MurreNet: Modeling Holistic Multimodal Representations Between Histopathology and Genomic Profiles for Survival Prediction

Poster ID: XXXX

MICCAI 2025 Daejeon REPUBLIC OF KOREA

Mingxin Liu[1], Chengfei Cai[1,2], Jun Li[1], Pengbo Xu[3], Jinze Li[1], Jiquan Ma[4], Jun Xu[1,*]

[1] Jiangsu Key Laboratory of Intelligent Medical Image Computing, Nanjing University of Information Science and Technology

[2] Taizhou University, [3] Harbin Medical University, [4] Heilongjiang University | Contact: mxliu.mercy@gmail.com, jxu@nuist.edu.cn*

Motivations

(1) Multimodal features often contain **overlapping and redundant information** that will hinder reliable fusion, making it challenging to identify complementary information across modalities.

(2) Disentangled features are insufficient without proper guidance, may still **lack semantic structure**, posing a challenge to effective **meaningful alignment across feature spaces**.

(3) Simple feature fusion **overlook the hierarchical and intricate dependencies** between heterogeneous modalities, leading to suboptimal cross-modality representation learning.

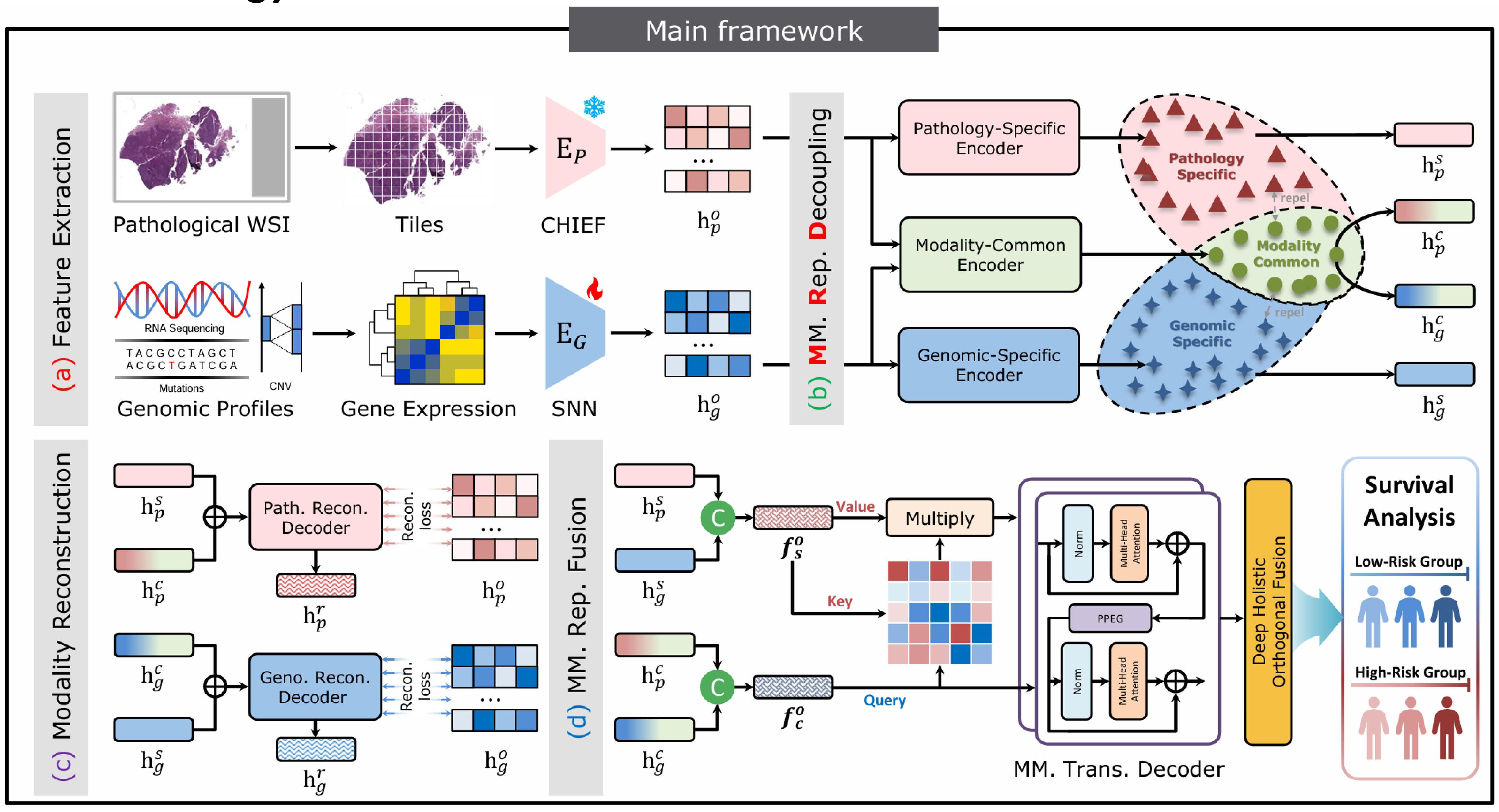
Contributions

