initialize G independent populations $\{P_1, \dots, P_G\}$
2. For each population $p \in \{P_1, \dots, P_G\}$:
 - Initialize N wolves as binary vectors of length d
 - Evaluate fitness $\text{Fit}(s)$ for each wolf using cross-validated classifier
3. For $t = 1$ to T :
 - For each population p :
 - Identify $\alpha_p, \beta_p, \delta_p$ wolves (best three solutions)
 - For each remaining wolf i in the population p :
 - Update continuous position using standard GWO equations
 - Apply sigmoid/threshold function to obtain binary position
 - Evaluate fitness $\text{Fit}(S_i)$
 - Update leadership roles dynamically based on fitness
 - If $t \bmod M == 0$:
 - Select elite wolves from each population
 - Migrate elites to neighboring population and replace worst solutions
4. Collect all α_p wolves across populations
5. Select global best solution S^* with minimum objective $\text{Fit}(S^*)$.

Return S^* .

End.

Figure 1. Pseudocode of the MPC-GWO feature selection algorithm for gene expression datasets.

Given this intrinsic randomness [22], evaluation based solely on training data is insufficient; independent test sets or cross-dataset validation are instead used to assess generalization and robustness of selected biomarker panels. Feature-level consistency across stochastic runs is not required, and averaging discrete feature subsets is generally not meaningful [23]. Contemporary studies therefore emphasize external validation of representative or best-performing solutions over enforcing deterministic feature reproducibility [24].

2.5. Experimental Settings

The experimental framework illustrated in Figure 2 was designed to evaluate MPC-GWO for feature selection, following a unified pipeline of data preprocessing, MPC-GWO-based feature subset optimization, classifier training, and performance evaluation. In MPC-GWO, multiple grey wolf populations operate in parallel, each maintaining a hierarchical structure (one α , one β , one δ , and two ω wolves). Each wolf encodes a candidate feature subset via binary representation. Intra-population search is guided by the standard GWO leadership mechanism, while inter-population cooperation is achieved through periodic elite exchange. The optimization objective jointly considers classification performance and feature subset compactness. Four performance metrics were used: accuracy (Acc), area under the ROC curve (AUC), sensitivity (Sn), and specificity (Sp), as defined in [25,26].

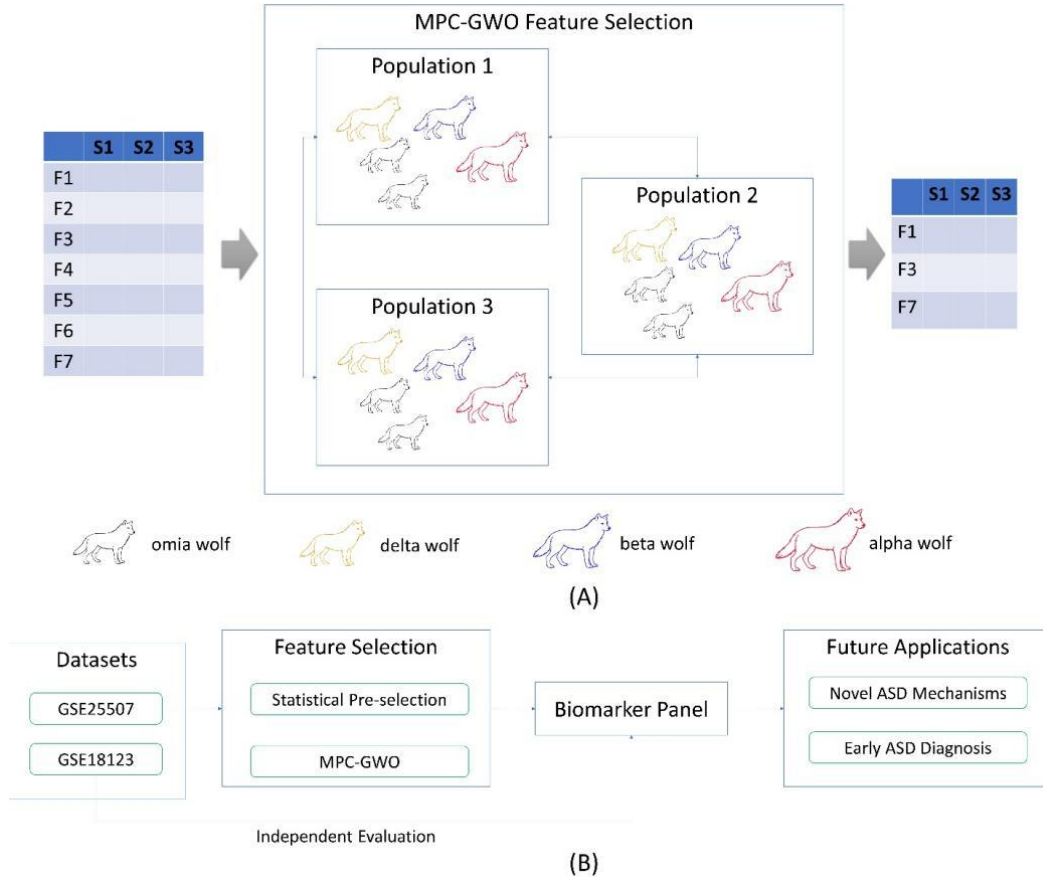


Figure 2. MPC-GWO framework: (A) architecture; (B) workflow. Wolf icon via ChatGPT; all other elements by authors. S1/S2... = samples; F1/F2... = features.

