OmniCLIC: a unified Omics Contrastive Learning framework for effective integration and classification of multi-omics data

Mingzhou Zhang,[†] Xuzeng Liu,[†] Wenyan Chu,[‡] Hang Zhang,[†] and Yunhe Wang[*,†]

[†]*School of Artificial Intelligence, Hebei University of Technology, Tianjin, 300401, China*

[‡]*School of Electrical Engineering, Hebei Vocational University of Technology and Engineering, Xing Tai 054000, China*

E-mail: wangyh082@hebut.edu.cn

Abstract

Integrating multi-omics data for cancer subtype classification remains a critical yet challenging task due to the high dimensionality, heterogeneity, and limited interpretability of omics features. To address these limitations, we propose OmniCLIC, a unified Omics Contrastive Learning and Integration Classification framework that enables end-to-end multi-omics integration, feature learning, and prediction. Three key components are proposed in OmniCLIC: OmniNet, a customized MLP with feature-wise scaling, neural tangent parameterization, and regularization strategies for omics-specific representation learning; a contrastive learning module that jointly optimizes supervised contrastive and cross-entropy losses; and OCDN, a decision-level fusion module that captures inter-omics correlations via a cross-modal correlation tensor to enhance generalization. We evaluate OmniCLIC on four benchmark cancer datasets, where it consistently outperforms state-of-the-art methods in accuracy and robustness

across both binary and multi-class multi-omics datasets. Furthermore, OmniCLIC enables biologically meaningful interpretation by identifying key molecular features through its built-in scaling layers. Functional enrichment analyses on the selected features reveal subtype-specific pathways, such as epithelial morphogenesis and PI3K-Akt signaling, aligning with known cancer biology.

Introduction

In recent years, advancements in biological sequencing technologies have greatly accelerated the generation of multi-omics data, including mRNA expression, DNA methylation, microRNA expression, protein expression, and more.[1,2] These heterogeneous yet complementary data types capture distinct layers of biological information, enabling a comprehensive exploration of cellular mechanisms from multiple molecular perspectives. Compared to single-omics approaches, multi-omics analyses offer a more holistic view of biological systems and have demonstrated strong potential in addressing challenges such as biological heterogeneity and uncovering hidden molecular mechanisms through cross-omics correlation.[3,4] By integrating diverse molecular signals, multi-omics studies help bridge the gap between genotype and phenotype, thereby laying a solid foundation for downstream functional and clinical research. This integration is particularly critical for elucidating the complex interplay between molecular networks and clinical phenotypes, offering novel insights into disease diagnosis, stratification, and personalized treatment strategies.[5,6]

To address the dual hurdles of dimensionality and heterogeneity of multi-omics data, existing multi-omics integration strategies have evolved into three main categories-Early Fusion, Intermediate Fusion, and Decision Fusion—each with its own trade-offs in capturing feature specificity and inter-omics interactions,[7] each balancing feature specificity against inter-omics interaction modeling. Early Fusion combines features from different omics sources into a single input vector, allowing traditional machine learning models, such as K-Nearest Neighbors (KNN),[8] Support Vector Machines (SVM),[9] to be directly applied. This approach

is valued for its simplicity and computational efficiency,[10] and it enables early-stage cross-omics feature interactions during training.[11–13] However, it suffers from limitations such as the curse of dimensionality and potential imbalance due to differing feature scales or distributions across omics types, where high-dimensional modalities may disproportionately influence model learning. Intermediate Fusion seeks to mitigate these issues by first learning modality-specific representations through a unified architecture. This strategy captures both intra-omics patterns and inter-omics correlations while maintaining a manageable representation size.[14–16] Nonetheless, it can struggle when dealing with heterogeneous data sources that do not naturally conform to a common representation space or shared learning framework. In contrast, Decision Fusion adopts a late-stage integration strategy, wherein separate models are trained for each omics modality, and their predictions are aggregated during inference.[17,18] This approach effectively respects the intrinsic characteristics of each omics type, avoids the pitfalls of direct feature-level fusion, and supports more robust and interpretable decision-making. While Decision Fusion preserves modality-specific characteristics, its reliance on independent model training limits its ability to capture emergent cross-modal relationships-motivating the need for multi-omics data modeling approaches.

Traditionally, tree-based models such as XGBoost,[19] CatBoost,[20] and LightGBM[21] have shown strong performance on structured datasets, owing to their robustness, interpretability, and computational efficiency. However, recent advances in deep learning have led to the development of neural network-based models specifically designed for tabular data, which now offer competitive—sometimes even superior—performance.[22] Representative methods include TabNet,[23] TabTransformer,[24] SAINT,[25] and RealMLP.[26] Although neural networks may still face challenges in certain tabular tasks compared to tree-based models, they bring a number of unique advantages particularly well-suited for multi-omics analysis. These include the ability to learn rich feature representations via embedding mechanisms;[27,28] flexible loss function design for complex learning objectives;[29,30] inherent compatibility with multi-modal and unstructured data integration;[31] support for online and incremental learning;[32]

and scalability through distributed training and hardware acceleration.[33] Furthermore, their high expressive power enabled by large-scale parameterization[34] makes them ideal for modeling complex nonlinear dependencies among biological features. Nonetheless, beyond the challenges of dimensionality and heterogeneity, multi-omics data often suffers from subtle inter-class differences, where samples from distinct phenotypes may exhibit highly similar molecular signatures. This leads to classification ambiguity and limits the discriminative capacity of conventional learning methods.

Contrastive learning was first introduced in SimCLR[35] and later extended to Supervised Contrastive Learning[36] for labeled data. Recent studies have demonstrated the effectiveness of contrastive learning in multi-omics integration, particularly in scenarios where labeled data is available. For instance, CLCLSA[37] leverages contrastive learning to maximize mutual information between different omics modalities, while MOCSS[38] applies contrastive learning to align shared representations across omics data for improved clustering and cancer subtyping. This dual utility-capturing correlations while enhancing discriminability—makes contrastive learning a natural fit for multi-omics integration, forming the cornerstone of our proposed framework. However, the high-dimensional feature space and heterogeneous distributions in multi-omics data pose feature ambiguity and representation misalignment across modalities challenges

To address those issues, we introduce OmniCLIC-a unified framework designed to tackle multi-omics integration, feature learning, and classification in an end-to-end manner. In OminCLIC, we leverage RealMLP's neural tangent parameterization and feature-wise scaling to address modality-specific heterogeneity. Moreover, we integrate supervised contrastive learning to explicitly address the challenge of subtle inter-class differences in multi-omics data, where samples from distinct phenotypes often exhibit highly similar molecular profiles. Furthermore, OmniCLIC incorporates the Omics Correlation Discovery Network (OCDN), a decision-level fusion module that constructs a cross-modal correlation tensor to capture inter-omics dependencies. OmniCLIC comprises three core modules:

- OmniNet-based feature learning: A multi-layer perceptron architecture specifically designed for multi-omics data. OmniNet incorporates innovative components such as feature-specific scaling layers, neural tangent parametrization (NTP), Leaky ReLU activation, layer normalization, and dropout regularization to effectively address the challenges posed by high-dimensional and heterogeneous omics features.

- Initial Prediction with Contrastive Learning: To enhance the discriminative power of learned features, supervised contrastive learning is employed alongside cross-entropy loss, which jointly optimizes feature representations and classification performance.

- Multi-omics OCDN Fusion Prediction: Finally, the Omics Cross-Domain Network (OCDN) fuses predictions from different omics modalities at the decision level by constructing a cross-omics correlation tensor, thereby improving the model's generalization ability and prediction accuracy.

Method

This section provides a detailed overview of the OmniCLIC framework. As shown in Figure 1, OmniCLIC integrates omics-specific feature learning, contrastive training, and multi-omics fusion through three core modules: (1) OmniNet-based Omics Feature Learning, (2) Initial Prediction with Contrastive Learning, and (3) Multi-omics Fusion via the Omics Cross-Domain Network (OCDN). The architectural and implementation details of these modules are further depicted in Figure 2. Together, these components operate synergistically to improve the effectiveness and robustness of multi-omics data integration and classification.

