In the following analysis, the time complexity of each model is described using these notations: n denotes the number of samples; d represents the total feature dimension after concatenation, where d_v is the dimension of the v-th view; V is the number of omics); c is the number of classes; I is the number of iterations; N_t is the number of trees in a random forest; d_{sub} is the feature subset dimension for a single tree; h_i refers to the dimension of the i-th hidden layer; k is the number of neighbors in a K-NN graph; $|E|$ is the number of edges in a graph; and B is the batch size.

- SVM (Support Vector Machine): utilizes the kernel trick to map concatenated multi-omics features into a high-dimensional space and finds a maximum-margin hyperplane for classification, with a training complexity of $O(n^2 \cdot d)$, making it highly sensitive to the sample size n.
- RF (Random Forest): classifies by aggregating votes from a multitude of decision trees built on the concatenated multi-omics features, featuring a training complexity of $O(N_t \cdot n \log(n) \cdot d_{sub})$, which is computationally efficient and easily parallelizable.
- DeepMO: first concatenates multi-omics features and then employs a Stacked Denoising Autoencoder (SDAE) for unsupervised feature learning and dimensionality reduction before classification, with a training complexity of approximately $O(I \cdot n \cdot \sum(h_i \cdot h_{i+1}))$, which is linear with respect to the sample size n.
- AMGL: constructs a K-NN graph for each view and then automatically learns weights to fuse them into a single consensus graph Laplacian matrix through iterative optimization, leading to a complexity of $O(I \cdot (V \cdot n^2 + n^2 \cdot c))$, which is prohibitively expensive due to its quadratic dependence on n.
- MVAR: jointly learns linear projection matrices for each view and an adaptive class regression target, with a complexity of $O(I \cdot (V \cdot d_v^3 + n \cdot d \cdot c))$; its primary bottleneck is the cubic dependence on feature dimension d_v from matrix inversion operations.
- MLAN: simultaneously learns a similarity matrix and a linear projection by adaptively identifying neighbors for each data point during optimization, resulting in a complexity of $O(I \cdot n^2 \cdot (d + c))$, which suffers from a quadratic dependence on sample size n due to the all-pairs similarity learning.
- MOGONET: trains a separate Graph Convolutional Network (GCN) for each omics-specific graph to learn view-exclusive node embeddings, which are then fused via a learnable View Correlation Discovery Network (VCDN), resulting in a complexity of $O(I \cdot V \cdot |E| \cdot h_0 \cdot h_1)$ or $O(I \cdot V \cdot n \cdot k \cdot h_0 \cdot h_1)$ for K-NN graphs, but its view-isolated computation limits early-stage cross-view interactions.
- MoGCN: first employs Similarity Network Fusion (SNF) to integrate multiple similarity networks (one for each omic) into a single, robust patient-to-patient graph, and then applies a GCN to this fused graph to learn embeddings for classification, with its complexity dominated by SNF at $O(V \cdot n^2 \cdot k + I_{snf} \cdot n^2 \cdot k)$ plus the GCN part at $O(I_{gcn} \cdot |E| \cdot h_0 \cdot h_1)$. The reliance of SNF on n^2 computations makes it less scalable than methods that avoid dense similarity fusion.
- SUPREME: follows a similar view-specific GCN paradigm but introduces a pre-training stage using Graph Autoencoders to initialize GCN weights, with the final features being fused through weighted averaging or concatenation; its complexity is $O(I_{pre} \cdot V \cdot |E| \cdot h_{in}^2 + I_{fine} \cdot V \cdot |E| \cdot h_{in} \cdot h_{out})$, where the computational bottleneck remains consistent with other GCN-based methods but with added pre-training costs.
- ERL-MVSC: (based on title inference) likely guides multi-view learning by designing a learnable regularization term to constrain the embedding spaces of different views, and its complexity is primarily determined by its backbone network, approximately $O(I \cdot n \cdot$ Complexity(Backbone)), which is on par with standard deep learning models.
- ECML: models the output of each view as "evidence" under Dempster-Shafer theory and leverages a contrastive loss to explicitly quantify and utilize inter-view conflict and consensus for a reliable weighted fusion, with a complexity of $O(I \cdot n \cdot$Complexity(Backbone)$+ I \cdot B \cdot V^2)$, adding a negligible contrastive loss computation overhead to a standard deep learning pipeline.
- **MOKAN:** The second sentence describes the model's complexity using the defined symbols. The time complexity of our model scales linearly with the number of samples n within each training epoch, specifically $\mathcal{O}(n \cdot G \cdot (d \cdot h + h \cdot c))$. The key differentiator from standard MLPs is the linear dependence on the grid size G, which impacts both the computational time and the parameter space.