(1) **M**ultimodal **R**epresentation **D**ecomposition module (**MRD**) is present to **explicitly decompose input data into modality-specific and -shared representations**, thereby reducing redundancy between modalities.

(2) A novel regularization strategy is designed to impose constraints on **distribution similarity, difference**, and **representativeness** of the disentangled modality representations.

(3) **D**eep **H**olistic **O**rthogonal **F**usion (**DHOF**) is proposed to integrate the modality-specific and -common representations for modeling **holistic interactions across heterogenous modalities**.

Methodology



Multimodal Representation Decoupling

(1) Extract specific representations via modality-specific encoders (2-Linear-layer MLP):

$$\mathrm{h}_p^s, \mathrm{h}_g^s = \mathrm{MLP}(\mathrm{h}_p^o), \mathrm{MLP}(\mathrm{h}_g^o)$$

(2) Calculates multimodal co-attention matrix:

$$\mathbb{A} = \mathrm{Linear}(\mathrm{h}_p^o)^\top \cdot \mathrm{Linear}(\mathrm{h}_g^o),$$

(3) Generate modality-common representations via modality-common encoder Γ_c :

$$\mathrm{h}_p^c = \Gamma_c(\mathrm{h}_p^o, \mathrm{h}_g^o) = \mathrm{MLP}(\mathbb{A}^\top) * \mathrm{h}_p^o,$$

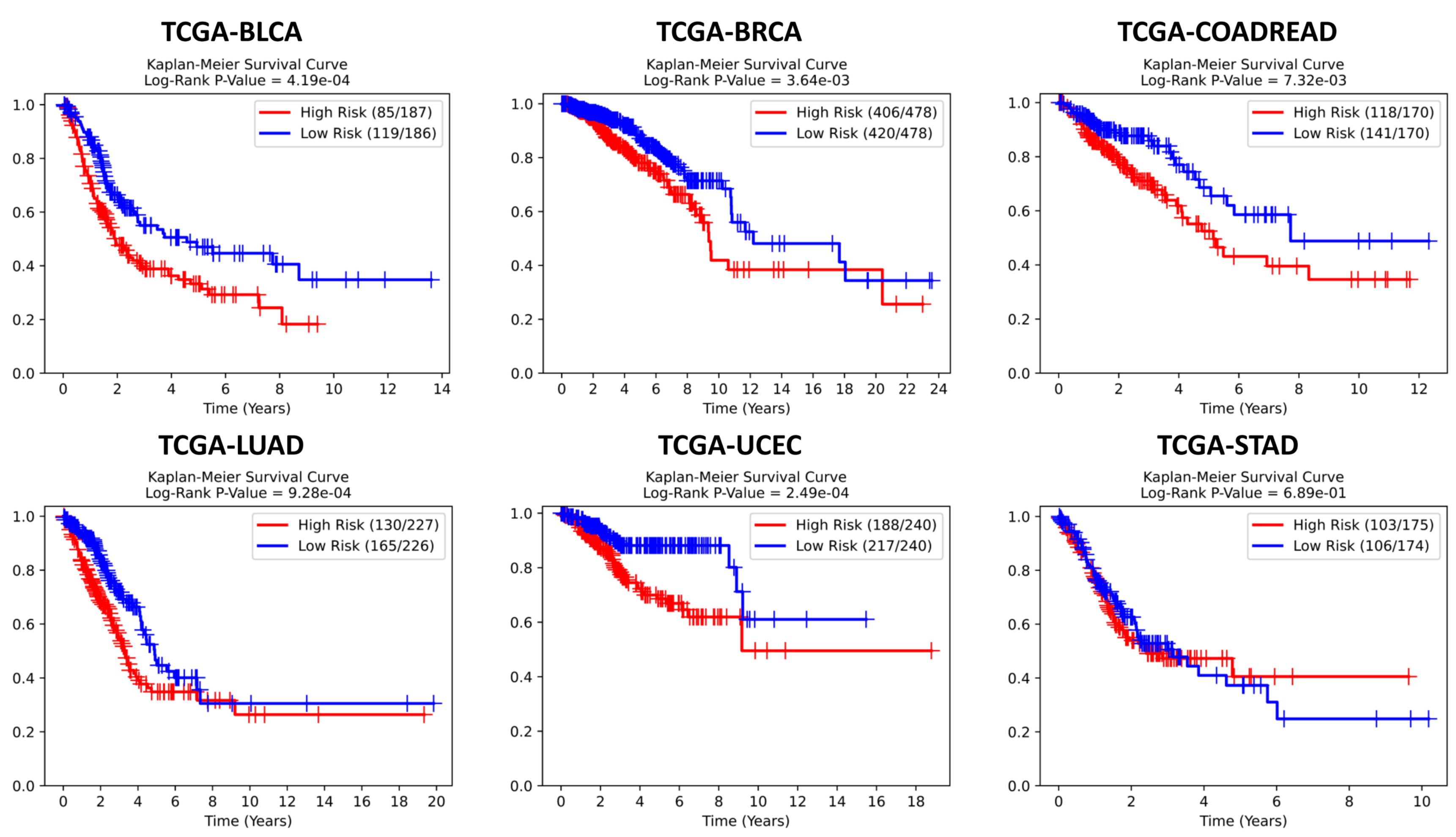
$$\mathrm{h}_g^c = \Gamma_c(\mathrm{h}_p^o, \mathrm{h}_g^o) = \mathrm{MLP}(\mathbb{A}) * \mathrm{h}_g^o,$$