OmniNet-based feature learning

To overcome the challenges of extracting informative features from high-dimensional, heterogeneous multi-omics data while mitigating the curse of dimensionality, we propose OmniNet—a

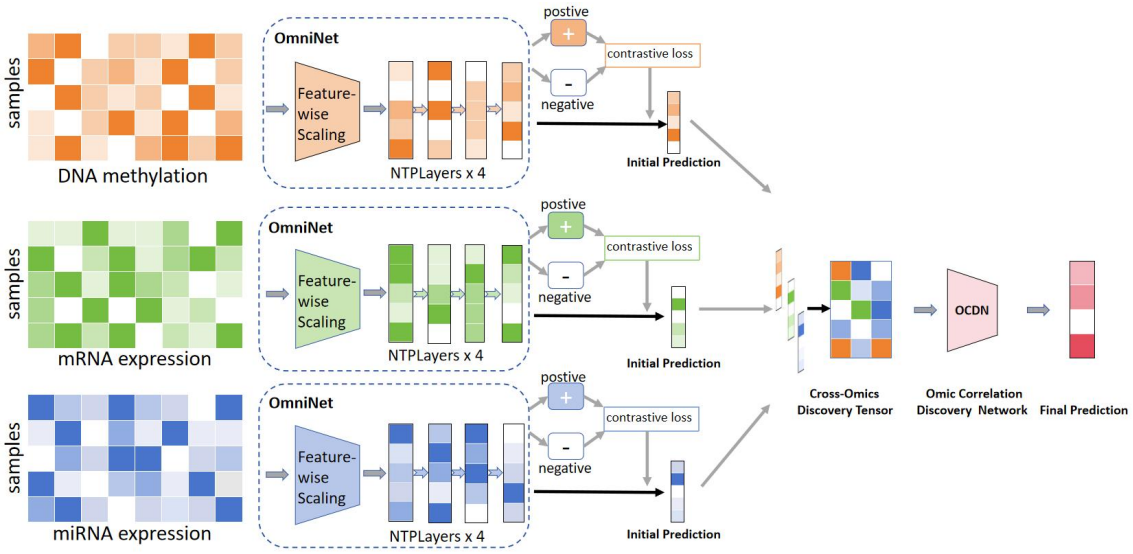

Figure 1: The framework of OmniCLIC consisting of three main modules: (1) OmniNet-based Omics Feature Learning, (2) Initial Prediction with Contrastive Learning, and (3) Multi-omics OCDN Fusion Prediction

unified architecture within OmniCLIC that simultaneously performs dimensionality reduction, feature recalibration, and modality-specific representation learning. This design enables balanced feature extraction across omics modalities while preserving biologically meaningful specificity.

OmniNet is built upon a customized RealMLP backbone and serves as the core engine for omics-specific feature learning in multi-omics data integration. It achieves strong performance by incorporating four complementary innovations. One key component is the feature-specific scaling layer, which enables automatic weighting of input features, particularly important in multi-omics settings, where different modalities contribute unequally to predictive outcomes. The input–output relationship in this layer is defined as follows:

$$\text{ScalingLayer}(x) = x \odot s, \quad s \in \mathrm{R}^{d_{\text{in}}} \tag{1}$$

where s is a learnable scaling vector initialized to 1.

Following feature-specific scaling, the input data is passed through four Neural Tangent Parametrization (NTP) layers,[26] which serve to reduce dimensionality and enhance training stability. NTP is a principled initialization and scaling scheme tailored for deep and wide neural networks. It is grounded in the Neural Tangent Kernel (NTK) theory,[39] which characterizes the training dynamics of infinitely wide neural networks under gradient descent. In contrast to conventional initialization methods, NTP ensures that gradients remain stable across layers, preventing vanishing or exploding gradients and thereby improving convergence behavior. The weights are scaled by $\frac{1}{\sqrt{d_l}}$ using the following equation:

$$z^{(l+1)} = \frac{1}{\sqrt{d_l}} W^{(l)} x^{(l)} + b^{(l)} \tag{2}$$

Where $x^{(l)} \in \mathrm{R}^{d_l}$ denotes the input feature vector of the l-th layer, $W^{(l)} \in \mathrm{R}^{d_{l+1} \times d_l}$ is the weight matrix, $b^{(l)} \in \mathrm{R}^{d_{l+1}}$ is the bias vector, and d_l represents the input dimension of the l-th layer. The $\frac{1}{\sqrt{d_l}}$ factor ensures stable gradient flow regardless of input dimension, crucial for

high-dimensional multi-omics datasets. We used the Leaky ReLU activation function as the non-linear transformation in our network. Unlike the standard ReLU, Leaky ReLU allows a small, non-zero gradient when the input is negative, thereby mitigating the risk of neuron inactivation ('dying ReLU' problem) and enhancing model learning capacity and stability.

To enhance training stability and accelerate convergence of OmniNet, we incorporated Layer Normalization (LayerNorm) into the network architecture. LayerNorm normalizes the activations within each layer to have zero mean and unit variance, effectively mitigating internal covariate shift and stabilizing gradient flow during training. The operation is formally defined as:

$$X_{\text{out}} = \text{LayerNorm}(X_{\text{in}}) = \frac{X_{\text{in}} - \mu}{\sigma} \cdot \gamma + \beta \tag{3}$$

where X_{in} is the input to the layer normalization, μ is the mean of X_{in} over the last few dimensions (usually the feature dimensions), σ is the standard deviation of X_{in} over the same dimensions, and γ and β are learnable parameters that allow the model to scale and shift the normalized values. We then employed dropout regularization[40] to further prevent overfitting and improve generalization of OmniCLIC.

Initial Prediction with Contrastive Learning

While OmniNet effectively captures omics-specific patterns, high-dimensional multi-omics data often exhibit class ambiguity, with samples from distinct phenotypes displaying similar molecular profiles. To address this challenge, we introduce supervised contrastive learning to explicitly promote intra-class compactness and inter-class separability in the latent space. This approach reduces overfitting to noisy features and enhances model robustness, particularly important for distinguishing cancer subtypes with subtle molecular differences.

Following the extraction of robust omics-specific features by OmniNet, we further refine these representations using supervised contrastive learning. During training, OmniNet is optimized not only with standard cross-entropy loss but also with a supervised contrastive

loss, which encourages the model to learn more discriminative features. Specifically, this loss function pulls together the representations of samples from the same class (positive pairs) and pushes apart those from different classes (negative pairs), thereby enhancing intra-class consistency and inter-class separability in the embedding space.

Supervised contrastive learning utilizes label information to construct positive pairs that samples belonging to the same class and negative pairs that samples from different classes. The corresponding loss function is defined as follows:

$$L_{\text{sup}} = \sum_{i \in I} \frac{-1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(z_i \cdot z_p / \tau)}{\sum_{a \in A(i)} \exp(z_i \cdot z_a / \tau)} \tag{4}$$

where i denotes the set of indices for all samples in the batch (including augmented samples), $P(i)$ denotes the set of positive samples sharing the same class label as sample i (including its augmented versions), $A(i)$ is the set containing all sample indices except i, z_i represents the L2-normalized feature vector of the i-th sample and τ denotes the temperature parameter.

Multi-omics OCDN Fusion Prediction

We propose the Omics Correlation Discovery Network (OCDN) for decision-level fusion, aiming to effectively integrate predictions from multiple omics modalities by modeling their interdependencies. OCDN constructs a three-dimensional correlation tensor to explicitly model combinatorial dependencies among omics-specific predictions, thereby uncovering latent biological relationships that contribute to phenotypic outcomes.

After generating initial predictions from each omics-specific OmniNet, OCDN performs decision-level fusion to integrate these outputs. Unlike early fusion approaches, which often suffer from feature dimensionality explosion, and late fusion strategies that may overlook complex inter-omics dependencies, OCDN offers a more expressive and flexible framework for multi-omics integration. By aggregating predictions from diverse omics modalities, OCDN effectively captures their complementary information while preserving modality-

specific strengths. This ensemble-inspired design enhances the robustness and accuracy of the final prediction, aligning with the principle that combining specialized learners yields superior performance over individual models.

