

PCLSurv: a prototypical contrastive learning-based multi-omics data integration model for cancer survival prediction

Zhimin Li^{1,†}, Wenlan Chen^{2,†}, Hai Zhong^{3,*}, Cheng Liang^{1,3,*}

¹School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China

²School of Computer Science and Engineering, Central South University, Changsha 410083, China

³Department of Radiology, the Second Hospital of Shandong University, Jinan 250033, China

*Corresponding authors. Cheng Liang, School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China.

E-mail: alcs417@sdnu.edu.cn; Hai Zhong, Department of Radiology, the Second Hospital of Shandong University, Jinan 250033, China.

E-mail: 18753107255@163.com.

†Zhimin Li and Wenlan Chen contributed equally to this work and are joint first authors.

Abstract

Accurate cancer survival prediction remains a critical challenge in clinical oncology, largely due to the complex and multi-omics nature of cancer data. Existing methods often struggle to capture the comprehensive range of informative features required for precise predictions. Here, we introduce PCLSurv, an innovative deep learning framework designed for cancer survival prediction using multi-omics data. PCLSurv integrates autoencoders to extract omics-specific features and employs sample-level contrastive learning to identify distinct yet complementary characteristics across data views. Then, features are fused via a bilinear fusion module to construct a unified representation. To further enhance the model's capacity to capture high-level semantic relationships, PCLSurv aligns similar samples with shared prototypes while separating unrelated ones via prototypical contrastive learning. As a result, PCLSurv effectively distinguishes patient groups with varying survival outcomes at different semantic similarity levels, providing a robust framework for stratifying patients based on clinical and molecular features. We conduct extensive experiments on 11 cancer datasets. The comparison results confirm the superior performance of PCLSurv over existing alternatives. The source code of PCLSurv is freely available at <https://github.com/LiangSDNULab/PCLSurv>.

Keywords: multi-omics integration; prototypical contrastive learning; survival prediction

Introduction

Cancer remains one of the leading causes of mortality worldwide, posing a significant burden on public health [1]. Accurate prediction of patient survival is critical for guiding clinical decision-making, optimizing treatment strategies, and improving patient outcomes. For instance, survival predictions can aid in identifying high-risk patients who may require more aggressive interventions, thereby contributing to more efficient resource allocation and personalized care. Despite progress in medical research and cancer therapies, precise survival prediction remains a major challenge.

With the rapid development of high-sequencing techniques, various single omics data, including gene expression [2], microRNAs [3], copy number alterations [4, 5], DNA methylation [6], proteomics [7], etc., have been widely investigated in survival prediction. These datasets offer rich biological information that can be leveraged to improve survival predictions and enhance understanding of cancer biology. Gene expression data have been demonstrated to have significant potential in prognostic classification [8]. DNA methylation patterns, which represent key epigenetic changes in gastric cancer initiation and progression, have spurred the development signature to improve survival prediction

[9]. However, single-omics data are inherently limited in their ability to capture the complex interplay between different molecular layers. This complexity cannot be fully understood using single-omics data alone, necessitating the integration of multi-omics data. To address these limitations, multi-omics integration combines information from diverse molecular sources, providing a more comprehensive view of cancer biology and enhancing the accuracy and robustness of survival prediction models [10]. Current computational approaches for addressing these challenges can be broadly divided into three categories: statistical models, traditional machine learning models, and deep learning models [11].

Statistical models, such as Cox proportional hazards regression [12], are still widely used in survival analysis due to their interpretability and suitability for small- to medium-sized datasets. With the advent of high-dimensional multi-omics data, extensions like penalized Cox models (e.g. Lasso-Cox [13] and Elastic Net Cox [14]) have been employed to address feature selection and reduce over-fitting. These methods retain the simplicity of the Cox model while improving performance by applying regularization techniques in more complex data settings. Bayesian linear models are

Received: December 17, 2024. Revised: February 23, 2025. Accepted: March 4, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com

also gaining attention, particularly for spatial cancer data. Recent methods incorporating priors such as horseshoe and point mass mixture priors have improved variable selection and predictive performance [15]. The use of graph-structured priors has allowed models to identify key pathways and genes across heterogeneous cohorts [16]. TBLMM uses Bayesian model and capture both linear and nonlinear interactions in multi-omics data [17]. However, traditional models often rely on assumptions of proportional hazards and linear relationships, limiting their ability to fully capture the complex, nonlinear interactions inherent in cancer data, particularly across diverse omics.

Representative machine learning methods for survival prediction include three-stage ensemble learning, XGBLC, and Ada-RSIS. The genetic algorithm-aided three-stage ensemble method enhances risk prediction by utilizing gene-pairing and genetic algorithms to optimize learner combinations [18]. The XGBLC model combines XGBoost with Lasso-Cox, improving the analysis of high-dimensional data by specialized gradient statistics and loss functions [19]. Ada-RSIS enhances multi-omics data integration by learning shared and individual subspaces, ensuring complementarity and risk-aware representations through adaptive weighting [20]. Traditional machine learning methods in survival analysis struggle with high-dimensional, noisy data and often require feature selection, leading to potential loss of important information. Additionally, they face challenges in integrating multimodal data and modeling complex, nonlinear relationships, which limits their performance and generalizability.

Deep learning models have gained popularity due to their ability to model nonlinear relationships and integrate complex multimodal datasets [21]. ML-ordCOX combines the Cox model with adaptive multitask learning to improve survival prediction using multi-omics data [22]. PathCNN constructs an interpretable convolutional neural network applied to integrated multi-omics data, demonstrating potential in predicting survival in glioblastoma [23]. HFBSurv introduces a hierarchical multimodal fusion framework that employs factorized bilinear models to integrate genomic and imaging features [24]. DFSC, combining deep forests with self-supervised learning, improves survival predictions for adaptive learning from high-dimensional genomic data [25]. Meta-learning with DeepLIFT for model interpretability, using Cox risk loss, enhances survival prediction and enables the exploration of various enrichment pathways [26]. FGCNSurv combines factorized bilinear models with graph convolutional networks, utilizing multi-graph fusion and advanced feature extraction to improve prediction accuracy [27]. CAMR introduces a cross-aligned multi-omics representation learning network that generates modality-invariant representations to enhance cancer survival predictions [28]. CSAM-GAN, integrating generative adversarial networks and attention modules, facilitates multi-omics feature selection and survival prediction by refining input features through sequential channel-spatial attention and an auto-encoder architecture [29]. Prior knowledge-guided multi-level graph neural networks integrate multi-omics data and pathway information to enhance tumor risk prediction accuracy [30]. Autosurv develops a deep learning framework for prognosis prediction by integrating multi-omics data while identifying key features that differentiate high- and low-risk patients leveraging a specially designed VAE [31]. The CLCox model classifies samples into different groups, where samples with the same label are considered positive sample pairs, and those with different labels are considered negative sample pairs thereby enhancing cancer prognosis prediction [32]. MMOSurv leverages meta-learning to perform effective multi-omics survival prediction for

specific cancer types. By utilizing meta-knowledge from related cancers, it adapts a deep Cox model with minimal training samples. MMOSurv effectively harnesses meta-information derived from the similarities and relationships between different omics data across related cancer datasets, thereby enhancing survival prediction for the target cancer with a limited number of multi-omics training samples [33]. The Prototypical Information Bottleneck and Disentangling framework is designed to address both intra-modal and inter-modal redundancies. It reduces intra-modal redundancy by introducing the Prototypical Information Bottleneck module, while the Prototypical Information Disentangling module disentangles redundant information across modalities, thereby enhancing survival prediction accuracy [34].