Results & Analysis

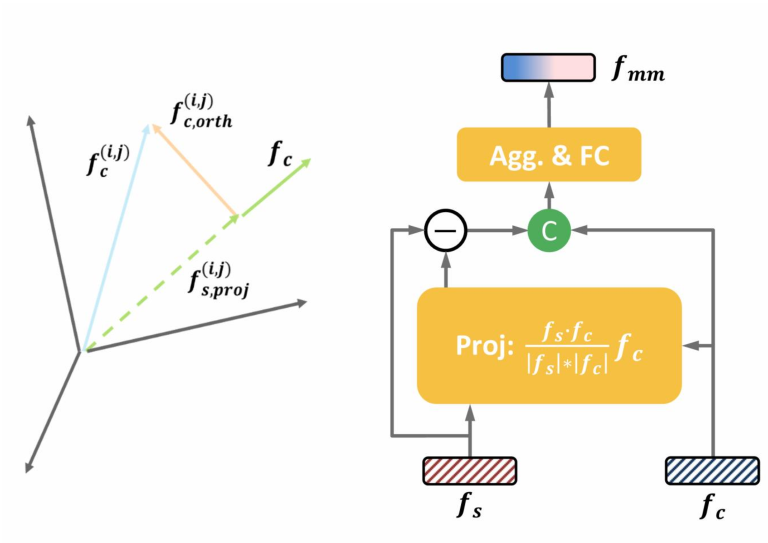
Comparison with SOTAs for Survival Prediction