For each sample, we form a 3D cross-omics correlation tensor $T_j \in \mathrm{R}^{c\times c\times c}$ by combining the predicted probability distributions $\hat{y}_{j,ai}, i = 1, 2, 3$ from three omics modalities, where c is the number of classes.

$$T_{j,a1a2a3} = \hat{y}_{j,a1}\hat{y}_{j,a1}\hat{y}_{j,a1} \tag{5}$$

Then, the cross-omics discovery tensor $T_{j,a1a2a3}$ is reshaped into a vector of dimension C^3. Finally, the vector is input into a two-layer fully connected layer with output dimension C for final label prediction. The OCDN is trained using cross-entropy loss, that is:

$$\begin{aligned} L_{ce} &= \sum_{j=1}^{n_{tr}} L_{ce}\left(\mathrm{OCDN}(T_{j,a1a2a3})\right) \\ &= \sum_{j=1}^{n_{tr}} -\sum_{j=1}^{n_{tr}} \log\left(\frac{e^{\mathrm{OCDN}(T_{j,a1a2a3})_{y_j}}}{\sum_{k=1}^{C} e^{\mathrm{OCDN}(T_{j,a1a2a3})_k}}\right) \end{aligned} \tag{6}$$

where L_{ce} denotes the cross-entropy loss function, and $\mathrm{OCDN}(T_{j,a1a2a3})_k$ denotes the k -th element in the vector $\mathrm{OCDN}(T_{j,a1a2a3}) \in \mathrm{R}^C$.

In OmniNet, we design a joint loss function that integrates conventional cross-entropy loss with supervised contrastive loss, forming a complementary optimization strategy. The cross-entropy loss focuses on refining the classification decision boundaries, ensuring accurate label prediction. In parallel, the contrastive loss enhances representation learning by encouraging intra-class compactness and inter-class separation in the embedding space, thereby promoting more discriminative and robust features. These two loss functions are then linearly combined to form the final objective, as follows:

$$L_{total} = \alpha \cdot L_{\mathrm{sup}} + (1-\alpha)\cdot L_{\mathrm{ce}} \tag{7}$$

where L_{ce} is the cross-entropy Loss and $alpha$ is a carefully designed proportionality coeffi-

cient used to dynamically balance the contributions of the two loss functions.

Our model OmniCLIC adopts an end-to-end training strategy, where omics-specific OmniNets and OCDN are jointly optimized. Each OmniNet extracts features and produces initial predictions enhanced by supervised contrastive learning, which are then integrated by the OCDN for final prediction. Compared to a modular training approach, where fixed OmniNet outputs were used to train OCDN and led to poor generalization, joint training enables continuous adaptation of feature representations, enriching the input to OCDN and preventing premature convergence. This dynamic optimization significantly improves cross-omics integration and predictive performance.

Experiments

Experimental settings

Datasets

To validate the effectiveness of our proposed OmniCLIC approach, we conducted experiments on four benchmark multi-omics datasets: BRCA, GBM, OV, and KIRP (summarized in Table 1). These datasets exhibit typical tabular structures with samples as rows and molecular features as columns, characterized by high-dimensionality where feature dimensions far exceed sample sizes. Each dataset integrates diverse omics modalities, including mRNA, miRNA, and DNA methylation profiles, reflecting the heterogeneity of real-world biological data. The BRCA dataset targets PAM50 breast cancer subtype classification (five classes), while GBM focuses on brain tumor subtyping (five classes), OV on ovarian cancer classification (four classes), and KIRP on binary kidney cancer subtype discrimination. Notably, all compared algorithms were evaluated on the same dataset splits, which follow a similar partitioning strategy to MOGONET[41] but with dataset-specific ratios to ensure sufficient training samples while maintaining test set representativeness. This approach pre-

serves the core principle of MOGONET's partitioning methodology while adapting to the unique sample distributions of each dataset, ensuring fair comparison with state-of-the-art methods. As shown in Table 1, these high-dimensional datasets pose significant challenges for multi-omics integration, necessitating robust frameworks to handle feature complexity and inter-modality dependencies.

Table 1: Summary of Multi-Omics Datasets

Dataset	Classes	Feature Count			Total
		mRNA	miRNA	Methylation	Samples
BRCA	5	59,427	1,881	24,371	754
GBM	5	11,345	324	11,191	246
OV	4	11,344	321	11,191	286
KIRP	2	16,175	393	16,244	255

Implementation details

Through extensive experimentation, training is terminated at 4,000 epochs to achieve satisfactory convergence, with the learning rate for OmniNet set to 1×10^{-3}. As illustrated in Figure 2, each omics-specific OmniNet begins by applying a learnable scaling layer to reweight input features. This is followed by dimensionality reduction to 256 units via an initial NTPLinear layer, and then three additional NTPLinear layers of the same dimensionality, each equipped with LeakyReLU activation ($\alpha = 0.01$) and dropout for regularization. Layer normalization is incorporated throughout to stabilize training, and the use of independent parameters for each omics modality ensures precise extraction of modality-specific features. Subsequently, the learned features are optimized via supervised contrastive learning, where the loss function emphasizes hard negatives through a temperature parameter τ, used in conjunction with a cross-entropy loss. Dataset-specific optimal values for τ and α are selected in alignment with theoretical principles of contrastive learning. Model optimization is performed using the Adam optimizer with a cosine-annealed learning rate schedule, initialized at a base rate of 1×10^{-3}. Finally, multi-omics fusion is carried out by modeling cross-omics interactions using a $C \times C \times C$ tensor, constructed from the outer products of

omics-level probability vectors. This interaction tensor is decoded by a two-layer LeakyReLU network ($\alpha = 0.01$) to generate final predictions, enabling effective decision-level integration of complex inter-omics relationships.

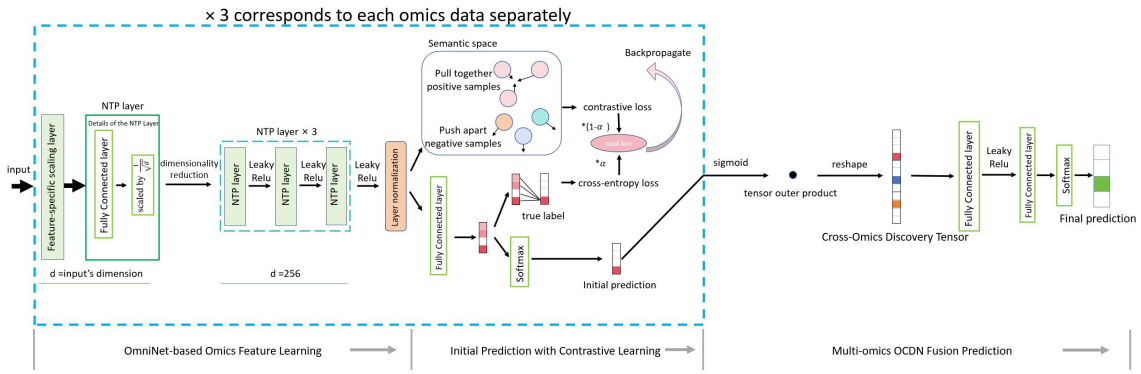

Figure 2: Implementation details of OmniCLIC.

Benchamark methods

We evaluate OmniCLIC against several representative machine learning methods, which categorized into seven traditional machine learning methods and three state-of-the-art deep learning models. The traditional machine learning methods include:

1. **K-Nearest Neighbors (KNN)**[8]: A non-parametric instance-based algorithm using $k = 5$ nearest neighbors with Euclidean distance (L_2 norm) as the distance metric.

2. **Neural Network (NN)**[42]: A multilayer perceptron with two hidden layers (64 and 32

units respectively), ReLU activation, trained using Adam optimizer with 500 maximum iterations.

3. **Logistic Regression (LR)**[43]: A linear classifier with multinomial loss and L_2 regularization, optimized via L-BFGS with 1000 maximum iterations.

4. **Random Forest (RF)**[44]: An ensemble of 100 decision trees with unlimited maximum depth and Gini impurity criterion, using random state 42 for reproducibility.