3. Results and Discussion

The MPC-GWO framework demonstrated a consistent improvement in classification accuracy over the course of the optimization process. During the early iterations, accuracy values exhibited noticeable fluctuations, and the observations show extensive exploration of the high-dimensional feature space. As the optimization progressed, the algorithm gradually stabilized and converged to an optimal feature subset, with convergence typically achieved within approximately 50 iterations.

3.1. Model Optimization

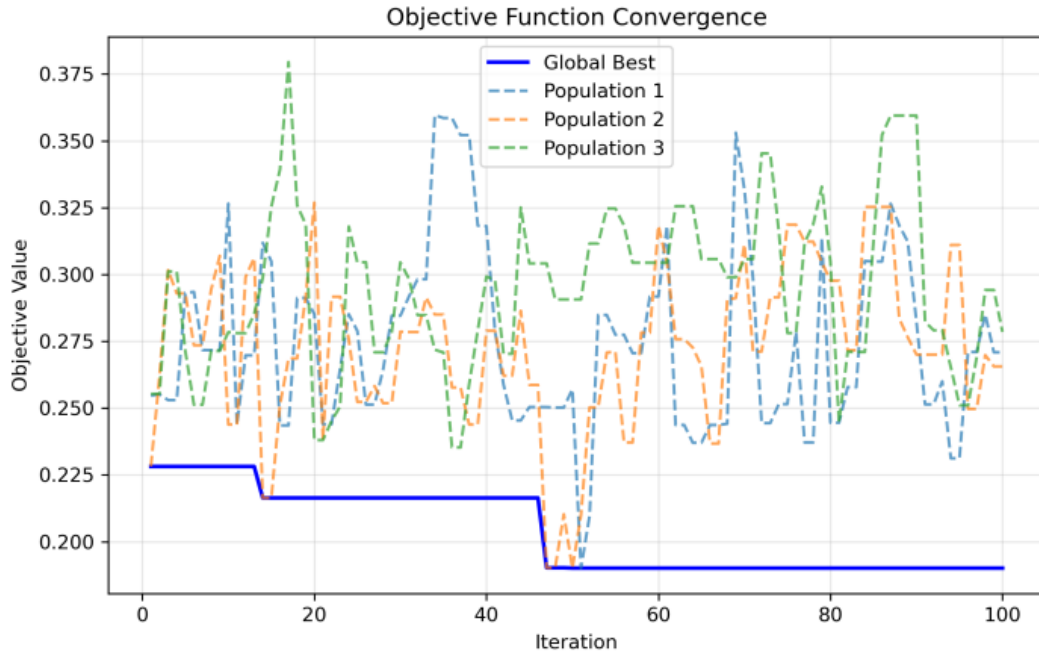


Figure 3. MPC-GWO convergence on GSE25507: best fitness vs. iterations. Three cooperative wolf populations jointly explore feature space for optimal gene subset.

The optimization results show that the MPC-GWO model can quickly find useful feature combinations for distinguishing individuals with ASD from controls. As the model runs, its performance improves rapidly at the beginning, which means it is able to explore many possible gene combinations efficiently. After several rounds of optimization, the model reaches a stable solution, indicating that it has identified a reliable set of genes and no longer needs major adjustments.

More importantly, MPC-GWO improves classification accuracy while significantly reducing the number of genes used. After narrowing the feature space to 200 candidate genes, the model ultimately selects only 24 genes—12.0% of the candidates—achieving an accuracy of 0.808 (AUC = 0.823, Sn = 0.842, Sp = 0.766). This is noticeably higher than the 0.713 accuracy obtained using all features, indicating that many genes contribute little or even interfere with prediction. By focusing on a compact and informative gene set, MPC-GWO builds a simpler, more effective model, demonstrating its ability to identify potential genetic markers for autism.