Although existing methods have made great achievements for cancer survival prediction, they often struggle to effectively integrate data from different omics, thereby neglecting the complementarity of multidimensional information. Additionally, existing methods often fail to fully leverage the semantic structural relationships between different omics data, missing the opportunity to mine the underlying similarities and correlations within the data. In this study, we propose PCLSurv, a novel deep learning framework for accurate cancer survival prediction using multi-omics data. Our model first uses autoencoders to extract view-specific features for each omics, and then applies sample-level contrastive learning to capture distinct yet consistent characteristics across views for each sample. These representations are fused via a bilinear fusion module (BFM) to obtain the cross-viewed information. Specifically, to capture high-level semantic similarities, our model leverages prototypical contrastive learning (PCL) to bring similar samples closer by aligning them with shared prototypes, while encouraging separation from unrelated prototypes, thereby enhancing the model's ability to distinguish between patient groups with different survival profiles. Extensive experiments across 11 cancer datasets confirm the superior performance of our model.

Materials and methods

Given the cancer multi-omics data, represented by $\{X^1, X^2, \dots, X^V\}$, where $X^v \in \mathbb{R}^{N \times d^v}$, $v \in [1, V]$ corresponds to the v th omics data matrix for N patients, and d^v is the feature dimension. For each patient i , we have an observed time-to-event O_i and a survival status indicator δ_i where $\delta_i = 1$ if the event is observed (uncensored) and $\delta_i = 0$ if the event is censored. In this work, we propose a prototypical contrastive learning-based framework to predict cancer patient risk scores by leveraging multi-omics data. Our method can effectively fuse the omics-specific representations and obtain semantically discriminative representations for survival prediction. The framework of our proposed PCLSurv is presented in Fig. 1 and we will introduce each module of our model below.

Reconstruction module

We first extract latent features for each omics data by using autoencoders, which effectively preserves essential omics-specific details for downstream tasks. Specifically, for the v th omics ($v \in [1, V]$), we adopt an encoder E^v to map the reduced data into the embedding space and obtain its low-dimensional latent representations $Z^v \in \mathbb{R}^{N \times d}$. The latent representations are then restored to their original dimensions via a corresponding decoder D^v . The encoder and decoder process can be simply demonstrated

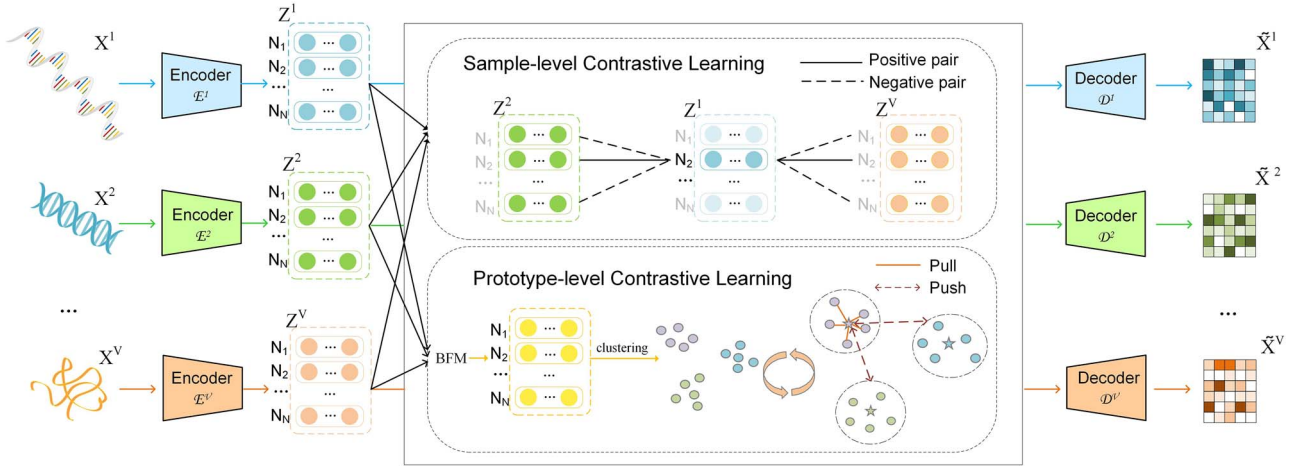


Figure 1. Overview of the PCLSurv framework for cancer survival prediction using multi-omics data. Autoencoders extract view-specific features for each omics type. Sample-level contrastive learning captures both distinct and consistent characteristics across views, while PCL aligns similar samples with shared prototypes and separates unrelated ones. BFM module integrates multi-omics features into a unified representation.

as follows:

$$Z^v = E^v(X^v) \quad (1)$$

$$\tilde{X}^v = D^v(Z^v) \quad (2)$$

We use mean squared errors as the loss function to maintain the fidelity of the latent representations:

$$L_{re} = \sum_{v=1}^V \|\tilde{X}^v - X^v\|_2^2 \quad (3)$$

Sample-level contrastive learning module

While the reconstruction module effectively reduces feature dimensions, it only retains essential information for each omics data reconstruction and fails to capture distinct yet consistent characteristics across multi-omics. In response to this limitation, we introduce the sample-level contrastive learning module, which is designed to align samples from different omics that exhibit shared characteristics, while differentiating them from unrelated samples. As a matter of fact, the primary objective of this module is to optimize the feature space by pulling positive sample pairs closer together while pushing negative sample pairs apart.

Specifically, for a given sample x_i , we consider $\{(z_i^v, z_i^w) \mid v \neq w\}$ as positive sample pairs whereas samples from different views, i.e. $\{(z_i^v, z_j^w) \mid i \neq j\}$ as negative pairs. We adopt the commonly used cosine similarity measure to calculate the similarity between pairs of sample representations:

$$\text{sim}(z_i^v, z_j^w) = \frac{(z_i^v)^T (z_j^w)^T}{\|z_i^v\| \|z_j^w\|} \quad (4)$$

where $v, w \in [1, V]$ and $i, j \in [1, N]$. To maximize the similarity of positive pairs and simultaneously minimize that of negative pairs, we formulate the sample-level contrastive loss for all samples as follows:

$$L_{con} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\sum_{v,w \in [1,V], v \neq w} \exp(\text{sim}(z_i^v, z_i^w)/\tau)}{\sum_{j=1}^N \sum_{v,w \in [1,V]} \exp(\text{sim}(z_i^v, z_j^w)/\tau)} \quad (5)$$

where τ is the temperature parameter. By leveraging this contrastive loss, our approach encourages the model to learn

representations that bring similar samples closer together in the feature space while pushing dissimilar samples further apart, enhancing the model's ability to distinguish between relevant biological patterns across different omics.

Bilinear fusion module

While the sample-level contrastive learning module effectively aligns samples from distinct omics types based on pairwise similarities, it is limited in its capacity to model higher-order interactions among omics [35]. We overcome this limitation by using a BFM, which allows for fine-grained cross-omics feature interactions and further fuses the omics-specific features to obtain a unified representation capturing complex interdependencies.