5. **XGBoost**[19]: A gradient boosting implementation with 100 trees, maximum depth of 6, learning rate of 0.1, and 80% subsampling ratio for both rows and columns.

6. **CatBoost**[20]: A gradient boosting variant with 1000 iterations, tree depth of 6, learning rate of 0.03, and ordered boosting for categorical features.

The deep learning methods, including CLCLSA, MOGONET, and MMDynamics, are detailed as follows:

1. **CLCLSA**[37]: A deep learning method for multi-omics integration with incomplete data, leveraging cross-omics autoencoders, contrastive learning, and self-attention mechanisms.

2. **MOGONET**[41]: A graph neural network framework that performs late fusion at the decision level rather than early feature fusion.

3. **MMDynamics**[45]: A dynamical fusion framework that models feature-level and modality-level informativeness.

Evaluation metrics

Multi-omics classification performance is evaluated through four evaluation metrics, with distinct applications for binary and multiclass tasks. For binary classification, we use Accuracy (ACC), F1 score, and Area Under the Curve (AUC). For multiclass classification, we employ ACC, F1-weighted, and F1-macro. They are detailed as follows:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R}, \quad P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN} \tag{9}$$

$$\text{AUC} = \int_0^1 \underbrace{\frac{TP}{TP + FN}}_{TPR} \, d \underbrace{\frac{FP}{FP + TN}}_{FPR} \tag{10}$$

$$F_{\text{weighted}} = \sum_{i=1}^{C} \frac{N_i}{N_{\text{total}}} F_{1_i}, \quad F_{\text{macro}} = \frac{1}{C} \sum_{i=1}^{C} F_{1_i} \tag{11}$$

Where TP is the number of positive samples correctly predicted as positive by the model. TN is the number of negative samples correctly predicted as negative by the model. FP is the number of negative samples the model predicts as positive, and FN is the number of positive samples the model predicts as negative. C is class count and N_i class-specific sample size. F1-weighted emphasizes majority classes, while F1-macro ensures equal class weighting, which is critical for imbalanced datasets.

Experimental results

Comparison with existing algorithms

To demonstrate the effectiveness of OmniCLIC, we conducted comparative experiments against baseline algorithms across four multi-omics datasets, including BRCA, GBM, OV, and KIRP. The results for those multi-omics datasets are presented in Tables 2-3. As observed from the tables, on the BRCA dataset, OmniCLIC achieved 93.18% accuracy, 93.23% F1-weighted, and 92.39% F1-macro, outperforming the second-best method, MMDynamics, by 3.07% in accuracy and 4.48% in F1-macro. Traditional machine learning methods such as KNN and Random Forest exhibited significantly lower performance, underscoring the advantages of deep learning in handling high-dimensional omics data. Although

Table 2: Performance comparison of OmniCLIC and baseline methods on the BRCA and GBM datasets

Method	**BRCA Dataset**			**GBM Dataset**		
	ACC (%)	**F1-weighted (%)**	**F1-macro (%)**	**ACC (%)**	**F1-weighted (%)**	**F1-macro (%)**
KNN	52.32	39.22	23.83	50.00	46.73	41.66
NN	70.20	71.18	56.48	70.00	70.16	74.05
RF	63.58	54.47	28.85	38.00	36.44	32.55
LR	87.42	87.29	77.80	72.00	70.57	74.26
XGBoost	78.81	76.46	54.63	56.00	55.35	58.76
CatBoost	76.82	70.91	54.86	50.00	47.58	46.20
CLCLSA	87.07	87.41	83.86	82.40	81.47	83.35
MOGONET	89.00	88.96	84.69	86.60	86.59	88.08
MMDynamics	90.11	89.89	87.91	87.20	86.97	88.07
OmniCLIC (Ours)	**93.18**	**93.23**	**92.39**	**92.00**	**92.03**	**93.06**

Table 3: Performance comparison of OmniCLIC and baseline methods on the OV and KIRP Datasets

Method	**OV Dataset**			**KIRP Dataset**		
	ACC (%)	**F1-weighted (%)**	**F1-macro (%)**	**ACC (%)**	**F1 (%)**	**AUC (%)**
KNN	54.78	52.76	52.99	75.49	85.55	61.30
NN	71.30	71.00	70.85	79.41	86.45	78.50
RF	60.00	59.04	58.93	74.51	85.23	83.48
LR	70.43	70.31	70.20	80.39	87.34	82.44
XGBoost	66.09	64.88	64.73	80.39	87.95	82.34
CatBoost	66.09	64.79	64.76	76.47	86.52	81.61
CLCLSA	80.87	80.70	80.55	88.24	92.40	89.45
MOGONET	86.09	86.07	86.05	86.27	91.03	86.05
MMDynamics	87.01	86.79	86.80	86.30	91.41	85.91
OmniCLIC (Ours)	**92.61**	**92.59**	**92.59**	**91.77**	**94.63**	**90.72**

MOGONET, a graph-based multi-omics method, achieved 89.00% accuracy, it fell short of OmniCLIC due to its limited capacity to effectively manage the tabular structure inherent in high-dimensional omics data. For the GBM dataset OmniCLIC achieved 92.00% accuracy, 92.03% F1-weighted, and 93.06% F1-macro, again outperforming both MMDynamics and MOGONET. On the OV dataset, OmniCLIC obtained 92.61% accuracy, 92.59% F1-weighted, and 92.59% F1-macro, surpassing MMDynamics by 5.60% in accuracy. For the binary KIRP dataset, OmniCLIC achieved 91.77% accuracy, 94.63% F1 score, and 90.72% AUC, outperforming CLCLSA by 3.53%.

Experimental results across the four multi-omics datasets demonstrate that OmniCLIC consistently outperforms all baseline methods in evaluation metrics, including accuracy, F1-weighted, and F1-macro scores. These findings underscore OmniCLIC's superior performance in the multi-omics classification task and highlight its robustness and generalizability across cancer types with varying molecular heterogeneity and class imbalance, from multi-omics binary classification problems to more complex multi-class scenarios.

Parameter analysis

We performed a grid search across four multi-omics datasets to evaluate the effect of τ and α hyperparameters on OmniCLIC. Notably, α and τ are key parameters of the contrastive learning module in OmniCLIC. Specifically, α regulates the weight between the supervised contrastive loss and cross-entropy loss, while τ controls the temperature parameter in the contrastive loss function. Specifically, τ was varied in the range $[0.11, 0.47]$ with a step size of 0.04, and α was varied in the range $[0.1, 0.9]$ with a step size of 0.1. The optimal configurations were $\alpha = 0.6$ and $\tau = 0.27$for BRCA, $\alpha = 0.9$ and $\tau = 0.19$ for GBM, $\alpha = 0.2$ and $\tau = 0.35$ for OV, and $\alpha = 0.5$ and $\tau = 0.27$ for KIRP, highlighting the importance of dataset-specific tuning for the contrastive learning framework.