3.2. Convergence Behavior Analysis

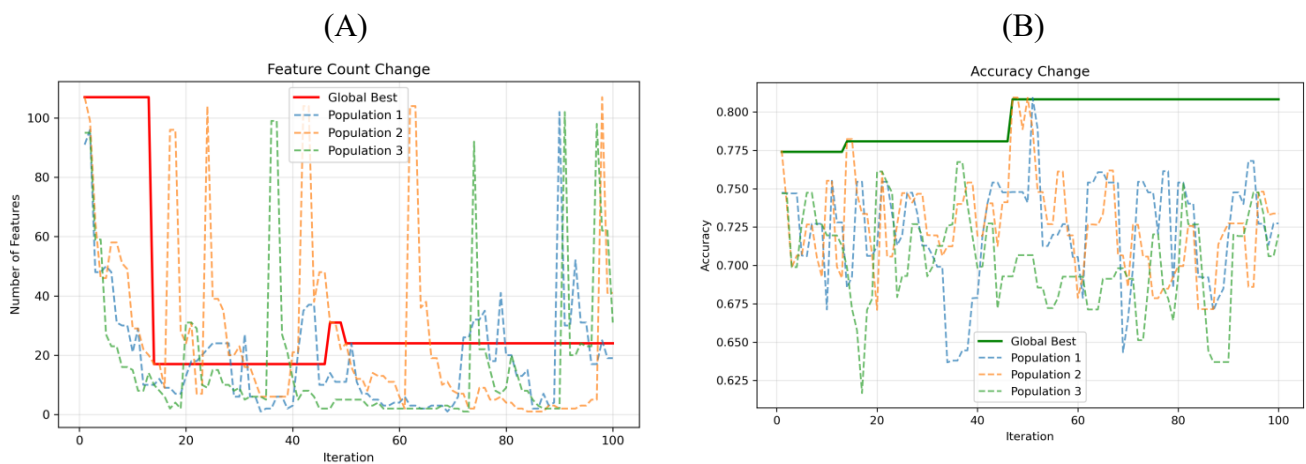


Figure 4. MPC-GWO convergence on GSE25507: (A) subpopulation fitness trends; (B) feature count & accuracy stability over iterations. Three cooperative wolf populations identify the optimal feature subset.

Figure 4 (A) shows that the three subpopulations exhibit distinct search behaviors in the early optimization stages, with diverging curves that indicate exploration from different directions. This diversity expands the search space and reduces the risk of missing good solutions, helping MPC-GWO avoid premature convergence. As optimization proceeds, the global solution curves gradually converge, suggesting effective information sharing across subpopulations. Figure 4 (B) further shows that later iterations become stable, with only minor changes in accuracy and feature count. This smooth, steady behavior demonstrates that MPC-GWO reliably reaches a final solution without large fluctuations or performance loss, even with complex high-dimensional gene expression data, ensuring the selected gene subset is robust for downstream classification.