Given two omics representations z_i^v and z_i^w of a sample i , its cross-omics representation h_i^{vw} is defined as

$$\begin{aligned} h_i^{vw} &= z_i^{vT} W_i z_i^w \\ &= \sum_{g=1}^G z^{vT} a_g^v b_g^w z^w \\ &= e^T (A_i^{vT} Z^v \odot B_i^{wT} Z^w) \end{aligned} \quad (6)$$

$$A_i^v = [a_1^v, \dots, a_G^v] \in \mathbb{R}^{d \times G}, B_i^w = [b_1^w, \dots, b_G^w] \in \mathbb{R}^{d \times G} \quad (7)$$

where G is the latent dimensionality of the factorized matrices A_i^v and B_i^w , \odot denotes the Hadamard product, and $e \in \mathbb{R}^G$ is a vector with all ones. To obtain the output feature H^{vw} , we need to learn the weight matrices [36] $A^v = [A_1^v, \dots, A_N^v] \in \mathbb{R}^{d \times G \times d'}$ and $B^w = [B_1^w, \dots, B_N^w] \in \mathbb{R}^{d \times G \times d'}$, both of which are three-order tensors. Additionally, Equation(6) can be rewritten as follows:

$$H^{vw} = \text{SumPooling}(\tilde{A}^v Z^v \odot \tilde{B}^w Z^w, G) \quad (8)$$

where $\text{SumPooling}()$ function performs and pools on Z by using 1D non-overlapping windows of size G , $\tilde{A}^v \in \mathbb{R}^{d \times Gd'}$ and $\tilde{B}^w \in \mathbb{R}^{d \times Gd'}$ are 2D matrices reshaped from A^v and B^w , respectively. Based on Equation (8), we compute the fused feature for all omics pairs and then average these representations to obtain a unified representation H . H is further combined with the averaged omics-specific representations from all views, yielding the final representation

Z, as formalized in the following equation:

$$Z = H \oplus \frac{1}{V} \left(\sum_{v=1}^V Z^v \right) \quad (9)$$

where \oplus denotes the concatenation operation. As a result, bilinear fusion enhances the model's ability to learn interactions between different omics datasets, enriching the latent space with valuable higher-order information beyond what concatenation or linear methods can achieve.

Prototypical contrastive learning module

To enhance feature clustering and improve class separability, we integrate PCL into our model, enabling it to capture higher-level abstractions by associating each sample with a representative prototype. Specifically, a prototype is defined as the central representation of a sample group in the feature space, typically calculated as the mean of the samples within a given group. Compared with sample-level contrastive learning, PCL focuses on capturing high-level semantic similarities for patient groups by leveraging prototypes. Specifically, to derive semantically representative prototypes, we apply the k-means algorithm to the representation Z learned from the BFM. This methodology results in the formation of a prototype set c , wherein each prototype $c_l \in c$ corresponds to the centroid for the l th cluster, representing the semantic center of that cluster.

For each sample z_i , we treat the prototype of the cluster it belongs to as a positive counterpart, while the remaining $K-1$ prototypes are treated as negative counterparts. Consequently, this approach effectively prevents semantically similar samples within the same cluster from being pushed apart. To further enhance the robustness of our model and capture multi-level semantic similarities, we perform clustering on the representations M times, each using different cluster numbers $K = \{k_m\}_{m=1}^M$. Taken together, the prototypical contrastive loss is defined as

$$L_{pc} = \frac{1}{M} \sum_{i=1}^N \sum_{m=1}^M -\log \left(\frac{\exp(\text{sim}(z_i \cdot c_o^m)/\sigma)}{\sum_{l=1}^{k_m} \exp(\text{sim}(z_i \cdot c_l^m)/\sigma)} \right) \quad (10)$$

where c_o^m is the prototype of the cluster to which z_i belongs and c_l^m is the prototype of the l th cluster in the m th clustering. σ is a hyperparameter controlling the concentration level. Since prototypes act as reference points that embody the common features or characteristics of each group, the introduction of PCL helps the model generalize better by focusing on overarching patterns rather than individual sample noise.

Overall loss function

The Cox proportional hazards model is widely used in survival analysis to examine the relationship between the survival times of subjects and predictor variables. We enhance survival prediction by using the fused omics representation Z as input to the hazards model. The survival risk score for the i th patient and the Cox partial likelihood loss for survival prediction are defined as follows, respectively:

$$s_i = \beta^T z_i \quad (11)$$

$$L_{cox} = - \sum_{i=1}^N \delta_i \left[\beta^T z_i - \log \left(\sum_{j: O_j > O_i} \exp(\beta^T z_j) \right) \right] \quad (12)$$

Algorithm 1 The pseudocode of PCLSurv

Input: Multi-omics datasets $\{X^v\}_{v=1}^V$, O , δ : X^v is the v -th omics data, O is observed time, δ is censoring indicator; Maximum iteration T ; number of clusters $K = \{k_m\}_{m=1}^M$

Output: survival outcome: risk score s

- 1: Initialize model parameters
- 2: **for** epoch = 1 to T **do**
- 3: Compute latent representations Z^v for each omics dataset by Eq.(1)
- 4: Compute reconstruction representations \tilde{X}^v by Eq.(2)
- 5: Compute reconstruction loss by Eq.(3)
- 6: Construct positive and negative pairs for sample-level contrastive learning
- 7: Compute contrastive loss by Eq.(5)
- 8: Fuse multi-omics features using Eq.(8) and Eq.(9), obtaining the fused representation Z
- 9: **for** $m = 1$ to M **do**
- 10: cluster Z into k_m clusters, and obtain prototypes
- 11: **end for**
- 12: Compute prototypical contrastive loss by Eq.(10)
- 13: Compute risk score s_i by Eq.(11)
- 14: Compute Cox loss by Eq.(12)
- 15: Compute the overall loss by Eq.(13)
- 16: Perform backpropagation and update parameters
- 17: **end for**

where δ denotes the event status indicator, O denotes observed time-to-event and $i, j \in [1, N]$.

By consolidating the reconstruction loss, sample-level contrastive loss, prototypical contrastive loss, and the Cox partial log-likelihood loss into a unified framework, our model learns robust, multi-omics representations that enhance survival prediction by capturing both individual and group-level characteristics relevant to patient outcomes, allowing these structures to mutually reinforce one another. The overall loss function is demonstrated in Equation (13) and we summarize the workflow of our model in Algorithm 1.

$$L = L_{re} + L_{con} + L_{pc} + L_{cox} \quad (13)$$

Results

In our study, extensive experiments validate the effectiveness of PCLSurv for cancer survival prediction.

Data and preprocessing

We use 11 multi-omics cancer datasets to comprehensively evaluate the performance of PCLSurv, including acute myeloid leukemia (AML), breast invasive carcinoma (Breast), colon adenocarcinoma (Colon), glioblastoma multiforme (GBM), kidney renal clear cell carcinoma (Kidney), liver hepatocellular carcinoma (Liver), lung squamous cell carcinoma (Lung), skin cutaneous melanoma (Melanoma), ovarian serous cystadenocarcinoma (Ovarian), sarcoma (Sarcoma), and an Integrated dataset. The first 10 datasets are widely used for assessing cancer prognosis analysis and can be downloaded from The Cancer Genome Atlas [37]. The integrated dataset is constructed by combining multi-omics data from 8 of the 10 aforementioned cancer types [38]. For each dataset, we use two types of omics data: gene expression and miRNA expression. Specifically, for all datasets, miRNAs with

Table 1. Overview of the datasets utilized in the experiments

Datasets	Number of samples	Number of genes	Number of miRNAs
AML	160	2000	558
Breast	619	2000	891
Colon	219	2000	613
GBM	273	2000	534
Kidney	182	2000	796
Liver	365	2000	852
Lung	329	2000	878
Melanoma	428	2000	901
Ovarian	285	2000	616
Sarcoma	255	2000	838
Integrated	2557	2000	643

zero variance are filtered out, and the top 2000 genes with the highest variance are selected for downstream tasks. Each type of omics data is further normalized to have zero mean and unit variance. Detailed information regarding the 11 datasets is listed in Table 1.