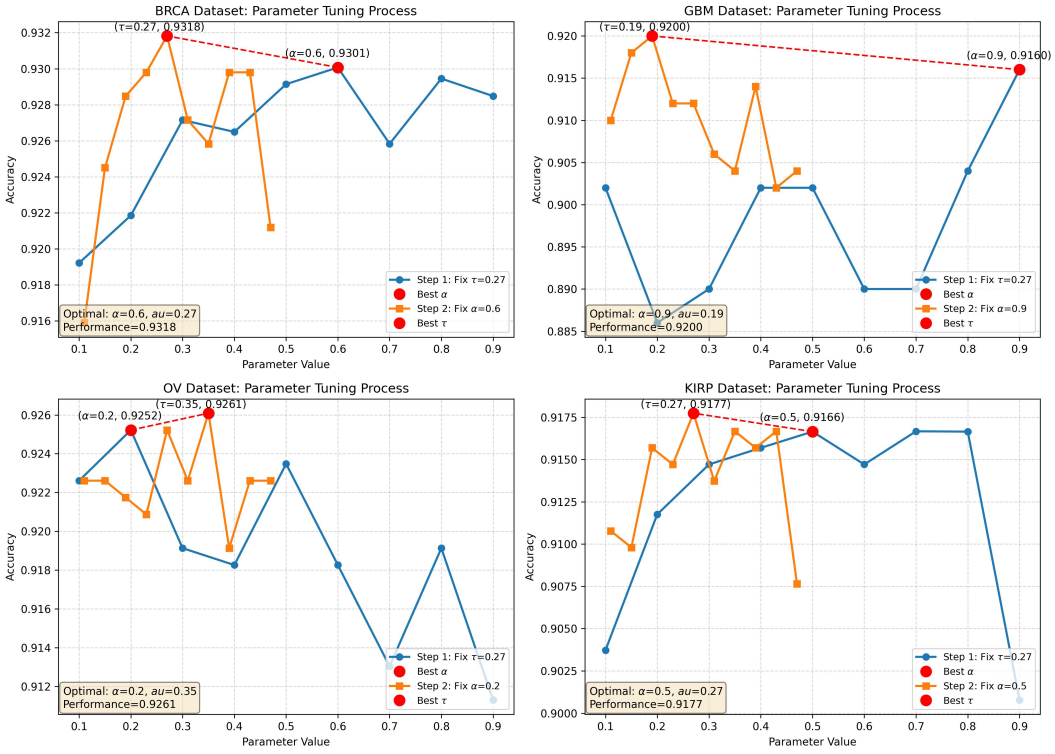

Figure 3: Performance comparison of different parameter settings in OmniCLIC of τ and α.

Ablation study

In order to evaluate the contribution of each component of our proposed algorithm, OmniCLIC, we further conducted ablation studies on four multi-omics datasets including BRCA, GBM, OV, and KIRP. Specifically, we designed four variants for each dataset: (1) OmniCLIC-OmniNet, where OmniNet was replaced with a fully connected layer; (2) OmniCLIC-Contras, where the contrastive learning loss was removed, and only cross-entropy loss was used to train OmniNet; (3) OmniCLIC-OCDN, where OCDN was replaced by a fully connected layer that concatenates the classification results from all OmniNets; and (4) Full OmniCLIC, representing the complete framework used for comparison.

The experimental results in Table 4 and Table 5 validate the necessity of all components within the OmniCLIC framework. The full OmniCLIC consistently achieves the best performance across all datasets, while each ablated variant reveals distinct limitations: (1) OmniCLIC-OmniNet results in the most significant performance decline (6.2–22.6%), reflecting the reduced capacity for effective feature representation; (2) OmniCLIC-Contras leads to moderate performance degradation (1.2–3.8%), highlighting the importance of contrastive learning in enhancing feature discriminability, especially for ambiguous subtypes; (3) OmniCLIC-OCDN causes a 2.9–4.5% drop in accuracy, with the largest decline observed in the OV dataset, where modeling cross-omics interactions is particularly crucial. These results underscore the complementary nature of the components of OmniCLIC that contrastive learning improves feature separability, which is effectively leveraged by OCDN, while OmniNet provides robust representations that are fundamental to the overall framework.

Table 4: Ablation study results of OmniCLIC on BRCA and GBM

Method	**BRCA Dataset**			**GBM Dataset**		
	ACC (%)	**F1-weighted (%)**	**F1-macro (%)**	**ACC (%)**	**F1-weighted (%)**	**F1-macro (%)**
OmniCLIC-OmniNet	90.46	90.41	88.72	89.60	89.57	90.82
OmniCLIC-Contras	91.98	92.03	90.74	90.32	90.37	91.54
OmniCLIC-OCDN	92.25	92.28	92.10	89.30	89.36	90.74
Full OmniCLIC	**93.18**	**93.23**	**92.39**	**92.00**	**92.03**	**93.06**

Table 5: Ablation study results of OmniCLIC on OV and KIRP

Method	OV			KIRP		
	ACC (%)	**F1-weighted (%)**	**F1-macro (%)**	**ACC (%)**	**F1 (%)**	**AUC (%)**
OmniCLIC-OmniNet	91.47	91.44	91.47	89.60	93.27	89.81
OmniCLIC-Contras	91.65	91.65	91.65	91.08	94.16	90.57
OmniCLIC-OCDN	89.13	89.15	89.15	90.98	94.13	89.48
Full OmniCLIC	**92.61**	**92.59**	**92.59**	**91.77**	**94.63**	**90.72**

Performance of OmniCLIC across different omics types

We also evaluated OmniCLIC's performance based on the average accuracy across seven omics configurations: three single-omics modalities (mRNA, DNA methylation, and miRNA), three two-omics combinations (mRNA + methylation, mRNA + miRNA, and methylation + miRNA), and the full integration of all three omics (mRNA + methylation + miRNA). As shown in Figure 4, the three-omics configuration achieved the highest average accuracy of 92.55%, outperforming all two-omics combinations—mRNA + methylation (87.83%), mRNA + miRNA (90.40%), and methylation + miRNA (86.61%), as well as single-omics inputs—mRNA (90.25%), methylation (75.10%), and miRNA (82.08%). Notably, the mRNA + miRNA combination (90.40%) yielded performance closest to the full three-omics model, likely due to their direct regulatory interactions. In contrast, methylation alone resulted in the lowest accuracy (75.10%), highlighting the limited predictive power of epigenetic data when used in isolation. These findings are consistent with OmniCLIC's classification performance advantages and demonstrate that integrating multi-omics data captures complementary biological signals across transcriptional, epigenetic, and post-transcriptional layers, thereby enhancing classification performance beyond what single or paired omics modalities can achieve.

Investigating the interpretability of OmniCLIC

We then extracted the latent representations learned by OmniNet and visualized them using dimensionality reduction techniques to further explore the interpretability of OmniCLIC.

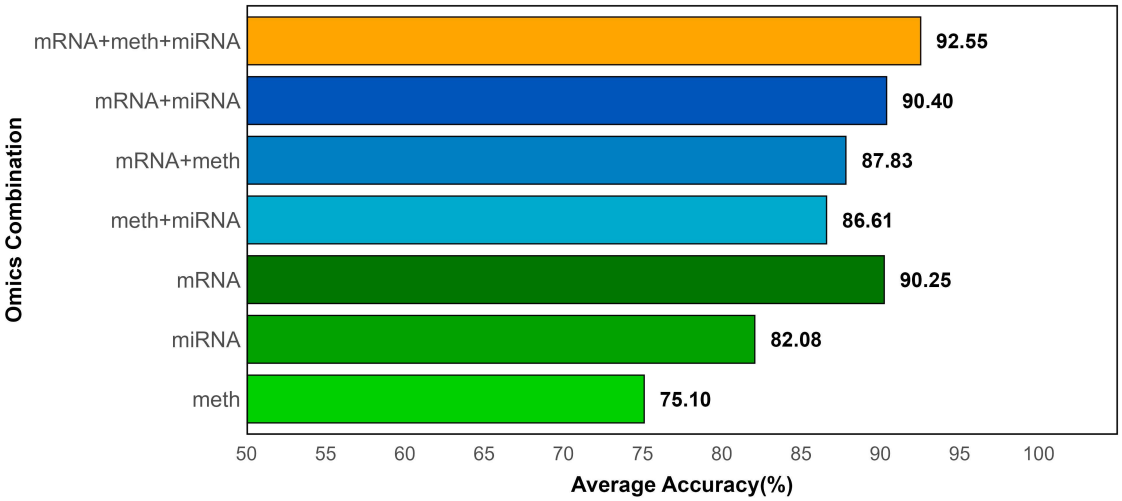

Figure 4: Comparison of average accuracy across different omics types.