3.3. Comparison with Single-Pack GWO

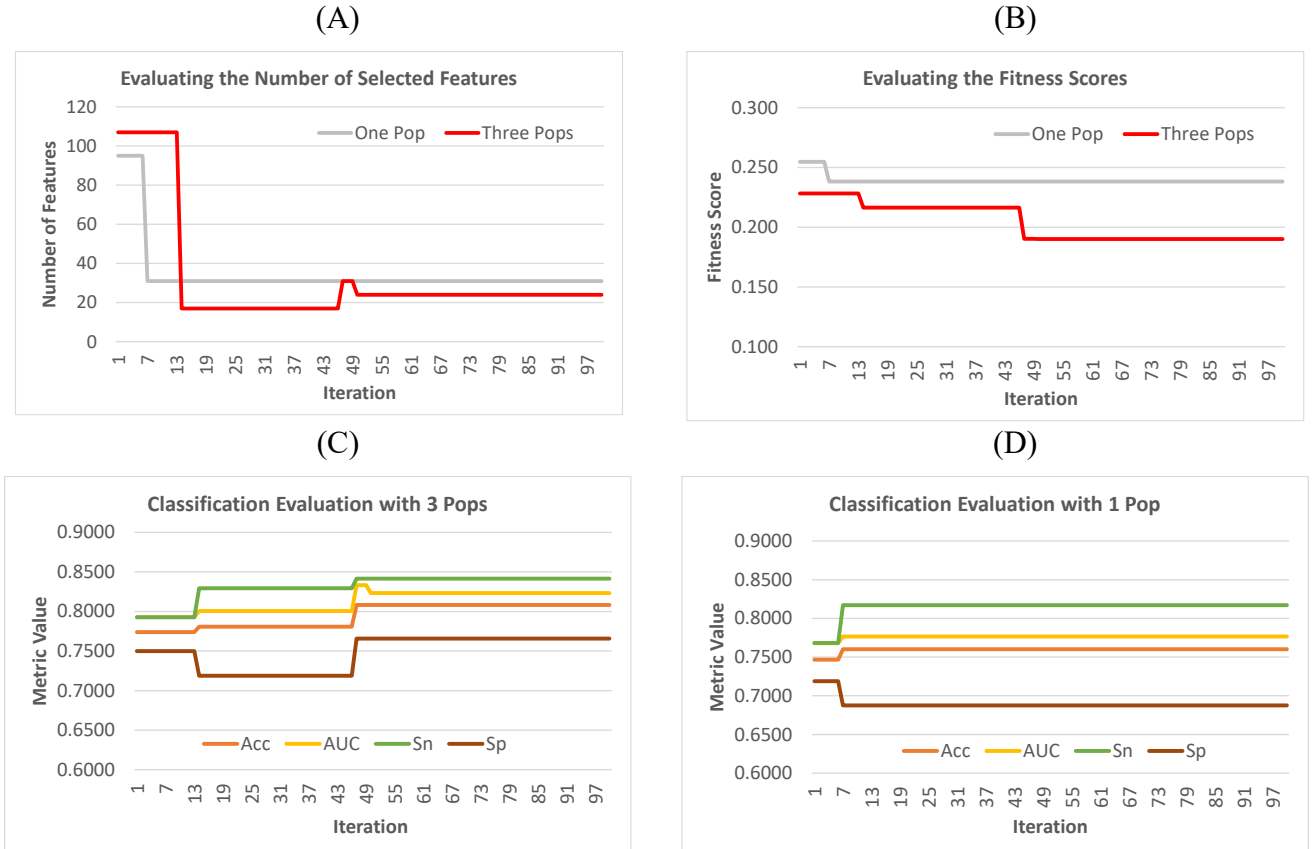


Figure 5. MPC-GWO vs. single-pack GWO on GSE25507: (A) feature count; (B) fitness score; (C) three-pack performance; (D) single-pack performance.

Figure 5 (A) and (B) show that the three-pack GWO exhibits more effective search behavior than the single-pack variant. The three-pack strategy rapidly reduces the number of selected features from 200 to 24 and maintains a stable subset size in later iterations, whereas the single-pack GWO converges earlier to a larger subset of 31 features. Moreover, the three-pack GWO achieves a continuously improving fitness score of 0.190, outperforming the single-pack fitness of 0.238. These results indicate that the multi-pack mechanism enhances the balance between exploration and exploitation, helping the algorithm escape local optima and achieve superior optimization performance.

As shown in Figure 5 (C) and (D), the improved search capability of the three-pack GWO translates directly into better classification performance. It achieves higher values across all four metrics: Acc (0.808 vs. 0.760), AUC (0.823 vs. 0.777), Sn (0.842 vs. 0.817), and Sp (0.766 vs. 0.688). These consistent improvements indicate that the selected feature subset is more discriminative and robust. Overall, the results confirm that the three-pack strategy enhances both optimization quality and downstream classification reliability on the GSE25507 dataset.

3.4. Performance Comparison of T-test selected Features

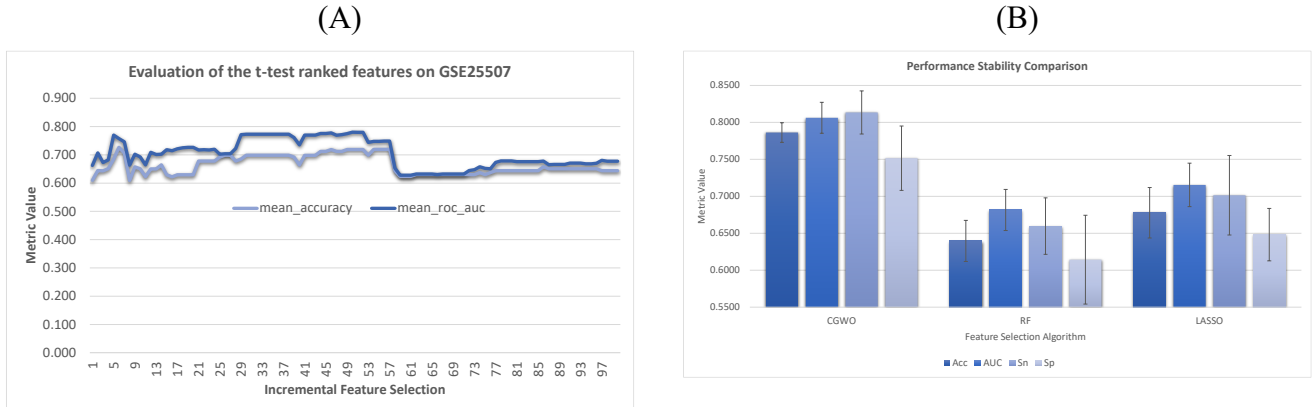


Figure 6. MPC-GWO vs. other methods on GSE25507: (A) accuracy vs. #t-test-ranked features (ascending p-value; lower p = stronger ASD association). (B) Comparison with RF and LASSO on Acc, AUC, Sn, Sp.

Figure 6 (A) shows how classification accuracy changes as features ranked by the t-test are gradually added. Accuracy generally improves with more features, with the fastest gains seen when using only a few top-ranked genes, indicating that highly ranked genes are informative for distinguishing autism from control samples. However, overall performance remains limited. With 24 genes (the number selected by MPC-GWO), the t-test method achieves only Acc = 0.678 and AUC = 0.720. Even at its best point, it reaches Acc = 0.719 and AUC = 0.780 but requires at least 50 features—still below MPC-GWO’s performance. As lower-ranked features are added, accuracy fluctuates and stops improving, suggesting noise or redundancy. These results highlight a key limitation of univariate filter methods like the t-test: they evaluate each feature independently, ignoring interactions between features. In high-dimensional gene expression data, such interactions are often important. Therefore, multivariate optimization-based methods like MPC-GWO are better suited for selecting compact, effective feature subsets, as they account for feature dependencies and overall model performance.