In summary, our KAN model presents a trade-off. It maintains linear scaling with sample size N per epoch (unlike non-linear SVMs or potentially dense GCNs) but introduces an additional computational factor G compared to standard MLPs.

TABLE I
ALGORITHM PARAMETER CONFIGURATIONS

Algorithm	Key Parameter Configuration
SVM	The Radial Basis Function (RBF) kernel was used, and the tolerance parameter `tol` (tolerance for the stopping criterion) was set to 0.1.
RF	The number of decision trees used in RF was tuned to be 20.
AMGL	The AMGL method is designed to be largely parameter-free, particularly concerning the fusion of multiple views. It employs an iterative optimization process.
MVAR	In the MVAR algorithm, the importance weights (α) for different data views are automatically learned and adjusted during model training. The parameter `r` was fixed at 2, and regularization λ was fixed at 1000.
MLAN	The number of adaptive neighbours, k, used for constructing the similarity matrix S, was fixed at $k = 9$. The maximum number of iterations was 30. View weights (`Wv`) were automatically learned, and trade-off λ was fixed at 25.
DeepMo	All encoders and the classifier were optimized using the Adagrad optimizer with a learning rate of 0.01. The classifier employed a weight decay of 0.001, encoder hidden dimension of 512, and dropout of 0.5. Trained for 100 epochs with a batch size of 64.
MOGONET	Adam optimizer was used. During pre-training (500 epochs), the learning rate for GCNs and view-specific classifiers lr_e_pretrain was 1e-3. During joint training (2500 epochs), the GCN learning rate was 5e-4 and VCDN learning rate was 1e-3. GCN dropout was 0.5. Dataset-specific graph adj_parameter and GCN hidden dims were used.
ERL-MVSC	The ERL-MVSC algorithm was implemented with its view weights (λ) being automatically learned and updated iteratively (with a fixed exponent $\alpha = 2$). Fixed trade-offs were $\beta = 0.01$, $\gamma = 1$, $\delta = 1e5$. Max iterations: 40.
MoGCN	We adopted its default hyperparameter configuration: a learning rate of 0.001, weight decay of 0.01, a hidden layer dimension of 64, a dropout of 0.5, and $K = 20$ neighbors for the affinity matrix.
SUPREME	The search space for the learning rate included [0.01, 0.001, 0.0001], and for the hidden layer size, it included [32, 64, 128, 256]. Each combination was trained for 50 epochs. The best configuration from the validation set was re-evaluated.
ECML	For ECML, training was conducted with a batch size of 100 for 500 epochs. The learning rate was set to 0.003. Additionally, the model employed an annealing schedule for a hyperparameter λ, which was varied from 0 to 1 over the initial 50 epochs.

To examine the influence of various omics data on cancer subtypes, we calculate importance scores for each omics type and identify the top 30 features within each category (methylation, mRNA, and miRNA) as cancer biomarkers. The heatmaps show the expression levels of the top 30 genes selected by MOKAN are shown in Fig. 1-3, clearly illustrating the expression differences of the top 30 feature genes within each omics category across different subtypes, providing important clues for further investigation into the functional characteristics of these genes within each subtype. The higher the importance score, the greater the impact of these features on cancer subtype classification.

Among the biomarkers derived from methylation data, the TIMER 2.0 database analysis revealed that NFIX was down-regulated in breast invasive carcinoma (BRCA). Studies indicated that NFIX was down-regulated in BRCA. This down-regulation suggested that NFIX played a significant role in the initiation and progression of the disease. Its low expression was closely associated with malignant progression and poor prognosis in BRCA patients [1]. The TCGA cohort analysis divided BRCA patients into high and low NFE2L3 mRNA expression groups based on the median expression level. Research showed that the expression levels of NFE2L3 were strongly linked to both the initiation and progression of BRCA [2]. NFIL3 was closely related to BRCA, particularly through its involvement in the hypoxic tumor microenvironment. Hypoxia, a common feature of tumors, activated cellular signaling pathways that promoted tumor survival, proliferation, and metastasis. Studies suggested that high levels of NFIL3 expression under hypoxic conditions enhanced tumor cell adaptability, facilitating tumor growth and increasing resistance to treatment. As a result, NFIL3 was identified as an independent prognostic factor in BRCA, with its overexpression correlating with tumor progression and poor prognosis [3]. Additionally, ODF1 exhibited copy number variation (CNV) across all molecular subtypes of breast cancer. This suggested that the increased gene copy number of ODF1 was strongly associated with various subtypes of breast cancer, particularly linking it to genomic instability [4].