Evaluation metrics

In our study, we use two primary metrics to evaluate model performance [39]: the concordance index (C-Index) and the area under the curve (AUC). The C-Index measures the predictive accuracy of survival models by assessing the concordance between the predicted and actual rankings of survival times. The C-Index can be mathematically expressed as

$$\text{C-index} = P\{\beta^T z_i > \beta^T z_j | O_j > O_i, \delta_i = 1\} \quad (14)$$

On the other hand, the AUC value is used to evaluate the model's ability to distinguish between classes in binary classification tasks by calculating the area under the receiver curve. The AUC can be formulated as

$$\text{AUC} = \frac{1}{N} \sum_{i \in Y} \sum_{j: O_i < O_j} \sum_{t} I(\beta^T z_i > \beta^T z_j) \quad (15)$$

where Y represents the set of all event times in the dataset, t denotes the cumulative count of comparable pairs for each event time, and $I()$ is the indicator function. Both the C-index and AUC values are bounded between 0 and 1. A value of 0.5 indicates random guessing, while a value of 1 represents perfect agreement between predicted and actual rankings for the C-index and perfect classification performance for AUC.

Implementation details

Each omics, the encoder network is set as a fully connected multilayer perceptron with layers structured as d^v -200-50-50 for all datasets, where d^v is the dimension of the input data. The decoder network mirrors the encoder structure with dimensions 50-50-200- d^v accordingly. All hidden layers utilize the ReLU activation function [40]. The number of prototypes k is iteratively selected from the range $\{2, \dots, 10\}$. The two temperature parameters τ and σ in the sample-level contrastive learning and PCL are set to 0.2 and 1, respectively. A detailed sensitivity analysis of these hyperparameters is provided in the Supplementary Figs S2–S5. The batch size is set as 64. The model is trained using the Adam optimizer, with the weight decay of 5×10^{-4} and the learning rate

of 2×10^{-3} . All experiments are implemented using Python 3.8.18 and PyTorch 1.7.1. The models are trained and evaluated on a workstation featuring an NVIDIA GeForce RTX 2080 Ti GPU.

Evaluation of survival prediction performance

In this subsection, we assess PCLSurv's survival prediction performance by benchmarking it against several state-of-the-art methods, including both single-omics and multi-omics approaches. The single-omics methods comprise RSF [41], En-cox [14], DeepHit [42], and DeepSurv [43], while the multi-omics methods include HFBSurv [24], CAMR [28], CustOmics [44], and FGCNSurv [27]. RSF uses random forests to analyze survival data. En-cox utilizes a cocktail algorithm combining coordinate descent. Both DeepHit and DeepSurv use neural networks to model survival data, allowing them to capture complex relationships between covariates. HFBSurv adopts a hierarchical multi-layer fusion strategy for multi-omics data integration. CAMR integrates the idea of a generative adversarial network. CustOmics employs variational autoencoders to learn latent representations. FGCNSurv uses graph convolutional neural networks to analyze sample relationships in survival prediction. For fair comparison, all of the above cancer survival prediction models are evaluated using exactly the same multi-views data throughout the experiment. Besides, we perform five-fold cross-validation to mitigate the impact of experimental randomness. We apply all methods on the 11 cancer datasets and the C-Index values obtained by each method are summarized in Table 2. For methods with single-omics data as input, we present two C-Index values by using gene expression and miRNA data, respectively. As demonstrated, PCLSurv consistently outperforms existing state-of-the-art methods across the datasets. It achieves the highest C-index values for 11 cancer datasets, suggesting that PCLSurv is particularly robust in capturing survival information from multi-omics data. Multi-omics models like PCLSurv, CAMR, and FGCNSurv generally outperform single-omics methods such as En-cox, DeepHit, and DeepSurv, especially in cancers like Colon, Lung, and Sarcoma. This indicates the superior performance of integrating multi-omics data types for cancer survival prediction. These findings highlight the superior predictive capability of PCLSurv in estimating survival outcomes across diverse cancer types, leveraging multi-omics data. Similarly, PCLSurv achieves higher AUC values (Table 3) than other models, further confirming its effectiveness in survival outcome. From the comparisons in both C-index and AUC values, PCLSurv demonstrates superior performance over traditional and some other advanced methods, particularly in datasets where multi-omics data integration plays a crucial role. The results suggest that multi-omics models are essential for improving survival predictions across various cancers, highlighting the importance of leveraging complementary information from different biological layers to enhance model accuracy.

To further assess the performance of PCLSurv, we evaluate whether there is a statistically significant difference in the survival time between the predicted high-risk and low-risk patient groups. Patients are stratified into these groups based on the median predicted risk scores, and the log-rank test is applied to determine if there is a significant difference in the survival times between the two groups. A more pronounced difference between the high-risk and low-risk groups indicates better model performance. Figure 2 presents the Kaplan–Meier survival curves for the predicted two groups from PCLSurv and the comparison methods, based on the Integrated dataset, along with the corresponding log-rank P-values. Among the single-omics methods,

Table 2. Comparison of survival prediction models across different cancers in terms of C-index

Dataset	PCLSurv	HFBSurv	CAMR	CustOmics	FGCNSurv	RSF	En-cox	DeepHit	DeepSurv
AML	0.751 ±0.04	0.650±0.07	0.654±0.05	0.609±0.09	0.681±0.04	0.635,0.663	0.653,0.560	0.613,0.580	0.638,0.593
Breast	0.777 ±0.06	0.677±0.03	0.667±0.10	0.685±0.07	0.756±0.06	0.712,0.561	0.745,0.670	0.590,0.569	0.640,0.497
Colon	0.780 ±0.04	0.773±0.07	0.687±0.17	0.597±0.11	0.710±0.10	0.517,0.516	0.433,0.548	0.552,0.528	0.504,0.504
GBM	0.654 ±0.01	0.563±0.06	0.659±0.06	0.569±0.05	0.610±0.03	0.520,0.534	0.537,0.534	0.533,0.536	0.553,0.562
Kidney	0.800 ±0.08	0.574±0.03	0.702±0.10	0.658±0.15	0.770±0.06	0.726,0.700	0.709,0.631	0.615,0.560	0.687,0.630
Liver	0.670 ±0.04	0.524±0.06	0.625±0.09	0.647±0.06	0.652±0.06	0.626,0.625	0.570,0.558	0.561,0.482	0.597,0.545
Lung	0.683 ±0.04	0.584±0.08	0.651±0.08	0.605±0.09	0.605±0.06	0.607,0.575	0.512,0.575	0.597,0.546	0.555,0.531
Melanoma	0.651 ±0.03	0.622±0.03	0.604±0.03	0.517±0.04	0.640±0.05	0.591,0.552	0.590,0.567	0.548,0.528	0.602,0.568
Ovarian	0.680 ±0.03	0.594±0.02	0.645±0.02	0.576±0.05	0.609±0.05	0.564,0.556	0.536,0.454	0.570,0.596	0.617,0.497
Sarcoma	0.755 ±0.04	0.626±0.05	0.657±0.03	0.550±0.09	0.703±0.05	0.689,0.670	0.596,0.606	0.610,0.506	0.647,0.667
Integrated	0.715 ±0.01	0.663±0.01	0.686±0.05	0.575±0.06	0.715±0.01	0.713,0.701	0.600,0.561	0.503,0.551	0.711,0.688

Table 3. Comparison of survival prediction models across different cancers in terms of AUC