Methods	Modality P.	Modality G.	BLCA ($n = 373$)	BRCA ($n = 956$)	COADREAD ($n = 340$)	LUAD ($n = 453$)	UCEC ($n = 480$)	STAD ($n = 349$)
MLP		✓	0.611 ± 0.030	0.619 ± 0.053	0.675 ± 0.049	0.619 ± 0.044	0.675 ± 0.069	0.607 ± 0.045
SNN [11]		✓	0.625 ± 0.054	0.621 ± 0.051	0.606 ± 0.087	0.615 ± 0.044	0.703 ± 0.072	0.603 ± 0.052
Coxnet [7]		✓	0.591 ± 0.061	0.593 ± 0.067	0.601 ± 0.059	0.605 ± 0.044	0.535 ± 0.051	0.523 ± 0.055
ABMIL [8]	✓		0.640 ± 0.032	0.642 ± 0.030	0.703 ± 0.064	0.607 ± 0.041	0.701 ± 0.038	<u>0.638 ± 0.028</u>
DSMIL [13]	✓		0.636 ± 0.031	0.657 ± 0.037	0.684 ± 0.071	0.591 ± 0.026	0.729 ± 0.063	0.630 ± 0.032
CLAM [19]	✓		0.618 ± 0.042	0.593 ± 0.077	0.643 ± 0.108	0.580 ± 0.029	0.647 ± 0.037	0.634 ± 0.054
TransMIL [20]	✓		0.640 ± 0.039	0.590 ± 0.058	0.684 ± 0.024	0.592 ± 0.048	0.705 ± 0.027	0.604 ± 0.054
DTFD-MIL [28]	✓		0.610 ± 0.025	0.625 ± 0.060	0.633 ± 0.061	0.590 ± 0.033	0.663 ± 0.016	0.580 ± 0.033
M3IF [14]	✓	✓	0.662 ± 0.043	0.639 ± 0.062	<u>0.710 ± 0.043</u>	0.628 ± 0.050	0.724 ± 0.103	0.607 ± 0.044
MCAT [5]	✓	✓	0.677 ± 0.062	<u>0.691 ± 0.041</u>	0.649 ± 0.052	0.675 ± 0.039	0.658 ± 0.055	0.586 ± 0.027
CMTA [30]	✓	✓	0.683 ± 0.016	0.667 ± 0.033	0.678 ± 0.040	0.648 ± 0.036	0.740 ± 0.066	0.584 ± 0.023
MoME [25]	✓	✓	0.650 ± 0.033	0.659 ± 0.032	0.692 ± 0.038	0.650 ± 0.024	0.705 ± 0.041	0.630 ± 0.039
MOTCat [26]	✓	✓	<u>0.683 ± 0.025</u>	0.675 ± 0.021	0.618 ± 0.030	0.667 ± 0.027	0.685 ± 0.052	0.596 ± 0.028
SurvPath [9]	✓	✓	0.673 ± 0.011	0.685 ± 0.024	0.650 ± 0.024	<u>0.676 ± 0.036</u>	0.737 ± 0.049	0.622 ± 0.045
PORPOISE [6]	✓	✓	0.651 ± 0.046	0.624 ± 0.048	0.702 ± 0.045	0.642 ± 0.042	<u>0.737 ± 0.097</u>	0.612 ± 0.038
MurreNet	✓	✓	**0.710 ± 0.030**	**0.718 ± 0.016**	**0.725 ± 0.044**	**0.691 ± 0.040**	**0.752 ± 0.075**	**0.651 ± 0.062**

Kaplan-Meier Survival Curves



Deep Holistic Orthogonal Fusion



(1) Calculates the projection $f_{s,proj}^{(i,j)}$ of each feature in specific representation onto the common representation f_c:

$$f_{s,orth}^{(i,j)} = f_s^{(i,j)} - f_{s,proj}^{(i,j)},$$

(2) Compute orthogonal component:

$$f_{s,proj}^{(i,j)} = \frac{\sum_{k=1}^{K} f_{s,k}^{(i,j)} \cdot f_{c,k}}{\sum_{k=1}^{K} (f_{c,k})^2} \cdot f_c$$

Training Objectives

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{sim} + \beta \mathcal{L}_{diff} + \gamma \mathcal{L}_{recon} + \mathcal{L}_{surv}$$

$$\mathcal{L}_{sim} = \mathcal{L}_1(\mathrm{h}_p^o, \mathrm{h}_p^c) = \frac{1}{n} \left\| \mathrm{h}_p^o - \mathrm{h}_p^c \right\|_1 \quad (\alpha = \beta = 1 \times 10^{-4}, \gamma = 1)$$

$$\mathcal{L}_{diff} = \mathcal{D}_{KL}(\mathrm{h}_p^c, \mathrm{h}_p^s) + \mathcal{D}_{KL}(\mathrm{h}_g^c, \mathrm{h}_g^s) = \sum_m^{\{p,g\}} \left[\mathrm{h}_m^c \cdot \log(\frac{\mathrm{h}_m^c}{\mathrm{h}_m^s}) \right]$$

$$\mathcal{L}_{recon} = \frac{1}{2} \left(\mathcal{L}_{MSE}(\mathrm{h}_p^o, \mathrm{h}_p^r) + \mathcal{L}_{MSE}(\mathrm{h}_g^o, \mathrm{h}_g^r) \right) = \frac{1}{2} \sum_m^{\{p,g\}} (\mathrm{h}_m^o - \mathrm{h}_m^r)^2$$

Acknowledgements. This work was supported by the National Key R&D Program of China (No.2023YFC3402800), National Natural Science Foundation of China (Nos.82441029, 62171230, 62101365, 92159301, 62301263, 62301265, 62302228, 82302291, 82302352, 62401272), Jiangsu Provincial Department of Science and Technology's Major Project on Frontier-Leading Basic Research in Technology (No.BK2023200).

Paper IMIC Lab