Figure 6 (B) shows that MPC-GWO outperforms RF and LASSO across all evaluated metrics, achieving the highest mean accuracy (0.808) and AUC (0.823), along with improved sensitivity (0.842) and specificity (0.766). RF achieves moderate performance (Acc = 0.678, AUC = 0.737, Sn = 0.706, Sp = 0.642), while LASSO shows lower discriminative ability (Acc = 0.732, AUC = 0.756, Sn = 0.802, Sp = 0.642), possibly due to excessive sparsity that overlooks informative feature interactions. Across 20 random runs, MPC-GWO selects the smallest average number of genes (29.4), fewer than RF (75.9) and LASSO (45.1), demonstrating a balanced trade-off between compactness and predictive power. Despite its stochastic nature and longer runtime (≈ 75.4 s), MPC-GWO exhibits stable performance with lower AUC variability (Std = 0.021) compared to RF (0.028) and LASSO (0.029). Overall, the cooperative multi-population strategy enables MPC-GWO to explore the combinatorial feature space more effectively and identify robust, discriminative gene subsets for ASD classification.

3.5. Independent Evaluation of the ASD Detection Model



Figure 7. 24-gene ASD signature from GSE25507 was validated on GSE18123 without retraining.

Figure 7 illustrates that the 24-gene signature optimized on GSE25507 achieved an accuracy of 0.808 and AUC of 0.823. Since MPC-GWO is a stochastic optimizer, external validation is necessary to rule out performance due to random chance. When applied directly to the independent GSE18123 dataset without retraining, the model maintained an accuracy of 0.674 and AUC of 0.733 (sensitivity = 0.610, specificity = 0.733), still well above random expectation. This confirms that MPC-GWO captures reproducible biological signals that generalize beyond the discovery dataset.

3.6. Permutation-Based Significance Assessment of MPC-GWO Classification Performance

We conducted a permutation test to evaluate whether the MPC-GWO classifier achieved performance beyond chance. Using the original labels, the MPC-GWO model achieved the performance of Acc=0.808 and AUC=0.823 with a runtime of **52.54 seconds**. We then randomly permuted the labels 1000 times while keeping the feature matrix unchanged and reran the full MPC-GWO pipeline. Across all 1000 permutations, 0 runs produced the performance greater than or equal to Acc=0.808 and AUC=0.823. The empirical p-value was therefore 0.001.

These results indicate that the observed predictive performance is highly unlikely to arise from random label structure under the current modeling pipeline. The large separation between the observed AUC (0.823) and the null distribution supports the statistical significance of the MPC-GWO-based feature selection and classification framework. The use of permutation-based evaluation aligns with recent recommendations for robust validation of machine learning models in biomedicine discussed in [27].

3.7. Enriched Functions in the ASD Biomarker Panel

Table 1. Significantly enriched GO molecular functions of the 24-gene ASD panel via Fisher’s exact test using DAVID [28].

Category	Term	Count	Fold Enrichment	Fisher Exact
GOTERM_BP_DIRECT	nervous system development	4	6.68	2.56E-03
UP_SEQ_FEATURE	DOMAIN:Ig-like C2-type	2	21.67	3.78E-03
DISGENET	Neurodevelopmental Disorders	2	20.85	3.90E-03
GOTERM_CC_DIRECT	extracellular space	5	3.05	1.95E-02

The functional enrichment results for the 24-gene biomarker panel point to a coherent neurodevelopmental signal. The most significant biological process term, nervous system development (fold enrichment = 6.68), aligns with convergent evidence that ASD risk genes and ASD-associated transcriptional dysregulation concentrate in pathways governing neuronal differentiation, circuit assembly, and synaptic maturation [29]. The DISGENET [30] Neurodevelopmental Disorders category further supports that the selected genes overlap with loci previously implicated in neurodevelopmental phenotypes, consistent with the broad genetic heterogeneity of ASD but with convergence on shared developmental mechanisms [31]. Methodologically, these enrichments were derived using DAVID with Fisher’s exact test, which is a standard workflow for hypothesis generation from compact gene lists [28].

At the molecular and cellular component levels, enrichment of Ig-like C2-type domains and extracellular space suggests a prominence of cell-surface and extracellular signaling components that can mediate axon guidance, synapse formation, and neuron–glia interactions. C2-set immunoglobulin-like domains typify immunoglobulin superfamily cell-adhesion molecules such as contactins. Although CASPR/CNTNAP proteins are structurally distinct members of the neuroligin superfamily, CNTNAP2 exhibits strong genetic and mechanistic associations with ASD-relevant neurodevelopmental processes [32]. The extracellular space term is also compatible with growing evidence that extracellular matrix and extracellular milieu regulation influence synaptic stabilization and circuit plasticity, and have been implicated in ASD-relevant models and human tissue profiling

[33]. All these enrichments support an interpretation in which the panel captures neurodevelopmental biology with emphasis on cell-adhesion and extracellular mechanisms that are plausibly linked to ASD pathophysiology.