A comparison between the clustering quality of OmniCLIC embeddings and that of raw multi-omics input was conducted. As illustrated in Figure 5, the learned embeddings exhibit superior cluster separation and achieve higher Adjusted Rand Index (ARI) scores[46] than the raw data, confirming OmniCLIC's effectiveness in capturing meaningful and discriminative representations for complex multi-omics data.[47]

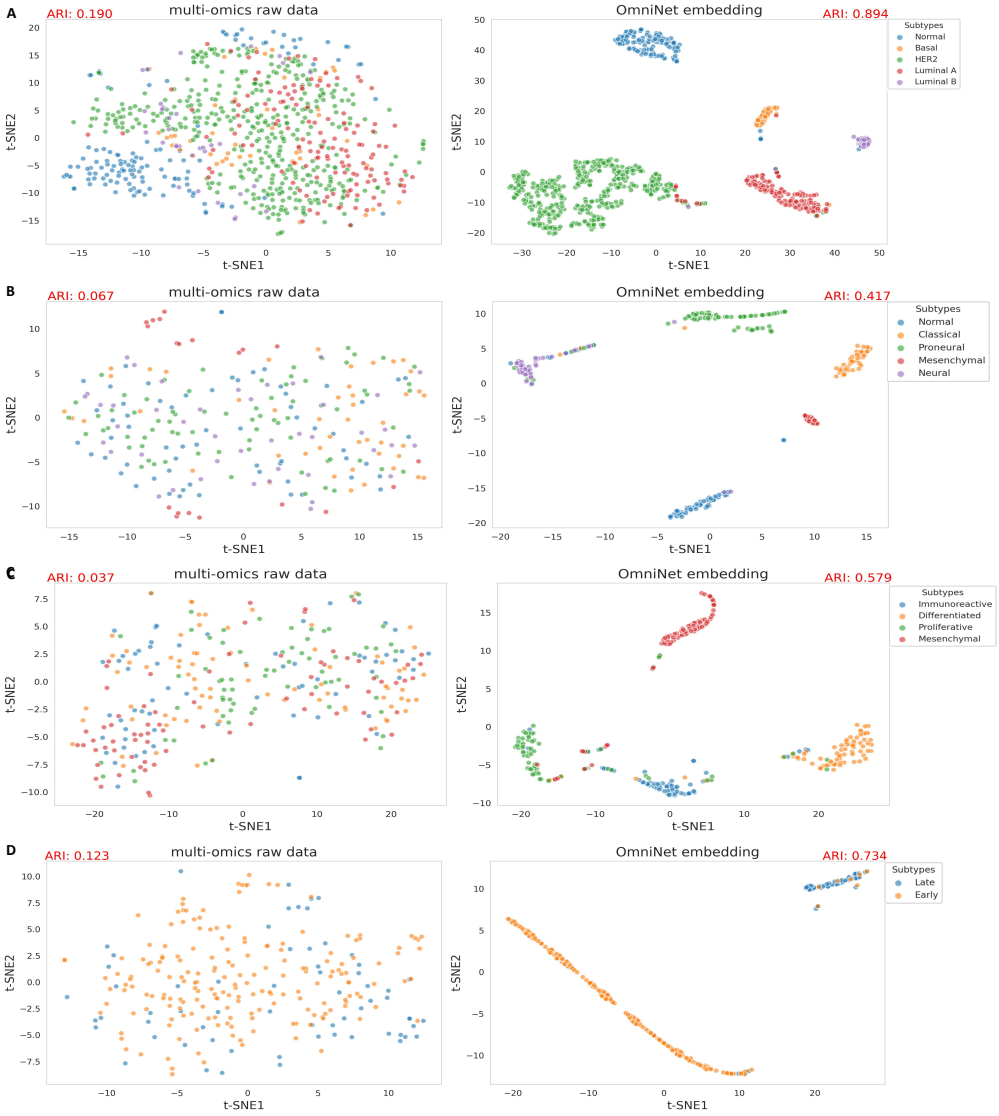

Figure 5: Visualization comparison between raw data (left) and OmniNet's embeddings. (A) Visualization on the BRCA dataset (5 classes), (B) visualization on the GBM dataset (5 classes), (C) visualization on the OV dataset (4 classes), (D) visualization on the KIRP dataset (2 classes).

Key genes analysis of BRCA breast cancer subtypes

To validate the biological interpretability of OmniCLIC, we propose a streamlined approach that eliminates the need for additional feature selection models. Specifically, we utilize the learnable parameters from the scaling layer within the trained OmniNet module of OmniCLIC as importance scores to identify key genes relevant for cancer subtype classification. Unlike existing methods that rely on external models or post hoc feature attribution techniques,[41,48] our approach integrates feature selection directly into the training process. By leveraging the scaling layer's parameters, which are dynamically optimized during training, OmniCLIC captures the relative importance of genes in an end-to-end manner. This not only simplifies the analysis pipeline but also enhances interpretability by aligning feature selection with the model's decision-making process. Consequently, our method provides a unified framework for both classification and biologically meaningful feature prioritization in multi-omics data analysis.

Based on the method, we first identified the top 100 most significant CpG methylation sites from the BRCA dataset by ranking feature importance scores derived from the scaling layer parameters of OmniNet. These CpG sites were subsequently mapped to their corresponding genes using the Illumina HumanMethylation450 annotation file (hg19 build). The resulting gene set was then subjected to Gene Ontology (GO) enrichment analysis to identify overrepresented biological processes, as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The analytical outcomes are visualized in Figure 6, comprising a GO enrichment bubble plot, a KEGG enrichment bar plot, and a KEGG network diagram.

As illustrated in Figure 6 (A), the GO term "morphogenesis of an epithelium" emerged as the most significantly enriched biological process, with the highest gene count in the bubble chart. This result suggests a prominent involvement of epithelial development and remodeling pathways in the breast cancer subtypes analyzed. The dysregulation of epithelial morphogenesis—encompassing cell adhesion, polarity establishment, migration, and cytoskeletal reorganization—has long been recognized as a key driver of tumor progression and hetero-

geneity. Particularly, aberrant epithelial structuring is closely linked to aggressive breast cancer phenotypes marked by epithelial–mesenchymal transition (EMT), loss of apical-basal polarity, and enhanced metastatic potential. These findings are consistent with mechanistic insights from prior studies. For instance, Gray et al.[49] demonstrated that cytoskeletal reprogramming within epithelial morphogenesis pathways facilitates cancer cell invasion, while work by Rodriguez-Boulan and Macara[50] emphasized the contribution of polarity proteins to structural disorganization in malignant tissues.

Figure 6 (B) highlights that the PI3K-Akt signaling pathway and Neuroactive ligand-receptor interaction pathway are the most enriched KEGG terms, as indicated by their high gene counts. The prominence of the PI3K-Akt pathway reaffirms its well-established role in tumorigenesis, regulating essential cellular processes such as proliferation, apoptosis inhibition, and metabolic reprogramming via effectors like mTOR and GSK3β.[51] In parallel, enrichment of the Neuroactive ligand–receptor interaction pathway underscores the influence of neurotransmitter and hormone signaling in oncogenic contexts. Notably, G-protein-coupled receptors mediating these ligand-receptor interactions often converge on the PI3K-Akt axis, forming a synergistic signaling network implicated in both cancer development and neuronal activity.[52,53] This dual enrichment suggests a potential interplay between neuroendocrine regulation and cancer signaling circuits.

As shown in Figure 6 (C), the KEGG network analysis revealed that the Human cytomegalovirus (HCMV) infection pathway constitutes the largest cluster in terms of gene connectivity and count. This observation implies a potentially significant role of viral-related mechanisms in breast cancer pathogenesis. HCMV has been reported to hijack host cellular machinery through viral proteins such as IE1, IE2, and US28, which disrupt tumor suppressor functions (p53, Rb), activate oncogenic pathways (PI3K/Akt, NF-κB), and trigger EMT, thereby promoting tumor progression and immune evasion.[54] The involvement of viral components in the altered gene landscape of breast cancer reinforces the growing recognition of infection-related oncogenesis.

Collectively, these findings demonstrate that the feature importance scores derived from OmniCLIC not only enable accurate classification but also reveal biologically meaningful patterns, offering mechanistic insights into subtype-specific tumor behavior.