Among the biomarkers derived from mRNA data, ANKRD30A was closely related to breast cancer due to its breast-specific expression pattern. The gene ANKRD30A was highly expressed in luminal subtypes of breast cancer, which were generally associated with a better prognosis. This made ANKRD30A a potential biomarker for early breast cancer detection and prognosis prediction [5]. A study on ANKLE1 identified this gene as a new hotspot mutation for breast cancer, playing a critical role in DNA damage response and repair. Whole Exome Sequencing (WES) revealed that mutations in ANKLE1, such as the variant rs2363956, were associated with breast cancer. The GT genotype was found to offer a protective effect compared to the GG genotype. Furthermore, experiments on breast cancer cell lines demonstrated that ANKLE1 expression increased in response to DNA

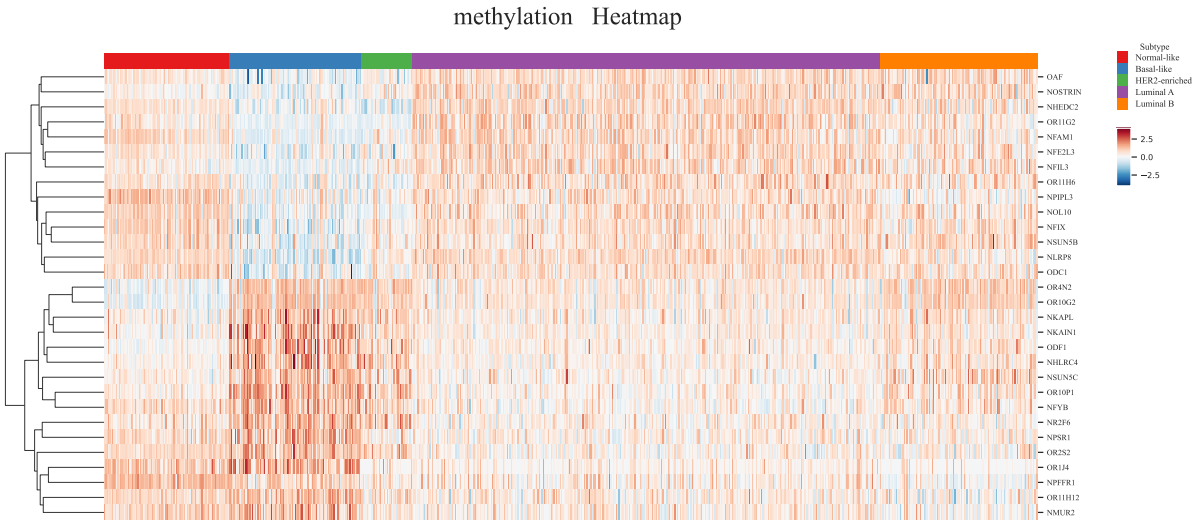

Fig. 1. The heatmap shows the expression levels of the top 30 genes selected by MOKAN in methylation across five breast cancer subtypes.