3.8. Biological Association of the ASD Biomarkers

Table 2. 24-gene ASD panel: gene symbols, names, PubMed ASD record counts, and key supporting references.

Feature	Gene	Title	Literature	Example PMID
201773_at	ADNP	activity-dependent neuroprotector homeobox	144	[31]
1556382_a_at	NAA15	N(alpha)-acetyltransferase 15, NatA auxiliary subunit	10	[34]
208487_at	LMX1B	LIM homeobox transcription factor 1, beta	6	[35]
37950_at	PREP	prolyl endopeptidase	6	[36]
46665_at	SEMA4C	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4C	1	[37]
219594_at	NINJ2	ninjurin 2	1	[38]
225142_at	KDM7A	lysine (K)-specific demethylase 7A	1	[39]

Analysis of the 24-gene biomarker panel highlighted seven genes with prior literature evidence linking them to ASD or broader neurodevelopmental phenotypes (Table 2). Among these, ADNP shows the strongest and most established association, with extensive PubMed support (144 records) and representative genetic evidence consistent with its recognized role in syndromic neurodevelopmental disorders [31]. Several additional genes in the panel have moderate but recurrent ASD-relevant evidence, including NAA15 (10 records) and LMX1B (6 records). These genes are plausibly connected to core developmental processes through regulation of transcriptional programs and protein modification, suggesting that the panel captures upstream regulatory components that can influence neuronal differentiation and maturation. Likewise, PREP (6 records), a peptidase implicated in neurocognitive phenotypes, provides a mechanistic link to neural function that is compatible with ASD-related alterations in neurodevelopment and cognition [36].

In contrast, SEMA4C, NINJ2, and KDM7A currently have limited direct ASD literature support (each with 1 record), but their functional annotations suggest biologically interpretable connections to ASD-relevant pathways rather than incidental findings. SEMA4C is a semaphorin-family transmembrane signaling molecule containing immunoglobulin-related domains, aligning with established roles of semaphorin-mediated guidance cues in neural circuit formation and synaptic organization, which are repeatedly implicated in ASD models [37]. NINJ2, a cell-surface adhesion-related molecule, may similarly reflect processes in neurite outgrowth, axonal interaction, or neuroinflammatory interface mechanisms [38]. KDM7A is a histone demethylase, indicating a potential epigenetic regulatory axis. Epigenetic modulation is a recurrent theme in ASD biology and can plausibly influence neurodevelopmental gene expression programs even when direct gene-level ASD evidence remains sparse [39].

Collectively, Table 2 suggests that the 24-gene panel spans both (i) genes with established ASD relevance (particularly ADNP, and to a lesser extent NAA15, LMX1B, and PREP) and (ii) genes that are more appropriately interpreted as pathway-level contributors or candidate modulators (e.g., SEMA4C, NINJ2, KDM7A) that warrant further validation. This mixture is consistent with the genetic and mechanistic heterogeneity of ASD, where robust predictive signatures may include both direct risk genes and genes positioned in convergent neurodevelopmental pathways such as transcriptional control, epigenetic regulation, and cell-surface guidance/adhesion signaling.

4. Conclusion

This study proposes a cooperative multi-population Grey Wolf Optimization (MPC-GWO) framework for feature selection in high-dimensional biological data. By enabling information exchange across populations via periodic elite migration, MPC-GWO prevents premature convergence and improves exploration. Applied to ASD transcriptomic data, it reduced features from 54,675 probes to a 24-gene signature, improving classification performance (Acc=0.808, AUC=0.823) on the discovery cohort and retaining predictive signal in an independent cohort (Acc=0.674, AUC=0.733). Functional analysis linked selected genes to ASD-related pathways. Robustness was confirmed via independent validation, repeated runs, and a label permutation test (empirical $p=0.001$).

Limitations include higher computational cost, sensitivity to hyperparameters, binary discretization loss, and run-to-run variability—emphasizing statistical validation over strict feature reproducibility. Future work will focus on larger heterogeneous cohorts, multi-modal data integration, and broader biomedical applications beyond ASD to establish MPC-GWO as a scalable optimization-driven feature selection framework.