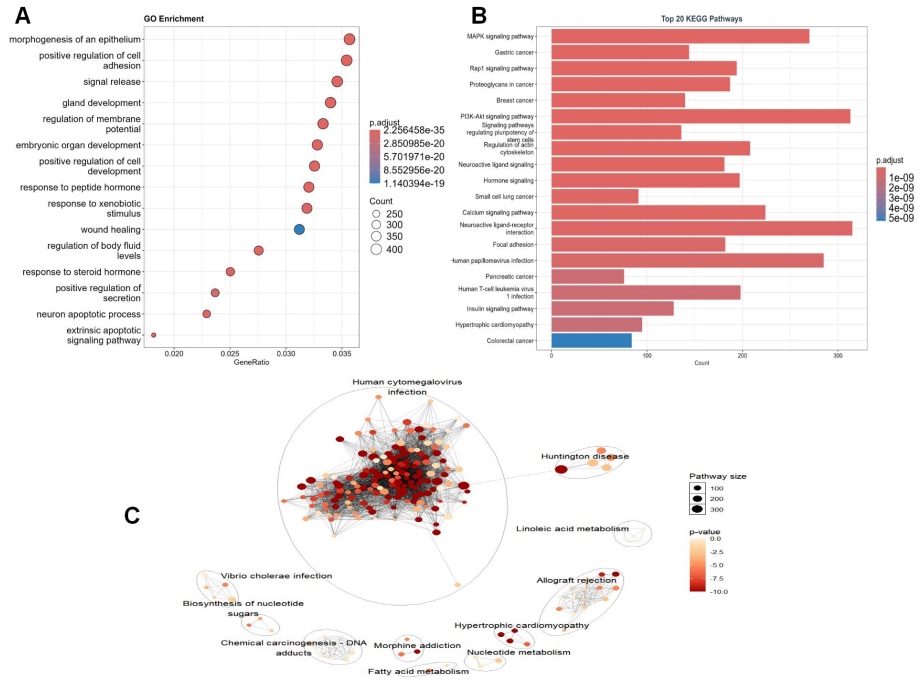

Figure 6: Functional Enrichment of Omic Signatures. (A) GO enrichment bubble plot showing top biological processes. (B) KEGG pathway enrichment bar plot. (C) KEGG network analysis of pathway interactions.

Survival Analysis

We also performed survival analysis on the BRCA dataset to validate the interpretability of OmniCLIC. As illustrated in Figure 7, the four molecular subtypes identified by our model exhibit markedly distinct Kaplan–Meier survival trajectories. The log-rank test yielded a highly significant p-value (¡ 0.001), indicating strong survival stratification across the predicted subtypes. These findings are consistent with the seminal work,[55] which originally demonstrated through gene expression profiling that Basal-like and ERBB2-enriched (HER2+) breast cancer subtypes are associated with the poorest prognoses, while Luminal A and B subtypes

generally confer more favorable outcomes. In particular, Basal-like tumors—characterized by elevated basal keratin expression and frequent TP53 mutations—exhibited rapid declines in survival, whereas Luminal subtypes, driven by estrogen receptor signaling, showed slower disease progression and longer patient survival.

The strong alignment between OmniCLIC's predicted subtypes and established clinical outcomes highlights the biological and clinical relevance of its classifications. This result underscores OmniCLIC's capacity not only to deliver high classification accuracy but also to provide interpretable outputs that reflect real-world patient heterogeneity. The model offers a robust and biologically grounded framework for multi-omics-based cancer subtyping with tangible translational implications.

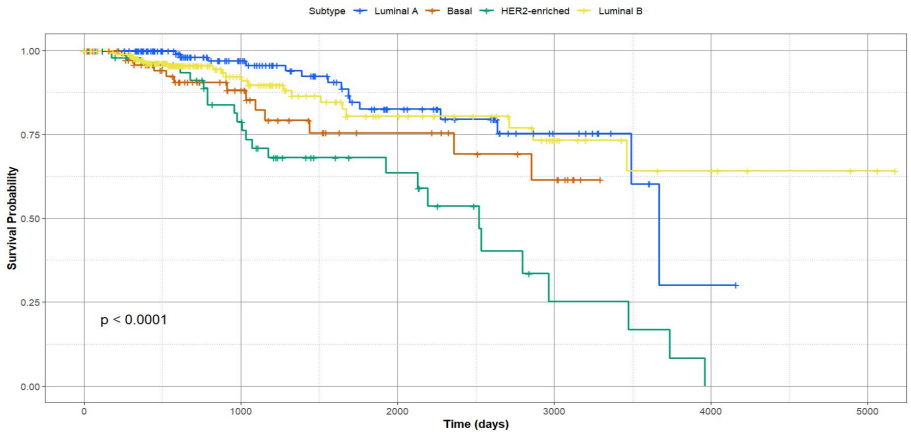

Figure 7: Survival analysis of PAM50 subtypes in the BRCA dataset using OmniCLIC.

Conclusions

In this study, we introduced OmniCLIC, a novel framework for multi-omics data integration and classification while effectively leveraging the complementary strengths of deep learning and contrastive learning. Extensive experiments on diverse real-world datasets demonstrate that OmniCLIC consistently surpasses traditional machine learning methods and current multi-omics integration techniques across multiple evaluation metrics. Moreover, OmniCLIC provides biologically interpretable insights by effectively identifying key molecular features that correlate with clinically relevant cancer subtypes. This interpretability underscores the framework's potential to enhance classification accuracy and facilitate deeper understanding of underlying disease mechanisms. Future research directions include expanding OmniCLIC to incorporate more than three omics modalities simultaneously and adapting the framework for semi-supervised learning scenarios with limited labeled data, thereby broadening its applicability and enhancing its robustness in complex biomedical contexts.

References

(1) Vasaikar, S. V.; Straub, P.; Wang, J.; Zhang, B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic acids research* **2018**, *46*, D956–D963.

(2) Wu, X.; Xu, W.; Deng, L.; Li, Y.; Wang, Z.; Sun, L.; Gao, A.; Wang, H.; Yang, X.; Wu, C.; others Spatial multi-omics at subcellular resolution via high-throughput in situ pairwise sequencing. *Nature biomedical engineering* **2024**, *8*, 872–889.

(3) Hasin, Y.; Seldin, M.; Lusis, A. Multi-omics approaches to disease. *Genome biology* **2017**, *18*, 1–15.

(4) Hoadley, K. A.; Yau, C.; Hinoue, T.; Wolf, D. M.; Lazar, A. J.; Drill, E.; Shen, R.; Taylor, A. M.; Cherniack, A. D.; Thorsson, V.; others Cell-of-origin patterns dominate

the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* **2018**, *173*, 291–304.

(5) Kreitmaier, P.; Katsoula, G.; Zeggini, E. Insights from multi-omics integration in complex disease primary tissues. *Trends in Genetics* **2023**, *39*, 46–58.

(6) Chen, X.; Huang, Y.; Wee, L.; Zhao, K.; Mao, Y.; Li, Z.; Yao, S.; Li, S.; Liang, Y.; Huang, X.; others Integrated analysis of radiomics, RNA, and clinicopathologic phenotype reveals biological basis of prognostic risk stratification in colorectal cancer. *Science Bulletin* **2024**, *69*, 3666–3671.

(7) Benkirane, H.; Pradat, Y.; Michiels, S.; Cournède, P.-H. CustOmics: A versatile deep-learning based strategy for multi-omics integration. *PLOS Computational Biology* **2023**, *19*, e1010921.

(8) Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE transactions on information theory* **1967**, *13*, 21–27.

(9) Boser, B. E.; Guyon, I. M.; Vapnik, V. N. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. 1992; pp 144–152.

(10) Chai, H.; Zhou, X.; Zhang, Z.; Rao, J.; Zhao, H.; Yang, Y. Integrating multi-omics data through deep learning for accurate cancer prognosis prediction. *Computers in biology and medicine* **2021**, *134*, 104481.

(11) Chaudhary, K.; Poirion, O. B.; Lu, L.; Garmire, L. X. Deep learning–based multi-omics integration robustly predicts survival in liver cancer. *Clinical cancer research* **2018**, *24*, 1248–1259.

(12) Choi, J. M.; Chae, H. moBRCA-net: a breast cancer subtype classification framework based on multi-omics attention neural networks. *BMC bioinformatics* **2023**, *24*, 169.

(13) Hira, M. T.; Razzaque, M. A.; Angione, C.; Scrivens, J.; Sawan, S.; Sarker, M. Integrated multi-omics analysis of ovarian cancer using variational autoencoders. *Scientific reports* **2021**, *11*, 6265.