Dataset	PCLSurv	HFBSurv	CAMR	CustOmics	FGCNSurv	RSF	En-cox	DeepHit	DeepSurv
AML	0.841 ±0.07	0.656±0.03	0.659±0.04	0.611±0.14	0.738±0.08	0.640,0.684	0.662,0.616	0.638,0.626	0.678,0.619
Breast	0.815 ±0.05	0.690±0.09	0.717±0.02	0.776±0.09	0.754±0.11	0.627,0.562	0.685,0.639	0.562,0.569	0.663,0.536
Colon	0.753 ±0.14	0.743±0.06	0.696±0.25	0.667±0.12	0.712±0.13	0.477,0.463	0.631,0.689	0.572,0.558	0.546,0.598
GBM	0.707 ±0.02	0.592±0.02	0.650±0.06	0.605±0.07	0.665±0.05	0.443,0.506	0.582,0.608	0.590,0.533	0.550,0.569
Kidney	0.833 ±0.08	0.581±0.09	0.756±0.10	0.675±0.20	0.832±0.07	0.738,0.724	0.720,0.677	0.616,0.598	0.626,0.639
Liver	0.689 ±0.05	0.580±0.01	0.661±0.14	0.653±0.07	0.668±0.06	0.563,0.585	0.557,0.571	0.531,0.579	0.513,0.516
Lung	0.726 ±0.06	0.702±0.08	0.676±0.13	0.660±0.12	0.630±0.09	0.563,0.581	0.538,0.553	0.525,0.596	0.547,0.532
Melanoma	0.697 ±0.04	0.660±0.02	0.612±0.03	0.585±0.04	0.683±0.06	0.465,0.519	0.538,0.540	0.508,0.572	0.643,0.620
Ovarian	0.729 ±0.05	0.598±0.04	0.671±0.06	0.637±0.05	0.597±0.07	0.548,0.557	0.597,0.604	0.523,0.566	0.607,0.531
Sarcoma	0.813 ±0.01	0.593±0.04	0.643±0.04	0.582±0.14	0.754±0.07	0.658,0.606	0.593,0.544	0.647,0.569	0.604,0.608
Integrated	0.761 ±0.01	0.698±0.02	0.680±0.10	0.592±0.09	0.760±0.02	0.656,0.665	0.585,0.568	0.542,0.528	0.637,0.683

DeepSurv, a deep learning-based approach, outperforms traditional methods like En-cox and RSF, showing more significant log-rank P-values for both gene expression and miRNA data. Moreover, multi-omics approaches, including FGCNSurv and our proposed PCLSurv, achieve even more significant results than the single-omics methods. Notably, PCLSurv method achieves the most significant log-rank P-value of $3.09\text{e-}19$, surpassing all other multi-omics methods, highlighting its superior ability to distinguish between high-risk and low-risk groups and its effectiveness in survival prediction. Supplementary Fig. S1 displays Kaplan-Meier survival curves by PCLSurv across all datasets. The clear separation between the curves, combined with significant log-rank P-values, confirms that PCLSurv effectively differentiates between high-risk and low-risk groups, demonstrating robust performance in survival prediction across various datasets.

Evaluation of clinical associations with survival prediction

To assess the correlation between PCLSurv's survival prediction and clinical factors, we perform both univariate and multivariate Cox proportional hazards analyses [45] on Breast dataset. Specifically, according to the prediction scores obtained by PCLSurv, patients are classified into high-risk and low-risk categories. We then conduct analyses focusing on the following key clinical factors: age at diagnosis, histologic grade, tumor size (T stage), lymph node involvement (N stage), metastasis (M stage), and the predicted risk groups. Table 4 shows that in both analyses, the predicted risk from our model is the most significant factor impacting survival among the six factors considered.

In the univariate analysis, the risk group predicted by PCLSurv serves as a highly effective survival indicator, with a P-value

low than 0.005. For the multivariate analysis, we further explore the association between clinical factors and survival outcomes, assessing the contribution of each predictor to mortality risk. Notably, in this multivariate context, the risk group variable remains the statistically significant factor, with the high-risk group showing a slightly elevated mortality risk compared to the low-risk group (HR = 3.21). These results suggest that the PCLSurv-predicted risk effectively captures critical prognostic information, integrating the effects of the various clinical characteristics. Therefore, the model's predicted survival risk stands out as a key prognostic factor in survival prediction.

To evaluate the survival outcomes associated with radiotherapy across different risk categories, we further divide patients within each risk group into two subgroups according to the radiotherapy treatment information on Breast dataset. For the high-risk group, the numbers of patients with and without radiotherapy treatment are 139 and 170, respectively, while for the low-risk group, the corresponding numbers are 153 and 157. Kaplan-Meier survival curves are constructed to compare the survival times of treated and untreated patients in each subgroup. As shown in Fig. 3, patients in both high-risk and low-risk groups who received treatment demonstrate longer survival times compared to those who did not, highlighting the potential benefits of radiotherapy across risk categories. In addition, radiotherapy has a more significant impact on the survival time of patients in high-risk groups compared to those in low-risk groups.

Ablation study

To evaluate the contribution of each component in PCLSurv, we perform a comprehensive ablation study, systematically modifying or removing key modules to observe their effects on survival

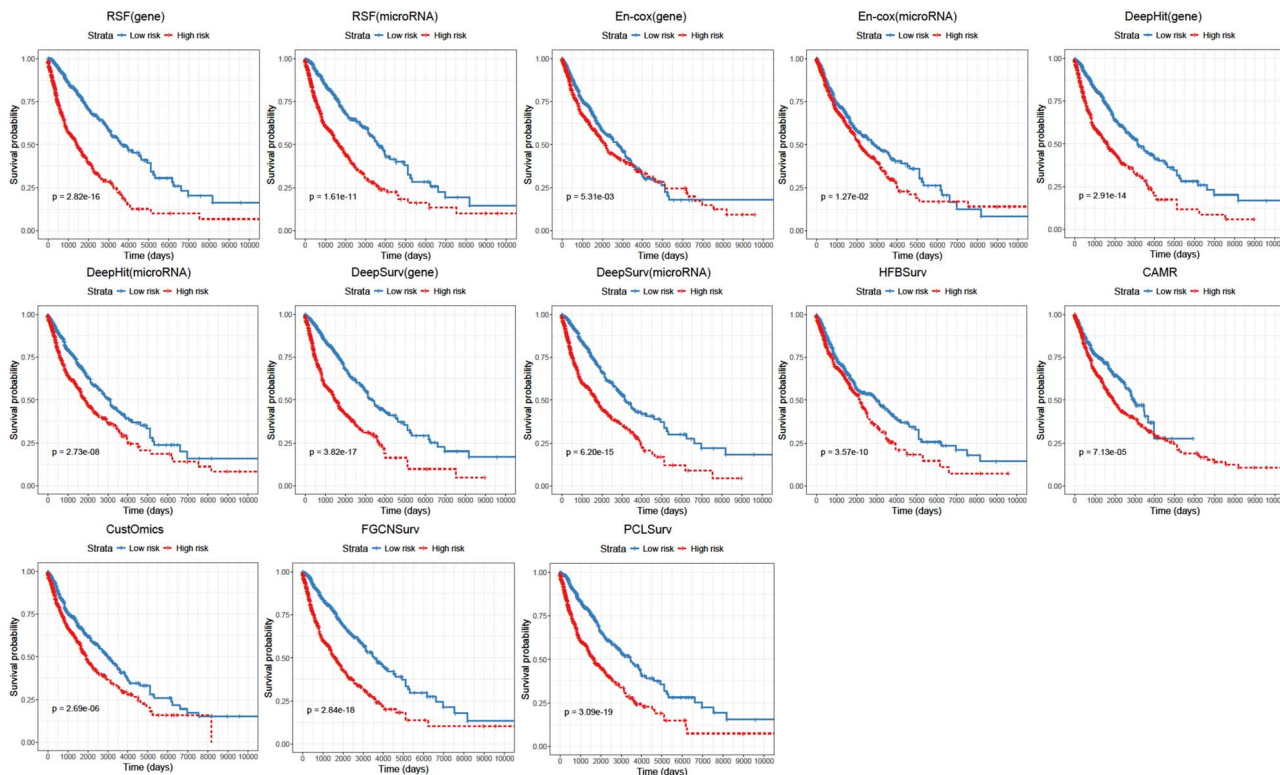


Figure 2. Performance of the comparison methods using Kaplan–Meier curve for the Integrated datasets.