References

- [1] G. Li, Y. Cui, and J. Su, “Adaptive mechanism-based grey wolf optimizer for feature selection in high-dimensional classification,” *PLOS One*, vol. 20, no. 5, p. e0318903, May 2025, doi: 10.1371/journal.pone.0318903.
- [2] Md. M. Rahman, O. L. Usman, R. C. Muniyandi, S. Sahran, S. Mohamed, and R. A. Razak, “A Review of Machine Learning Methods of Feature Selection and Classification for Autism Spectrum Disorder,” *Brain Sci.*, vol. 10, no. 12, p. 949, Dec. 2020, doi: 10.3390/brainsci10120949.
- [3] Q. Al-Tashi, S. J. Abdul Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, “Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection,” *IEEE Access*, vol. 7, pp. 39496–39508, 2019, doi: 10.1109/ACCESS.2019.2906757.
- [4] Y. Wang, Y. Yin, H. Zhao, J. Liu, C. Xu, and W. Dong, “Grey wolf optimizer with self-repulsion strategy for feature selection,” *Sci. Rep.*, vol. 15, no. 1, p. 12807, Apr. 2025, doi: 10.1038/s41598-025-97224-8.
- [5] T. R. Golub *et al.*, “Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring,” *Science*, vol. 286, no. 5439, pp. 531–537, Oct. 1999, doi: 10.1126/science.286.5439.531.
- [6] N. Nouri, G. Gaglia, H. Mattoo, E. De Rinaldis, and V. Savova, “GENIX enables comparative network analysis of single-cell RNA sequencing to reveal signatures of therapeutic interventions,” *Cell Rep. Methods*, vol. 4, no. 6, p. 100794, Jun. 2024, doi: 10.1016/j.crmeth.2024.100794.
- [7] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer,” *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.
- [8] J. Y. Khaseeb, A. Keshk, and A. Youssef, “Improved Binary Grey Wolf Optimization Approaches for Feature Selection Optimization,” *Appl. Sci.*, vol. 15, no. 2, p. 489, Jan. 2025, doi: 10.3390/app15020489.
- [9] J. Huang, X. Deng, and L. Hu, “Enhanced grey wolf optimizer with hybrid strategies for efficient feature selection in high-dimensional data,” *Inf. Sci.*, vol. 705, p. 121958, Jul. 2025, doi: 10.1016/j.ins.2025.121958.
- [10] Z. Wang, L. Shu, S. Yang, Z. Zeng, D. He, and S. Chan, “Multi-population dynamic grey wolf optimizer based on dimension learning and Laplace Mutation for global optimization,” *Expert Syst. Appl.*, vol. 265, p. 125863, Mar. 2025, doi: 10.1016/j.eswa.2024.125863.
- [11] A. Maleki, M. Roayaei, and S. Mirjalili, “Enhancing leadership-based metaheuristics using reinforcement learning: A case study in grey wolf optimizer,” *Knowl.-Based Syst.*, vol. 330, p. 114471, Nov. 2025, doi: 10.1016/j.knosys.2025.114471.
- [12] M. Banaie-Dezfouli, M. H. Nadimi-Shahraki, and Z. Beheshti, “BE-GWO: Binary extremum-based grey wolf optimizer for discrete optimization problems,” *Appl. Soft Comput.*, vol. 146, p. 110583, Oct. 2023, doi: 10.1016/j.asoc.2023.110583.