(14) Zhang, X.; Zhang, J.; Sun, K.; Yang, X.; Dai, C.; Guo, Y. Integrated multi-omics analysis using variational autoencoders: application to pan-cancer classification. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2019; pp 765–769.

(15) Zhang, X.; Xing, Y.; Sun, K.; Guo, Y. OmiEmbed: a unified multi-task deep learning framework for multi-omics data. *Cancers* **2021**, *13*, 3047.

(16) Zhu, H.; Wang, Y.; Ma, Z.; Li, X. A comparative study of swarm intelligence algorithms for ucav path-planning problems. *Mathematics* **2021**, *9*, 171.

(17) Han, Z.; Zhang, C.; Fu, H.; Zhou, J. T. Trusted Multi-View Classification. 2021; `https://arxiv.org/abs/2102.02051`.

(18) Zheng, X.; Tang, C.; Wan, Z.; Hu, C.; Zhang, W. Multi-level confidence learning for trustworthy multimodal classification. Proceedings of the AAAI conference on artificial intelligence. 2023; pp 11381–11389.

(19) Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016; pp 785–794.

(20) Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A. V.; Gulin, A. CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems* **2018**, *31*.

(21) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Light-

gbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems* **2017**, *30*.

(22) Shwartz-Ziv, R.; Armon, A. Tabular data: Deep learning is not all you need. *Information Fusion* **2022**, *81*, 84–90.

(23) Arik, S. Ö.; Pfister, T. Tabnet: Attentive interpretable tabular learning. Proceedings of the AAAI conference on artificial intelligence. 2021; pp 6679–6687.

(24) Huang, X.; Khetan, A.; Cvitkovic, M.; Karnin, Z. Tabtransformer: Tabular data modeling using contextual embeddings. *arXiv preprint arXiv:2012.06678* **2020**,

(25) Somepalli, G.; Goldblum, M.; Schwarzschild, A.; Bruss, C. B.; Goldstein, T. SAINT: Improved Neural Networks for Tabular Data via Row Attention and Contrastive Pre-Training. 2021; `https://arxiv.org/abs/2106.01342`.

(26) Holzmüller, D.; Grinsztajn, L.; Steinwart, I. RealMLP: Advancing MLPs and default parameters for tabular data. ELLIS workshop on Representation Learning and Generative Models for Structured Data.

(27) Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). 2019; pp 4171–4186.

(28) Wang, Y.; Bian, C.; Wong, K.-C.; Li, X.; Yang, S. Multiobjective deep clustering and its applications in single-cell RNA-seq data. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **2021**, *52*, 5016–5027.

(29) Kendall, A.; Gal, Y.; Cipolla, R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; pp 7482–7491.

(30) Wang, Y.; Li, X.; Wong, K.-C.; Chang, Y.; Yang, S. Evolutionary multiobjective clustering algorithms with ensemble for patient stratification. *IEEE Transactions on Cybernetics* **2021**, *52*, 11027–11040.

(31) Elmadany, N. E. D.; He, Y.; Guan, L. Multimodal learning for human action recognition via bimodal/multimodal hybrid centroid canonical correlation analysis. *IEEE Transactions on Multimedia* **2018**, *21*, 1317–1331.

(32) Polikar, R.; Upda, L.; Upda, S. S.; Honavar, V. Learn++: An incremental learning algorithm for supervised neural networks. *IEEE transactions on systems, man, and cybernetics, part C (applications and reviews)* **2001**, *31*, 497–508.

(33) Chahal, K. S.; Grover, M. S.; Dey, K.; Shah, R. R. A hitchhiker's guide on distributed training of deep neural networks. *Journal of Parallel and Distributed Computing* **2020**, *137*, 65–76.

(34) Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; others Language models are few-shot learners. *Advances in neural information processing systems* **2020**, *33*, 1877–1901.

(35) Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. International conference on machine learning. 2020; pp 1597–1607.

(36) Khosla, P.; Teterwak, P.; Wang, C.; Sarna, A.; Tian, Y.; Isola, P.; Maschinot, A.; Liu, C.; Krishnan, D. Supervised contrastive learning. *Advances in neural information processing systems* **2020**, *33*, 18661–18673.

(37) Zhao, C.; Liu, A.; Zhang, X.; Cao, X.; Ding, Z.; Sha, Q.; Shen, H.; Deng, H.-W.; Zhou, W. CLCLSA: Cross-omics linked embedding with contrastive learning and self attention for integration with incomplete multi-omics data. *Computers in biology and medicine* **2024**, *170*, 108058.

(38) Chen, Y.; Wen, Y.; Xie, C.; Chen, X.; He, S.; Bo, X.; Zhang, Z. MOCSS: Multi-omics data clustering and cancer subtyping via shared and specific representation learning. *Iscience* **2023**, *26*.

(39) Jacot, A.; Gabriel, F.; Hongler, C. Neural Tangent Kernel: Convergence and Generalization in Neural Networks. Advances in Neural Information Processing Systems. 2018.

(40) Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* **2014**, *15*, 1929–1958.

(41) Wang, T.; Shao, W.; Huang, Z.; Tang, H.; Zhang, J.; Ding, Z.; Huang, K. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nature communications* **2021**, *12*, 3445.

(42) McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* **1943**, *5*, 115–133.

(43) Fisher, R. A. The use of multiple measurements in taxonomic problems. *Annals of eugenics* **1936**, *7*, 179–188.

(44) Breiman, L. Random forests. *Machine learning* **2001**, *45*, 5–32.

(45) Han, Z.; Yang, F.; Huang, J.; Zhang, C.; Yao, J. Multimodal Dynamics: Dynamical Fusion for Trustworthy Multimodal Classification. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022; pp 20675–20685.

(46) Hubert, L.; Arabie, P. Comparing partitions. *Journal of classification* **1985**, *2*, 193–218.

(47) Halkidi, M.; Vazirgiannis, M. Clustering validity assessment: Finding the optimal partitioning of a data set. Proceedings 2001 IEEE international conference on data mining. 2001; pp 187–194.

(48) Li, M.; Guo, H.; Wang, K.; Kang, C.; Yin, Y.; Zhang, H. AVBAE-MODFR: A novel deep learning framework of embedding and feature selection on multi-omics data for pan-cancer classification. *Computers in Biology and Medicine* **2024**, *177*, 108614.

(49) Gray, R. S.; Cheung, K. J.; Ewald, A. J. Cellular mechanisms regulating epithelial morphogenesis and cancer invasion. *Current opinion in cell biology* **2010**, *22*, 640–650.

(50) St Johnston, D.; Sanson, B. Epithelial polarity and morphogenesis. *Current opinion in cell biology* **2011**, *23*, 540–546.

(51) He, Y.; Sun, M. M.; Zhang, G. G.; Yang, J.; Chen, K. S.; Xu, W. W.; Li, B. Targeting PI3K/Akt signal transduction for cancer therapy. *Signal transduction and targeted therapy* **2021**, *6*, 425.

(52) Dankoski, E. C.; Wightman, R. M. Monitoring serotonin signaling on a subsecond time scale. *Frontiers in integrative neuroscience* **2013**, *7*, 44.

(53) Heine, M.; Thoumine, O.; Mondin, M.; Tessier, B.; Giannone, G.; Choquet, D. Activity-independent and subunit-specific recruitment of functional AMPA receptors at neurexin/neuroligin contacts. *Proceedings of the National Academy of Sciences* **2008**, *105*, 20947–20952.

(54) Taher, C.; de Boniface, J.; Mohammad, A.-A.; Religa, P.; Hartman, J.; Yaiw, K.-C.; Frisell, J.; Rahbar, A.; Söderberg-Naucler, C. High prevalence of human cytomegalovirus proteins and nucleic acids in primary breast cancer and metastatic sentinel lymph nodes. *PloS one* **2013**, *8*, e56795.

(55) Sørlie, T.; Perou, C. M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M. B.; Van De Rijn, M.; Jeffrey, S. S.; others Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences* **2001**, *98*, 10869–10874.