Table 4. Univariate and multivariate Cox proportional hazards analysis

Variable	Univariate			Multivariate		
	Hazard ratio	95% CI	P value	Hazard ratio	95% CI	P value
Age (≤ 50 / > 50)	1.26	0.66–2.40	0.48	1.83	0.93–3.60	0.08
Grade ($\leq II$ / $> II$)	2.06	0.81–5.22	0.13	1.69	0.63–4.56	0.30
T stage ($\leq T2$ / $> T2$)	1.15	0.65–2.06	0.63	1.34	0.72–2.50	0.36
N stage ($\leq N1$ / $> N1$)	2.56	1.32–4.94	0.01	2.80	1.39–5.66	<0.005
M stage (M0/MX)	0.62	0.22–1.75	0.36	0.43	0.15–1.25	0.12
Risk group (PCLSurv)	3.21	1.71–6.04	<0.005	3.83	1.97–7.46	<0.005

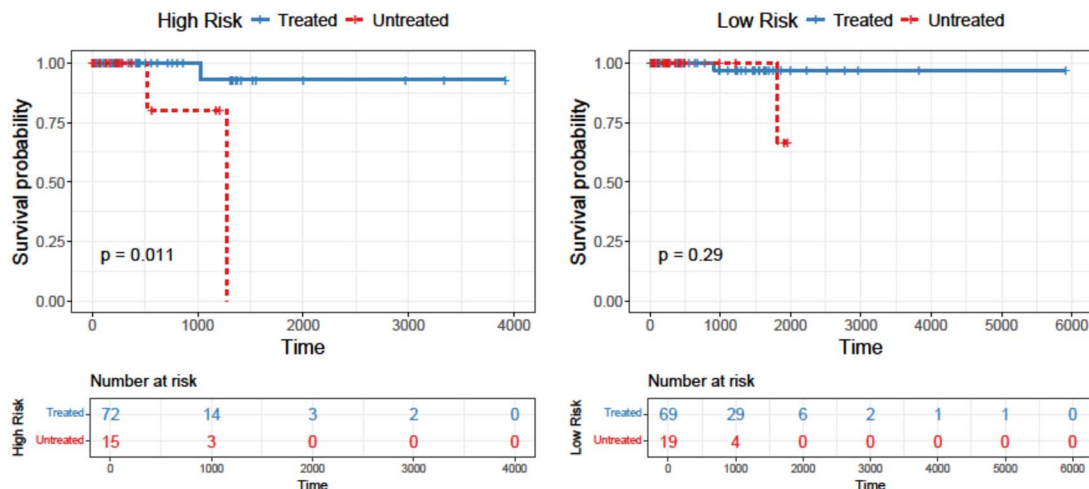


Figure 3. Kaplan-Meier survival curves for breast cancer patients stratified by radiotherapy.

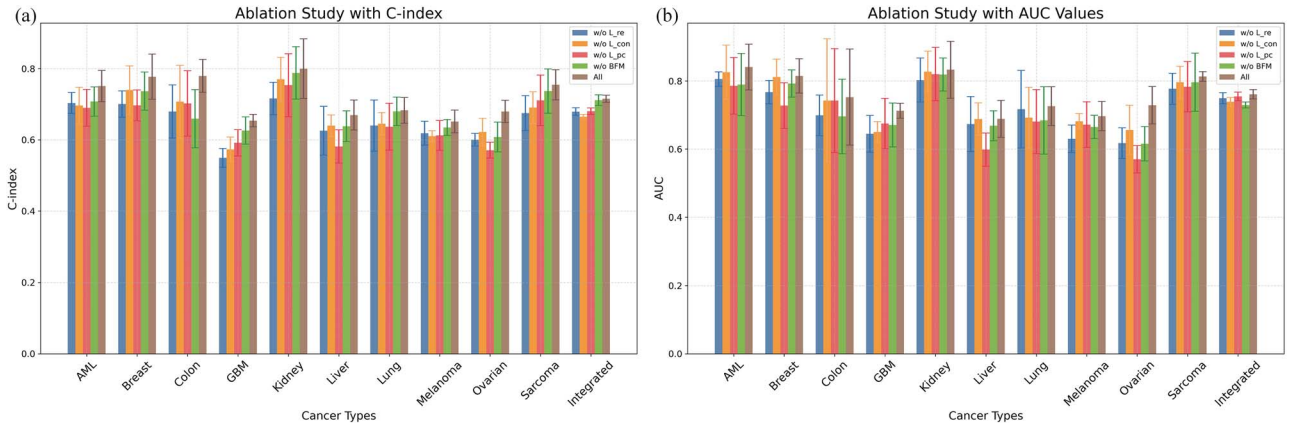


Figure 4. Ablation study. (a) Ablation study with C-index. The figure presents the impact of removing key modules from the model on C-index performance across datasets. (b) Ablation study with AUC. The figure illustrates how the exclusion of specific model modules affects AUC values across datasets.

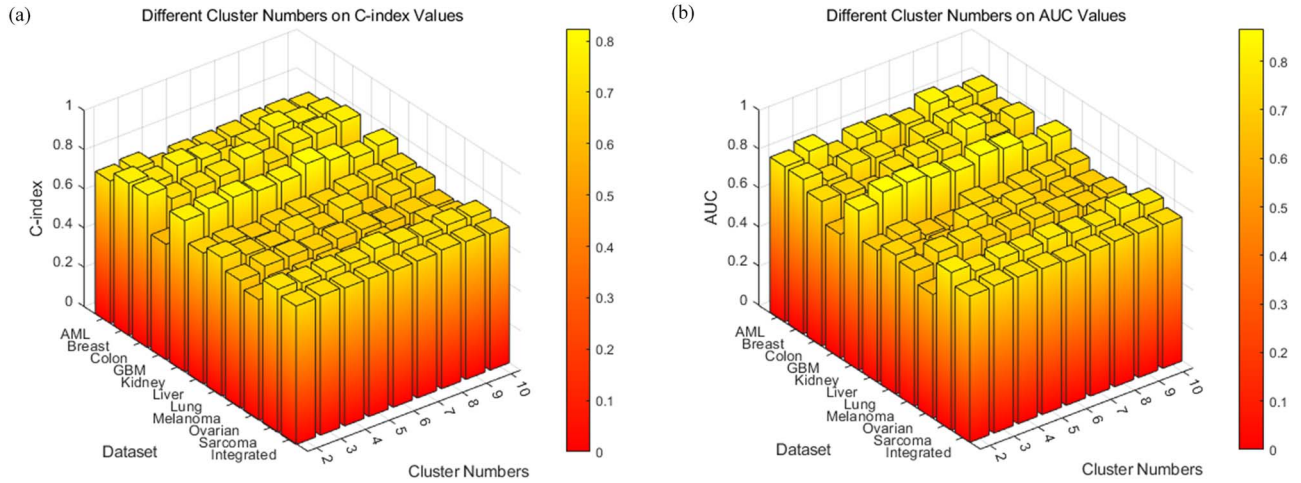


Figure 5. Impact of different cluster numbers. The figure depicts the performance of the model, on 11 cancer datasets as the number of clusters varies from 2 to 10. (a) Impact of different cluster numbers on C-index values. (b) Impact of different cluster numbers on AUC values.

prediction performance. Our model includes four components: reconstruction module, sample-level contrastive learning module, BFM, and PCL module. For each variant, we track its performance using both the C-index and AUC (Fig. 4).