- [13] M. D. Alter *et al.*, “Autism and Increased Paternal Age Related Changes in Global Levels of Gene Expression Regulation,” *PLoS ONE*, vol. 6, no. 2, p. e16715, Feb. 2011, doi: 10.1371/journal.pone.0016715.
- [14] S. W. Kong *et al.*, “Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders,” *PloS One*, vol. 7, no. 12, p. e49475, 2012, doi: 10.1371/journal.pone.0049475.
- [15] I. Hussain, M. Qureshi, M. Ismail, H. Iftikhar, J. Zywołek, and J. L. López-Gonzales, “Optimal features selection in the high dimensional data based on robust technique: Application to different health database,” *Heliyon*, vol. 10, no. 17, p. e37241, Sep. 2024, doi: 10.1016/j.heliyon.2024.e37241.
- [16] S. Berto *et al.*, “Association between resting-state functional brain connectivity and gene expression is altered in autism spectrum disorder,” *Nat. Commun.*, vol. 13, no. 1, p. 3328, Jun. 2022, doi: 10.1038/s41467-022-31053-5.
- [17] Z. Chen *et al.*, “Feature selection may improve deep neural networks for the bioinformatics problems,” *Bioinformatics*, vol. 36, no. 5, pp. 1542–1552, Mar. 2020, doi: 10.1093/bioinformatics/btz763.
- [18] S. Nogueira, K. Sechidis, and G. Brown, “On the Stability of Feature Selection Algorithms,” *J. Mach. Learn. Res.*, vol. 18, no. 174, pp. 1–54, 2018.
- [19] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, Oct. 2007, doi: 10.1093/bioinformatics/btm344.
- [20] J. Li *et al.*, “Feature Selection: A Data Perspective,” *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–45, Nov. 2018, doi: 10.1145/3136625.
- [21] B. Xue, M. Zhang, W. N. Browne, and X. Yao, “A Survey on Evolutionary Computation Approaches to Feature Selection,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, pp. 606–626, Aug. 2016, doi: 10.1109/TEVC.2015.2504420.
- [22] B. H. Nguyen, B. Xue, and M. Zhang, “A survey on swarm intelligence approaches to feature selection in data mining,” *Swarm Evol. Comput.*, vol. 54, p. 100663, May 2020, doi: 10.1016/j.swevo.2020.100663.
- [23] U. M. Khaire and R. Dhanalakshmi, “Stability of feature selection algorithm: A review,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1060–1073, Apr. 2022, doi: 10.1016/j.jksuci.2019.06.012.
- [24] Y. Wu, H.-C. Huang, and L.-X. Qin, “Making External Validation Valid for Molecular Classifier Development,” *JCO Precis. Oncol.*, no. 5, pp. 1250–1258, Nov. 2021, doi: 10.1200/PO.21.00103.
- [25] K. Li *et al.*, “TaiChiNet: PCA-based Ying-Yang dilution of inter- and intra-BERT layers to represent anti-coronavirus peptides,” *Expert Syst. Appl.*, vol. 282, p. 127786, Jul. 2025, doi: 10.1016/j.eswa.2025.127786.
- [26] R. Zhang *et al.*, “PepHarmony: a multi-view contrastive learning framework for integrated sequence and structure-based peptide representation,” *Neural Netw.*, vol. 194, p. 108148, Feb. 2026, doi: 10.1016/j.neunet.2025.108148.
- [27] X. Mi, B. Zou, F. Zou, and J. Hu, “Permutation-based identification of important biomarkers for complex diseases via machine learning models,” *Nat. Commun.*, vol. 12, no. 1, p. 3008, May 2021, doi: 10.1038/s41467-021-22756-2.
- [28] B. T. Sherman *et al.*, “DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update),” *Nucleic Acids Res.*, vol. 50, no. W1, pp. W216–W221, Jul. 2022, doi: 10.1093/nar/gkac194.
- [29] H. R. Willsey, A. J. Willsey, B. Wang, and M. W. State, “Genomics, convergent neuroscience and progress in understanding autism spectrum disorder,” *Nat. Rev. Neurosci.*, vol. 23, no. 6, pp. 323–341, Jun. 2022, doi: 10.1038/s41583-022-00576-7.
- [30] J. Piñero *et al.*, “The DisGeNET knowledge platform for disease genomics: 2019 update,” *Nucleic Acids Res.*, p. gkz1021, Nov. 2019, doi: 10.1093/nar/gkz1021.
- [31] X. Zhou *et al.*, “Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes,” *Nat. Genet.*, vol. 54, no. 9, pp. 1305–1319, Sep. 2022, doi: 10.1038/s41588-022-01148-2.
- [32] Q. Zhang *et al.*, “Contactin-associated protein-like 2 (CNTNAP2) mutations impair the essential α -secretase cleavages, leading to autism-like phenotypes,” *Signal Transduct. Target. Ther.*, vol. 9, no. 1, p. 51, Mar. 2024, doi: 10.1038/s41392-024-01768-6.

- [33] L. E. Rexrode *et al.*, “Molecular profiling of the hippocampus of children with autism spectrum disorder,” *Mol. Psychiatry*, vol. 29, no. 7, pp. 1968–1979, Jul. 2024, doi: 10.1038/s41380-024-02441-8.
- [34] H. Cheng *et al.*, “Truncating Variants in NAA15 Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies,” *Am. J. Hum. Genet.*, vol. 102, no. 5, pp. 985–994, May 2018, doi: 10.1016/j.ajhg.2018.03.004.
- [35] A. R. Burmeister *et al.*, “Maternal Herpesviridae infection during pregnancy alters midbrain dopaminergic signatures in adult offspring,” *Neurobiol. Dis.*, vol. 169, p. 105720, Jul. 2022, doi: 10.1016/j.nbd.2022.105720.
- [36] E. Diez-Itza, A. Viejo, and M. Fernández-Urquiza, “Pragmatic Profiles of Adults with Fragile X Syndrome and Williams Syndrome,” *Brain Sci.*, vol. 12, no. 3, p. 385, Mar. 2022, doi: 10.3390/brainsci12030385.
- [37] A. Khan, P. Sharma, S. Dahiya, and B. Sharma, “Plexins: Navigating through the neural regulation and brain pathology,” *Neurosci. Biobehav. Rev.*, vol. 169, p. 105999, Feb. 2025, doi: 10.1016/j.neubiorev.2024.105999.
- [38] I. M. W. Silva, J. Rosenfeld, S. A. Antoniuk, S. Raskin, and V. S. Sotomaior, “A 1.5Mb terminal deletion of 12p associated with autism spectrum disorder,” *Gene*, vol. 542, no. 1, pp. 83–86, May 2014, doi: 10.1016/j.gene.2014.02.058.
- [39] T. Guo *et al.*, “Histone Modifier Genes Alter Conotruncal Heart Phenotypes in 22q11.2 Deletion Syndrome,” *Am. J. Hum. Genet.*, vol. 97, no. 6, pp. 869–877, Dec. 2015, doi: 10.1016/j.ajhg.2015.10.013.