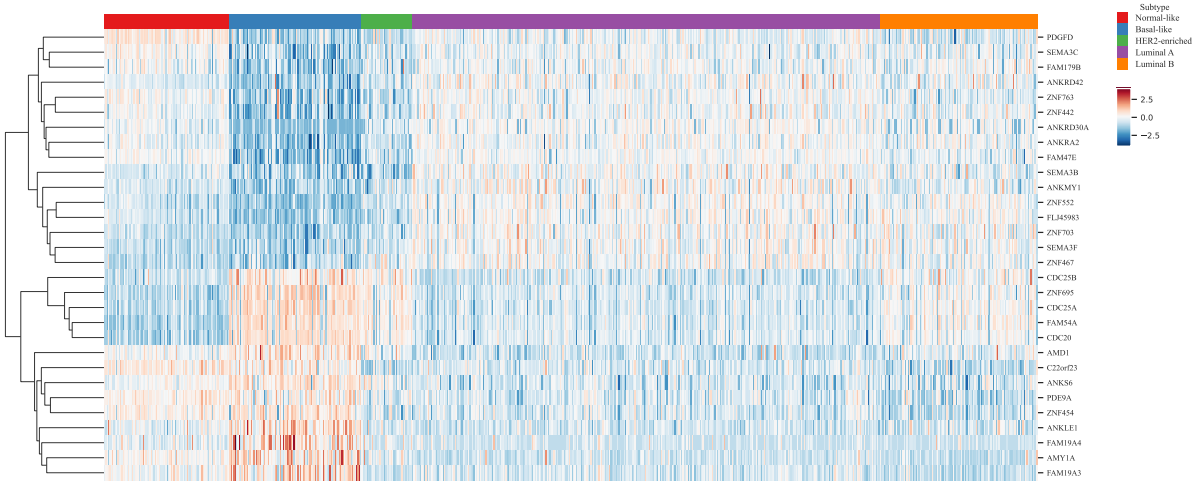

Fig. 2. The heatmap shows the expression levels of the top 30 genes selected by MOKAN in mRNA across five breast cancer subtypes.

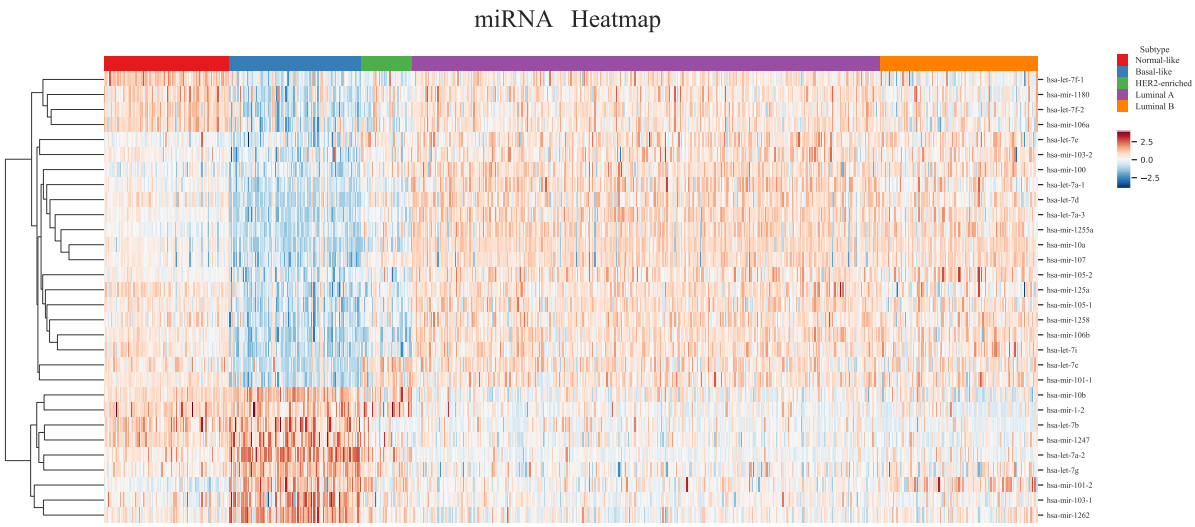

Fig. 3. The heatmap shows the expression levels of the top 30 genes selected by MOKAN in miRNA across five breast cancer subtypes.

damage, underscoring its importance in maintaining genomic stability [6]. CDC25B played a significant role in breast cancer by promoting unchecked cell division and contributing to chemotherapy resistance, particularly to drugs like doxorubicin. Its overexpression in breast cancer disrupted normal cell cycle control, enabling cancer cells to bypass DNA damage responses and continue proliferating, which diminished the effectiveness of treatments. This highlighted CDC25B as a potential target for overcoming drug resistance in breast cancer therapy [7].