The first experiment removes the reconstruction module, resulting in a substantial decline in predictive performance. Specifically, for the Breast dataset, the C-index drops from 0.7773 to 0.7305, while the AUC value decreases from 0.8149 to 0.7674. Similar trends are observed across other datasets. These findings suggest that the reconstruction module plays a role in maintaining data integrity while projecting high-dimensional data into a more discriminative space. In the second experiment, we remove the sample-level contrastive learning module to evaluate its impact on model performance. This modification results in a moderate reduction in both C-index and AUC values on all datasets. Notably, in certain datasets, such as Colon and GBM, the exclusion of the sample-level contrastive learning module causes a more substantial deterioration in performance. Then, we evaluate the impact of excluding the prototype contrastive learning module. The removal of this component degrades performance metrics across all datasets. Both the C-index and AUC exhibit significant declines. Lastly, we replace the BFM with a simple summation operation. As expected, this adjustment also leads to a decline in

performance, highlighting the effectiveness of the BFM in capturing complex interactions across omics data. Supplementary Tables S1 and S2 provide a detailed summary of the ablation experiments across 11 datasets, offering a clear comparison of performance metrics for each scenario. The consistent decline in both C-index and AUC following the removal of each module underscores the essential contributions of these components to the model's predictive accuracy. In conclusion, the ablation study reveals that all four modules are indispensable for robust survival prediction.

Parameter analysis

In our PCL module, the number of clusters k is set to obtain prototypes at different levels for capturing semantic similarities. We analyze the impact of k on the model performance. We report the C-index and AUC values for cluster numbers, ranging from 2 to 10 in Fig. 5. Supplementary Tables S3 and S4 provide detailed results of the analysis experiments across 11 datasets. The results reveal minimal performance fluctuations across this range, indicating robust model performance under various clustering configurations. Notably, values within this range yield favorable results with only slight variations depending on the dataset. This stability demonstrates that our method, leveraging PCL, effectively adapts

to diverse clustering scenarios, enhancing its utility for cancer survival prediction.

Conclusion

In this study, we introduce PCLSurv, a novel method for cancer survival prediction based on prototype contrastive learning. PCLSurv is designed to generate lower-dimensional representations that retain essential, non-redundant information critical for predicting cancer patient survival. Specifically, PCLSurv clusters samples according to their semantic structure following a stepwise fusion process, thereby extracting more informative latent features from multimodal data. Within this framework, we incorporate learnable prototype representations that align similar information across different samples, yielding latent representations well-suited for survival analysis. Experimental results demonstrate that PCLSurv significantly outperforms existing prediction methods. Additionally, Kaplan–Meier curve analysis and Cox proportional hazards analysis validate the efficacy of PCLSurv in cancer survival prediction. In summary, PCLSurv can flexibly extend to handling multi-omics datasets and serve as a valuable tool for advancing research in the field.

Despite these promising results, there is still room for improvement. First, the performance of PCLSurv is limited by the constraints of the current multimodal cancer datasets. Expanding our study to include a larger patient cohort could enhance predictive accuracy. Furthermore, PCLSurv cannot be applied to incomplete multi-omics datasets as it requires cross-view counterparts to construct positive pairs for sample-level contrastive learning. Addressing this limitation would require methodological advances, such as alternative pair construction strategies or imputation techniques, which warrant future exploration.

Key Points

- We propose PCLSurv, a PCL-based multi-omics data integration model for cancer survival prediction. PCLSurv integrates autoencoders for feature extraction, sample-level contrastive learning for cross-view consistency, and a BFM for feature integration. PCL enhances the model by aligning similar samples with shared prototypes and separating unrelated ones, improving the capture of high-level semantic relationships.
- We extensively evaluate PCLSurv against several state-of-the-art methods on 11 cancer datasets, demonstrating superior performance in cancer survival prediction, with significantly improved predictive accuracy and robustness.
- Cox proportional hazards analyses demonstrate that PCLSurv-predicted risk scores are significantly associated with patient survival, validating its clinical relevance. Radiotherapy analysis suggests that risk scores can help stratify patients based on their treatment response, highlighting PCLSurv's potential in personalized therapy.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions.

Author contributions

Z.L. and C.L. conceived the ideas, Z.L. and W.C. preprocessed the raw multi-omics datasets and conducted the experiments, Z.L. and H.Z. analyzed the results. Z.L., W.C., H.Z., and C.L. wrote and reviewed the manuscript.

Supplementary data

Supplementary data are available at *Briefings in Bioinformatics* online.

Conflict of interest: All authors declared no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China [62372279], the Natural Science Foundation of Shandong Province [ZR2023MF119], and the Jinan Clinical Medical Science and Technology Innovation Plan Project(202430031).

References

1. Siegel RL, Miller KD, Wagle NS. et al. Cancer statistics, 2023. *CA Cancer J Clin* 2023;**73**:17–48. <https://doi.org/10.3322/caac.21763>
2. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Eng J Med* 2009;**360**:790–800. <https://doi.org/10.1056/NEJMra0801289>
3. Stahlhut C, Slack FJ. MicroRNAs and the cancer phenotype: profiling, signatures and clinical implications. *Genome Med* 2013;**5**:1–12.
4. Steele CD, Ammal Abbasi SM, Islam A. et al. Signatures of copy number alterations in human cancer. *Nature* 2022;**606**:984–91. <https://doi.org/10.1038/s41586-022-04738-6>
5. Doré S, Ali M, Sorin M. et al. Exploring the prognostic significance of arm-level copy number alterations in triple-negative breast cancer. *Oncogene* 2024;**43**:2015–24. <https://doi.org/10.1038/s41388-024-03051-y>
6. Eckhardt A, Drexler R, Schoof M. et al. Mean global DNA methylation serves as independent prognostic marker in IDH-wildtype glioblastoma. *Neuro Oncol* 2024;**26**:503–13. <https://doi.org/10.1093/neuonc/noad197>
7. Kim Y, Jeon J, Mejia S. et al. Targeted proteomics identifies liquid-biopsy signatures for extracapsular prostate cancer. *Nat Commun* 2016;**7**:11906. <https://doi.org/10.1038/ncomms11906>
8. Zaccaria GM, Altini N, Mezzolla G. et al. SurvIAE: survival prediction with interpretable autoencoders from diffuse large B-cells lymphoma gene expression data. *Comput Methods Programs Biomed* 2024;**244**:107966. <https://doi.org/10.1016/j.cmpb.2023.107966>
9. Peng Y, Qiyang W, Wang L. et al. A DNA methylation signature to improve survival prediction of gastric cancer. *Clin Epigenetics* 2020;**12**:1–16. <https://doi.org/10.1186/s13148-020-0807-x>
10. Lan W, Liao H, Chen Q. et al. DeepKEGG: a multi-omics data integration framework with biological insights for cancer recurrence prediction and biomarker discovery. *Brief Bioinform* 2024;**25**:bbae185. <https://doi.org/10.1093/bib/bbae185>
11. Pang J, Liang B, Ding R. et al. A denoised multi-omics integration framework for cancer subtype classification and survival prediction. *Brief Bioinform* 2023;**24**:bbad304. <https://doi.org/10.1093/bib/bbad304>

12. Cox DR. Regression models and life-tables. *J R Stat Soc B Methodol* 1972;**34**:187–202. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>
13. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997;**16**:385–95. [https://doi.org/10.1002/\(SICI\)1097-0258\(19970228\)16:4<385::AID-SIM380>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3)
14. Yang Y, Zou H. A cocktail algorithm for solving the elastic net penalized Cox's regression in high dimensions. *Stat Interface* 2013;**6**:167–73. <https://doi.org/10.4310/SII.2013.v6.n2.a1>
15. Jinjian M, Liu Q, Kuo L. et al. Bayesian variable selection for the Cox regression model with spatially varying coefficients with applications to Louisiana respiratory cancer data. *Biom J* 2021;**63**: 1607–22. <https://doi.org/10.1002/bimj.202000047>
16. Madjar K, Zucknick M, Ickstadt K. et al. Combining heterogeneous subgroups with graph-structured variable selection priors for Cox regression. *BMC Bioinformatics* 2021;**22**:586. <https://doi.org/10.1186/s12859-021-04483-z>
17. Hai Y, Ma J, Yang K. et al. Bayesian linear mixed model with multiple random effects for prediction analysis on high-dimensional multi-omics data. *Bioinformatics* 2023;**39**:btad647. <https://doi.org/10.1093/bioinformatics/btad647>
18. Zhu S, Kong W, Zhu J. et al. The genetic algorithm-aided three-stage ensemble learning method identified a robust survival risk score in patients with glioma. *Brief Bioinform* 2022;**23**:bbac344. <https://doi.org/10.1093/bib/bbac344>
19. Ma B, Yan G, Chai B. et al. XGBLC: an improved survival prediction model based on XGBoost. *Bioinformatics* 2022;**38**:410–8. <https://doi.org/10.1093/bioinformatics/btab675>
20. Zhao Z, Feng Q, Zhang Y. et al. Adaptive risk-aware sharable and individual subspace learning for cancer survival analysis with multi-modality data. *Brief Bioinform* 2023;**24**:bbac489. <https://doi.org/10.1093/bib/bbac489>
21. Wang F-A, Zhuang Z, Gao F. et al. TMO-Net: an explainable pre-trained multi-omics model for multi-task learning in oncology. *Genome Biol* 2024;**25**:1–24.
22. Bichindaritz I, Liu G, Bartlett C. Integrative survival analysis of breast cancer with gene expression and DNA methylation data. *Bioinformatics* 2021;**37**:2601–8. <https://doi.org/10.1093/bioinformatics/btab140>
23. Jung Hun O, Choi W, Ko E. et al. PathCNN: interpretable convolutional neural networks for survival prediction and pathway analysis applied to glioblastoma. *Bioinformatics* 2021;**37**:i443–50.
24. Li R, Xingqi W, Li A. et al. HFBSurv: hierarchical multimodal fusion with factorized bilinear models for cancer survival prediction. *Bioinformatics* 2022;**38**:2587–94. <https://doi.org/10.1093/bioinformatics/btac113>
25. Chen Z, Wei Q. Developing an improved survival prediction model for disease prognosis. *Biomolecules* 2022;**12**:1751. <https://doi.org/10.3390/biom12121751>
26. Cho HJ, Shu M, Bekiranov S. et al. Interpretable meta-learning of multi-omics data for survival analysis and pathway enrichment. *Bioinformatics* 2023;**39**:btad113. <https://doi.org/10.1093/bioinformatics/btad113>
27. Wen G, Li L. FGCNSurv: dually fused graph convolutional network for multi-omics survival prediction. *Bioinformatics* 2023;**39**:btad472. <https://doi.org/10.1093/bioinformatics/btad472>
28. Xingqi W, Shi Y, Wang M. et al. CAMR: cross-aligned multimodal representation learning for cancer survival prediction. *Bioinformatics* 2023;**39**:btad025.
29. Shi M, Li X, Li M. et al. Attention-based generative adversarial networks improve prognostic outcome prediction of cancer from multimodal data. *Brief Bioinform* 2023;**24**:bbad329. <https://doi.org/10.1093/bib/bbad329>
30. Yan H, Weng D, Dongguo Li YG. et al. Prior knowledge-guided multilevel graph neural network for tumor risk prediction and interpretation via multi-omics data integration. *Brief Bioinform* 2024;**25**:bbae184. <https://doi.org/10.1093/bib/bbae184>
31. Jiang L, Chao X, Bai Y. et al. Autosurv: interpretable deep learning framework for cancer survival analysis incorporating clinical and multi-omics data. *NPJ Precis Oncol* 2024;**8**:4. <https://doi.org/10.1038/s41698-023-00494-6>
32. Sun A, Franzmann EJ, Chen Z. et al. Deep contrastive learning for predicting cancer prognosis using gene expression values. *Brief Bioinform* 2024;**25**:bbae544. <https://doi.org/10.1093/bib/bbae544>
33. Wen G, Li L. MMOSurv: meta-learning for few-shot survival analysis with multi-omics data. *Bioinformatics* 2024;**41**:btae684. <https://doi.org/10.1093/bioinformatics/btae684>
34. Zhang Y, Xu Y, Chen J. et al. Prototypical information bottlenecks and disentangling for multimodal cancer survival prediction. In: *The Twelfth International Conference on Learning Representations*. OpenReview.net, Appleton WI.
35. Lan W, Zhou G, Chen Q. et al. Contrastive clustering learning for multi-behavior recommendation. *ACM Trans Inf Syst* 2024;**43**: 1–23. <https://doi.org/10.1145/3698192>
36. Yu Z, Yu J, Fan J. et al. Multi-modal factorized bilinear pooling with co-attention learning for visual question answering. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1821–30. Los Alamitos, CA: IEEE Computer Society, 2017.
37. Zhu Y, Qiu P, Ji Y. TCGA-assembler: open-source software for retrieving and processing TCGA data. *Nat Methods* 2014;**11**: 599–600. <https://doi.org/10.1038/nmeth.2956>
38. Chen W, Wang H, Liang C. Deep multi-view contrastive learning for cancer subtype identification. *Brief Bioinform* 2023;**24**:bbad282. <https://doi.org/10.1093/bib/bbad282>
39. Shao W, Han Z, Cheng J. et al. Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE Trans Med Imaging* 2019;**39**:99–110. <https://doi.org/10.1109/TMI.2019.2920608>
40. Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 315–23. Microtome Publishing, Brookline, MA, JMLR Workshop and Conference Proceedings, 2011.
41. Ishwaran H, Kogalur UB, Blackstone EH. et al. Random survival forests. *Ann Appl Stat* 2008;**2**:841–60. <https://doi.org/10.1214/08-AOAS169>
42. Lee C, Zame W, Yoon J. et al. DeepHit: a deep learning approach to survival analysis with competing risks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. **32**. Palo Alto, CA: AAAI Press, 2018. <https://doi.org/10.1609/aaai.v32i1.11842>
43. Katzman JL, Shaham U, Cloninger A. et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol* 2018;**18**: 1–12.
44. Benkirane H, Pradat Y, Michiels S. et al. Customics: a versatile deep-learning based strategy for multi-omics integration. *PLoS Comput Biol* 2023;**19**:e1010921. <https://doi.org/10.1371/journal.pcbi.1010921>
45. Fulong Y, Quan F, Jinyuan X. et al. Breast cancer prognosis signature: linking risk stratification to disease subtypes. *Brief Bioinform* 2019;**20**:2130–40.