Among the biomarkers derived from miRNA data, hsa-let-7d was identified as one of the differentially expressed microRNAs in breast cancer. Analysis of 77 breast cancer samples and 17 control samples revealed that hsa-let-7d was significantly upregulated. This miRNA was one of 13 specifically related to breast cancer, selected from the Human MicroRNA Disease Database (HMDD). The study suggested that hsa-let-7d, along with other miRNAs, may have influenced the progression of breast cancer by regulating target gene networks. The differential expression of these miRNAs was closely linked to breast cancer progression, making them potential diagnostic biomarkers or therapeutic targets for future treatments. Additionally, the study constructed a miRNA-target gene network, demonstrating how these miRNAs regulated tumor-related genes [8]. hsa-let-7b played a significant role in regulating breast cancer progression. It acted as a tumor suppressor by inhibiting the expression of key oncogenes, such as HK2, which was involved in aerobic glycolysis, a critical process for cancer cell metabolism and growth. The study found that let-7b-5p downregulated HK2, reducing glycolysis and suppressing breast cancer cell growth and metastasis, both in vitro and in vivo. Furthermore, let-7b-5p was shown to be downregulated in breast cancer tissues compared to normal tissues, with its reduced expression correlating with more aggressive tumor behavior. These findings suggested that upregulating let-7b-5p could have been a promising therapeutic approach, particularly for breast cancers that overexpress HK2 [9].

In summary, MOKAN effectively identifies biological processes and pathways linked to specific cancers. Additionally, it uncovers the regulatory interactions between methylation or mRNAs or miRNAs and these pathways, offering clear insights into the model interpretation.

REFERENCES

[1] Y. Ren, Y. Zhao, Y. Shan, S. Li, N. Su, Z. Cui, and Z. Yin, "Circular rna hsa_circ_0049657 as a potential biomarker in non-small cell lung cancer," *International Journal of Molecular Sciences*, vol. 24, no. 17, p. 13237, 2023.

[2] M. Liu, H. Wei, J. Yang, X. Chen, H. Wang, Y. Zheng, Y. Wang, and Y. Zhou, "Multi-omics analysis of molecular characteristics and carcinogenic effect of nfe2l3 in pan-cancer," *Frontiers in Genetics*, vol. 13, p. 916973, 2022.

[3] Y. Li, H. Yu, X. Han, and Y. Pan, "Analyses of hypoxia-related risk factors and clinical relevance in breast cancer," *Frontiers in Oncology*, vol. 14, p. 1350426, 2024.

[4] J. M. Fernandez-Muñoz, M. E. Guerrero-Gimenez, L. A. Ciocca, M. J. Germanó, and F. C. M. Zoppino, "Mutational landscape of hsp family on human breast cancer," *Scientific Reports*, vol. 14, no. 1, p. 12471, 2024.

[5] C. Chen, L. Yang, M. Rivandi, A. Franken, T. Fehm, and H. Neubauer, "Bioinformatic identification of a breast-specific transcript profile," *PROTEOMICS–Clinical Applications*, vol. 14, no. 6, p. 2000007, 2020.

[6] D. Bakshi, A. Katoch, S. Chakraborty, R. Shah, B. Sharma, A. Bhat, S. Verma, G. R. Bhat, A. Nagpal, S. Vaishnavi *et al.*, "Ankle1 as new hotspot mutation for breast cancer in indian population and has a role in dna damage and repair in mammalian cells," *Frontiers in Genetics*, vol. 11, p. 609758, 2021.

[7] C. Christowitz, T. Davis, A. Isaacs, G. Van Niekerk, S. Hattingh, and A.-M. Engelbrecht, "Mechanisms of doxorubicin-induced drug resistance and drug resistant tumour growth in a murine breast tumour model," *BMC Cancer*, vol. 19, pp. 1–10, 2019.

[8] T. Zheng, X. Zhang, Y. Wang, and X. Yu, "Predicting associations between micrornas and target genes in breast cancer by bioinformatics analyses," *Oncology Letters*, vol. 12, no. 2, pp. 1067–1073, 2016.

[9] R. Murria Estal, S. Palanca Suela, I. de Juan Jiménez, C. Egoavil Rojas, Z. García-Casado, M. J. Juan Fita, A. B. Sanchez Heras, A. Segura Huerta, I. Chirivella Gonzalez, D. Sánchez-Izquierdo *et al.*, "Microrna signatures in hereditary breast cancer," *Breast Cancer Research and Treatment*, vol. 142, pp. 19–30, 2013.